Proceedings of the ASME/ASCE/IEEE 2018 Joint Rail Conference
JRC2018
April 18-21, 2018, Pittsburgh, PA, USA

JRC2018-6109

RAIL DEFECT DETECTION USING FIBER OPTIC SENSORS AND WAVELET ALGORITHMS

Saied Taheri, PhD
Virginia Triad Transport Technologies, V3T
Blacksburg, VA, USA

Behzad Moslehi, PhD
Intelligent Fiber Optic Systems Corporation, IFOS
Santa Clara, CA, USA

Vahid Sotoudeh, PhD
Intelligent Fiber Optic Systems Corporation, IFOS
Santa Clara, CA, USA

Brad M. Hopkins, PhD
Independent Consultant
St. Louis, MO, USA

ABSTRACT

Early detection of rail defects can avoid derailments and costly damage to the train and railway infrastructure. Small breaks, cracks or corrugations on the rail can quickly propagate after only a few train cars have passed over it, creating a potential derailment. The current technology makes use of a dedicated instrumented car or a separate railway monitoring vehicle to detect large breaks. These cars are usually equipped with accelerometers mounted on the axle or side frame. The simple detection algorithms use acceleration thresholds which are set at high values to eliminate false positives. As a result, rail surface defects that produce low amplitude acceleration signatures may not be detected, and special track components that produce high amplitude acceleration signatures may be flagged as defects.

This paper presents the results of a feasibility study conducted to develop new and more advanced sensory systems as well as signal processing algorithms capable of detecting various rail surface irregularities. A dynamic wheel-rail interaction model was used to simulate train dynamics as a result of rail defects and to assess the potential of this new technology on rail defect detection. In a future paper, we will present experimental data in support of the proposed model and simulations.

INTRODUCTION

Current technology for rail defect detection is uses an instrumented car [1], a railway monitoring vehicle, or an instrumented in service train [2]. Variety of sensors such as video cameras, accelerometers, strain gauges, acoustic sensors, LIDAR, and ultrasonic wave sensors are used for this purpose [3]. Although these systems work well for large breaks, they cannot detect small defects are sensor accuracy and bandwidth. With the application of fiber optic sensor system, both of these limitations are eliminated. It is anticipated that with this new system, alongside the rail defect detection algorithms, small cracks will be detected.

Fiber-optic sensors provide significant benefits to railway applications when compared to conventional electronic sensors. Optical fiber has negligible size and weight compared to copper wiring, reducing the physical burden imposed by the sensing system. Tens or even hundreds of fiber-optic sensors – measuring multiple parameters including temperature, strain, vibration, etc. – can be multiplexed along a single strand of fiber. Optical fiber's intrinsic immunity to electromagnetic and radiofrequency interference (EMI/RFI, respectively) is a distinct advantage in railway applications, in which trains produce intense electromagnetic fields (especially electrically powered trains). Fiber-optic sensors offer high signal stability and durability during high vibration loads as experienced by trains. The distance between sensing points and the length of the sensing cable do not significantly affect system measurement accuracy because optical fiber has very minimal signal attenuation.

Proper training of the defect detection algorithm requires a large data set consisting of operating conditions and physical parameters. To generate this training data, a dynamic wheel-rail interaction model was developed that relates defect geometry to the side frame vertical acceleration signature. As a result, this present work seeks to generate a dynamic wheel/rail interaction model capable of generating artificial training data representing a wide range of operating conditions.

In the past, several wheel/rail interaction models have been derived. Fermer studied the vertical interaction between the train and track for various pad stiffness values [4]. Dukkipati studied the dynamic oscillations and forces resulting from various static loads, rail stiffness and damping, tie spacing, and ballast damping [5]. Lei presented a model for studying the

vertical response between the wheel and rail due to random rail surface irregularities [6]. Wu presented a finite element rail model with train sprung mass modeled as a static load. The model was used to study the effects of non-linear pad and ballast stiffness characteristics on wheel/rail impact loads [7]. Lundqvist studied the impact loads in rails that arise from unsupported ties [8]. Sun used a VAMPIRE model to determine the wheel impact forces resulting from a turnout or crossing [9].

The dynamic wheel/rail model presented in this paper considers a combined lumped-distributed parameter system. The rail is modeled using Euler-Bernoulli beam theory with appropriate boundary conditions. A multi-body single wheel representation is used to model the train, and several lumped parameter pad-tie-ballast systems are used to model the dynamics beneath the rail. Non-linear stiffness characteristics are considered for the pad and ballast. The combined lumped and distributed parameter system is solved simultaneously in the Matlab environment by utilizing the ode23() function to solve for all of the lumped parameter systems and a fifth order B-spline approximation to solve for the distributed parameter Euler-Bernoulli beam.

This paper presents the simulation model, its validation and the data processing algorithm for use with strain signals/acceleration from a fiber optic sensor system mounted to the side frame of a bogie. The defect detection algorithm is designed to detect and classify rail defects. The system can be installed on revenue trains that is going about its normal day-to-day operation. This can reduce or eliminate the need for testing with separate instrumented cars or railway monitoring vehicles which can be a time consuming and costly process.

The algorithm in this paper works in two primary phases: 1) a defect detection phase, and 2) a defect classification phase. In [10], a generic defect detection and classification algorithm using wavelets was presented. The detection phase utilizes the wavelet transform to locate irregularities in the signal. In recent years, wavelet analysis has shown itself to be useful in event detection. Wavelets are compactly supported, which allows them to be used for decomposing a signal into a time-scale (time-frequency) representation. Traditional digital signal processing makes use of the Fourier transform, which gives the frequency content of a signal but yields no information regarding the location in time or space at which each frequency occurs. The wavelet transform provides localization in both the time and frequency domains, which makes it extremely beneficial in the analysis of non-stationary or transient signals [11-12] like those expected in rail defect detection. Wavelet analysis has been used in various event detection applications, including crack identification in beams [13] and plates [14], machinery health monitoring [15], detection of climate change [16], and event location in pacemakers [17].

