Stochastic dynamic simulation of railway vehicles collision using data-driven modelling approach

Shaodi Dong¹ · Zhao Tang¹ · Michelle Wu^2 · Jianjun Zhang²

Received: 14 October 2021/Revised: 15 March 2022/Accepted: 16 March 2022/Published online: 28 April 2022 © The Author(s) 2022

Abstract Using stochastic dynamic simulation for railway vehicle collision still faces many challenges, such as high modelling complexity and time-consuming. To address the challenges, we introduce a novel data-driven stochastic process modelling (DSPM) approach into dynamic simulation of the railway vehicle collision. This DSPM approach consists of two steps: (i) process description, four kinds of kernels are used to describe the uncertainty inherent in collision processes; (ii) solving, stochastic variational inferences and mini-batch algorithms can then be used to accelerate computations of stochastic processes. By applying DSPM, Gaussian process regression (GPR) and finite element (FE) methods to two collision scenarios (i.e. lead car colliding with a rigid wall, and the lead car colliding with another lead car), we are able to achieve a comprehensive analysis. The comparison between the DSPM approach and the FE method revealed that the DSPM approach is capable of calculating the corresponding confidence interval, simultaneously improving the overall computational efficiency. Comparing the DSPM approach with the GPR method indicates that the DSPM approach has the ability to accurately describe the dynamic response under unknown conditions. Overall, this research demonstrates the feasibility and usability of the proposed DSPM approach for stochastic dynamics simulation of the railway vehicle collision.

Zhao Tang tangzhao@swjtu.edu.cn

² National Centre for Computer Animation, Bournemouth University, Dorset, UK **Keywords** Dynamic simulation · Railway vehicle collision · Stochastic process · Data-driven stochastic process modelling

1 Introduction

1.1 Background

The dynamic analysis of railway vehicles collision, as a vital issue in passive safety, has been an indispensable research topic in recent years [1, 2]. Although the real train crash experiments are the most reliable and preferred approach, the incurrence of extensive human labour and substantial costs for full-size tests has led researchers to adopt scaled tests instead [3, 4]. Nowadays, the dynamic simulation methods including the finite element (FE) simulation and multi-body dynamic simulation have significantly contributed to cutting down research and labour costs [5, 6]. However, the application of these dynamic simulation methods still needs to be improved in the actual implementation process due to the following reasons: (1) Complex modelling. It is generally expensive to employ the traditional simulation method due to the complexity of the modelling process and the length of time it takes to simulate the system. (2) Stochastic factors. It is inevitable that there will be stochastic factors in the experiments and simulation processes [7], resulting in a reduction in accuracy and the appearance of a random distribution (Fig. 1a).



¹ State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China

1.2 State-of-the-art development

To address the challenge of balancing accuracy and efficiency, data-driven modelling methods as a reverse approach have recently received more interest in the engineering field. For instance, Tang et al. [8–10] proposed a data-driven collision modelling method where useful force-displacement curve models are extracted from the existing FE simulation data, to predict the collision dynamic response under various collision conditions. Müller et al. [11] proposed a machine learning method to quickly predict and estimate the severity of crushing caused by collisions based on the FE simulation data. Li et al. [12] proposed a vehicle collision mathematical model (VCMM) to predict the collision response of vehicles at unknown velocities, and the accuracy and efficiency of the method are verified by comparison with the FE simulation results.

Besides, many researchers are committed to dynamic simulations considering stochastic factors, and have put forward several modelling methods for stochastic dynamic simulation. For example, Luo et al. [13] applied mathematics normal distribution to rationally describe different stochastic parameters, and then accurately predicted the evolution of wheel profile wear and related vehicle dynamic for high-speed trains. Lu et.al [14] discussed in detail the dynamic effects of vehicle speed, load, road surface roughness and tire stiffness by analysing tire dynamic load and dynamic load coefficient. Souffran et al. [15] analysed the relationship between stochastic characteristics and variables (i.e. vehicle speed, acceleration, and road gradient), and then proposed a methodology for modelling real-world vehicle missions. The application of a novel probability method is another popular approach for assessing vehicle dynamic. In this regard, Xu et al. [16] proposed a probability reduction method, considering the random combination of different random variables in vehicle-track coupled systems; Hao et al. [17] applied a specific prior distribution to quantify the probability by checking all modal activities between consecutive data points. In addition, Gaussian process regression (GPR) [18] has gradually become a baseline stochastic processes approach due to its relatively high accuracy of description; however, this approach needs to store a full training dataset to realize predictions (Fig. 1b), which means that the complexity will grow proportionally with the size of the dataset.

To improve efficiency, a data-driven approach combined with variational inference is developed to analyse the stochastic process [19–22], and an efficient prediction function is derived from the mini-batch training data (Fig. 1c). The data-driven stochastic method has been widely used in engineering, such as health monitoring [23], real-time monitoring [24], vehicle speed prediction [25], and their applicability has been independently verified.

Although the above-mentioned methods have yielded improvements in terms of effectiveness, there are seldom studies to set up empirical or physical models incorporating

Fig. 1 Stochastic process analysis: **a** graph showing the stochastic process in engineering [5]; **b** graphical model of the Gaussian process regression (GPR), where the connectivity between values of the function f_i is denoted by a loop around the plate [16]; **c** graphical model representing the stochastic processes using stochastic variational inference, with inducing variables u working as global variables and dependencies between the observations [18, 19]



stochastic process analysis in collision dynamic simulations, resulting in traditional dynamic simulation methods either requiring a lot of calculation time or lacking consideration of randomness and physical interpretability. Therefore, the existing stochastic process analysis methods for collision dynamic simulations of railway trains are still facing many challenges in maintaining high efficiency and accuracy while ensuring physical interpretability.

1.3 Motivation, contribution and structure

Considering different kinds of stochastic variability noise factors [26–28] in the dynamic simulation of railway vehicles collision, we propose a new data-driven stochastic process modelling approach to reduce the process disturbance and acquire accurate results using a simpler model. The major contributions of our study include:

- Developing four kinds of kernels in the stochastic process model to describe different disturbances, namely, the trend uncertainty, the time uncertainty, the smaller irregularities and the data sources uncertainty.
- Using stochastic variational inferences and mini-batch algorithms, the model could effectively explain every single detail of the stochastic process.
- Proposing a DSPM approach for collision dynamic simulation of railway vehicles to accurately describe the stochastic processes.
- Obtaining a higher computing accuracy and efficiency of the DSPM approach compared with that of GPR and FE methods.

