

ADHESION LEVEL IDENTIFICATION IN WHEEL-RAIL CONTACT USING DEEP NEURAL NETWORKS

Sanaullah Mehran Ujjan

ME Scholar, NCRA- Condition Monitoring Lab.

Mehran University of Engineering and Technology, Jamshoro, (Pakistan).

E-mail: mehranujjan44@gmail.com ORCID: <https://orcid.org/0000-0003-1879-8754>

Imtiaz Hussain Kalwar

Head of Department, Electrical Engineering.

DHA Suffa University, Karachi, (Pakistan).

E-mail: imtiaz.hussain@dsu.edu.pk ORCID: <https://orcid.org/0000-0002-7947-9178>

Bhawani Shankar Chowdhry

Professor Emeritus.

Mehran University of Engineering and Technology, Jamshoro, (Pakistan).

E-mail: bhawani.chowdhry@faculty.muet.edu.pk ORCID: <https://orcid.org/0000-0002-4340-9602>

Tayab Din Memon

Chairman, Department of Electronics.

Mehran University of Engineering and Technology, Jamshoro, (Pakistan).

E-mail: tayabdin82@gmail.com ORCID: <https://orcid.org/0000-0001-8122-5647>

Dileep Kumar

Research Assistant, NCRA- Condition Monitoring Lab.

Mehran University of Engineering and Technology, Jamshoro, (Pakistan).

E-mail: dileepkalani1994@gmail.com ORCID: <https://orcid.org/0000-0002-6211-1078>

Recepción: 29/01/2020 **Aceptación:** 17/04/2020 **Publicación:** 30/04/2020

Citación sugerida Suggested citation

Ujjan, S. M., Kalwar, I. H., Chowdhry, B. S., Memon, T. D., y Kumar, D. (2020). Adhesion level identification in wheel-rail contact using deep neural networks. *3C Tecnología. Glosas de innovación aplicadas a la pyme. Edición Especial, Abril 2020*, 217-231. <http://doi.org/10.17993/3ctecno.2020.specialissue5.217-231>

ABSTRACT

Robust and accurate adhesion level identification is crucial for proper operation of railway vehicle. It is necessary for braking and traction forces characterization, development of maintenance strategies, wheel-rail wear predictions and development of robust onboard health monitoring systems. Adhesion being the function of many uncertain parameters is difficult to model, whereas data driven algorithms such as Deep Neural networks (DNNs) are very good at mapping a nonlinear function from cause to effect. In this research a solid axle Wheel-set was modeled along with different adhesion conditions and a dataset was prepared for the training of DNNs in Python. Furthermore, it explored the potential of DNNs and various data driven algorithms on our noisy sequential dataset for classification task and achieved 91% accuracy in identification of adhesion condition with our final model.

KEYWORDS

Wheel-rail contact, Adhesion, Solid-axle Wheel-set, Deep learning, Neural networks, Classification, Time-series data.

1. INTRODUCTION

Adhesion level identification is an important task for proper operation of a railway vehicle as traction effort is a direct function of adhesion coefficient. Hence different magnitude of tracking and braking forces is required in different contact conditions, making adhesion level identification crucial for tracking and braking forces characterization (Spiryagin *et al.*, 2014). Both overestimation or under estimation can cause either derailment or rapid noise at wheel rail interface (Olofsson, 2009).

Adhesion coefficient quantitatively relates ratio of traction effort to the normal load on the wheel and qualitatively represents contact condition between wheel and rail.

Many cases of derailment have been reported in Pakistan (“A Timeline of Neglect: Train Incidents in Pakistan”, 2019) due to various causes but with a suitable mechanism of adhesion level identification and adhesion control at least one factor (over/under estimation of adhesion condition) can be tackled down.

Adhesion problem has sought attention of many researchers and many solutions ranging from mathematical control theory to statistical and genetic algorithms (Bibi, Chowdry, & Shah, 2018) have been proposed and applied (Shrestha, Wu, & Spiryagin, 2019). Although, here we compare two of them, model based and data driven work on adhesion, but later being rarely applied we aim to express the potential of data driven models in railway domain.

