Condition monitoring opportunities using vehicle-based sensors

This item was submitted to Loughborough University’s Institutional Repository by the/an author.


Metadata Record: https://dspace.lboro.ac.uk/2134/6643

Version: Published

Publisher: © Professional Engineering Publishing

Please cite the published version.
This item was submitted to Loughborough’s Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to: http://creativecommons.org/licenses/by-nc-nd/2.5/
Condition monitoring opportunities using vehicle-based sensors

C P Ward¹, P F Weston², E J C Stewart³, H Li³, R M Goodall¹*, C Roberts², T X Mei³, G Charles⁵, and R Dixon¹

¹Control Systems Group, Electronic and Electrical Engineering, Loughborough University, Loughborough, UK
²Birmingham Centre for Rail Research and Education, The University of Birmingham, Edgbaston, Birmingham, UK
³Department of Engineering and Technology, Manchester Metropolitan University, Manchester, UK
⁴School of Computing, Science and Engineering, The University of Salford, Salford, UK
⁵School of Engineering, The University of Nottingham, University Park, Nottingham, UK

The manuscript was received on 16 March 2010 and was accepted after revision for publication on 21 July 2010.

DOI: 10.1243/09544097JRRT406

Abstract: Recent increases in railway patronage worldwide have created pressure on rolling stock and railway infrastructure through the demand to improve the capacity and punctuality of the whole system, and this demand must also be balanced with reducing life-cycle costs. Condition monitoring is seen as a significant contributor in achieving this. The emphasis of this article is on the use of sensors mounted on rolling stock to monitor the condition of infrastructure and the rolling stock itself. This is set in the context of modern rolling stock being fitted with high-capacity communication buses and multiple sensors, resulting in the potential for advanced processing of collected data. This article brings together linked research that uses a similar set of rolling stock sensors, and discusses: general usage and benefits, a track defect detection method, running gear condition monitoring, and absolute train speed detection.

Keywords: condition monitoring, real-time, vehicle-based sensors, track defects, parameter estimation, low adhesion, wheel–rail profile, vehicle speed, inertial sensors

1 INTRODUCTION

The last two decades have seen widespread increases in railway patronage worldwide, meaning there is significant pressure to improve the capacity and punctuality of rail services, while reducing life-cycle costs. Condition monitoring systems are seen as a significant contributor in achieving such improvements.

Broadly speaking, four different types of monitoring systems exist: infrastructure-based infrastructure monitoring, rolling-stock-based infrastructure monitoring, rolling-stock-based rolling stock monitoring, and infrastructure-based rolling stock monitoring [1]. To date, the use of such approaches for fault detection and diagnosis purposes has been relatively straightforward. Dedicated measurement trains are normal in many railway administrations for assessing the condition of the track [2], and there are some examples where simplified versions of these measuring systems have been fitted to service vehicles. For rolling stock monitoring, some existing fleets have enhanced their on-board data logging capabilities with global positioning system (GPS) and communications equipment in order to analyse real-time data from in-service train sets (for example, see references [3] to [5]).

This article is concerned with monitoring based on measurements made by sensors fitted to vehicles (i.e. the second and third types described in the previous paragraph, in particular focused upon in-service vehicles). Electronic and software systems now form an essential element of railway vehicle technology; since it is common for such systems to be connected to a central control processor through a train communication bus, there is significant potential for advanced processing techniques that can extract more sophisticated system/sub-system knowledge, either from the sensor data currently available or from additional sensors connected to the communications bus. Modern wireless communication systems also provide the opportunity for transferring either the raw sensor data or information derived from such data to track-based

*Corresponding author: Department of Electric and Electronic, University of Loughborough, Loughborough LE11 3TU, UK. email: r.m.goodall@lboro.ac.uk
information systems that can provide further analysis. However, the emphasis here is upon the potential for advanced vehicle-based processing.

The purpose of the article is to describe linked research activities all using a relatively common set of vehicle-based sensors combined with advanced processing concepts, which are aimed towards different technical objectives. Section 2 summarizes the opportunities from a sensing technology viewpoint to provide an overview, and then sections 3 to 5 describe three specific processing options – identification of track defects, monitoring of running gear condition, and determination of absolute train speed. The concluding section summarizes where these developments can contribute at a systems level and identifies the longer-term trends arising from the use of advanced processing concepts such as these.

2 BOGIE-MOUNTED SENSOR OPTIONS FOR MONITORING IN GENERAL – TYPE, LOCATION, NUMBER, AND REQUIREMENTS

Sensors mounted on in-service vehicles can be used to identify certain track defects, monitor the running gear condition, and determine the absolute train speed, all during normal revenue service.

In the case of identifying track defects, the alternative approach is to use specialist track inspection trains. However, due to capacity constraints and the availability and cost of the inspection train itself, it is difficult to monitor the whole rail network in a timely and cost-effective manner.

Table 1 shows an appropriate sensor set for use on in-service vehicles, and this is also illustrated by the diagram given in Fig. 1. This is quite a comprehensive list of sensing possibilities, and it is therefore unlikely that all would be fitted to any particular bogie, although they have generally been chosen as relatively low-cost items. It is difficult to quantify costs pre-production, and so the term ‘low-cost’ is being used qualitatively, but the expectation is that the cost of the sensors themselves will be marginal with respect to the cost of a modern bogie. The proposed sensor set comprises predominantly inertial sensors that are cheap and easy to fit, often in a single box. It is also worth noting that developments in Micro-Electro-Mechanical-Systems technology and increasing use within the automotive industry continue to drive down the costs of these sensors.

