Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-19 (2022) 103–108

Data-Driven Wheel Slip Diagnostics for Improved Railway Operations

Bernadin Namoano*, Cristobal Ruiz-Carcel*, Christos Emmanouilidis , and Andrew G Starr***

**Cranfield University, United Kingdom (email: {bernadin.namoano; c.ruizcarcel; a.starr}@cranfield .ac.uk)*
***University of Groningen, Netherlands, c.emmanouilidis@rg.nl)*

Abstract: Wheel slip activity detection is crucial in railway maintenance, as it can contribute to avoiding wheel damage but also track deteriorations leading to significant maintenance costs, trains delays, as well as the risk of accidents. Wheel slip activity is characterised by lower adhesion between track and wheel, especially in braking conditions, locking the wheels. It is complex to model or predict, being influenced by a multitude of factors including ambient conditions, global vehicle load, track and axle quality, leaves and objects present on the rail, steep incline, oxidation of the rails, and braking forces applied to the wheels. This paper presents a combined wavelet and tuned Long-Short Term Memory (LSTM) approach for the detection of wheel slip from time series data collected from real-world trains. Results provide evidence of superior performance over methods such as decision trees and random forests, naïve Bayes, k-nearest neighbours, logistic regression, and support vector machines.

Keywords: Streaming analytics; railway diagnostics, rime series classification, train wheel slip detection.

1. INTRODUCTION

One significant problem encountered in railway maintenance arises as a result of wheel slip activity. Often characterised by lower adhesion between the track and the wheel, specially during hard braking conditions, such slip can severely affect the railway infrastructure. The direct effects of wheel slip activities on the infrastructure are the track surface and the train wheels damage. Hence, damage can induce a significant maintenance cost, cause train delays, and increase the risk of accident as high severity defects can lead to trains going off-tracks. Additionally, wheel slip activity can initiate or accelerate incipient crack growth on the rails and the surrounding assets, and generate further vibration to the train body, generating noise also leading to passenger discomfort.

Early detection is crucial not only to address these issues but also to increase the associated assets' lifetime (Krummenacher et al., 2018). Wheel slip is particularly complex to model or predict as it is influenced by a multitude of factors including ambient conditions, global vehicle load, track and axle quality, leaves and objects present on the rail, steep incline, oxidation of the rails, and braking forces applied to the wheels. In general, two main types of wheel slip can be observed:

- The friction between the track and the wheel is so low that the wheel slides without moving the vehicle; the slipping axle speed shows short periods of fast acceleration (Garnham & Davis, 2009).
- During braking or acceleration, the wheels have transient periods of slippage. The slipping axles often lock in short periods, showing brief periods of fast speed drops in axle speed profiles. An example is shown in Figure 1 where the train moved between two stops. Two phases can be observed. The acceleration phase, where axle speeds increase to roughly a maximum speed of 55 miles per hour (mph) whilst the brake pressure is zero. This is followed by a deceleration phase, where short braking events are applied to the axles. Wheel slip activity can be seen between the time window [60-100] on both front and back axles.

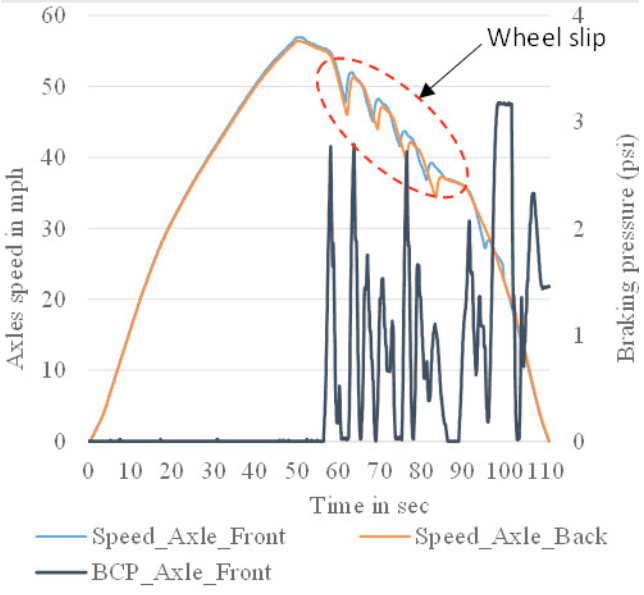

Figure 1 Wheel slip activity in axle speed profile.

