Smart Structures and Systems, *Vol. 21, No. 5 (2018) 705-713*
DOI: https://doi.org/10.12989/sss.2018.21.5.705

Online condition assessment of high-speed trains based on Bayesian forecasting approach and time series analysis

Lin-Hao Zhang[1a], You-Wu Wang[1,2b], Yi-Qing Ni[1,2c] and Siu-Kai Lai[*1]

[1]*Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*
[2]*Hong Kong Branch of Chinese National Rail Transit Electrification and Automation Engineering Technology Research Center, Hong Kong*

(Received December 20, 2017, Revised March 29, 2018, Accepted March 31, 2018)

Abstract. High-speed rail (HSR) has been in operation and development in many countries worldwide. The explosive growth of HSR has posed great challenges for operation safety and ride comfort. Among various technological demands on high-speed trains, vibration is an inevitable problem caused by rail/wheel imperfections, vehicle dynamics, and aerodynamic instability. Ride comfort is a key factor in evaluating the operational performance of high-speed trains. In this study, online monitoring data have been acquired from an in-service high-speed train for condition assessment. The measured dynamic response signals at the floor level of a train cabin are processed by the Sperling operator, in which the ride comfort index sequence is used to identify the train's operation condition. In addition, a novel technique that incorporates salient features of Bayesian inference and time series analysis is proposed for outlier detection and change detection. The Bayesian forecasting approach enables the prediction of conditional probabilities. By integrating the Bayesian forecasting approach with time series analysis, one-step forecasting probability density functions (PDFs) can be obtained before proceeding to the next observation. The change detection is conducted by comparing the current model and the alternative model (whose mean value is shifted by a prescribed offset) to determine which one can well fit the actual observation. When the comparison results indicate that the alternative model performs better, then a potential change is detected. If the current observation is a potential outlier or change, Bayes factor and cumulative Bayes factor are derived for further identification. A significant change, if identified, implies that there is a great alteration in the train operation performance due to defects. In this study, two illustrative cases are provided to demonstrate the performance of the proposed method for condition assessment of high-speed trains.

Keywords: high-speed train; in-service monitoring; condition assessment; Bayesian forecasting; time series analysis

1. Introduction

High-speed rail (HSR) is currently emerging as an environmentally friendly mode of transport that can provide comfortable and convenient intercity passenger services. Under good management planning and traffic flow, it is greatly beneficial for a huge volume of people to strengthen social networks and business activities. To maintain desirable service quality and operation safety, it is paramount to develop innovative monitoring techniques that not only target the challenging practical issues induced by train dynamics but also aim to enhance the system reliability and ride quality of high-speed trains.

Vibration is still an unavoidable problem to high-speed trains caused by rail irregularity, rail/wheel contact forces, vehicle dynamics, and aerodynamic instability (Remennikov and Kaewunruen 2008). These factors can be either correlated with or independent of each other. A two-stage suspension system (i.e., primary and secondary suspension systems) is usually adopted in high-speed trains. The efficacy of this double mechanical vibration isolation system depends on its damping and stiffness properties. Despite the rapid advancement of technology, vibration transmitted from bogies to car bodies via the suspension system is still a handicap. Although it is considered to pose only a mild risk of motion sickness given sufficiently good design and practice of train vehicles, vibration of passenger coaches can greatly affect ride comfort. In this connection, people exposed to train vehicle movements are fully suffering from dynamic and ambient influences, and both physiological and psychological components can be the determining factors. Several specifications such as GB5599 and UIC 513R, that stipulate the acceptable ride index and vibration level of railway vehicles, are available.

Researches available in the literature have reported that the vibration magnitude at a low-frequency range of 5-10 Hz is mainly induced by wheel-rail contact bouncing at two sides of abrasion concave (Huang *et al.* 2013). Moreover, low-frequency vibration can efficiently transfer energy from rail/wheel into car body via the suspension system (Wang *et al.* 2016), and such low-frequency vibration effects can lead to uncomfortable feeling of passengers (Kim *et al.* 2003). In

∗Corresponding author, Assistant Professor
E-mail: sk.lai@polyu.edu.hk
[a] Ph.D. Student
[b] Postdoctoral Fellow
[c] Professor and Director

http://www.techno-press.com/journals/sss&subpage=7
ISSN: 1738-1584 (Print), 1738-1991 (Online)

reality, various wheel conditions of train vehicles may result in different levels of ride discomfort that can act as an indicator reflecting the operation and service quality of high-speed trains. A vibration-based indicator in terms of the Sperling index has been widely used to describe the ride quality (Nakagawa *et al.* 2010).

The dynamic performance of high-speed trains is greatly affected by route status, rail/wheel conditions and ambient environments. In particular, wheels act as one of crucial components in high-speed trains; their operational behavior is decisive to the train's service life and running quality. It is known that abnormal impact loads can be generated by the defect of wheels, including wheel flats, wheel spalling, corrugation and polygonization (Vernersson 1999, Nielsen and Johansson 2000, Barke and Chiu 2005a, Remennikov and Kaewunren 2008). All these problems can be classified as out-of-roundness of wheels. Imperfections of wheel treads can bring a detrimental impact, in which the generated dynamic loads may cause serious damage to both tracks and vehicle components (Nielsen and Johansson 2000, Barke and Chiu 2005a). Most of the conventional inspection techniques used for detection of wheel defects are lack of accuracy and are usually conducted in an offline manner. A literature review summarizing various techniques in relation to fault/defect detection applications in railway engineering is available (Barke and Chiu 2005b).

