Time series data interpretation for "wheel-flat" identification including uncertainties

Wen-Jun Cao[1,2]*, Shanli Zhang[1], Numa J. Bertola[2,3], I. F. C. Smith[2,3], C. G. Koh[1]

[1]Department of Civil and Environmental Engineering, National University of Singapore, 117576, Singapore

[2] ETH Zurich, Future Cities Laboratory, Singapore-ETH Centre, 138602, Singapore

[3] Applied Computing and Mechanics Laboratory (IMAC), School of Architecture, Civil and Environmental Engineering (ENAC), Swiss Federal Institute of Technology (EPFL), CH-1015, Lausanne, Switzerland

*Email: caowenjun@u.nus.edu/ wenjun.cao@arch.ethz.ch

Abstract

Train wheel flats are formed when wheels slip on rails. Crucial for passenger comfort and the safe operation of train systems, early detection and quantification of wheel-flat severity without interrupting railway operations is a desirable and challenging goal. Our method involves identifying the wheel-flat size by using a model updating strategy based on dynamic measurements. Although measurement and modeling uncertainties influence the identification results, they are rarely taken into account in most wheel-flat detection methods. Another challenge is the interpretation of time series data from multiple sensors. In this paper, the size of the wheel flat is identified using a model falsification approach that explicitly includes uncertainties in both measurement and modeling. A two-step important-point-selection method is proposed to interpret high-dimensional time series in the context of inverse identification. Perceptually important points, which are consistent with the human visual identification process, are extracted and further selected using joint entropy as an information-gain metric. The proposed model-based methodology is applied to a field train-track test in Singapore. The results show that the wheel flat size identified using the proposed methodology is within the range of true observations. Additionally, it is also shown that the inclusion of measurement and modeling uncertainties is essential to accurately evaluate the wheel-flat size because identification without uncertainties may lead to an underestimation of the wheel-flat size.

Keywords

Wheel-flat identification; measurement and modeling uncertainty; time-series data; model-based; joint entropy

1. Introduction

"Wheel flat" occurs when a wheel “locks up” while the train is moving, for example, due to emergency braking. Early detection of wheel flat is important to not only reduce high vibrations due to impact force but also to reduce the operational risks of railway operations. However, it is challenging to quantify the wheel-flat size without interrupting railway operations. Over the past two decades, various monitoring systems have been proposed which use the measured responses of the train-track system to detect the occurrence of wheel flats [1,2]. As it is difficult to detect wheel flats directly, indirect methods are based on strain [3,4], acceleration [5,6], rail deflection [7] and rail-seat force [8].

A significant amount of research has focused on the determination of unknown wheel-condition properties using model-based data-interpretation techniques and field measurements. It is inevitable that uncertainties exist in measurement and modeling, especially in train-track system modelling which involves rather complicated wheel-rail interaction[9].

Residual minimization is the most widely used method for parameter identification, where an optimal set of model-parameter values is obtained by minimizing the discrepancy between model predictions and measurements [10,11].The main drawback of this method is that uncertainties are assumed to be either negligible or zero mean. This leads to parameter values that are consistent only with the measurements used in the identification process. They may fail to predict measurements under other loads or at other locations. Instead of obtaining such a solution, Bayesian model updating [12–14], which has been proposed and developed since 1990’s, is able to provide statistical descriptions of predictions. However, the accuracy of the predictions depends on the knowledge of uncertainty dependency, which is usually not available when there is a combination of correlated and systematic errors as is often the case in civil engineering practice.

Compared to model-free interpretation, model-based interpretation is able to provide a physical insight into the causes of observed (measured) behavior, thus providing better quantitative information of wheel-flat sizes instead of merely providing an alert without knowing the wheel flat severity. Furthermore, model-based interpretation provides support for obtaining optimal maintenance strategies.

Recently, Goulet and Smith [15] proposed a population-based method called error-domain model falsification (EDMF), where several model instances having a unique set of plausible

parameter values are generated. EDMF involves falsifying model instances from a set of plausible model instances based on measurement and modeling uncertainties. Specifically, a model instance is falsified if the difference between its predictions and measurements is greater than the threshold derived from combining modelling and measurement uncertainties. This application has already been successfully applied to performance monitoring of bridges [16–18], wind engineering [19], water supply networks [20]and post-earthquake assessments [21]. This method outperforms residual minimization and Bayesian model updating by correctly identifying parameter values in the presence of systematic errors without the need to consider correlations between errors [15].

In wheel-flat detection, dynamic responses of interest are recorded when a railway train is running along a track. Measured values for each sensor are time series data of large size and high dimension when the data is denoted as a vector [22]. Traditional approaches usually employ the whole time series for the purpose of matching, resulting in expensive computation. However, in order to perform real-time wheel profile identification in the context of high-frequency train service, the data-interpretation process must be carried out rapidly.

Reducing the number of data points helps reduce the computational time. Resampling [23] is one of the traditional methods but its main drawback is that resampling may distort the shape of the time series if the dimensional reduction is significant. Another approach involves dividing the time series into segments and using some scalar metric (e.g. mean value) of each segment to reconstruct the time series using a fixed segment length [24] or an adapted segment length [25]. The disadvantage of this approach is that data points are no longer accurate after reconstruction, thereby adding additional uncertainty into the system identification framework. Additionally, they may smoothen out salient data points which contain important information. Other approaches map time series into another domain using discrete Fourier transform (DFT) [26], discrete wavelet transform (DWT) [27] or singular value decomposition (SVD) [28]. However, these approaches often lose important data points (e.g. peak values) and fail to capture the shape of the time series.

Chung et al. [29] introduced perceptually important point (PIP) identification to extract important points in the human visual identification process. The importance of each point of the time series is evaluated by its contribution to the overall shape of the series. This method is used for pattern matching of technical patterns in financial applications. Similar ideas could be found in other works [30,31]. As an extension of this idea, Fu et al. [32] proposed a tree structure to support incremental updating.

In the context of system identification, Papadimitriou [33] and Robert-Nicoud et al. [34] introduced information entropy (also called Shannon entropy [35]) as an information-gain metric to evaluate the expected performance of potential sensor locations. These methodologies do not account for mutual information between sensor locations. The joint-entropy metric was thus

introduced to accommodate mutual information and it was successfully applied within the model-falsification framework [36,37]. Despite the potential of these methods to select important points for wheel-flat detection, the joint-entropy metric has not yet been proposed.

The goal of this work is twofold: (1) perform wheel-flat detection considering both measurement and modeling uncertainties through the model-falsification framework, and (2) propose an approach to accept high-dimensional time series data as input into identification with reasonable computational times.

This paper is organized as follows. First, the framework of the error-domain model falsification and the difficulties when applied to time series data are presented in Section 2. Then, the proposed framework is explained in Section 3. Section 4 presents a case study in Singapore train system, and conclusions are summarized in Section 5.

2. Error-domain model falsification

In this study, a model falsification approach is adopted to provide a robust and rigorous framework for wheel-flat identification that explicitly includes both modeling and measurement uncertainties. Error-domain model falsification is a population-based identification methodology. The task can be described as identifying unknown model-parameter values $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_n]$, using measurement data $\boldsymbol{y} = [y_1, y_2, \dots, y_m]$ where n is the number of the parameters and m is the number of measurements. $g_i(\boldsymbol{\theta})$ predicts ith structural response; $\boldsymbol{\theta}^*$ denotes the true values of parameters. The true prediction should be equal to the difference between $g_i(\boldsymbol{\theta})$ and modeling uncertainty $\epsilon_{i,model}$. It is also equal to the difference between measurement y_i and measurement uncertainty $\epsilon_{i,meas}$ (shown in Equation 1). By rearranging both uncertainties on the right-hand-side, Equation (2) is obtained.

$$g_i(\boldsymbol{\theta}) - \epsilon_{i,model} = y_i - \epsilon_{i,meas} \tag{1}$$

$$g_i(\boldsymbol{\theta}) - y_i = \epsilon_{model} - \epsilon_{meas} = U_c \tag{2}$$

The combination of both modeling and measurement uncertainties is treated as a random variable U_c described by a probability density function $f_{U_c}(\epsilon)$ [38]. In this paper, U_c is called the combined error. For candidate models (CMS), the discrepancy between its prediction and the corresponding measurement must fall inside the threshold bounds. Otherwise, the model instance is falsified.

$$T_{i,low} \leq g_i(\boldsymbol{\theta}) - y_{i,} \leq T_{i,high} \; \forall i \in \{1, \dots, m\} \tag{3}$$

Threshold boundaries are defined using Šídák correction to maintain the same reliability of identification for all numbers of sensors. Threshold bounds represent the narrowest interval in which the integral of $f_{U_{c,i}}(\epsilon)$ is $\phi^{1/m}$ and are given in Equations (4) and (5). The confidence level ϕ is usually taken as 0.95 in civil engineering [15].

$$T_{min,i} = F_{U_{c,i}}^{-1}\left(\frac{1}{2}\left(1-\phi^{1/m}\right)\right) \quad (4)$$

$$T_{max,i} = F_{U_{c,i}}^{-1}\left(1-\frac{1}{2}\left(1-\phi^{1/m}\right)\right) \quad (5)$$

where $F_{U_{c,i}}^{-1}(\cdot)$ is the inverse cumulative distribution function of $U_{c,i}$.

