Data-driven fault diagnosis of bogie suspension components with on-board acoustic sensors

Felix Sorribes-Palmer[1], Bernd Luber[2], Josef Fuchs[3], Thomas Kern[4] and Martin Rosenberger[5]

[1,2,3] *Virtual Vehicle Research GmbH, Inffeldgasse 21a, 8010, Graz, Austria*
felix.sorribespalmer@v2c2.at
bernd.luber@v2c2.at
josef.fuchs@v2c2.at

[4,5] *Siemens Mobility Austria GmbH, Eggenberger Str. 31, 8020 Graz, Austria*
thomas.a.kern@siemens.com
martin.rosenberger@siemens.com

ABSTRACT

This paper proposes a data-driven approach for fault-detection and isolation of bogie suspension components with on-board acoustic sensors. The fault detection technique is based on the acoustic emissions variation due to structural modal coupling changes in the presence of faulty components. A suspensions component failure introduces an imbalance into the system, resulting in dynamics interferences between the motions. These interferences modify the energy introduced into the system as well as its acoustic emissions. The unknown arbitrary track irregularities generate together with a variable train speed a random nonstationary vehicle excitation. Speech recognition techniques were used to generate features that consider this phenomenon. Frequency spectrums were analysed in different operating conditions to design efficient features. The robustness of the methodology was verified with data from two different test measurement campaigns on a test ring, where the influence of the sensor locations for the fault classification process was studied. The proposed methodology achieved good fault classification performance on the investigated use cases, removed dampers and 50% damper degradation on primary and secondary vertical suspension.

1 INTRODUCTION

The main goal of condition based maintenance (CBM) is to support maintenance practitioners to make right and fast decisions as well as analysing the system performance based on time series data acquired from different sensors depending on component functionality (Atamuradov, Medjaher, Dersin, Lamoureux, & Zerhouni, 2017). Interval-based maintenance inspections of vehicle components can lead to unexpected failures and reduced availability of the vehicles. Through CBM concepts with on-board monitoring systems for monitoring chassis components, the availability of the vehicles can be decisively increased through precise knowledge of the component status. Suspension component faults influence the vehicle dynamic behaviour, which can be monitored using different on-board sensors. A sketch of a common approach with inertial sensor used on vehicle suspension monitoring is shown in Figure 1 (Zoljic-Beglerovic, Golkani, Steinberger, & Horn, 2018; Zoljic-Beglerovic, Stettinger, Luber, & Horn, 2018). Primary and secondary suspension condition can be estimated using inertial sensors like gyroscopes and accelerometers which are placed near dampers and springs to estimate their condition. This implies a considerable number of sensors, that must be robust enough to withstand the high fatigue loads at the less damped components like axle boxes.

For this reason, inertial sensors which are the most used in on-board health monitoring systems for railway vehicles are being located over the first or second suspension level of the vehicle (bogie and/or carbody), were induced accelerations are smaller. But this results in a more challenging task of fault detection. With acoustic sensors this limitation could be overcome as they can be installed far from the structural loads and still register the fault acoustic emissions. Furthermore, high frequency acoustic emissions can be detected before vibration response is detectable by inertial sensors in the same location, increasing fault detection horizon.

.

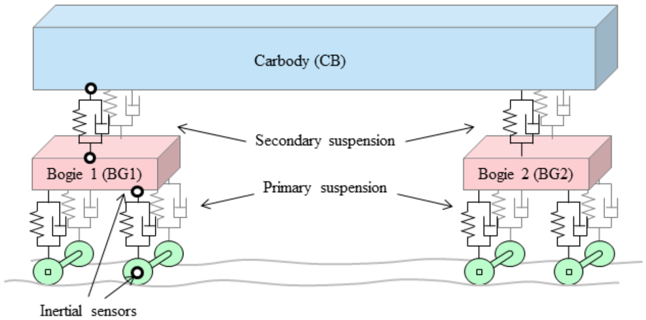


Figure 1 Example of condition monitoring with on-board inertial sensors.

1.1 State of the art of railway condition monitoring

A review of health monitoring technologies on railway industry is presented in (Chong, Lee, & Shin, 2010), where wayside detection methods, advanced integrated sensor methods, operating principles and functions are analysed. Bogie performance standards from a regulatory perspective and existing technologies that are currently in use in railroad revenue service are summarized in (Shahidi, Maraini, Hopkins, & Seidel, 2014; Ward, Goodall, Dixon, & Charles, 2013). Reducing estimation uncertainty of track and train component condition is critical for a robust and reliable prognostics. This can be achieved with on-board monitoring of rail and bogie components.