In general, wayside detection systems can be pursued to provide a plenty of information for assessment of vehicle performance. For instance, the wheel impact monitor (WIM) as a non-contact method can quantify the extent of wheel defects. However, the modeling of forces exerted on tracks is required to be known before testing. In addition, the conventional inspection techniques for the detection of wheel defects are often one-off and cannot achieve long-term continuous monitoring. Some investigations have been devoted to formulating dynamic models for the simulation and detection of wheel defects (Bian *et al.* 2013, Alexandrou *et al.* 2016). Due to the high complexity of train system and rail structure, the model-based techniques are rather cumbersome and time-consuming. The adopted assumptions or simplifications in modeling can greatly affect the accuracy and reliability of the formulated models and model-based fault detection. With the emergence of advanced sensing technologies (Laory *et al.* 2013), the use of data-driven techniques (also referred to as model-free approaches) fully based on the monitoring data of system responses and environments has increased in popularity for health monitoring of structural systems in recent years. There are a number of merits that can be gained from structural health monitoring (SHM) systems (Spencer *et al.* 2004, Ko and Ni 2005, Brownjohn 2007, Lynch 2007, Ni *et al.* 2009, Ou and Li 2010).

Successful implementation of online SHM systems to various infrastructures has been reported, such as long-span bridges (Wong 2004, Ni *et al.* 2011, Yun *et al.* 2011), high-rise buildings (Lin *et al.* 2005, Kijewski-Correa *et al.* 2006, Ni *et al*. 2017), underground tunnels (Mohamad *et al.* 2012, Ding *et al.* 2013, Ye *et al.* 2013), and rail infrastructure (Barke and Chiu 2005b, Hu *et al.* 2015, Wang *et al.* 2018). The use of online SHM systems allows comparing static and dynamic monitoring data to provide a prompt evaluation of the structural condition (Niu *et al*. 2012, Zhang *et al*. 2015). Over the past two decades, a variety of SHM-based damage identification and condition assessment methods, such as artificial neural network (Ni *et al.* 2002, Geem *et al.* 2007), support vector machine (Yang *et al.* 2005, Bornn *et al.* 2009), statistical pattern recognition (Worden and Manson 2007, Sohn and Oh 2009), principal component analysis (Elangovan *et al.* 2011, Gharibnezhad *et al.* 2015), reliability-based evaluation (Catbas *et al.* 2008, Xia *et al.* 2012), Bayesian inference (Jiang and Mahadevan 2008, Kuok and Yuen 2012), and Gaussian process-based modeling (Dervilis *et al.* 2016, Wan and Ni 2018) have been proposed.

Ni *et al.* (2015) recently developed a Bayesian approach for successive evaluation of ride comfort, where the online monitoring data of acceleration responses are acquired from an on-board sensing system to evaluate the Sperling index evolutionarily. The Bayesian forecasting framework enables the prediction of conditional probabilities, while the Sperling index is a good indicator competent to quantify the running performance of high-speed trains. Hitherto, there is only a paucity of research devoted to the evaluation of ride comfort with the use of online monitoring data (Kim *et al.* 2004). In the present study, the integration of Bayesian forecasting approach with time series analysis is proposed for online condition assessment of an in-service high-speed train capitalizing on the monitoring data acquired from an on-board sensing system deployed on the train.

The dynamic behavior of an in-service train, which would be affected by the wheel quality, can be captured by the on-board sensing system deployed on it. The monitoring data used in the present study are those acquired from an instrumented in-service high-speed train before and after wheel lathing (a process of making out-of-round wheels perfectly round again in the depot), with the purpose to identify potential change in wheel quality on the basis of the monitoring-derived Sperling index sequence. The dynamic linear model (DLM) and Bayesian forecasting approach are applied to achieve this target. In the framework of Bayesian inference, both Bayes factor and cumulative Bayes factor are derived to determine whether the current observation is an outlier or a significant change due to the occurrence of wheel defects. The evaluation results by the proposed approach can be updated with reduced uncertainties when new monitoring data are available. The present work is also beneficial to scheduling the condition-based maintenance plan for high-speed train wheels.

2. Instrumentation system

An on-board sensing system comprising piezoelectric-type and optical fiber sensors, which are deployed at both motor and trailer bogies, axle boxes, gearboxes and floor of coaches for acceleration, strain, temperature and noise measurement, has been implemented on a high-speed passenger train to continuously collect online monitoring data during its routine operation (Wang *et al.* 2016). Fig. 1

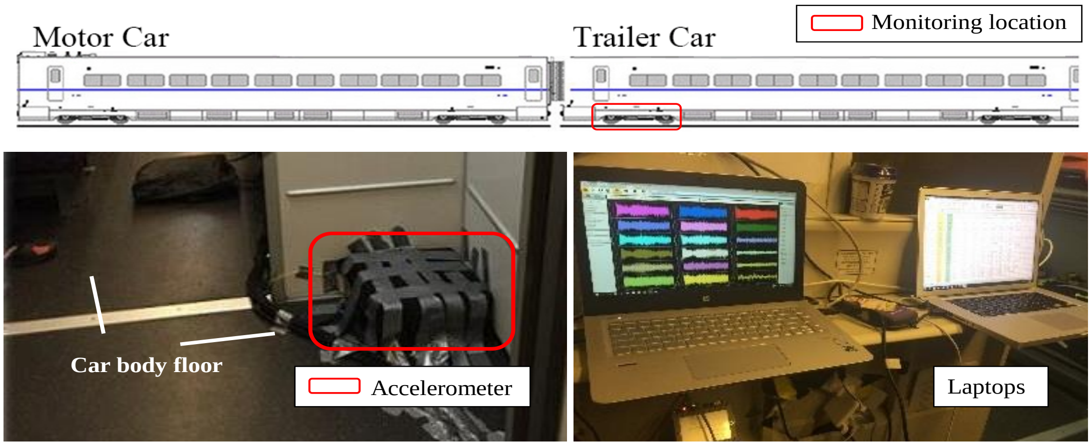

Fig. 1 Accelerometers and data acquisition unit for online monitoring of an in-service high-speed train

illustrates the accelerometers deployed at the floor level of a coach and the data acquisition unit during the monitoring. In accordance with the specifications (China National Bureau of Standards 1985, International Union of Railways 1994), the acceleration data collected at different locations of the coach floor are used to calculate the Sperling index and evaluate the ride quality in real time. During the on-board monitoring, the train was operated at maximum speeds of 200 to 250 km/hr.

