Estimation of Wheel/Rail Contact Forces Based on an Inverse Technique

Tao Zhu*, Shoune Xiao, Guangwu Yang, Mingmeng Wang

State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China
* Corresponding author: 232665368@qq.com; hamilyzt@yahoo.cn

Abstract For actual operating conditions, wheel/rail contact forces of high-speed train are very difficult to directly measure. Minimizing the role of driving force between wheel and rail is a key point to ensure railway wheel-rail transport systems in good condition and efficient operation in the long-term. A time-domain inversion method for dynamic loads was proposed. Based on the state space equation, dynamic programming methods and the Bellman principle of optimality, the main theoretical derivation of the inversion mathematical model was given. With a high-speed vehicle system as the research object, accelerations of axle box as input conditions, the vertical and horizontal wheel/rail forces were identified. Inverse results were compared with SIMPACK simulation results which had the same kinetic parameters. The results indicate that the vertical and horizontal wheel/rail forces had the same trend with SIMPACK simulation results. Results from the inverse model were also compared with experiment data. The inverse model has high inverse accuracy, and can be used for real-time monitoring of the running train wheel/rail contact forces.

Keywords inverse technique, wheel/rail contact force, SIMPACK simulation, experiment data

1. Introduction

The estimation of dynamic forces acting on a structure is a problem that has been treated with only partial success. Methods for such estimation include in two categories, direct methods and indirect methods. Direct methods use the placement of force transducers into the load paths at the point of force application. Indirect methods use other sensor types placed at locations on the structure that may not necessarily correspond to the force input locations. Many situations require indirect methods because the forces cannot be measured[1]. For example, the train is subject to a wheel/rail impact load when operating because of rail irregularities and crossing turnouts.

Currently, various methods for inverse identification problem associated with indirect force measurements have been proposed, see for example Ref. [2-4] for an overview. Among them, are two main methods: the frequency domain[5,6]and the time domain method [7,8].

The running stability of a vehicle depends on the wheel/rail interaction. Wheel/rail contact forces play an important role to keep the vehicle stable on straight track and make it able to negotiate through curves smoothly. The possibility of gaining information about wheel–rail contact forces in real time and on-board normal rolling stock vehicles has significant value. But due to the complexity of the inverse identification problem in railway vehicle systems, not much research has been performed in this area. Some papers on this subject focused on impact detection at the contact point[9,10]. Commercially available systems for monitoring wheel-rail health are based on strain measurements at a chosen location on the track, and the track strain will be measured when the train passed [11]. A big disadvantage of the application of such a system is the necessity to locate strain measurement points at many locations on the track, not only time consuming but also expensive[12]. There is a great need to formulate a method that can be based on measurements on the vehicle but not on the track.

This paper presents a new method to identify the time history of input excitation based on the dynamic programming equation. The forces were identified in the time domain by a recursive formula; the response of the structure was reconstructed by using the identified forces for comparison; and the objective function between the identified and measured values were minimized. The dynamic programming technique possesses inherent limitations that cannot be avoided, however, it still effectively solves these problems during the identification process, and greatly reduces the influence from insufficient known qualities and improper boundary conditions, and obtain decent results comparable with the exact forces. The mathematical model is then applied to estimate the wheel-rail vertical load of a high-speed train, and the inversion results are compared with the rolling and vibrating test-bed and the very detailed SIMPACK model simulation results.

2. Basis of load identification theory

The finite element model of an n-DOFs linear elastic time-invariant structure, the dynamic governing equation is given by:

$$M\ddot{X}(t) + C\dot{X}(t) + KX(t) - F(t) = 0 \tag{1}$$

where M , C , and K are the system mass, damping, and stiffness matrices, respectively; $X(t)$ is the displacement vectors of the structure; and F is the vector of the input excitation forces.

Using the state space formulation, Eq. (1) is converted into a set of first order differential equations as follows:

$$\dot{x} = Ax + Bf \tag{2}$$

For the load identification problem, the known responses of the system M , C , and K are used to solve the unknown input vector $f(s)$ which is in discrete form. In order to facilitate the computer solution, these differential equations are then rewritten as discrete equations using the standard exponential matrix representation.

$$x_{i+1} = Cx_i + Df_i \tag{3}$$

$$y_i = Qx_i \tag{4}$$

where, $C = e^{Ah}$ is the exponential matrix, and together with matrix $D = A^{-1}(C - I)$ is the input influence matrix which represents the dynamics of the system and associates with load. Q is a $m \times 2n$ selection matrix related the measurements to the state variables. x_{i+1} denotes the values at the $(i+1)th$ time step of the computations.

