Adhesion estimation at the wheel–rail interface using advanced model-based filtering

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


Additional Information:


Metadata Record: [https://dspace.lboro.ac.uk/2134/10121](https://dspace.lboro.ac.uk/2134/10121)

Version: Accepted for publication

Publisher: © Taylor and Francis

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/
RESEARCH ARTICLE

Adhesion estimation at the wheel-rail interface using advanced model based filtering

C.P. Ward\textsuperscript{a*}, R.M. Goodall\textsuperscript{a}, R. Dixon\textsuperscript{a} and G.A. Charles\textsuperscript{b}

\textsuperscript{a}Control Systems Group, School of Electronic, Electrical and Systems Engineering, Loughborough University, Leicestershire, LE11 3TU, UK; \textsuperscript{b}Faculty of Engineering, Coates Building, University of Nottingham, Nottingham, NG7 2RD, UK

(January 2012)

The railway industry in the UK is currently expanding the use of condition monitoring of railway vehicles. These systems can be used to improve maintenance procedures or could potentially be used to monitor current vehicle running conditions without the use of cost prohibitive sensors. This paper looks at a novel method for the online detection of areas of low adhesion in the wheel/rail contact that cause significant disruption to the running of a network, particularly in the autumn season. The proposed method uses a Kalman-Bucy filter to estimate the creep forces in the wheel-rail contact area; post-processing is then applied to provide information indicative of the actual adhesion level. The algorithm uses data that, in practice, would be available from a set of modest cost inertial sensors mounted on the vehicle bogie and wheel-sets. The efficacy of the approach is demonstrated using simulation data from a non-linear dynamic model of the vehicle and its track interface.

Keywords: Condition monitoring, wheel-rail contact, low adhesion, advanced filtering, real-time estimation, fault diagnosis/detection

1. Introduction

The railway industry in the United Kingdom has seen something of a renaissance in the past 20 years with annual increases in passenger numbers \cite{1} and tonnage of freight hauled, putting demands on rolling stock to be available for a greater proportion of the time. This results in a need to optimise service timetables along with vehicle maintenance and is therefore one of the key drivers for the application of condition based maintenance and the utilisation of optimised condition based timetabling.

Condition monitoring systems to detect faults and estimate running conditions (on-board railway vehicles) in real time are vital to enable the transition to this smarter way of running the railways. The industry is starting the transition to this philosophy with systems such as the Bombardier ORBITA, \cite{2}, that provides real time information about the running condition of a fleet of rail vehicles, using similar principles to those applied in the aerospace industry. In academia many vehicle-based condition monitoring schemes have also been investigated, such as: monitoring of suspension components, \cite{3}; wheel-rail profile estimation, \cite{4}; vehicle speed estimation, \cite{5} and creep force detection, \cite{6}.

This paper proposes a novel approach to the detection of areas of low adhesion in the wheel/rail interface that are created by such things as: leaf contaminant;
dew formation; lubricant dispersal; ice; and many other reasons, [7]. These areas can create large cost and punctuality issues for train operators and users alike. They are often ephemeral and manifest in timetables needing to be rescheduled for reduced station stop and signal approach speeds, especially in autumn when the problem is at its height. This mitigation can have the effect of reducing capacity on already crowded networks.

This conceptually novel technique relies upon the detection of lateral and yaw creep forces in the contact patch using advanced filter as the rail vehicle is in normal running, i.e. no braking or traction force is applied. This means that a picture of the entire network can be collated as the data is generated continuously and a potentially dangerous wheel slip/slide event does not have to be triggered in order to detect low adhesion, as is the case with many current adhesion detection systems that log activity of wheel slip/slide protection systems.

The proposed new approach will therefore provide real time knowledge (derived from inertial measurement data onboard a rail vehicle and advanced model based filtering) of the current adhesion conditions across the network that can be utilised in many ways: as a means of alerting drivers to changes in adhesion; for targeted deployment of mitigation efforts such as railhead cleaning; rescheduling of timetables to account for extended vehicle stopping distances; and many more besides. A number of ideas have been proposed for this application such as: multiple Kalman filters to estimate creep coefficients, [8]; inverse modelling for the estimation of creep forces, [9]; and as first proposed in [6] and further developed in [10, 11], a Kalman-Bucy filter (KBF) estimation of creep forces and interpretation through post-processing.

The highlighted approach is therefore a two step algorithm: the first step is estimation of creep forces in the contact area through inertial sensor measurements (simulated in this example) and model based KBF; the second step uses parameter estimation to process the creep force signals into usable adhesion level information. Creep forces provide the guidance mechanism of the wheelset system [12] and are dependent upon adhesion conditions, therefore if adhesion reduces so does the creep force. However these forces are a function of many factors such as: vehicle speed; track irregularity size; and wheel-rail profile. They can not be used directly and two methods of post processing using parameter estimation with varying levels of additional track irregularity information are shown.

This paper covers the simulation modelling and advanced processing required to implement the method: section 2 covers the nonlinear simulation modelling that integrates an advanced method of creep force generation and is used to generate all of the appropriate test signals; section 3 covers the high level algorithm concept for adhesion estimation; section 4 explores the creep force estimation technique through KBF and generation of appropriate filter design models and results; section 5 covers the parameter estimation methods of creep force post processing, along with results.

2. Simulation modelling

The aim of this process is to determine the creep forces present in the wheel/rail contact as a rail vehicle is operating in normal traffic. Hence two distinct models are required: a simulation model (covered in this section) that contains all of the significant nonlinearities associated with the system dynamics and is used to generate all of the test measurements; and a estimator design model (covered in section 4) that is a simplified linearised version of the simulation model and is used for processing purposes.
As in previous studies the simulation model used here is considered only in a lateral and yaw sense as the vertical and longitudinal effects are substantially uncoupled from lateral/yaw effects and can be neglected [12]. Another key concept of the process is to determine the adhesion conditions before any braking or accelerative torque is applied therefore not making the longitudinal dynamics signal necessary. The vehicle type modelled is a British Mk.3 coach and is created as a full length vehicle, with two bogies, two wheelsets per bogie, all with lateral and yaw degrees of freedom. It can be separated into two interacting sections: the wheel/rail contact where the complex non-linear contact mechanics occur; and the linear Newtonian mechanics of the specific vehicle’s suspension systems and associated geometries.

