

producing delay in each cycle, changes the place of the starting shock wave ahead (point A to point B), making the queue length in the back of the queue larger (point C to point D). For instance, if 2 seconds is determined for this lost time (as the length of the line AB), then more than 2 seconds is needed to discharge the added queue in this period (the line EF or similarly the line LG). By adding the time L_s to the end of the displayed red interval, a wave, BD, is formed and considered instead of AC. This wave extends the maximum queue from point C to point D. The line EF shows the incremental effect of lost time and is defined as the minimum green time that must be provided to clear the queue length due to the start-up lost time. We call the line EF as L_{ec} . Since the lines EF and CG are the same length, we can formulate L_{ec} as below:

$$L_{ec} = \overline{CL} + \overline{LG} = L_s + \left(\frac{L_s}{SW_R} + \frac{L_s}{SW_N} \right) \left(\frac{SW_N \cdot SW_I}{SW_N - SW_I} \right)$$

By simplifying this equation we have:

$$L_{ec} = \frac{q \cdot L_s}{S - q}$$

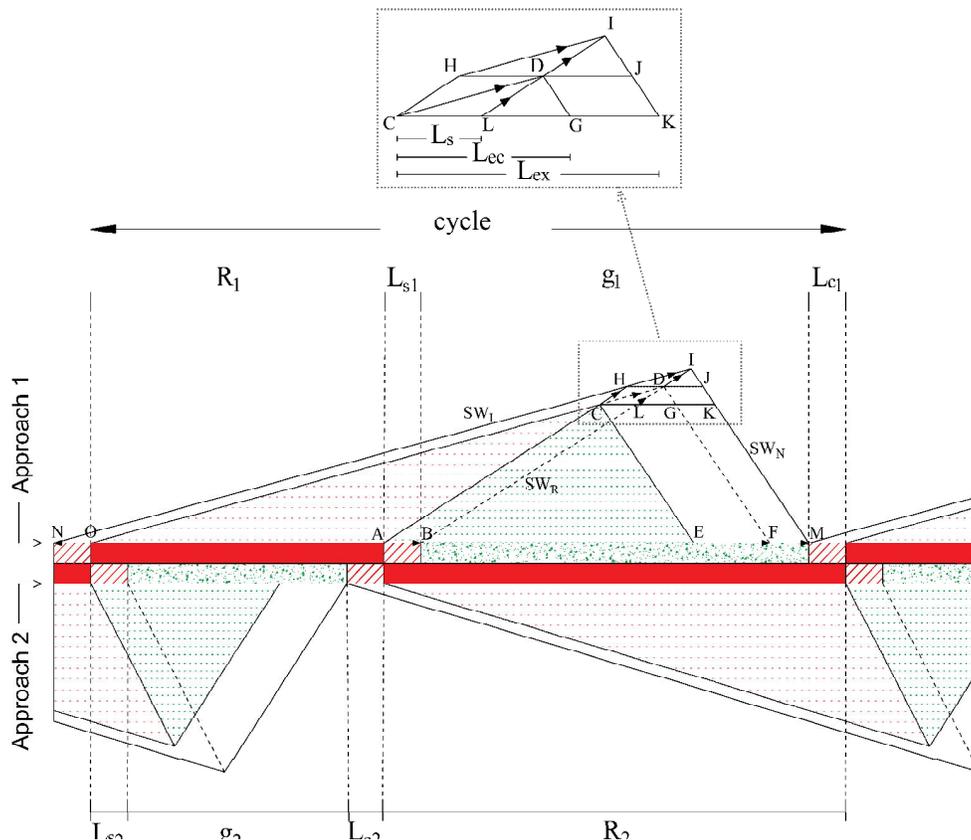


Figure 2. The effect of lost times on cycle intervals in the shock wave analysis for two approaches.