Sensor information from a suitably instrumented bogie can identify certain irregularities and defects on the track. For example, a pitch rate gyro can be used to obtain the mean vertical alignment of the track at wavelengths longer than those corresponding to the bogie pitch mode. Axlebox accelerometers can be used to measure shorter wavelength vertical irregularity. Similarly, the bogie roll rate gyro gives
an approximation of the track cross level for longer wavelengths, while the difference between the axlebox
vertical motions can be used for shorter wavelengths.
The absolute roll can be estimated using a combination of lateral sensing accelerometer and roll and yaw rate gyros on the bogie. The use of axlebox accelerometers and roll rate gyro allows the twist from the design
transitions to be included in the absolute twist estimate. In the processing of the acquired data, the
speed of the vehicle is also necessary to perform the conversion between time and displacement along
the track.

Sensors measuring the dynamic response of a bogie
to excitations from track irregularity and other inputs
can also be used to identify variations in performance arising from faults and/or wear in the mechanical
components (springs, dampers, etc). Of course,
(through, e.g. inverse models) the same sensor set can also be used to identify the track inputs which excite the bogie dynamics.

3 DETECTING TRACK DEFECTS

Perfect track alignment results in wheelsets, bogies,
and vehicle bodies following smooth trajectories
through space, following gradual changes in height
and sweeping through horizontal curves, typically
accompanied by a suitable tilt on canted track. Imperfect track results in deviations being superimposed
on the otherwise smooth trajectories that can be used
to identify random track irregularities, or discrete
defects. Attempts to monitor track irregularity have
used vertically and/or laterally sensing accelerometers
mounted on the body [6], bogie [7], or axlebox [8, 9]. In
some cases, a measure derived from acceleration (such
as frequency-weighted rms) is used to identify poor
ride quality that is generally associated with poor track
geometry. The most extensive and expensive option is
to fit a full track geometry measuring system with an
inertial measurement unit on the bogie coupled with
optical line sensors.

3.1 Sensors

Axlebox-mounted accelerometers can provide short
wavelength information about the vertical profile, ideal for detecting corrugation and bad rail joints
[8]. The vertical acceleration associated with a 30 m
wavelength irregularity with an amplitude of 2 mm,
on a vehicle travelling at 45 m/s (100 mph), is only
0.18 m/s². An axlebox-mounted accelerometer typically
has a range of 100 g (1000 m/s²). Hence, there
is likely to be a problem with poor signal-to-noise
ratio. A practical solution to this problem is to use
very high-quality accelerometers with a smaller oper-
ating range mounted on the bogie, and to measure the
displacement down to left and right axleboxes using
displacement sensors. The bogie is subject to smaller
accelerations as the primary suspension filters out
some of the high-frequency, high-acceleration signals.
This provides the ability to obtain results at speeds
down to about 15 km/h.

Even though the bogie is isolated from the rails
by the primary suspension, the bogie orientation
and motion inevitably tend to follow the track. For
vertical track irregularities with wavelengths longer
than the bogie wheelbase, at which frequency the
filtering effect of the primary suspension is small, bogie vertical motion and vertical track irregularity are very similar. Hence, it is possible to monitor track vertical irregularity by using a bogie-mounted, vertically sensing accelerometer to reconstruct the vertical path taken by the bogie. In practice, the double integration of the drifting offset in the sensor output and of noise within the sensor and generated in the analogue-to-digital conversion process results in uncontrollably large errors at increasing wavelengths. This means that some high-pass filtering is required so that only sufficiently accurately reconstructed wavelengths remain. The position on the bogie at which the accelerometer is attached considerably affects the results obtained. One sensible location is above the centre of a wheelset.

A pitch rate gyro attached to a bogie can also measure the vertical trajectory taken by the bogie [10]. A vertical curvature signal is obtained by dividing the pitch rate (measured by the pitch rate gyro) by the vehicle speed, similar to dividing the curvature by the square of the vehicle speed. This curvature signal can be doubly integrated in the spatial domain to give a vertical alignment (irregularity) from which various quantities can be obtained using various high-pass filters. As with an accelerometer, the long wavelength information is lost in the errors from double integration. However, the signal-to-noise ratio turns out to be more favourable using a pitch rate gyro, partly because the pitch rate signal increases linearly with speed instead of with the square of the vehicle speed, and also because the signal falls off less quickly at low frequencies. Hence, a pitch rate gyro provides results at longer wavelengths at low vehicle speeds than an accelerometer of similar quality (and cost). This is particularly advantageous when mounted on an in-service vehicle that makes frequent station stops, as compared to a track recording vehicle that can travel without slowing down too often. In addition, the location of the pitch rate gyro on the bogie is much less important than that of an accelerometer.

Similar results are obtained using a yaw rate gyro to monitor lateral irregularity rather than a laterally sensing accelerometer. However, the bogie does not follow the lateral alignment of the track as closely as it follows the vertical alignment. There is usually a very stiff primary natural suspension and the wheelsets can move laterally with respect to the track, which of course does not happen in the vertical direction. These lateral and yaw kinematic motions are at significantly lower frequencies than the primary vertical suspension, which means that this is a significant effect. However, because of the dynamics relating lateral irregularity to the path taken by the bogie (kinematic wavelength), it is theoretically possible to determine a transfer function to negate this effect and to return to an estimate of lateral track irregularity [11].