An overview of different fault detection and diagnosis approaches for health monitoring and a comparison between model-based and data-driven approaches with the advantages and disadvantages of each one is presented in (Tidriri, Chatti, Verron, & Tiplica, 2016).

A particular model-based technique is the observer-based method, which is an effective component failure detecting system but presents difficulties when applied to systems which are subjected to unknown disturbances and model uncertainties (Ngigi, Pislaru, Ball, & Gu, 2012). Kalman filter (KF) is one of the most popular methods used to model linear system with additive Gaussian noise for estimation and prediction purposes (Jesussek & Ellermann, 2015; Tsunashima, Mori, Tsunashima, & Mori, 2010; Zoljic-Beglerovic, Golkani, et al., 2018; Zoljic-Beglerovic, Stettinger, et al., 2018). Extended Kalman filter (EKF), Unscented Kalman filter (UKF) and Particle Filters (PF) offer accurate estimation results for nonlinear stochastic systems (P. Li et al., 2007; P. Li & Goodall, 2004). More examples of model-based applications on train dynamics monitoring are presented in (Bozzone, Pennestrì, & Salvini, 2011; Liu, Alfi, & Bruni, 2016; Ward et al., 2010). For systems that are dynamically complex and/or nonlinear, model-based approaches may lead to the use of high-order and/or linearized multiple models which can be difficult to implement. A component failure (e.g. a damper) in either of the suspensions will introduce and imbalance into the system, resulting in dynamics interferences between the motions (Mei & Ding, 2009).

Data-driven approaches attempt to build degradation models using condition monitoring data collected via installed sensors to predict future health state instead of building physical models. They can detect faulty situations faster, are easier to implement and require less a priori knowledge. They are suitable for complex and large-scale systems where the developing and validation of a model becomes more difficult (Atamuradov et al., 2017). The success of data-driven models depends on the quality of the training data and the understanding of the problem, which can be given by an expert or by a model-based approach. Applications examples of data-driven models on bogie components monitoring can be found on (C. Li, Luo, Cole, Spiryagin, & Sun, 2017; Shahidi, Maraini, & Hopkins, 2016; Shahidi, Maraini, Hopkins, & Seidel, 2015).

Hybrid approaches relax the need of accurate models of complex systems and use convenient statistics simplifying the monitoring task (Atamuradov et al., 2017). An example of vehicle dynamics monitored using hybrid approaches is presented in (Gasparetto, Alfi, & Bruni, 2013).

A method to monitor the running stability in a high-speed railway bogie and detect different faults in the critical components to vehicle stability, particularly wheel wear and degradation of yaw dampers is investigated in (Gasparetto et al., 2013). Three different wheel profiles and four different yaw damper conditions are considered in the classification. Prony method is used to identify the characteristic exponents of the system combined with random decrement technique to approximate the auto-correlation of the acceleration signals; the residual stability margin and the hunting mode frequency are used by the classification algorithms to estimate the condition of the yaw dampers and of wheel-rail conicity.

1.2 Acoustic sensors in CBM application

Acoustic emissions are commonly used to monitor components with high maintenance cost like: axle cracks, axle bearings (Amini, Entezami, & Papaelias, 2016), engine valves (Ali, Hui, Hee, & Leong, 2018), diesel injection systems (Elamin, 2013), railway wheels (Anastasopoulos, Bollas, Papasalouros, & Kourousis, 2010), rails ((Bergseth, Höjer, Lyu, Nilsson, & Olofsson, 2019; Jensen, Chauhan, Haddad, Song, & Junge, 2015; Lanza di Scalea et al., 2017; Thompson et al., 2018)), etc.

Vehicle structure components act as signal filters for inertial sensors, wave-propagation attenuation at high frequency is higher in the structure than in air, due to geometric spreading of the wave front, internal friction, dissipation of energy into adjacent media and velocity dispersion (Pollock, 2018). When faults are detectable at low frequencies structural vibrations, it could be already too late to conduct a preventive action. The weak abnormality features of early faults are generally masked by background noise or other interference making them difficult to detect (Gong et al., 2018).

Installing acoustic sensors on the carbody minimizes their exposure to fatigue loads and enhances the overall robustness of the application extending the useful life of the monitoring system. At the same time, this remote location characteristic results in the challenge of pinpointing which exact component may be malfunctioning. Furthermore, a major drawback is the difficulty to perform validation experiments and simulations. To obtain reliable noise information, a real train with emerging or advanced faults must be instrumented and operated ((Bernal, Spiryagin, & Cole, 2019)).