1.1. MODEL BASED WORK

Several attempts to correctly estimate the adhesion have been reported in recent years related to model based approaches (Shrestha *et al.*, 2019; Malvezzi *et al.*, 2013; Hussain, 2012; Ward *et al.*, 2012) where a mathematical relation between excitation and response of the system is constructed. Adhesion has been expressed as a function of slip velocity and acceleration by Carl and Brook (1985), and Malvezzi *et al.* (2013). Function of longitudinal vehicle velocity and pressure in contact zone by Spiryagin *et al.* (2016), function of lateral creep force and yaw movement by Ward *et al.* (2012).

However modelling a physical phenomenon such as adhesion with too many uncertain or even unknown factors is a time taking task. As any mathematical model inherently has the tendency to ignore some parameters in the process but ignoring adhesion effective parameters for the sake of simplicity and robustness of the model can cause difficulties especially when the model is later on used for complex tasks e.g. development of a Real-time on-board health monitoring system in locomotives (Shah *et al.*, 2020).

Nonetheless, with the current state of knowledge about adhesion it remains a riddle to be solved. With the beginning of Era of Big data and Deep learning models whose implementations in many domains yielded state of the art results. In 2018 Rail safety and standards board announced 300,000 euros for encouraging the researchers to develop data driven solutions for solving the adhesion riddle.

1.2. DATA DRIVEN WORK

Only few examples of implementation of data driven solutions for solving adhesion have been reported up till now, Castillo *et al.* (2016) have used Artificial Neural Networks (ANNs) to estimate adhesion states in an ABS system. Li, Feng, and Wei (2015) have used optimized Recursive ANN with the focus on optimizing the utilization rate of adhesion available. Gajdar, Rudas, and Suda (1997) have used classical BP Neural network to estimate friction coefficient of the rail wheel which is mathematically different but closely associated with adhesion coefficient. Data is extracted through a simulated wheel rail condition. Malvezzi *et al.* (2013) have used Neural networks in combination with an adhesion model to identify adhesion coefficient. Experimental data is acquired during tracking and braking tests. Zhang *et al.* (2017) have used a deep auto sparse encoder to estimate adhesion status of a locomotive however the parameters chosen for training the Autoencoder are not directly measurable and have been acquired through an SMC estimator, so the accuracy of the data driven model heavily depends upon the SMC observer. From the above mentioned data driven approaches are experimentally verified and show viable results (Castillo *et al.*, 2016; Li *et al.*, 2015). However, very few approaches solely focus on adhesion identification and ignore the complexities associated with it, also any data driven algorithm should be able to identify adhesion conditions from directly measurable parameters rather than relying on estimated parameters (Zhang *et al.*, 2017; Gajdar *et al.*, 1997).

In this research, a data driven method based on DNNs is employed, which solely focuses on adhesion condition identification. Main objective of this study has been to produce a self-reliant data driven solution to the adhesion riddle. Data driven model should predict adhesion condition from noisy sensor data, without relying on estimation theory to get clearer data at their input. This will enhance possibility of integration of this model in an onboard condition monitoring system.

Rest of the paper proceeds as follows. Section 2 and 3 describe the simulation of the wheel-rail interaction and different adhesion conditions, Section 4 includes data recording and dataset preparation, Section 5 explains the process of implementation of DNNs, Section 6 compares and discusses the results obtained with different data driven algorithms, and Section 7 concludes the paper.

2. SIMULATING WHEEL-RAIL INTERACTION

In order to acquire data for the training of DNNs, a Wheel-rail contact model which manifests the running of a wheel on a track in different adhesion conditions was needed so a nonlinear model of a solid axle Wheel-set established in Hussain, Mei, and Ritchings (2013) was simulated in MATLAB/Simulink.