The acceleration responses in both lateral and vertical directions collected at the interior floor of a trailer car are used in this study. The acceleration data were acquired at a sampling rate of 1000 Hz. By calculating at each time slot the Sperling index from the measured accelerations, the ride quality index sequence is obtained which characterizes the operation condition of the high-speed train traveling along the whole route. To verify the effectiveness of the proposed Bayesian forecasting and time series analysis procedure for train condition assessment, the acceleration monitoring data acquired from the train before and after wheel lathing are used in the study.

3. Methodology

The proposed method for online condition monitoring of in-service trains is presented in the sub-sections below. Following this method, the online monitoring data of accelerations acquired from the in-service train at each time segment are synthesized to obtain the Sperling index sequence. Next, a time series analysis is employed to model the dynamic process of the comfort index evolution over time for the purpose of eliciting a continuous monitoring and assessment procedure. To provide an early warning on damage or deterioration prior to structure failure or costly repair, Bayesian forecasting is incorporated into time series analysis to diagnose potential anomaly of the operation status. As such, a built-in detection algorithm is constructed. When a significant change is detected, potential defects of the train are alarmed.

3.1 Sperling index

The Sperling index is a specific indicator that can be used to quantify the comfort feeling of passengers due to external factors of a running train vehicle (Nakagawa *et al.* 2010). This index is capable of reflecting the running train status under various operation conditions. Making use of the acquired acceleration data, the Sperling index can be calculated using Eq. (1). In general, the smaller the values are, the better the ride quality is. The formula for computing the Sperling index is (Zhou *et al.* 2009)

$$W=\left(W_1^{10}+W_2^{10}+W_3^{10}\cdots+W_n^{10}\right)^{0.1}=\left(\sum_{i=1}^{n}W_i^{10}\right)^{0.1} \tag{1}$$

where n is the number of frequency contents $\{f_1, f_2, f_3, \ldots, f_n\}$ of the measured time-domain acceleration after applying the fast Fourier transform (FFT); W_i stands for the Sperling index value of each contributed component:

$$W_i=7.08\left(\frac{A_i}{f_i}F(f_i)\right)^{0.1} \tag{2}$$

where A_i is the acceleration at the frequency f_i (Hz) and $F(f_i)$ is the frequency modification factor that is a piecewise-linear function of f_i (China National Bureau of Standards 1985). The Sperling index is obtained for every 100 seconds to form a ride quality index sequence which represents the operation condition of the monitored train in each time segment as illustrated in Fig. 2.

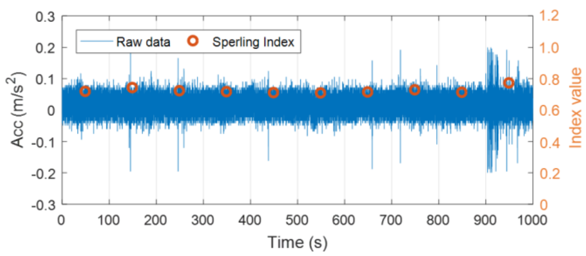

Fig. 2 Original signals and obtained Sperling indices in different time segments

3.2 Dynamic linear model

The dynamic linear model (DLM) (West and Harrison 1999, Petris *et al.* 2009) can be regarded as a special case of generic state-space models. It is commonly used for time series analysis in the context of Bayesian inference. The DLM is an integration of the state equation given in Eq. (3) with varying parameters and the observation equation given in Eq. (4):

$$Y_t = F_t^T \theta_t + v_t, \qquad v_t \sim N(0, V_t) \tag{3}$$

$$\theta_t = G_t \theta_{t-1} + \omega_t, \qquad \omega_t \sim N(0, W_t) \tag{4}$$

where F_t is a $p\times1$ vector of known parameters; G_t is a $p\times p$ matrix of known constants; θ_t is a $p\times1$ vector of unknown parameters; v_t is an observation noise term to represent the measurement error that corrupts the observation Y_t; ω_t is an evolution noise term to denote a stochastic change in the state vector θ_t; $\{V_t\}$ and $\{W_t\}$ are two independent sequences of independent Gaussian random vectors with mean zero and known variance; and Y_t is the observation series at time t. With Eqs. (3) and (4), the evolving relationship between the measurement and unknown state parameters can be obtained.

In this study, the Sperling index sequence obtained using the measured vibration data is viewed as the measurement in the DLM formulation, where a second-order DLM is employed to model the dynamic process. The underlying rationale is to decompose a time series with the observed measurement into several mathematical elements including values and gradients (Lipowsky *et al.* 2010). Thus, Eqs. (3) and (4) can be rewritten as

$$Y_t = \mu_t + v_t, \qquad v_t \sim N(0, \sigma_{obs}^2) \tag{5}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \omega_{1t}, \qquad \omega_{1t} \sim N(0, \sigma_{level}^2) \tag{6}$$

$$\beta_t = \beta_{t-1} + \omega_{2t}, \qquad \omega_{2t} \sim N(0, \sigma_{trend}^2) \tag{7}$$

where the state vector has two elements, μ_t and β_t, which denote the current level of the parameter values and the current rate of change of the parameters (i.e., gradients), respectively. In the form of a state-space model, they can be written as

$$\theta_t = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}, \quad F_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad G_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad W_t = \begin{pmatrix} \sigma_{level}^2 & 0 \\ 0 & \sigma_{trend}^2 \end{pmatrix}, \quad V_t = \sigma_{obs}^2 \tag{8}$$

where σ_{level} and σ_{trend} are the variances of the parameter values and the change rate of the parameters, respectively. The measurement variance is equivalent to the square of its standard deviation σ_{obs}.