It is challenging to deal with time series data using EDMF. First, each time series typically contains thousands of values. The point-to-point comparison between two time series is computationally demanding. Additionally, EDMF is a population-based method. The initial model set is usually very large, requiring thousands and even millions of model instances to represent the possible combinations of model-parameter values. As shown in Equations (4) and (5), the threshold boundaries become larger with the increase in comparison points. For example, when m changes from 1 to 1000, ϕ = 0.95, $\phi^{1/m}$ changes from 0.95 to 0.99995. In this regard, Vernay et al. [39] computed the moving-average time series of measurements by replacing the measured value with the average value of its neighboring steps for the assessment of airflow in urban areas. The optimal averaging window is calculated to be short enough to be able to assume constant mean wind conditions and being long enough to avoid fluctuations of flow variables associated with turbulence. However, this approach is not applicable in the wheel-flat quantification. Additional averaging uncertainty would be relatively large compared with other uncertainties and this would adversely influence the accuracy of identified wheel-flat size.

3. Methodology

The overall goal of the methodology is to quantify the wheel-flat size using model class $g(\cdot)$ and measured data time series $\boldsymbol{P} = (P_1, P_2, \ldots, P_m)$ with trains running on rail track. The simulation model includes modeling of train-track interaction system and modeling of wheel flats. The model of the train-track interaction system is assumed to be calibrated and validated before being adopted in the wheel-flat quantification. The unknown parameter set $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_{\mathrm{d}}]$ involves the unknown geometric information of the wheel flats, the location of wheel when wheel flat hits the rail etc. Modeling and measurement uncertainties $\boldsymbol{U_g}$, $\boldsymbol{U_y}$ are quantified based on sensor knowledge, previously estimated uncertainties in the modeling method and engineering judgement.

Instead of carrying out EDMF at every data point and seeking solutions of the whole time series matching, a data-point-selection method is proposed to reduce the data size and improve the falsification efficiency. The task can be expressed as follows:

Given the p-dimensional variable $\boldsymbol{P} = (P_1, P_2, \ldots, P_m)$, *find a lower dimensional representation of it,* $\boldsymbol{S} = (S_1, S_2, \ldots, S_k)$ *with* $k \leq m$, *that captures the most important information in the original data.*

As will be explained in Section 4, the identification of wheel flat in this paper uses rail pad sensors (RPS) which measure forces due to the moving wheels on the rail pads. An example of RPS forces due to three good wheels (i.e. without wheel flat) and one bad wheel (i.e. with wheel flat) is shown in Figure 1. The force is positive if it is compression. Bogie 1 and Bogie 2 are traveling from the left to the right. In Bogie 1 (Figure 1 a-b-c) which has two good wheels (A and B), the rail force reaches twice the maximum values (I and II) when Wheel A and Wheel B cross the sleeper which is installed with a rail-pad sensor. In Bogie 2 (Figure 1 d-e-f) which has a good wheel (C) and a bad wheel (D), it is easily observed that there is a significant difference in the shape of the rail-force time history. The impact force induced by a wheel flat is dramatically greater than the force induced by a good wheel. Therefore, the shape of the time series is important for wheel flat detection.

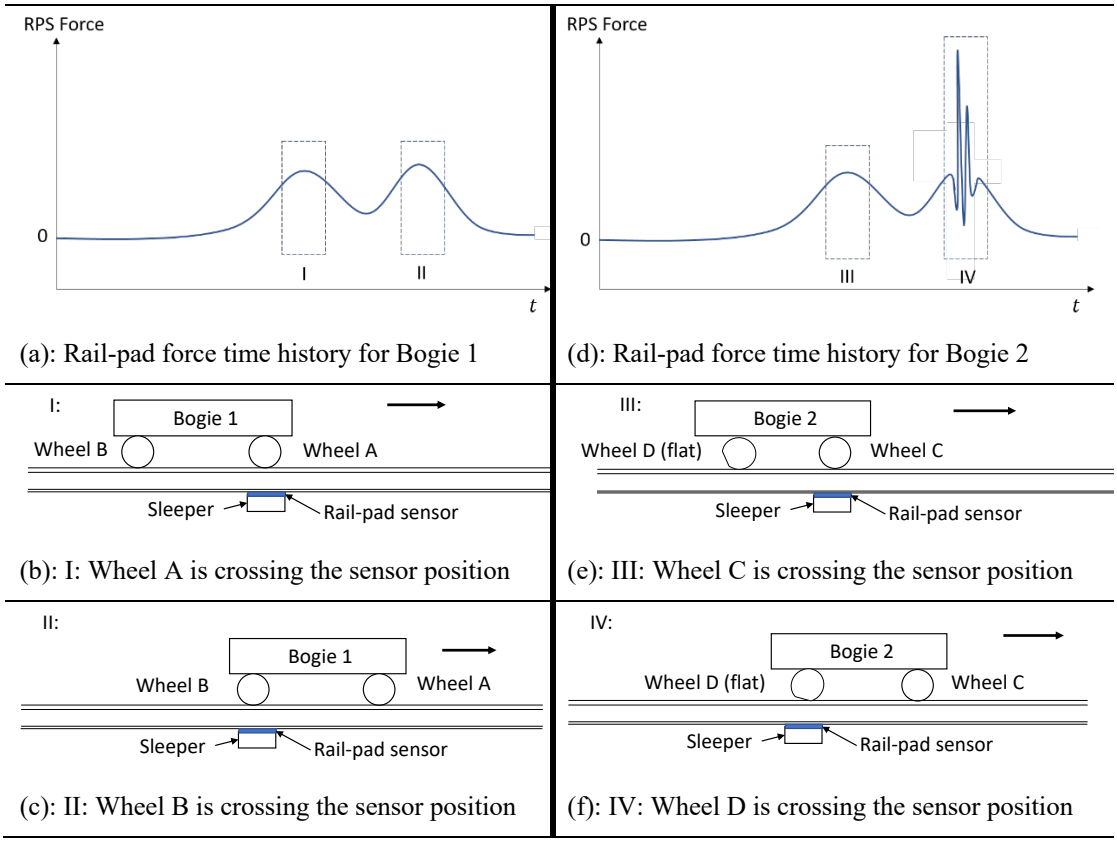


Figure 1: Sensor responses for a wheel in good condition (a-b-c) and a wheel flat (d-e-f).

Time series IV records the dynamic impact force when the wheel flat is crossing the sensor position. An enlarged view of IV is shown in Figure 2, where the process is separated into several phases. Phase 1: the wheel flat surface is not interacting with the rail and the induced

dynamic force is the same as the good wheel response; Phase 2: the wheel starts to separate from the rail due to the existence of wheel flat and the wheel impacts on the top of the rail, resulting in two peak forces (referred to as P1 force and P2 force). The P1 force is a high-amplitude impact followed by a lower amplitude P2 force [40]. Phase 3: the wheel-flat surface leaves the contact area and does not interact with the rail again. It is clear that the shape and amplitude of the time series are important to describe the contact between the wheel flat and the rail track reflecting the presence and size of the wheel flat.

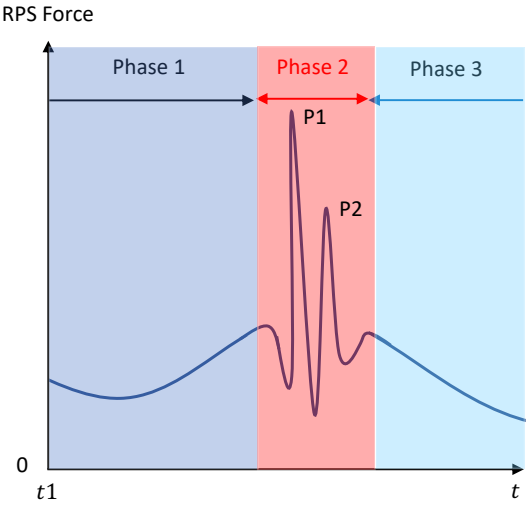


Figure 2: Rail-force time history of IV

From the model prediction perspective, the quality of information provided by several points is reflected in their capacity to distinguish between the models. A population of model instances is generated with the initial ranges defined by users based on engineering judgement and experience. These models generated using the finite element method predict (or simulate) the responses of train running on rail track. The responses are time series consisting of thousands of data points. To measure the ability of selecting data points to discriminate between model instances, the joint-entropy framework is adopted. Entropy values at each data point are calculated using the distribution of predicted model-instances responses. The initial population of model instances is used to investigate the expected information gain provided by each data point using the joint-entropy framework. Points with large entropy values indicate high disorder in predictions.

In summary, data points are defined as important when two requirements are satisfied. First, they contribute significantly to the overall shape of the time series. Second, they have great potential to identify unknown parameter values.

Accordingly, the proposed important-point selection contains two steps. First, perceptual important points (PIP) PIP $= (Q_1, Q_2, \dots, Q_j), j \leq m$ are selected based on measured data. Second, entropy-based important points EIP $= (S_1, S_2, \dots, S_k), k \leq j$ are selected based on

initial-model-set predictions among PIP. The selection process stops when the joint-entropy evaluation of the important-point set reaches its maximum values.