Figure 1 shows diagram of the modeled solid-axle Wheel-set, F_{xr} and F_{xl} are longitudinal creep forces, ω_r and ω_l are angular velocities of right and left wheel respectively and V is longitudinal velocity of the Wheelset under the effect of creep forces.

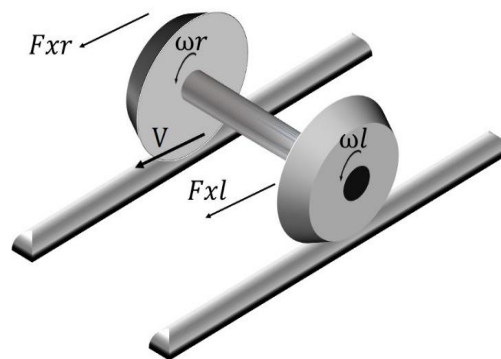


Figure 1. Solid Axle Wheel-set on a straight track.

Below are the equations of motions of this Wheel-set model.

$$V = \left(\left(\frac{1}{M_v} \right) * \int (F_{xr} + F_{xl}) dt \right) + v_0 \quad (1)$$

$$\omega_R = \left(\left(\frac{1}{I_r} \right) * \int (T_m - T_s - T_r) dt \right) + \omega_0 \quad (2)$$

$$\omega_L = \left(\left(\frac{1}{I_l} \right) * \int (T_s - T_l) dt \right) + \omega_0 \quad (3)$$

$$\theta_s = \int (\omega_R - \omega_L) dt \quad (4)$$

Here θ_s is angle of twist when both wheels have different angular velocities, where as F_{xr} , F_{xl} creep forces of both wheels and can be calculated as:

$$F_{xr} = \frac{F_r * \gamma_{xr}}{\gamma r} \quad (5)$$

$$F_{xl} = \frac{F_l * \gamma_{xl}}{\gamma l} \quad (6)$$

Where F_r and F_l are total creep forces of left and right wheel calculated from adhesion coefficient μ and normal load N .

$$F_r = \mu_r * N \quad (7)$$

$$F_l = \mu_l * N \quad (8)$$

μ_r and μ_l are adhesion coefficients representing the different adhesion levels, these coefficients are calculated by Polach model (Polach, 2005).

3. SIMULATING DIFFERENT ADHESION CONDITIONS

In order to manifest the different adhesion conditions or adhesion coefficient values in (8) (9) on which the Wheel-set model was to run, we used Polach model to generate different adhesion conditions (Polach, 2005), Figure 2 shows different adhesion conditions on which the simulations are carried out.

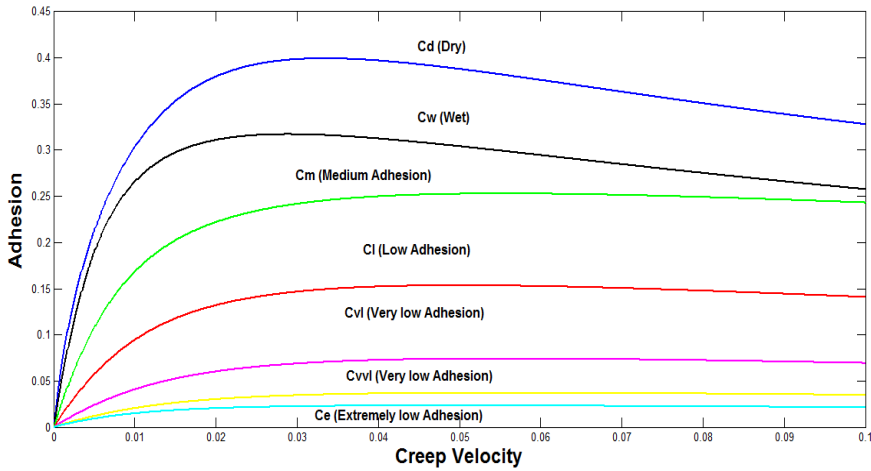


Figure 2. Adhesion curves from Polach Model.