Also, the clearance lost time from the previous cycle (line NO) adds more delay to this effect (line FM), and the total effect is shown by L_{ex} (line EM). According to Figure 2, the total extra delay time because of the lost time in the approach is equal to the line CK (similarly EM in Figure 2). Supposing that start-up and clearance lost times are identical, and that the triangles ΔCDG and ΔHIJ are equal, then the lines LG , DJ , and GK are all the same length. Therefore, the minimum total extra time required to clear the extra queue length due to both the start-up lost time and clearance lost time is given by:

$$L_{ex} = L_{ec} + \overline{LG} = L_s + 2\overline{LG}$$

By simplifying it, we have:

$$L_{ex} = \frac{L_s(S+q)}{S-q}$$

The equation above is used in the constraints of the proposed optimization model.

The Proposed Optimization Model. In signal timing optimization problems, three objective functions are frequent: capacity maximization, cycle time minimization and delay minimization (Wong et al. 2003). In this paper, to find the optimum cycle time, we aim to minimize the shock wave delay function.

To optimize the signal timing of an isolated traffic signal, we consider a signalized intersection with two approaches (a one-lane eastbound, and a one-lane northbound). According to Figure 2, in the shock wave delay function, for the first approach we replace r_1 by $R_1+L_{s1}+L_{c1}$, and for the other approach we replace r_2 by $R_2+L_{s2}+L_{c2}$. In the delay function, we could also replace $t_{m1}+t_{c1}$ by $R_2-(L_{s1}+L_{c1})$ and $t_{m2}+t_{c2}$ by $R_1-(L_{s2}+L_{c2})$. However, since the results are similar, the change was not applied in the delay function.

Then, the two shock wave delay functions below are used in the optimization model for two approaches:

$$D_1 = \frac{X_{m1}}{2} \cdot \left[(R_1 + L_{s1} + L_{c1}) \cdot (k_{j1} - k_{a1}) + (t_{m1} + t_{c1}) \cdot (k_{d1} - k_{a1}) \right]$$

$$D_2 = \frac{X_{m2}}{2} \cdot \left[(R_2 + L_{s2} + L_{c2}) \cdot (k_{j2} - k_{a2}) + (t_{m2} + t_{c2}) \cdot (k_{d2} - k_{a2}) \right]$$

Based on Equation 2, the objective function, which is the average delay for two approaches, is given by:

$$d = \frac{D_1 + D_2}{(R_1 + R_2)(q_1 + q_2)}$$

Where $R_1+R_2=C$.

If we substitute D_1 and D_2 in the equation above and simplify the equation, d is obtained as below:

$$d = \frac{q_1 S_1 (S_2 - q_2) (R_1 + L_{s1} + L_{c1})^2 + q_2 S_2 (S_1 - q_1) (R_2 + L_{s2} + L_{c2})^2}{2(S_1 - q_1)(S_2 - q_2)(R_1 + R_2)(q_1 + q_2)}$$

The significance of this equation is that to calculate d , only the data of arrival and departure flow rates (i.e. q_1 , S_1 , q_2 , and S_2) in both intersection approaches is needed, while L_{s1} , L_{c1} , L_{s2} , and L_{c2} are all initially assumed as a fixed known value, about 2 seconds each. Then, our proposed optimization on the basis of lost time analysis is written as below:

$$\text{Min} \left\{ d = \frac{q_1 S_1 (S_2 - q_2) (R_1 + L_{s1} + L_{c1})^2 + q_2 S_2 (S_1 - q_1) (R_2 + L_{s2} + L_{c2})^2}{2(S_1 - q_1)(S_2 - q_2)(R_1 + R_2)(q_1 + q_2)} \right\}$$

s.t.

$$(1) R_1 \geq L_{ex2} + L_{s2} + L_{c2}$$

$$(2) R_2 \geq L_{ex1} + L_{s1} + L_{c1}$$

$$(3) R_1 \geq t_{m2} + t_{c2} + L_{s2} + L_{c2}$$

$$(4) R_2 \geq t_{m1} + t_{c1} + L_{s1} + L_{c1}$$

$$(5) g_1 \geq L_{ex1}$$

$$(6) g_2 \geq L_{ex2}$$

The defined constraints are obtained from the geometry in Figure 2. In the above setting, the first and second constraints provide the minimum duration for the displayed red times based on the other approach's lost times and lost times effects. Constraints 3 and 4 provide minimum duration based on the required time to form and clear the queue in the other approach, with the purpose of keeping the intersection in the undersaturated condition. The third constraint has been derived from the combination of two constraints below:

$$R_1 \geq g_2 + L_{s2} + L_{c2} \quad \text{and} \quad g_2 \geq t_{m2} + t_{c2}$$

And the similar way for the fourth constraint:

$$R_2 \geq g_1 + L_{s1} + L_{c1} \quad \text{and} \quad g_1 \geq t_{m1} + t_{c1}$$

Finally, in constraints 5 and 6, L_{ex1} and L_{ex2} have been considered as a minimum for the green times in each intersection.

