High-Speed Rail Suspension System Health Monitoring Using Multi-Location Vibration Data

Ning Hong, Lishuai Li, Weiran Yao, Yang Zhao, Cai Yi, Jianhui Lin, and Kwok Leung Tsui

***Abstract*— A novel data-driven framework to monitor the health status of high-speed rail suspension system by measuring train vibrations is proposed herein. Unlike existing methods, this framework does not rely on sophisticated dynamic models or high-fidelity simulations; it combines the power of data and domain knowledge to generate a model that can be trained quickly and adapted easily to different rail systems. In addition, the framework includes a module to generate a training dataset, tackling a typical challenge in real-world system monitoring, namely, the lack of labeled data due to practical limits. Based on the multi-output support vector regression (MSVR), the proposed framework can monitor the stiffness and damping coefficients of the suspension system using vibration signals measured on trains in real time. The framework comprises three modules. First, a simple suspension system dynamics model is built to generate a training dataset. Furthermore, key features are extracted from frequency response curves to reflect the impact of spring and damper degradation. Subsequently, a supervised learning model based on the MSVR is built to predict the stiffness and damping coefficients of suspension systems from features extracted in the second module. Once the model is built, real-time monitoring can be achieved by feeding the vibration signals as they are collected during operations. The proposed framework was evaluated on simulation data for its accuracy and tested on real-world operational data for its practicability.**

***Index Terms*— Health monitoring, suspension system, high-speed rail, multi-output support vector regression, vibration signal, field data, data-driven approach.**

I. Introduction

As HIGH-SPEED railway transports are developing at a considerable speed in many regions worldwide, especially China, the safety and reliability issues of rolling stocks are receiving increasing attention. Because of the higher speed, more intensive use, and more complex structures compared to ordinary trains, safety concerns are accentuated for high-speed trains. The possible failure of train structures during their operation may cause immense loss of life or monetary damage, resulting in irretrievable and catastrophic consequences [1]. Therefore, the current maintenance strategy of high-speed rail systems is highly conservative. Maintenance events are scheduled based on fixed intervals (i.e., running mileage or time) that include a large safety margin, causing increased operational cost.

Real-time health monitoring for high-speed rail systems could address this issue – maintaining a high standard of safety while reducing cost. It could provide an early warning of a fault or component degradation, avoid in-service failures, and schedule maintenance when required. Therefore, unnecessary maintenance can be reduced significantly. Furthermore, operations can be optimized based on health monitoring results. In support of real-time health monitoring development, this study develops a data-driven framework to monitor the status of methods for the suspension system of high-speed rail systems. We focus on the suspension system because it is fundamental in rail vehicles and its reliability is directly related to vehicle safety. It supports the bogie and car body, reduces the forces generated by the track unevenness on the wheels, and controls the behavior of the car body with respect to the track surface for rider comfort.

Many studies have been performed to develop methods for the condition monitoring or fault detection of railway suspension systems, where two distinctive approaches are primarily applied: the model-based approach and the data-driven approach. Surveys on the modern techniques used for the health monitoring of railway vehicle dynamics are provided in [2]–[4]. Most model-based methods transform the vehicle's equations of motion into state-space representations. Subsequently, the changes of the model's parameters (stiffness or damping) can be detected using a Kalman filter as well as its variations [5]–[10]. A dynamic model of the suspension system for a three-car CRH2 (China Railway High-speed 2) multiple unit was established and an improved total measurable fault information residual method was proposed to estimate the health condition of vehicles in [11]. An interacting multiple-model algorithm was introduced for detecting faults in suspension systems in [12]. The parameter estimation method applies a particle filter to monitor the health condition of the components [13]. A number of other model-based

Manuscript received December 12, 2018; revised April 12, 2019; accepted May 31, 2019. Date of publication June 20, 2019; date of current version June 29, 2020. This work was supported in part by the Hong Kong Research Grants Council Theme-Based Research Scheme under Grant T32-101/15-R, and in part by the Early Career Scheme under Project 21202716. The Associate Editor for this paper was M. Zhou. *(Corresponding author: Lishuai Li.)*

N. Hong, L. Li, Y. Zhao, and K. L. Tsui are with the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong (e-mail: neil.hong@my.cityu.edu.hk; yangzhao9-c@my.cityu.edu.hk; lishuai.li@cityu.edu.hk; kltsui@cityu.edu.hk).

W. Yao is with the Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: weiran@cmu.edu).

C. Yi is with the School of Automobile and Transportation, Xihua University, Chengdu 610039, China (e-mail: justin.yi@163.com).

J. Lin is with the Traction Power State Key Laboratory, Southwest Jiaotong University (SWJTU), Chengdu 614200, China (e-mail: lin13008104673@126.com).

Digital Object Identifier 10.1109/TITS.2019.2921785

approaches have been proposed for rail vehicle suspension system fault detection or health monitoring in [14]–[20]. Although these model-based methods have proven their diagnostic usefulness in simulation studies, they require precise suspension and inertial parameter values. This causes two problems: (i) they are cumbersome to use when being applied to different rail systems; (ii) the results of the model-based methods could be primarily distorted if the parameter values are inaccurate. Another issue of the model-based approach is that the development of an appropriate model requires the detailed knowledge of the system concerned. For systems that are dynamically complex and/or nonlinear, model-based approaches may require the use of high-order and/or linearized multiple models that can be unsuitable for practical implementations [21].

Alternatively, data-driven techniques can be used for condition-based monitoring (CBM), and the fault-detection and identification (FDI) of suspension components. Data-driven methods do not utilize complex vehicle dynamic models and typically require only some vibration/acceleration signals acquired from the vehicle. For example, the peak-to-peak amplitude of pass-band filtered acceleration signals was used to detect faults in rotary components of the traction system in Shinkansen bogies [22]. The level of interactions of the dynamic behaviors between the two suspensions was used as a key indicator of the suspension conditions [21]. Similarly, relations between the specific motions of a vehicle's bogie (angular and translational) were used as important features for the CBM of suspension systems proposed in [23], [24]. Moreover, using multivariate statistical approaches has garnered increasing interest, such as principle component analysis, multi-block partial least squares, independent component analysis, and canonical variate analysis, to monitor system health conditions [25]–[27]. Generally, these data-driven approaches have fewer application limitations, and can be easily extended to different systems if a large amount of training data is available. These data-driven approaches present two typical limitations. One limitation is that the results are less interpretable than the results from the model-based approaches – indicators of system health status or fault categories can be computed yet detailed health information is difficult to obtain, i.e., the physical degradation parameters. Another limitation is that few of these studies validate the model using the real operational data of high-speed trains.

To bridge the gap between the model-based and data-driven approaches, this paper presents a novel domain-knowledge guided data-driven approach that is simple but effective for the health monitoring of vehicle suspensions, which does not require numerous precise information of the underlying physical model. Based on vibration data collected on the trains, the proposed approach can estimate the health status of major components in the suspension system via the multi-output support vector regression (MSVR) model [28]–[30]. The primary idea of our proposed method is to extract features from the vibration data that are relevant to the suspension system degradation, and predict the stiffness and damping coefficients of suspension systems based on a simple dynamics model. The key components of the proposed method include 1) a feature extraction method based on a simple dynamics model to select the relevant information in the multi-location vibration data, 2) a module to generate a training dataset to handle the issue of no labeled data due to practical limits, and 3) a supervised learning framework based on the MSVR model to reconstruct system health conditions.

The main advantages of our approach compared with existing methods are:

1) Easy-to-implement and adaptable to other systems – using simple dynamics model, not sophisticated dynamic models, compared with the model-based approach;

2) Interpretable results – reporting the health status of major components in the suspension system using indicators (e.g. spring stiffness & damping coefficient) with physical meanings, not a few fault labels, compared with pure data-driven methods;

3) Unbound to the availability of real-world labeled data - a novel way to generate training datasets via a simple dynamic model and impact analysis, which addressed a typical challenge in real-world system monitoring, namely, the lack of labeled data due to practical limits.

This paper is structured as follows: The proposed framework is described in Section II. Model evaluation and testing are presented in Section III. In Section III, we evaluate the proposed approach and apply it to real operational data. Finally, conclusions and future works are presented in Section IV.

II. Method

We herein propose a novel data-driven framework for the prediction of the health status of high-speed rail suspension systems. The inputs of the framework are vibration signals collected on trains during its operation. The outputs are the stiffness and damping coefficients of train suspension systems. Spring stiffness and damping coefficients are chosen as model output because they are key indicators reflecting the health condition of a suspension system. The primary and the secondary suspensions are often simplified to spring–damper elements [5]–[9], [31]. The performance of the spring-damper elements is monitored by spring stiffness and damping coefficients.