Here are the equations used in simulating different adhesion conditions based on Polach model (Polach, 2005).

$$\mu = \mu_0 (1 - A) e^{(-BVc)} + A \tag{9}$$

$$F = \frac{2Q\mu}{\pi} \left(\frac{\varepsilon}{\varepsilon^2 + 1} + \arctan \varepsilon \right) \tag{10}$$

$$\varepsilon = \frac{2 C \pi a^2 b}{3 Q \mu} S \tag{11}$$

In (9) μ is adhesion coefficient and A, B are curve tuning parameters whereas V_c is creep velocity, In (10) F is creep force, Q is wheel load, and ε is gradient of tangential stress in the area of adhesion which is calculated in (11) from a, b half axes of contact ellipses, contact shear stiffness C, and total creep s.

4. DATA RECORDING AND PREPARTION OF DATASET

On each of the contact conditions manifested by Polach model, Wheel-set was run for 5 minutes and parameters which are practically measurable through sensors namely, longitudinal velocity V , angular velocity of both wheels and the integration of the difference between them θ_s was generated using simulations in presence of constant track disturbance.

Figure 3 shows the one sample of theta recorded during 0.63 seconds of simulation on 7 different adhesion conditions, length of the sample 0.63 seconds is chosen with a trial and error process, lower sample size being more preferable as it would increase the response time of the DNN, however 0.63 was identified as limit without deteriorating performance of the model, magnitude of θ_s seems to lower with decrease in adhesion, but frequency seems to increase. However signal's characteristics are not part of discussion here as data driven algorithms will be feed individual values of θ_s and should classify and identify them as different signals

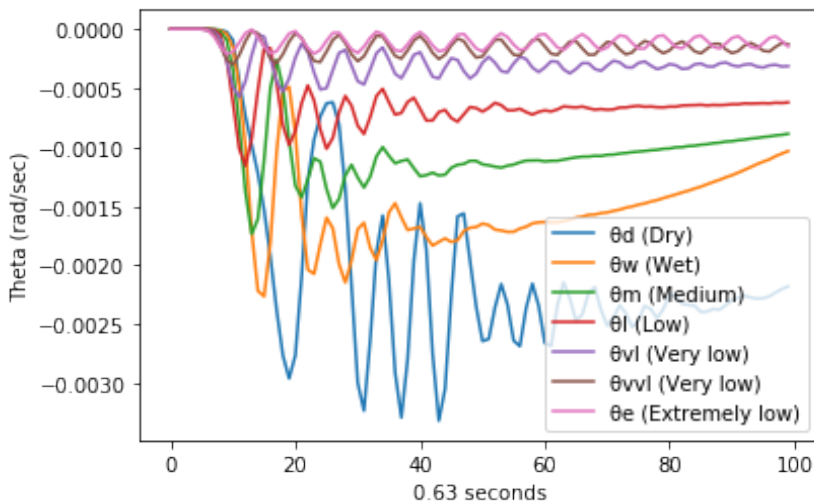


Figure 2. Single Sample of θ_s on different Adhesion conditions.

Keeping in mind the robustness required in the model, recorded parameters were analyzed and θ_s was chosen as the parameter to train the DNNs as it is a scaled and meaningful representation of angular velocities of both wheels, hence reducing the no of features and simplifying the process of convergence of DNN.

Recorded sequence data was converted into a comma separated file (CSV) and preprocessed using Python programming in order to be used for training the DNN, Sequences of θ_s recorded over different adhesion conditions were spilt on equal interval of 0.63 seconds, labeled and stacked into a matrix of $n \times m$ shape as shown below:

$$data = \begin{bmatrix} \theta_{11} & \theta_{12} & \square & \theta_{1m} \\ \theta_{21} & \theta_{22} & \square & \theta_{2m} \\ \square & \square & \square & \square \\ \theta_{n1} & \theta_{n2} & \square & \theta_{nm} \\ 0 & 1 & \square & 6 \end{bmatrix} \tag{12}$$

Here n=100 no of features in a single sample of theta for 0.63 seconds, where m= 3290 total no of examples from all classes, 470 from each class or adhesion condition. Last row represents labels associated with each class or adhesion condition [0, 6]. Data matrix was shuffled and spilt using 70/30 rule into Train set which was to be used for training the DNN and Test set on which the validation of the trained model was done.