The defect classification phase of the algorithm uses the output from the defect detection phase as an input into an artificial neural network (ANN). A similar method has been used in [18] and [19] to detect faulty wheel bearings with a wayside acoustic system. Features are extracted using the discrete wavelet transform and subsequently classified using a genetic algorithm. However, with the current problem of locating rail defects in real time with a limited accelerometer signal, the features present in the signal will be much different than the features present in an acoustic signal for faulty wheel bearings. The neural network in this paper uses the original signal and the first four scales of the wavelet transform as inputs and classifies defects as either a rail break or an impending rail break.

NOMENCLATURE

A	Non-negative constant
d	Number of vanishing moments
$f(t)$	Original signal
j	Discrete scale
J	Maximum scale
N	Length of signal
s	Continuous scale index
t	Time
T	Threshold value
u	Continuous position index
$Wf(s,u)$	Continuous wavelet coefficients
α	Lipschitz exponent
$\psi(t)$	Wavelet function
$\varphi(t)$	Scaling function
A_c	Cross sectional area of rail
b_i	Location of rail tie
c_b	Ballast damping
c_p	Pad damping
c_1	¼ primary suspension damping
c_2	¼ secondary suspension damping
C_H	Hertzian contact spring coefficient
E	Young's Modulus of rail
I	Area moment of inertia of rail
i	Index for rail tie
k_b	Ballast stiffness
k_p	Pad stiffness
k_1	¼ primary suspension stiffness
k_2	¼ secondary suspension stiffness
M_B	¼ bolster mass
M_c	1/8 carbody mass
M_s	½ tie mass
M_w	½ axle + ½ sideframe + wheel mass
r	Wheel radius
t	Time
u_c	Vertical displacement of carbody
u_B	Vertical displacement of bolster
u_w	Vertical displacement of wheel
u_R	Vertical displacement of rail
$u_{s,i}$	Vertical displacement of rail tie
V	Forward speed of the train
x	Coordinate for traveling direction of train
δ	Dirac delta function
ρ	Density of rail

WAVELETS IN EVENT DETECTION

A wave with concentrated, finite amount of energy is mathematically considered to be a wavelet, $\psi(t)$, if and only if its Fourier transform, $\Psi(\omega)$, satisfies:

$$\int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < +\infty \tag{1}$$

Equation (1) suggests that:

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0, \tag{2}$$

which states that a wavelet must have zero mean value. Therefore, when wavelets are used as a basis to represent a signal, they are able to provide bands of both time and frequency information because of their compact support.

The scaled and shifted versions of the mother wavelet form the set of basis functions over which the signal is decomposed. The continuous wavelet transform (CWT) of a function $f(t)$ is defined as:

$$Wf(s,u) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} f(t) \psi^* \left(\frac{t-u}{s} \right) dt, \tag{3}$$

where $\psi^*(t)$ is the complex conjugate of the wavelet. Equation (3) shows that the continuous wavelet transform is the inner product of the signal, $f(t)$, with scaled (s) and shifted (u) versions of the mother wavelet function $\psi(t)$. The wavelet transform is able to accurately detect changes and irregularities in the signal. This can be seen by investigating the use of the scaling function, $\varphi(t)$. For every wavelet function, $\psi(t)$, which acts as a detail, high-pass filter, there is also a scaling function, $\varphi(t)$, that acts as smoothing, low-pass filter. If the wavelet function is chosen to be the first derivative of the scaling function with respect to time, equation (3) can be written as:

$$\begin{aligned} Wf(s,u) &= \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} f(t) \frac{d}{du} \varphi^* \left(\frac{t-u}{s} \right) dt \\ &= \frac{d}{du} \left(\frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} f(t) \varphi^* \left(\frac{t-u}{s} \right) dt \right) \end{aligned} \tag{4}$$

The wavelet transform is therefore proportional to the first derivative of $f(t)$ smoothed by $\varphi(t)$. Smoothing and differentiating the signal results in a new signal that contains information about the regularity of $f(t)$ in its peaks. This means that the modulus maxima of $Wf(s,u)$ are the locations of the sharp irregularities in $f(t)$.

The ability of the CWT to locate irregular events in a signal can further be seen in its ability to estimate the local Lipschitz exponent of a signal, which is a mathematical measure of the degree of regularity in the signal. Mallat [20] showed that the regularity of a point in a signal can be estimated by observing the decay of the wavelet coefficient values across scales. Mallat showed that a function $f(t)$ is Lipschitz α at point u if and only if there exists some non-negative constant A such that

$$|Wf(s,u)| \le As^{\alpha} \tag{5}$$

Taking the logarithm of both sides of this equation gives:

$$\log|Wf(s,u)| \le \log(A) + \alpha \log(s) \tag{6}$$

The wavelet coefficient values with scale values can then be fitted to equation (6) using a linear least-squares regression to determine the value of α at each location u.

In performing the wavelet transform, a wavelet should be chosen with the proper number of vanishing moments, which is related to its ability to represent a signal with certain regularity. A wavelet has d vanishing moments if it satisfies the criteria

$$\int_{-\infty}^{\infty} t^p \psi(t) dt = 0 \text{ for } 0 \le p \le d, \tag{7}$$

where d is an integer. A wavelet with d vanishing moments is orthogonal to a polynomial of degree up to $d - 1$, and can accurately represent a signal with regularity α for $d \ge \alpha$.

ARTIFICIAL NEURAL NETWORKS IN CLASSIFICATION

Artificial Neural Networks (ANN) have the ability to classify information produced from non-linear systems. They have been used in recent years in health monitoring applications such as wheel bearing fault classification [18], [19] and impedance-based structural health monitoring [21]. The power of the ANN in fault classification lies in its ability to be trained to recognize various input patterns and categorize them accordingly. The basic structure of an ANN is shown in Figure 1.