### 3.2 Processing

The processing chain was found to result in the best performance for detecting vertical and lateral irregularity described in detail in reference [10]. In summary, it consists of obtaining curvature by dividing the time-domain pitch rate by the instantaneous vehicle speed. These time-domain curvature samples are resampled into the spatial domain, using the speed of the vehicle. In this article, they are 0.125 m apart. The resulting spatial-domain curvature samples are doubly integrated with respect to distance along the track and then high-pass filtered to obtain estimates of vertical irregularity with wavelengths less than 35 m or 70 m, for example. The processing to obtain lateral irregularity is identical but uses data from the yaw rate gyro.

### 3.3 Examples

Some examples of vertical and lateral track irregularities obtained from sensors mounted on the bogies of a Tyne and Wear Metro vehicle and a Class 175 mainline vehicle have been published [10, 11]. More recently, an inertial measurement system, comprising three accelerometers and three rate gyros, has been mounted on the bogie of a Class 508 vehicle as part of an energy monitoring system. Custom-designed and built electronics sample the sensors at 8192 Hz and the results are down sampled to allow data to be saved 256 times a second to local flash memory. A tacho signal to provide train speed is also recorded. Data will be collected continuously from an in-service vehicle.

In reference [10], the vertical and lateral alignment are considered over wavelengths longer than the bogie wheelbase, where the bogie pitch and yaw follow the track slope and heading fairly closely. However, one can see details at shorter wavelengths by examining the lateral and vertical curvature signals. As the wheelsets encounter an irregularity such as a dipped joint, or a disruption to the heading, the leading and the trailing wheelsets are affected in turn, spatially separated by the bogie wheelbase.

Figure 2 shows the vertical 35 m alignment over 100 m of jointed track, obtained as described in reference [10], together with the vertical curvature signal. The vertical alignment shows the presence of dipped joints spaced approximately 18 m apart (consistent with 60 ft rail sections). The vertical curvature shows a characteristic pattern associated with dipped joints. In particular, there are step changes in curvature when the leading wheelset reaches the bottom of the dip and a step change in the opposite direction, approximately 2.5 m later, when the trailing wheelset reaches the same point. The magnitude of the step change is related to the severity of the dip. Hence, the joints can be monitored not only from the point of view of how
deep they are from the perspective of the bogie but also in terms of dip angle.

A second example concerns the motion of the bogie through a pair of back-to-back switches and crossings forming a crossover, where the train is travelling at approximately 5 m/s. Figure 3 shows the 35 m lateral alignment through a crossover and over approximately 100 m of plain track beyond. Labels A to F are used for alignment with later figures. As is typical, the lateral irregularity around the switch and crossing work is significantly higher than that on the plain track. The rotational symmetry of the physical crossing that would be expected is apparent about the zero point midway between C and D. Figure 4 shows the bogie yaw angle through the same crossover. The start and end of the crossover are at A and F, respectively, roughly 75 m apart, based on the tacho signal that may have significant inaccuracies at low speeds.

The maximum (negative) yaw is nearly 0.077 radians, which is consistent with a 1 in 13 crossing angle. Other small features are visible, but it is difficult to interpret their cause.

Figure 5 shows the vertical and lateral curvature seen from the yaw and pitch rate gyros together with the tacho signal, through the same crossing. The figure shows a greater level of detail about the switch and crossing work than is available in Figs 3 and 4.

From the lateral curvature, the progress over the crossover is characterized by curving to the left (negative curvature) as the bogie leaves the first track, then going more or less straight, and finally curving to the right to join the other track. It is possible to infer the causes of some of the features in the vertical and lateral curvatures. At A, the bogie begins to be turned away from the initial track, and so this must be where the switch blade tips are located. There is a disturbance in the vertical curvature at C and D, similar to those
seen previously for jointed track but of smaller magnitude. This is interpreted as the two crossings where the vertical path of the bogie is disturbed as the leading and then the trailing wheelsets pass. The lateral curvature shows some mild transient effects over the crossings. The vertical curvature suggests that the toe of the trailing switch is approximately at E; but the exact position is not clear from the lateral curvature. The distances from switch toe to heal (A to B) and switch toe to crossing (A to C) are reasonably consistent with a standard 1 in 13 crossing.

3.4 Summary and future work

Sensors mounted only on the bogie do not allow left and right vertical rail irregularity to be separated, and do not give any information about gauge or twist. Information about vertical and lateral track alignment over 35 m or 70 m can be obtained from yaw and pitch rate gyros mounted on the bogie, in combination with a tacho signal. While this alignment information is good for establishing general track condition, it does not reveal shorter wavelength details such as those found at dipped joints and through switch and crossing work. However, these particular details have been shown to be observed in the curvature signal derived from the pitch and yaw rate gyros.

The next phase of in-service trials is intended to provide data from which changes over time may be observed. Data will also be obtained at different speeds over the same sections of track, which will allow the consistency of the curvature information to be assessed.

4 MONITORING DYNAMIC PERFORMANCE CHARACTERISTICS

The bogie, with its associated suspension components and wheelsets, consumes a large proportion of the maintenance budget for rolling stock. The bogie has a variety of tasks, principally to provide guidance both on straight track and through curves, to ensure dynamic stability, and to provide ride comfort for the passengers. Failure mode studies have shown that the majority of vehicle faults emanate from faulty wheel profiles and suspension components [12, 13].