The main contribution of the present study is the development of a suitable methodology for automatic wheel slip activity detection for the second type of slip. It applies a combination of univariate axle speed analysis using a wavelet approach and long short-term memory neural networks (LSTM) (Hochreiter & Schmidhuber, 1997). The proposed method is able to accurately distinguish between normal and wheel slip conditions using the trained model and operational measurement data. The analysis is conducted using diesel multiple unit (DMU) trains data as shown in Figure 2.

10.1016/j.ifacol.2022.09.191

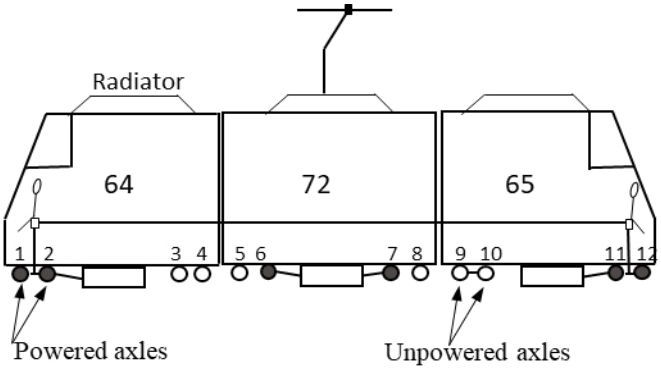

Figure 2 DMU train powered axles investigated.

Each train is comprised of three carriages (marked with 64, 72, and 65). The leading carriages have their front axles powered (axles 1, 2, 11, and 12) and the middle one has the 6 and 7 axles powered. When braking, brakes pressures are applied to the powered axles of all the carriages. The studied wheel slip activity occurs mainly in the powered axles. The results of the application of the methodology are compared to other classical methods using real world train operational data sets.

The paper is structured as follows. Section 2 discusses related work, placing this study in the context of the broader literature in the field. Section 3 focuses on the proposed methodological approach, followed by details on the comparative evaluation methods. Section 4 presents results obtained with developed and compared methods and a discussion and recommendation on the obtained results. Section 5 is the conclusion.

2. RELATED WORK

2.1 Wheel slip and the wheel – track interface

Understanding the wheel-rail contact (Enblom, 2009) is an important step to understand wheel slip behaviour. During acceleration, there is an increase in the rotating force applied on the vehicle axles. The level of friction (rolling resistance) between the wheel and the rail is less than the amount of force generated by the increased torque, which often results in over-rotation, also known as wheel slip activity. This wheel slip activity causes damage to both the track and the wheels. Another type of wheel behaviour that can occur under braking conditions is under-rotation. It is different from the aforementioned wheel slip, as the force being applied by the brakes is a resistance force. Specifically, the brake pad is pressing against the wheel and is opposed to the turning force from the torque under traction. If the amount of force applied via the brakes exceeds the resistive force of the wheel on the track (also including other variables such as vehicle weight and train speed), then the wheel will under-rotate. This lock-up causes a flat spot on the wheels. A flat spot is a serious incident for a train wheel as the slapping effect causes significant damage to the wheel as well as the track.

