Proceedings of the Sixth International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 7, Paper 4.13
Civil-Comp Press, Edinburgh, United Kingdom, 2024
ISSN: 2753-3239, doi: 10.4203/ccc.7.4.13

Study of Load Spectrum Distribution by Non-Parametric Fitting Method of High-Speed Pantograph

H. Wei, N. Zhou and X. Zhi

**State Key Laboratory of Rail Transit Vehicle System, Southwest Jiaotong University
Chengdu, China**

Abstract

The service status of the pantograph is highly correlated with the operation status of high-speed trains, which affects the current collection quality, operational safety, and reliability of the pantograph catenary system. Currently, most of research of pantograph catenary system is focused on the interaction rather than the structures. Based on existing failure cases of pantograph, this study takes upper arm of pantograph as the object, and takes time-domain load data as the matrix to study the load spectrum distribution fitting method based on non-parameter method. Firstly, a rigid-flexible hybrid pantograph and catenary coupling model is established for dynamics simulation, and the upper arm load data is obtained. Based on the rain flow counting method, the time-domain data is transformed into statistics data with amplitude and mean. Then, based on the diffusion kernel density estimation method, the amplitude distribution is fitted, and the fitting results of the Gaussian kernel density estimation method are compared. Finally, the accuracy of the model is verified by goodness of fit tests, the result indicates that the diffusion kernel density estimation method can well fit the load spectrum of the upper arm which can provide guidance for fatigue reliability design of pantographs.

Keywords: pantograph catenary system, high speed pantograph, upper arm, load spectrum distribution, diffusion kernel density estimation method, non-parametric fitting.

1 Introduction

The electric multiple units, through contact between pantograph and contact wire to achieve current collection, the pantograph catenary system (PCS), as one of the important coupling systems in the dynamic of coupled systems in high-speed railways [1], in the process of the system operation, due to the periodic arrangement of the contact wire, the repeated action of cyclic loads on structural components will inevitably bring about structural fatigue problems. what’s more, the complex serving environment, such as difference in temperature, the humidity, the corrosion and so on, which introduces uncertainties to it [2]. The above will bring huge challenges to the stable operation of the PCS.

According to recent reports in past few years, there have been many cases of pantograph frame failure. From 2013 to 2019, the Shenzhen Metro in China experienced 214 cases of pantograph failures, 60.7% of them are crack failures. In June 2020, there were 12 cases of crack failures in a pantograph in the Guangzhou Metro in China [3]. These failures have a significant impact on the safety of train operations. The load spectrum is an important index to evaluate the fatigue characteristic, so it is necessary to study the load spectrum of pantograph.

The raw data are simulation data on computer or collected data on pantograph commonly. However, their length is always limited, therefore, a mathematical description is needed to extrapolate the load spectrum from limited length to desired length. Therefore, it is necessary to use mathematical methods to fit the distribution of the initial load spectrum, which is an important step in the process of compiling load spectra.

For general mechanical structure, the load spectra usually follow normal distribution, Weibull distribution, or other distribution functions. The fitting shape of this method is often determined by several parameters, so it is also known as parameter estimation method. For example, An et al. [4] studied the reliability modelling of gears under normal distribution random loads, while Su et al. [5] studied the fatigue strength of wind turbine planetary gear systems under Weibull distribution random loads. However, the fitting of a single function often has the best fitting effect for specific load types or even specific load ranges, so there may be limitations in certain application.

Dressler [6] first proposed an extrapolation approach based on non-parametric density estimation. Fink et al. [7] utilized multivariate kernel density estimation method for novelty detection. Chen et al. [8] conducted research on establishing a standardized load spectrum for bogie frames of a high-speed train, considering the distributions, actual operating conditions, and calibrated damage theory based on kernel density estimation method. However, reference [9] indicate that Gaussian kernel density estimation (GKDE) suffered from boundary bias. In 2010, Botev et al. [10] proposed diffusion-based kernel density estimation (DKDE) to address the boundary problem of GKDE.