The rest of this paper is organized as follows. In Sect. 2, the FE model of railway vehicles is established for collision simulation. A detailed introduction of the DSPM approach is given in Sect. 3. In Sect. 4, an overview of how the DSPM approach can be applied to stochastic dynamic simulation is provided. Next, two collision scenarios are utilized to evaluate the performance of our proposed model, and a detailed comparative analysis of three simulation methods, namely, the DSPM approach, GPR method and FEA method, is elaborated in Sect. 5. In Sect. 6, the challenges faced by the DSPM approach are analysed. Finally, conclusions and discussion on this research work are drawn in Sect. 7.

2 FE simulation

2.1 Modelling

In this paper, we take a representative subway vehicle as the main research object, which has a maximum design speed of 100 km/h. Generally, the car body of the vehicle mainly consists of two side walls, an underframe, a roof frame, and two end walls. In addition, the structure of the body-in-white (BIW) and the chassis should consider static and dynamic loads. The layout of a spacious space in the end area is designed as energy-absorbing areas to prevent large deformations, and the passenger area is designed in the middle part. Figure 2a shows the three views (i.e. bottom, left, and front) of the FE model of the railway vehicle collision. There are eleven different cross-sectional characteristics of thickness for the plates and beams, which are all indicated by different colours for easy inspection.

The FE simulations of railway vehicle collision scenarios are conducted using full-scale detailed models which contain a total of 182 different components. The car body has a length, width, and height of 24,000, 3200, and 3100 mm, respectively, and a distance 20,000 mm between the centres of the bogies. Specifically, the front section and the end of the driver's cab are deformable and consist of quadrilateral shell elements with an accuracy of 20 mm, and the rest of the vehicle consists of quadrilateral shell elements with an accuracy of 50 mm. The composition of the FE model is shown in Fig. 2b. In FE modelling, the mass of the bogie and appropriate boundary conditions are considered, as the bogie (Fig. 2c) is the critical component supporting the vehicle. In this case, the FE model is composed of 819,389 quadrilateral shell elements (approximately 95.5% of the total number of elements) and 37,742 rigid connection elements (welding, joining, etc.). In addition, the weight of the railway vehicle and BIW is 46 t and 7.03 t (about 15.2% of the total weight), respectively.

In the FE model, there are two connection methods: one is the common node; the other is Reb2. The static and dynamic friction coefficients of the wheel-rail contact are taken as 0.15 and 0.1, respectively; and the other corresponding friction coefficients are taken as 0.2 and 0.15, respectively. The contact forms in this FE model include (i) surface-to-surface contact (the energy absorbing device and the rigid wall, wheel-rail, etc.) and (ii) single-surface contact (the energy-absorbing device itself, etc.).

The FE model of the energy-absorbing devices, containing anti-climbing teeth, energy-absorbing beams, aluminium honeycombs, ribs, etc., is established. In this FE model, the aluminium honeycomb, the energy-absorbing beam, and rib structure are refined into a unit with a size of 3, 4, and 4 mm, respectively, and the total mass is 75 kg. Then the FE simulation of the energy-absorbing device impacting the front fixed rigid wall at an initial velocity V_0 was implemented(Fig. 3a). In the FE model of railway vehicles, the Z degree of freedom of the bogie-to-car body is constrained to avoid the overall Euler buckling deformation, and the rigid wall is constrained with 6 degrees of



Fig. 2 FE model of railway vehicles collision: a three views of the FE model; b composition of the FE model; c FE model of the bogie

freedom. Finally, the FE model of the lead car impacts the rigid wall at initial velocity V_0 was established, as shown in Fig. 3b. In addition, the material parameters of the components in these two FE models are listed in Table 1.

Meanwhile, there are nonlinear connections in the FE model that exist for the coupler buffer device and for the anti-climbing device. As shown in Fig. 4a, the solid line and the dashed line represent the loading and unloading process of the coupler buffer device during the collision, respectively. Using nonlinear hysteresis characteristic curves (Fig. 4b), a mathematical model of the anti-climbing device can be built to simulate its mechanical characteristics.

2.2 Model verification

Considering that railway vehicle collision is a complex problem in terms of material, contact, and geometric nonlinearity [29], we conducted the crash experiment on the energy-absorbing device to verify the accuracy of the FE model. A frontal collision experiment was conducted with the energy-absorbing device at an initial speed of 16.21 m/s, with a rebound speed of 0.95 m/s. Similarly, the instantaneous velocity V_0 of the energy-absorbing device before collision in the FE simulation was set to 16.21 m/s. The transient impact force test system was used to measure the impact force in real time and the high-speed camera system was used to record the whole impact process. Afterward, sequence image tracking was performed on the energy absorption device.

Finally, dynamic responses such as the displacement, velocity, acceleration, and impact force during the collision process were obtained. The FE model of the energy-absorbing device (Fig. 5a–b) and the representative FE model of railway vehicle (Fig. 5c–d) were carried out on the Ls-Dyna platform in Windows with Core i7, 3.4 GHz processor, and 8 GB RAM. Figure 6 illustrates detailed



Fig. 3 Collision condition: a energy-absorbing component to a rigid wall; b lead car to a rigid wall

 Table 1 Material parameters of the components employed in FE simulation

Component	Material setting	Yield strength (MPa)	Density $(\times 10^{-9})$	Young's modulus E (GPa)	Poisson's ration <i>v</i>	
Bogie framework	Mat.20	_	46.8	206	0.30	
Rail	Mat.20	-	7.85	306	0.30	
Endwell	Mat.24	125	2.70	68	0.30	
Car body	Mat.24	205	2.70	68	0.30	
Underframe	Mat.24	245	2.70	68	0.30	
Chassis	Mat.24	205	5.74	68	0.30	
Beams	Mat.24	235	7.85	210	0.30	
Honeycombs	Mat.3	292	2.70	69	0.30	

comparisons between experimental and FE simulation results of the energy-absorbing device.

As it can be seen in Fig. 6, collision dynamic responses of the FE simulation are highly consistent with the results of the experimental test, indicating the accuracy of the simulation results. The maximum displacement of the experimental test and FE simulation is 461.99 and 459.58 mm, respectively, and the maximum error is 0.52% (Fig. 6a). Figure 6b shows that the collision velocity of the experimental test and FE simulation drops rapidly to zero in the initial stage, and then it rebounds at a low velocity until the collision process ends. The acceleration results of the experiment are consistent with those of the FE simulation, as shown in Fig. 6c. Figure 6d shows that after SAE 1400 Hz filtering, the corresponding average values of the simulated collision force and the tested forces are 491.90 and 450.81 kN, respectively, and the error between them is 6.8%; the variation trend of the simulated collision force agrees well with the experimental collision force. Figure 6e–f shows that the absorbed energy of the experimental test and the FE simulation is 199.69 and 188.36 kJ, respectively, and the error is 6.01%, which is within the acceptable range. As a result, the FE model can effectively and truly reflect the collision process of railway vehicles.