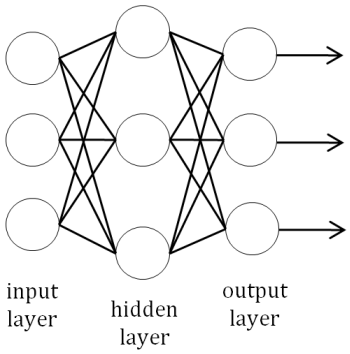

Figure 1. Basic structure of an ANN, with input layers, hidden layers, and output layers shown

Figure 1 shows three layers to the network: an input layer, a hidden layer, and an output layer. The user specifies the number of inputs, hidden nodes, and outputs. At each hidden node, all of the inputs are weighted and summed and then some activation function is applied to the weighted, summed total. This number is the output of the hidden layer. The outputs from the hidden layer are sent to the output layer, where the numbers are weighted and summed and then have some activation function applied to them. These numbers are then the output of the ANN. The mathematical operations applied at each node are relatively simple, but when this is done with many nodes it all combines to make a complex network capable of recognizing patterns in the input signals.

The network can be trained to determine a set of weights that will produce accurate, desirable results. One method of training the network is called supervised learning, where the error is calculated between the actual ANN output and the desired output. A minimization routine is then used to adjust the weights accordingly to minimize the error. Supervised learning can be used when the user has a set of training data with known desired output. The training data along with desired output are fed into the ANN to train it and adjust the weights accordingly. After the network has been trained, it can then be used for classification.

The only downside to supervised training is that it requires a relatively large amount of data to span the entire space of inputs that will occur during actual use. When a limited amount of data is available to the user, artificial training data can be generated using actual data as a basis. Artificial training data is not always an accurate substitute for obtaining a wider range of actual training data that fully covers all possibilities for inputs; however, it is often used in practice and can still produce good results.

SENSORY SYSTEM

This paper proposes a cost-effective, multipoint, multifunction fiber-optic sensor system for broken rail and defect detection. Figure 2 shows details of the fiber Bragg grating (FBG) sensor array installation on a freight car wheel-axle boxes. A tri-axial set of FBG sensors will be installed on each of the left and right axle boxes. Signals from the FBG sensor arrays will be sampled at 1 MS/s with 16-bit resolution. Recorded data will be processed for rail breakage identification.

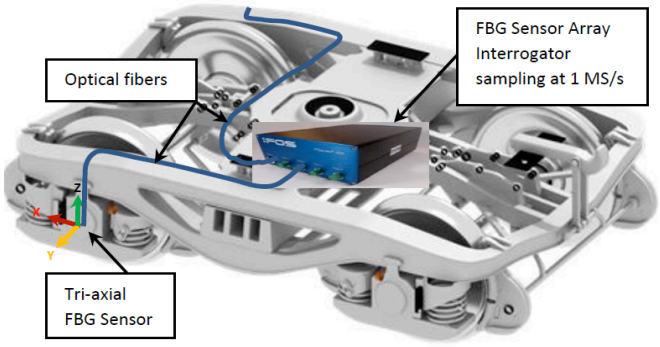

Figure 2. FBG sensor array installation on the freight car wheel-axle boxes

Investigative Approach

FBG sensor: The basic principle of an FBG-based sensing system lies in the resonant frequency of the FBG. The spectral width defined by the pass-band of such a grating forms a channel that reflects a characteristic wavelength, allowing the remainder of the spectrum to pass though the fiber without significant optical loss, as shown schematically in Figure 3. FBGs thus operate by acting as a selective filter that reflects a single wavelength, called the Bragg wavelength, λ_B. The Bragg wavelength is related to the grating pitch, Λ, and the mean refractive index of the FO core, n, by $\lambda_B = 2\Lambda n$. Both n and Λ vary with changes in strain or temperature.

The dynamic response of FBGs must be taken into account for their effectiveness in high-bandwidth monitoring applications.

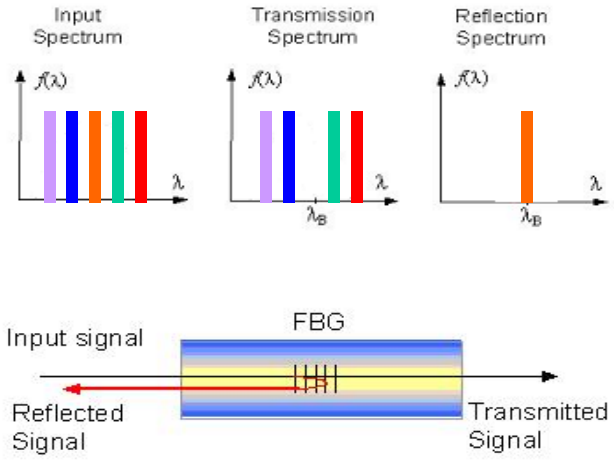

Figure 3. FBG functional principle

As shown by Weaver-Timoshenko-Young, the longitudinal resonance frequency of an FBG, f, in Hz is given by:

$$f = \frac{1}{2}\sqrt{\frac{E}{\rho\, l^2}} \tag{8}$$

where E is the optical fiber Young's Modulus (69 GPa), ρ is the fiber mass density (2648 kg/m^3) and l is the free length of the fiber. Therefore, the resonance frequency of an FBG sensor that is 1-cm in length would be approximately 255 kHz. Embedding FBG-based sensors in composite materials or surface-mounting the FBG with adhesive covering its entire length can reduce the free length of the fiber by orders of magnitude. These techniques have the favorable effect of increasing the sensor resonance frequency by the same degree, thereby making it suitable for high-bandwidth (1+ MHz) monitoring purposes, e.g. for acoustic emissions or transient events such as train wheel contact forces.

High-Speed FBG Sensor Hardware/Software Interrogation System: The FBG interrogator is an integral part of the overall system. As shown in Figure 4, the interrogator consists of an Amplified Spontaneous Emission (ASE) broadband light source, an optical circulator coupler, an array waveguide grating (AWG), photodiode array and photonic signal processing (PSP) electronics.

The ASE illuminates an array of gratings whose reflections are then individually optically filtered by the AWG, which essentially consists of an array of narrow-band optical transmission filters integrated on a planar waveguide, permitting the fast and simultaneous interrogation of multiple wavelengths without any mechanically moving parts or laser scanning. As the strain on the FBG sensor increases or decreases its reflected wavelength shifts upward or downward, causing a change of power ratio between two adjacent channels.