This article is the first to apply the GKDE method to fit the load spectrum distribution of pantograph, for the preparation of the upper arm load spectrum. Firstly, a rigid flexible hybrid pantograph and catenary coupling model was established to obtain upper arm load data. Then, based on the rain flow counting method, the time-domain load of the upper arm was subjected to cyclic statistics to obtain the corresponding amplitude, mean, and number of cycles. Finally, based on DKDE methods, the distribution of the load spectrum of the load data is fitted, and several goodness evaluation methods were used to describe its fitting accuracy, the results were compared with the GKDE method.

2 Methods

In order to obtain the load spectrum distribution of the upper arm, dynamic simulation was used to obtain the time-domain load, due to the inability of the widely used multi rigid body model and lumped mass model to obtain the spatial load at the hinge point of the upper arm during operation, based on the modal reduction method, the upper arm and collector head guidance are flexible, while the remaining parts still use rigid bodies. A rigid-flexible hybrid pantograph model was established, combined with the catenary model established based on finite element method and the Hertz contact algorithm, to establish a PCS coupling dynamic model. The modelling process is shown in the Figure 1.

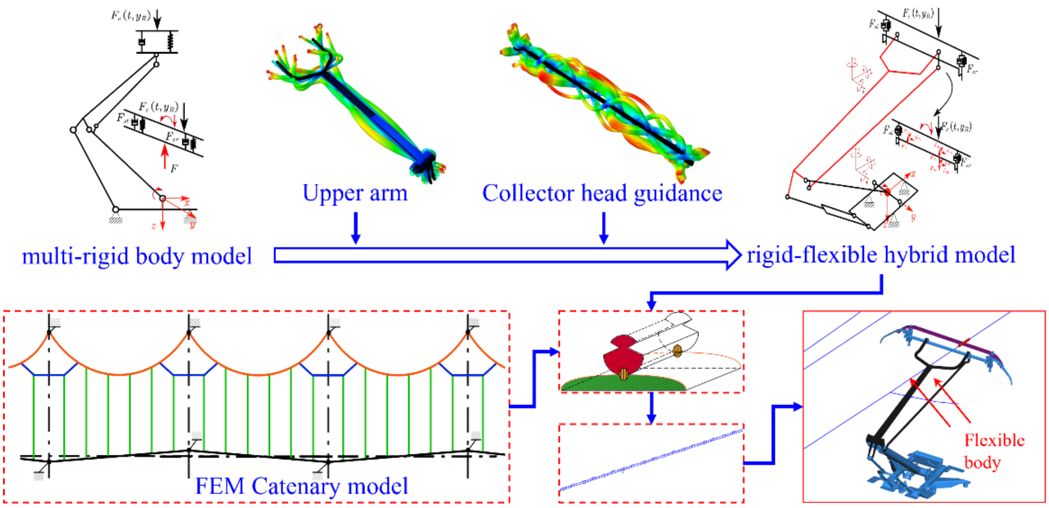

Figure 1: Modelling process of PCS model

The data obtained from the dynamic simulation is load time domain data, which is not statistically significant and cannot be analysed for distribution. Therefore, by using a reasonable load statistical method, the time domain sequence can be transformed into a load spectrum that describes both load level and number of cycles. The rain flow counting method can well reflect the cyclic hysteresis behaviour of material stress-strain [11], and is a relatively mainstream statistical method, as shown in Figure 2.