A key challenge lies in the lack of labeled data for training the data-driven model from real operations and simulations. In real operations, the health status of the suspension systems cannot be evaluated unless they are removed from the train to the test bench. Because the operational cost related to such testing is extremely high, the suspension systems are typically inspected visually, maintained regularly, and replaced at a pre-specified age or usage. Therefore, the labeled data of suspension system health conditions from real operations are rare. Alternatively, high-fidelity simulations are often used to generate labeled data, yet it is difficult to match the simulation data with the vibration data collected from real operations. Detailed technical specifications are required for the high-fidelity simulations to be similar enough to the trains from which vibration data are collected. Different trains would require different simulations and different models, thereby rendering them cumbersome to use. To address these challenges, the proposed framework includes a simple vehicle

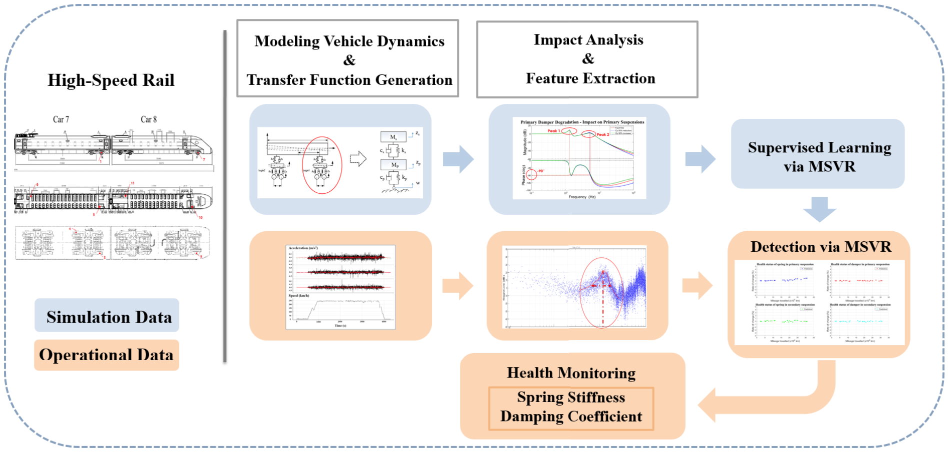

Fig. 1. Integral framework for suspension system health monitoring.

dynamics model that is used to generate labeled data and develop feature extraction techniques for processing vibration data. The vehicle dynamics model can be easily adapted to other trains because it only requires a few key parameters of the train system.

Overall, three modules are included in the proposed framework: 1) modeling vehicle dynamics and transfer function generation, 2) impact analysis and feature extraction, and 3) supervised learning via MSVR, as illustrated in Fig. 1. Herein, we introduce each module briefly. The details are provided in the following sub-sections.

1) Modeling vehicle dynamics and transfer function generation. The first module consists of the modeling and transfer function generation of a high–speed rail's primary and secondary suspension systems. The suspension system is simplified as a two degree-of-freedom (2-DOF) model. This module is the foundation of the following two modules. It facilitates the second module: impact analysis to enable us learn how to select and construct features. Also, it serves as a training dataset generator for the third module: Supervised learning via MSVR.

2) Impact analysis and feature extraction. The second module investigates the impact of springs and damper degradation on frequency response curves. Based on the impact analysis results, features are selected and constructed as input for the proposed method. The vehicle parameters of a specific bogie type under study are applied to obtain the theoretically based Bode plots of the primary and secondary suspension systems. Through simulation analysis, the impact of specific component degradation is analyzed on frequency response curves. We find that the shape and location of the peak with the largest magnitude in the response curve are most sensitive to changes in spring stiffness and damping coefficient of a suspension system. Thus, the positions, heights, and widths of the peaks at the resonant frequency are chosen as important features to be measured and extracted for building the regression model in the third step. A signal processing algorithm was applied to perform the smoothing and peak detection of the magnitude response curves.

3) Supervised learning via MSVR. The third module pertains to obtaining the health status of each suspension system component (spring and damper in primary and secondary suspensions) based on the features extracted from the second module. The MSVR method was applied to model the components' parameters to obtain the point estimates of the components' condition. The output of the MSVR model is the relative changes of the spring and damper parameters, and the input is the relative changes of the position, width, and height of the peak at the resonant frequency.

Once the model is built, it can be used for real-time monitoring. Vibration signals collected on the trains can be processed using the feature extraction technique and analyzed via the MSVR model to predict the health conditions of the suspension systems in real time.

A. Modeling Vehicle Dynamics and Transfer Function Generation

The objective of this module is to build a simple dynamics model that can provide insights into how vibration signals change because of suspension system degradation. The model is not required to generate raw vibration signals that match the operational data exactly. Only underlying features that are applicable to the operational data are required. The results of this module will be used to develop the second module for feature extraction and the third module for supervised learning.

Rail vehicles generally contain both a primary (between the wheelset and bogie frame) and a secondary (between the bogie frame and vehicle body) suspension to support the car body and bogie, to isolate the forces generated by the track unevenness at the wheels, and to control the behavior of the car body with respect to the track surface for providing ride stability, comfort, and safety [32]. The vertical suspension system considered in this research is shown in Fig. 2 [33], including four air springs, four secondary dampers, eight primary springs and eight primary dampers.

In this research, the suspension system is simplified as a 2-DOF model, as depicted in Fig. 3, after ignoring the pitch and rolling. A quarter train model (one of the four wheels in one bogie) was used to simplify the problem to a one-dimensional spring–damper system [34]–[37].

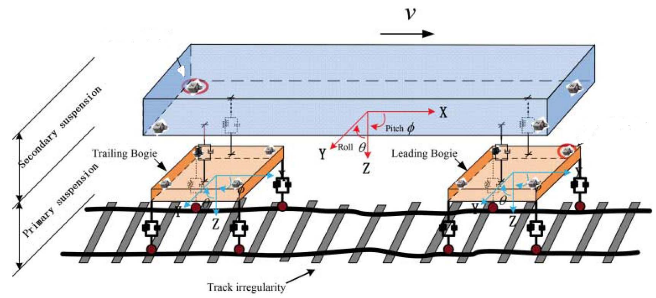

Fig. 2. Vertical suspension system of the rail vehicle.

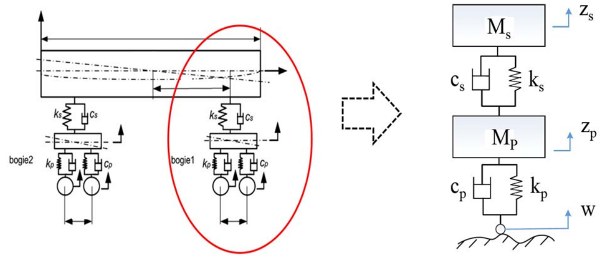

Fig. 3. Simplified 2-DOF vertical suspension system of the rail vehicle.

TABLE I
PARAMETERS OF THE VEHICLE SUSPENSION SYSTEM

	Description	Unit
M_p	Mass of bogie frame ¼	kg
M_s	Mass of car body ¼	kg
k_p	Primary spring stiffness per wheelset	kN/m
k_s	Secondary spring stiffness per bogie	kN/m
c_p	Primary damping coefficient per wheelset	kNs/m
c_s	Secondary damping coefficient per bogie	kNs/m

The standard dynamic equations of the 2-DOF model for both the car body and bogies are described as follows:

$$M_p\ddot{z}_p + (c_p + c_s)\dot{z}_p - c_s\dot{z}_s + (k_p + k_s)z_p - k_s z_s - c_p\dot{w} - k_p\dot{w} = 0 \quad (1)$$

$$M_s\ddot{z}_s - c_s(\dot{z}_p - \dot{z}_s) - k_s(z_p - z_s) = 0 \quad (2)$$

where z_p, z_s, and w denote the vertical displacement of the car body, bogie, and wheel, respectively. The parameters of the vertical vehicle suspension system are given in Table I.

To generate the transfer function of such a system, assume that all of the initial conditions are zeroes, and that these equations represent the situation when the high-speed rail's wheel passes a bump. Using the Laplace transform, the dynamic equation (1) and (2) are expressed in the form of transfer functions. The Laplace transform takes a function of a positive real variable t (time) to a function of a complex variable s (frequency). Subsequently, we can obtain the transfer function $G_p(s)$ of the primary suspension system, and $G_s(s)$ of the secondary suspension system as follows:

$$G_P(s) = \frac{z_P(s)}{w(s)} = \frac{c_P s + k_P}{M_P s^2 + [c_P + c_S(1-\Delta)]s + [k_P + k_S(1-\Delta)]}$$

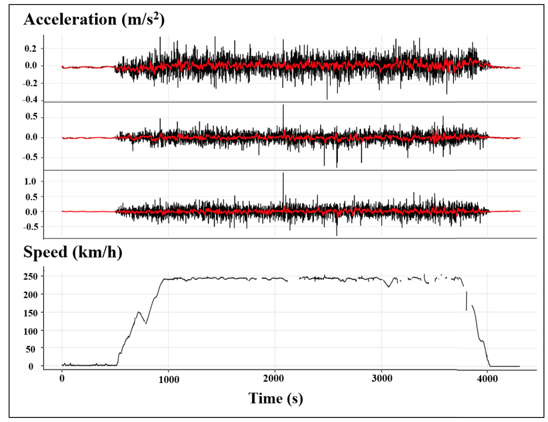

Fig. 4. An example of the longitudinal, lateral, and vertical vibration signals (from top to bottom) from the axle box, and speed data for one trip; the curves highlighted in red is the vibration signals smoothed by the moving average method.

$$\text{where } \Delta = \frac{c_s s + k_s}{M_s s^2 + c_s s + k_s} \quad (3)$$

$$G_S(s) = \frac{z_s(s)}{z_p(s)} = \frac{c_s s + k_s}{M_s s^2 + c_s s + k_s} \quad (4)$$

where, s is a complex number frequency parameter, called the Laplace variable. Therefore, the transfer functions of our simplified primary and secondary suspension systems are obtained as Equation (3) and (4), which can be used to derive the magnitude response and phase response of a system.