Regarding the strategy of keeping the intersection in the undersaturated condition, it should be noted that the degree of saturation is obtained by the volume-to-capacity (v/c) ratio (see equation below) and is dependent on the signal timing, or more specifically, the green to cycle time ratio.

$$\frac{v}{c} = \frac{q}{S \cdot \frac{g}{C}}$$

In each approach, v/c can be less than 1.0, 1.0 or more than 1.0, where they are called undersaturated, saturated, and oversaturated, respectively. If we aim to keep the intersection in the undersaturated condition, we must optimize the cycle time so the v/c ratio in both approaches will be less than 1.0, and the total ratio less than 2.0. For example, if S , q_1 , and q_2 are supposed 1800 *veh/hr*, 90 *veh/hr*, and 1620 *veh/hr*, and if $g/C=0.5$ is assumed, then the v/c ratios of the approaches will be 0.1 and 1.8, respectively. Hence, with this timing, one approach is undersaturated

and the other is oversaturated. Since the total v/c ratio of the approaches is less than 2.0 (i.e. $0.1+1.8=1.9$), our method is able to find the optimum cycle time. In this case, suppose that the g/C ratios for two approaches are 0.075 and 0.925. Then the v/c ratios will be 0.67 and 0.97. This way, both approaches are in the undersaturated condition. Thus, our method works in the case that the total v/c ratio of both approaches is less than 2.0.

Convexity. An important feature of the convex optimization problem is that the optimal solution is unique (Luenberger et al. 1984). A minimization problem that involves a convex objective function and a convex feasible region is referred to as a convex optimization problem. Our proposed two-variable (R_1 and R_2) optimization model is convex and provides one feasible solution. The first derivative test implies that the objective function is twice differentiable and continuous. The second derivative test of convexity, by calculating the determinant of Hessian for a given set of parameters, shows that the function is positive definite (convexity condition) in the feasible convex domain ($R_1 > 0$ and $R_2 > 0$) as below:

$$Hessian = \begin{bmatrix} \frac{2}{411} \left(\frac{319R_2^2 - 1832R_2 + 5104}{(R_1 + R_2)^3} \right) & \frac{-2}{411} \left(\frac{319R_1R_2 - 916R_1 + 916R_2 - 5104}{(R_1 + R_2)^3} \right) \\ \frac{-2}{411} \left(\frac{319R_1R_2 - 916R_1 + 916R_2 - 5104}{(R_1 + R_2)^3} \right) & \frac{2}{411} \left(\frac{319R_1^2 + 1832R_1 + 5104}{(R_1 + R_2)^3} \right) \end{bmatrix}$$

$$\det(Hessian) = \frac{2560}{137(R_1 + R_2)^4} > 0$$

EVALUATION

To evaluate the proposed method, we consider hypothetical parameters as below:

$$S_1 = S_2 = 1800 \text{ veh / hr} = 0.5 \text{ veh / sec}$$

$$k_{d1} = k_{d2} = 60 \text{ veh / km} = 0.06 \text{ veh / m}$$

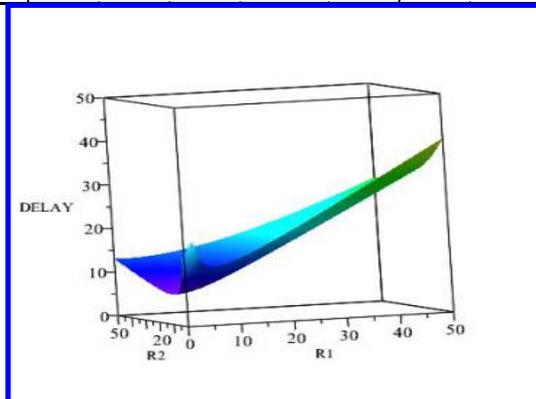
$$k_{j1} = k_{j2} = 120 \text{ veh / km} = 0.12 \text{ veh / m}$$