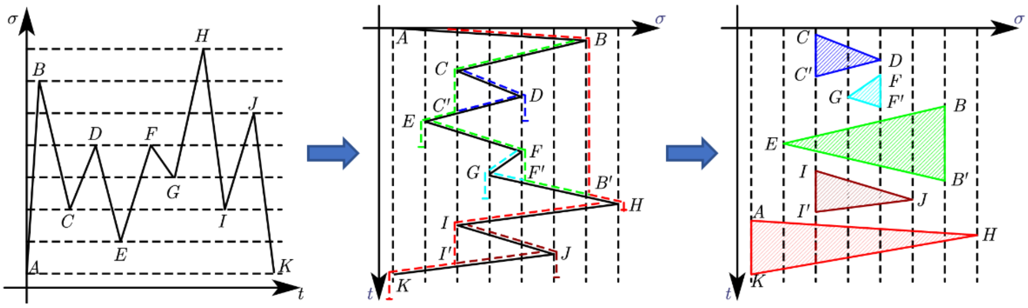

Figure 2: Rain flow counting method

Parametric estimation methods rely on the inherent distribution of the samples, but the non-parametric not. GKDE employs kernel functions as base function, and based on the inherent distribution characteristics of the data, it automatically allocates the base functions to the horizontal axis. The vertical axis values of the allocated kernel functions are linearly superimposed to obtain the final probability density curve.

Let X_1, X_2, ..., X_n be a set of sample points from the random variable X, with the probability distribution denoted as f (x). The probability density function obtained through GKDE can be expressed as Equation (1),

$$\hat{f}(x;h)=\frac{1}{nh}\sum_{i=1}^{n}K(x,X_i,h)=\frac{1}{nh}\sum_{i=1}^{n}K\left(\frac{x-X_i}{h}\right) \tag{1}$$

where, K is the kernel function, which is a symmetric, bounded probability density function with respect to the vertical axis. Therefore, it must satisfy the following conditions, as Equation (2) shows,

$$\begin{cases} K(-x)=K(x) \\ \sup\left|K(x)\right|<\infty \\ \int_{-\infty}^{\infty}K(x)\,dx=1 \end{cases} \tag{2}$$

where, n is the sample size, while h is the bandwidth coefficient, which is obtained using the plug-in method. According to the rule of thumb, it is assumed that $f(x)$ belongs to a family of normal distribution functions characterized by a standard deviation σ, therefore, the bandwidth h can be expressed as Equation (3) [12],

$$\hat{h}=1.06\sigma n^{-\frac{1}{5}} \tag{3}$$

In regards to kernel density estimation, the distribution model faces issues with inaccurate bandwidth calculations and boundary errors. Therefore, based on GKDE, Botev et al. [10] proposed the DKDE, which can be expressed as Equation (4):

$$\hat{f}(x;t)\approx a_k\sum_{k=0}^{N-1}\exp\left(-\frac{k^2\pi^2 t}{2}\right)\cos\left(\frac{k\pi(2j+1)}{2N}\right) \tag{4}$$

where, N is the number of sample segments; a_k is the simplified expression of the coefficients, which can be expressed as Equation (5),

$$a_k = \begin{cases} 1, k = 0 \\ 2\sum_{j=0}^{N-1} P(j)\cos\left[\frac{\left[(2j+1)k\pi\right]}{2N}\right], k \neq 0 \end{cases} \tag{5}$$

The domain R is divided into N segments, so x can be represented in the [0,1] range, as Equation (1) shows,

$$x = \frac{2j+1}{2N},\ j = 1, 2, \ldots,\ N-1 \tag{6}$$

The bandwidth selection for the DKDE method adopts the improved plug-in method, which eliminates the need to assume that f (x) satisfies the normal distribution function family for calculating the standard deviation σ, as required by the GKDE method. This approach avoids errors caused by model selection [10], as Equation (7) shows,

$$t = \xi\gamma^{[l]}(t) \tag{7}$$

Using the improved plug-in method to calculate the bandwidth $\sqrt{t}$, the corresponding actual domain [LL, LU] bandwidth can be expressed as Equation (8),

$$h_{DKDE} = (L_U - L_L)\sqrt{t} \tag{8}$$

where, the lower and upper limits of the actual domain [L_L, L_U] are defined as Equation (9) and (10),

$$L_L = \begin{cases} \min(L_i) & , \forall L_i \geq 0 \\ 5\min(L_i) & , \exists L_i < 0 \end{cases} \tag{9}$$

$$L_U = \begin{cases} \max(L_i) & , \forall L_i \geq 0 \\ 5\max(L_i) & , \exists L_i < 0 \end{cases} \tag{10}$$

The actual probability density can be represented as Equation (11),

$$\hat{f}_L(X_j) = \frac{\hat{f}((2j+1)/2N)}{L_U - L_L},\ j = 0,\ 1, \ldots,\ N-1 \tag{11}$$

3 Results

Based on the above method, a 20 spans catenary was established, and the hinge load on the upper arm of the pantograph was obtained as shown in the Figure 3. Analysis of the data shows that there are significant differences in the hinge loads on left and right sides, including the values, directions, and waveforms.