2.2 Wheel slip activity modelling

Wheel slip activity has been broadly studied through wheel-rail adhesion condition analysis. For example, substance contaminants (oil, snow, grease, water, leaves, wear, oxides, ballasts, and dust), present on the wheel-rail interfaces, significantly influence the adhesion coefficients that represent longitudinal and normal force ratio (Iwnicki, 2006),(Bosso et al., 2019). These adhesion coefficients need to be in line with defined safety standards. For this reason, manufacturers build assets, such as braking systems, wheels, and rails, that undergo various tests in different degradation conditions, in order to ensure that the equipment meets such standards (UIC, 2005), and (Barna, 2011). The off-line analysis is often model-driven, where vehicle dynamics such as braking distance, vehicle velocity range, wheel lock limit or slide limit, and compressed air pressure (Barna, 2011), are modelled. Additional analysis on improving adhesion through studying the effects of contaminants has been been repoted (Lewis & Olofsson, 2009) (Garnham & Davis, 2009). Limitations of the investigated model-driven methods are their implementation challenges on real-world applications, as assumptions such as constant vehicle mass, speed, and friction coefficients are not realistic. Moreover, models may include many parameters, which add supplementary difficulty in estimating them. In contrast, the tackled wheel slip activity problems are likely to show patterns in speed profiles. Therefore, there is a need to investigate the development of more data-driven approaches for wheel slip detection, especially methods that are able to take integrate short the short past information for the detection.

3. METHODOLOGY

The signals selected to drive the detection are axle speeds and brake pressure measurements. The proposed data processing methodology is shown in Figure 3.

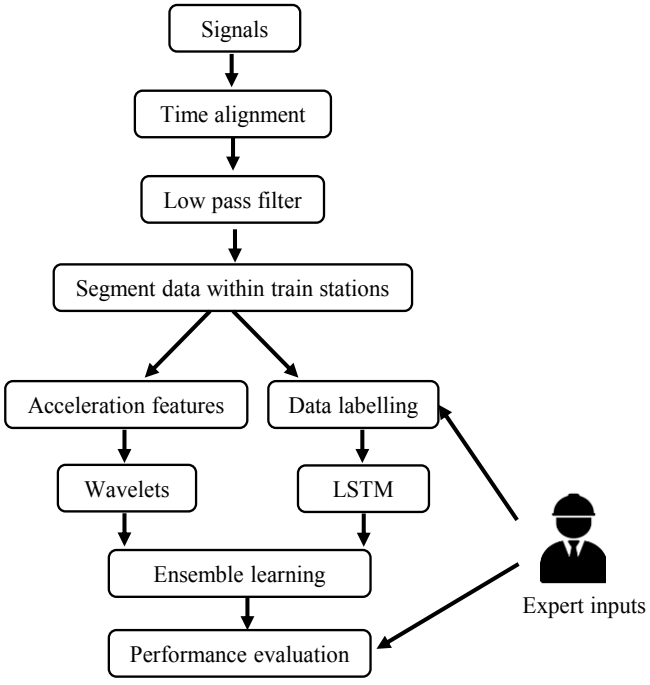

Figure 3 Methodology workflow

It includes a workflow that: (a) performs data alignment, to ensure appropriate time alignment of data derived from separate measurement sensor; (b) data filter to reduce signal noise and focus the analysis on signal aspects that are more likely to convey useful information; (c) data segmentation, to

derive time series segments enabling the analysis to focus on each identified segment, thus offering better contextual basis for the processing; (d) feature extraction and data labelling before applied further signal processing and modelling to prepare for the machine learning task; (e) the modelling and machine learning task; (f) combination of modelling for ensemble learning; and (g) performance evaluation. Expert knowledge is applied to check the validity of data labelling and performance evaluation. Each step is described in the following sections.

3.1 Data cleaning, time Alignment, and filtering

Axle speeds, braking pressure, braking types, and GPS coordinate and speeds, data are collected at a frequency of 20 Hz on different carriages of urban trains operating under different conditions and geographical locations. This offers sufficient sampling and resolution to observe the targeted wheel slip phenomena without missing out on relevant slippage events.Figure 2 shows the axles enumerated from 1 to 12 with axles 1, 2, 6, 7, 11, and 12 being the powered axles. Data attributes comprise GPS location, axle speed, the mechanical and electrodynamic brake pressures applied for each axle, and the leading vehicle information. Additional computed attributes are added in the investigation. These attributes are longitudinal and latitudinal acceleration, calculated from GPS coordinates; acceleration, derived from axle speeds; and brake levels regarding both emergency brakes and regular brakes.