The goal is to find the unknown forcing term f that will cause the system described in Eq.(3) to best match the measurements $\hat{y}_i$. The mathematical representation of a best match is to minimize the least squares error between $\hat{y}_i$ and y_i. This is expressed in matrix-vector notation with the inner product of two vectors (.,.). The least error squares are now expressed as:

$$E = \sum_{i=1}^{N} (y_i - \hat{y}_i)^T \lambda_1 (y_i - \hat{y}_i) + (f_i)^T \lambda_2 (f_i) \tag{5}$$

where T is the transpose of a matrix, y_i and $\hat{y}_i$ are the output variables of the system for the identification formula and measurement, respectively. λ_1 and λ_2 are symmetric positive definite matrices that provide the flexibility of weighting the measurement and the forcing terms. The second term is known as the regularization parameter and the method is called the Tikhonov method. The value of λ_2 is very important for the result, fortunately, there exists a method that can be used to estimate the optimum value of λ_2, see the reference [13].

To minimize the least-squares error E in Eq.(6) over the sequence of the forcing vector, the dynamic programming method and Bellman's Principle of Optimality are applied. This leads to defining the minimize value of E for any initial x and the number of stages, n . Thus:

$$F_n(x) = \min_{f_i} E_n(x, f_i) \tag{7}$$

The recurrence formula can be derived by applying the Principle of Optimality:

$$F_{n-1}(x) = \min_{f_{n-1}}[(Qx_{n-1} - \hat{y}_{n-1})^T \lambda_1 (Qx_{n-1} - \hat{y}_{n-1}) + (f_{n-1})^T \lambda_2 (f_{n-1}) + F_n(Cx_{n-1} + Df_{n-1})] \tag{8}$$

This equation represents the classic dynamic programming structure in that the minimizing at any point is determined by selecting the decision f_{n-1} to minimize the immediate cost (first and second terms) and the remaining cost resulting from the decision (the third term). The solution is obtained by starting at the end of the process, $n = N$, and working backward toward $n = 1$. At the end point $n = N$, the minimum is determined from:

$$F_N(x) = \min_{f_N}[(Qx_N - \hat{y}_N)^T \lambda_1 (Qx_N - \hat{y}_N) + (f_N)^T \lambda_2 (f_N)] \tag{9}$$

At this end point the minimum is obtained by choosing $f_N = 0$ which gives:

$$F_N(x) = \min_{f_N}[(Qx_N - \hat{y}_N)^T \lambda_1 (Qx_N - \hat{y}_N)] \tag{10}$$

Eq.(10) can be expanded to:

$$F_N(x) = (x_N, Q^T \lambda_1 Q x_N) - 2(x_N, Q^T \lambda_1 \hat{y}_N) + (\hat{y}_N, \lambda_1 \hat{y}_N) \tag{11}$$

Eq.(11) can be changed as:

$$F_N(x) = (x_N, R_N x_N) + (x_N, S_N) + q_N \tag{12}$$

where $R_N = Q^T \lambda_1 Q$, $S_N = -2Q^T \lambda_1 \hat{y}_N$, $q_N = (\hat{y}_N, \lambda_1 \hat{y}_N)$.

Eq.(12) shows that F_N is quadratic in x_N. It can be proven inductively that all of the F_n are quadratic in x_n, thus for any n we can write:

$$F_n(x) = (x_n, R_n x_n) + (x_n, S_n) + q_n \tag{13}$$

Substituting Eq.(13) into Eq.(8) and minimizing the equation, the optimal forcing term f^*_{n-1}:

$$(2\lambda_2 + 2D^T R_n D) f^*_{n-1} = -D^T S_n - 2D^T R_n C x_{n-1} \tag{14}$$

For simplification the Eq.(14), let：

$$V_n = (2\lambda_2 + 2D^T R_n D)^{-1} \tag{15}$$

$$H_n = 2D^T R_n \tag{16}$$

Eq.(23) can now be written as:

$$f^*_{n-1} = -V_n D^T S_n - V_n H_n C x_{n-1} \tag{17}$$

These are recurrence formulas required to determine the optimal solution of Eq. (6).