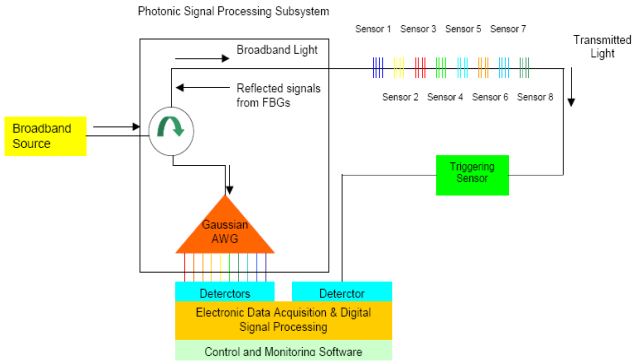

Figure 4: Optical Processor

DYNAMIC WHEEL/RAIL INTERACTION MODEL

To span a wide class of signals for training the defect detection algorithm, data is needed for various size rail defects and other surface irregularities. Since it would be too time consuming and expensive to collect such a large amount of real-world data, a dynamic wheel/rail interaction model was developed that allows for quick generation of training data. A diagram of the model is shown in Figure 5. The train is modeled as a single wheel passing over the rail with vertical dynamics considered. The train consists of three masses: 1/8 of the carbody, ¼ of a bolster, and a wheel, with primary suspension connecting the wheel and bolster and secondary suspension connecting the bolster and carbody. Both suspensions are modeled with a linear spring and a linear damper in parallel. The dynamic forces between the wheel and the rail are modeled using a non-linear Hertzian contact spring. The rail is modeled as a distributed parameter Euler-Bernoulli (EB) beam. The pad is modeled as a non-linear spring and linear damper in parallel, the rail tie is modeled as a mass, and the ballast is modeled as a non-linear spring and linear damper in parallel. The boundary conditions for the EB beam are dependent on the rail surface irregularity that is being modeled. For small surface defects and process noise, a fixed-fixed rail is used with the defect modeled as an additional vertical displacement input at the location of the irregularity. For a broken rail, two rails are used in the simulation, where the wheel travels from fixed-free EB beam to a free-fixed EB beam. In all cases, a single point contact is assumed between the wheel and the rail. It was found in the literature that 10 equally spaced (0.65 m) lumped parameter sleeper bays is appropriate to apply the fixed boundary conditions as previously described [22], [9].

Equations of Motion

The equations of motion for the model in Figure 5 are as follows. The equation of motion for the carbody is:

$$M_c\ddot{u}_c = c_2(\dot{u}_B - \dot{u}_c) + k_2(u_B - u_c) \tag{9}$$

The equation of motion for the bolster is:

$$\begin{aligned} M_B\ddot{u}_B = c_2(\dot{u}_c - \dot{u}_B) + k_2(u_c - u_B) + \ldots \\ \ldots + c_1(\dot{u}_w - \dot{u}_B) + k_1(u_w - u_B) \end{aligned} \tag{10}$$

The equation of motion for the wheel is:

$$M_w\ddot{u}_w = c_1(\dot{u}_B - \dot{u}_w) + k_1(u_B - u_w) + C_h(u_R - u_w) \tag{11}$$

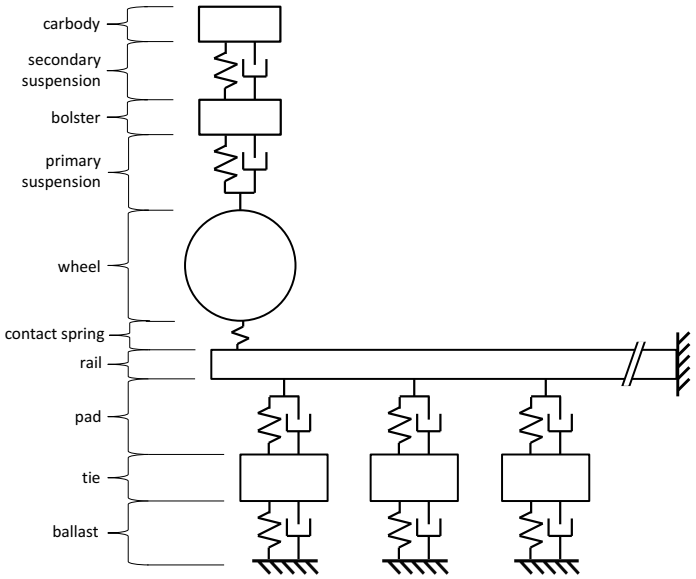

Figure 5. Wheel/rail interaction model: distributed parameter Euler-Bernoulli beam and discrete lumped parameter inputs across the beam.

The equation of motion for the rail is:

$$EI\frac{\partial^4 u_R(x,t)}{\partial x^4}+\rho A\frac{\partial^2 u_R(x,t)}{\partial t^2}=\ldots$$

$$\ldots=\sum_{i=1}^{N}\left[\left(k_{p,i}\left(u_{s,i}-u_R\right)+c_{p,i}\left(\dot{u}_{s,i}-\dot{u}_R\right)\right)\delta\left(x-b_i\right)\right]+\ldots \quad (12)$$

$$\ldots+C_h\left(u_R-u_w\right)\delta\left(x-Vt\right)$$

The equation of motion for each lumped parameter sleeper bay is:

$$M_s\ddot{u}_s=c_p(\dot{u}_R-\dot{u}_s)+k_p(u_R-u_s)-c_b\dot{u}_s-k_bu_s \quad (13)$$

Equations (9) – (13) are numerically solved using a Matlab script. Equations (9), (10), (11), (13) are solved using the ode23 function in Matlab and equation (12) is solved using a fifth order B-spline approximation [23]. The lumped parameter systems (train and sleeper bays) and the distributed parameter system (rail) are all solved for simultaneously by iterating at each time step until the values converge to error less than 0.001%.

System Parameters

Table 1 shows the parameters that were used in the wheel/rail interaction model. Values for the masses, dampers, and springs are actual values from the instrumented freight car used in this study, and values for the rail and foundation are taken from UIC 60 rail parameters as published in [7]. The parameter values in Table 1 accurately represent the true physical parameters of the train/rail system.

Table 1. Parameter values used in wheel/rail interaction model.