3 DSPM approach

In FE simulation, collision dynamic responses are often obtained by inputting relevant characteristics such as initial velocity. Inspired by this modelling theory, a data-driven stochastic process model was built to estimate Y = f(X), where $f(\cdot)$ is the base function that converts the input relevant vector X into the output collision dynamic response Y. The modelling process of DSPM approach mainly can be divided into three stages, including data collection and extraction, model training, and model prediction. As shown in Fig. 7, the training data were collected from FE simulation, and then the collected data was extracted and recognized at the stage of data collection and extraction. In the process of model training, parameters of the collected data were initialized by the maximum



Fig. 4 Mechanical characteristics of the nonlinear connection: a characteristics of the coupler buffer device; b nonlinear hysteresis characteristic curve of the anti-climbing energy-absorbing device



Fig. 5 Collision experiments and FE simulation: **a** FE model of energy absorbing device, t = 0 ms; **b** FE model simulation results of energy absorption device at an initial speed of 60 km/h, t = 50 ms; **c** FE model of railway vehicle, t = 0 ms; **d** FE simulation results of the railway vehicle at 60 km/h, t = 50 ms

likelihood estimation (MLE) algorithm using mini-batches. In the process of model prediction, the initial parameters from the trained model were used for model prediction. Finally, the estimated output values, including mean and deviation at the new point, can be obtained.

3.1 Stochastic process analysis

In the DSPM approach, we suppose that there exist some sets of sample data $X = (x_1, x_2, ..., x_n)$ and $X' = (x'_1, x'_2, ..., x'_n)$ with the corresponding set of observed outputs $Y = (y_1, y_2, ..., y_n)$, and the prior assumption can be expressed as



Fig. 6 Comparison results: a displacement; b velocity; c acceleration; d collision force; e force-displacement curve; f absorbed energy

$$y(\boldsymbol{X}) \sim \mathcal{GP}(\boldsymbol{u}, k(\boldsymbol{X}, \boldsymbol{X}'; \boldsymbol{\theta})), \tag{1}$$

where $y(\cdot)$ is the output function, $\mathcal{GP}(\cdot)$ represents the traditional Gaussian process, \boldsymbol{u} is the mean value, $k(\boldsymbol{X}, \boldsymbol{X}'; \boldsymbol{\theta})$ is a covariance function with the vector of hyper-parameters $\boldsymbol{\theta}$, which also represents the kernel function [30]. In addition, let's postulate there is a dataset $\{\boldsymbol{H}, \boldsymbol{Z}\}$ with

$$\mathbf{Z} \sim \mathcal{N}(\boldsymbol{m}, \boldsymbol{S}),\tag{2}$$

where Z is the hypothetical data, $\mathcal{N}(m, S)$ represents a

normal distribution with mean value m and covariance matrix S. Here, $H = \{h^i\}_{i=1}^n$ and $Z = \{z^i\}_{i=1}^n$. Then, a stochastic process f(X) can be obtained by the conditional distribution, as follows:

$$f(\boldsymbol{X}) = y(\boldsymbol{X})|\boldsymbol{m}, \boldsymbol{S},\tag{3}$$

$$f(\mathbf{X}) \sim \mathcal{GP}(\mu(\mathbf{X}; \boldsymbol{\theta}, \boldsymbol{m}), \Sigma(\mathbf{X}, \mathbf{X}'; \boldsymbol{\theta}, \mathbf{S})),$$
 (4)

where



Fig. 7 Framework and data flow of DSPM approach

$$\mu(\boldsymbol{X};\boldsymbol{\theta},\boldsymbol{m}) = k(\boldsymbol{X},\boldsymbol{H};\boldsymbol{\theta})k(\boldsymbol{H},\boldsymbol{H};\boldsymbol{\theta})^{-1}\boldsymbol{m},$$
(5)

$$\Sigma(\boldsymbol{X},\boldsymbol{X}';\boldsymbol{\theta},\boldsymbol{S}) = k(\boldsymbol{X},\boldsymbol{X}';\boldsymbol{\theta}) - k(\boldsymbol{X},\boldsymbol{H};\boldsymbol{\theta})k(\boldsymbol{H},\boldsymbol{H};\boldsymbol{\theta})^{-1}k(\boldsymbol{H},\boldsymbol{X}';\boldsymbol{\theta}) + k(\boldsymbol{X},\boldsymbol{H};\boldsymbol{\theta})k(\boldsymbol{H},\boldsymbol{H};\boldsymbol{\theta})^{-1}Sk(\boldsymbol{H},\boldsymbol{H};\boldsymbol{\theta})^{-1}k(\boldsymbol{H},\boldsymbol{X}';\boldsymbol{\theta}).$$
(6)

3.2 Parameter Initialization with MLE

To update and get the hyper-parameters θ and noise variance parameters η_*^2 , we apply the gradient ascent on the log-marginal-likelihood. The log-likelihood function of the initial parameter can be expressed as follows [31]:

$$\mathcal{L}(\boldsymbol{\theta}, \eta_*^2) = \frac{1}{2} \boldsymbol{m}^{\mathrm{T}} k(\boldsymbol{H}, \boldsymbol{H}; \boldsymbol{\theta})^{-1} \boldsymbol{m} + \frac{1}{2} \log |k(\boldsymbol{H}, \boldsymbol{H}; \boldsymbol{\theta})| + \frac{n}{2} \log(2\pi).$$
(7)

Let $\partial \mathcal{L}(\theta, \eta_*^2)/\partial \mu(H; \theta, 0) = 0$ and $\partial \mathcal{L}(\theta, \eta_*^2)/\partial \Sigma(X, X'; \theta, S) = 0$. We can initialize the training procedure by hypothesizing that $m_0 = 0$ and $S_0 = k(H, H; \theta_0)$, where θ_0 is the initial set of hyperparameters, m_0 is the initial mean value, and S_0 is the initial covariance matrix. Then $\mu(H; \theta, m) = m$ and $\Sigma(H, H; \theta, S) = S$ of MLE can be obtained by optimizing the hyper-parameters and noise variance parameters of the model.