Other bogie component faults like wheel profile wear, rolling contact fatigue, wheel flats and wheel corrugation can cause track and vehicle component damage ((Nielsen & Johansson, 2000)). At the same time, they can influence vehicle dynamics and introduce noise into the monitoring process of suspension components with acoustic sensors. An on-board approach to monitor the wheel wear state studying the vehicle dynamics with inertial sensors is proposed to elongate the inspection intervals in (Luber, Müller, Sorribes-Palmer, Pietsch, & Six, 2018). Worn rail profiles can be detected analysing the noise emitted from contact between the wheel and the rail, using two microphones located into the inner and outer wheel-rail contact ((Höjer, Bergseth, Olofsson, Nilsson, & Lyu, 2016)). A sensor setup on the axle box for wheel defect detection and bearing monitoring using accelerometers and acoustic emission sensors is presented in (Amini, 2016).

Sensors	Advantages	Disadvantages
Inertial	Capture vehicle dynamics and structural modes Easier to be used in simulation validation and parameter influence study	Sensors on axle-box and bogie frame are exposed to higher fatigue loads
Acoustic	Capture acoustic modes Higher frequencies can be analysed to maximize the fault detection horizon Sensors exposed to smaller vibration loads Detect several faults with just one sensor	Acoustic emissions from faulty components are mixed with exterior noise sources (aerodynamic noise, other faults; trains, bridges, tunnels, other faults, etc) Enough energy must be introduced into the system in the fault frequency range High excitation uncertainty

Table 1. Pros and cons of inertial and acoustic sensors for bogie component monitoring.

The main objectives of the present work are:

- perform condition monitoring of train vertical suspension dampers with on-board acoustic sensors
- present a framework for fault detection and isolation (diagnosis) of bogie components
- find the right operating conditions to detect the expected difference between normal and faulty damper to perform classification
- validate and verify robustness of the methodology and explore capabilities of acoustic sensors for fault detection applications

The paper is organised as follows: in Section 2, the fault-detection and isolation (FDI) approach is introduced. In Section 3, the test data from experiments, operating conditions and targeted faults are described. Model classification results are presented and discussed in Section 4. Finally, conclusions are drawn in Section 5.

2 PROPOSED FDI APPROACH

Acoustic emissions from different vertical suspension faults were monitored with on-board acoustic sensors and processed to extract features and generate data-driven models to perform FDI. The main steps performed on the data-driven models' generation are sketched in Figure 2.

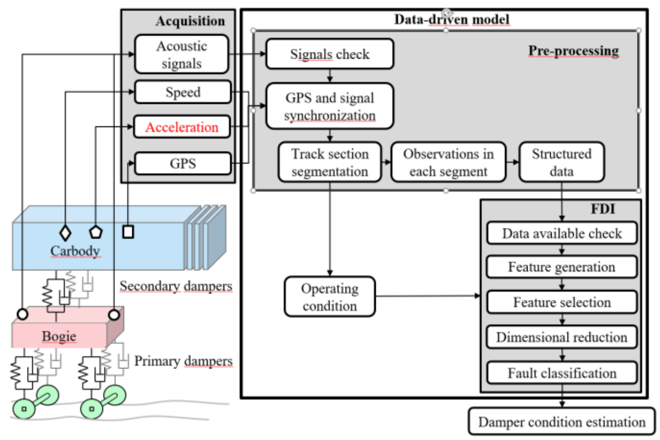


Figure 2. Main steps of the data-model generation for FDI of bogie components.

In the pre-processing phase, signal unit check, outlier removal (sensor malfunction/ wrong measurements), handle missing values and filters are performed. Also signal synchronization with GPS data for later segmentation on different track section types is carried out. All the passages through the track sections are considered as observations and organized in a structured data frame.

Inside the feature generation block, noise identification for signal denoising with spectrum subtraction are performed. Separation techniques between harmonic and percussive signal are considered to detect impact hearing events and modal excited response.

Generally, signal-to-noise ratio increases with increasing recording time and decreases with increasing test speed, as a result of the increased standard deviation of the incoherent portion of the wheel-generated excitation ((Lanza di Scalea et al., 2017)). For this reason, the influence of signal length of each observation on the classification performance has been studied.

Among the features used in this work are statistical based in time domain; and in frequency domain, energy frequency bands from discrete Fourier transform (DFT) and power spectral density (PSD) are used. Due to the non-stationarity of the excitation also time-frequency features like mel frequency coefficients (MFCCs) have been used.

In a first attempt, simple features like energy content from PSD frequency bands are used to explore influencing parameters and to analyse the limitations of the features for fault classification. The probability to excite the targeted suspension mode depends on the track condition and train speed. The hypothesis that suspension dampers faults modify the energy introduced into the system, increasing the energy content at higher frequencies is assumed. The modal coupling with other structural components is also influenced and with-it the acoustic emissions. This effect is chased to be retained into features for fault classification. Keeping this in mind, this work proposes to monitor also higher acoustic modes to consider cases in which excitation energy could have not been introduced in the fault targeted frequency band but in a coupled mode of higher frequencies. PSDs ratio between faulty and normal dampers acoustic signal were analysed to find the frequency bands at which faulty behaviour could be detected.