3.3 Bayesian forecasting

Bayesian inference is a promising approach for statistical analysis. It can be derived from the well-known Bayes' rule by computing the posterior probability of the model parameters from prior information and measurement mixtures. Using the Bayesian inference, it is easy to update a previous evaluation with a new likelihood when new observation data are available. To integrate the Bayesian inference into time series analysis, the following notation is introduced

$$P(\theta_t | D_t) \sim N(m_t, C_t), \quad D_t = Y_1, Y_2, ..., Y_t \tag{9}$$

where $\{Y_1, Y_2, \ldots, Y_t\}$ stand for the observations.

The above equation presents the PDF of the state parameter θ at time t based on the measurements. This PDF is normally distributed with a mean of m and a variance of C. If the state at time t is given, the posterior distribution of the state parameter and the prediction for one-step forecast at time t+1 can be obtained by the following equations (Petris *et al.* 2009)

$$P(\theta_{t+1} | D_t) \sim N(a_{t+1}, R_{t+1}) \tag{10}$$

$$P(Y_{t+1} | D_t) \sim N(f_{t+1}, Q_{t+1}) \tag{11}$$

where $P(\theta_{t+1} | D_t)$ and $P(Y_{t+1} | D_t) \sim N(f_{t+1}, Q_{t+1})$ are the posterior distribution of the state parameters and the one-step forecasting distribution, respectively. The means and variances of the above two distributions can be calculated as follows

$$a_{t+1} = G_{t+1} \cdot m_t \tag{12}$$

$$f_{t+1} = F_{t+1}^T \cdot a_{t+1} \tag{13}$$

$$R_{t+1} = G_{t+1} C_t G_{t+1}^T + W_{t+1} \tag{14}$$

$$Q_{t+1} = F_{t+1}^T R_{t+1} F_{t+1} + V_{t+1} \tag{15}$$

If we obtain the measurement Y_{t+1} at time t+1 in this inference process, the corrections are due to Eq. (9) and the distribution of the state parameters is updated to $P(\theta_{t+1}|D_{t+1}) \sim N(m_{t+1}|C_{t+1})$ where the mean and the variance are given by

$$m_{t+1} = a_{t+1} + A_{t+1} e_{t+1} \tag{16}$$

$$C_{t+1} = R_{t+1} - A_{t+1} A_{t+1}^T Q_{t+1} \tag{17}$$

where $e_{t+1}=Y_{t+1}-f_{t+1}$ and $A_{t+1}=R_{t+1}F_{t+1}/Q_{t+1}$. In the subsequent sub-sections, a special logic called the Bayes factor will be adopted to identify outliers and significant changes.

3.4 Detection algorithm

3.4.1 Outlier detection

In the present study, the detection logic is based on the calculation of Bayes factor to identify potential outliers and defects. As mentioned before, the Bayesian forecasting approach enables the generation of PDFs for the next observation. In this process, the detection is carried out by comparing the measurement with the current model (forecast distribution for the current time, denoted as “Model 0”) and the alternative model (achieved by shifting a prescribed offset of mean value h of the current model, denoted as “Model 1”). Thus the Bayes factor is obtained from the ratio of the two models as

$$H_t = \frac{\text{PDF value of Model 1}}{\text{PDF value of Model 0}} \tag{18}$$

Since the PDFs of both models are normally distributed, the Bayes factor can be further obtained as

$$H_t = \exp\left(\frac{\pm 2h \cdot (Y_t - f_t) - h^2}{2Q_T^2}\right) \tag{19}$$

where H_t is a monotonic straight line in terms of a logarithmic scale, which means that a larger Bayes factor will provide a better fit between the measurement and the alternative model. In the case of H_t =1, it implies that the probability measures from both models are equal. On the other hand, the plus and minus signs in Eq. (19) are used to detect the outliers with positive and negative deviations, respectively (Lipowsky *et al.* 2010). According to Jeffreys (1961), the threshold value for outlier detection can be set as H_{min} = 10. However, the shift value of h should be determined in accordance with the required confidence level. In this study, $h = 1.645\sigma_t$ at a 90% confidence level is used. Thus, an uncertainty limit (*ucl*) can be determined by the following equation

$$ucl = \frac{\ln(H_{\min})}{h}\sigma_t^2 + \frac{h}{2} \tag{20}$$

In the case of $h = 1.645\sigma_t$ and H_{min} = 10, the uncertainty limit is equal to $2.22\sigma_t$. The observation would be identified as an outlier when its deviation from the mean value of the current model is larger than *ucl*.

3.4.2 Change (defect) detection

To discriminate between outliers and significant changes, the cumulative Bayes factor is developed, which is defined as the product of k consecutive Bayes factors

$$H_t(k) = \prod_{t-k+1}^{t} H_t \qquad k = 1, 2, \ldots, l_{\max} \tag{21}$$

where $l_{\max}$ denotes the maximum number of Bayes factors considered. The maximum cumulative Bayes factor is calculated by the following equation

$$L_t = H_t(l_t) = \max(H_t(k)) \qquad 1 \le l_t \le l_{\max} \tag{22}$$

where l_t refers to the run length by counting the number of recent and consecutive observations that contribute to the maximum value of L_t. It is recursively calculated by

$$l_t = \begin{cases} 1, & if \quad L_{t-1} \le 1 \\ 1 + l_{t-1}, & if \quad L_{t-1} > 1 \end{cases} \tag{23}$$

The threshold value of l_t is $l_{\min} = 4$, as suggested by Pole *et al.* (1994). In this procedure, a notification of change at time t when $L_t > H_{\min}$ is denoted as the time of notification (i.e., $TON = t$). When a change is detected and recorded, it is defined as the time of change occurrence (i.e., $TOC = t - l_t + 1$).