Using Pearson product-moment correlation coefficient to measure the relationship between identification results and actual results, usually expressed by γ . The equation can be expressed by:

$$\gamma = \frac{\sum_{i=1}^{n}(F_{Si} - \bar{F}_S)(F_{Ii} - \bar{F}_I)}{\sqrt{\sum_{i=1}^{n}(F_{Si} - \bar{F}_S)^2}\sqrt{\sum_{i=1}^{n}(F_{Ii} - \bar{F})^2}} \tag{18}$$

where, F_{Si} is the SIMPACK simulation value at each time point, F_{Ii} is the identification value at each time point, $\bar{F}_S$ is the standardization variables of the SIMPACK simulation value, $\bar{F}_I$ is the standardization variables of the identification value.

3. Laboratory verification

First, a laboratory test is performed at TPL at Southwest Jiaotong University using the rolling and vibrating test-bed. The car body vertical acceleration, two bogie frames accelerations and four axle boxes accelerations are measured. Unfortunately, because of the limitation of test conditions, we can not measure the vertical and lateral interface forces directly. So we use a set of measured vertical acceleration response as inputs into the inverse vehicle model to identify other components of the vehicle acceleration responses, and compare with the measured results, by this way to verify the inversion model.

Test scenario is shown in Figure 1, the velocity of the rolling and vibrating test-bed is 250 km/h, the form of rail incentive is actual measured line spectrum of Wu-Guang line.

Figure 1. Test scene of the rolling and vibrating test-bed

Using car body, two bogie frames and the first axle box (numbered from left to right) measured vertical acceleration as input into the inverse vehicle model, the fourth axle box acceleration response and the fourth wheel-set vertical force are identified. See figures 2 and 3.

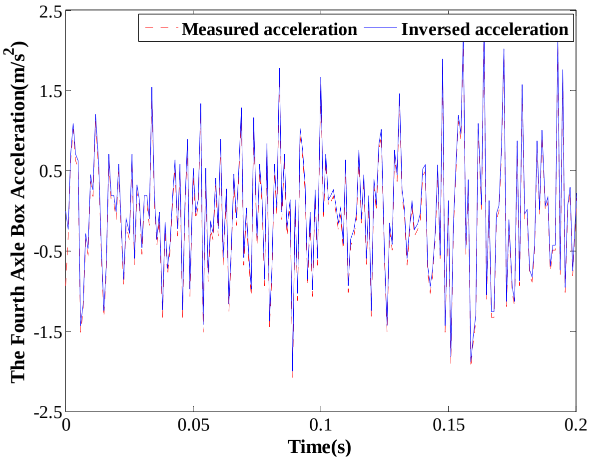

Figure 2. Measured and inversed accelerations for the fourth axle box

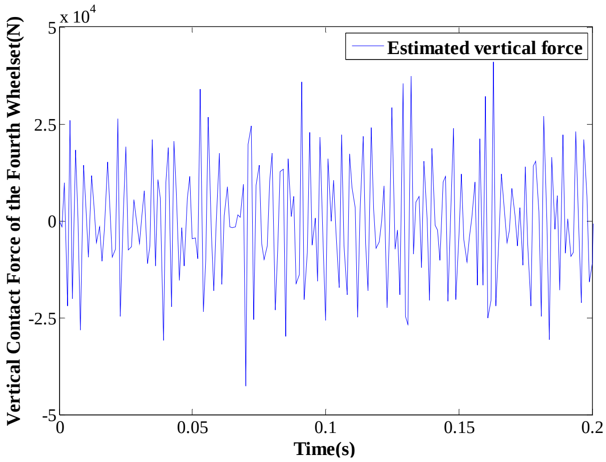

Figure 3. The estimated vertical dynamic contact force for the fourth wheel-set

From figure 2, the acceleration of the fourth axle box which identified by the inverse model is very similar to the measured value, and its correlation coefficient is 0.9756, which can be thought as height correlation. Figure 3 shows the inversed vertical dynamic contact force for the fourth wheel-set, unfortunately, it is unable to be compared with measurement value due to the limitation of test-bed. It is worth noting that, due to the limitation of accelerometers in low frequencies, figure 3 is just the vertical dynamic contact force for the fourth wheel-set, the real vertical force should add the weight reaction force.