2.1. Creep force modelling

Fundamentally, creep forces provide the guidance mechanism for wheelsets. These forces are generated in reaction to the creeps (or slips) in the rolling contact of the wheel-rail interface in normal running. These are relative velocities of the wheel and the rail in the contact area and are defined as

\[ s_i = \frac{w_i}{V}, \quad i = x, y \]  

where \( V \) is the forward velocity of the wheelset, \( w_i \) is the creep (slip) velocity in the relevant direction (where \( x \) is longitudinal direction and \( y \) is lateral direction), where this is defined as

\[ w_i = V_w - V_r, \quad i = x, y \]  

where \( V_w \) is the velocity of the wheel through the contact patch, and \( V_r \) is the velocity of the rail through the contact patch. Creep generation is a highly nonlinear process, but normal practice for wheel-rail contact modelling is to linearise the creep forces generated in the model based upon Kalker coefficients, [13]. Due to the importance here of modelling the non-linear adhesion characteristics up to and beyond the creep saturation, use is made of the more advanced contact force model developed in [14]. This model is essentially a practical curve fitting mechanism that provides close correlation to collected testing data. The total creep force in the contact area (excluding spin effects due to their small contribution to the overall wheelset dynamics in the straight line running and high speed curving of this example. It is noted that flange contact forces will become more important under tighter curving regimes and through switch and crossings, though this is a point that will be addressed in later studies) is calculated as

\[ F = \frac{2Q\mu}{\pi} \left( \frac{\epsilon}{1 + \epsilon^2} + \arctan \epsilon \right) \]  

where \( Q \) is the wheel load, with

\[ \epsilon = \frac{2}{3} \frac{C\pi a^2 b}{Q\mu} s \]
where $C$ is the proportionality coefficient of the contact shear stiffness (N/m$^3$). Kalker coefficients can be used for this purpose, where for the longitudinal direction

$$\epsilon_x = \frac{1}{4} G\pi abc_{11} s_x$$

(5)

where $s_x$ is the longitudinal creep. With $s_y$ as the lateral component of the total slip $s$ is

$$s = \sqrt{s_x^2 + s_y^2}$$

(6)

The forces $F_x, F_y$ in the longitudinal and lateral directions are

$$F_i = F_i \frac{s_i}{s}, \quad i = x, y$$

(7)

and the adhesion coefficients

$$f_i = \frac{F_i}{Q}, \quad i = x, y$$

(8)

The friction coefficients rely upon the slip velocity, where

$$\mu = \mu_0 \left[ (1 - A)e^{-Bw} + A \right]$$

(9)

$A$ is the ratio of limit friction coefficient at infinity slip velocity $\mu_{\infty}$ to the maximum friction coefficient $\mu_0$

$$A = \frac{\mu_{\infty}}{\mu_0}$$

(10)

For large creep applications the force is calculated using reduction factors, $k_A$ in the area of adhesion and $k_S$ in the area of slip, as

$$F = \frac{2Q\mu}{\pi} \left( \frac{k_A \epsilon}{1 + (k_A \epsilon)^2} + \arctan(k_S \epsilon) \right), \quad k_S \leq k_A \leq 1$$

(11)

where, the gradient of the area of adhesion to the area of slip corresponds to a reduction in the Kalker coefficient, as

$$k = \frac{k_A + k_S}{2}$$

(12)

Therefore allowing for a change in initial gradient with a change in adhesion. Experimentation has shown that, contrary to expectation from theoretical models such as that of Kalker [13], the initial slope of the creep curve varies with different adhesion levels, [7, 15]. Four levels of adhesion are defined in this study as ‘dry’, ‘wet’, ‘low’ and ‘very low’ conditions, relating to 55%g, 30%, 6%g and 3%g potential deceleration rates. Neither ‘dry’ nor ‘wet’ conditions are generally a problem but the ‘low’ and ‘very low’ conditions will affect braking performance and hence cause operational problems with rail vehicles. The accompanying constants are given in Table 1 and the creep curves are given in Figure 1 for the wheel loadings of the particular vehicle type. This varying slope means that different adhesion levels can be detected without the creep forces becoming saturated. The effect of varying the
adhesion levels on the running system is shown in Figure 2. This shows the sum

of the lateral creep moments and gravitational stiffnesses moment for the front bogies front wheelset for the same system disturbance (i.e. the lateral position of the track). For ‘dry’ and ‘wet’ conditions the forces are visually identical. This is because the track excitation is insufficient to reach saturation of the creep forces, and also the initial slope of the characteristic is unchanged. However for ‘low’ and ‘very low’ conditions the creep forces generated reduces substantially, confirming that detection of changes of adhesion level is feasible.

The PSD of each of the cases is shown in Figure 3. This repeats the observations of the time plots in that the amplitude of the key frequencies from 0 to 4 Hz reduces as the adhesion condition worsens.
2.2. System dynamics modelling

Previous studies [6] modelled the system as a half vehicle body constrained in yaw, with one bogie and two wheelsets. The model is extended as in [10] to a full vehicle body, with two bogies and four wheelsets to represent a British Mk.3 coach, schematics of the primary and secondary suspensions are shown in Figures 4 and 5 respectively, with corresponding physical parameters shown in Table 2. These equations encompass the lateral and yaw dynamics of the wheelsets, the bogies and
the vehicle body. Dynamic equations adapted for the Newtonian vehicle dynamics from [8] are given in Appendix A.

The tests are performed at a vehicle speed of 20 m/s, on straight track, with track irregularity taken from a recording car on the British Great Western mainline. The wheel profile used is a new S1002, with the corresponding rail profile being a new UIC60.