Currently, railway condition monitoring for bogie-related applications is primarily through signal processing and knowledge-based assessments [14]. There is potential for increased performance of these techniques if a priori knowledge of the system is used in the form of a system model [15]. Therefore, all of the techniques presented here use a form of model-based estimation.

Presented in this section are a number of techniques for real-time parameter detection for three safety critical aspects of the bogie. Firstly, the suspension parameters are estimated, then wheel–rail adhesion forces, and finally, three approaches to wheel–rail profile estimation are described.

4.1 Suspension parameter estimation

Suspension parameter estimation has previously been reported in references [16] to [18]. These articles describe the use of model-based Rao–Blackwellized Particle Filters to determine the condition of secondary lateral and yaw dampers. In addition, the effective conicity of a wheel–rail combination is estimated.

Simulation work showed that with a full idealized sensor set on the wheelsets, the bogie and the body measuring all of the lateral and yaw accelerations (Fig. 6) plus a detailed knowledge of the track disturbance, the suspension parameters, and conicity could be estimated with confidence. Eliminating the sensors on the wheelsets marginally reduced the quality of the estimates. When uncertainty was added to the lateral track disturbance signal, the suspension parameter estimates were largely unaffected. The conicity estimates, however, failed to converge to the expected value of $\lambda = 0.15$ and settled to an incorrect value with a steady-state offset that depended on the assumption of the input disturbance. This observation was repeated with estimates from data gathered on a Class 175 Coradia vehicle as can be seen in Fig. 7.

Estimation of the suspension component parameters validated the use of a model-based approach that was subsequently adopted to estimate adhesion levels, as described in the next subsection. The poor conicity estimates obtained in this study motivated a search for new techniques that will be described in section 4.3.

4.2 Low-adhesion estimation

Low-adhesion conditions are a continuing problem for many railways. These conditions result in significant disruption to timetabled operations, particularly during the leaf fall season, and in some cases, can lead to signals being passed at danger. Anything that can help identify such conditions is potentially an important contribution to railway technology. Such knowledge could be used within a train, or transmitted to a national adhesion management system to inform drivers of local reductions in adhesion characteristics. This subsection summarizes work on a novel approach based on monitoring bogie dynamic performance [19].

The work proposes that adhesion conditions can be estimated in real time from dynamic measurements of the vehicle's lateral dynamics without the need to apply braking. This technique relies on the assumption that forces generated for guidance are the same as those used for braking performance.
As with the previous section, the model used was a plan view, half body vehicle, and single bogie with two wheelsets, as shown in Fig. 6. It is assumed that ideal sensors are present that can measure the lateral and yaw accelerations of the wheelsets, the lateral and yaw accelerations of the bogie, and the lateral acceleration of the vehicle body. Wheel–rail contact forces are typically calculated as a function of creep in the contact patch and are linearized using Kalker coefficients. This generalization is normally used for stability and control calculations; however, force non-linearities are important in this case to understand creep forces up to and beyond adhesion saturation. Use was made of the contact force model of Polach [20], which is effectively a curve fitting mechanism.

Experimentation has shown that the initial slope of the creep curve varies with different adhesion properties. Figure 8 shows the creep curves for varying conditions for fixed contact patch size and load. Although the theory predicts that at zero creepage
the gradient should remain constant for different adhesion conditions, a number of experiments have shown that this is not the case [21, 22]. This characteristic is very important to the concept of low adhesion detection, because it governs the ability to detect adhesion level differences during normal unsaturated dynamic running.

A pragmatic approach to finding the contact force is to estimate its value in the wheel–rail contact. This model ignores the complex non-linear relations and instead considers the system as a rigid body ‘floating’ on a series of contact points. Initial application of this technique on longitudinal forces at each wheel found these individual forces to be unobservable, and therefore this technique was applied to the net lateral force and net yaw moment at each wheelset.

Simulations performed showed that there is some difference between the estimated creep and the real creep forces due to the Kalman filter not distinguishing between creep and gravitational forces. Varying adhesion conditions can be detected by looking at the power spectral density (PSD) of the estimated creep force and creep moment time samples. This can be observed by the peak of the PSD for the creep forces reducing as the adhesion condition worsens (Fig. 9). In this half vehicle simulation, the trailing wheelset in the bogie displayed the largest difference in creep forces between adhesion conditions. Further tests will be required on representative models of a full rail vehicle composition to determine which wheelsets will provide the best signals for adhesion detection.

4.3 Wheel–rail profile estimation

The wheel–rail contact interaction is one of the most important elements of the rail system. The characteristic of this contact governs the straight-line stability and the cornering performance of the rolling stock. However, this geometric relationship changes with time as the wheel and railhead wear. Currently, the conditions of these two components are monitored on a scheduled basis and are measured independently. The concepts described in the following subsections are for real-time assessment of the wheel–rail contact geometry.

The speculative nature of this work requires a simple model of a single wheelset with suspended mass (Fig. 10). The wheelset dynamics have lateral, roll, and yaw degrees of freedom, with vertical and longitudinal modes omitted due to the small level of coupling between the planes. The suspended mass has a lateral degree of freedom only.