3.3 Model prediction

The mean m and covariance matrix S in the stochastic process are key parameters to our DSPM approach, and the

parameters can be estimated after parameter initialization. It is convenient to apply efficient mini-batch training to replace the original dataset. Therefore, we can update the mean m and covariance S of the hypothetical dataset $\{H, Z\}$ by adopting the posterior distribution of a mini-batch of data $\{I, J\}$ with size M, as shown Eq. (8) and Eq. (9):

$$\mu(\boldsymbol{H};\boldsymbol{\theta},\boldsymbol{m}) + \Sigma(\boldsymbol{H},\boldsymbol{I};\boldsymbol{\theta},\boldsymbol{S}) \big(\Sigma(\boldsymbol{I},\boldsymbol{I};\boldsymbol{\theta},\boldsymbol{S}) + \sigma_{\epsilon}^{2} \boldsymbol{I} \big)^{-1} [\boldsymbol{J} - \mu(\boldsymbol{I};\boldsymbol{\theta},\boldsymbol{m})] \rightarrow \boldsymbol{m},$$
(8)

$$\begin{split} & \Sigma(\boldsymbol{H},\boldsymbol{H};\boldsymbol{\theta},\boldsymbol{S}) \\ & -\Sigma(\boldsymbol{H},\boldsymbol{I};\boldsymbol{\theta},\boldsymbol{S}) \big(\Sigma(\boldsymbol{I},\boldsymbol{I};\boldsymbol{\theta},\boldsymbol{S}) + \sigma_{\epsilon}^{2}\boldsymbol{I} \big)^{-1} \Sigma(\boldsymbol{I},\boldsymbol{H};\boldsymbol{\theta},\boldsymbol{S}) \\ & \to \boldsymbol{S}. \end{split}$$
(9)

In this manner, we can predict the solution at a new test point \mathbf{x}^* through the mean $\mu(\mathbf{x}^*; \theta, \mathbf{m})$ and the predicted variance $\Sigma(\mathbf{x}^*, \mathbf{x}^*; \theta, \mathbf{S})$ according to the mean value \mathbf{m} and covariance matrix \mathbf{S} of the hypothetical data, where μ and Σ are obtained from Eq. (5) and Eq. (6), respectively. For more information, please refer to the pseudocode in Table 2.

4 Stochastic dynamic simulation of railway vehicle collisions

A detailed overview of stochastic dynamic simulation of railway vehicle collision using the DSPM approach is presented in Fig. 8. As can be seen from the schematic diagram, the framework can be divided into four sections, i.e. data collection and normalization, stochastic process modelling, DSPM training, and stochastic dynamic simulation and prediction. Specifically, the section of data collection and normalization includes modules of raw data collection and data preprocessing; the section of stochastic process modelling contains modules of core extractor, stochastic model training, and hyper-parameters determination; the section of DSPM training includes modules of model training and evaluation; the section of stochastic dynamic simulation and prediction contains modules of model validation and prediction, etc. Besides, an outline of the main procedures of each key section is also provided.

4.1 Data collection and normalization

Accordingly, the database of our DSPM approach has l calculation cases. Case i is calculated at the initial velocity $\mathbf{v}_0^k(k = 1, 2, ...l)$; then the dynamic response curve can be extracted and expressed as $\mathbf{p}_i(t) = (\mathbf{d}_i(t), \mathbf{v}_i(t), \mathbf{f}_i(t), \mathbf{E}_i(t))(t = 1, 2, ..., \varepsilon)$, where $\mathbf{d}_i(t), \mathbf{v}_i(t), \mathbf{f}_i(t)$ and $\mathbf{E}_i(t)$ represent the displacement, velocity, interface force, and internal energy at the time t, respectively, while ε corresponds to the maximum value of time. If we express the dynamic response curve in case i as $g_k(\cdot)$, we can then write

$$\boldsymbol{p}_{i}(t) = g_{k}(t, \boldsymbol{v}_{0}^{i})(t = 1, 2, \dots, \varepsilon; i = 1, 2, \dots, l).$$
(10)

Denoting the new velocity value by v_0^{New} , the undetermined dynamic response curve for a new case can be expressed as

$$\boldsymbol{p}^{\text{New}}(t) = g^{\text{New}} \left[t^{\text{New}}, \boldsymbol{\nu}_0^{\text{New}} \right], \tag{11}$$

where $p^{\text{New}}(t)$ represents the time series of the dynamic response when a new initial velocity v_0^{New} and time t^{New} are given.

After gathering the sample data, we need to normalize the data to reduce the sensitivity of the DSPM approach. Then the feature extraction is performed on the sample data after post-processing to create the training data.

4.2 Stochastic process modelling

In this paper, we illustrate four complex kernels with hyperparameter optimization to represent different noise properties [30–33].

• A temporal trending term is to be explained by a radial basis function (RBF) kernel [32]. The RBF kernel is given by

$$k_{\text{RBF}}(\boldsymbol{X}, \boldsymbol{X}'; \boldsymbol{\theta}) = \gamma^2 \exp\left(-\frac{1}{2} \sum w^2 (\boldsymbol{X} - \boldsymbol{X}')^2\right), \quad (12)$$

where γ^2 represents a variance parameter, and *w* represents a scalar of an isotropic variant of the kernel.

• A time component is expressed by the periodic ExpSineSquared kernel with a periodicity parameter q(q > 0) [33], and the kernel is given by

Table 2 Pseudocode of our data-driven stochastic process modelling approach

Input: All available input sample data X = (x₁, x₂, ..., x_n), X' = (x'₁, x'₂, ..., x'_n) with the corresponding set of observed outputs, Y = (y₁, y₂, ..., y_n).
Output: The mean value and variance value of the new sample data.
Steps:
Step 1: Preprocess the sample dataset-normalization and outlier removal.
Step 2: Define a stochastic process by the conditional distribution of a hypothetical dataset.
Step 3: Assume four kinds of Kernel functions to explain different noises in the data-driven stochastic process modelling.
Step 4: Parameters initialization and update through taking steps proportional to the gradient of the logMLE.
L(θ, η²_{*}) = ¹/₂m^Tk(H, H; θ)⁻¹m + ¹/₂log|k(H, H; θ)| + ⁿ/₂log(2π)
Step 5: Employ the posterior distribution of the observed mini-batch of data {I, J}.
Step 6: Repeat Step 4 and Step 5 until optimum parameters are updated using AdamOptimizer. Then a new test point x* can be predicted by giving the mean μ(x*; θ, m) and variance Σ(x*, x*; θ, S).