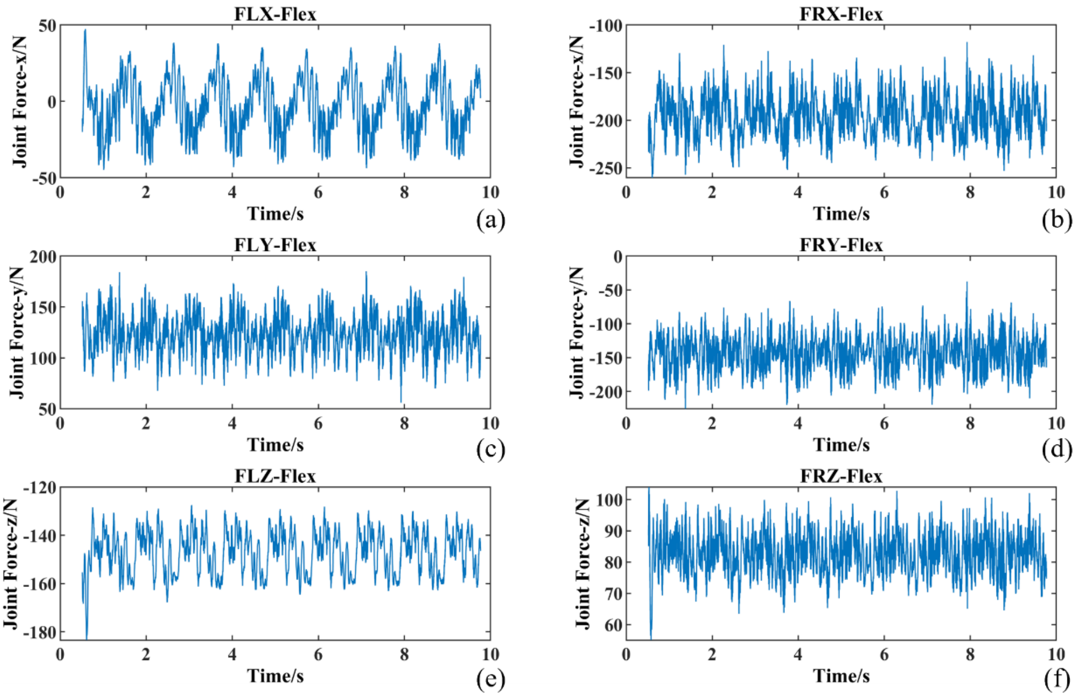

Figure 3: Load data of upper arm

Taking the vertical load on right side (FRZ) of the upper arm as an example, the rain flow counting method was used to obtain the amplitude, mean, and the number of cycles. The statistical result is shown in the Figure 4.

After the dynamic simulation calculation and rain flow statistical processing, the basic data for load spectrum compilation has been obtained. Therefore, based on the GKDE and DKDE methods described in Section 2, spectral distribution fitting was performed on the load amplitude of rain flow statistics. The results are shown in Figure 5 and 6.