Sections 4.3.2 to 4.3.4 review simulation results of model-based parameter estimation techniques. Included are: the use of Kalman filtering for estimating effective conicity [19]; least squares estimation of conicity using a piecewise cubic polynomial (PCP) function [23, 24]; and, using local linear recursive least squares estimation of the rolling radii and contact angles. The next subsection touches on the source of models used for each of these methods, although each technique makes its own assumptions and approximations.
4.3.1 System model

A non-linear model for the wheelset lateral and yaw dynamics is taken from reference [25] and is shown in Appendix 2 as equations (11) and (12). The accompanying equations for the suspended mass, the lateral suspension force, and the yaw suspension force are shown in equations (13) to (15), respectively. Static non-linearities are present in the wheelset dynamics in the form of the contact geometry described by the left and right rolling radii and contact angles \((r_L, r_R, \delta_L, \delta_R)\), which are functions of the relative lateral wheel–rail displacement. Industrial practice is to linearize the system of equations about the central portion of the running surface and create a function known as conicity, often denoted \(\lambda\). This assumes point contact and can be represented by four relationships

\[
\begin{align*}
\frac{1}{2}(r_L - r_R) &= \lambda y, \\
\frac{1}{2}(r_L + r_R) &= r_0 \\
\frac{1}{2}(\delta_L - \delta_R) &= 0, \\
\frac{1}{2}(\delta_L + \delta_R) &= \lambda
\end{align*}
\] (1)

If substituted into the non-linear wheelset model, full linear equations can be generated, equations (16) and (17) from reference [25]. Further simplifications to this model can be made by ignoring the smaller creep force terms, equation (18) from reference [26]. Parameters and states are given for all of the models in Appendix 1.

It is assumed that with all of the models there are ideal sensor sets present, and this includes measurement of the wheelset lateral, yaw and roll accelerations, and the lateral acceleration of the suspended...
mass. Further, it is assumed that the lateral rail irregularity and the gauge width variation are known for use in the system identification methods.

### 4.3.2 Conicity estimation through Kalman filtering

In the initial stages of the work, a Kalman filter [27] was used to estimate a generic smooth continuous conicity parameter applied to linear equations (16) and (17). The state of the Kalman filter was augmented to include the conicity parameter so that it could be estimated, therefore making the problem non-linear and an extended Kalman filter was used. However, this simplistic approach failed to converge [19].

The Kalman filter was augmented further by estimating the unknown track disturbance ($d$). Figure 11 shows the real and estimated values of the disturbance and conicity parameters. The estimates in simulation are acceptable, but uncertainty increases at lower conicity values. The estimation process can be improved further by adding another equation which provides additional dynamic information about how the conicity function varies with lateral displacement. This is added into the state equations in the form

$$\dot{\lambda} = \frac{d\lambda}{dy}$$  \hspace{1cm} (2)

$$\dot{d} = 0$$  \hspace{1cm} (3)

This process stores a lookup table of corresponding conicity relative to the wheel–rail position. The table is then used to give the Kalman filter some knowledge of the variability of the conicity. Figure 12 shows the results of this process. Around small displacements there is a very good fit, because a cluster of information is available but at larger displacements, where there may be fewer data points available, the fit fails.

### 4.3.3 Conicity estimation by system identification

An alternative to using the Kalman filter as a parameter estimator is to use system identification. The disadvantage with this approach is that a detailed knowledge of the input to the system from the track disturbance is required and this might not be measurable in practice.
If the system can be modelled as a ‘grey box’ regression model defined as

\[ \hat{y}(i) = X^T(i)\theta + \omega(i) \tag{4} \]

where \( \hat{y} \) is the estimated output variable, \( X \) is the vector of regressors, \( \theta \) is the vector of unknown parameters, \( \omega \) are the combined known parameter and regressor terms, and \( i \) is the discrete sample number, then the parameter estimate can be obtained using least squares estimation [28–30]. Non-linear terms can be added into the regressor matrix, such as higher order terms [31] or multiple PCP functions [32]. The PCP technique is a multi-section smoothing function that enables complex non-linear shapes to be approximated.

Figure 13 shows the results of using the PCP technique to estimate conicity as a function of relative wheel–rail position applied to the model of equation (18) [23]. There is a very close fit to the non-linear shape of the conicity function. It should be noted that the conicity may in practice be discontinuous and therefore not so well matched when using the smooth PCP technique.

4.3.4 Contact geometry estimation by system identification

When the identification technique of the previous section was applied to the non-linear model of equations (11) and (12), the conicity estimates were of poor quality, due to the estimation model being insufficient to fit to the complex dynamics of the simulation model. An alternative approach is to estimate the rolling radii and contact angles directly.

The unknown non-linear parameters present in the system equations are the four geometric combinations \( (r_L + r_R), (\delta_L + \delta_R), \) and \( (\delta_L - \delta_R) \). P8 wheel profiles and 113A railhead shapes in various states of wear were used for the study. Due to the complex discontinuous non-linear nature of these combinations, a piecewise linear approach was adopted rather than attempting to fit parameters across the entire range of lateral displacement. To achieve this, the collected dynamic data are first separated into a number of discrete sections that represent a restricted range of the relative wheel–rail displacement. Individual identifications are performed on each of these sections.