In short, there are two steps in the outliner identification and change (defect) detection. The first step is to calculate the Bayes factor and to judge if the observation is an outlier or not. The criterion for catching an outlier is presented as $H_t > H_{\min}$ and $H_{t-1} \le H_{\min}$. If this condition is not satisfied, the procedure will turn to the change (defect) detection. In the second step, the cumulative Bayes factor L_t and the run length l_t will be calculated. In contrast to the outlier detection, the significant change (defect) is triggered by the following two criteria (Lipowsky *et al.* 2010):

(i) The occurrence of two consecutive Bayes factors is $H_t > H_{\min}$, which is equivalent to $L_t > H_{\min}^2$; and

(ii) The concurrence of $L_t > H_{\min}$ and $l_t > l_{\min}$ is resulted.

Once a change is detected, the retrospective refinement will be conducted by re-setting the time to *TOC* and by adjusting the mean value of Model 0 to the measurement ($f_t = Y_t$). The flowchart of the algorithm is given in Fig. 3.

4. Analysis results and discussion

The measured vibration signals of the car body at the floor level are processed by the proposed technique. In terms of the Sperling index, the original data collected are converted into the ride quality index sequence at time intervals of 100s, which reflects the operation condition of the high-speed train. In the assessment procedure, a second-order DLM is formulated to model the dynamic process.

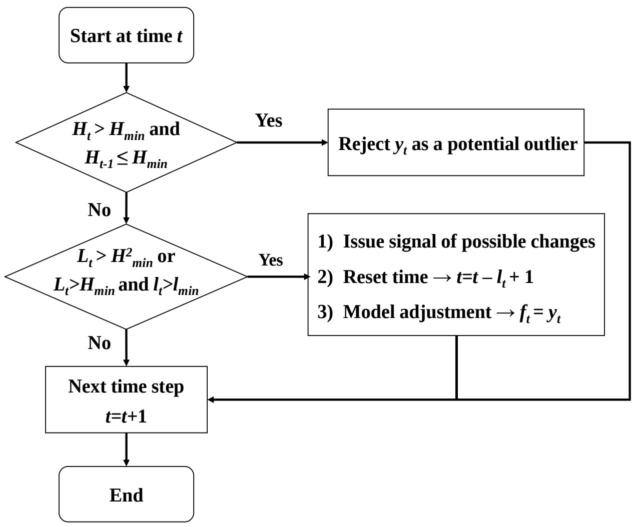

Fig. 3 Flowchart of the proposed algorithm

Then, the Bayesian forecasting approach is applied to conduct one-step prediction prior to the next observation. Two metrics, Bayes factor and cumulative Bayes factor, are worked out for outlier detection and change detection. The vibration signals acquired from the car body under different wheel quality conditions (before and after wheel lathing) are used. The proposed procedure is applied to identify the alteration in operation performance stemming from wheel defects. Figs. 4 and 5 show two cases in accordance with the monitoring data acquired from different intervals of the rail line.

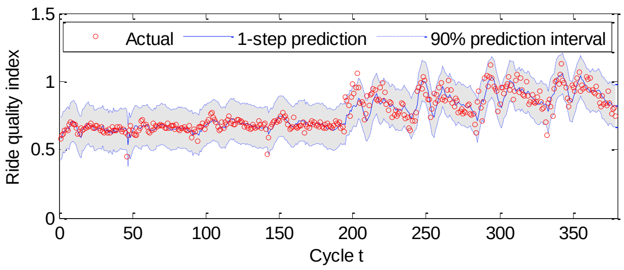

Fig. 4 Original model for ride quality index prediction in Case 1

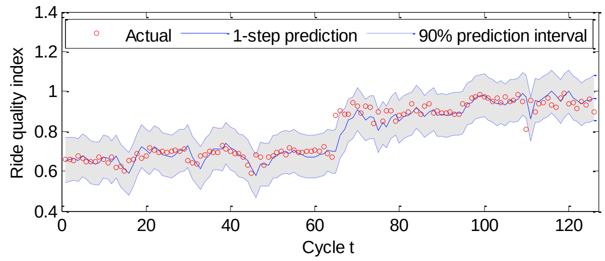

Fig. 5 Original model for ride quality index prediction in Case 2

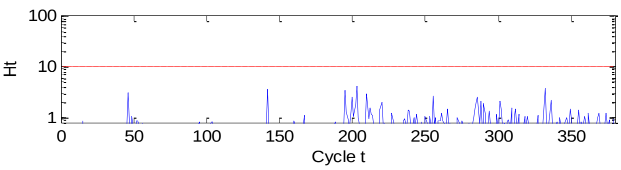

Fig. 6 Bayesian dynamic linear model for outlier detection in Case 1

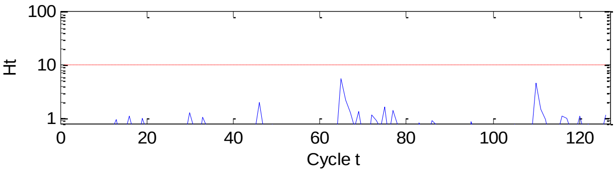

Fig. 7 Bayesian dynamic linear model for outlier detection in Case 2

The condition assessment results by applying the proposed procedure to the above two sets of sequences are provided in Figs. 6 to 11. As an example, the data at the time interval $t \in [0, 63]$ shown in Fig. 5 were measured from the train after wheel lathing, while the other correspond to the status before wheel lathing. In applying the proposed procedure, the shift parameter is set as $h = 1.645\sigma$, resulting in an uncertainty limit of $ucl = 2.22\sigma$.