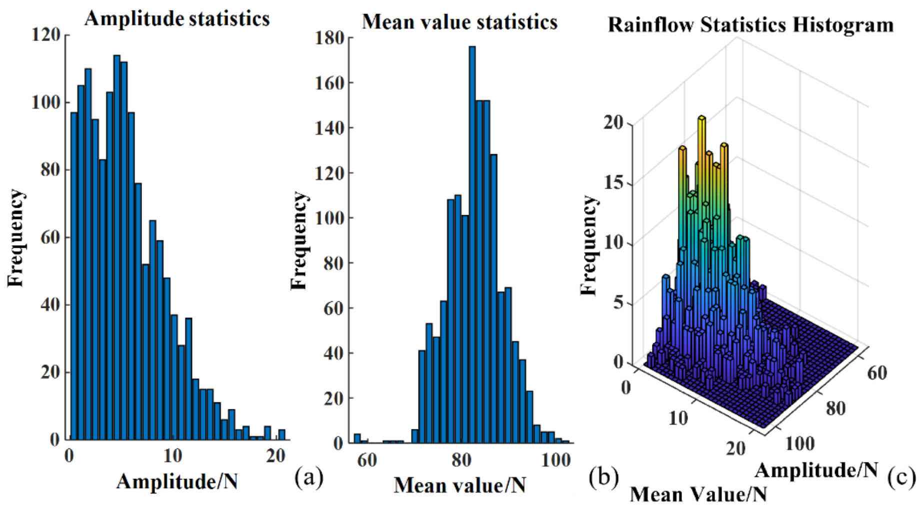

Figure 4: The statistical result of FRZ

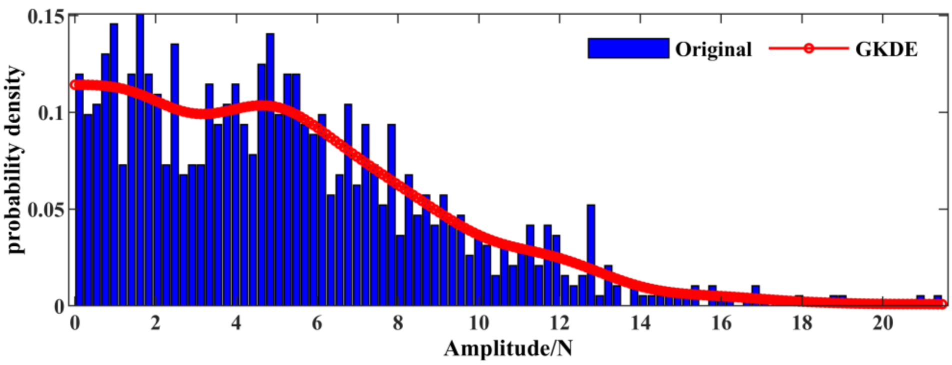

Figure 5. Spectral distribution fitting by GKDE.

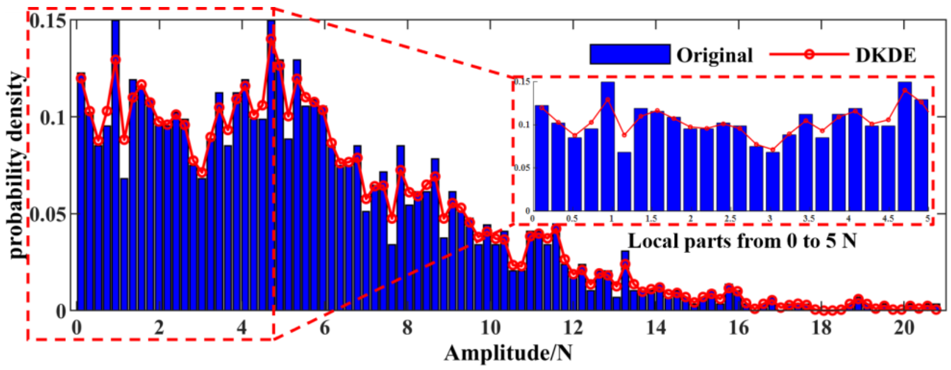

Figure 6: Spectral distribution fitting by DKDE.

Observing the overall and local parts in Figure 6, the DKDE method significantly improves boundary errors and adaptability to subtle features compared to the GKDE's. To provide a comprehensive and accurate depiction of the goodness-of-fit for fitting methods, multiple indicators were employed to assess the fitting results. Four indicators such as the chi-square value (χ^2), the Root Mean Square Error (RMSE), Kolmogorov-Smirnov test statistic (KS), and R-squared value (R2) were introduced [46,47]. The test results are presented in Table 1.