End



Fig. 8 Framework of stochastic dynamic simulation of railway vehicle collision

$$k_{\text{ESS}}(\boldsymbol{X}, \boldsymbol{X}'; \boldsymbol{\theta}) = \gamma^2 \exp\left(-2w^2 \sin^2(\pi(\boldsymbol{X} - \boldsymbol{X})^2/q)\right)^2.$$
(13)

Smaller irregularities are to be explained by a Rational Quadratic kernel, which could be better expressed than a RBF kernel component. The Rational Quadratic can accommodate several length scales [34], and can be parameterized by a scale mixture parameter α > 0, the definition is as follows:

$$k_{\rm RQ}(\boldsymbol{X}, \boldsymbol{X}'; \boldsymbol{\theta}) = \gamma^2 \exp\left(1 + \frac{\alpha}{2} w^2 (\boldsymbol{X} - \boldsymbol{X}')^2\right)^{-\alpha}.$$
 (14)

At last, we applied the White kernel as the part of a sum-kernel which explains the noise as independently and identically normally distributed [35]. Overall, using these four kernels, it is possible to explain the noise present in the real-world collision process.

4.3 Model training

The training datasets are generally divided into three sample sets, i.e. training sample set, validation sample set, and testing sample set. Using the DSPM approach, we can predict the confidence intervals using just two sample sets, namely, training sample data and validating sample data. To improve the accuracy of our model, we employ the kfold cross-validation method [36, 37], where k-1 subsets of training datasets are used for training and validating. Specifically, in this paper, the FE simulation results from different initial impact velocities may be used as training data for our model. Different training sessions are performed to train and validate the model; then the hyperparameters referred to in Sect. 3 can be obtained according to the validation results. Consequently, a stochastic model with confidence in 95% is selected during the training sessions of the model, and then the relationship between the inputs and the outputs is built by training the datadriven stochastic process model.

4.4 Simulations and evaluation

To validate this model, parameters are optimized to minimize the mean square error (MSE) between the output value and the expected value. Then an optimized datadriven model is built until the MSE meets our demands, which can not only predict the dynamic characteristics but also give the dynamic response distribution considering stochastic factors.

Let $X_t = (x_t^{(1)}, x_t^{(2)}, \dots, x_t^{(M)})^T$ and $Y_t = (y_t^{(1)}, y_t^{(2)}, \dots, y_t^{(M)})^T$ denote the inputs and outputs of test samples, respectively, where *M* is the size of the test samples, and $\hat{y}_t^{(i)} = \hat{g}(x_t^{(i)})$ denote the *i*th predicted output of the data-driven model.

We employ four different metrics to quantitatively characterize the prediction accuracy of our proposed model, among which the MSE [38] is calculated as follows:

MSE =
$$\frac{1}{M} \sum_{i=1}^{M} \left(\hat{y}_{t}^{(i)} - y_{t}^{(i)} \right)^{2}$$
. (15)

By taking the squared root of MSE, we can obtain the root mean squared error (RMSE) [39] as in Eq. (16):

RMSE =
$$\sqrt{\frac{1}{M} \sum_{i=1}^{M} \left(\hat{y}_{t}^{(i)} - y_{t}^{(i)}\right)^{2}}.$$
 (16)

Another metric we apply to assess the model's prediction accuracy is the mean absolute error (MAE) [40] as defined in Eq. (17):

MAE
$$= \frac{1}{M} \sum_{i=1}^{M} \left| \hat{y}_{t}^{(i)} - y_{t}^{(i)} \right|,$$
 (17)

Finally, the goodness of fit R^2 [41] can be obtained in accordance with Eq. (18):

$$R^{2} = 1 - \frac{\sum \left(y^{(i)} - \hat{y}^{(i)}\right)^{2}}{\sum \left(y^{(i)} - \bar{y}\right)^{2}}.$$
(18)

In these approaches, we used the initial conditions as the input to feed the model, and the predicted mean and deviation results as the output. Once this model has been trained, it can be used to predict the mean values and deviations of dynamic response under varying initial conditions. Finally, two illustrative examples were provided, and the prediction results were compared with those obtained by two baseline methods, i.e. the FE simulation and GPR, to illustrate the accurateness of our model. A detailed discussion regarding the stochastic model is presented in Sect. 5.

5 Illustrative example analysis

We used two collision scenarios in this section to investigate the dynamic simulations of railway vehicle collisions, one in which the lead car collides with a rigid wall and one in which the lead car collides with another lead car. An analysis of the basic trend and distribution range of the results obtained from the stochastic process analysis was conducted to validate the accuracy of the DSPM method. All the cases in this paper were carried out on the Windows platform with Core i7, 3.4 GHz processor, and 8 GB RAM.

5.1 Collision scenario set up

We designed two different collision events in order to encompass the vast majority of aspects of real train collision scenarios. The first scenario involves the collision between the lead car and a rigid wall, and the second one is the collision between one lead car and another lead car, where the moving lead car A crashes with the static lead car B. The specific collision scenario is shown in Fig. 9.

The initial velocities are generally limited to below 100 km/h in line on actual conditions [42]. For a lead car, the effect of head-on-train impact (case 2) with an initial velocity v is almost equivalent to one of rigid-wall impact (case 1) with an initial velocity of 2v. Since drivers can trigger an emergency brake before the collision happens, the initial velocity is less likely to reach the maximum running velocity. Therefore, the initial velocity range for case 1 was set from 10 to 60 km/h in practical terms, and the collision time was set to 300 ms for case 1 to simulate the scenario of a high-speed collision, covering a history of 300 sample states (sampling period = 1 ms). Similarly, case 2 belongs to the general cases where the initial velocity ranges from 20 to 50 km/h [43], for which the collision time was set to 100 ms, covering 1000 sample states (sampling period = 1 ms). After the data was generated, it was fed into the DSPM approach described in Sect. 3 to predict the accurate dynamic response curve with deviations. Table 3 lists the parameters of the two collision scenarios in our work.

5.2 Model training

In case 1, we applied the FE method and the DSPM approach to calculate the dynamic response of railway vehicles, including displacement, velocity, interface force, internal energy, and kinetic energy at six different impact velocities (Table 3). Comparisons of dynamic responses for



Fig. 9 Two collision scenarios: a lead car to rigid wall; b lead car to lead car

displacement, velocity, interface force, internal energy, and kinetic energy are depicted in Fig. 10.

Figure 10 illustrates that the displacement of the lead car gradually increases over time, and then decreases slowly, with a tendency to become flat after 0.3 s. Within 0.15 s after the collision, the single car's velocity gradually decreases until the vehicle comes to a standstill. The internal energy gradually increases and reaches its maximum value at 0.1 s, while the kinetic energy gradually decreases to a minimum value at the same time. The interface force curve generally goes through three stages: rapid increase, then stabilize, and finally rapid decrease. When the impact velocity exceeds 50 km/h, the interface force rises rapidly to a peak and quickly decreases, showing the evolution of the interface force from three stages to two stages. As can be seen, the fluctuation of the interface force is within 95% confidence interval, and the dynamic responses of our model are fundamentally consistent with the trend of the FE analysis results for validation and prediction. Figure 10f shows that the goodness-of-fit (R^2) of displacement, velocity, interface force, internal energy, and kinetic energy for case 1 is 0.998, 0.992, 0.961, 0.996, and 0.987, respectively, implying that this model can fit well with these dynamic response.