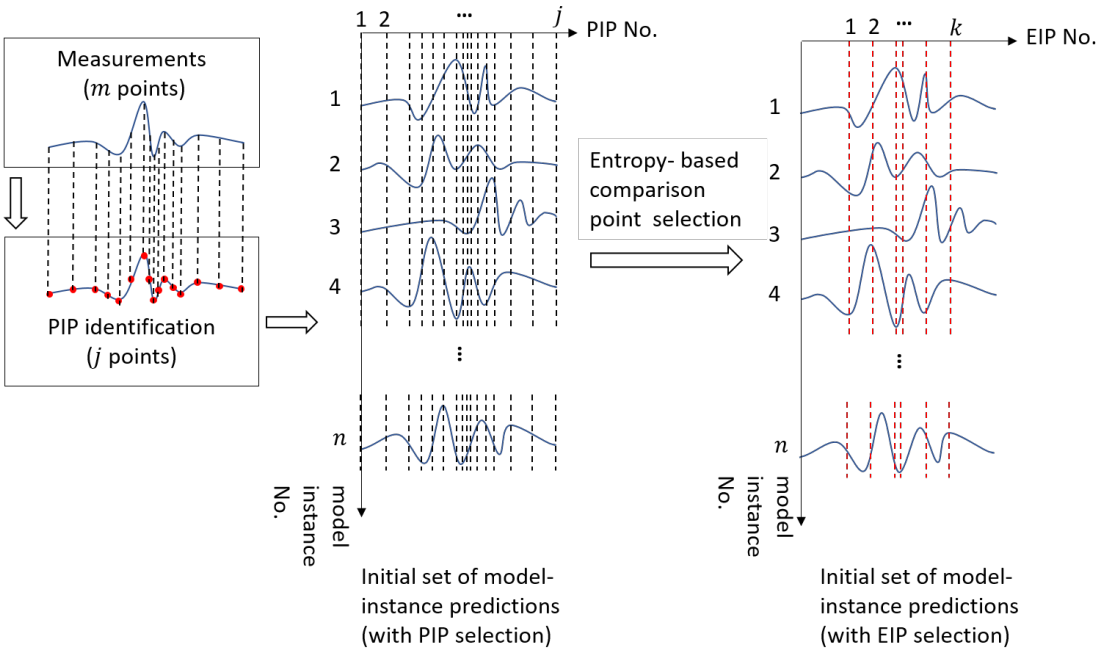


Figure 3: Schematic of the proposed important point selection

3.1 Perceptual-important point selection

The aim of this procedure is to obtain the time-series pattern without sacrificing the accuracy of the selected-data-point values. Usually, the time series pattern is characterized by the first point, the last point and some salient points such as peak values. It is consistent with the human visual identification process. However, only obvious and salient points can be captured by human visual identification. To reduce the time series to different levels of detail, the contribution of each data point to the overall shape is quantified and evaluated. The data point selected based on perceptual importance is named perceptual important data point (PIP). The concept was first introduced in the application of tracking stock performance in finance [29].

There are three ways to evaluate the importance of the PIP in a time series: Euclidean distance, perpendicular distance and vertical distance (Figure 3). Perpendicular distance and vertical distance calculate the perpendicular and vertical distances respectively between the test points and the line connecting the two adjacent PIPs. Euclidean distance is the sum of the distances between the test point and its adjacent points. Fu et al. [41] shows that Euclidean distance has the worst performance in efficiency and effectiveness compared to the other two distances.

As a result, in this paper, only perpendicular-distance and vertical-distance calculations are performed.

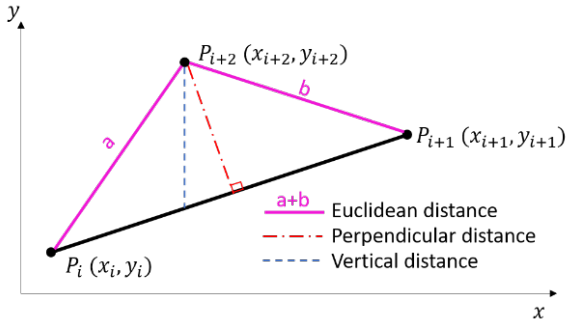


Figure 4: Distance-based data importance evaluation

The PIP selection procedure can be summarized as follows:

1. Given the p-dimensional variable $\boldsymbol{P} = (P_1, P_2, \dots, P_m)$, the first two PIPs are the first and last data point $Q_1 = P_1, Q_2 = P_m$.
2. Q_3 is the data point that has the largest distance (perpendicular distance/vertical distance) to the line joining Q_1 and Q_2. Then the data series is divided into two segments, i.e. $[Q_1, Q_3]$ and $[Q_3, Q_2]$.
3. Q_4 is in either of these two segments, and it is the data point providing the largest distance to its adjacent points.
4. The process continues until all required j data points are collected. The selected points are PIP $= (Q_1, Q_2, \dots, Q_j)$. The number of j is fixed by engineering judgement.

3.2 Entropy-based important-point selection

Information entropy, originally proposed in the field of information theory, is a metric to evaluate the information content in a set of predictions. In EDMF, n initial model instances are sampled by taking the unique values of the unknown model parameters. In the present study, time series responses (each time series has been reduced from m data points to j data points after the first step selection) are calculated based on the model class selected.

For each data point (e.g. data point No. t), there is a discrete variable vector $X_t = [x_{1,t}, x_{2,t}, \dots, x_{n,t}]$ generated by n initial model instances, representing the range of model-instance predictions at data point t. To account for uncertainties, the width of intervals is equal to the sum of measurement and modeling uncertainties, following the EDMF framework (Section 2). The frequency count in each interval corresponds to the number of model predictions that lie within threshold bounds and thus could not be differentiated from the measured value. The width of the interval is a constant at each comparison point. The probability p_i of a model instance falls inside the ith interval is calculated as (m_i/n), where

m_i is the number of model instances that lie within ith interval and n is the total number of initial model instances. Information entropy of data point t is calculated as:

$$H(X_t) = -\sum_{i=1}^{l_t} p_i \, log_2(p_i) \, , i \in \{1, \dots, l_t\} \tag{6}$$

where the range of predictions of data point X_t is divided into l_t intervals with their respective probability $[p_1, p_2, \dots, p_l]$. The information entropy $H(X_t)$ reaches its maximum value when model-instance predictions are uniformly distributed into intervals, and its minimum value when all model instances fall into a single interval, meaning that model-instance predictions of this particular data point could not be differentiated.

When several data points are involved in a set, the redundant information gain between them arises due to the mutual information data points may provide. Papadopoulou et al. [37] introduced the joint-entropy metric to account for the mutual information, in the information entropy of a set of data points. The joint entropy assesses the information entropy between sets of predictions. For a set of two data points (e.g. data points t and t+1), the joint entropy is defined as:

$$H\left(X_{t,t+1}\right) = -\sum_{i=1}^{l_t}\sum_{g=1}^{l_{t+1}} p_{i,g} \log_2\left(p_{i,g}\right), i \in \{1, \dots, l_t\}, g \in \{1, \dots, l_{t+1}\} \tag{7}$$

where $p_{i,g}$ is the joint probability of predictions falling into the ith interval of data point t and the gth interval for data point $t + 1$.

From the information perspective, the joint entropy $H\left(X_{t,t+1}\right)$ follows the union calculation, i.e. the information provided by data point t and t+1 can be calculated by subtracting the mutual information $I\left(X_{t,t+1}\right)$ from the sum of information provided by each data point.

$$H\left(X_{t,t+1}\right) = H(X_t) + H(X_{t+1}) - I\left(X_{t,t+1}\right) \tag{8}$$

Important data points are selected to achieve the maximum joint entropy which corresponds to the maximum expected information gain. High entropy indicates high sensitivity to the parameter inputs and high potential to discriminate between the model predictions using field measurements. Nevertheless, there is no guarantee of direct connection between the high joint-entropy value with the smaller candidate model set. This is because the entropy framework provides only a statistical advantage in terms of optimal data-point selection as the true value of the model parameters is unknown.

The methodology presented in Figure 5 is adopted. Important points are selected from PIP to EIP, one at a time. In the starting state, EIP is a null set and the PIP set contains all the perceptually important data points identified in the first step. First, the entropy of each data point is calculated with their respective prediction distribution. The data point with the highest

information entropy is selected and is added into the EIP set. In the second selection, each time, the data point with the largest information gain is selected from the remaining PIP set into the EIP set. If the difference between the joint entropy $\mathbf{H}^2$ (the superscript is the number of entropy-calculation iterations) and the joint entropy obtained with the previous data-point set $\mathbf{H}^1$ is smaller than a stop criterion α, this means that sufficient data points have been added to the set IP. Otherwise, the data-point selection process is repeated until the criterion is met. The criterion α is set to stop searching for redundant data points when new data points provide little new information and increase the computational time. The value of α is set as $\alpha = 1\% \times \log_2(1/n)$, n is the total number of simulations, where $\log_2(1/n)$ is the maximum entropy that can be achieved.