As the data come from different sensors in different locations on the train, desynchronisation and time lag is often observed. In order to derive data directly usable for modelling and machine learning, data alignment needs to be established. The first level of alignment is made by synchronising the GPS time recorded with the train axles sensors time. Subsequently, a second level time alignment is made by applying Dynamic Time Wrapping (DTW) (Keogh & Ratanamahatana, 2005) with the L2-norm in the axle speeds, to derive a high level of time alignment in the stream data sequences.

In the filtering process, noise and disturbances on the braking and speed signals are reduced using low-pass filtering, and specifically exponential moving average techniques with a window of five time-steps, representing 0.25 seconds, an appropriate time granularity given the nature of the process.

3.2 Data Segmentation, Feature Extraction and Labelling

In the segmentation step, data are broken into events representing significant points in time during journeys. A journey represents the starting of a train from a station to its stop at the next station. To achieve an appropriate segmentation, two variables are taken into account. The first variable is axle speeds that drop to zero for a minimum period of two minutes. While other event duration might be set resulting in different event recordings, the value was set after discussing with domain experts, indicating that it is typically events that exceed such duration that is likely to initiate or contribute to evolving damage. The second variable is the release of the doors (unlock, opening, or closing). Combining these two variables, a heuristic rule such as "if the doors open and the stopping period is longer than two minutes, then it is at a station" can be used to segment the data. To be more accurate, train station GPS coordinates are used to compute the distance between the train position and the closest station. This distance should be lower than an empirically-set threshold of approximately 200 meters for the situation to qualify as the train being at a station. Such rules of thumb are basic examples of how contextualizing data rather than blind data-driven processing should be sought whenever possible.

For the univariate speed analysis, slope angle features are extracted from the axle speed profiles by computing an angular function $\varphi(t)$:

$$\varphi(t) = \tan^{-1}\frac{Y_{t+k} - Y_t}{X_{t+k} - X_t} \tag{1}$$

where tan^{-1} represents the inverse tangent function, Y_t represents the speed at the time t, and X_t represents the time. Considering quantization error and noise, the k value has been set to 5 instead of 1, representing a time variation of 0.25 seconds. The data involved in this work was labelled as normal and faulty behaviour by industry experts and this is relevant to incorporating expert knowledge in Figure 3.

3.3 Univariate Wavelet Analysis

In Figure 1, the axle speed profile indicates that wheel slip activities are present during the deceleration phase. Wheel slips represent a fast transition between normal to slip conditions for short time periods. A natural method to investigate fast transition changes is through wavelet analysis. A basic function ψ is a wavelet, also known as a mother wavelet if it satisfies the admissibility condition:

$$\int_{-\infty}^{+\infty} \frac{|\tilde{\Psi}(w)|}{|w|} dw < \infty \tag{2}$$

where $\tilde{\Psi}$ is the Fourier transform of ψ. This implies that the wavelet should integrate to zero. Many wavelets have been proposed over time, for example, Haar, Daubechies, Gaussian, and Morlet wavelets(Misiti et al., 2006). They differ from each other by their vanishing moments, which give different local regularities. The Morlet wavelet is used in the context of the present work, as it was seen to relate well with how human visión operates (Daugman, 1985). The scaling (dilatation) of the wavelet ψ with a factor α is defined as follows:

$$\psi_t(t) = \frac{1}{\alpha}\psi\left(\frac{t}{\alpha}\right) \tag{3}$$

The wavelet transforms of a signal S is denoted by C and for finite energy, at the scale (dilation) α at time t is defined by:

$$\mathrm{C}(\alpha, t) = \mathrm{S(t)} * \psi_\alpha(t) \tag{4}$$

where * is the convolution operation. The coefficients $C(\alpha, t)$ measure the variations of the signal S at the dilatation α. From a qualitative point of view, large $C(\alpha, t)$ values indicate irregularity of the signal S around the position t at the

dilatation α. Conversely, small values of the coefficients indicate signal regularity.