In case 2, we also applied the FE method and DSPM approach to calculate the dynamic responses at three different impact velocities, including displacement, velocity, interface force, internal energy, and kinetic energy of the lead car A. The compared results between the FE method and DSPM approach of lead car A are depicted in Fig. 11.

As shown in Fig. 11, after 1000 ms of collision, the displacement of lead car A shows an increasing trend, while its velocity presents a decreasing trend. The interface

Table 3 Parameters of the collision scenario

Collision scenario No.	Collision scenario	Velocity (km/h)
1	Lead car to rigid wall	10, 20, 30, 40, 50 and 60
2	Lead car to lead car	20, 40 and 50

force approaches zero after experiencing cyclic fluctuations. The internal energy gradually increases and reaches the maximum value at 0.5 s after collision, while the kinetic energy gradually decreases to the minimum value at the same time. Figure 11a-e also shows that the dynamic response results of our model are in essence consistent with the trend of the FE method results, and the fluctuation of the dynamic response results is within 95% confidence interval (marked as orange area). Figure 11f shows that the goodness-of-fit (R^2) of displacement, velocity, interface force, internal energy, and kinetic energy of lead car A in case 2 is 0.999, 0.944, 0.950, 0.993, and 0.927, respectively, indicating that our model has good performance in fitting the collision dynamic response curves of railway vehicles. As a result, the comparisons between the DSPM approach and the FE method illustrate that the dynamic responses obtained from the FE method are essentially within the range predicted by the DSPM approach, especially displacement, velocity, internal energy, and kinetic energy.

5.3 Model validation

To verify the model, we applied the DSPM to predict and analyse the dynamic response of rail vehicles at a velocity of 35 km/h in case 1 and case 2, and then the results were compared with that of the FE method. In this section, we focus to introduce the collision dynamic response of displacement, velocity, internal energy, and kinetic energy. The predicted results in case 1 is shown in Fig. 12, and the predicted results in case 2 is shown in Fig. 13.

The predicted mean values (marked as red line in Fig. 12) and deviation ranges (marked as orange in Fig. 12) of the dynamic responses are obtained through training 1806 samples in case 1. Figure 12 shows that even though the relative error of displacement and internal energy is relatively large at the beginning due to the numerical changes around the initial stage, the predicted collision dynamic responses are consistent with the trend of the FE results, and the fluctuation of the dynamic response results



Fig. 10 Dynamic responses of the lead car in case 1:. a displacement; b velocity; c interface force; d internal energy; e kinetic energy; f average goodness-of-fit of case 1



Fig. 11 Dynamic response of lead car A in Case 2: a displacement; b velocity; c interface forces; d internal energy; e kinetic energy; f average goodness-of-fit in case 2

is also within 95% confidence interval. Meanwhile, the average relative error of displacement, velocity, internal energy, and kinetic energy are 0.022, 0.125, 0.037 and 0.25, respectively, which can verify the accuracy capability of the DSPM approach.

In case 2, the predicted results of lead car A, including the mean-values (marked as red line in Fig. 13) and deviation ranges (marked as orange in Fig. 13) are obtained through training 3303 samples, and then the comparison between the DSPM and FEM results is analysed. The relative errors increase with the fluctuations of the predicted results which is still within 95% confidence interval (marked as orange). Moreover, the average relative error of DSPM approach in displacement, velocity, internal energy, and kinetic energy are 0.04, 0.03, 0.07, and 0.15, respectively, implying that the dynamic response can be accurately described with this approach.



Fig. 12 Predicted results and relative errors in case 1: a displacement; b velocity; c internal energy; d kinetic energy

5.4 Comparative analysis

According to the derived optimum parameters, the corresponding predicted results of case 1 and case 2 can be obtained. The model hyperparameters of mini-batch and data size are set as 1000 and 500, respectively, and the maximum iterations are set as 1000. In addition, the final predicted results of this DSPM approach are depicted and the average performance indicators are listed in Table 4. The box plot illustrating the errors is presented in Fig. 14a (case 1) and Fig. 15a (case 2), and the goodness-of-fit (R^2) results of our method and GPR method are shown in Fig. 14b (case 1) and Fig. 15b–c (case 2).

Through comparative analysis of Table 4, Fig. 14 and Fig. 15, the following results are obtained:

• FE simulation is the most commonly used method in engineering, but its simulation time is almost 10 times

longer than that of other two methods, so this approach suffers from a high computational cost.

- Generally, GPR method is used to deal with stochastic processes, because of its strong fitting ability. Otherwise, it is difficult to apply and promote in engineering due to its insufficient predictive ability and high computational complexity.
- Our proposed DSPM approach with lower error and higher R^2 , has an obvious advantage of efficiency and accuracy compared with FE and GPR method.

Overall, as can be seen from these two collision cases, the proposed DSPM approach shows its higher prediction accuracy and extrapolating capability. Moreover, the comparative results indicate that our proposed method is generally more efficient and robust than the FE and GRP methods.



Fig. 13 Predicted results and relative errors of lead car A in case 2: a displacement; b velocity; c internal energy; d kinetic energy

6 Challenges faced by DSPM approach

6.1 Insufficient training samples

In fact, we are still facing some challenges, such as how to obtain valuable training samples. In the future, few-shot learning such as meta-transfer learning may be introduced to obtain more valuable training samples.

6.2 Inappropriate training strategy

Due to the challenging training requirements, the datadriven approach is not yet applicable in many circumstances. In the future, the training strategy should be systematically studied, such as hidden layers, network nodes, transfer functions and other hyperparameters, etc.

6.3 Poor interpretability

Interpretability has always been a key issue affecting the credibility and usability of data-driven approaches. In the future, the hybrid modelling method that combines data science and physical principles will be an important research prospect.