Parameter	Symbol	Value	Units
1/8 carbody mass	M_c	16.5×10^3	kg
¼ bolster mass	M_B	184	kg
½ axle + ½ sideframe + wheel mass	M_w	1072	kg
¼ primary suspension stiffness	k_1	4.52×10^6	N/m
¼ primary suspension damping	c_1	5.5×10^4	N/m/s
¼ secondary suspension stiffness	k_2	1.75×10^9	N/m
¼ secondary suspension damping	c_2	8.85×10^3	N/m/s
Hertzian contact spring	C_H	3.5×10^8	$N/m^{3/2}$
Young's Modulus of rail	E	2.1×10^{11}	N/m^2
Area moment of inertia of rail	I	30.55×10^{-6}	m^4
Cross sectional area of rail	A_c	7.69×10^{-3}	m^2
Density of rail	ρ	7850	kg/m^3
Pad stiffness	k_p	$52 + 6.24 \times 10^8 x_p^2$	MN/m
Pad damping	c_p	3.0×10^6	N/m/s
½ tie mass	M_s	40	kg
Ballast stiffness	k_b	$22.75 + 2.60 \times 10^8 x_b^2$	MN/m
Ballast damping	c_b	6.0×10^5	N/m/s
Wheel radius	r	0.4826	m
Forward speed	V	15.65	m/s

Since there is no suspension between the side frames and the wheels, ½ the axle mass, ½ the sideframe mass, and the wheel mass are all lumped together in the parameter M_w. Therefore, in the simulations, the vertical acceleration of the wheel will be observed and considered to be analogous to the side frame acceleration in the experiments.

Broken Rail Model

Figure 6 shows the approach to modeling a broken rail. The broken rail consists of a fixed-free beam of length L_1 on the left side and a free-fixed beam of length L_2 on the right side with a gap of size g between the two. Both the left and right beams are supported by a finite number of lumped parameter pad-tie-ballast systems. The vibration responses of the left and right rails are completely independent of one another. The broken rail model assumes single point contact between the wheel and rail at all times. Therefore, the wheel can only be in contact with the left or right rail, and not both. It is also assumed that there is no externally applied drive or brake torque to the wheel and zero slip between the wheel and rail. During simulation, the train starts at some initial position on the left rail (x_{w0}) and travels right at a constant forward speed. It is assumed that both the left and right rails are un-deflected at time $t = 0$. Therefore, as the train moves to the right, the end of the left rail will gradually become more deflected in the negative direction. Because of this negative deflection, part of the wheel will come into contact with the right rail at some small distance from the free end of the left rail.

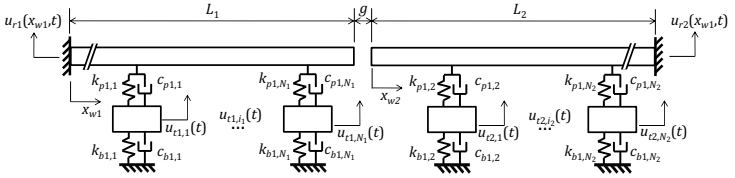

Figure 6. Diagram of broken rail model.

Figure 7 shows the wheel as it is rolling along the left rail just as it comes into contact with the right rail. The wheel remains in contact with the left rail until it reaches a point at distance x_e from the end of the rail. At this point the x location of the wheel on the left rail is $x_{w1} = L_1 - x_e$ and the deflection of the left rail is $u_{r1}(L_1 - x_e, t)$. At this point, the wheel comes into contact with the right rail at point B and immediately loses contact with the left rail at point A.

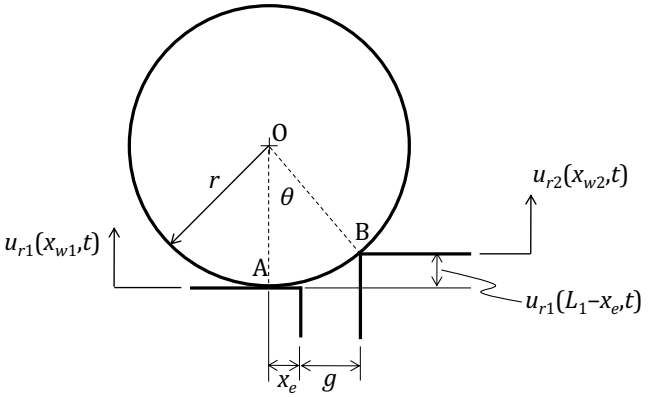

Figure 7. Wheel driving over broken rail.

Figure 8 shows the geometric relationships at the point the wheel contacts the right rail as shown in figure 7. This diagram can be used to determine equations for the broken rail model. Although the Hertz contact model utilizes the deformation at the wheel-rail contact to determine the contact force, this deformation is small and can be considered negligible in deriving the geometric relationships. Therefore, it is assumed that $\overline{OA} = \overline{OB} = r$. The angle $\theta = \angle AOB$ and the distance $h_b = u_{r1}(L_1-x_e,t)$ are also assigned. By viewing AOB as a triangle in a Cartesian coordinate plane, the following coordinates can be assigned to each point: A, (0,0); B, (x_e+g,h_b); O, $(0,r)$. Then, using Pythagoras' theorem the relationship for the distance between point O and point B can be written:

$$r^2 = (0 - (x_e + g))^2 + (r - h_b)^2 \tag{14}$$

Multiplying out and rearranging terms gives the polynomial:

$$x_e^2 + (2g)x_e + (g^2 + h_b^2 - 2rh_b) = 0 \tag{15}$$

The quadratic equation is then used to solve for x_e:

$$x_e = -g + \text{sqrt}(2rh_b - h_b^2) \tag{16}$$

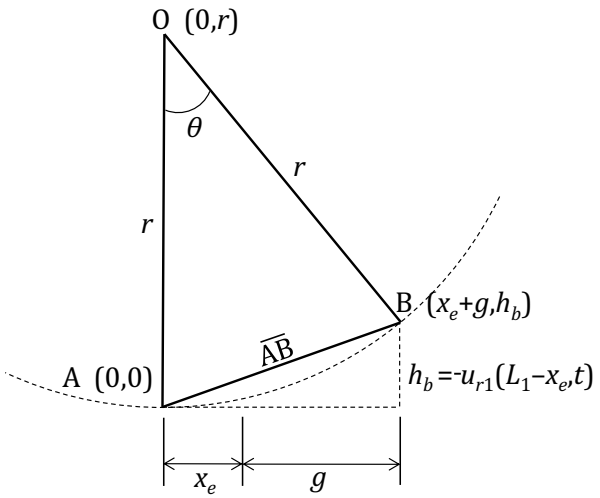

Figure 8. Geometric relationships between wheel and rails at location of broken rail.