### 3. Adhesion estimation method

The basic concept of analysing the creep force estimates to produce a meaningful understanding of adhesion level is shown in the schematic of Figure 6 that was first proposed as a potential method in [16]. The basic algorithm flow is: inertial vehicle based measurements and their integrals are used in the estimation of the creep forces via advanced model based filtering (it should be noted that all signals here are obtained through simulation, but that they are representative of signals which could be measured in practice); the estimates are then post processed by parameter identification; the output is then interpreted as a level of adhesion.

Covered in section 4 is the method of estimating creep forces in the wheel rail
contact, and section 5 covers two techniques for interpreting the creep force levels by parameter estimation. This paper considers two processing scenarios, the first without knowledge of the lateral track irregularity which would results in a more straightforward implementation but makes the estimation problem more difficult, the second with this knowledge.

4. Creep force estimation technique

The well known KBF [17] is used to estimate the creep forces, combined with the gravitational stiffness, therefore requiring a simplified design model equivalent of the simulation model. Previous study [6] demonstrated that a KBF cannot distinguish between the creep forces and the gravitational stiffness so these need to be combined in the design model. The design model uses simplified versions of equations A1 to A8, where the complex nonlinearities of the contact mechanics are combined into a single state, these are now

\[ m_{FF}\ddot{y}_{FF} = F_{FF} + F_{syFF} \]  
\[ I_{FF}\ddot{\psi}_{FF} = M_{FF} + M_{s\psi FF} \]  
\[ m_{FR}\ddot{y}_{FR} = F_{FR} + F_{syFR} \]  
\[ I_{FR}\ddot{\psi}_{FR} = M_{FR} + M_{s\psi FR} \]  
\[ m_{RF}\ddot{y}_{RF} = F_{RF} + F_{syRF} \]  
\[ I_{RF}\ddot{\psi}_{RF} = M_{RF} + M_{s\psi RF} \]  
\[ m_{RR}\ddot{y}_{RR} = F_{RR} + F_{syRR} \]  
\[ I_{RR}\ddot{\psi}_{RR} = M_{RR} + M_{s\psi RR} \]

where for the purposes of the filter combining the creep forces and gravitational stiffnesses, the following assumptions are made

\[ \dot{F}_{FF} = \dot{F}_{FR} = \dot{F}_{RF} = \dot{F}_{RR} = 0 \]  
\[ \dot{M}_{FF} = \dot{M}_{FR} = \dot{M}_{RF} = \dot{M}_{RR} = 0 \]

These relationships mean that there is no dynamic understanding of the creep forces and gravitational stiffnesses inherent in the estimator models. The state vector is augmented with non-dynamic force states that are used to estimate the size of the ‘force deficit’ left by the estimator model. This method therefore relies
upon a use of a model of the suspension, which correctly represents the dominant
dynamics (forces), as the foundation of the estimator design.

The KBF is based upon state space methods, where the design model state
equation is defined as

\[ \dot{x} = A_k x + B_k u + z \]  (23)

where \( x \) is the state vector, \( \dot{x} \) is the rate of change of the state vector, \( z \) is the
Gaussian noise source on each of the state vectors, \( A_k \) is the state matrix and \( B_k \)
is the input matrix. The output equation of the design model is defined as

\[ y = C_k x + D_k u + v \]  (24)

where \( y \) is the output vector, \( v \) is the Gaussian noise on the output vector, \( C_k \) is
the output matrix and \( D_k \) is the input matrix.

The filter algorithm can be separated into two sections. The first section calcu-
lates how much to adapt the filter to changes in the system being measured, this
can either be calculated iterative in the loop, using

\[ K = PC_k^T R^{-1} \]  (25)

\[ \dot{P} = A_k P + P A_k^T - K R K^T + Q \]  (26)

where \( K \) is the ‘Kalman gain’, \( P \) is the error covariance. The second option is to
calculate \( P \) offline by setting \( \dot{P} = 0 \), in this case equation 26 is now in the form of
a Riccati equation and one of the well known solvers can be used.

The ‘Kalman gain’ is then used in the second section of the filter to update the
estimates. The estimated state and output are then calculated simultaneously as

\[ \hat{y} = C_k \hat{x} + D_k u \]  (27)

\[ \hat{x} = A_k \hat{x} + B_k u + K (y - \hat{y}) \]  (28)

where \( \hat{y} \) is the estimated output and \( \hat{x} \) is the estimated state.

Design choices are made by selecting covariance matrices of the state \( Q \) and the
output \( R \). These define the level of noise that is present in the states and mea-
surements respectively, and represent a tradeoff between confidence in the design
model and confidence in the signals measured. As previously mention it should be
noted that the signals here are from the simulation model and represent signals
that are likely to be measured in practice.

In this system, the design model is chosen such that the system input (track
irregularity) is not included due to this term only occurring in the estimated creep
force state. The filter then becomes output only as it uses measurements from the
rail vehicle alone, meaning \( B_k = D_k = 0 \). Further simulation studies have looked
into the sensing requirements for the method, and conclusions are that: only half
the vehicle needs to be instrumented; no measurements are required on the vehicle
body; and measurements are required either side of the primary suspension in the
axle-box and on the bogie. This latter point is due to the level of filtration through
the primary suspension, meaning that if measurements (in this case simulated) are
taken only from the bogie, key components of the dynamics of the wheelset will be
lost meaning poor estimation of the creep forces. The example outputs shown in Section 4.1 are for an estimation model applied to signals generated from the front half of the simulation model, with a full measurement set for the two wheelsets, the bogie and the lateral dynamics of the vehicle body. The state vector for the estimation model is therefore defined as

$$x = [y_{FF} \dot{y}_{FF} \psi_{FF} \ddot{y}_{FF} y_{FR} \dot{y}_{FR} \psi_{FR} y_{BF} \dot{y}_{BF} \psi_{BF} y_{V} \dot{y}_{V} \cdots$$

$$\cdots F_{FF} F_{FR} M_{FF} M_{FR}]^T$$

with the corresponding output vector as

$$y = [y_{FF} \dot{y}_{FF} \psi_{FF} \ddot{y}_{FF} y_{FR} \dot{y}_{FR} \psi_{FR} y_{BF} \dot{y}_{BF} \psi_{BF} y_{V} \dot{y}_{V}]^T$$