It was also appreciated that the input from the track disturbance should not be modelled as an idealized source [33], as this would have different frequency content to that found in a real system. Consideration was also given to the gauge width variation as this adds a degree of uncertainty into the model parameters. Input data used were from the Paddington to Bristol line with lateral disturbance and gauge width variation sampled every 0.2 m by a track recording vehicle. The standard deviation of the track disturbance and gauge width variation are of similar magnitudes at approximately 2 mm. Figure 14 shows \( (r_L + r_R) \) generated during simulation, demonstrating that the parameters being identified are distributed over a region rather than being a single-valued function of lateral displacement. This is because the gauge width variation adds uncertainty to the relationship. The gauge width variation also excites the system dynamically, because of the asymmetry of the wheelset tread shapes used, though this is a secondary effect.

In practice, for real-time applications, the least squares algorithm from the previous section can be run recursively in a similar manner to a Kalman filter. The method potentially saves computation expense, because the calculation is not dealing with the entire dataset for each iteration, just the latest data. This makes it feasible for applications where the processing power may be limited.

---

**Fig. 13** Least squares estimated conicity function with PCP function

**Fig. 14** Uncertainty in the parameters due to gauge width variation in simulation
Two ‘grey box’ multiple input single output identifications are performed, where for the lateral acceleration equation, from equation (11)

\[
\Omega_1 = \left[ \left( \frac{-2f_{22}}{mv} \right) \dot{y} + \left( \frac{k_y}{m} \right) y + \left( \frac{f_y}{m} \right) \dot{y}_m \right] + \left( \frac{k_y}{m} \right) y_m + \left( \frac{2f_{23}}{mv} \right) \psi + \left( \frac{W}{m} \right) \phi \right]
\]

(5)

\[ X_1 = [\psi, \dot{\psi}, 1] \]  

(6)

and for the yaw acceleration equation, from equation (12)

\[
\Omega_2 = \left[ \left( \frac{-2f_{23}}{I_v} \right) \dot{y} + \left( -\frac{2f_{21}f_{11}}{I_v} - \frac{2f_{33}}{I_v} - f_y \right) \psi \right]
\]

(7)

\[ X_2 = [\psi, \dot{\psi}, 1] \]  

(8)

Figure 15 shows the estimates when the gauge width variation is fixed at zero for \((r_L + r_R)\) and indicates good parameter convergence. Figure 16 shows the estimates with gauge width variation present. There is some spreading of the estimates because of the gauge width variation. Although not shown, the estimates of \((\delta_L + \delta_R)\) fail to converge. The failure to converge may be due to this parameter being most closely related to the variation in gauge width, which is no longer zero.

4.4 Summary and future work

This section covered a number of methods for the real-time estimation of critical components associated with the bogie. The first concept was a particle filter method for the estimation of suspension components. This showed that a reduced sensor set, on the bogie and the body, could be used to estimate damper coefficients, but was unable to estimate the effective conicity. This was demonstrated with real data collected from a Class 175 train. The next step is to apply this technique in real time to a suspension parameter estimation problem.

The second concept was a method for estimating wheel–rail adhesion using a Kalman filter. Simulations showed that, in principle, very low-adhesion conditions can be detected. This technique will be applied to a real system in the future.

The final concept was estimation of the wheel–rail profile. Kalman filtering and non-linear identification were applied to the conicity estimation problem. A piecewise-linear identification was also applied to the direct estimation of the contact geometry. All the techniques demonstrated the potential of the concept, with the main disadvantage being that the disturbance input from the rail is required, and that this may be difficult to measure. Possible alternative solutions for this are a combined state/parameter estimation loop to first estimate the disturbance, or the use of frequency analysis of output signals along a known section of track, both of which are currently being investigated.

5 VEHICLE SPEED MONITORING

The speed of a railway vehicle is normally measured through the use of inductive sensors detecting the teeth of a gearbox or slots of a slotted wheel [34]. However, significant problems arise when the conditions cannot be satisfied. When wheel slip/slide occurs (e.g. due to excessive tractive effort and/or extremely low adhesion, the measurement becomes unreliable as large errors may be introduced from the wheel slip/slide regardless how accurate the measurement of the wheel speed is due to the much

Fig. 15 Recursive least squares estimate of the rolling radii sums with no gauge width variation

Fig. 16 Recursive least squares estimate of the rolling radii sums with gauge width variation
increased speed difference between the axle and the vehicle). The use of unbraked and unpowered axle can solve the problem, but is not always desirable as the reduction on the number of axles for traction and/or braking may compromise the train control in low-adhesion conditions. Also, the tachometry-based technique requires regular re-calibration as the wheel diameter is reduced due to wear. It is also worth noting that the new technique could be used in conjunction with tachometry to avoid the need for re-calibration.

Although a number of new methods have been proposed, most noticeably those based on spatial measurement/filtering using sensing technologies such as Doppler radar [35], Eddy current sensors [36], and image processing [37], those have not been applied for traction and braking control applications because the measurement accuracy and reliability achievable are not considered sufficient.

5.1 Speed measurement using bogie-based inertial sensors

The idea for monitoring vehicle speed included in this article was first proposed in 2008 [38, 39] and was followed by more detailed studies that tackle specific technical issues [40, 41]. The new concept was conceived from the observation that the absolute vehicle speed determines the time delay in bounce motion between any two wheelsets as all railway wheelsets travel on and pass any point of a track one after another. The time delay in the wheelset motions in response to the track irregularities if/when detected can therefore be used to determine the vehicle speed. The measurement can be highly accurate if the speed does not change rapidly and/or the time required to detect the delays is sufficiently small not to introduce significant errors in rapid acceleration or deceleration.