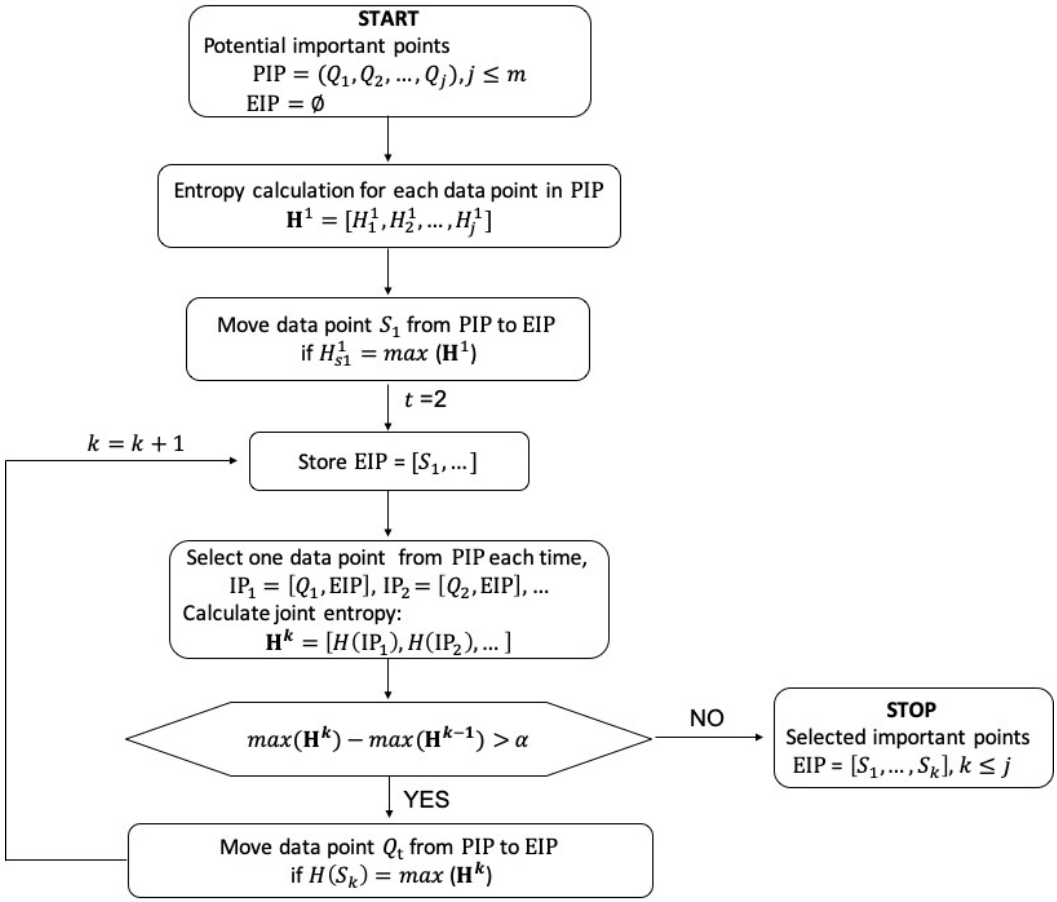


Figure 5: The flowchart of entropy-based important-point selection

After using the proposed important-point-selection approach, the lower dimensional time series $\boldsymbol{S} = (S_1, S_2, \dots, S_k)$ replaces the high-dimensional time series $\boldsymbol{P} = (P_1, P_2, \dots, P_m)$. The new measurement and initial model sets at the selected data points are denoted as $\boldsymbol{y}'$ and $\boldsymbol{\Omega}' = [\boldsymbol{\theta}, \boldsymbol{g}']$. The corresponding modeling and measurement uncertainties at selected data points are denoted as $\mathbf{U}_{\boldsymbol{g}'}$, $\mathbf{U}_{\boldsymbol{y}'}$. Candidate models (CMS) with identified parameter values are found after carrying out EDMF. The general flowchart is shown in Figure 6.

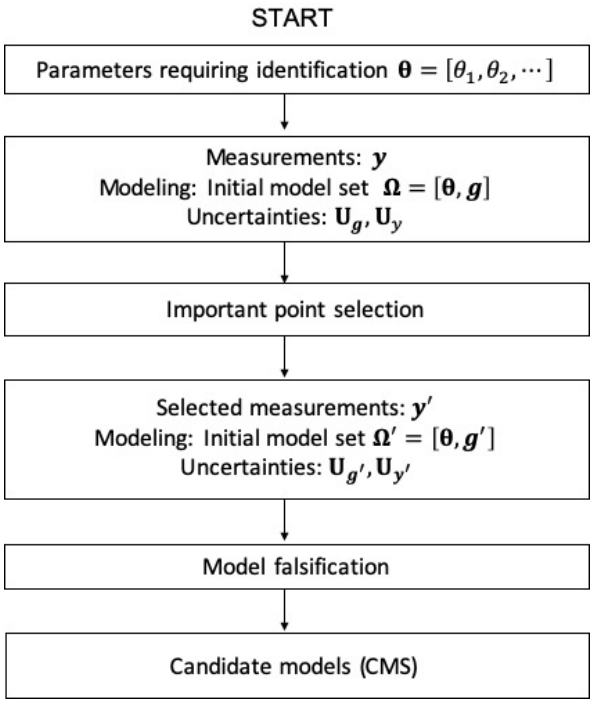


Figure 6: General flowchart of flat-size quantification

4. Case study

4.1 System description and field test

A field test was carried out at a test track in a train depot in Singapore. The test track is a ballasted track with timber sleepers spacing of 0.7 m. This test adopts a monitory system based on rail-pad-force measurement [42]. Zhang et al. [43] have invented a multilayered sensing device (herein referred to as the "rail pad sensor") which converts the rail pad into a load sensor by attaching the multilayered sensing device to the surface of the rail pad. Ten rail pads were replaced with rail pad sensors on five consecutive sleepers (shown in Figure 7 and Figure 8). A train with 12 bogies ran on the track at 50 km/h. The impact force due to the wheel flat takes place within the instrumented length between Sleeper 1 and Sleeper 5. A laser sensor was placed at the side of rail track 1, 0.9 m ahead of rail pad sensor 1 (RPS1) (shown in Figure 7) to record the time corresponding to each wheel passing the laser sensor.

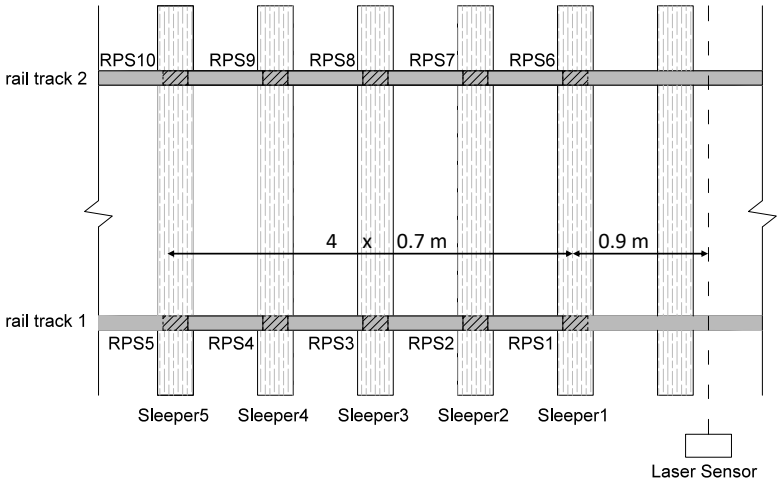


Figure 7: Sensor placement in the test

(a) (b)

Figure 8: (a): Photograph of the tested rail track; (b): Installation of rail pad sensor

4.2 System modeling

Train-track dynamics involves three parts, i.e. railway track, train and wheel-rail interaction (including the effect of wheel flat). This study adopts the Timoshenko-Rayleigh beam model to simulate the rail and standard Kelvin contact-point model (a spring and viscous dashpot in parallel connecting two points) to simulate the rail fastening system (K_p, C_p) and the sleeper support (K_b, C_b). A discretized mass-spring-damper system is employed to model the train. A pitch plane 4-DOF a quarter coach model is used [44]. Interaction between the leading and trailing bogies and their wheelsets in a train body is considered to be negligible. The bogie frame mass (M_b) is connected to the wheel unsprung mass (M_w) through the primary suspension springs and dampers (K_1, C_1). The quarter train body weight is simplified as a constant load

$(M_c g)$ acting on the bogie. A single Nonlinear Hertzian spring (K_{H1}, K_{H2}) is used to model the wheel-rail interaction. Figure 9 shows schematically the train-track interaction model.

A total of 34 sleeper bays are included in the track model to reach a condition that at mid-region of the track model, the reflected waves from the clamped-end boundary are negligible. In this case study, half the bogie and one rail are modelled as the loading is assumed to be evenly distributed between the two rails. The train-track model has already been validated by repeating two well-known experiments [45,46] in previous work [47] where the simulated force using this model agrees well with measured data [47].

The tested train is an empty train carrying no goods and customers. The parameter values of the train system are provided by the train operator. However, the parameter values of the track system vary from site to site. Their values have been calibrated using an independent field test with only good wheels running on the rail [47]. Also, modeling half the bogie is justified in this field test. The impact force generated by the wheel flat at one side of the rail has a negligible effect on the response of rail-pad sensor on the opposite side [47]. Moreover, in the field test, when each bogie passed over, the response of the rail pad sensor immediately went back to a very low level (within environmental noise). This indicates that the effect of other wheels is negligible small [47].