4. The application of the inverse model in high-speed train

The commonly used simulation package, SIMPACK, was used to develop a wagon model based on the same parameters as the inverse model. The parameters of the SIMPACK model were the same as the inverse model which was used to generate wheel-rail forces and accelerations at the axle box. These accelerations will be as the input conditions for the inverse mathematical model. In order to make the SIMPACK model replace a practical field test, the vehicle model developed with SIMPACK needs to be a very refined model which includes the nonlinearity of the wheel-rail contact geometry, the nonlinearity of the wheel-rail creep rate and creep forces, the nonlinearity of the vehicle suspension components, and so on.

Taking into account the complexity of the car body systems, as well as many non-linear factors, we need to simplify the body. In this paper, for the vertical and lateral stochastic vibration inverse modeling of the car body, about twenty-seven degrees of freedom are considered.

The measured track irregularity from Beijing to Tianjin was used as the input to the SIMPACK with a simulated velocity of 70 m/s. The resulting axle box accelerations were then used as inputs for the inverse model. The wheel-rail forces were estimated using the MATLAB package. The outputs of the inverse model were the axle box accelerations and wheel-rail reacting forces. The vertical and lateral contact forces of the third wheel-set of the inverse model and the SIMPACK simulation were compared, see Figures 4-5. At the same times, the derailment index which is got from the inverse model and the SIMPACK simulation are also compared. See figure 6.

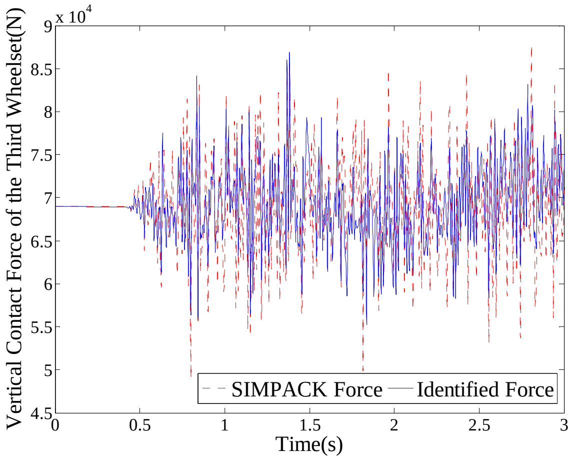

Figure 4. Wheel-rail vertical forces comparison

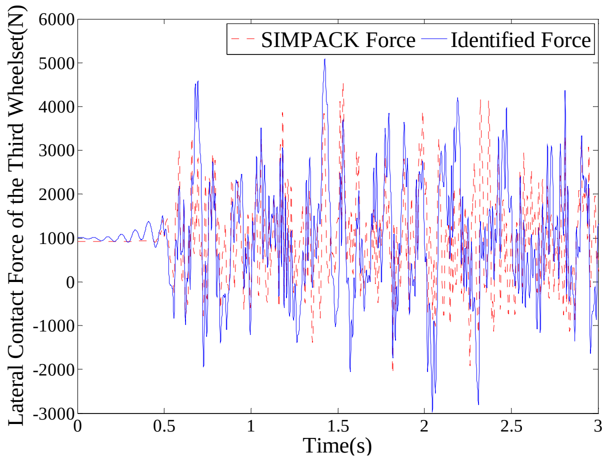

Figure 5. Wheel-rail lateral forces comparison

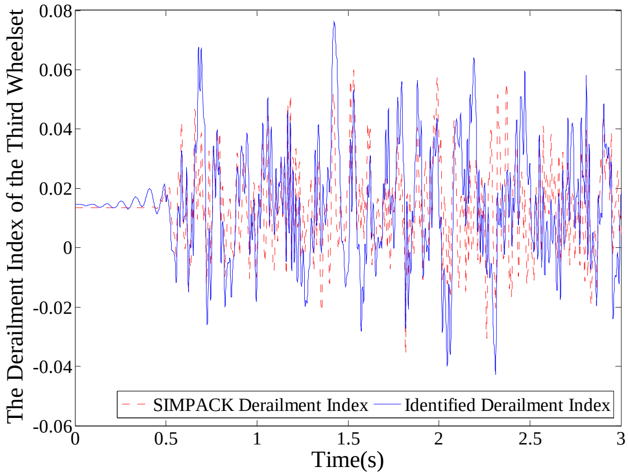

Figure 6. Index coefficient comparison

To compare the results of the SIMPACK simulation and those obtained from the inverse identification method, as shown in Figures 4-5, we can observe that the inverse forces are consistent with the simulation results. Their correlation coefficients are 0.6984 and 0.6235, respectively. By comparing the derailment index, it can be found that the tendencies of the result are also quite accordant.