A measurement scheme that takes into consideration the practical requirements of the rail industry was proposed as illustrated in Fig. 17. The need for the installation of any sensors on the wheelsets/axle boxes, which presents a harsh working environment for sensors, is removed. This is replaced by using inertial sensors mounted on a bogie frame, which is preferred because the reliability requirement and also the cost of the sensors will be much lower compared with axle-mounted ones. An intelligent data processing method is then developed to derive (or estimate) the wheelset bounce motions from the measured bogie bounce and pitch vibrations. The data processing is made as simple as possible to reduce potential difficulties in practical implementation, but sufficiently effective to obtain the estimation of the wheelset motions suitable for detecting time delays between them [38]. The detection of the time shift is achieved by computing the cross-correlations of the two filtered signals and detecting the time shift at the peak correlation value. It is then straightforward to calculate the vehicle speed \( V_m \) from the detected time delay \( T_{\text{delay}} \) and semi wheel space \( L_b \) using equation (9)

\[
V_m = \frac{2L_b}{T_{\text{delay}}}
\]  

(9)

5.2 Measurement performance

As the proposed technique is independent from the wheel rotational speed of the wheels/wheelsets, the measurement is not affected by wheel slip/slide and it does not require knowledge of wheel radius. There is also a clear advantage of the low sensing requirement. The use of inertial sensors on the bogies can be expected to become standard installations in future railway vehicles, not least as required for tilting control systems, so that there is little extra cost involved for practical implementation of the technique. The measurement technique is also tolerant to sensor errors (e.g. the effect of sensor noise tends to be filtered out by the cross-correlation computations as the noise components are uncorrelated between different signals [40]).

A moving window of sampled data for cross-correlation calculations as illustrated in equation (2) will enable the continuous detection and update of the train speed and is used in the study for performance assessment. The number of shifted intervals \( m \) is incremented from \(-N\) to \(N\), where \(N\) is the total number of sampled data used for each time window \( T_{\text{wdw}} = N \cdot T_s \) of the running cross-correlation

\[
R_{xy}(m) = \sum_{i=1}^{N-m} x(i+m)y(i)
\]  

(10)
In each computation step (for each time window), the detection of the peak point of cross-correlations and the corresponding delay intervals will be carried out which is in turn used to calculate the speed \( (V_m) \). Repeating the computation steps, as the time window moves, provides a regular and fast update of the measured speed signal at the rate of the sampling interval \( (T_s) \).

There are two factors that may potentially affect the accuracy of the scheme. One is the truncation error introduced due to the finite sampling rate for collecting/processing data, and the other is due to a delay in the speed detection during vehicle acceleration or deceleration. However, the use of reasonably selected sampling rate and data length for the new measurement technique will be able to deliver a high level of performance as illustrated in the two examples given below.

Figure 18 shows the estimated speed of the proposed measurement method, where a comprehensive vehicle model for a conventional bogie vehicle in a multi-body simulation package (Vampire) environment is run at the two different speeds of 50 km/h and 100 km/h on a track profile that includes track irregularities superimposed on a low-speed curve consisting of a constant curve section of 400 m in curve radius and 6° in cant angle connected to a straight track section via a transition at either end. Curved tracks present one of the most difficult conditions for the proposed measurement scheme, as there will be variations in the wheel space (increased between the outer wheels and decreased for the inner wheels).

The measurement for the vehicle speed of 50 km/h varies within a small range of the real speed and the maximum measurement error is 0.22 km/h. The measurement for the speed of 100 km/h gives also a small error range with a maximum error of 0.65 km/h. The measurement errors here are largely due to the truncation caused by the discrete time processing. In this case, a sampling interval of 1 ms is used. For the wheel space of 2.6 m for the vehicle used in this study, the truncation errors are expected to be within \([-0.22, +0.05\text{ km/h}]\) at the vehicle speed 50 km/h and \([-0.57, +0.65\text{ km/h}]\) at the speed of 100 km/h, which are confirmed by the simulation results. It is possible to reduce the error further by using a smaller sampling interval, but this would have to be balanced with the increased computational demand [41].

Figure 19 shows the second type of error in the form of a measurement delay when the vehicle accelerates (from 40 to 48 m/s, or from 144 to 172.8 km/h) at the rate of 1 m/s\(^2\), where the truncation error (in a random manner) is superimposed with a steady state error (or offset) due to the delay. This is because the cross-correlation technique detects the average time shift between the two signals for a given time window and therefore is likely to cause a measurement error during acceleration or deceleration. This type of error tends to be more significant at low speeds, as longer time windows are required to provide an accurate measurement of the time shift [38].

5.3 Summary and future work

Reliable and accurate measurement of the vehicle ground speed even in adverse conditions such as wheel slip/slide does not have to be achieved through the expensive equipment and/or complex systems. It is possible to provide an effective solution with an innovative use of inertial sensors as demonstrated in this study, even though this type of sensors is not normally associated with speed measurement. The performance of the new measurement method can be substantially better than the requirement specified in the UIC standard for wheel slide control (UIC-5014-05) [40]. Possible applications for the proposed measurement solution include.

1. Replacement or supplement of the conventional axle-based speed sensors to provide more accurate
measurement for traction/braking control systems in all operation conditions including wheel slip or slide.

2. A tool for accurate calibration of the axle-based sensors, the accuracy of which is affected by the changes in wheel contact radius due to wear/re-profiling, etc.