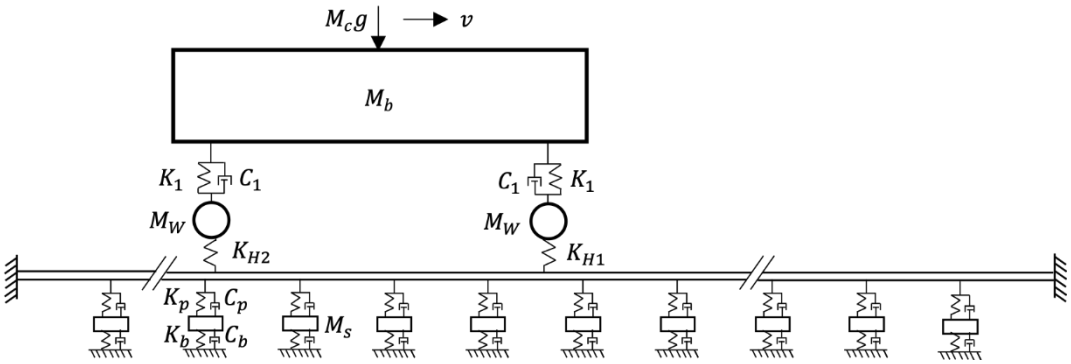


Figure 9: Structural model of the train-track interaction system

This study focuses on the completely worn flat (see Figure 10) which is the most common type of wheel flat [48,49]. The dynamic train-track interaction results in a high-frequency impact load on the crossing nose as the downward motion of the vertical wheel trajectory is reversed and the wheel is accelerated upwards. A mathematical expression of the wheel center trajectory $z_{tr}(x)$ is given in Equation (9). It is a superposition of two harmonic shapes with c as the control parameter of curvature.

$$z_{tr}(x) \approx \frac{d}{2}\left(1 - cos\left(\frac{2\pi x}{L}\right)\right) - \frac{c}{4}\left(1 - cos\left(\frac{4\pi x}{L}\right)\right), 0 \le x \le L \quad (9)$$

where L is the length of wheel flat, x is the longitudinal position of wheel center.

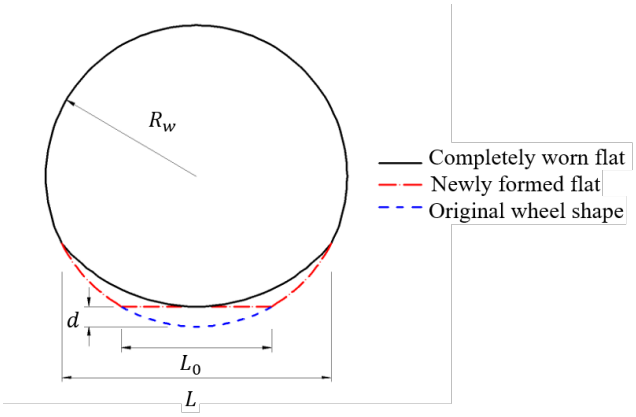


Figure 10: Schematic drawing of wheel flat

Two key parameters to describe the geometry of wheel flat are its length (L) and depth (d). Another parameter c is a second-order term to capture the trajectory curvature. Because z_{tr} is a non-negative value, it can be deduced that $c < d/2$. The impact position X_p is introduced as another model parameter to identify the location where wheel flat hits the rail (and hence which wheel has the flat). The origin of the x-axis is 10.5 m on the left from the position of RPS5. The initial ranges of these four parameters, which are large enough to include all possible sizes of wheel flats based on engineering knowledge, are listed in Table 1.

Table 1: Initial ranges of parameters

	L (mm)	d (mm)	c (mm)	X_p (m)
Initial ranges	[10, 200]	[0.1, 1]	[0, 0.5]	[12, 13]

Sources of modeling and measurement uncertainties are given in Table 2. Uncertainties are described using uniform distributions and are generally expressed as ratios of the predictions/measured values, except for sensor resolution and environment noise which are taken as the absolute values. The rail pad sensors used in this study have been tested both in the laboratory and in field tests [42,47]. The sensor accuracy is taken from those tests. The environmental noise is measured in the field when no train is running on the rail track.

Uncertainties are inherent in modeling, with various assumptions made about the behavior of the track, vehicle and wheel flats. Consequently, the model is only able to represent approximately the real behavior. For parameter values in the train system, a sensitivity study is carried out to quantify their effects on the force estimation. The ranges of the investigated properties are listed in Table 3. 100 samples based on Latin hypercube sampling are generated. Variations of those parameters lead to [-15%, 5%] error of the force estimation.

As mentioned above, the track system has already been calibrated before applied in wheel-flat detection. The parameter values in the track system have been updated using the responses

of rail pad sensor with good wheels running on the rail. After calibration, the simulated responses is [-5%, 5%] error of the measured responses [47].

Mesh uncertainty is estimated based on a mesh refinement study [50]. Uncertainty of finite element models is referenced from previous work [16,51]. Additional uncertainty accounts for all other sources that individually have negligible influence e.g. round off of numbers [17]. All sources of uncertainties are summarized in Table 2.

Table 2: Uncertainty sources for wheel flat detection

Uncertainty sources		**Force**	
		Min	Max
Modeling uncertainties	Train system	-15%	5%
	Finite element model	-8%	5%
	Mesh refinement	-3%	3%
	Additional uncertainties	-1%	1%
	Model calibration accuracy (track system)	-5%	5%
Measurement uncertainties	Sensor accuracy	-5%	5%
	Environment noise	-1 kN	1 kN
	Sensor resolution	-50 N	50 N
	Additional uncertainty	-2%	2%

Table 3: Ranges of parameter values in the train system

Properties	Ranges	Properties	Ranges
Wheel mass (M_w)	[95%, 105%]	Suspension stiffness (K_1)	[-90%, 110%]
Bogie mass (M_b)	[95%, 105%]	Suspension Damping (D_1)	[-90%, 110%]
Train body mass (M_c)	[95%, 105%]		

4.3 Results

In this study, the wheel with flat is on one side of the track which was instrumented with five rail pad sensors RPS1 to RPS5. Unfortunately, RPS3, RPS4 and RPS5 did not work properly due to installation problem. Only the measurement signals of RPS1 and RPS2 are used for wheel flat detection. The flowchart of time series processing is shown in Figure 11. First, according to the relative position of the laser sensor to the rail pad sensors (Figure 7), the signals of RPS1 and RPS2 are extracted. Based on the knowledge of good wheels and wheels with flats

presented in Figure 1, the signals related to possible wheel flats are extracted to test the proposed methodology. Important data points are selected based on the proposed two-step approach (Section 3). Finally, the selected data points will be used to carry out the model-based flat-size quantification following the flowchart in Figure 6.

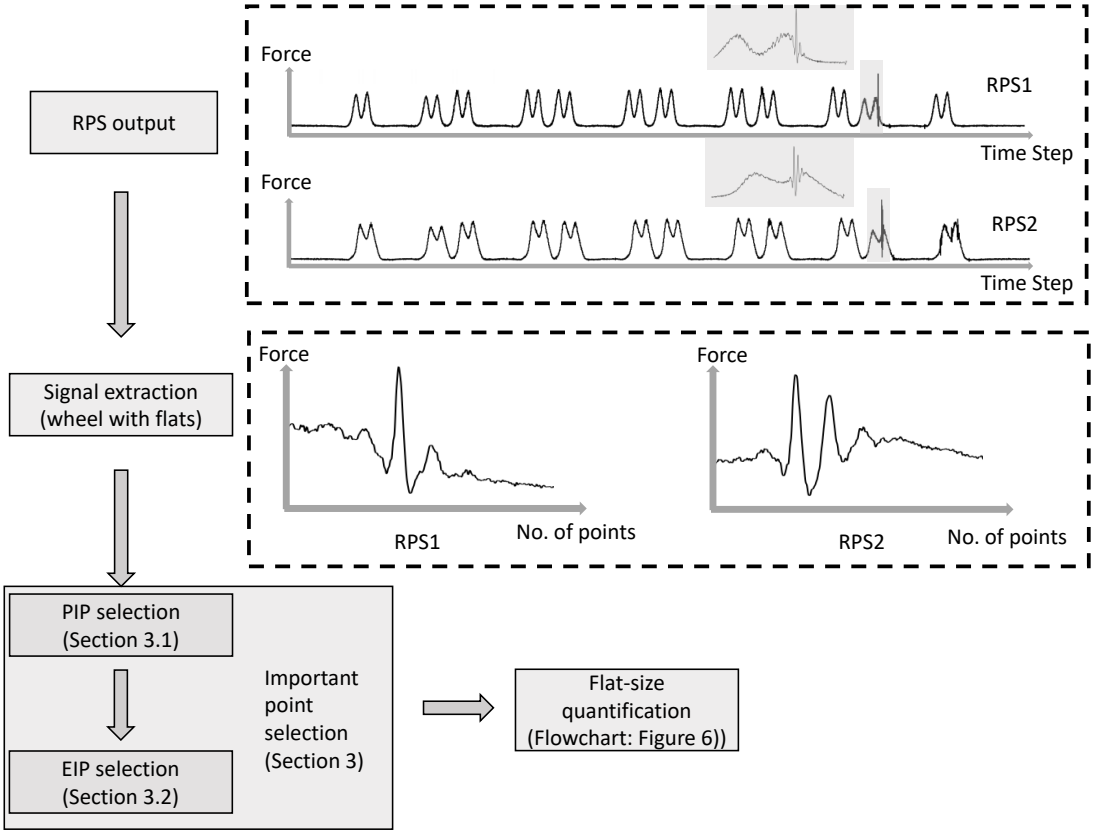
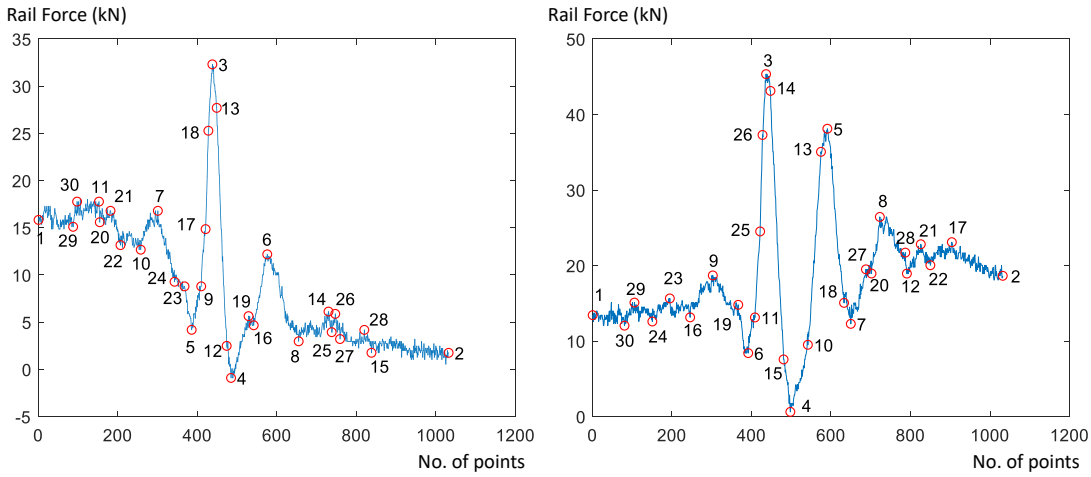