3.4 Multivariate LSTM analysis

Traditional feed-forward neural networks are unable to take into consideration previous output predictions. This may be a major issue in most time series classification problems, as time series data points (also called observations) often have short or long term dependencies. Recurrent neural networks (Lukoševičius & Jaeger, 2009) tackle this issue in sequential learning problems with a feedback loop between the neural network layers. However, recurrent neural networks show deficiencies known as exploding and vanishing gradient problems (Pascanu et al., 2012) in long term dependency learning. Long Short-Term Memory (LSTM) networks were introduced to address exploding or vanishing challenges where gradients grow to infinite or quickly decay to zero (Hochreiter & Schmidhuber, 1997). Additional notions of "cell" and "gate" that allow gradient update selections were added to the classical neural networks to build LSTM networks (Hochreiter & Schmidhuber, 1997). One of the main justification for the application of the LSTM method in wheel slip activity detection is that there is often a time lag between the cause of the wheel slip (hard braking, track condition) and the output observations. In addition, a wheel slip can cause severe damage leading to intermittent faults. Therefore, it is both very short term and longer term phenomena that may influence the modelled outcome. Moreover, LSTM networks are able to handle multivariate attributes with complex non-linear relationships, which is also in line with the present problem context.

3.5 Ensemble learning

Ensemble learning with voting or boosting methods produce stronger learners out of weaker ones (Drucker et al., 1994). A straightforward ensemble learning method is used for wheel slip activity recognition. Wavelet analysis is combined with LSTM methods with a voting method for efficient detection. Because wavelet analysis detects the activity when there is a fast change in shape curvatures, earlier detection is not possible. Therefore, when the LSTM detects wheel slip activity, the wavelet analysis is used as confirmation of this within the next two seconds post-detection.

3.6 Evaluation Methods

Established classification performance metrics are employed in this study. Specifically, four performance metrics are used to assess the classification approaches, namely precision, recall, the F1-score, and the area under the curve (AUC). The precision metric measures the percentage of matching the annotation accuracy, whereas the recall metric measures the percentage of annotated changes to be the true one. The F1-score is the test accuracy metric representing the accuracy of all identified cases. The AUC indicator measures the method's ability to separate normal and faulty conditions, and the AUC measurement is used in the present context since data are imbalanced.

$$Precision(P) = \frac{Number\ \text{of}\ correct\ annotations}{Number\ of\ annotation}\ \text{x}100 \quad (5)$$

$$Recall\ (R) = \frac{Number\ of\ Correct\ annotations}{Number\ of\ true\ annotations}\text{x}100 \quad (6)$$

$$f1\ score(F1) = 2\frac{P.R}{P+R} \quad (7)$$

3.7 Experiments

Experiments were carried out using real world data provided through an industrial partner, displayed in Figure 4, Figure 5, and Figure 6. It consists of 12,000,000 observations representing six months of the urban train operations and 26 attributes, including the axle speed used for the wavelet analysis. Table 1 presents the parameters empirically set for each method. The data has been divided into training, testing, and validation sets; 60% of the data has been used for training, 20% for testing, and 20% for the validation sets. The LSTM network is implemented using Keras and Tensorflow libraries.