For operational data, the exact transfer function is difficult to obtain. Thus, we estimate the magnitude information of a transfer function via measured vibration data of the suspension systems. Considering that the data are time domain vertical acceleration data, as shown in Fig. 4, we first perform mean removal and normalization. Subsequently, an algorithm called TFESTIMATE [38] is applied to estimate the transfer function of the primary and secondary suspension systems via a nonparametric system identification method, Welch's averaged periodogram method [39]. Given the input signal from the axle box and the output signal from the bogie frame, the function can compute and plot the transfer function estimate of the primary suspension system. Similarly, the transfer function of the secondary suspension system can be obtained based on the vibration signals from the bogie frame and car body.

B. Impact Analysis and Feature Extraction

After formulating the transfer function of the suspension system, the next step is to generate the Bode plots and investigate the impact of spring and damper degradation on the magnitude of the frequency responses to guide feature construction.

Applying the vehicle parameters in Table I, the theoretically based Bode plots of the primary and secondary suspension systems are generated, as shown in Fig. 5. The frequency at which the response curve exhibits the maximum magnitude is known as the resonant frequency. At this frequency, the slope of the magnitude curve is zero; meanwhile, the output phase angle shifts by 90 degrees from the input. Some systems contain multiple, distinct resonant frequencies [40]. For the primary suspension system (Fig. 6), it contains two resonant

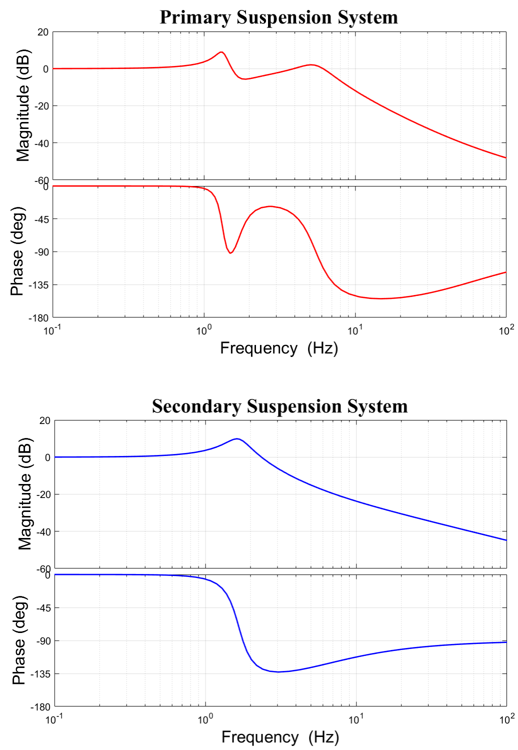

Fig. 5. Bode plots of primary and secondary suspension systems.

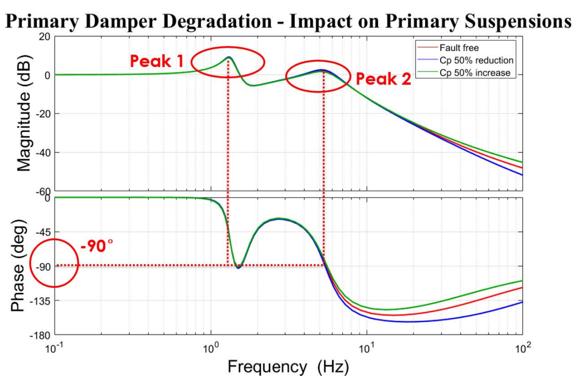

Fig. 6. Peaks at resonant frequencies of the primary suspension system.

frequencies because it is a 2-DOF system and the number of secondary suspension systems is one for the same reason.

To quantify the impact on the frequency responses of the suspensions' specific components degradation and payload changes, a series of frequency response curves were generated by changing the spring stiffness, damping coefficient, and mass of car body simultaneously. As shown in Fig. 7, we find that the shape and location of the peak with the largest magnitude in the response curve are most sensitive to changes in spring stiffness and damping coefficient of a suspension system. This is aligned with domain knowledge - The frequency at which the response curve exhibits the maximum magnitude is the resonant frequency, which is an inherent attribute of a system. For high-speed rail suspension systems, the resonant frequency relates directly to the performance of a suspension system. When the resonant frequency changes, it implies that some intrinsic parameters of the system have changed. Therefore, we use Position, Height and Width of the largest peak in magnitude frequency response curves as our model input.

The features inside all the curves and the relative changes in the component parameters form the database for training the supervised learning module in the next step. Using Fig. 7 (1) as an example, when reducing the primary spring stiffness k_P by half and the remaining primary damping coefficient c_P unchanged (highlighted in blue), the peak in the first resonant frequency shifts to the left, while its height and width increase. Therefore, the output vector of this training set is $y_i = \begin{bmatrix} -50\% \\ 0\% \end{bmatrix}$, and the input vector of this training set $x_i = \begin{bmatrix} position\% \\ width\% \\ height\% \end{bmatrix}$ consists of the relative changes (use the features from 100% healthy condition as the standard) in the features extracted from the curve. Here, i represents one particular training set.

For the operational data, the findpeaks [41] algorithms are used to locate and measure the peaks of the magnitude response curve. It detects peaks and subsequently estimates the position, height, and width of each peak by least-squares curve fitting at the top part of the peaks, assuming they are locally Gaussian or Lorentzian, when the frequency response curve is subject to noise.

C. Supervised Learning via MSVR

In the last module, we develop the support vector regression (SVR) method to predict the health status of the suspension system components based on the features extracted from the magnitude frequency response curves. SVR is chosen because its performance is the most accurate and reliable among advanced regression methods that can solve this problem. The commonly used advanced regression methods are SVR, Gaussian process regression (GPR), Random Forest (RF), artificial neural network (ANN), and linear regression (LR). Except for ANN and RF, all other commonly used regression methods are tested and evaluated. SVR-based methods perform better in general. Details on the performance comparison of different regression methods can be found in Section III. C. ANN is not chosen in the proposed method because ANN requires a large number of training data and difficult to tune the parameters. SVR requires less training data and is relatively easier for others when they want to reproduce the estimation results. RF is not tested in the research, because it has been observed to over-fit for some datasets with noisy regression tasks [42]. MSVR is chosen over SVR because our model has multiple outputs: spring stiffness and damping coefficient.

Let us consider a set of independent samples denoted as $L = \{(x_i, y_i)\}_{i=1}^{M}$, where each $x_i \in \Re^m$ represents a vector of m extracted features of the peak, $y_i \in \Re^n$ is a vector of n associated relative variations of the component parameters, and M is the number of samples for modeling. For brevity, we aggregate all $x_i s$ $(i = 1, 2, \ldots, M)$ into an $m \times M$ feature matrix$x$ and all $y_i s(i = 1, 2, \ldots, M)$ into an $M \times n$ target vector y, such that $L = \{x, y\}$. Our goal is to infer from the