Figure 11: Time series processing for flat-size quantification

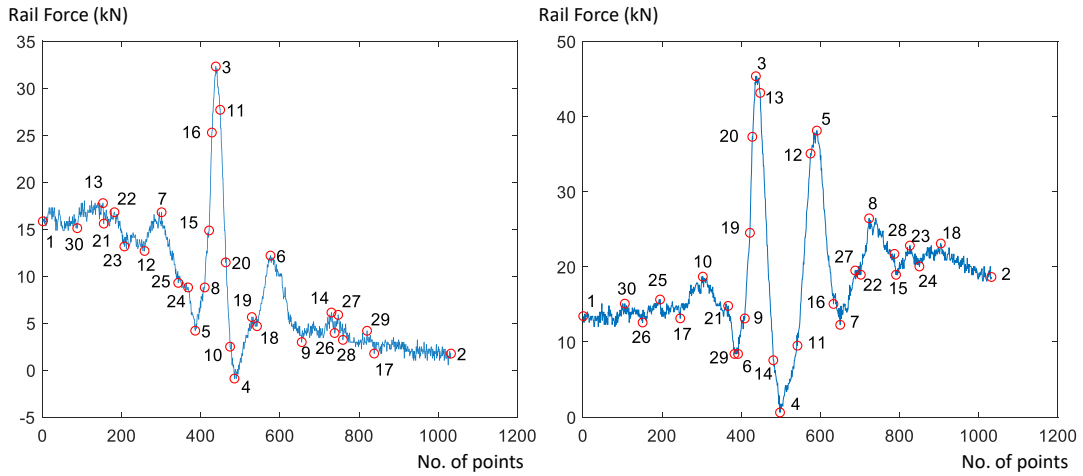
4.3.1 First step

Adopting the PIP selection strategy as explained above, $j = 30$ data points are selected for each time series, hence there are 60 points in total for the two RPS. Both vertical distance and perpendicular distance are compared. The identified PIP and identified order are shown in Figure 12 and Figure 13 which use vertical distance and perpendicular distance, respectively, as the importance measurement. The identified order is consistent with the importance of the data point. The earlier the point is identified, the more important it is.



(a) (b)

Figure 12: Identified PIP using perpendicular distance for (a) RPS1 time series and (b) RPS2 time series



(a) (b)

Figure 13: Identified PIP using vertical distance for (a) RPS1 time series and (b) RPS2 time series

While the identification order is not the same, there is little difference in the points identified in Figure 12 and Figure 13. This observation is consistent with the time series application in finance [32].

4.3.2 Second step

Since there is little difference in the identified points using vertical distance and perpendicular distance, here, the vertical distance is selected to be used in the next step. In this case study, 50,000 samples are generated. For each data point, the information entropy at each PIP is calculated based on the distribution of model-instance predictions. The results are shown in Figure 14. A dozen PIPs have an information-entropy value equal to zero. This is because

the initial condition of the train-track system is set as the same for each simulation and corresponds to first points in the time series that have very low signal-to-noise ratio. During the first seconds, for each simulation, before the wheel flat hits the rail, the rail pad forces are the same. Among all PIPs in Figure 14, data point 58 (which is data point 28 in the second time series) is selected as the first EIP due to the large information entropy.

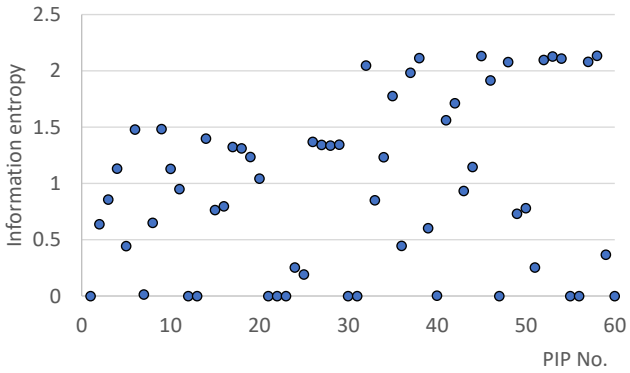


Figure 14: Information entropy at each PIP

Following the entropy-based selection, one data point is selected each time. Figure 15 shows the results of selected points and the order of selection. It is observed that EIP are distributed mostly after the first peak is achieved. This shows that before the wheel flat hits the rail, the responses are similar. Therefore, these data points have a low possibility to discriminate between model-instance predictions. It is after the wheel flat hits the rail track that the response force begins to take effect to distinguish model predictions.

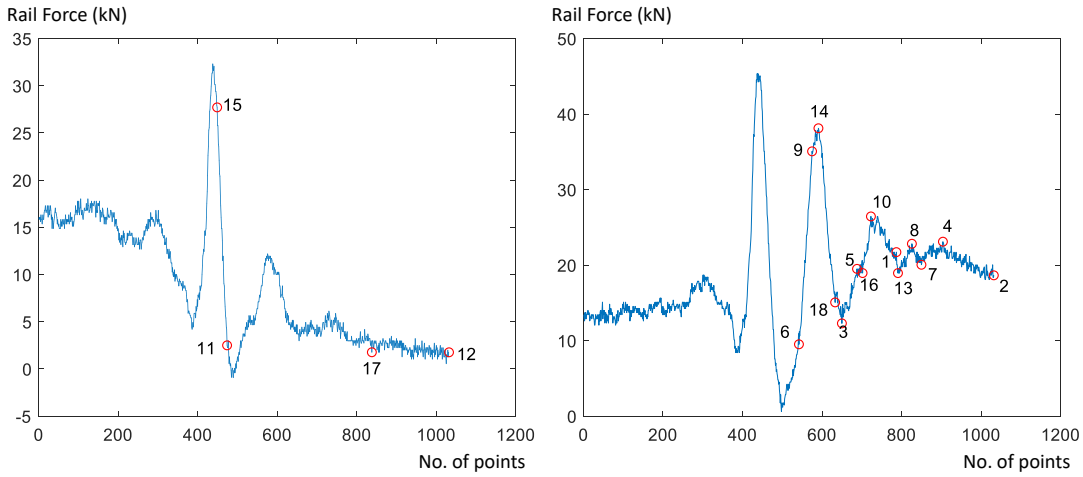


Figure 15: Identified EIP for (a): RPS1 time series and (b): RPS2 time series

In Figure 15, only four points are chosen from RPS1 time series while 14 points are selected from RPS2 time series. To investigate the effects of these two sensors, the entropy-based point-selection process is carried out for (a) RPS1, (b) RPS2 and (c) RPS1+RPS2. In this

process, the joint entropy is calculated until all possible measurement points are selected and results are presented in Figure 16.