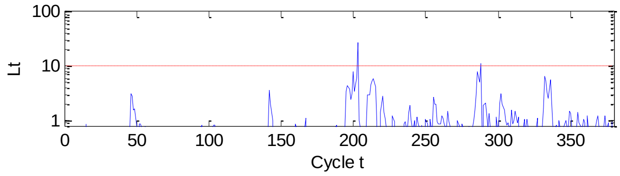

Fig. 8 Bayesian dynamic linear model for change detection in Case 1

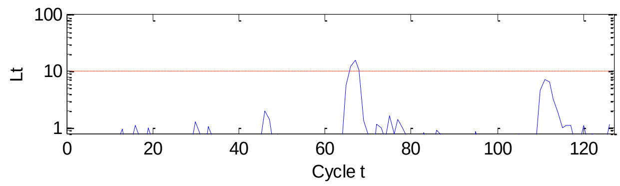

Fig. 9 Bayesian dynamic linear model for change detection in Case 2

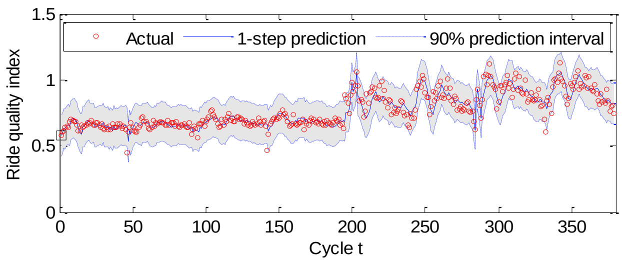

Fig. 10 Refined model for ride quality index prediction in Case 1

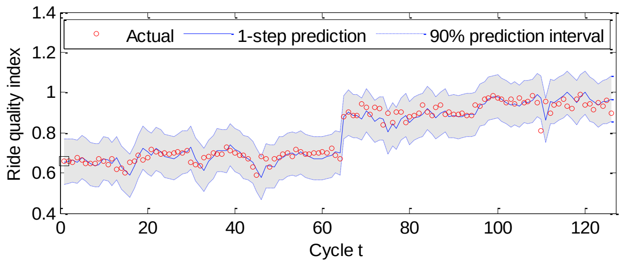

Fig. 11 Refined model for ride quality index prediction in Case 2

In Figs. 4 and 5, the actual values (red points) in terms of the Sperling indices obtained at equal time intervals are from the original vibration signals, while the results of one-step forecasting (guess of mean value) obtained by DLM are presented by a blue solid line and a 90% prediction interval is denoted by a grey shadow. The Bayes factor and the accumulative Bayes factor are used for outlier detection and change detection, respectively. The detection results for the above two cases are shown in Figs. 6 to 9. It is observed that, although some observations fall outside of the 90% prediction interval, the corresponding Bayes factors are within the threshold of $H_{\min} = 10$ as shown in Figs. 6 and 7, implying that there is no outlier detected. This observation can be explained since the measurements are held by a 95% prediction interval induced by setting $H_{\min} = 10$ ($ucl = 2.22\sigma$) in the outlier detection.

On the other hand, Figs. 8 and 9 show the results of the cumulative Bayes factor for change detection. It can be seen that the evaluation results exceed the threshold line of $L_{\min} = 10$ at $t = 201$ (i.e., $TON_1 = 201$) and $t = 67$ (i.e., $TON_2 = 67$) in the two cases, respectively. Meanwhile, the corresponding run lengths for the two cases are l_t (= 6) > l_{min} (= 4) and l_t (= 5) > l_{min} (= 4), respectively. In other words, both situations fulfill the conditions as there is a significant change. The identified occurrence time instants of change in these two cases are $t = 196$ (i.e., $TOC_1 = TON_1 – l_t+1 = 196$) and $t = 63$ (i.e., $TOC_2 = TON_2 – l_t+1 = 63$), respectively. These results are in good agreement with the actual time point of inflicting wheel lathing.

In accordance with the real-time identification of change occurrence, the model refinement by altering the guessed mean value is carried out by the proposed strategy. The newly predicted results (blue solid line) are shown in Figs. 10 and 11. It is clearly observed that the prediction at the time of change occurrence in the refined models agrees well with the actual measurement, and it offers better performance than the original model.

5. Conclusions

Operation safety and ride comfort are imperiled by the technical demands for increasing the running speed of high-speed trains, in which the quality of wheels is a dominant factor affecting the running stability. Besides, the dynamic response of high-speed trains is also subject to the influence of ambient conditions. Integrating a data-driven strategy for online monitoring and assessment enables engineers to trace the health status of high-speed trains. To achieve this target, it is highly desired to develop innovative techniques for real-time condition identification and assessment.

In the present work, an on-board sensing system has been installed on an in-service high-speed train to collect a variety of monitoring data during the routine operation. A vibration-based indicator in terms of the Sperling index is adopted for pre-processing the measured signals. The obtained ride quality index sequences are used to indicate the operational condition of the train. A novel technique that integrates Bayesian forecasting with time series analysis is used to conduct a probabilistic assessment of the train condition. Two metrics, Bayes factor and cumulative Bayes factor, are applied respectively for outlier detection and change detection. The present study is demonstrated by using the monitoring data acquired under different wheel conditions (before and after wheel lathing). Two illustrative cases are presented to examine the effectiveness of the proposed method.

Acknowledgments

The work described in this paper was supported by the funding from the Research Grants Council of the Hong Kong Special Administrative Region, China (Grant No. PolyU 152024/17E). The authors would also like to appreciate the funding support by the Ministry of Science and Technology of China (Grant No. 2018YFE0190100), the Innovation and Technology Commission of Hong Kong SAR Government (Grants No. K-BBY1 and 1-BBYJ) to the Hong Kong Branch of Chinese National Rail Transit Electrification and Automation Engineering Technology Research Center and by the National Natural Science Foundation of China (Grant No. 11602210).

References

Alexandrou, G., Kouroussis, G. and Verlinden, O. (2016), “A comprehensive prediction model for vehicle/track/soil dynamic response due to wheel flats”, *Proc. Instit. Mech. Eng., Part F: J. Rail Rapid Transit*, **230**(4), 1088-1104.

Barke, D.W. and Chiu, W. K. (2005a), “A review of the effects of out-of-round wheels on track and vehicle components”, *Proc. Instit. Mech. Eng., Part F: J. Rail Rapid Transit*, **219**(3), 151-175.