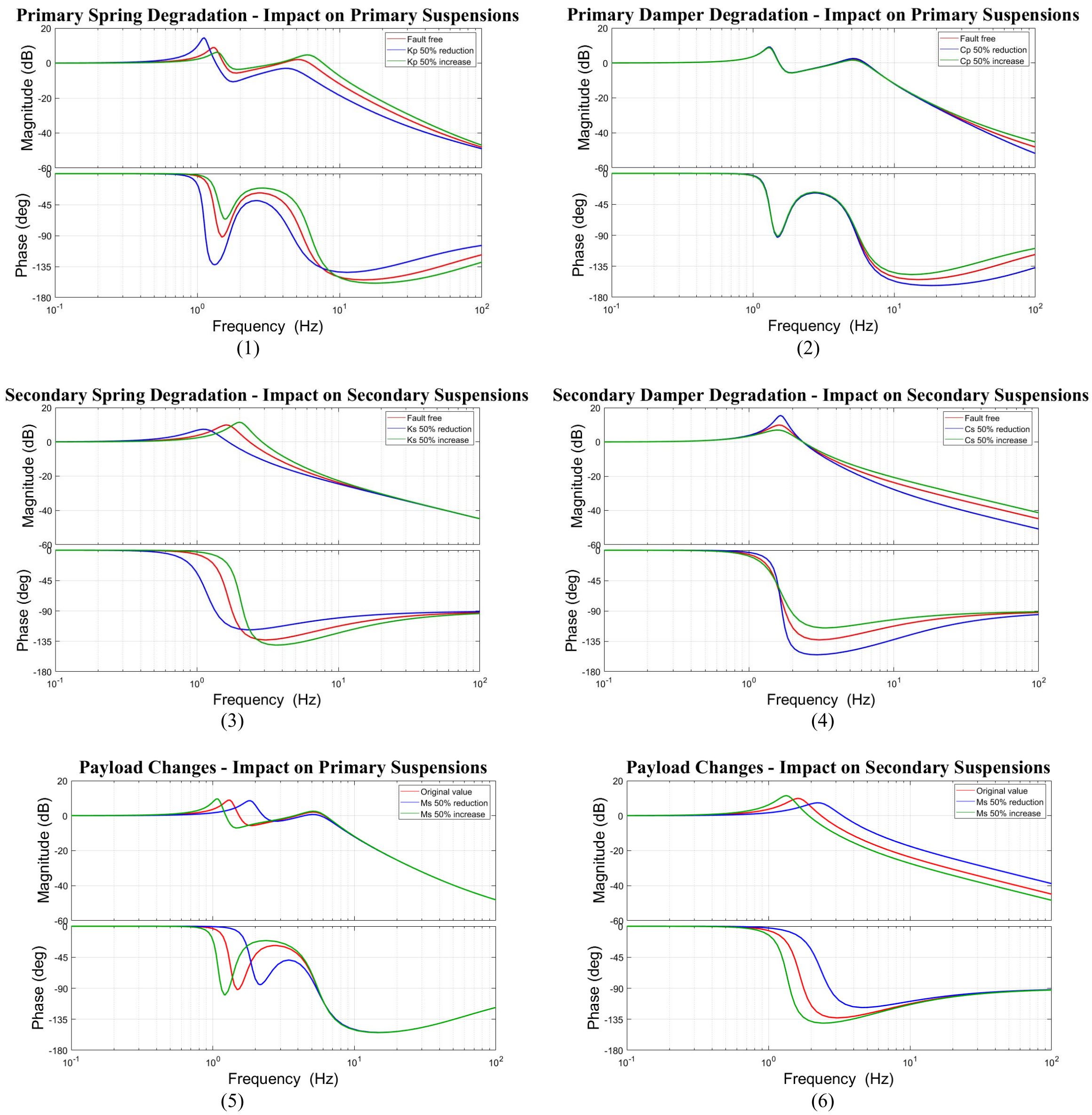

Fig. 7. Changes in frequency responses of damper and spring with increasing and decreasing parameters for the suspension system.

set of samples L the function $f(\cdot)$, such that

$$y = f(x) \tag{5}$$

Considering that the output components' parameters are multidimensional, the MSVR was applied. The key idea of MSVR is to expand the Vapnik ε-insensitive loss function to a multi-dimensional output case, that is, a hyper-spherical insensitive zone, to handle all the outputs together [29], [43]. Therefore, MSVR can improve the generalization performance of the decision model especially when only scarce samples are available [43]. Here, a brief introduction of MSVR will be provided as follows:

In this study, each input vector is $x_i = \begin{bmatrix} position_i \\ width_i \\ height_i \end{bmatrix}$, which is the relative variations of the position, width, and height of the peak at the resonant frequency. Further, each output vector $y_i = \begin{bmatrix} spring\ stiffness_i \\ damping\ coefficient_i \end{bmatrix}$ is the relative variations of the spring and damper parameters. The goal of the MSVR model is to develop a function $f(x)$ that can accurately predict the relative change rate values based on the input features of future data:,

$$f(x) = W \cdot \phi(x) + b \tag{6}$$

where $\cdot$ denotes the dot product, $W$is a set of weight vectors, b refers to a set of biases, and $\phi(\cdot)$ represents the nonlinear mapping from the primal space to the high-dimension feature space.

Subsequently, MSVR is formulated as the minimization of the following function by searching the regressors W and b [28], [43]:

$$\min_{W,b} L_p = \frac{1}{2}\sum\nolimits_{j=1}^{Q}\left\|w^j\right\|^2 + C\sum\nolimits_{i=1}^{n} L(u_i) \tag{7}$$

where $W = [w^1, \cdots, w^Q]$, $b = [b^1, \cdots, b^Q]^T$, $u_i = \|e_i\| = \sqrt{e_i^T e_i}$, and $e_i^T = y_i^T - \phi(x_i)^T W - b^T$; C is a positive real regularized parameter that determines the tradeoff between the regularization and the error reduction term. The ε-insensitive quadratic loss function $L(u)$ is defined as follows:

$$L(u) = \begin{cases} 0, & u < \varepsilon \\ u^2 - 2u\varepsilon + \varepsilon^2, & u \geq \varepsilon \end{cases} \tag{8}$$

As (5) cannot be solved straightforwardly, an iterative reweighted least-squares (IRWLS) method [28] was adopted to obtain the desired solution. To construct the IRWLS procedure, we approximated (1) using a first-order Taylor expansion of $L(u)$ over the previous solution, thus resulting in the following equation [29]:

$$L'_p(W, b) = \frac{1}{2}\sum\nolimits_{j=1}^{Q} \left\|w^j\right\|^2 + \frac{1}{2}\sum\nolimits_{i=1}^{n} a_i u_i^2 + \tau \tag{9}$$

where,

$$a_i = \begin{cases} 0, & u_i^k < \varepsilon \\ \dfrac{2C(u_i^k - \varepsilon)}{u_i^k}, & u_i^k \geq \varepsilon \end{cases} \tag{10}$$

and τ is the sum of constant terms that does not depend on either W or b, and the superscript k denotes the number of iterations. Therefore, the IRWLS procedure is now constructed; it searches the next step solution linearly along the descending direction based on the previous solution. Further, the procedure can be summarized in the following steps [29]:

1) Set $k = 0$, $W^k = 0$, $b^k = 0$, and compute u_i^k and a_i.
2) Compute the solution to (7), and label it as W^s and b^s. Define a descending direction for (5) as

$$P^k = \begin{bmatrix} W^s - W^k \\ (b^s - b^k)^T \end{bmatrix}. \tag{11}$$

3) Obtain the next step solution, i.e.,

$$\begin{bmatrix} W^{k+1} \\ (b^{k+1})^T \end{bmatrix} = \begin{bmatrix} W^k \\ (b^k)^T \end{bmatrix} + \eta^k P^k$$

Subsequently, use a backtracking algorithm to compute the step size η^k.

4) Compute u_i^{k+1} and a_i, set $k = k + 1$, and return to step 2 until convergence.

More details regarding MSVR can be found in [28], [29].

III. Evaluation and Testing

As mentioned above, a major challenge in our research and other high-speed rail health monitoring research is the lack of truth values of the components' health status in the suspension system against time/mileage; this implies that we cannot utilize the traditional method to validate our proposed approach.

In our research, the approach evaluation and testing includes two parts:

1) Using the simulation data generated from impact analysis for evaluation: an evaluation of the proposed method is performed based on the simulation data generated from Module 2. Noises are added with a signal-to-noise ratio of the specific range to test the robustness and sensitivity of the proposed method.
2) Using the real tracking data for testing: by applying our proposed approach in the field data of high-speed train vibration signal, we can obtain the estimated values regarding the health status of the primary components in the suspension system.

TABLE II
Sensor Configuration in Vehicle Suspension Systems of Car 7 and Car 8

CAR	Sensing Point			Sampling Frequency
Car-7 (Trail Car)	Axle box	Left	**1**	5000 Hz
		Right	2	
	Bogie frame	Left	**3**	2000 Hz
		Right	4	
	Car body	Left	**5**	
		Right	6	
Car-8 (Motor Car)	Axle box	Left	7	5000 Hz
		Right	8	
	Bogie frame	Left	9	2000 Hz
	Car Body	Left	10	
		Right	11	

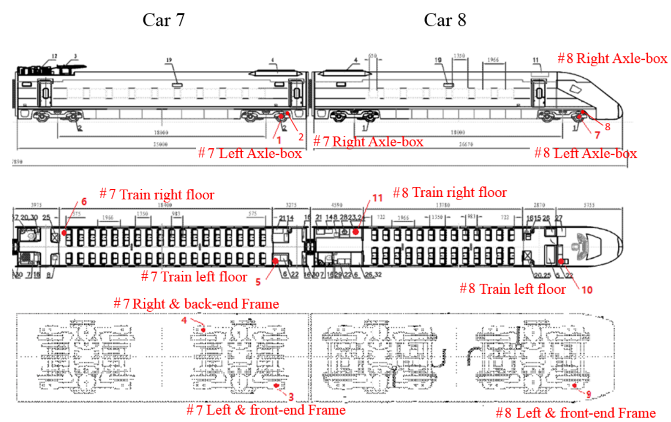

Fig. 8. Position of 11 sensors on two cars, left is the trail car, and right is the motor car.