5. DNN IMPLEMENTATION

DNN was implemented using the Tensorflow and Keras, High End APIs of Python used for implementation of Deep learning models. We went through various Hyper parameters configurations of the DNN, no of neurons in each layer, no of layers, activation functions, loss functions, regularization parameters, and optimizers. Figure 4 shows the configuration of the model which exhibited the highest validation accuracy and less generalization error.

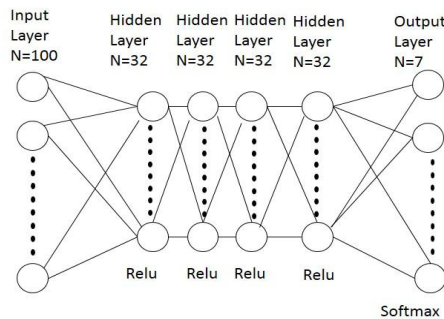


Figure 2. Block diagram of DNN.

The four densely-connected (each input is connected to every neuron in succeeding hidden layer) each with 32 neurons with a relu (rectifying linear unit) activation were used. Each neuron represents a matrix multiplication and summation operation, every neuron a_n^l superscript l representing the no of layer and n is number of neuron are calculated as in (13) and then relu activation function is applied as in (14).

$$a'_n = \sum_{n=1}^N w_n \theta_n \quad (13)$$

$$f(a) = \begin{cases} 0 & \text{if } a < 0 \\ a & \text{if } a \geq 0 \end{cases} \quad (14)$$

In (13) w_n represents weights of the NN which are randomly initialized from Gaussian distribution and are optimized over iterations using an optimization function reported in Kingma and Ba (2015). Equation (15) shows the way Adam optimizer way of updating its weight, it can be thought of a combination between stochastic gradient descent and RMSprop (Root Mean Square proportion).

$$w_t = w_{t-1} - \eta \frac{\widehat{m}_t}{\sqrt{\widehat{v}_t + \varepsilon}} \quad (15)$$

Here η is the step size like learning rate, \widehat{m}_t and \widehat{v}_t are exponentials of moving averages of weights. Loss function used here is categorical cross entropy which is one of the most used ways of loss calculation in multiclass classification problems. Softmax at the output is a stack of multiple sigmoid activation functions in order to calculate the probability of each class.

6. RESULTS AND DISCUSSION

Having formulated problem as a multiclass classification one, multiple data driven algorithms were tried, Table 1 shows validation accuracy achieved by different data driven algorithms.

Table 1. Accuracy of different data driven algorithms on our dataset.

Algorithm	Accuracy
Decision Tree	80%
Logistic Regression	19%
K-nearest Neighbours	62%
Support Vector Machine-Polynomial	71%
Support Vector Machine-Gaussian	35%
Deep Neural Network (DNN/MLP)	91%

Table 1 shows that DNN outperformed traditional data driven methods by a large margin on our sequence classification task and achieved reasonable accuracy to work with. Keeping in mind the future real time implementation, and reasonable accuracy achieved on a simpler

model other Deep Models such as Convolutional and Recurrent Neural Networks are not compared here as those may be bulky and an overkill in terms of computational cost of those models.

Confusion matrix in Figure 5 shows the accuracy achieved by DNN on individual classes (adhesion conditions) and Table 2 show accuracy as well as precision and recall of the DNN achieved by model.