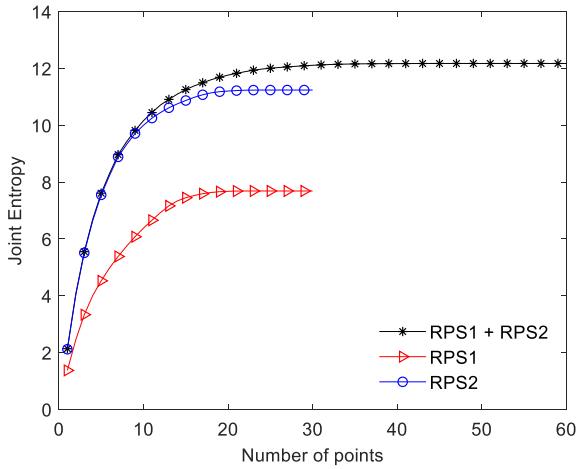


Figure 16: Joint-entropy of the selected PIPs using only RPS1 response, only RPS2 response, and both RPS1 response and RPS2 response

The joint-entropy of two sensors is larger than the value obtained when only one sensor is involved. This result shows that each sensor provides unique information. In Figure 16, the maximum joint entropy values are 7.68, 11.24, 12.17 when using RPS1, RPS2, and both respectively. Consistent with the selected EIP in Figure 15, Figure 16 shows that the information provided by RPS1 is less than the information provided by RPS2.

4.3.3 Identification results

Using the selected EIP, identification using the model-falsification method is carried out. The initial model set includes 50,000 model instances generated by Hammersley sampling [52]. Modeling and measurement uncertainties are shown in Table 2. After using 18 selected EIP, 49,985 models are falsified leading to 15 candidate models (0.04% of all model instances). The results are shown in Figure 17. The first four vertical axes represent the model-parameter values. Each grey line represents a candidate model instance. To validate the parameter identification result presented in this study, L_0 was measured on site with a measuring tape. Because it was not a newly formed flat, there was no clear boundary of L_0 in the measurement. The measured L_0 is estimated to be in the range of [30 mm, 60 mm]. Using the identified values of d, L_0 is calculated by $L_0 = \sqrt{8dR_w}$, where the wheel radius R_w=400mm in this case study. The identified L_0 is in the range of [41 mm, 54 mm]. This value range is consistent with the measurement, as shown in the last vertical axis in Figure 17.

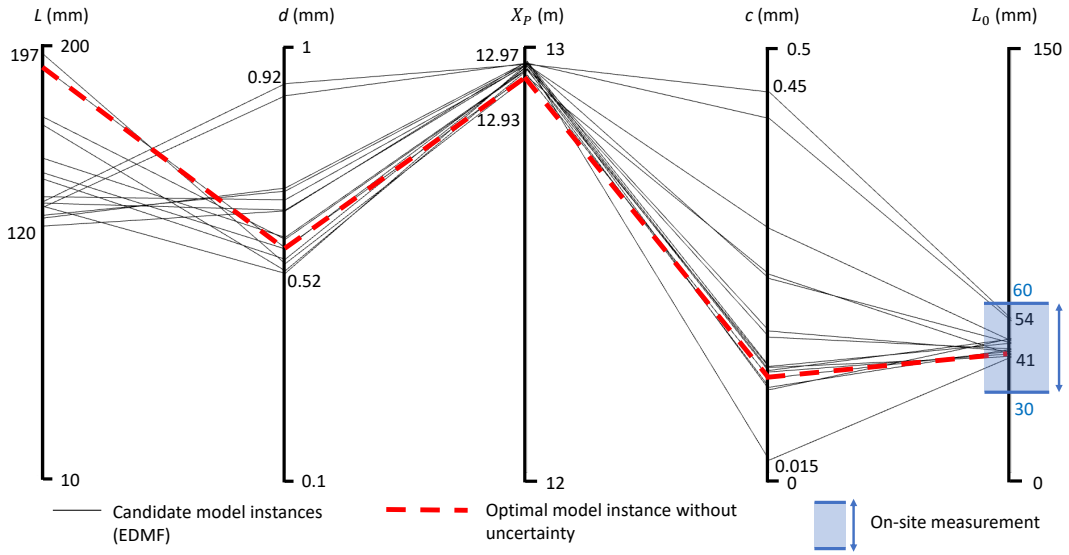


Figure 17: Identified parameter results

According to Railway Group Standard GM/RT2466 [53], for passenger vehicles (wheel radius R_w> 330 mm, speed v ≤ 125 mile/h), the wheelset should be (and was) removed from service within 24 hours of the fault being found.

4.4 Comparison with no uncertainty considered

In this section, no uncertainty is considered for comparison purpose. Among 50,000 model instances, the optimal model instance is obtained when the objective function is to minimize the residual between the model prediction and the in-situ measurement using: $\boldsymbol{\theta}^* = \min_{\boldsymbol{\theta}} \sqrt{\sum_1^k (g_i(\boldsymbol{\theta}) - y_i)^2}$. The optimal parameter values are highlighted via a dashed line in Figure 17. The obtained value of L_0 (43 mm) is near the lower bound of the values obtained considering uncertainties. This result shows that, without considering uncertainties, the wheel-flat size may be underestimated, and it thus shows the importance of considering uncertainties.

4.5 Comparison with random data point selection

As previously mentioned, the entropy framework provides a statistical advantage in terms of optimal data-point selection. Using EIP, 15 candidate models are obtained. To investigate the performance of the entropy framework, the number of candidate models is compared with the ones in which random data points are selected as comparison data points. First, 2,000 cases of 18 data points out of 60 PIPs are generated randomly. Then the candidate models are calculated based on EDMF. The distribution of the number of candidate models (CMS) is shown in Figure 18. The mean value of numbers of CMS using random data points is 180. As

mentioned in Section 4.3.3, using the proposed point selection approach, the number of CMS is only 15. This result shows that EIP statistically outperforms random selection for the same number of data points.

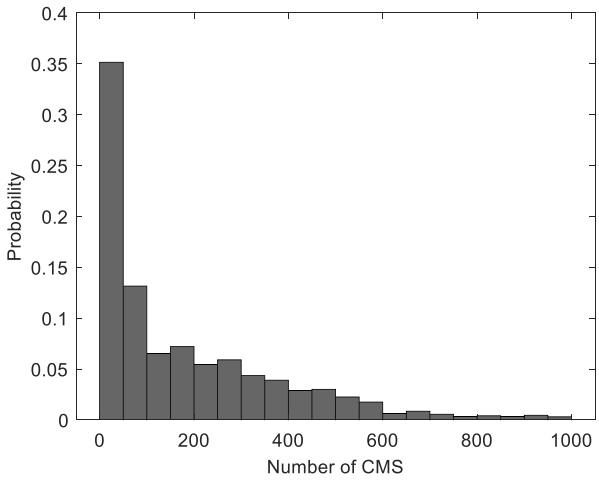


Figure 18: The distribution of CMS numbers using random comparison points.

4.6 Comparison with equidistant data-point selection

As demonstrated in Section 4.3.2, 18 data points are selected based on the proposed approach. In this section, for comparison purposes, we also choose 18 data points from the responses of RPS1 and RPS2 in total, where one point is selected every 128 data points. The modeling uncertainties and measurement uncertainties adopted are the same as in Table 2. After carrying out EDMF, the identified range of L_0 is found to be [26 mm, 56 mm]. This range is twice that of the one obtained using the proposed approach in Section 4.3.3. Thus, the proposed two-step approach of important-point selection is able to narrow down the range of wheel flat length more efficiently than equidistant data-point selection.

5. Conclusions

This study focuses on the identification of wheel-flat size using dynamic signals of rail pad sensors. Both modeling and measurement uncertainties are explicitly included by adapting the model-falsification framework to this task. The paper also presents a two-step important-point selection method to accommodate high-dimensional time series.

The specific conclusions are:

- To accurately assess the wheel-flat size, it is essential to include both modeling and measurement uncertainties. As shown in the case study, without including them, the wheel-flat size is potentially underestimated. This may increase the risk of not taking early action to rectify the bad wheel.

- The proposed two-step approach of important-point selection has been applied successfully to deal with the time series data. In the case study, it helps reduce the number of comparison points from 2,046 to only 18, thereby significantly reducing computational time. It is also verified that the proposed approach statistically outperforms the random data selection approach and

Beyond the rail-pad-force-based monitoring system in the case study, the proposed approach is applicable to other monitoring systems that compare measured time series data with model simulations.

Acknowledgements

This research was conducted at the Future Cities Laboratory at the Singapore-ETH Center (SEC). The SEC was established as a collaboration between ETH Zurich and National Research Foundation (NRF) Singapore (FI 370074016) under the auspices of the NRF's Campus for Research Excellence and Technological Enterprise (CREATE) programme.

References

1. Barke D. Structural Health Monitoring in the Railway Industry: A Review. *Struct Heal Monit* 2005; 4: 81–93.
2. Lewis R, Olofsson U. *Wheel-Rail Interface Handbook*. Elsevier Ltd., 2009. Epub ahead of print 2009. DOI: 10.1533/9781845696788.
3. Wei C, Xin Q, Chung WH, et al. Real-Time Train Wheel Condition Monitoring by Fiber Bragg Grating Sensors. *International Journal of Distributed Sensor Networks* 2011; 8: 409048.
4. Filograno ML, Corredera P, Rodriguez-Plaza M, et al. Wheel Flat Detection in High-Speed Railway Systems Using Fiber Bragg Gratings . *IEEE Sensors Journal* 2013; 13: 4808–4816.
5. Skarlatos D, Karakasis K, Trochidis A. Railway wheel fault diagnosis using a fuzzy-logic method. *Applied Acoustics* 2004; 65: 951–966.
6. Jia S, Dhanasekar M. Detection of Rail Wheel Flats using Wavelet Approaches. *Structural Health Monitoring* 2007; 6: 121–131.
7. Bracciali A, Lionetti G, Pieralli M. Effective wheel flats detection through a simple device. In: *proceedings of the 2002 techrail workshop, Paris*. 2002, pp. 513–521.