A. Operational Data Description

The Test and Control Group of the State Key Laboratory of Traction Power in Southwest Jiaotong University continuously tracked the signal on one CRH1A (CRH, China Railway High-speed)-type train operating in China for 10 months from August 2015 to June 2016. The bogie frame of CRH1A EMU was designed by employing the AM96 bogie that used a modern high-speed bus bogie classic U-frame without a bolster structure, as a prototype. The data contain 25 entire trips collected along three different routes, 11 trips on the Changsha–Huaihua line, three trips on the Guangzhou–Zhuhai line, and 16 trips on the Sanya–Haikou loop line.

The vibration signals of one trail car (referred as Car 7) and one motor car (referred as Car 8) were collected through 11 onboard acceleration sensors on the car body, bogie frame, and axle box. Meanwhile, the speed data of the train was also collected via a GPS device and the operation speed

TABLE III
AM96 VEHICLE PARAMETERS

	Description	Unit
M_p	Mass of bogie frame ¼	1725 kg
M_s	Mass of car body ¼	6300 kg
k_p	Primary spring stiffness per wheelset	1.30×10^6 kN/m
k_s	Secondary spring stiffness per bogie	0.69×10^6 kN/m
c_p	Primary damping coefficient per wheelset	3.7×10^3 kNs/m
c_s	Secondary damping coefficient per bogie	22.6×10^3 kNs/m

of this CRH1A-type high-speed rail was up to 250 km/h. Fig. 8 shows the information on the positions of the sensors, and Table II details the sensor configuration in the vehicle suspension systems of Car 7 and Car 8.

B. Modeling Vehicle Dynamics and Training Data Generation for CRH1A EMU

The principal parameter values of the AM96 bogie [44], [45] are given in Table III. By applying the vehicle parameters of Boogie-AM96, the theoretically based transfer function of the primary and secondary suspension system can be obtained:

$$G_S(s) = \frac{226s + 6900}{63s^2 + 226s + 6900} \tag{12}$$

$$G_P(s) = \frac{9324s^3 + 3.31\times10^6 s^2 + 1.28\times10^7 s + 3.59\times10^8}{4347s^4 + 81870s^3 + 5.52\times10^6 s^2 + 1.28\times10^7 s + 3.59\times10^8} \tag{13}$$

To generate the training dataset, we changed the spring stiffness and damping coefficient from 50% to 200% of its initial value with an interval of 1%, to obtain the associated Bode plots. Meanwhile, to render the method reliable under external factor changes, we generated training data with the mass of the car body varying from 50% to 200% of its initial value with an interval of 10%.

In addition, to test the robustness and sensitivity of the proposed method, we added white Gaussian noise with a specific signal-to-noise ratio, which ranges from 26 dB to 54 dB, using algorithms called AWGN [46]. Meanwhile, for simulating real conditions, we did the similar operation and the signal-to-noise ratio was set as 34 dB to expand the training dataset.

Subsequently, the findpeaks algorithms were applied to perform the smoothing and peak detection of these magnitude response curves. The simulation range and features extracted for further modeling are summarized in Table IV.

C. Evaluation and Testing of MSVR for CRH1A EMU

To compare with MSVR, we tested three additional regression methods: 1) the standard support vector regression (SVR) method, which only contains a one-dimensional output; 2) the multivariate Gaussian process regression (MV-GPR) method [47], which has been proven to be a powerful and effective method for nonlinear regression problems owing to its simple structure in obtaining and expressing uncertainties in predictions, its capability in capturing a wide variety of behaviors by parameters, and its natural Bayesian interpretation [48]; 3) the multivariate least-squares linear regression (MV-LR) [49], which is the most typically used method owing to its fast operation and ease of implementation.

The metrics used to evaluate the models are the mean absolute percentage error (MAPE) and root mean square error (RMSE). Models with a low MAPE and low RMSE are preferred.

- **Mean Absolute Percentage Error (MAPE):** It is a measure of prediction accuracy of a forecasting method in statistics, and is defined by the formula

$$M = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right| \tag{14}$$

  where A_t is the actual value and F_t is the forecast value.
- **Root Mean Square Error (RMSE):**It is a measure of the differences between values predicted by a model or an estimator and the values observed, and is defined by the formula

$$R = \sqrt{\frac{\sum_{t=1}^{n} (\widehat{y}_t - y_t)^2}{n}} \tag{15}$$

  where y_t is the actual value and $\widehat{y}_t$ is the predicted values.

Table V reports the parameter settings for the models and the parameter tuning are based on cross-validation results. For all MSVR and SVR model training, the Gaussian radial basis function (RBF) kernel was used and defined as $K(x_i, x_j) = exp(\left\| x_i - x_j \right\|^2 / \sigma^2)$, where σ is the kernel parameter. Therefore, three hyperparameters are used for MSVR: regularization parameter C, kernel parameter σ, and error parameter ε. For the MSVR and SVR models, the parameters are optimized using the five-fold leave-one-out cross validation in the predefined ranges, and the best combination is selected according to the least RMSE. For MV-GPR models, the squared exponential covariance function is selected. Further, model hyperparameters $\theta_{mvgpr} = \{l, \sigma_f^2, \sigma_n^2\}$ are tuned similarly, where l is a characteristic length-scale, σ_f^2 is a signal variance, and σ_n^2 is the noise variance.

The performances of the MSVR, SVR, MV-GPR, and MV-LR models trained by the simulation data are summarized in Table VI. We found that the MSVR models performed significantly better, and was more accurate by 10% or higher than the MV-GPR and MV-LR models for all the evaluation metrics. Regarding the single outputs from the SVR models, the M-SVR model also improved the results by approximately 2% accuracy. This may be because, in the MSVR model, it considers the nonlinear relations not only between features but also among the output component parameters themselves. Meanwhile, the estimation accuracy of all the MSVR models could achieve 91% or more, implying that the relationship between the features extracted and the health status of each

TABLE IV
OVERVIEW OF SIMULATION ANALYSIS – COMPONENT DEGRADATION IMPACT

Simulation Group	Components	Standard Value	Simulation Range	Response Curves
I	Primary spring stiffness	1.30×10^6 kN/m	50–200% (1%)	364816
	Secondary spring stiffness	0.69×10^6 kN/m	50–200% (1%)	
	Mass of car body	6300 kg	50–200% (10%)	
II	Primary damping coefficient	3.7×10^3 kNs/m	50–200% (1%)	364816
	Secondary damping coefficient	22.6×10^3 kNs/m	50–200% (1%)	
	Mass of car body	6300 kg	50–200% (10%)	

TABLE V
PARAMETERS SETTINGS FOR MSVR, SVR, AND MV-GPR MODELS

Model	Kernel Function	Hyper-parameter	Training Range
MSVR	**RBF**	$\theta_{msvr} = \{C, \sigma_{RBF}, \varepsilon\}$	$\{[1,100],[10^{-3},100],[10^{-6},10^{-3}]\}$
SVR	**RBF**	$\theta_{svr} = \{C', \sigma'_{RBF}, \varepsilon'\}$	$\{[1,100],[10^{-3},100],[10^{-6},10^{-3}]\}$
MV-GPR	**Squared Exponential**	$\theta_{mvgpr} = \{l, \sigma_f^2, \sigma_n^2\}$	$\{[10^{-2},10],[10^{-3},100],[10^{-4},10^{-1}]\}$

TABLE VI
PERFORMANCES OF MSVR, SVR, MV-GPR, AND MV-LR MODELS IN TERMS OF MAPE AND RMSE

Simulation Group	Output	Metrics	MSVR	SVR	MV-GPR (MEAN)	MV-LR
I	Primary Spring Stiffness	MAPE (%)	**5.89**	7.34	14.76	19.69
		RMSE	**0.0579**	0.0721	0.1438	0.1953
	Primary Damper Coefficient	MAPE (%)	**8.97**	10.29	17.26	21.71
		RMSE	**0.0871**	0.1013	0.1703	0.2151
II	Secondary Spring Stiffness	MAPE (%)	**4.37**	6.96	13.27	18.06
		RMSE	**0.0426**	0.0673	0.1311	0.1791
	Secondary Damper Coefficient	MAPE (%)	**7.53**	9.92	16.43	20.84
		RMSE	**0.0739**	0.0972	0.1632	0.2073

suspension's component is well built for the real tracking of data detection.

D. Application to Data From Real Operations

We applied the proposed method on the vibration data from real operations, as described in Section III. A.