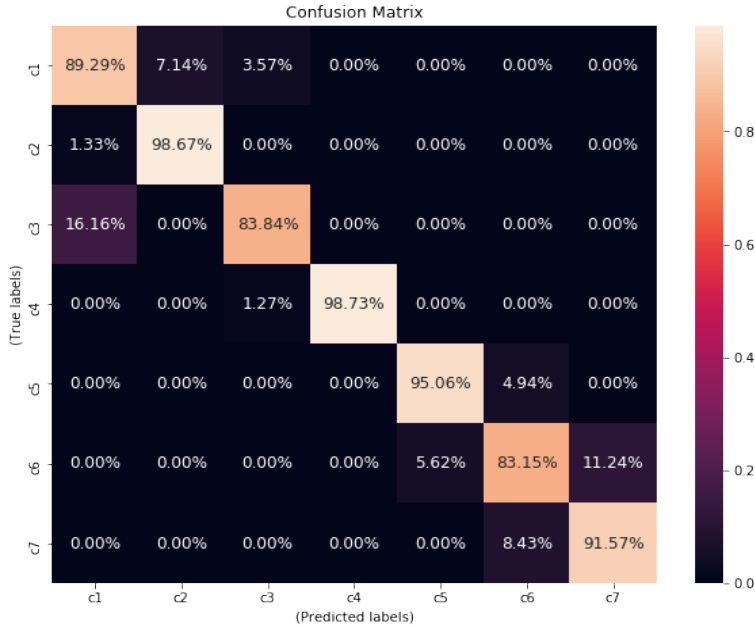


Figure 2. Confusion Matrix (DNN).

Table 1. Classification report (DNN).

	Precision	Recall	F1-score
Cd (Dry)	0.82	0.89	0.85
Cw (Wet)	0.93	0.99	0.95
Cm (Medium)	0.95	0.84	0.89
Cl (Low)	1.00	0.99	0.99
Cvl (Very low)	0.94	0.95	0.94
Cvvl (Very low)	0.87	0.83	0.85
Ce (Extremely low)	0.88	0.92	0.90
Accuracy			0.91
Macro average	0.91	0.91	0.91
Weighted average	0.91	0.91	0.91

7. CONCLUSION AND FUTURE WORK

This paper demonstrates the potential of data driven model a Deep Neural Network solely focusing on identification of contact condition or adhesion condition. It was observed in the results that deep neural networks performed well on the task of inferring adhesion condition from directly measurable parameters and achieved 91% accuracy on data. However experimental data collection in railway is expensive and open access datasets are a rarity, providing labels to that data is another problem in its entirety. Point to the future is to attempt further on adhesion riddle using data driven model that may overcome the hurdles of less availability of data and should be able to work in an unsupervised setting.

ACKNOWLEDGMENT

We acknowledge the support of the ‘Haptics, Human Robotics, and Condition Monitoring Lab’ established in Mehran University of Engineering and Technology, Jamshoro under the umbrella of National Center of Robotics and Automation funded by the Higher Education Commission (HEC), Pakistan.

REFERENCES

- A Timeline of Neglect: Train Incidents in Pakistan.* (2019, October 31). The Express Tribune. https://tribune.com.pk/story/2090607/1-timeline-neglect-train-incident-pakistan/?__cf_chl_jschl_tk__=67b51aba1d9fc784047d7a9762c146c04cc3533c-1580365929-0-AcF2YBpMeT3YhfRtd7XjbqIHW9ee6w56yPvQP0YDpiMvDN3xSyqEQ9ABBwqUJBVwBWTQynuWfHkSJxYVI48aCK-CgNuuVfT1N3OQ
- Bibi, R., Chowdry, B. S., & Shah, R. A.** (2018). PSO Based Localization of Multiple Mobile Robots Employing LEGO EV3. In *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, Pakistan.* <https://doi.org/10.1109/ICOMET.2018.8346452>
- Carl, B., & Brook, S.** (1985). United States patent. *Geothermics*, 14(4), 595–599. [https://doi.org/10.1016/0375-6505\(85\)90011-2](https://doi.org/10.1016/0375-6505(85)90011-2)