Table 1. Learning experiment details

	Parameters
Data	**Number of observations**: 12,000,000 **Training**: 7,200,000 **Testing**: 2,400,000 **Validation**: 2,400,000 **Number of attributes**: 27
Wavelet+LSTM	**Mother wavelet:** Morlet **Input signals**: axle speeds **Hidden neurons per layer**: 50, 90,60 **Activation**: reLu, sigmoid, sigmoid **Solver**: m-adam **Adam regularisation parameter**: 0.05
Decision trees	**Minimum number of instance leave:** 50 **Maximal tree depth:** 1000 **Minimum subset split:** 80
K-NN	**Number of neighbours:** 50 **Metric:** Mahalanobis **Weight:** Uniform weight
Naive Bayes	-
Logistic regression	**Regularisation:** Ridge (L2-norm)
SVM	**Cost:** 1.00 **Kernel:** Radial Basis Function (RBF) **Numerical tolerance:** 0.0010
Random forest	**Number of trees:** 80 **Maximal tree depth:** 1000 **Minimum subset split:** 70

4. RESULTS AND DISCUSSION

In the wavelet analysis, an example of the computed $\varphi(t)$ is shown in Figure 4. A fast transition to slip mode shows a sharp increase and decrease of the φ values. The scalogram of the convolution of φ with the Morlet wavelet is shown in Figure 5. The normalised $C(\alpha, t)$ values, the scale, and the time are highlighted in this figure. It can be seen that high normalised $C(\alpha, t)$ values, marked in yellow, show irregularities in the speed signal, which represent wheel slip activities. Wavelet analysis shows efficient detection when wheel slip activity occurs during the speed deceleration phase. However, when it occurs during the acceleration phase, the speed profile may not be significantly affected. Figure 6 shows an example of this case. The LSTM method compensates for this deficiency through the learning of the multivariate attributes. An

interesting observation of the LSTM result is that the method is able to detect wheel slip conditions several seconds before the slipping event starts. This is related to the causal relationship between the braking pressure and the resulting slip. Such timely detection can allow, for example, automatic activation controllers to soft release the brakes or inject more sand to reduce the slip activity.

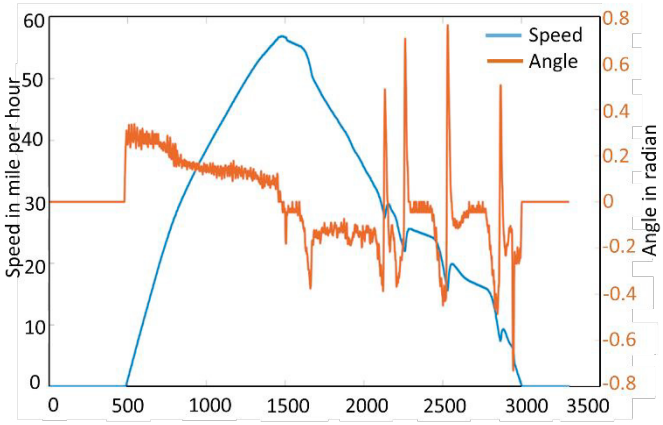

Figure 4 Angle value computed from the axles speed.

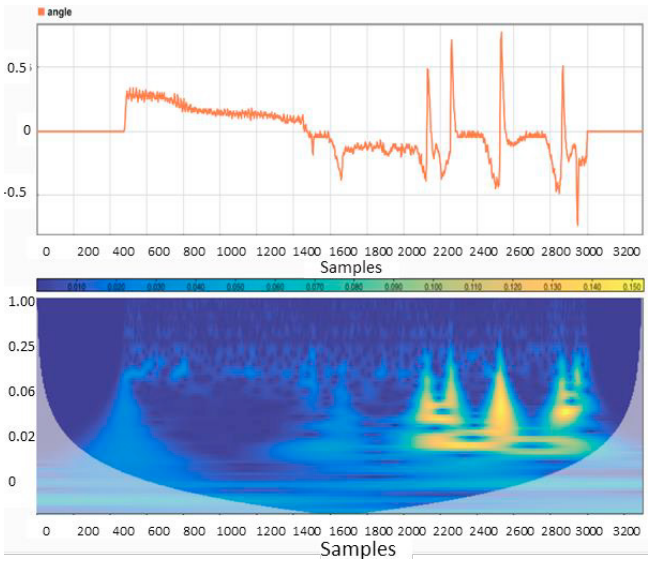

Figure 5 Scalogram of the wavelet.