$$L_{s1} = L_{c1} = L_{s2} = L_{c2} = 2 \text{ sec}$$

To investigate the model for different situations, we use different values of q and k_a for both approaches, where q is ranging from 90 veh/hr (indicating low v/c of 0.1 for $g/C=0.5$) to 1620 veh/hr (indicating high v/c of 1.8 for $g/C=0.5$). The results for 9 combinations using Maple software (version 18.00) are presented in Table 1. In this Table, the results from Webster's method are also provided for evaluation purposes. Since the shock wave delay model has been proven accurate for calculating delay, we used this model to calculate the delay related to signal timing provided by Webster's method. In Table 1, we show displayed green (G) times instead of displayed red (R) times, as is prevalent in signal timing studies. The 3-D graph of the optimization solution for a given case is also presented in Figure 3.

Table 1. Results of the Proposed Signal Timing Method and Webster's Method.

Case	Approach 1		Approach 2		The proposed method					Webster's method				
	q_1 (veh/hr)	k_{a1} (veh/km)	q_2 (veh/hr)	k_{a2} (veh/km)	C_{opt} (sec)	G_{opt1} (sec)	G_{opt2} (sec)	G_{opt1}/C_{opt} (sec)	d (sec)	C_{opt} (sec)	G_{opt1} (sec)	G_{opt2} (sec)	G_{opt1}/C_{opt} (sec)	d (sec)
1	90	2	90	2	12.42	6.21	6.21	0.50	4.42	18.88	9.44	9.44	0.50	5.04
2	90	2	1620	34	160.00	12.00	148.00	0.08	11.58	340.00	21.47	318.53	0.06	17.51
3	180	4	360	8	14.69	6.44	8.25	0.44	4.98	24.29	9.43	14.86	0.39	5.81
4	180	4	1260	26	41.14	8.11	33.03	0.20	7.60	85.00	13.62	71.38	0.16	9.97
5	270	6	720	16	18.03	6.70	11.33	0.37	5.94	37.78	12.12	25.66	0.32	7.90
6	450	10	450	10	16.00	8.00	8.00	0.50	6.00	34.00	17.00	17.00	0.50	8.65
7	450	10	900	20	32.00	12.00	20.00	0.38	9.33	68.00	24.00	44.00	0.35	15.22
8	450	10	1260	26	160.00	44.00	116.00	0.28	33.47	340.00	91.37	248.63	0.27	65.78
9	810	18	900	18	160.00	76.00	84.00	0.48	41.89	340.00	161.26	178.74	0.47	84.55

**Figure 3. 3D Graph of Signal Timing Optimization in a Given Case.**

DISCUSSION

The results given in Table 1 show that our model, in all cases, produces lower optimum cycle time. Generally, green to cycle time ratios are comparable in both methods, with a maximum difference of approximately 12%. Also, in all cases, the average delay of our proposed method is lower than Webster's method, where the constraints are not considered when finding the delay of both methods. When the initial arrival flow rates are low in both approaches (i.e. the total v/c is less than 0.5), there is no significant difference between the results of the two methods (e.g. cases 1 and 3) but when the total v/c is around 2.0 (i.e. 1.9 in cases 2, 8, and 9) the difference is remarkable (an optimum cycle time of 160 seconds vs. 340 seconds). Webster's method is considered inefficient in high v/c ratios (Ahn 2004). This is evident in the delays in cases 8 and 9, which are almost twice the amount of delay in our solution. Furthermore, as recommended by the Canadian capacity guide for signalized intersections, the maximum cycle time used will depend largely on the control method, but should not exceed 160 seconds (Teply et al. 2008). Thus, in high v/c ratios, the results of our model is superior, where our method produces exactly 160 seconds at the highest total v/c ratios. It is prudent to mention that the total ratio of higher than 2.0 means both approaches are in the oversaturated condition. While these conditions (e.g. bottlenecks) cannot be prevented, various control policies are

presented by researchers to manage these situations, which are beyond the scope of this research.