Fitting method	RMSE	KS	R^2	χ^2	$\chi^2{}_c$
GKDE	0.0122	0.0147	0.9084	76.93	120.99
DKDE	0.0049	0.0020	0.9863	102.03	

Table 1: Goodness of fit test result

Based on the comparison of test results for various methods, in the parameter methods, the DKDE has the smaller RMSE value, KS value, while the R^2 value is closer to 1, and both DKDE and GKDE pass the chi-square test. However, except for the χ^2 value, the remaining indicators show that the performs of DKDE better than that of GKDE.

4 Conclusions and Contributions

In order to study the fitting method of the load spectrum distribution of the pantograph, this paper establishes rigid-flexible hybrid pantograph and catenary coupling model, and obtains spatial load data of the upper arm hinge position. Due to weak adaptability of parametric method, the DKDE method is used to fit the load spectrum distribution of the pantograph for the first time, and compared with the GKDE method. Through four indicators, goodness-of-fit test is performed, indicating that the DKDE method can effectively process the load spectrum data of the pantograph. It belongs to a completely data-driven method and can provide technical support for the fatigue and life design of the pantograph.

Acknowledgements

This research was supported by the National Natural Science Foundation of China [Grant No.52072319], the independent project of State Key Laboratory of Traction Power [Grant No. 2023TPL-T05] and the National Natural Science Foundation of Sichuan Province [Grant No. 2023NSFSC0396], which are highly appreciated by the authors.

References

[1] W. H. Zhang, D. Zou, M. Y. Tan, et al. Review of pantograph and catenary interaction. Frontiers of Mechanical Engineering, 13(2), 311-322, 2018.

[2] J. Pombo, & J. Ambrósio. Environmental and track perturbations on multiple pantograph interaction with catenaries in high-speed trains. Computers and Structures, 124, 88–101, 2013.

[3] C. Yi, D. Wang, & L. Zhou. A simulation investigation on the influence of pantograph crack defect on graphite contact strip wear. Engineering Failure Analysis, 131, 105889, 2022.

[4] Z. W. An, & D. M. Sun. Reliability modelling of gear under stochastic loads with normal distribution when considering dependence of failure modes. Journal of Lanzhou University of Technology, 43(4): 37-41, 2017.

[5] M. M. Su, A. Mutellip, & G. Y. Sun. Analysis of Planetary Gear Fatigue Strength of Wind Turbine Under Weibull Distribution Random Load. Journal of Mechanical Transmission, 44(9): 101-105, 2020.

[6] K. Dressler, B. Gründer, M. Hack, et al. Extrapolation of rain flow matrices. SAE Technical Paper, 960569: 200, 1996.

[7] O. Fink, E. Zio, & U. Weidmann. Novelty detection by multivariate kernel density estimation and growing neural gas algorithm. Mechanical Systems and Signal Processing, 50-51, 427-436, 2015.

[8] D. Y. Chen, Q. Xiao, M. H. Mou, et al. Study on establishment of standardized load spectrum on bogie frames of high-speed trains. Acta Mechanica Sinica, 35, 812-827, 2019.

[9] T. Berry, & T. Sauer. Density estimation on manifolds with boundary. Computational Statistics & Data Analysis, 107, 1-17, 2017.

[10] Z. I. Botev, J. F. Grotowski & D. P. Kroese. Kernel density estimation via diffusion. The Annals of Statistics, 38(5), 2916-2957, 2010.

[11] J. J. Sun, S. G. Sun, Q. Li, et al. Compilation Method of Test Load Spectrum for Bogie Frame. Journal of the China Railway Society, 43(6): 29-36, 2021.

[12] S. Ma, S. G. Sun, B. J. Wang, et al. Estimating load spectra probability distributions of train bogie frames by the diffusion-based kernel density method. International Journal of Fatigue, 132, 105352, 2020.