8. Woschitz H. Development of a rail-strain-pad using FBG sensors. In: *SHMII-5 2011 - 5th International Conference on Structural Health Monitoring of Intelligent Infrastructure.* 2011, pp. 1–10.

9. Remennikov AM, Kaewunruen S. A review of loading conditions for railway track structures due to train and track vertical interaction. *Struct Control Heal Monit* 2008; 15: 207–234.

10. Hiatt M, Mathiasson A, Okwori J, et al. Finite element model updating of a PSC box girder bridge using ambient vibration test. *Adv Mater Res* 2010; 168–170: 2263–2270.

11. Okasha NM, Frangopol DM, Orcesi AD. Automated finite element updating using strain data for the lifetime reliability assessment of bridges. *Reliab Eng Syst Saf* 2012; 99: 139–150.

12. Beck JL, Katafygiotis LS. Updating Models and Their Uncertainties. I: Bayesian Statistical Framework. *J Eng Mech* 1998; 124: 455–461.

13. Yuen KV. *Bayesian Methods for Structural Dynamics and Civil Engineering*. 2010. Epub ahead of print 2010. DOI: 10.1002/9780470824566.

14. Wan HP, Ren WX. Stochastic model updating utilizing Bayesian approach and Gaussian process model. *Mech Syst Signal Process*. Epub ahead of print 2016. DOI: 10.1016/j.ymssp.2015.08.011.

15. Goulet J-A, Smith IFC. Structural identification with systematic errors and unknown uncertainty dependencies. *Comput Struct* 2013; 128: 251–258.

16. Cao W-J, Koh CG, Smith IFC. Enhancing static-load-test identification of bridges using dynamic data. *Eng Struct* 2019; 186: 410–420.

17. Goulet J-A, Kripakaran P, Smith IFC. Multimodel structural performance monitoring. *J Struct Eng* 2010; 136: 1309–1318.

18. Pasquier R, Goulet J-A, Acevedo C, et al. Improving fatigue evaluations of structures using in-service behavior measurement data. *J Bridg Eng* 2014; 19: 04014045.

19. Vernay DG, Raphael B, Smith IFC. Improving simulation predictions of wind around buildings using measurements through system identification techniques. *Build Environ* 2015; 94: 620–631.

20. Moser G, Paal SG, Smith IFC. Leak Detection of Water Supply Networks Using Error-Domain Model Falsification. *J Comput Civ Eng* 2018; 32: 04017077.

21. Reuland Y, Lestuzzi P, Smith IFC. A model-based data-interpretation framework for post-earthquake building assessment with scarce measurement data. *Soil Dyn Earthq Eng* 2019; 116: 253–263.

22. Fu TC. A review on time series data mining. *Eng Appl Artif Intell* 2011; 24: 164–181.

591 23. Åström KJ. On the choice of sampling rates in parametric identification of time
592 series. *Information Sciences* 1969; 1: 273–278.
593 24. Keogh EJ, Pazzani MJ. A simple dimensionality reduction technique for fast
594 similarity search in large time series databases . *Lecture Notes in Computer Science*
595 *(including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in*
596 *Bioinformatics)* 2000; 1805: 122–133.
597 25. Chakrabarti K, Keogh E, Mehrotra S, et al. Locally adaptive dimensionality
598 reduction for indexing large time series databases . *ACM Transactions on Database*
599 *Systems (TODS)* 2002; 27: 188–228.
600 26. Chu KKW, Wong MH. Fast Time-Series Searching with Scaling and Shifting.
601 *Proc Eighteenth ACM SIGACT-SIGMOD-SIGART Symp Princ Database Syst* 1999;
602 237–248.
603 27. Popivanov I, Miller RJ. Similarity Search Over Time-Series Data Using
604 Wavelets. *Proc 18th Int Conf Data Eng* 2002; 212–221.
605 28. Korn F, Jagadish H V., Faloutsos C. Efficiently supporting ad hoc queries in
606 large datasets of time sequences. *ACM SIGMOD Rec* 1997; 26: 289–300.
607 29. Chung F, Fu T, Luk RW, et al. Flexible Time Series Pattern Matching Based
608 on Perceptually Important Points. In: *International Joint Conference on Artificaial*
609 *Intelligence (IJCAI) Workshop on Learning from Temporal and Spatial Data*. 2001, pp.
610 1–7.
611 30. Pratt KB, Fink E. Search for patterns in compressed time series. *Int J Image*
612 *Graph* 2002; 2(01): 89–106.
613 31. Douglas DH, Peucker TK. Algorithms for the Reduction of the Number of
614 Points Required to Represent a Digitized Line or its Caricature. In: *Cartographica: the*
615 *international journal for geographic information and geovisualization*. 1973, pp. 112–
616 122.
617 32. Fu T, Chung F, Luk R, et al. Representing financial time series based on data
618 point importance. *Engineering Applications of Artificial Intelligence* 2008; 21: 277–300.
619 33. Papadimitriou C. Pareto optimal sensor locations for structural identification.
620 *Comput Methods Appl Mech Eng* 2005; 194: 1655–1673.
621 34. Robert-Nicoud Y, Raphael B, Smith IFC. Configuration of measurement
622 systems using Shannon's entropy function. *Comput Struct* 2005; 83: 599–612.
623 35. Shannon CE. The mathematical theory of communication. *Bell Syst Tech J*
624 1948; 27: 379–423.
625 36. Bertola NJ, Papadopoulou M, Vernay D, et al. Optimal multi-type sensor
626 placement for structural identification by static-load testing. *Sensors (Switzerland)* 2017;
627 17: 2904.

628 37. Papadopoulou M, Raphael B, Smith IFC, et al. Hierarchical sensor placement
629 using joint entropy and the effect of modeling error. *Entropy* 2014; 16: 5078–5101.
630 38. Robert-Nicoud Y, Raphael B, Burdet O, et al. Model identification of bridges
631 using measurement data. *Comput Civ Infrastruct Eng* 2005; 20: 118–131.
632 39. Vernay DG, Raphael B, Smith IFC. Augmenting simulations of airflow around
633 buildings using field measurements. *Adv Eng Informatics* 2014; 28: 412–424.
634 40. Mandal NK, Dhanasekar M, Sun YQ. Impact forces at dipped rail joints. *Proc*
635 *Inst Mech Eng Part F J Rail Rapid Transit* 2016; 230(1): 271–282.
636 41. Fu T chung, Chung F l., Luk R, et al. Stock time series pattern matching:
637 Template-based vs. rule-based approaches. *Eng Appl Artif Intell* 2007; 20: 347–364.
638 42. Zhang S, Koh CG, Kuang KSC. Proposed rail pad sensor for wheel-rail contact
639 force monitoring. *Smart Mater Struct* 2018; 27: 115041.
640 43. Zhang S, Koh CG, Kuang KSC. *Patent No. 11201806959U. A sensor for load*
641 *measurement*. 2018.
642 44. Nielsen JCO, Igeland A. Vertical dynamic interaction between train and track-
643 influence of wheel and track imperfections. *J Sound Vib* 1995; 187: 825–839.
644 45. Newton SG, Clark RA. Investigation into the dynamic effects on the track of
645 wheelflats on railway vehicles. *J Mech Eng Sci* 1979; 21: 287–297.
646 46. Zhai WM, Cai CB, Wang QC, et al. Dynamic effects of vehicles on tracks in
647 the case of raising train speeds. *Proc Inst Mech Eng Part F J Rail Rapid Transit* 2001;
648 215: 125–135.
649 47. Zhang S. *Train wheel monitoring by rail pad sensor and identification*
650 *algorithms*. Ph.D. Thesis, National University of Singapore, 2016.
651 48. Steenbergen MJMM. The role of the contact geometry in wheel-rail impact
652 due to wheel flats. *Veh Syst Dyn* 2007; 46: 713–737.
653 49. Steenbergen MJMM. The role of the contact geometry in wheel-rail impact
654 due to wheel flats: Part II. *Veh Syst Dyn* 2008; 46: 713–737.
655 50. Martínez-Casas J, Giner-Navarro J, Baeza L, et al. Improved railway wheelset–
656 track interaction model in the high-frequency domain. *J Comput Appl Math* 2017; 309:
657 642–653.
658 51. Pasquier R, Smith IFC. Iterative structural identification framework for
659 evaluation of existing structures. *Eng Struct* 2016; 106: 179–194.
660 52. Hammersley JM. Monte Carlo methods for solving multivariable problems.
661 *Ann N Y Acad Sci* 1960; 86: 844–874.
662 53. Rail-Safety-and-Standards-Board. Railway Group Standard GM/RT2466.
663 2010; 1–56.
664