7 Conclusion

In this paper, we propose a novel DSPM approach for collision dynamic simulation to accurately predict and analyse the dynamic response with a confidence interval in the collision process. The main conclusions are summarized as follows:

(1) Considering different kinds of stochastic noise, the DSPM approach could provide high-precision

Table 4 Comparisons between our DSPM approach and FE and GPR methods

Collision scenario No	Dynamic response	Method	Training set				Testing/simulation set					
			Time (s)	R^2	MAE	MSE	RMSE	Time (s)	R^2	MAE	MSE	RMSE
Case 1: lead car	Displacement	FE	-	-	=	-	-	21,336.6	-	-	-	_
to rigid wall		GPR	2581.9	0.999	0.012	0.0004	0.02	0.22	0	0.31	0.166	0.408
		DSPM	861	0.998	0.025	0.001	0.034	0.09	0.99	0.03	0.001	0.039
	Velocity	FE	-	-	-	-	-	21,336.6	-	-	-	-
		GPR	1965.46	0.998	0.023	0.0013	0.03	0.23	- 2.08	0.735	0.732	0.859
		DSPM	803.19	0.992	0.064	0.007	0.088	0.025	0.995	0.050	0.003	0.058
	Interface force	FE	-	-	-	-	-	21,336.6	-	-	-	-
		GPR	1693.08	0.992	0.039	0.007	0.087	0.234	- 1.88	0.741	0.612	0.782
		DSPM	792.57	0.961	0.249	0.138	0.371	0.036	0.93	0.162	0.039	0.199
	Internal energy	FE	-	_	_	-	_	21,336.6	-	-	-	-
		GPR	2161.42	0.999	0.014	0.0006	0.025	0.28	- 0.02	0.250	0.090	0.300
		DSPM	790.81	0.996	0.037	0.003	0.059	0.055	0.993	0.020	0.0005	0.024
	Kinetic energy	FE	_	_	_	-	_	21,336.6	_	-	_	_
	25	GPR	2077.19	0.998	0.018	0.001	0.036	0.23	- 0.03	0.473	0.318	0.564
		DSPM	786.6	0.987	0.068	0.012	0.110	0.042	0.993	0.038	0.0021	0.04
Case 2: lead car	Displacement of car A	FE	_	_	_	-	_	36,300	_	_	_	_
to lead car		GPR	6348.61	0.999	0.015	0.0006	0.024	1.15	- 0.19	0.58	0.50	0.707
		DSPM	795.12	0.999	0.008	0.0001	0.013	0.16	0.993	0.052	0.003	0.055
	Displacement of car B	FE	_	_	_	_	_	36,300	_	_	_	_
		GPR	5949.48	0.997	0.02	0.002	0.02	1.37	- 1.39	0.563	0.354	0.595
		DSPM	796.15	0.999	0.010	0.0002	0.016	0.13	0.993	0.046	0.0034	0.058
	Velocity of car A	FE	_	_	_	_	_	36,300	_	_	_	_
		GPR	5750.14	0.996	0.032	0.0036	0.06	1.31	-0.81	0.617	0.409	0.639
		DSPM	793.0	0.944	0.144	0.05	0.23	0.168	0.861	0.130	0.031	0.17
	Velocity of car B	FE	_	_	_	_	_	36,300	_	_	_	_
		GPR	6696.0	0.999	0.012	0.0003	0.017	1.12	- 0.10	0.648	0.555	0.745
		DSPM	798.02	0.930	0.159	0.069	0.26	0.157	0.872	0.165	0.036	0.191
	Interface force	FE	_	_	_	_	_	36.300	_	_	_	_
	of car A & car B	GPR	6342.57	0.992	0.037	0.0077	0.087	1 41	- 0.08	0.629	0 496	0 704
		DSPM	4120.0	0.950	0.52	0.499	0.70	0.179	0.843	0.38	0.25	0.50
	Internal energy	FE	_	_	_	_	_	36.300	_	_	_	_
	of car A	GPR	6048.78	0.999	0.009	0.0003	0.019	1.21	- 3.89	0.741	0.563	0.750
		DSPM	794 04	0.993	0.047	0.006	0.082	0.175	0.851	0.119	0.018	0.136
	Internal energy	FE	_	_	_	_	_	36 300	_	_	_	_
	of car B	GPR	7571.8	0 999	0.012	0.0007	0.027	1 18	- 178	0.720	0 548	0 740
		DSPM	796 69	0.990	0.052	0.0007	0.027	0.18	0.873	0.09	0.009	0.097
	Kinetic energy	FF	190.09	0.770	0.052	0.007	0.09	36 300	0.075	0.07	0.007	0.077
	of car A	GPR	50/0 /8	0.007	0.02	0.002	0.04	1 37	_ 1 30	0 563	0 354	0 595
		DSPM	701 129	0.997	0.02	0.002	0.04	0.184	- 1.59	0.505	0.007	0.393
	V:	EE	/71.130	0.727	0.150	0.07	0.200	26 200	0.015	0.114	0.027	0.105
	of car B	GPP	-	-	-	-	-	1 25	- _ 26	- 0.53	- 0.380	-
		DCDM	700 00	0.997	0.024	0.002	0.032	0.211	- 2.0	0.55	0.309	0.02
		DSPM	/98.08	0.885	0.190	0.114	0.338	0.211	0.847	0.1/3	0.038	0.194



Fig. 14 Comparison between DSPM and CPR method in case 1: a box-plot; b R^2



Fig. 15 Comparative results analysis in case 2: a box plot; b R^2 of car A; c R^2 of car B

predicted results of dynamic response in collision dynamic simulations of railway vehicles.

- (2) The DSPM approach improves calculation efficiency by encoding a large amount of data into some minibatch data in complicated collision dynamic simulations of railway vehicles.
- (3) The comparison results demonstrated the superior performance of the DSPM approach in terms of efficiency and accuracy than FE and GRP methods.

Moreover, the scaling of data-driven stochastic processes is, and will remain, an important research area, and it may open up other applications of this approach in other stochastic dynamic simulations of railway vehicles. In the future, we plan to focus our attention on addressing the faced challenges, and refining our approach by making further modifications to the DSPM approach, including obtaining valuable training samples, training strategy analysis, improving model interpretability. Acknowledgements We would like to extend our deep gratitude to all those who have offered practical, cordial, and selfless support in writing this manuscript, especially Yuru Li offered help with FE modelling. Meanwhile, this work is supported by the National Key Research and Development Project (No. 2019YFB1405401) and the National Natural Science Foundation of China (No. 5217120056).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