The primary tuning parameters here are the $Q$ and $R$ matrices. As these cannot readily be known for real applications, they are assumed to be diagonal matrices. Although this suggests that the individual states and measurements are statistically independent, which may not be the case, it has the advantage of making the tuning task manageable using engineering judgement rather than needing a priori statistical knowledge that will not be available in practice. In this case, the state noise is selected as

$$Q = diag[1 1 1 1 1 1 1 1 1 1 1 1 e^9 1e^9 1e^9 1e^9]$$

The high values associated with the last four positions in the matrix assign uncertainty to the assumptions of equations 21 and 22, allowing the filter to adapt the state estimates to the creep force levels. As previously stated the $R$ matrix is again defined as a diagonal matrix as

$$R = diag[1e^{-8} 1e^{-8} 1e^{-8} 1e^{-8} 1e^{-8} 1e^{-8} 1e^{-8} 1e^{-8} 1e^{-8} 1e^{-8} 1e^{-8} 1e^{-8}]$$

due to the assumed low noise levels on the signals generated through the simulation model. It should be noted that tuning of these matrices is heuristic in nature and the gains associated will be varied for data gathered in later stages of simulation and experimental testing. Unknown noise characteristics of the measured signals and the model of the suspension system may not be as close to reality as in simulation due mainly to nonlinearity effects.

### 4.1. Creep force estimation results

Two key groups of tests are performed to test the efficacy of the creep force estimation technique

1. Static adhesion level tests at the four defined adhesion levels: ‘dry’; ‘wet’; ‘low’; ‘very low’
2. Step adhesion changes, from the ‘dry’ to the ‘very low’ condition part way through the simulation

The static level tests are performed to ensure that the designed KBF can work in a range of conditions that will be experienced in reality. The step tests ensure that the filter can adapt in real time to adhesion changes that will occur as the vehicle traverses different sections of track. Each of the tests is additionally performed at three different levels of lateral track irregularity, full scale, half scale and double scale. This is to ensures that the technique can function on multiple section of a rail


<table>
<thead>
<tr>
<th>Irregularity level</th>
<th>Max (m)</th>
<th>Min (m)</th>
<th>STD (m)</th>
<th>VAR (m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>0.0134</td>
<td>-0.0073</td>
<td>0.0015</td>
<td>2.29e−6</td>
</tr>
<tr>
<td>Half</td>
<td>0.0067</td>
<td>-0.0036</td>
<td>7.57e−4</td>
<td>0.57e−6</td>
</tr>
<tr>
<td>Double</td>
<td>0.0269</td>
<td>-0.0146</td>
<td>0.003</td>
<td>9.16e−6</td>
</tr>
</tbody>
</table>

**Figure 7. Creep yaw torque estimations, static adhesion tests**

The network with different track maintenance levels and that for the post processing the track irregularity changes are not misinterpreted as adhesion changes. The lateral track irregularity signal is taken from track recording car data obtained from the 200 km/h Great Western mainline in the UK and is scaled to obtain the half and double sized signals. Maximum/minimum values, standard deviations (STD) and the variance (VAR) of each of the signals is shown in Table 3.

#### 4.1.1. Static adhesion test

As mentioned the KBF gains will be static in implementation and will have to adapt to different adhesion characteristics present on the railhead, therefore a series of tests were undertaken at different static levels of adhesion. Figure 7 shows a section of estimated creep yaw torque data for the full scale track irregularity (half and double scale track irregularity tests are omitted here for clarity) for all four of the adhesion conditions.

Using visual inspection this shows that for all of the adhesion conditions the KBF produces a good estimation of the simulated combined creep and gravitational torques. A numeric assessment of estimation accuracy uses the well known coefficient of determination or $R^2$ [18], where

$$R^2 = 1 - \frac{\sigma^2(\epsilon)}{\sigma^2(y)} \quad (33)$$

and where $\sigma^2(\epsilon)$ is the variance of the residuals ($\epsilon = y_m - \hat{y}_k$, where $\hat{y}_k$ is the estimated output) and $\sigma^2(y_m)$ is the variance of the measured output. The coefficient of determination yields a number in the range $0 \leq R^2 \leq 1$ which tends to unity when the estimation matches the actual signal. It is often multiplied by 100 and expressed as a percentage.
Table 4 summarises the results of the tests analysed using the coefficient of determination at the three track irregularity sizes. The general trends are that for the ‘dry’, ‘wet’ and ‘low’ adhesion conditions at the three lateral track irregularity levels, the estimation level is reasonably consistent at around 90%. For all three irregularity levels, as the adhesion drops to the ‘very low’ condition the estimation confidence reduces, to around 85% for the full and double scale track irregularity levels and to around 70% for the half scale excitation. This drop may be due to the simulation model entering the saturation part of the creep curve for this condition, therefore the signal becomes more nonlinear and the linear KBF is not able to adapt as well to this condition. However it still tracks the reduction in the creep forces so will be usable for the detection of low adhesion.

4.1.2. Step adhesion change tests

As mentioned the KBF estimator is expected to adapt in real time to changes in the adhesion level, and Figure 8 shows how the KBF performs for a step change in the adhesion level. The test is performed for three varying levels of lateral track irregularity. Subplot (a) demonstrates the change in the adhesion level from dry to very low conditions at 15 seconds, subplots (b) to (d) show how the estimator adapts to this adhesion change for full scale track irregularity, half scale track irregularity and double scale track irregularity respectively. All three demonstrate by visual inspection that the filter can adapt well and quickly to a large change in adhesion condition of the simulation model, meaning that the linear filter is flexible enough to cope with what are large nonlinear variations in measured signals.