Barke, D.W. and Chiu, W.K. (2005b), “Structural health monitoring in the railway industry: A review”, *Struct. Health Monit.*, **4**(1), 81-93.

Bian, J., Gu, Y. and Murray, M. (2013), “Numerical study of impact forces on railway sleepers under wheel flat”, *Adv. Struct. Eng.*, **16**(1), 127-134.

Bornn, L., Farrar, C.R., Park, G. and Farinholt, K. (2009), “Structural health monitoring with autoregressive support vector machines”, *J. Vib. Acoust.*, **131**(2), 021004.

Brownjohn, J.M. (2007), “Structural health monitoring of civil infrastructure”, *Philos. Trans. A Math. Phys. Eng. Sci.*, **365**(1851), 589-622.

Catbas, F.N., Susoy, M. and Frangopol, D.M. (2008), “Structural health monitoring and reliability estimation: Long span truss bridge application with environmental monitoring data”, *Eng. Struct.*, **30**(9), 2347-2359.

China National Bureau of Standards (1985), *Railway Vehicles Specification for Evaluation of Dynamic Performance and Accreditation Test*, GB5599, Beijing, China.

Dervilis, N., Shi, H., Worden, K. and Cross, E.J. (2016). “Exploring environmental and operational variations in SHM data using heteroscedastic Gaussian processes”, *Dynamics of Civil Structures*, (Eds., S. Pakzad, and C. Juan), Springer, Cham.

Ding, L.Y., Zhou, C., Deng, Q.X., Luo, H.B., Ye, X.W. and Ni, Y. Q., Guo, P. (2013), “Real-time safety early warning system for cross passage construction in Yangtze riverbed metro tunnel based on the internet of things”, *Automat. Constr.*, **36**, 25-37.

Elangovan, M., Devasenapati, S.B., Sakthivel, N.R. and Ramachandran, K I. (2011), "Evaluation of expert system for condition monitoring of a single point cutting tool using principle component analysis and decision tree algorithm", *Exp. Syst. Appl.*, **38**(4), 4450-4459.

Geem, Z.W., Tseng, C.L., Kim, J. and Bae, C. (2007), "Trenchless water pipe condition assessment using artificial neural network", *Pipelines 2007: Advances and Experiences with Trenchless Pipeline Projects*, 1-9.

Gharibnezhad, F., Mujica, L.E. and Rodellar, J. (2015), "Applying robust variant of principal component analysis as a damage detector in the presence of outliers", *Mech. Syst. Signal Pr.*, **50-51**, 467-479.

Hu, Q., Lam, H.F. and Alabi, S.A. (2015), "Use of measured vibration of in-situ sleeper for detecting underlying railway ballast damage", *Int. J. Struct. Stab. Dyn.*, **15**(8), 1540026.

Huang, Z.W., Cui, D.B., Du, X. and Jin, X.S. (2013), "Influence of deviated wear of wheel on performance of high-speed train running on straight tracks", *J. China Railw. Soc.*, **35**(2), 14-20.

International Union of Railways (1994), *Guidelines for Evaluating Passenger Comfort in Relation to Vibration in Railway Vehicles*, UIC 513R, Paris, France.

Jeffreys, H. (1961), *Theory of Probability*, Clarendon Press, Oxford, UK.

Jiang, X. and Mahadevan, S. (2008), "Bayesian wavelet methodology for structural damage detection", *Struct. Control Health Monit.*, **15**(7), 974-991.

Kijewski-Correa, T., Kilpatrick, J., Kareem, A., Kwon, D.K., Bashor, R., Kochly, M., Young, B.S., Abdelrazaq, A., Galsworthy, J., Isyumov, N., Morrish, D., Sinn, R.C. and Baker, W.F. (2006), "Validating wind-induced response of tall buildings: Synopsis of the Chicago full-scale monitoring program", *J. Struct. Eng.*, **132**(10), 1509-1523.

Kim, Y.G., Kwon, H.B., Kim, S.W., Kim, C.K. and Kim, T.W. (2003), "Correlation of ride comfort evaluation methods for railway vehicles", *Proc. Instit. Mech. Eng., Part F: J. Rail Rapid Transit*, **217**(2), 73-88.

Kim, Y.G., Kim, S.W., Park, C.K., Kim, K.H. and Park, T.W. (2004), "Analysis on the characteristics of the ride comfort for high speed trains on the high-speed line/conventional line", *Transactions of the Korean Society for Noise and Vibration Engineering*, **14**(10), 999-1006.

Ko, J.M. and Ni, Y.Q. (2005), "Technology developments in structural health monitoring of large-scale bridges", *Eng. Struct.*, **27**(12), 1715-1725.

Kuok, S.C. and Yuen, K.V. (2012), "Structural health monitoring of Canton Tower using Bayesian framework", *Smart Struct. Syst.*, **10**(4), 375-391.

Laory, I., Trinh, T.N., Posenato, D. and Smith, I.F. (2013), "Combined model-free data-interpretation methodologies for damage detection during continuous monitoring of structures", *J. Comput. Civ. Eng.*, **27**(6), 657-666.

Lin, C. C., Wang, C. E., Wu, H. W. and Wang, J. F. (2005), "On-line building damage assessment based on earthquake records", *Smart Mater. Struct.*, **14**(3), S137-S153.

Lipowsky, H., Staudacher, S., Bauer, M. and Schmidt, K.J. (2010), "Application of Bayesian forecasting to change detection and prognosis of gas turbine performance", *J. Eng. Gas Turb. Power*, **132**(3), 031602.

Lynch, J.P. (2007), "An overview of wireless structural health monitoring for civil structures", *Philos. T. A Math. Phys. Eng. Sci.*, **365**(1851), 345-372.

Mohamad, H., Soga, K., Bennett, P.J., Mair, R.J. and Lim, C.S. (2012), "Monitoring twin tunnel interaction using distributed optical fiber strain measurements", *J. Geotech. Geoenviron. Eng.*, **138**(8), 957-957.