First, data preprocessing was conducted, including data cleaning and normalization. Data from the trail car (Car 7) was analyzed owing to the deficiency of data in the motor car (car 8). Only data from sensor 1 (axle box), sensor 3 (bogie frame), and sensor 5 (car body) were selected for the study as they directly measure the vibration input and output of the primary and secondary suspension systems. The vibration data collected from one entire trip were regarded as one dataset, which implies that 25 datasets were available for processing. Further, each dataset consisted of three data subsets from sensor 1, sensor 3, and sensor 5. Considering that the data were time domain vertical acceleration data, we performed mean removal and normalization (divided by sample standard deviation) for each data subset.

Subsequently, the transfer functions and frequency response curves were calculated based on the vibration data. The TFESTIMATE algorithm was applied to estimate the transfer function of the primary and secondary suspension systems. Given the input signal from sensor 1 (axle box) and the output signal from sensor 3 (bogie frame), the function could compute and plot the transfer function estimate of the primary suspension system. Similarly, the transfer function of the secondary suspension system could be generated. Subsequently, the magnitude frequency response curves of the primary and secondary suspensions were obtained for further peak finding and feature extraction (Fig. 9). In total, 60 magnitude–frequency response curves were generated, with one curve each for the primary and secondary suspensions generated for one data subset.

The magnitude frequency response curves of the secondary suspensions generated from our simulation data and real tracking data have a prominent resonant frequency at approximately 1 Hz, which is consistent with the literature. Zhai et al. [50] found that the natural vibration frequency of the car body's vertical suspension system (secondary suspension system) was

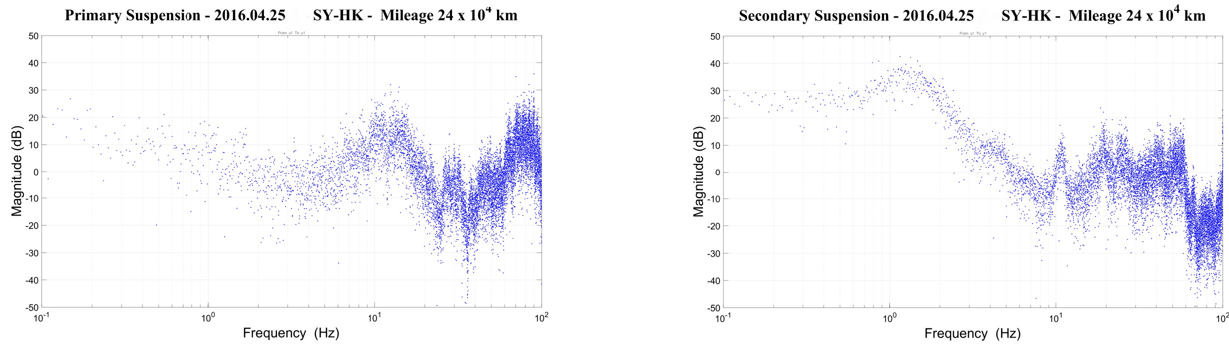

Fig. 9. Samples of magnitude frequency response curves for the primary and secondary suspension system.

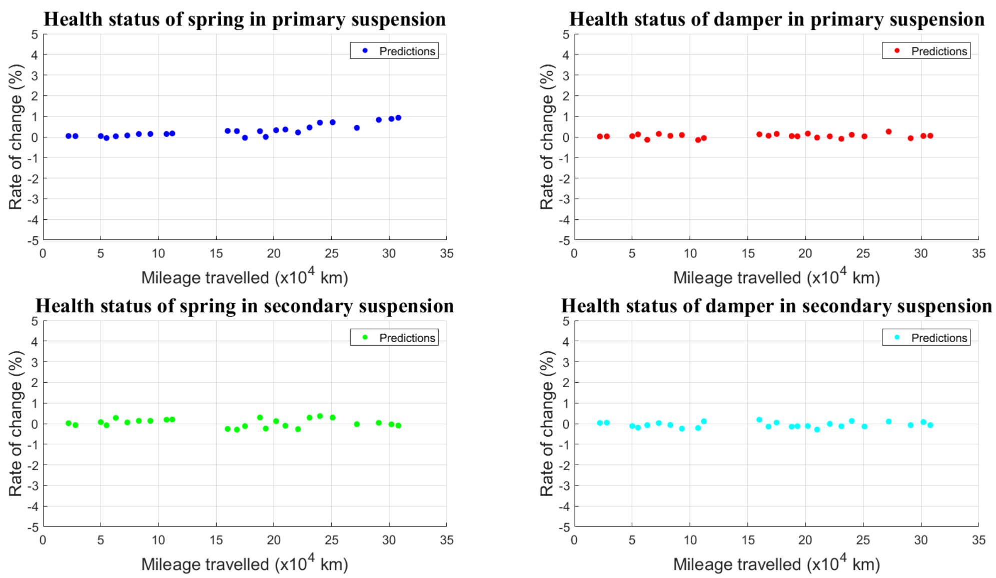

Fig. 10. Variation in the health status of the suspension system with mileage traveled.

approximately 1 Hz. They also found that the prominent resonant frequency of the railway vertical primary suspension system was approximately 11.46 Hz, which was studied and tested in [51].

Finally, by applying the MSVR models built before, we estimated the stiffness and damping variations of the suspension systems for each trip. The results are shown in Fig. 10. We found no obvious change in stiffness and damping coefficients in any suspension system in the first 30.8×10^4km mileage traveled. Meanwhile, the fluctuation range of the estimation results for the secondary suspension system is smaller than that for the primary suspension system. This result is reasonable, because the air spring and anti-yaw damper in the secondary suspensions have a longer useful life, less load, and less wear. Further, based on the specifications of TG/GL145-2016 CRH 1A/1B/1E EMU LEVEL III MAINTENANCE REGULATION, the air spring should be replaced when the train has traveled $360 \pm 10 \times 10^4$km. Meanwhile, the CRH1A train in our study merely traveled 30.8×10^4km, which is one-twelfth of the designed mileage, during the signal tracking period. For the primary suspension system, a slight change occurred, approximately 1%, in the spring stiffness, when the train traveled over 25×10^4km. This may be caused by the poorer lubricating property in the spring blades against the mileage traveled. This is also consistent with another study that reported the degradation rate of the suspension spring health of approximately 1% when the operational time is approximately one-twelfth of the designed lifetime [52], [53].

E. Discussion on Sensor Requirements

In this study, the operational vibration data were collected via accelerometer sensors. The performance of these sensors would affect the effectiveness of the proposed method. Setting comprehensive required specifications on the sensors is challenging within this study, yet some sensitivity analysis and discussion are provided in this section to inform future studies.

The performance and cost of specific accelerometers vary significantly depending on the type of mechanisms involved in the accelerometers, e.g. piezoelectric, piezoresistive, capacitive, Hall effect, magnetoresistive, and temperature sensors, etc. In addition, low-cost and low power micro-electro-mechanical systems (MEMS) devices are increasingly available. In this study, the operational vibration data were collected via accelerometer sensors. The sensors were installed by China State Key Laboratory of Traction Power for multiple research projects. Thus, the selection of the sensors and installation locations were not particularly optimized for this study.

TABLE VII
ZW9609A-2 ACCELEROMETER INFORMATION

TYPE	RANGE	PRECISION	SENSITIVITY	FREQUENCY RESPONSE	POWER	PRICE
ZW9609A-2	± 2g	0.1%	Vertical: 993.7 mV/g Lateral: 992.6 mV/g Longitudinal: 1003.3 mV/g	DC-2500 Hz (-3dB)	+8 ~ +20 V_{DC}	~700 USD

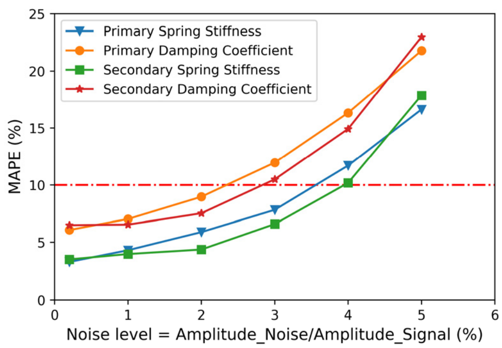

Fig. 11. The sensitivity analysis result of the framework.

The sensors are ZW9609A-2 integrated circuits piezoelectric accelerometers. The detailed information of the accelerometer is depicted in VII.

In terms of precision, a sensitivity analysis was performed to test the robustness and sensitivity of the proposed method subject to accelerometer measurement errors. When generating training dataset as described in Section III. B, we added white Gaussian noise with different specific signal-to-noise ratios to the simulated sensor measurements, using algorithms called AWGN. The sensitivity analysis result is showed in Fig. 11. The proposed method can provide relatively accurate results (MAPE $< 10\%$) when the sensor precision is higher than 2%.

Regarding the installation locations, our method requires three sensors to be installed on axle box, bogie frame, and car body, respectively, and they should be vertically aligned with the suspension system to be monitored. In this study, the installation of the accelerometers was not optimized for monitoring suspension systems. More sensors were installed than needed. We only picked Sensor 1, Sensor 3 and Sensor 5, to demonstrate how to use our method in real operations. The detailed sensor configuration is described in Fig. 8 and Table II.