Once the wheel comes in contact with the rail at point B, there is a loss of contact at point A and perfect adhesion at point B. The wheel then remains in contact with the right rail at point B until it rotates an angle θ about point B. At this point, the wheel then begins moving to the right along the right rail at a constant forward speed. The time that the wheel remains in contact with the right rail at point B is given by the amount of time it takes the wheel to rotate through an angle θ in degrees, which is:

$$t_b = (2\pi\theta)/\omega, \qquad \omega = 2\pi V_x/r \tag{17}$$

Since the rotational dynamics of the wheel are neglected and single point contact is assumed, as soon as the wheel contacts point B, the wheel is stationary in the longitudinal direction at point B for a time t_b. After this period of time, the wheel moves to the right at a speed of V_x. Since the contact force is modeled with a spring, and the wheel contacts point B at an angle θ, there will be a component of contact force in the longitudinal direction for the duration t_b. In this study, only the vertical dynamics of the train subsystem are of interest. Additionally, since $r \gg g$, $\theta \ll 1$, and $\cos\theta \approx 1$, the total force in the Hertz contact spring is approximately equal to the vertical component of the force in the spring.

The final model exists in the form of a Matlab script that takes train forward speed as an input, and gives the vertical displacement, velocity, and acceleration of the wheel, bolster, and carbody as outputs. The adjustable parameters of the model are: all train bodies and rail tie mass values, spring rates for suspension, pads, and ballast, damping rates for suspension, pads, and ballast, number of pad-tie-ballast systems, tie spacing, length of rail, Young's modulus of the rail, area moment of inertia of the rail, cross sectional area of the rail, density of the rail, and size of the crack/break in the rail.

SIMULATIONS AND RESULTS

To validate the effectiveness of the dynamic wheel/rail interaction model to generate artificial training for the defect detection algorithm, the data collected from the instrumented freight car was compared to the response of the model. The simulated and actual response are for the same parameters and operating conditions as listed in Table 1 while traveling over a rail break with rail separation of 102 mm. The simulation begins with the train starting at x_{w0} = 0.4 m for an un-deflected fixed-free beam. The train moves to the right at a constant forward speed of 15.65 m/s until the wheel contacts the right rail. The rail deflection and vertical displacement, velocity, and acceleration of all train bodies are observed. During the simulation, 200 beam elements are used for each rail in the fifth order B-spline approximation, and a 0.001 second time step is used.

Figure 9 shows the rail responses for the broken rail simulation. At t = 0.414 s, the wheel comes in contact with the

right rail and leaves the left rail. Therefore, the left rail response is shown from 0 s to 0.414 s and the right rail response is shown from 0.414 s to 0.5 s. Figure 10 shows the vertical displacement, vertical velocity, and vertical acceleration responses of the wheel, bolster, and car body. The wheel encounters the broken rail at 0.414 s, which is seen by the discontinuities in the train body responses at this time. In the case of this study, the wheel vertical acceleration is of primary interest as this is what the defect detection algorithm receives as an input.

Figure 11 shows a comparison between the simulation results (bottom right plot of figure 10) and actual data collected from the instrumented freight vehicle. The simulation was run at a sampling frequency of 1000 Hz, but was down-sampled to a rate of 256 Hz to match the sampling frequency of the collected data. The top plot shows the vertical acceleration response at the location of the rail break, and the bottom plot shows the FFT of the vertical acceleration signatures. The top plot shows a double peak for the actual data, which arises from the location of the accelerometer at the bogie side frame and both wheels passing over the broken rail. For means of comparison between the actual and simulated responses, the single oscillation shown in the bottom right plot of figure 10 is repeated twice for the simulated response in the top plot of figure 11.

Note that in the FFT plots the connected line between discrete frequency values is not necessarily an accurate representation of the true frequency content since no accurate assumption can be made as to what the true frequency content is between the discrete values. The connected line is used as a visual aid to better illustrate the comparison between the actual and simulated responses. The results show that the dynamic wheel/rail interaction model simulation closely matches the actual response of the train when driving over a broken rail.

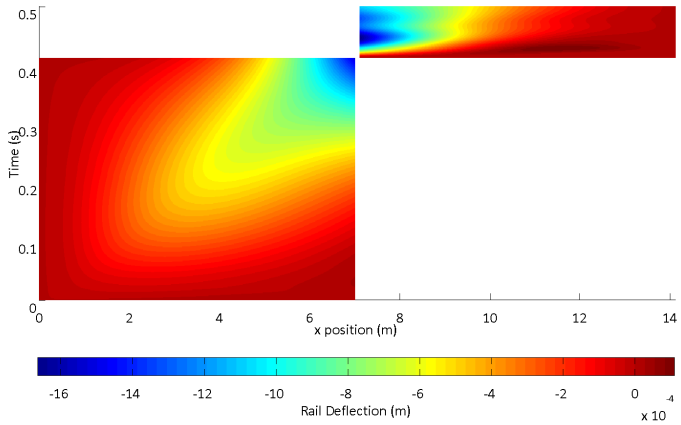

Figure 9. Left and right rail responses for broken rail simulation.