- Castillo, J. J., Cabrera, J. A., Guerra, A. J., & Simón, A.** (2016). A Novel Electrohydraulic Brake System with Tire-Road Friction Estimation and Continuous Brake Pressure Control. *IEEE Transactions on Industrial Electronics*, 63(3), 1863–1875. <https://doi.org/10.1109/TIE.2015.2494041>
- Gajdar, T., Rudas, I., & Suda, Y.** (1997). Neural Network based estimation of friction coefficient of wheel and rail. In *Proceedings of IEEE International Conference on Intelligent Engineering Systems, Proceedings, Budapest, Hungary*, 315–318. <https://doi.org/10.1109/ines.1997.632437>
- Hussain, I.** (2012). *Multiple model based real time estimation of wheel-rail contact conditions*. (PhD thesis). University of Salford. <http://usir.salford.ac.uk/id/eprint/38094/>
- Hussain, I., Mei, T. X., & Ritchings, R. T.** (2013). Estimation of wheel-rail contact conditions and adhesion using the multiple model approach. *Vehicle System Dynamics*, 51(1), 32–53. <https://doi.org/10.1080/00423114.2012.708759>
- Kingma, D. P., & Ba, J. L.** (2015). Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, San Diego, 1–15*. <https://arxiv.org/abs/1412.6980>
- Li, N., Feng, X., & Wei, X.** (2015). Optimized adhesion control of locomotive airbrake based on GSA-RNN. In *7th International Conference on Intelligent Human-Machine Systems and Cybernetics, Hangzhou, China, 2, 157–161*. <https://doi.org/10.1109/IHMSC.2015.222>
- Malvezzi, M., Pugi, L., Papini, S., Rindi, A., & Toni, P.** (2013). Identification of a wheel-rail adhesion coefficient from experimental data during braking tests. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 227(2), 128–139. <https://doi.org/10.1177/0954409712458490>
- Olofsson, U.** (2009). Adhesion and friction modification. In *Wheel-Rail Interface Handbook*. Woodhead Publishing Limited. <https://doi.org/10.1533/9781845696788.1.510>
- Polach, O.** (2005). Creep forces in simulations of traction vehicles running on adhesion limit. *Wear*, 258(7–8), 992–1000. <https://doi.org/10.1016/j.wear.2004.03.046>

- Shah, A. A., Chowdhry, B. S., Memon, T. D., & Kalwar, I. H.** (2020). Real Time Identification of Railway Track Surface Faults using Canny Edge Detector and 2D Discrete Wavelet Transform. *Annals of Emerging Technologies in Computing (AETiC)*, 4(2), 53–60. <https://doi.org/10.33166/AETiC.2020.02.005>
- Shrestha, S., Wu, Q., & Spiriyagin, M.** (2019). Review of adhesion estimation approaches for rail vehicles. *International Journal of Rail Transportation*, 7(2), 79–102. <https://doi.org/10.1080/23248378.2018.1513344>
- Spiriyagin, M., Cole, C., Sun, Y. Q., McClanachan, M., Spiriyagin, V., & McSweeney, T.** (2014). *Design and simulation of rail vehicles*. CRC Press. <https://doi.org/10.1201/b17029>
- Spiriyagin, M., Wolfs, P. J., Cole, C., Spiriyagin, V., Sun, Y. Q., & Mcsweeney, T.** (2016). Theoretical investigation of the effect of rail cleaning by wheels on locomotive tractive effort. In *CORE 2016, Maintaining the Momentum, Conference on Railway Excellence, Melbourne, Victoria*. <https://trid.trb.org/view/1468498>
- Ward, C. P., Goodall, R. M., Dixon, R., & Charles, G. A.** (2012). Adhesion estimation at the wheel-rail interface using advanced model-based filtering. *International Journal Of Vehicle Mechanics and Mobility*, 50(12), 1797–1816. <https://doi.org/10.1080/00423114.2012.707782>
- Zhang, C., Sun, J., He, J., & Liu, L.** (2017). Online Estimation of the Adhesion Coefficient and Its Derivative Based on the Cascading SMC Observer. *Journal of Sensors*, 2017, Article ID 8419295. <https://doi.org/10.1155/2017/8419295>