IV. CONCLUSION

A novel domain-knowledge-guided data-driven framework for the prediction of the health status of high-speed rail suspension systems was proposed. The proposed framework included three parts: 1) vehicle dynamics modeling, 2) impact analysis and feature extraction, and 3) building MSVR models. Simple modeling and transfer function generation of the high-speed rail's primary and secondary suspension systems were performed first. Next, the investigation of the impact of spring and damper degradation generated magnitude–frequency curves. The peak finding and measurement in the magnitude–frequency curves were performed by intelligent algorithms. The positions, heights, and widths of the peaks at the resonant frequency were the important features to be measured and extracted for building the regression model. Finally, the MSVR method was used to model the relationship between the components' parameters and the extracted features.

The proposed framework was evaluated and tested both with the simulation data and real tracking data. We found that the MSVR models achieved good prediction performance for both the simulation data and field data. Through analysis, the health monitoring results of the suspension system from the high-speed rail in practice using the proposed framework were reasonable and consistent with the other studies' results.

However, this framework still requires improvement because the accuracy of the health status estimation is closely linked to the vehicle dynamics modeling that is affected by the vehicle parameters for perfect health. Regarding the vehicle parameters for perfect health, some deviations could exist between the parameters from real high-speed rails and those provided by the manufacturer. Further, we still require the ground truth values of the suspension system's health status in the field data to improve and further validate our framework.

Future extensions of this research work includes the following: 1) modification of the framework to enable real-time online monitoring; 2) estimation of the remaining useful life and the design of the corresponding maintenance strategy; 3) extension of the application of the proposed approach to the suspension systems of other types of vehicles, such as metros, trucks, and sedans, as they can be subjected to similar vehicle system dynamics and monitored using similar types of features.

ACKNOWLEDGMENT

The authors would like to thank the editors and anonymous reviewers for their helpful comments and suggestions. The authors would also like to thank the Test and Control Group of the State Key Laboratory of Traction Power in Southwest Jiaotong University for their technical support.

REFERENCES

[1] M. Hong, Q. Wang, Z. Su, and L. Cheng, "*In situ* health monitoring for bogie systems of CRH380 train on Beijing–Shanghai high-speed railway," *Mech. Syst. Signal Process.*, vol. 45, no. 2, pp. 378–395, Apr. 2014.

[2] S. Bruni, R. Goodall, T. X. Mei, and H. Tsunashima, "Control and monitoring for railway vehicle dynamics," *Veh. Syst. Dyn.*, vol. 45, nos. 7–8, pp. 743–779, Jun. 2007.

[3] R. W. Ngigi, C. Pislaru, A. Ball, and F. Gu, "Modern techniques for condition monitoring of railway vehicle dynamics," *J. Phys., Conf. Ser.*, vol. 364, May 2012, Art. no. 012016.

[4] C. Li, S. Luo, C. Cole, and M. Spiryagin, "An overview: Modern techniques for railway vehicle on-board health monitoring systems," *Veh. Syst. Dyn.*, vol. 55, no. 7, pp. 1045–1070, Mar. 2017.

[5] K. Goda and R. M. Goodall, "Fault-detection-and-isolation system for a railway vehicle bogie," in *Proc. 18th Dyn. Veh. Roads Tracks AVSD Symp. IHELD Kanagawa Jpn.*, AUG. 2003, pp. 24–30.

[6] P. Li and R. Goodall, "Model-based condition monitoring for railway vehicle systems," *Control*, vol. 5, pp. 1–5, Sep. 2004.

[7] H. Fan, X. Wei, L. Jia, and Y. Qin, "Fault detection of railway vehicle suspension systems," in *Proc. 5th Int. Conf. Comput. Sci. Educ.*, Aug. 2010, pp. 1264–1269.

[8] X. Wei, H. Liu, and Y. Qin, "Fault diagnosis of rail vehicle suspension systems by using GLRT," in *Proc. Chin. Control Decis. Conf. (CCDC)*, May 2011, pp. 1932–1936.

[9] M. Jesussek and K. Ellermann, "Fault detection and isolation for a full-scale railway vehicle suspension with multiple Kalman filters," *Veh. Syst. Dyn.*, vol. 52, no. 12, pp. 1695–1715, Sep. 2014.

[10] H. Fang, N. Tian, Y. Wang, M. Zhou, and M. A. Haile, "Nonlinear Bayesian estimation: From Kalman filtering to a broader horizon," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 2, pp. 401–417, Mar. 2018.

[11] Y. Wu, B. Jiang, N. Lu, and D. Zhou, "ToMFIR-based incipient fault detection and estimation for high-speed rail vehicle suspension system," *J. Frankl. Inst.*, vol. 352, no. 4, pp. 1672–1692, Apr. 2015.

[12] Y. Hayashi, H. Tsunashima, and Y. Marumo, "Fault detection of railway vehicle suspension systems using multiple-model approach," *J. Mech. Syst. Transp. Logist.*, vol. 1, no. 1, pp. 88–99, 2008.

[13] P. Li, R. Goodall, P. Weston, C. Seng Ling, C. Goodman, and C. Roberts, "Estimation of railway vehicle suspension parameters for condition monitoring," *Control Eng. Pract.*, vol. 15, no. 1, pp. 43–55, Jan. 2007.

[14] S. X. Ding, T. Jeinsch, P. M. Frank, and E. L. Ding, "A unified approach to the optimization of fault detection systems," *Int. J. Adapt. Control Signal Process.*, vol. 14, no. 7, pp. 725–745, Nov. 2000.

[15] X. Wei, L. Liu, and M. Verhaegen, "Fault detection and estimation for LTI systems and its application to a lab robotic manipulator," in *Proc. Chin. Control Decis. Conf.*, Jun. 2009, pp. 1595–1600.

[16] X. Wei, S. Lin, and H. Liu, "Distributed fault detection observer for rail vehicle suspension systems," in *Proc. 24th Chin. Control Decis. Conf. (CCDC)*, May 2012, pp. 3396–3401.

[17] X. Wei, H. Liu, and L. Jia, "Fault detection of urban rail vehicle suspension system based on acceleration measurements," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2012, pp. 1129–1134.

[18] X. Wei, L. Jia, K. Guo, and S. Wu, "On fault isolation for rail vehicle suspension systems," *Veh. Syst. Dyn.*, vol. 52, no. 6, pp. 847–873, Apr. 2014.

[19] P. Xue, X.-D. Chai, and S.-B. Zheng, "Research on vehicle diagnosis based on state-space method," *Artif. Intell. Res.*, vol. 4, no. 2, p. 55, Jun. 2015.

[20] X. Y. Liu, S. Alfi, and S. Bruni, "An efficient recursive least square-based condition monitoring approach for a rail vehicle suspension system," *Veh. Syst. Dyn.*, vol. 54, no. 6, pp. 814–830, Apr. 2016.

[21] T. X. Mei and X. J. Ding, "Condition monitoring of rail vehicle suspensions based on changes in system dynamic interactions," *Veh. Syst. Dyn.*, vol. 47, no. 9, pp. 1167–1181, Aug. 2009.

[22] T. Oba, K. Yamada, N. Okada, and K. Tanifuji, "Condition monitoring for Shinkansen Bogies based on vibration analysis," *J. Mech. Syst. Transp. Logist.*, vol. 2, no. 2, pp. 133–144, 2009.

[23] T. X. Mei and X. J. Ding, "A model-less technique for the fault detection of rail vehicle suspensions," *Veh. Syst. Dyn.*, vol. 46, no. 1, pp. 277–287, Jan. 2008.

[24] T. Kojima and Y. Sugahara, "Fault detection of vertical dampers of railway vehicle based on phase difference of vibrations," *Quart. Rep. RTRI*, vol. 54, no. 3, pp. 139–144, Aug. 2013.

[25] C. Lee, S. W. Choi, and I.-B. Lee, "Variable reconstruction and sensor fault identification using canonical variate analysis," *J. Process Control*, vol. 16, no. 7, pp. 747–761, Aug. 2006.

[26] S. Yin, X. D. Steven, A. Naik, P. Deng, and A. Haghani, "On PCA-based fault diagnosis techniques," in *Proc. Conf. Control Fault-Tolerant Syst. (SysTol)*, Oct. 2010, pp. 179–184.

[27] X. Wei, Y. Guo, and L. Jia, "MBPLS-based rail vehicle suspension system fault detection," in *Proc. 26th Chin. Control Decis. Conf. (CCDC)*, May /Jun. 2014, pp. 3602–3607.

[28] M. Sanchez-Fernandez, M. de-Prado-Cumplido, J. Arenas-Garcia, and F. Perez-Cruz, "SVM multiregression for nonlinear channel estimation in multiple-input multiple-output systems," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2298–2307, Aug. 2004.