5. Creep force post processing using least squares parameter estimation

Creep forces alone cannot be used to assess the level of adhesion due to their dependence upon the lateral track irregularity size, wheel/rail profile and the vehicle speed. Therefore analysis of the estimated creep force and moment signals is required to create usable interpretations that are not dependent upon factors beyond adhesion changes.

Presented here are two techniques making use of parameter estimation through a well known least squares algorithm [18], that is summarised in Appendix B. The basic flow of the creep force post processing concept is shown in Figure 9 where: estimated creep forces and/or measured dynamics signals are windowed into suitable lengths of moving data in relation to the vehicles position, in this case simulated data (additionally stored and synchronised track irregularity data can be used if available); the least squares algorithm of Appendix B is used to estimate parameters from suitably selected model structures using the moving windowed data; the estimated parameters are interpreted into an ‘adhesion level’ for the position relating to the windowed data.

The first technique shown in subsection 5.1 is a black-box regressor parameter estimation and analyses creep force and moments signals without any additional knowledge of the system input (lateral track irregularity) meaning the analysis could be considered standalone if used in application due to no knowledge of difficult to measure states being required. The second technique in subsection 5.2
Figure 8. Creep force estimate of a step adhesion change: (a) adhesion level change; (b) full scale track irregularity; (c) half scale track irregularity; (d) double scale track irregularity.

Figure 9. Schematic of the generic creep force post processing algorithm.

demonstrates how the interpretation of low adhesion could be improved by incorporating recorded track irregularity levels into the estimation process by estimating parameters for a simplified linear version of a creep force model. This would require more input from end-users and therefore can no longer be called standalone. Each method is validated using the simulated example data generated in the previous section by the creep force estimation technique for the static and step change tests.

5.1. Output identification regressor analysis

The parameter estimation technique as proposed previously is utilised in this first creep force post-processing method to identify frequency based changes in the data from the creep force estimation algorithms. This is achieved by identifying creep force signals from regressor data of the estimated creep forces. The technique is summarised in Figure 10 where: moving window of the estimated creep force and moment data is collated; regressor parameter estimation of the windowed creep force data performed using the least squares algorithm; the estimated regressor model is assessed for eigenvalues which are time averaged over a further window.
to smooth the data; Eigen values changes are assessed either using logic or fuzzy logic to determine the adhesion level.

![Diagram](image)

Figure 10. Creep force post processing schematic, with no track irregularity information

This technique is therefore effectively standalone as it does not require any additional information about the lateral track irregularity disturbance or any input from an end user. The generic model formation for the parameter estimation technique is selected as

$$
\hat{F}_t = \theta_1 F_{t-1} + \theta_2 F_{t-2} + \cdots + \theta_n F_{t-n}
$$

(34)

where $\hat{F}_t$ is the creep force estimate through the identification process, $F_{t-n}$ is the time regressor at sample $n$ of the estimated creep force from the KBF algorithm, and $\theta_n$ is the estimated parameter for the corresponding regressor signal. If there are changes in the adhesion level, these will be reflected in changes in the parameters of the model. However parameter changes alone are difficult to interpret, hence the need to determine the eigenvalues of the regressor model for analysis.

The method is demonstrated here with a second order model that creates a corresponding single eigenvalue as this dimension model proved sufficient to estimate the creep force signal and also highlight changes in the adhesion level, where

$$
\hat{F}_t = \theta_1 F_{t-1} + \theta_2 F_{t-2}
$$

(35)

5.1.1. Results

Outputs from the creep force estimation process are analysed here in this first parameter identification post processing method, hence presented here are a number of sets of data for steady state adhesion levels and step changes in adhesion level from section 4.1.

The model as outlined in the previous section produces a single eigenvalue for analysis and the method utilises a 5 second moving window of creep force estimation data sampled at 100 Hz. The first set of tests analyses the simulation creep force estimates at constant adhesion levels. Figure 11 shows the variation with time of the eigenvalue for four set adhesion levels and full scale lateral track irregularity. Subplot (a) shows the eigenvalue signal to have a level of variation that means output values cross over but that the general trend is that as the adhesion level reduces the eigenvalue also reduces. Subplot (b) demonstrates how the signal can be smoothed if the Eigen value signal is averaged over a second moving window of two seconds. The adhesion levels are now clearly separate, threshold levels for which can be set to define adhesion level ‘bands’ and logic applied to determine the adhesion level.

The effects of changes in the track irregularity size on the post processing technique are shown in Figure 12. The analysis is applied to the ‘dry’ and ‘very low’ adhesion tests for full, half and double track irregularity sizes. This demonstrates that the outputs are dependent upon irregularity size but that the general trend of a reduction in the adhesion level means a reduction in the eigenvalue still holds, and that there is a definite separation in the eigenvalues for the extremes of the ad-
Figure 11. Creep force signal regressor parameter estimation eigenvalue analysis: (a) eigenvalue estimates; (b) time averaged eigenvalue

Adhesion levels present, meaning changes can be detected via simple threshold setting or more advanced fuzzy logic methods.

Figure 12. Creep force signal regressor parameter estimation eigenvalue analysis for varying track irregularity sizes: (a) eigenvalue estimates; (b) time averaged eigenvalue

Figure 13 shows how the analysis reacts to a step test change in adhesion level for three sizes of track irregularity: full scale; double scale; and half scale. Subplot (a) highlights how the adhesion change step reduction occurs at 15 seconds from the ‘dry’ condition to the ‘very low’ condition. Subplot (b), similarly to the previous example, shows the eigenvalue for a moving five second window of 100 Hz sampled data. The outputs are quite noisy and difficult to interpret, but show a definite change as the adhesion condition reduces. This signal is again time averaged over a moving two second window and is shown in subplot (c). This more clearly shows the change in the frequency content of the signal, therefore as mentioned thresholds for this change can be set to determine the adhesion condition. Careful selection is
required for the size of the sample windows, both the identification and the time averaging: too short and the output signal reacts to erroneous small changes; too long and the analysis may not react to important adhesion changes. The figure also shows that the above trends are repeated for the range of track irregularity levels.