3. It can be particularly useful in conditions such as tunnels and underground where the use of other devices (e.g. GPS) may be problematic.

6 CONCLUSIONS–TRENDS, OPPORTUNITIES, AND RESEARCH CHALLENGES

The article has illustrated potential monitoring opportunities arising from advanced processing of vehicle-mounted sensors: two applications relate to condition monitoring of both track and vehicles for enhanced maintenance and/or improved system reliability, and a third provides operational information for use on board the train in real time.

The emphasis has very much been upon developing the processing concepts and the associated algorithms, but it is of course necessary to convert the concepts into practical engineering solutions, principally involving (i) identification of minimized sensor configurations for lower cost and (ii) high-reliability solutions to ensure consistently verifiable information. The second point is particularly important, because without good-quality, reliable information the system will be difficult to validate where system integrity is involved, and may also be discredited where it is used to inform maintenance processes.

As the title implies, the article has strongly focused upon the on-train sensing and processing aspects, but one of the key issues that must be addressed prior to the deployment of practical condition monitoring systems is the management of data once it leaves the vehicle, in particular because the collection of in-service data will inevitably lead to the need to retain large quantities of data. Data need to be retained and processed in a number of different ways, ideally together with the context of how, when, and where they were collected. Initially it is important that the data are verified to ensure that they are correct and no sensor errors are present (off-set faults, noise, null values, communication errors, etc.). Secondly, straightforward robust algorithms need to be used to ensure that critical faults are identified in close to real-time. Thirdly, the data need to be stored for post-processing (the main focus of this article) to identify longer-term incipient faults. As confidence grows in the accuracy of condition monitoring algorithms it will not be necessary to retain data for long periods of time. As more data are collected from both infrastructure and vehicle-borne sensors, the railway industry needs to address the need for standards for data collection, such standards need to address the relationship between data collection by various means.

It is inevitable that condition monitoring technology will increasingly take advantage of new processing techniques as they emerge, both to extract higher-integrity information from existing sets of sensors and to provide lower-cost solutions using simpler, more robust sensor sets. It is therefore important to maintain a ‘watching brief’ on such theoretical developments, the challenge being to identify these possibilities, research their application within railway systems, and critically evaluate their prospective contributions from a business viewpoint.

ACKNOWLEDGEMENTS

The authors would like to express thanks to the following organisations that made this work possible: Rail Research United Kingdom (RRUK), Engineering and Physical Science Research Council (EPSRC), Tyne & Wear Metro, Alstom, Merseyrail, Department for Transport (DfT) and the Rail Technology Unit (RTU) at Manchester Metropolitan University (MMU).

© Authors 2010

REFERENCES


2 Foeillet, G. IRIS 320 is a global concept inspection vehicle merging engineering and R&D tools for infrastructure maintenance. In Proceedings of the Eight World Congress on Railway research, Seoul, South Korea, 18–22 May 2008.


7 Ackroyd, P., Angelo, S., Nejikovsky, B., and Stevens, J. Remote ride quality monitoring of Acela train set


30 Aström, K. J. Adaptive control, 1989 (Addison-Wesley).


APPENDIX 2

A2.1 Monitoring dynamic performance characteristics equations

Wheelset lateral dynamics

\[ m \ddot{y} + \frac{2f_{23}}{V} \left( \dot{y} + n_c \frac{r_c}{r_0} \dot{\psi} - V \ddot{\psi} \right) + 2f_{11} \left( 1 - \frac{r_c}{r_0} \right) \psi + 2f_{23} \left( \frac{\psi}{V} - \frac{\delta_1 - \delta_R}{2r_0} \right) + W \left[ \frac{\delta_1 - \delta_R}{2} + \psi \right] = F_s \] (11)

Wheelset yaw dynamics

\[ I \dddot{\psi} + \frac{I_{ay}}{r_0} \dot{\psi} + \frac{2f_{11}}{r_0} \left( \frac{r_c}{r_0} - \frac{r_l}{r_0} \right) - \frac{2f_{23}}{V} \dot{\psi} = -2f_{23} \left( r_c + r_l \right) \dot{\psi} + 2f_{23} + 2f_{11} \frac{F_s}{V} - f_{33} \left( \delta_1 - \delta_R \right) r_0 \] (12)

Suspended mass lateral dynamics

\[ \ddot{y} = \frac{1}{m_m} (F_s) \] (13)

A2.2 Full linear equations

Wheelset lateral dynamics

\[ m \ddot{y} + \frac{2f_{23}}{V} \left( \dot{y} + n_c \frac{\lambda \dot{y}}{T} - V \ddot{\psi} \right) + 2f_{23} \frac{\dot{\psi}}{V} + W \lambda \ddot{\psi} = F_s \] (16)

Wheelset yaw dynamics

\[ I \dddot{\psi} + \frac{I_{ay} V \lambda \dot{\psi}}{r_0} + \frac{2f_{11} \lambda \dot{\psi}}{r_0} = 2f_{23} \frac{\dot{\psi}}{V} \left( \dot{y} + \frac{r_c \dot{y}}{T} - V \ddot{\psi} \right) + 2f_{11} \frac{F_s}{V} \dot{\psi} = M_s \] (17)

A2.3 Simplified linear equation

Wheelset yaw dynamics

\[ I \dddot{\psi} + \frac{2f_{11} \lambda \dot{\psi}}{r_0} + \frac{2f_{11} \dot{\psi}}{V} = M_s \] (18)