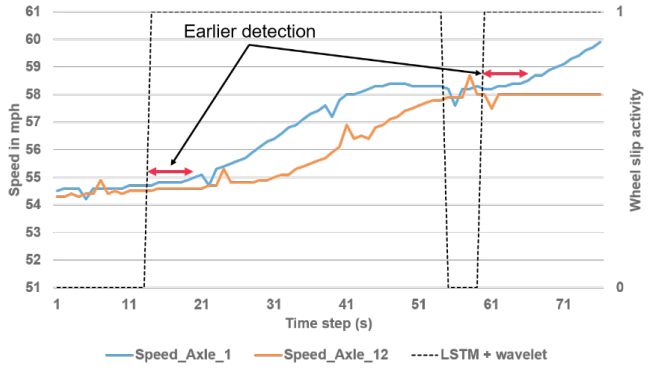

Figure 6 Wheel slip activity during an acceleration phase.

Table 2 shows the investigated methods of performance during the validation phase. The colour represents a visualisation of the performance for each metric, with green indicating better performance. Multiple classification methods were applied to this effect, including Decision trees, k-Nearest Neighbours (k-NN), Naïve Bayes, Logistic regression, SVM, and Random Forest, on top of the proposed joint Wavelet and LSTM approach of this paper. The Random Forest method shows the lowest performance in comparison to the remaining classifiers. It is followed by the Decision Tree and k-NN methods, which show similar F1 accuracy. Whilst the F1 score is higher (94% and 92% respectively), the AUC shows a low-performance score (50%). This is linked to the fact that the data are imbalanced and therefore the classifiers are unable to separate the classes very well. Naïve Bayes, Logistic regression, and SVM methods show the AUC's best performance score in comparison to the random forest, k-NN, or decision tree methods. The highest performance score is shown by the proposed method where the combined LSTM+wavelet approach outperformed all of the classifiers.

Table 2. Validation results of the classifications methods.

Model	AUC	F1	Precision	Recall
Decision tree	50	94	92	96
K-NN	50	93	92	95
Naive Bayes	86	86	87	85
Logistic regression	78	91	94	89
SVM	68	82	91	76
LSTM + Wavelet	98	98	99	97
Random forest	40	60	49	79

Limitations of the analysis are principally related to the employed classifier hyper-parameter settings and the very specific nature of the employed data (methods can be performed differently on other datasets). Numeric hyper-parameters have been tested empirically using a genetic algorithm to find appropriate settings. However, as the genetic algorithm does not guarantee global maxima, the obtained parameters can be biased, leading to evidence but not necessary conclusions. The identified hyperparameters have been set throughout the trials in order to employ the best performing settings, according to the evidence. A further limitation is related to the data annotation process. Although experts have annotated data, human-based annotations may have biases or errors. These in turn influence the learning and the classification results. Additionally, the experiments carried out do not take into consideration information such as track and wheel condition, vehicle mass, and traction forces, which are all contextually relevant parameters. The lack of such inputs limits the performance of the undertaken analysis.

5. CONCLUSION

This paper introduced a new method for detecting wheel slips activity on urban trains. The employed method involves a multi-stage data workflow, that includes data cleaning and alignment, data filtering, data labelling and feature extraction, and a simple ensemble approach that based on time series wavelet and LSTM learning, was shown to be successful in identifying wheel slip events when tested on real data from

operational trains. Focusing on the performance of the method, this compared favourably to alternative classifiers, namely Decision tree, K-NN, Naïve Bayes, Logistic Regression, SVM, and Random Forest. Limitations of the proposed method are the partial reliance on expert annotation and reliance on method parameters that have been tested empirically. Moreover, in this analysis, the absence of certain conditions such as weather information, track condition, traction load, and wheel condition limits analysis performance. Areas for improvement include tackling the challenge mentioned and incorporating external data for more efficient diagnosis. Long-term work includes designg a distributed online diagnosis system with fault isolation, allowing similar train systems to share newly discovered fault characteristics.