5.2. Estimation of traction coefficients

The previous technique made no requirements for knowledge of the track irregularity level but requires user-based knowledge of the system to impose thresholds on processed data to determine changes in the adhesion conditions that may be dependent upon a particular section of track. The concept behind this second post-processing technique is to use the estimated creep forces and knowledge of the track time signals to identify linearised traction coefficients. The basic flow of the technique is shown in Figure 14 where: the estimated creep forces and moments, along with measured dynamics data and stored synchronised track irregularity data is partitioned into suitably sized moving windows; the least square algorithm is applied to generate $\theta$ values for the grey box equations below; the $\theta$ values are interpreted into creep coefficients; the creep coefficients are then interpreted by logic or fuzzy logic to determine the current adhesion level.

The model format in this method uses a simplified linear model from [19] of the lateral and yaw creep forces for the front wheelset of the front bogie is represented
as

\[ F_{FF} = 2f_{22} \psi_{FF} - \frac{2f_{22}}{V} \dot{y}_{FF} \]  

(36)

\[ M_{FF} = \frac{2l \lambda f_{11}}{r_0} (y_{FF} - d_{FF}) - \frac{2l^2 f_{11}}{V} \dot{\psi}_{FF} \]  

(37)

where \( f_{11} \) is the longitudinal creep coefficient, \( f_{22} \) is the lateral creep coefficient, \( \lambda \) is the linearised conicity of the wheelset, \( l \) is the half width of the bogie, \( F_{FF} \) and \( M_{FF} \) are the lateral creep force and yaw creep torque respectively, \( y_{FF} \) is the lateral position of the wheelset, \( d_{FF} \) is the lateral track irregularity, \( \psi_{FF} \) is yaw position of the wheelset, \( V \) is the forward vehicle speed and \( r_0 \) is the rolling radius of the wheels. The assumption for the parameterisation model is that

\[ F_{FF} = \theta_a \psi_{FF} - \theta_b \dot{y}_{FF} \]  

(38)

\[ M_{FF} = \theta_c (y_{FF} - d_{FF}) + \theta_d \dot{\psi}_{FF} \]  

(39)

where \( \theta_a, \theta_b, \theta_c \) and \( \theta_d \) are the parameters to be estimated. Equation 39 shows that information of the lateral track irregularity is required, this is assumed to be as a time signal of the lateral irregularity that in practice will require careful synchronisation. Additional geometric knowledge is required of the assumed linearised conicity, in addition to the speed of the vehicle. The creep coefficients can then be found working backwards from the identified parameter values using

\[ f_{11a} = -\frac{\theta_c r_0}{2l \lambda}; \ f_{11b} = -\frac{\theta_d V}{2l^2} \]  

(40)

\[ \dot{f}_{11} = \frac{f_{11a} + f_{11b}}{2} \]  

(41)

\[ f_{22a} = \frac{\theta_d}{2}; \ f_{22b} = -\frac{\theta_b V}{2} \]  

(42)

\[ \dot{f}_{22} = \frac{f_{22a} + f_{22b}}{2} \]  

(43)

5.2.1. Results

The analysis was again performed for a number of tests with static and dynamically varying adhesion levels, all tests utilised a moving 5 second window of data. Static adhesion tests were performed for the four adhesion levels and for clarity at the full lateral track irregularity level. This is shown in Figure 15 which highlights a clear drop in the identified \( f_{22} \) level from the ‘dry’/‘wet’ level to the ‘low’ and ‘very low’ levels, meaning again that threshold levels can easily be defined to demonstrate large reductions in the available adhesion.

A further example of the analysis for the lateral creep coefficient is shown in Figure 16 for a step reduction in the adhesion level at 15 seconds. The analysis is
Figure 15. Parameter estimation of linearised traction coefficients from estimated creep forces and known lateral track irregularity levels for four levels of adhesion.

Figure 16. Parameter estimation of traction coefficients for a step change in the adhesion level; (a) adhesion change; (b) estimated lateral creep coefficient.

Again performed for a moving 5 second time window and is additionally demonstrated for three different levels of track irregularity. It can be seen that again a definite reduction in the estimated creep coefficient is observed and that the levels are now virtually independent of lateral track irregularity levels and with very little time lag. Therefore this second technique with knowledge of the lateral track irregularity level offers key performance benefits over the analysis without track knowledge, however it comes at the cost of precise synchronisation of signals that may be practically difficult, this is a consideration that will receive further development in future studies as this will be a key consideration in any application of the method.
6. Conclusions

Low adhesion in the wheel/rail interface of rail vehicles is a significant problem that causes cost and punctuality issues for railway networks worldwide. A two step method has been proposed for the detection of areas of low adhesion: the first step is creep force estimation using low cost inertial sensors and Kalman-Bucy filtering; the second step is the post processing of the creep force estimate signals via parameter identification to reduce their dependency upon track irregularity levels and produce an interpretable signal. Guidance creep forces reduce as the adhesion level reduces and the Kalman-Bucy filter method as presented has been shown to estimate creep forces with confidence for a variety of adhesion conditions, both in static and dynamically varying conditions. Two post processing techniques using parameter identification have been shown to reduce the dependency of the signal upon variable track irregularity levels, with the most advantageous results arising when information of the track irregularity level is used. This proven technique in simulation will now be further validated using multi-bodied simulation data and finally using full scale testing.

Acknowledgments

The authors would like to thanks RSSB who manage project T959 with funding provided by the UK Rail Industry Strategic Rail Programme.