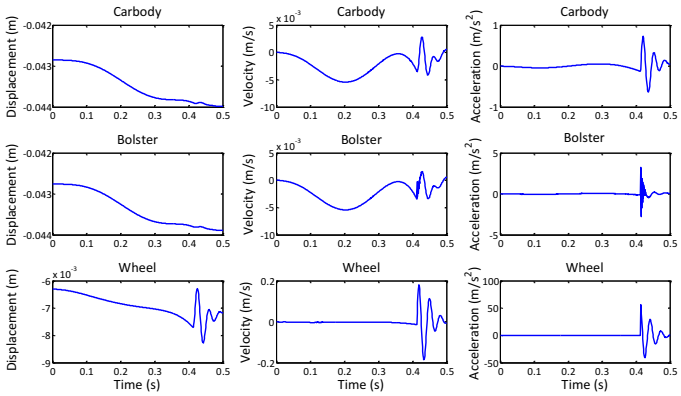

Figure 10. Vertical displacement, velocity, and acceleration of train wheel, bolster, and car body during broken rail simulation.

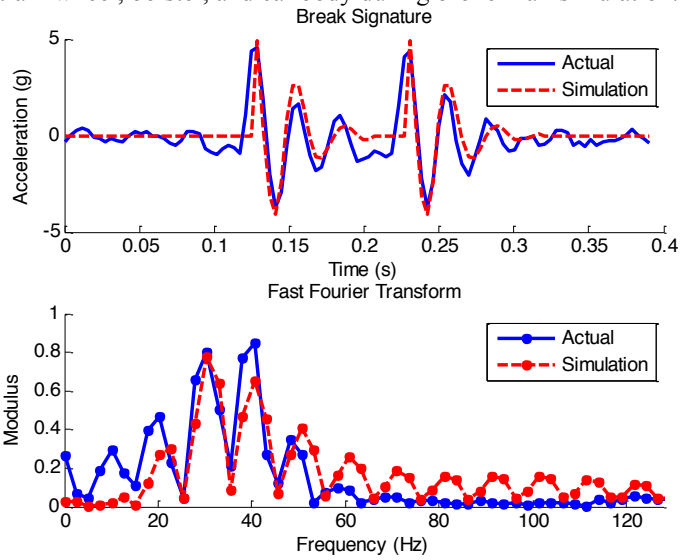

Figure 11. Vertical acceleration signature and frequency content for a 102 mm rail break. Actual and simulated responses are compared.

DEFECT DETECTION AND CLASSIFICATION ALGORITHM

There are three main steps to the rail defect detection algorithm: the wavelet transform, thresholding of the wavelet transformed signal, and classification using the artificial neural network. The wavelet transform and thresholding steps are responsible for locating irregularities in the original accelerometer signal and distinguishing them from all other process noise and sensor noise in the signal. The neural network step is responsible for classifying the defects from the first two steps, which in this case is distinguishing between an impending rail break and a rail break. The three steps to the defect detection and classification algorithm are discussed in greater detail in the following sections.

Wavelet Transform

The Daubechies family of wavelets has proven itself to be effective in event detection because it provides bases of orthonormal wavelets with a maximum number of vanishing moments for a given support length. To determine a suitable

number of vanishing moments for the analyzing wavelet, the local regularity of the test signal was calculated by using equation (6). It was found that a Daubechies wavelet with three vanishing moments (db3) was sufficient to span the regularity of the test signal (sampling rate of 256 Hz) in this study. The center frequency values at the first 5 scales of the db3 wavelet family for a signal sampled at 256 Hz are shown in Table 2. From observation of the test signal, it was found that the vertical acceleration signals from an impending rail break (surface crack) and a rail break contain dominant frequencies around 40 Hz. Additionally, after performing the wavelet transform and observing the results, it was found that no significant information concerning rail defects was located in scale 5 or higher. Therefore, the first 4 scales of the wavelet transform were used in identifying the defects present in the signal.

Table 2. scales of the db3 wavelet family

Scale	Center Frequency (Hz)
1	300
2	150
3	75
4	37.5
5	18.75

Thresholding of Wavelet Transformed Signal

Thresholding of the wavelet transformed signal serves the purpose of removing sensor noise and process noise from the signal, leaving only information concerning the location of defects. Process noise can be considered anything that causes a significant vertical acceleration signal but is not a defect, such as frogs, switches, and joints. It was observed from the test signal that hard threshold values could be applied at each scale to successfully remove all sensor and process noise and leave only the signatures generated from the impending break and the break. The value of the required threshold at each wavelet transform scale was observed to a have a sigmoidal shape. The equation for determining the hard threshold value at each scale is:

$$T(j) = \frac{P}{a + e^{-mj}} \tag{18}$$

where T is the threshold value, j is the discrete scale, and the constants in the equation are $P = 0.22$, $a = 0.025$, and $m = 1.25$. Table 3 shows the threshold values calculated for the first four scales of the wavelet transform using equation (18).

Table 3. Threshold values at the first four scales of the wavelet transform, calculated using equation (14)

Scale (j)	Threshold ($T(j)$)
1	0.71
2	2.05
3	4.53
4	6.93

The threshold values from Table 3 are applied to the wavelet transformed signal by using the following criteria:

$$y_{j,k} = \begin{cases} |d_{j,k}| & |d_{j,k}| \geq T(j) \\ 0 & |d_{j,k}| < T(j) \end{cases} \tag{19}$$

where $y_{j,k}$ is the output of the thresholding step that is sent to the next step as an input to the neural network.

The threshold values in Table 3 are not necessarily universal for all signals that will be encountered during typical operation. It can only be conclusively said that these threshold values are accurate for the class of signals represented by the test signal used in this study. For a more comprehensive set of signals spanning a wider range of operating conditions, including various forward speeds and payloads, new threshold values would have to be developed.

Neural Network Classification

The neural network receives five inputs: the raw signal and the first four threshold scales of the wavelet transform. The network has one hidden layer with 100 neurons and uses a radial basis function as an activation function [24]. The network has one output. The network was trained with learning data to produce an output of 0 when there is no defect, an output of 0.7 when there is an impending rail break/surface crack, and an output of 1.0 when there is a rail break.

Artificial training data was generated by using the test signal and adding various levels of Gaussian white noise to it. Four sets of training data were produced with signal to noise ratios of 97.86, 24.62, 10.86, and 6.13. The network was trained until the error converged to a value close to zero, and the resulting weights were used in the final neural network.