- Zhu T, Xiao S, Lei C et al (2021) Rail vehicle crashworthiness based on collision energy management: an overview. Int J Rail Transp 9(2):101–131
- Zhu T, Xiao SN, Hu GZ et al (2019) Crashworthiness analysis of the structure of metro vehicles constructed from typical materials and the lumped parameter model of frontal impact. Transport 34(1):75–88
- Li R, Xu P, Peng Y et al (2016) Scaled tests and numerical simulations of rail vehicle collisions for various train sets. Proc Inst Mech Eng Part F J Rail Rapid Transit 230(6):1590–1600
- 4. Lu S, Xu P, Yan K et al (2019) A force/stiffness equivalence method for the scaled modelling of a high-speed train head car. Thin-Walled Struct 137:129–142
- Cho HJ, Koo JS (2012) A numerical study of the derailment caused by collision of a rail vehicle using a virtual testing model. Veh Syst Dyn 50(1):79–108
- Scholes A, Lewis JH (1993) Development of crashworthiness for railway vehicle structures. Proc Inst Mech Eng Part F J Rail Rapid Transit 207(1):1–16
- 7. Montgomery DC, Runger GC, Hubele NF (2009) Engineering statistics. Wiley, New York
- Tang Z, Zhu Y, Nie Y et al (2017) Data-driven train set crash dynamics simulation. Veh Syst Dyn 55(2):149–167
- 9. Dong S, Tang Z, Yang X et al (2020) Nonlinear spring-massdamper modeling and parameter estimation of train frontal crash using CLGAN model. Shock Vib 2020:9536915
- Nie Y, Tang Z, Liu F et al (2018) A data-driven dynamics simulation framework for railway vehicles. Veh Syst Dyn 56(3):406–427
- Müller M, Botsch M, Böhmländer D et al (2018) Machine learning based prediction of crash severity distributions for mitigation strategies. J Adv Inf Technol 9(1):15–24
- Li YR, Zhu T, Xiao SN et al (2021) Application of the collision mathematical model based on a BP neural network in railway vehicles. Proc Inst Mech Eng Part F J Rail Rapid Transit 235(6):713–725
- Luo R, Shi H, Teng W et al (2017) Prediction of wheel profile wear and vehicle dynamics evolution considering stochastic parameters for high-speed train. Wear 392:126–138

- 14. Lu Y, Yang S, Li S et al (2010) Numerical and experimental investigation on stochastic dynamics load of a heavy duty vehicle. Appl Math Model 34(10):2698–2710
- Souffran G, Miègeville L, Guérin P (2012) Simulation of realworld vehicle missions using a stochastic Markov model for optimal powertrain sizing. IEEE Trans Veh Technol 61(8):3454–3465
- Xu L, Zhai W (2017) A new model for temporal–spatial stochastic analysis of vehicle–track coupled systems. Veh Syst Dyn 55(3):427–448
- Hao P, Boriboonsomsin K, Wu G et al (2017) Modal activitybased stochastic model for estimating vehicle trajectories from sparse mobile sensor data. IEEE Trans Intell Transp Syst 18(3):701–711
- Hoffman MD, Blei DM, Wang C et al (2013) Stochastic variational inference. J Mach Learn Res 14(5):1303–1347
- Urtasun R, Darrell T (2008) Sparse probabilistic regression for activity-independent human pose inference. In: 2008 IEEE conference on computer vision and pattern recognition. Anchorage, AK, USA. IEEE, pp 1–8
- 20. Hensman J, Rattray M, Lawrence N D (2012) Fast variational inference in the conjugate exponential family. arXiv preprint arXiv:1206.5162
- Hensman J, Fusi N, Lawrence ND (2013) Gaussian processes for big data. arXiv preprint arXiv:1309.6835
- 22. Raissi M, Babaee H, Karniadakis GE (2019) Parametric Gaussian process regression for big data. Comput Mech 64(2):409–416
- 23. Shokravi H, Shokravi H, Bakhary N et al (2020) A comparative study of the data-driven stochastic subspace methods for health monitoring of structures: a bridge case study. Appl Sci 10(9):3132
- Prudencio EE, Bauman PT, Williams SV et al (2013) A dynamic data driven application system for real-time monitoring of stochastic damage. Procedia Comput Sci 18:2056–2065
- Jiang B, Fei Y (2017) Vehicle speed prediction by two-level data driven models in vehicular networks. IEEE Trans Intell Transp Syst 18(7):1793–1801
- Maraun D, Huth R, Gutiérrez JM et al (2019) The VALUE perfect predictor experiment: evaluation of temporal variability. Int J Climatol 39(9):3786–3818
- Fortunato AB, Bertin X, Oliveira A (2009) Space and time variability of uncertainty in morphodynamic simulations. Coast Eng 56(8):886–894
- Haigermoser A, Luber B, Rauh J et al (2015) Road and track irregularities: measurement, assessment and simulation. Veh Syst Dyn 53(7):878–957
- 29. Oprea RA (2013) A constrained motion perspective of railway vehicles collision. Multibody SysDyn 30(1):101–116
- Rasmussen CE (2004) Gaussian processes in machine learning. In: Advanced lectures on machine learning. Springer, Berlin, pp 63–71
- Shimodaira H (2000) Improving predictive inference under covariate shift by weighting the log-likelihood function. J Stat Plann Inference 90(2):227–244
- Zhou Q, Jiang P, Shao X et al (2017) A variable fidelity information fusion method based on radial basis function. Adv Eng Inf 32:26–39
- Oloyede I (2018) Bayesian classification of high dimensional data with gaussian process using different kernels. Anale Seria Informatică 16(1):164–167
- 34. Longman FS, Mihaylova L, Yang L et al. (2019) Multi-band image fusion using Gaussian process regression with sparse rational quadratic kernel. In: 22th international conference on information fusion (FUSION). Ottawa, ON, Canada. IEEE, pp 1–8

- 35. Hida T, Kuo HH, Potthoff J et al (2013) White noise: an infinite dimensional calculus. Springer, New York
- Sonnenburg S, Rätsch G, Henschel S et al (2010) The SHOGUN machine learning toolbox. J Mach Learn Res 11:1799–1802
- Fushiki T (2011) Estimation of prediction error by using K-fold cross-validation. Stat Comput 21(2):137–146
- 38. Alpaydin E (2020) Introduction to machine learning. MIT Press, London
- Yang RJ, Wang N, Tho CH et al (2005) Metamodeling development for vehicle frontal impact simulation. J Mech Des 127(5):1014–1020
- 40. Chen Q, Song X, Yamada H et al. (2016) Learning deep representation from big and heterogeneous data for traffic accident

inference. In: Proceedings of the thirtieth AAAI conference on artificial intelligence. Phoenix, Arizona, USA, February 12–17, 2016

- Besbeas P, Morgan BJT (2014) Goodness-of-fit of integrated population models using calibrated simulation. Methods Ecol Evol 5(12):1373–1382
- 42. Gao G, Wang S (2019) Crashworthiness of passenger rail vehicles: a review. Int J Crashworthiness 24(6):664–676
- Li SY, Zheng ZJ (2015) Energy-absorbing structure design and crashworthiness analysis of high-speed trains. Explos Shock Waves 35(2):164–170 (in Chinese)