Appendix A. System dynamics modelling

Lateral and yaw plan view system dynamics modelling for a British Mk.3 coach.

\[ m_{FF} \ddot{y}_F = F_{L_y} + F_{R_y} + F_{s_y} + F_g \]  \hspace{1cm} (A1)

\[ I_{FF} \ddot{\psi}_F = F_{L_y} R_{L_x} - F_{L_x} R_{L_y} + F_{R_y} R_{R_x} - F_{R_x} R_{R_y} + M_{s_y} + M_g \]  \hspace{1cm} (A2)

\[ m_{FR} \ddot{y}_F = F_{L_y} + F_{R_y} + F_{s_y} + F_g \]  \hspace{1cm} (A3)

\[ I_{FR} \ddot{\psi}_F = F_{L_y} R_{L_x} - F_{L_x} R_{L_y} + F_{R_y} R_{R_x} - F_{R_x} R_{R_y} + M_{s_y} + M_g \]  \hspace{1cm} (A4)

\[ m_{RF} \ddot{y}_F = F_{L_y} + F_{R_y} + F_{s_y} + F_g \]  \hspace{1cm} (A5)
\[ I_{RF} \ddot{\psi}_{RF} = F_{LyRF} R_{LxRF} - F_{LxRF} R_{LyRF} \\
+ F_{RyRF} R_{RxRF} - F_{RxRF} R_{RyRF} \\
+ M_{s\psi_{RF}} + M_{g_{RF}} \]  

(A6)

\[ m_{RR} \ddot{y}_{RR} = F_{Lg_{RR}} + F_{Rg_{RR}} + F_{s_{y_{RR}}} + F_{g_{RR}} \]  

(A7)

\[ I_{RR} \ddot{\psi}_{RR} = F_{LyRR} R_{LxRR} - F_{LxRR} R_{LyRR} \\
+ F_{RyRR} R_{RxRR} - F_{RxRR} R_{RyRR} \\
+ M_{s\psi_{RR}} + M_{g_{RR}} \]  

(A8)

\[ m_{FB} \ddot{y}_{FB} = -(F_{sy_{FF}} + F_{sy_{FR}} + F_{sy_{VF}}) \]  

(A9)

\[ I_{FB} \ddot{\psi}_{FB} = -(M_{s\psi_{FF}} + M_{s\psi_{FR}} + M_{sy_{VF}} \\
+ L(F_{sy_{FF}} - F_{sy_{FR}})) \]  

(A10)

\[ m_{RB} \ddot{y}_{RB} = -(F_{sy_{RF}} + F_{sy_{RR}} + F_{sy_{VR}}) \]  

(A11)

\[ I_{RB} \ddot{\psi}_{RB} = -(M_{s\psi_{RF}} + M_{s\psi_{RR}} + M_{sy_{VR}} \\
+ L(F_{sy_{RF}} - F_{sy_{RR}})) \]  

(A12)

\[ m_{V} \ddot{y}_{V} = F_{sy_{VF}} + F_{sy_{VR}} \]  

(A13)

\[ I_{V} \ddot{\psi}_{V} = M_{s\psi_{VF}} + M_{s\psi_{VR}} \]  

(A14)

where \( F_{ijkl}, R_{ijkl}, M_{\psi_{kl}} \) are the forces (creep, gravitational and suspension), positions and moments, \( m_{kl} \) is the mass, \( I_{kl} \) is the moment of inertia, \( y_{kl} \) is the lateral position, \( \psi_{kl} \) is the yaw angle; where \( i = L(eft), R(ight), s(pension); j = x \) (longitudinal), \( y \) (lateral); \( k = F(ront \ bogie), R(rear \ bogie), V(ehicle); l = F(ront \ wheelset), R(rear \ wheelset), B(ogie) \)

The accompanying suspension forces and moments (for small angles) for the primary and secondary suspension are given by equations A15 to A26.

\[
F_{sy_{FF}} = (k_{y1} + k_{y1B}) y_{BF} - (k_{y1} + k_{y1B}) y_{FF} + k_{y1B}(L - D) \psi_{FF} \\
+ (k_{y1}L + k_{y1BD}) \psi_{BF} + (f_{y1} + f_{y1B}) \dot{y}_{BF} - (f_{y1} + f_{y1B}) \dot{y}_{FF} \\
+ f_{y1B}(L - D) \dot{\psi}_{FF} + (f_{y1}L + f_{y1BD}) \dot{\psi}_{BF}
\]  

(A15)
\[ M_{\psi FF} = k_{y1B}(L - D)y_{FF} - (k_{x1} + k_{x1B})l^2\psi_{FF} - k_{y1B}(L - D)^2\psi_{FF} - k_{\psi1B}\psi_{FF} - k_{y1B}(L - D)y_{BF} + (k_{x1} + k_{x1B})l^2\psi_{BF} - k_{y1B}(L - D)D\psi_{BF} + k_{\psi1B}\psi_{BF} + f_{y1B}(L - D)y_{FF} - (f_{x1} + f_{x1B})l^2\psi_{FF} - f_{\psi1B}\psi_{FF} - f_{y1B}(L - D)y_{BF} + (f_{x1} + f_{x1B})l^2\psi_{BF} - f_{y1B}(L - D)D\psi_{BF} + f_{\psi1B}\psi_{BF} \tag{A16} \]

\[ F_{\psi FR} = (k_{y1} + k_{y1B})y_{BF} - (k_{y1} + k_{y1B})y_{FR} - k_{y1B}(L - D)\psi_{FR} - (k_{y1}L + k_{y1B}D)\psi_{BF} + (f_{y1} + f_{y1B})\dot{y}_{BF} - (f_{y1} + f_{y1B})\dot{y}_{FR} - f_{y1B}(L - D)\dot{\psi}_{BF} - (f_{y1}L + f_{y1B}D)\dot{\psi}_{BF} \tag{A17} \]