CONCLUSION AND RECOMMENDATIONS

In this paper, we proposed an optimization model to find the optimum cycle and green splits by minimizing the shock wave delay model for a simple isolated signalized intersection as a fixed-time method. To do so, we modelled the lost time compliant with delay calculations in the shock wave model. The incremental effect of lost time in extending the required green times in each cycle affects the extension of the red time in the other approach. This way, cycle time is optimized, based on the minimum required green and red times in both approaches, by establishing an appropriate balance between the two approaches, while providing the minimum delay for individual vehicles in the intersection. In this method, the key strategy is to keep both approaches in the undersaturated condition. Therefore, our model works when the total v/c ratio of both approaches is less than 2.0; otherwise, where both approaches are oversaturated, other control policies should be considered.

The results show that our proposed method outperforms Webster's method, particularly in high v/c ratios. Without considering the lost time and its incremental effect, the comparable model produces very small values for the optimum cycle time, close to zero, which is impractical; the triangles in the model are originally formed based on the red time of each approach, hence the model tends to shrink the area of triangles to provide the minimum possible delay, which occurs when the red times are shrunk to near-zero values in both approaches.

Moreover, without the constraints presented in our paper, the ratio of green to cycle time will be similar to both the proposed method and Webster's method if we input the optimum cycle time as a known value to find the optimum green splits. Nevertheless, to optimize the cycle time and provide reasonable values, the proposed restrictions are required to provide low boundaries to prevent green times from reaching zero. In conclusion, considering the lost time as a constant, defining the relations in cycle time based on the displayed red times, the effective green times, and the lost times, and embedding the lost time in both the delay function and constraints aids the model to generate appropriate, real-world results.

This method, even without further development, can provide insight by supplying an appropriate approximation about optimum signal settings. It can be considered a starting point, while for future work it can be adapted and applied in coordinated intersections, network scale or traffic adaptive systems. Also, this model can be improved by adding queue length control constraints, pedestrian constraints, and number of vehicle stops, using them in a multi-objective optimization problem. To investigate the applicability and strength of the proposed optimization model, different types of traffic signals (e.g. with two or more lanes) in various conditions will be examined via access to real data.

REFERENCES

- Ahn, W. Y. (2004). "Dynamic signal optimisation for isolated road junctions" (Doctoral dissertation, University of London).

- Dion, F., Rakha, H., & Kang, Y. S. (2004). "Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections" *Transportation Research Part B: Methodological*, 38(2), 99-122.
- Federal Highway Administration (2008). "Traffic Signal Timing Manual" Technical Report FHWA-HOP-08-024, U.S. Department of Transportation.
- Lighthill, M. J., & Whitham, G. B. (1955). "On Kinematic Waves II. A Theory of Traffic Flow on Long Crowded Roads" *Proceedings of Royal Society of London*, Vol. 229, No. 1178, May, 317-345.
- Luenberger, D. G., & Ye, Y. (1984). "Linear and nonlinear programming" (Vol. 2) Reading, MA: Addison-wesley.
- May A. D. (1990). "Traffic Flow Fundamentals" Prentice Hall, Englewood Cliffs, New Jersey 07632.
- Michalopoulos, P. G., & Stephanopoulos, G. (1977). "Oversaturated signal systems with queue length constraints—I: Single intersection" *Transportation Research*, 11(6), 413-421.
- Michalopoulos, P. G., & Stephanopoulos, G. (1981). "An application of shock wave theory to traffic signal control" *Transportation Research Part B: Methodological*, 15(1), 35-51.
- Ponlathep, L. (2010). "A simple adaptive signal control algorithm for isolated intersections using time-space diagrams" *Intelligent Transportation Systems (ITSC)*, 13th International IEEE Conference, 273-278
- Pueboobpaphan, R., Yamamoto, F., Nakatsuji, T., & SUZUKI, H. (2005). "Relationship between uninterrupted and interrupted speed at isolated signalized intersection" *Journal of the Eastern Asia Society for Transportation Studies*, 6, 2194-2209.
- Richards, P. I. (1956). "Shock waves on the highway" *Operations research*, 4(1), 42-51.
- Rorbech, J. (1968). "Determining the length of the approach lanes required at signal-controlled intersections on through highways" *Transportation Research*, 2(3), 283-291.
- Roshandeh, A. M., Levinson, H. S., Li, Z., Patel, H., & Zhou, B. (2014). "New methodology for intersection signal timing optimization to simultaneously minimize vehicle and pedestrian delays" *Journal of Transportation Engineering*, 140(5), 04014009.
- Sermpis, D., & Kefallinos, D. (2010). "Real time intervention in traffic signals using fixed-time and traffic responsive signal control strategies" 12th WCTR
- Stephanopoulos, G., Michalopoulos, P. G., & Stephanopoulos, G. (1979). "Modelling and analysis of traffic queue dynamics at signalized intersections" *Transportation Research Part A: General*, 13(5), 295-307.
- Teply, S., Allingham, D. I., Richardson, D. B., & Stephenson, B. W. (2008). "Canadian capacity guide for signalized intersections"
- Webster F. V. (1958). "Traffic Signal Settings" Road Research Laboratory, 39, Her Majesty's Stationery Office, London, England.
- Wong, C. K., & Wong, S. C. (2003). "Lane-based optimization of signal timings for isolated junctions" *Transportation Research Part B: Methodological*, 37(1), 63-84.