Nakagawa, C., Shimamune, R., Watanabe, K. and Suzuki, E. (2010), "Fundamental study on the effect of high frequency vibration on ride comfort", *J. Mech. Syst. Transp. Logist.*, **3**(1), 287-293.

Ni, Y.Q., Lin, K.C., Wu, L.J. and Wang, Y.W. (2017), "Visualized spatiotemporal data management system for lifecycle health monitoring of large-scale structures", *J. Aerosp. Eng.*, **30**(2), B4016007.

Ni, Y.Q., Liu, X.Z., Zhao, W.Z. and Liang, S.L. (2015), "Outlier detection in sensor-assisted online ride comfort assessment of high-speed trains", *Proceedings of the 10th International Workshop on Structural Health Monitoring*, Stanford, California, USA.

Ni, Y. Q., Wang, B. S. and Ko, J. M. (2002), "Constructing input vectors to neural networks for structural damage identification", *Smart Mater. Struct.*, **11**(6), 825-833.

Ni, Y.Q., Wong, K.Y. and Xia, Y. (2011), "Health checks through landmark bridges to sky-high structures", *Adv. Struct. Eng.*, **14**(1), 103-119.

Ni, Y. Q., Xia, Y., Liao, W. Y. and Ko, J. M. (2009), "Technology innovation in developing the structural health monitoring system for Guangzhou New TV Tower", *Struct. Control Health Monit.*, **16**(1), 73-98.

Nielsen, J.C. and Johansson, A. (2000). "Out-of-round railway wheels-a literature survey", *Proc. Instit. Mech. Eng., Part F: J. Rail Rapid Transit*, **214**(2), 79-91.

Niu, Y., Fritzen, C.P., Jung, H., Buethe, I., Ni, Y.Q. and Wang, Y. W. (2015), "Online simultaneous reconstruction of wind load and structural responses – Theory and application to Canton Tower", *Comput.-Aided Civ. Infrastruct. Eng.*, **30**(8), 666-681.

Ou, J. and Li, H. (2010), "Structural health monitoring in mainland China: review and future trends", *Struct. Health Monit.*, **9**(3), 219-231.

Petris, G., Petrone, S. and Campagnoli, P. (2009), *Dynamic Linear Models with R*, Springer, Berlin, Germany.

Pole, A., West, M. and Harrison, J. (1994), *Applied Bayesian Forecasting and Time Series Analysis*, CRC Press, Boca Raton, Florida, USA.

Remennikov, A. M. and Kaewunruen, S. (2008), "A review of loading conditions for railway track structures due to train and track vertical interaction", *Struct. Control Health Monit.*, **15**(2), 207-234.

Sohn, H. and Oh, C.K. (2009), "Statistical pattern recognition", *Encyclopedia of Structural Health Monitoring*, (Eds., C. Boller, F.-K. Chang and Y. Fujino), John Wiley & Sons, Chichester, UK.

Spencer, B.F. Jr., Ruiz-Sandoval, M.E. and Kurata, N. (2004), "Smart sensing technology: opportunities and challenges", *Struct. Health Monit.*, **11**(4), 349-368.

Vernersson, T. (1999), "Thermally induced roughness of tread-braked railway wheels: Part 1: brake rig experiments", *Wear*, **236**(1), 96-105.

Wan, H.P. and Ni, Y.Q. (2018), "Bayesian modeling approach for forecast of structural stress response using structural health monitoring data", *J. Struct. Eng.*, DOI: 10.1061/(ASCE)ST.1943-541X.0002085.

Wang, J., Liu, X.Z. and Ni, Y.Q. (2018), "A Bayesian probabilistic framework for acoustic emission based rail condition assessment", *Comput.-Aided Civ. Infrastruct. Eng.*, **33**(1), 21-34.

Wang, X, Ni, Y.Q, Zhang, L.H. and Sun, Q. (2016), "Understanding of dynamic interaction of an in-service high-speed train via on-board monitoring", Presented on *the 1st International Workshop on Structural Health Monitoring for Railway System*, Qingdao, China.

West, M. and Harrison, J. (1997), *Bayesian Forecasting and Dynamic Models*, Springer-Verlag, New York, USA.

Wong, K.Y. (2004), "Instrumentation and health monitoring of cable-supported bridges", *Struct. Control Health Monit.*, **11**(2), 91-124.

Worden, K. and Manson, G. (2007), "The application of machine learning to structural health monitoring", *Philos. T, A Math. Phys. Eng. Sci.*, 365(1851), 515-537.

Xia, H.W., Ni, Y.Q., Wong, K.Y. and Ko, J.M. (2012), "Reliability-based condition assessment of in-service bridges using mixture distribution models", *Comput. Struct.*, **106-107**, 204-213.

Yang, B.S., Hwang, W.W., Kim, D.J. and Tan, A.C. (2005), "Condition classification of small reciprocating compressor for refrigerators using artificial neural networks and support vector machines", *Mech. Syst. Signal Pr.*, **19**(2), 371-390.

Ye, X.W., Ni, Y.Q. and Yin, J.H. (2013), "Safety monitoring of railway tunnel construction using FBG sensing technology", *Adv. Struct. Eng.*, **16**(8), 1401-1409.

Yun, C.B., Lee, J.J. and Koo, K.Y. (2011), "Smart structure technologies for civil infrastructures in Korea: recent research and applications", *Struct. Infrastruct. E.*, **7**(9), 673-688.

Zhang, F.L., Ni, Y.Q., Ni, Y.C. and Wang, Y.W. (2016), "Operational modal analysis of Canton Tower by a fast frequency domain Bayesian method", *Smart Struct. Syst.*, **17**(2), 209-230.

Zhou, J., Goodall, R., Ren, L. and Zhang, H. (2009), "Influences of car body vertical flexibility on ride quality of passenger railway vehicles", *Proc. Instit. Mech. Eng., Part F: J. Rail Rapid Transit*, **223**(5), 461-471.