\[ M_{\psi FR} = - k_{y1B}(L - D)y_{FR} - (k_{x1} + k_{x1B})l^2\psi_{FR} - k_{y1B}(L - D)^2\psi_{FR} - k_{\psi1B}\psi_{FR} + k_{y1B}(L - D)y_{BF} + (k_{x1} + k_{x1B})l^2\psi_{BF} - k_{y1B}(L - D)D\psi_{BF} + k_{\psi1B}\psi_{BF} - f_{y1B}(L - D)\dot{y}_{FR} - (f_{x1} + f_{x1B})l^2\dot{\psi}_{FR} - f_{\psi1B}\dot{\psi}_{FR} + f_{y1B}(L - D)\dot{y}_{BF} + (f_{x1} + f_{x1B})l^2\dot{\psi}_{BF} - f_{y1B}(L - D)D\dot{\psi}_{BF} + f_{\psi1B}\dot{\psi}_{BF} \tag{A18} \]

\[ F_{\psi RF} = (k_{y1} + k_{y1B})y_{BR} - (k_{y1} + k_{y1B})y_{RF} + k_{y1B}(L - D)\psi_{RF} + (k_{y1}L + k_{y1B}D)\psi_{BR} + (f_{y1} + f_{y1B})\dot{y}_{BR} - (f_{y1} + f_{y1B})\dot{y}_{RF} + f_{y1B}(L - D)\dot{\psi}_{RF} + (f_{y1}L + f_{y1B}D)\dot{\psi}_{BR} \tag{A19} \]

\[ M_{\psi RF} = k_{y1B}(L - D)y_{RF} - (k_{x1} + k_{x1B})l^2\psi_{RF} - k_{y1B}(L - D)^2\psi_{RF} - k_{\psi1B}\psi_{RF} - k_{y1B}(L - D)y_{BR} + (k_{x1} + k_{x1B})l^2\psi_{BR} - k_{y1B}(L - D)D\psi_{BR} + k_{\psi1B}\psi_{BR} + f_{y1B}(L - D)y_{RF} - (f_{x1} + f_{x1B})l^2\dot{\psi}_{RF} - f_{\psi1B}\dot{\psi}_{RF} - f_{y1B}(L - D)y_{BR} + (f_{x1} + f_{x1B})l^2\dot{\psi}_{BR} - f_{y1B}(L - D)D\dot{\psi}_{BR} + f_{\psi1B}\dot{\psi}_{BR} \tag{A20} \]

\[ F_{\psi RR} = (k_{y1} + k_{y1B})y_{BR} - (k_{y1} + k_{y1B})y_{RR} - k_{y1B}(L - D)\psi_{RR} - (k_{y1}L + k_{y1B}D)\psi_{BR} + (f_{y1} + f_{y1B})\dot{y}_{BR} - (f_{y1} + f_{y1B})\dot{y}_{RR} - f_{y1B}(L - D)\dot{\psi}_{RR} - (f_{y1}L + f_{y1B}D)\dot{\psi}_{BR} \tag{A21} \]
$M_{\psi VR} = -k_y 1B(L - D)\dot{y}_{RR} - (k_{x1} + k_{x1B})^2\psi_{BR} - k_y 1B(L - D)^2\psi_{RR}$
\begin{equation}
- k_{\psi 1B}\psi_{RR} + k_y 1B(L - D)\dot{y}_{BR} + (k_{x1} + k_{x1B})^2\dot{\psi}_{BR} \\
- k_y 1B(L - D)D\psi_{BR} + k_{\psi 1B}\psi_{BR} - f_y 1B(L - D)\dot{y}_{RR} \\
- (f_{x1} + f_{x1B})^2\dot{\psi}_{RR} - f_y 1B(L - D)^2\dot{\psi}_{RR} - f_{\psi 1B}\dot{\psi}_{RR} \\
+ f_y 1B(L - D)\dot{y}_{BR} + (f_{x1} + f_{x1B})^2\dot{\psi}_{BR} - f_y 1B(L - D)D\dot{\psi}_{BR} \\
+ f_{\psi 1B}\dot{\psi}_{BR} \tag{A22}
\end{equation}

$F_{syVF} = -k_y 2y_V - f_y 2\dot{y}_V + k_y 2y_{BF} + f_y 2\dot{y}_{BF} - k_y 2c\psi_V - f_y 2c\dot{\psi}_V \tag{A23}$

$M_{syVF} = -k_y 2c^2\dot{\psi}_V - f_y 2c^2\dot{\psi}_V + k_y 2c\dot{y}_V + f_y 2c\dot{y}_{BF} + f_y 2c\dot{\psi}_{BF} \tag{A24}$

$F_{syVR} = -k_y 2y_V - f_y 2\dot{y}_V + k_y 2y_{BR} + f_y 2\dot{y}_{BR} + k_y 2c\psi_V + f_y 2c\dot{\psi}_V \tag{A25}$

$M_{syVR} = -k_y 2c^2\dot{\psi}_V - f_y 2c^2\dot{\psi}_V - k_y 2c\dot{y}_V - f_y 2c\dot{y}_{BR} - k_y 2c\psi_{BR} - f_y 2c\dot{\psi}_{BR} \tag{A26}$

where $k_{mn}$ and $f_{mn}$ are the suspension stiffness and damper coefficients; with $m = y$ (lateral) or $\psi$ (yaw); $n = 1$ (primary suspension), 2 (secondary suspension).

**Appendix B. System identification**

System identification through parameter estimation is a concept that was first used to create dynamic models of a system from collected data instead of through first principles analysis [18] by the selection of parameters for an assumed model structure by statistical means. The essential principle is to minimise the sum squared error of the residual vector of a system over a defined window of data

$$r_t = y_t - \hat{y}_t \tag{B1}$$

where $r_t$ are the residuals, $y_t$ is the observed data and $\hat{y}_t$ is the predicted data. The standard system equation in matrix form is

$$y_t = X\theta + \epsilon \tag{B2}$$

where $y$ is the measure output, $X$ is the user selected matrix of parameter regressors, $\theta$ the selected parameter values and $\epsilon$ is the equation error. The least squares solution is of the form

$$(X^TX)\theta = X^Ty_t \tag{B3}$$

This can be rearranged to the easily handled form

$$\theta = (X^TX)^{-1}X^Ty_t \tag{B4}$$
References