CONCLUSIONS

A defect detection algorithm has been developed for locating and classifying impending rail breaks (surface crack) and rail breaks from an FBG tri-axial fiber optics sensor array mounted to the bogie side frame or axle of a freight car. The algorithm uses wavelets as a basis to process the signal to remove all sensor noise and process noise, where process noise is considered to be anything that shows up in the signal that was not generated from a defect, like the signatures generated from frogs, switches, and joints. The denoised signal is passed on to an artificial neural network to classify the determined defects as either an impending rail break or a rail break.

Additionally, a dynamic wheel/rail interaction model was developed for the purpose of generating training data for a rail defect detection algorithm. The algorithm of interest classifies bogie side frame vertical acceleration signatures to determine the nature of various surface irregularities. Training of the algorithm using actual data is unrealistic because of the wide range of operating conditions that is necessary to ensure accurate defect detection during use. The dynamic wheel/rail interaction model presented in this paper offers a solution for

generating artificial training data that spans a wide range of operating conditions.

The wheel/rail model presented in this paper simultaneously solves for a combined lumped-distributed parameter system. A bogie side frame vertical acceleration signature obtained from driving an instrumented freight car with known operating conditions over a broken rail was compared to the response generated using the model. The results show that the response of the model closely matches the collected data. Additionally, it was shown that the frequency content of the simulated response closely matches that of the actual response. This is important, as the defect detection algorithm that this model will be applied to utilizes a simultaneous time-frequency domain approach to analyze the signal. Future work with this model includes the generation of various vertical acceleration signatures for a wide range of operating conditions for proper training of the defect detection algorithm.

REFERENCES

[1] D. Li, Meddah, A., Lundber, W., "Instrumented Freight Car for Performance-based Track Inspection," TTCi Technology Digest, 2008.

[2] P. F. Weston, Ling, C.S., Goodman, C.J., Roberts, C., Li, P., Goodall, R.M., "Monitoring lateral track irregularity from in-service railway vehicles," Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, vol. 221, pp. 89 - 100, 2007.

[3] F. Lanza, Coccia, S., Bartoli, I., "On-line High-speed Rail Defect Detection - Phase III," U.S. Department of Transportation, Federeal Railroad Administration2005.

[4] M. Fermer, Nielsen, J.C.O., "Vertical interaction between train and track with soft and stiff railpads -- full-scale experiments and theory," Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, vol. 209, pp. 39 - 47, 1995.

[5] R. V. Dukkipati, Dong, R., "Idealized steady state interaction between railway vehicle and track," Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, vol. 213, pp. 15 - 29, 1999.

[6] X. Lei, Noda, N.-A., "Analyses of dynamic response of vehicle and track coupling system with random irregularity of track vertical profile," Journal of Sound and Vibration, vol. 258, pp. 147-165, 2002.

[7] T. X. Wu, Thompson, D.J., "The effects of track non-linearity on wheel/rail impact," Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, vol. 218, pp. 1 - 15, 2004.

[8] A. Lundqvist, Dahlberg, T., "Load impact on railway track due to unsupported sleepers," Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, vol. 219, pp. 67 - 77, 2005.

[9] Y. Q. Sun, Cole, C., McClanachan, M., "The calculation of wheel impact force due to the interaction between vehicle and turnout," Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, vol. 224, pp. 391-403, 2010.

[10] B. Hopkins, Taheri, S., "Track Health Monitoring Using Wavelets," presented at the ASME Rail Transportation Division Fall Technical Conference, Roanoke, VA, 2010.

[11] S. C. Burrus, Gopinath, R.A., Guo, H., Introduction to Wavelets and Wavelet Transforms: A Primer. Upper Saddle River, New Jersey: Prentice Hall, 1998.

[12] C. Varadharajan, "A Wavelet-Based System for Event Detection in Online Real-Time Sensor Data," Master of Science, Civil Engineering, Massachusetts Institute of Technology, Cambridge, MA, 2001.

[13] E. Douka, Loutridis, S., Trochidis, A., "Crack identification in beams using wavelet analysis," International Journal of Solids and Structures, vol. 40, pp. 3557 - 3569, 2003.

[14] E. Douka, Loutridis, S., Trochidis, A., "Crack identification in plates using wavelet analsysis," Journal of Sound and Vibration, vol. 270, pp. 279 - 295, 2004.

[15] Q. Miao, Huang, H., Fan, X., "Singularity detection in machinery health monitoring using Lipschitz exponent function," Journal of Mechanical Science and Technology, vol. 21, pp. 737-744, 2007.

[16] K. M. Lau, Weng, H. (1995) Climate signal detection using wavelet transofrm: how to make a time series sing. Bulletin of the American Meteorological Society. 2391 - 2402.

[17] M. Astrom, Olmos, S., Sornmo, L., "Wavelet based event detection in pacemakers," presented at the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Instanbul, Turkey, 2001.

[18] C. H. Choe, Wan, Y., Chan, A.K., "Neural pattern identification of railroad wheel-bearing faults from audible acoustic signals: Comparison of the FFT, CWT, and DWT features," in Proceedings of the SPIE 3078, 1997, pp. 480 - 486.

[19] P. Xu, Chan, A.K., "Fast and robust neural network based wheel bearing fault detection with optimal wavelet features," in International Joint Conference on Neural Networks, Honolulu, Hawaii, 2002.

[20] S. Mallat, Hwang, W.L., "Singularity detection and processing with wavelets," IEEE Transactions on Information Theory, vol. 38, pp. 617 - 643, 1992.

[21] J. Lopes, V., Park, G., Cudney, H.H., Inman, D.J., "Impedance-Based Structural Health Monitoring with Artificial Neural Networks," *Journal of Intelligent Material Systems and Structures,* vol. 11, pp. 206-214, 2000.

[22] K. Manabe, "Multiple-wheel induced vibration of rail with surface irregularity," Quarterly Report of RTRI, vol. 45, pp. 136-141, 2004.

[23] H. Caglar, Caglar, N., "Fifth-degree B-spline solution for a fourth-order parabolic partial differential equation," *Applied Mathematics and Computation,* vol. 201, pp. 597-603, 2008.

[24] S. Chen, Cowan, C.F.N., Grant, P. M., "Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks" *IEEE Transactions on Neural Networks,* vol. 2, pp. 302 - 309, 1991.