ACKNOWLEDGEMENTS

The authors wish to thank Unipart Rail, Instrumentel for providing the industrial data and domain-specific expertise and the Engineering and Physical Sciences Research Council (EPSRC) for the financial support through grant 2203091. Partial financial support towards change detection research was provided through H2020 project STAR, grant ID 956573.

REFERENCES

Barna, G. (2011). Control System of Wheel Slide Protection Devices for Rail Vehicles Meeting the Requirements of European Normative Documents System Sterowania Układów. *Proceedings of 13th International Conference on Quality, Safety and Ecology in Transport*, 11–18.

Bosso, N., Gugliotta, A., Magelli, M., Oresta, I. F., & Zampieri, N. (2019). Study of wheel-rail adhesion during braking maneuvers. *Procedia Structural Integrity*, *24*, 680–691. https://doi.org/10.1016/j.prostr.2020.02.060

Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, *2*(7), 1160. https://doi.org/10.1364/josaa.2.001160

Drucker, H., Cortes, C., Jackel, L. D., LeCun, Y., & Vapnik, V. (1994). Boosting and Other Ensemble Methods. *Neural Computation*, *6*(6), 1289–1301. https://doi.org/10.1162/neco.1994.6.6.1289

Enblom, R. (2009). Deterioration mechanisms in the wheel-rail interface with focus on wear prediction: A literature review. *Vehicle System Dynamics*, *47*(6), 661–700. https://doi.org/10.1080/00423110802331559

Garnham, J. E., & Davis, C. L. (2009). Rail materials. In *Wheel-Rail Interface Handbook* (pp. 125–171). Elsevier. https://doi.org/10.1533/9781845696788.1.125

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Iwnicki, S. (2006). *Handbook of Railway Vehicle Dynamics* (S. Iwnicki & S. Iwnicki, Eds.). CRC Press. https://doi.org/10.1201/9781420004892

Keogh, E., & Ratanamahatana, C. A. (2005). Exact indexing of dynamic time warping. *Knowledge and Information Systems*, *7*(3), 358–386. https://doi.org/10.1007/s10115-004-0154-9

Krummenacher, G., Ong, C. S., Koller, S., Kobayashi, S., & Buhmann, J. M. (2018). Wheel Defect Detection with Machine Learning. *IEEE Transactions on Intelligent Transportation Systems*, *19*(4), 1176–1187. https://doi.org/10.1109/TITS.2017.2720721

Lewis, R., & Olofsson, U. (2009). *Wheel—rail interface handbook*. Woodhead Publishing Limited. https://doi.org/10.1533/9781845696788

Lukoševičius, M., & Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, *3*(3), 127–149. https://doi.org/10.1016/j.cosrev.2009.03.005

Misiti, M., Oppenheim, G., Poggi, J.-M., & Misiti, Y. (2006). *Wavelets and Their Applications (Digital Signal and Image Processing Series)*. ISTE.

Pascanu, R., Mikolov, T., & Bengio, Y. (2012). On the difficulty of training Recurrent Neural Networks. *30th International Conference on Machine Learning, ICML 2013*, *PART 3*, 2347–2355.

UIC. (2005). *UIC Code 541-5 - Brakes - Specifications for the construction of various brake parts - Wheel Slide Protection device (WSP)* (3rd Edition).

2022-09-27

Data-driven wheel slip diagnostics for improved railway operations

Namoano, Bernadin

Elsevier

Namoano B, Ruiz-Carcel C, Emmanouilidis C, Starr AG. (2022) Data-driven wheel slip diagnostics for improved railway operations. IFAC-PapersOnLine, Volume 55, Issue 19, pp. 103-108
https://doi.org/10.1016/j.ifacol.2022.09.191