A Novel Approach to Assessing Railway Track Quality Based on Ensemble Empirical Mode Decomposition

Zaiwei Li^{1,2,3}; Xiaoyan Lei²; and Liang Gao³

¹School of Urban Rail Transportation, Shanghai Univ. of Engineering Science, Shanghai 201620, China; ²Engineering Research Center of Railway Environmental Vibration and Noise of Ministry of Education, East China Jiaotong Univ., Nanchang 330013, China; ³School of Civil Engineering and Architecture, Beijing Jiaotong Univ., Beijing 100044, China.

Abstract:

The quality of railway track geometry has a critical influence on the safety and comfort of train operations. In order to assess track geometry quality more effectively, the factor of track geometry wavelength should be considered. In this paper, a novel approach of assessment of track geometry quality based on intrinsic mode function (IMF), which derived from the measured data of track geometry by their characteristic wavelength scales, is presented. For this purpose, the frequency spectrum and the energy of each IMF of track geometry are calculated via fast Fourier transform method. The energy coefficient of each IMF of track geometry is defined as the ratio of each energy of IMFs to the total energy of the track geometry, and then the new parameter index is obtained with the corresponding calculation. In comparison with Track Quality Index (TQI) which is the main management method of track geometry in China, the new approach has been proved to be a general form of TQI. Finally, the track geometry data collected by track recording vehicles from Hefei-Wuhan high speed railway is studied using the new parameter index. The statistic analyzing results indicate that the new parameter index is the development and extension of Track Quality Index and can assess the track quality accurately and effectively.

Keywords: Railroad track; Geometry index; Ensemble empirical mode decomposition; Intrinsic Mode function; Track quality.

1 Introduction

High-speed railway has been developing rapidly for years in China. The higher train running speed inevitably makes higher requirements on the track condition, mainly in terms of track geometry condition. Anomalous track geometry conditions can lead to rapid degradation of track components and rolling stock, lading damage, passenger discomfort, and, in extreme cases, train derailments. Regular inspections of track condition provide the primary inputs for track maintenance planning and execution.

Nowadays, track recording vehicles is one of equipment often used to measure track geometry and provide the magnitude and locations of exceedances according to the local line speed to inform track maintenance (Weston et al. 2007).

Track geometry is defined as the deviation of the geometrical position of the rails against their theoretical position resulting from the intended track layout (Luber et al. 2010). The main parameters of the track geometry include gauge, profile, alignment, cross level and twist (Sadeghiand 2010).