5. Conclusion

A non-iterative recurrence algorithm for input estimation algorithm mathematical model has been established. Combined with Tikhonovo regularization algorithm, anti-noise ability of the inversion model is enhanced. Based on the response of the accelerations, the method can be applied to the estimation of vertical and lateral contact forces for an operating rail vehicle.

(1)The inversion model is verified by the experiment data of the laboratory test. Using some parts of accelerations which are measured from the rolling and vibrating test-bed to identify the other component of accelerations, and compare with laboratory tests. The results show that the inversion model can be used to identify the unknown output responses for interesting places.

(2) From the time domain, the comparison of the vertical and lateral contact forces results between inverse and SIMPACK models are given. The results show that, the inverse mathematical model has high relatively precision for inversing the wheel/rail contact forces of operation

high-speed vehicle. And their correlation coefficients are greater than 0.5, can be thought as significant correlation.

Since there exist many non-linear factors, such as wheel-rail contact geometry and creep effects, which not only must be taken into account in high-speed train modeling but also make the estimation process of the wheel-rail lateral contact forces are more complex than the vertical contact forces. More research is needed to expand the inverse model which considers the non-linear factors between the wheel and rail. Furthermore, the next step will involve field trial tests to verify the effectiveness of the inverse model.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (51275432); National Key Technology R&D Program in the 11th Five year Plan of china （2009BAG12A04-A11）; University Doctor Academics Particularly Science Research Fund (SWJTU09ZT23).

References

[1] T. Zhu, S. N. Xiao, and G. W. Yang, State-of-the-art development of load identification and its application in study on wheel-rail forces. Journal of the China Railway Society, 33 (2011) 29-36. (Chinese).

[2] K. K. Stevens, Force identification problems-an overview. In: Proceedings of SEM Spring Meeting, Houston, (1987) 838–844.

[3] B. J. Dobson, E. Rider, A review of the indirect calculations of excitation forces from measured structural response data. Proc. Inst. Mech. Engrs.: J. Mech. Engrg. Sci., 204 (1990) 69-75.

[4] L. J. L. Nordstrom, T. P. Nordberg, A critical comparison of time domain load identification methods. In: Proceedings of the Sixth International Conference on Motion and Vibration Control, 2 (2002) 1151-1156.

[5] J. Giergiel, T. Uhl, Identification of impact forces in mechanical systems. Arch. Mach. Des., 36 (1989) 321-336.

[6] J. Giergiel, and T. Uhl, Identification of the input excitation forces in mechanical structures. Arch. Transp., 1 (1989)8-24.

[7] B. J. Dobson, E. Rider, A review of the indirect calculation of excitation forces from measured structural response data. J. Mech. Eng. Sci. , 204 (1990) 69-75.

[8] T. Uhl, J. Pieczara, Identification of operational loading forces for mechanical structures. Arch. Transp. ,16 (2003) 109-126.

[9] S. Lechowicz, C. Hunt, Monitoring and managing wheel condition and loading. In: Proceeding of International Symposium for transportation recorders, Arlington, (1999)205-239.

[10] J. Nielsen, A. Johansson, Out of round railway wheels-literature curve. In: Proceedings Of the Institute of Mechanical Engineers – part F, 214 (2002) 79-91.

[11] A. Chudzikiewicz, Selected elements of the contact problems necessary for investigating the rail vehicle system. In: Kisilowski, J., Knothe, K. (eds.) Advanced railway vehicle system dynamics, WNT, Warszawa, 1991.

[12] A. Chudzikiewicz, Elements of vehicle diagnostics. ITE, Radom, 2002.

[13] P. C. Hansen, Analysis of discrete ill-posed problems by means of the L-curve. SIAM Rev. , 34(1992) 561-580.