[29] D. Tuia, J. Verrelst, L. Alonso, F. Perez-Cruz, and G. Camps-Valls, "Multioutput support vector regression for remote sensing biophysical parameter estimation," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 804–808, Jul. 2011.

[30] J. Rosentreter, R. Hagensieker, A. Okujeni, R. Roscher, P. D. Wagner, and B. Waske, "Subpixel mapping of urban areas using EnMap data and multioutput support vector regression," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 1938–1948, May 2017.

[31] W. Zhai, K. Wang, and C. Cai, "Fundamentals of vehicl–track coupled dynamics," *Veh. Syst. Dyn.*, vol. 47, no. 11, pp. 1349–1376, Oct. 2009.

[32] X. Wei, Y. Guo, L. Jia, and H. Liu, "Fault detection of rail vehicle suspension system based on CPCA," in *Proc. Conf. Control Fault-Tolerant Syst. (SysTol)*, Oct. 2013, pp. 700–705.

[33] X. Wei, L. Jia, and H. Liu, "A comparative study on fault detection methods of rail vehicle suspension systems based on acceleration measurements," *Veh. Syst. Dyn.*, vol. 51, no. 5, pp. 700–720, Feb. 2013.

[34] D. Hrovat, "Optimal active suspension structures for quarter-car vehicle models," *Automatica*, vol. 26, no. 5, pp. 845–860, Sep. 1990.

[35] G. Georgiou, G. Verros, and S. Natsiavas, "Multi-objective optimization of quarter-car models with a passive or semi-active suspension system," *Veh. Syst. Dyn.*, vol. 45, no. 1, pp. 77–92, Feb. 2007.

[36] W. Gao, N. Zhang, and J. Dai, "A stochastic quarter-car model for dynamic analysis of vehicles with uncertain parameters," *Veh. Syst. Dyn.*, vol. 46, no. 12, pp. 1159–1169, Oct. 2008.

[37] A. Agharkakli, G. S. Sabet, and A. Barouz, "Simulation and analysis of passive and active suspension system using quarter car model for different road profile," *Int. J. Eng. Trends Technol.*, vol. 31, no. 5, pp. 636–644, 2012.

[38] MATLAB. *Transfer Function Estimate—MATLAB Tfestimate*. Accessed: Dec. 09, 2018. [Online]. Available: https://www.mathworks.com/help/signal/ref/tfestimate.html

[39] H. Vold, J. Crowley, and G. T. Rocklin, "New ways of estimating frequency response functions," *Sound Vib.*, vol. 18, no. 11, pp. 34–38, 1984.

[40] Z. Sheng, S. Pfersich, A. Eldridge, J. Zhou, D. Tian, and V. C. M. Leung, "Wireless acoustic sensor networks and edge computing for rapid acoustic monitoring," *IEEE/CAA J. Autom. Sin.*, vol. 6, no. 1, pp. 64–74, Jan. 2019.

[41] T. C. O'Haver. *Peak Finding and Measurement*. Accessed: Dec. 09, 2018. [Online]. Available: https://terpconnect.umd.edu/ toh/spectrum/PeakFindingandMeasurement.htm

[42] M. R. Segal, "Machine learning benchmarks and random forest regression," UCSF, Center Bioinf. Mol. Biostatist., Apr. 2004. Accessed: Dec. 9, 2018. [Online]. Available: https://escholarship.org/uc/item/35x3v9t4

[43] W. Mao, X. Mu, Y. Zheng, and G. Yan, "Leave-one-out cross-validation-based model selection for multi-input multi-output support vector machine," *Neural Comput. Appl.*, vol. 24, no. 2, pp. 441–451, Feb. 2014.

[44] G. Kouroussis, D. P. Connolly, and O. Verlinden, "Railway-induced ground vibrations—A review of vehicle effects," *Int. J. Rail Transp.*, vol. 2, no. 2, pp. 69–110, Apr. 2014.

[45] A. S. Leblebici and S. Türkay, "Track modelling and control of a railway vehicle," *IFAC-PapersOnLine*, vol. 49, no. 21, pp. 274–281, Jan. 2016.

[46] MATLAB. *Add White Gaussian Noise to Signal—MATLAB awgn*. Accessed: Dec. 09, 2018. [Online]. Available: https://www.mathworks.com/help/comm/ref/awgn.html#d120e5590

[47] Z. Chen, B. Wang, and A. N. Gorban, "Multivariate Gaussian and student$-t$ process regression for multi-output prediction," Mar. 2017, *arXiv:1703.04455*. [Online]. Available: https://arxiv.org/abs/1703.04455

[48] P. Boyle and M. Frean, "Dependent Gaussian Processes," in *Advances in Neural Information Processing Systems*, L. K. Saul, Y. Weiss, and L. Bottou, Eds. Cambridge, MA, USA: MIT Press, 2005, pp. 217–224.

[49] N. E. Helwig. (Jan. 16, 2017). *Multivariate Linear Regression*. [Online]. Available: http://users.stat.umn.edu/ helwig/notes/mvlr-Notes.pdf

[50] W. Zhai, P. Liu, J. Lin, and K. Wang, "Experimental investigation on vibration behaviour of a CRH train at speed of 350 km/h," *Int. J. Rail Transp.*, vol. 3, no. 1, pp. 1–16, Feb. 2015.

[51] T.-C. I. Aravanis, J. S. Sakellariou, and S. D. Fassois, "Spectral analysis of railway vehicle vertical vibration under normal operating conditions," *Int. J. Rail Transp.*, vol. 4, no. 4, pp. 193–207, Aug. 2016.

[52] J. Luo, M. Namburu, K. Pattipati, L. Qiao, M. Kawamoto, and S. Chigusa, "Model-based prognostic techniques [maintenance applications]," in *Proc. AUTOTESTCON IEEE Syst. Readiness Technol. Conf.*, Sep. 2003, pp. 330–340.

[53] J. Luo, K. R. Pattipati, L. Qiao, and S. Chigusa, "Model-based prognostic techniques applied to a suspension system," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 38, no. 5, pp. 1156–1168, Sep. 2008.

Ning Hong received the B.S. degree in automation engineering from Zhejiang University, Hangzhou, China, in 2016. He is currently pursuing the Ph.D. degree with the Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong. His research interests focus on data analytics and machine learning, and their application on air transport and railway traffic.

Lishuai Li received the B.Eng. degree in aircraft design and engineering from Fudan University, and the M.Sc. and Ph.D. degrees in air transportation systems from the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology (MIT). She was a Consultant at McKinsey and Company with the Operations Practice in San Francisco. She is currently an Assistant Professor with the Department of Systems Engineering and Engineering Management, City University of Hong Kong. Her research interests are in the development of analytical methods for the design, management, operation of transportation systems, focusing on air transport.

Weiran Yao received the B.S. degree in aircraft airworthiness from Beihang University in 2016. He is currently pursuing the Ph.D. degree with the Department of Civil and Environmental Engineering, Carnegie Mellon University. He was a Research Assistant at the City University of Hong Kong in 2017, where he worked on using social media analytics for the design decision support of transit-oriented development.

Yang Zhao received the B.S. degree in statistics from the Shandong University of Science and Technology in 2011, and the Ph.D. degree from the City University of Hong Kong in 2015. She is currently the Scientific Officer with the Centre for Systems Informatics Engineering, City University of Hong Kong. Her research interests are in machine learning and statistics, especially their application to real problems.

Cai Yi received the Ph.D. degree in vehicle engineering from Southwest Jiaotong University, Chengdu, China, in 2015. She was at the Department of Systems Engineering and Engineering Management, City University of Hong Kong, as a Post-Doctoral Research Fellow, in 2017 and 2018, respectively. Her current research interests include condition monitoring and fault detection of vehicle components, vibration signal processing, and machinery state representation and prediction method.

Jianhui Lin is currently a Professor with the Traction Power State Key Laboratory, Southwest Jiaotong University. He had spent for over 30 years in safety monitoring during the rail transit operation, and experimental study on the technology of the large test rig, long engaged in security monitoring technology, fault diagnosis, signal analysis, and processing of application and research.

Kwok Leung Tsui is currently a Chair Professor of industrial engineering with the School of Data Science, City University of Hong Kong, and the Founder and the Director of the Center for Systems Informatics Engineering. Prior to joining the City University of Hong Kong, he was a Professor at the School of Industrial and Systems Engineering, Georgia Institute of Technology. His current research interests include data mining, surveillance in healthcare and public health, prognostics and systems health management, calibration and validation of computer models, process control and monitoring, and robust design and Taguchi methods. He was a recipient of the National Science Foundation Young Investigator Award. He is a fellow of the American Statistical Association, the American Society for Quality, and the International Society of Engineering Asset Management; a U.S. representative to the ISO Technical Committee on Statistical Methods. He was the Chair of the INFORMS Section on Quality, Statistics, and Reliability and the Founding Chair of the INFORMS Section on Data Mining.