Gauge is defined as the right angle distance between the two rails at a given location, measured 16mm below the top surface of the railhead. Profile is the track geometry of each rail or the track centre-line projected onto the longitudinal vertical plane. Alignment is the lateral deviation of the midpoint of the two rails from the nominal centre line of the track or alignment is the average of the lateral position of each rail to the track centre line. Cross Level is the difference in elevation between the top surfaces of the two rails measured at right angles to the track. Twist is the changing height difference in cross level over a predetermined distance. It is the measurement of large change of cross level over a fixed distance (Esveld 2001).

Over recent years, many researches and efforts have been devoted to establish a comprehensive methods for the evaluation of track conditions. Railways of different countries use various statistical or empirical methods to analyze the data obtained from automated inspections (Sadeghiand 2010). These assessment approaches can be mainly classified into three categories, standard deviations of track geometry data, power spectral density functions and complex method. The most commonly used methods on the basis of the standard deviations of track geometry data were developed respectively by the ORE (1981), USA railway (El-Sibaie et al. 2004), China railway (Wei at al. 1996), European railway research center (CEN 2005), Polish railway (Sadeghiand 2010), Swedish national railway (Anderson 2002), and Indian railway (Mundray 2003). The reason for applying this type method is to control the larger deviations of track geometry which usually cause the larger vehicle vibration responses. The quality and stability of the train can be guaranteed when the deviation values are in acceptable limited. However, it has been observed that in some cases this type approach is not sufficient. Some smaller magnitudes of the deviation may show a high correlation with the car body vertical acceleration, and the track geometry wavelength is the factor in this phenomenon. Therefore, the power spectral density functions have been used as diagnostic tool and classifier for track geometry. Corbin et al. (1983) analyzed the power spectral density of track geometry obtained from American railways, and then suggested analytical forms of power spectral density. Iyengar et al (1995) proposed the random field models for vertical irregularity obtained from Indian railways and discussed the vertical excitation power spectral density matrix of multi-axle vehicles. Fryba (1996) summarized various

power spectral density functions commonly used. Chen et al (2008) investigated the general track spectrum of Chinese main railway lines and proposed a scheme to grade track spectrum of Chinese main railway lines according to the running speed of railway. Although these approaches are based on the position of spectrum lines to assess the track quality, there are still some problems in the maintenance such as quantitative limited. Thus, the complex methods were proposed by the scholars from various countries. Hyslip (2002) explored the use of fractal analysis of track geometry data for indicating track geometry condition and evaluating the cause of substructure-related problems. Li et al. (2006) developed a real-time technique relating track geometry to vehicle performance based on the neural network approach. In addition, the methodology of neural network was also employed by Sadeghi et al (2012) to establish the correlations between the track structural conditions and the data obtained from automated inspections. Luber (2009) proposed a method for track geometry assessment by calculating a vehicle response on the basis of system identifications. Sadeghi et al (2009, 2010) suggested incorporating rail cant into the development of track geometry indexes and proposed a coefficient to each geometry parameter based on its role in the overall quality condition of the track. It could be found that these studies greatly promoted the development of the track quality assessment theory. However, there still exist a number of deficiencies in these studies such as inadequate consideration of the factor of track geometry wavelength and practical difficulties. Therefore, this paper attempts to develop a new assessment approach by taking the factor of track geometry wavelength into account. For this purpose, ensemble empirical mode decomposition is implemented, which has been used extensively in various fields such as mechanical engineering, space engineering, railroad engineering (Li et al, 2012).

An outline of the rest of the paper follows. In Section 2, the mathematical theory of empirical mode decomposition is overviewed. In Section 3, the characteristic of intrinsic mode functions of track geometry is presented and the relationship between track geometry and car body vibration acceleration based on empirical mode decomposition is studied. In Section 4, a new approach is proposed to assess track geometry quality by making use of intrinsic mode function according to the wavelength variations. In Section 5, the relationship between the new approach index and Track Quality Index used in China is investigated, and the effectiveness and feasibility of the proposed algorithm will be verified. Finally, in Section 6 conclusions are drawn.

2 Basic Theory of Ensemble Empirical Mode Decomposition

Empirical mode decomposition (EMD) can be applied to deal with non-linear or non-stationary data. The main idea of EMD is to decompose original time series data into a finite and small number of oscillatory modes based on the local characteristic time