Before feature selection and after outlier removal a robust scaler was applied. To find the most informative features, different methods of feature selection were used: Filter (SelectKBest, reliefF), wrapper (recursive feature elimination, sequential feature selection) and embedded methods (decision trees). The combination of these methods, using first filters and later wrappers, provided the best results to find the most efficient feature pool.

The hypothesis that similar excitation energy is introduced on the normal fault observations and faulty ones is assumed. Also, the hypothesis that in-service trains will travel at similar running speeds and experience similar track quality as on the test ring is assumed. In case that track quality is much different a proper rescaling must be performed. Overall and specific data-driven models were created for the different speed ranges and track sections, under the assumption that track quality is similar in all sections.

The amount of available data and the imbalance in each speed range as well as the number of features selected was analysed to avoid model overfitting. Observations were grouped into similar excitations regarding track irregularities and train speed ranges with enough available data. To overcome data imbalance of faulty components, several correction methods were analysed, and different scores were considered.

The data-driven models were trained using the 70% of the available data in the operating conditions range selected and tested in the other 30%. All test sets classification scores were averaged over 10 random repeated splits, with 5-fold stratified cross-validation.

To consider the imbalance between normal and faulty conditions, normalized Gini Coefficient (NGini) was used as classification scorer. Gini Coefficient represents twice the area under the curve of receiver operating characteristic curve over the random guess (Contributors, 2019). When it is normalized with the maximum possible Gini Coefficient, considering the imbalance between classes, it can vary between 1 (100% accuracy) and -1 (0% accuracy), and 0 means random guess (50% accuracy). NGini coefficient value of 0.7 has been considered a good model, but also the number of features used (Nfeatures), and the data explained with should be also considered. To compare data-driven models' generalization for different operating conditions an indicator was created (classifier overfit factor, COF). This indicator considers the number of features used, the amount of explained data (Samples), the NGini and its standard deviation, and is given by the equation:

$$\mathrm{COF} = \left(\frac{0.5}{1 + \mathrm{Nfeatures}/\mathrm{Samples}}\right)\left(\frac{\mathrm{NGini}}{0.05 + \mathrm{std(NGini)}}\right) \qquad (1)$$

Models are considered acceptable when COF value is over 1.3 (e.g. NGini=0.68, Samples=42, Nfeatures=3 and NGini standard deviation=0.2).

3 EXPERIMENTS AND DATA DESCRIPTION

In supervised learning, measured observations from normal and faulty component status must be available to generate data-driven models. This reference data cannot be retrieved from in-service vehicles by removing dampers, as this would compromise the safety of the passengers. To obtain the required data under controlled operating conditions, experimental tests were performed on the test ring in Wegberg-Wildenrath, Germany. The track is composed by 2 curves (Cur1 and Cur2) of 1700 m length (radius 540 m), 2 straight sections (Str1 and Str2) of 550 m length and 4 transitions (Tra1, Tra2, Tra3 and Tra4) of 400 m length. Information about track irregularities was not available. To analyse the influence of the acoustic sensor locations on the fault classification performance, two test campaigns were carried out.

A test campaign (test 1) with 8 microphone sensors sampling at 19200 Hz and 1 GPS sensor sampling at 1 Hz was performed. In this case, the track postprocessing was divided into straight (2 and 4) and curved (1 and 3) sections, see Table 2.

Test 1	1	-	2	-	3	-	4	-
Test 2	Cur1	Tra1	Str2	Tra2	Cur2	Tra3	Str1	Tra4

Table 2. Wegberg-Wildenrath test ring track sections considered in both tests.

Another test camping was carried out (test 2), with 12 microphones sampling at 8000 Hz, together with 2 GPS sensors. In the second campaign, the acoustic sensors were more exposed to aerodynamic noise, making more challenging to distinguish the faulty pattern from the noise inside the signal. This required to imply more complex methods to generate features for classification. The data-driven models generated for classification were focused on the fault of primary and secondary vertical suspension dampers, which are indicated in Table 3.

Fault code	Fault description
SVD00	Secondary vertical damper removed
SVD50	Secondary vertical damper with 50% degradation
PVD00	Primary vertical damper removed
PVD50	Primary vertical damper with 50% degradation

Table 3. Suspension dampers faults considered in the classification.

3.1 Data imbalance

A clear data imbalance between normal and faulty samples can be seen in the observations count in both tests shown in Table 4. There were no available measurements of fault PVD50 in test 1.

	Available samples	
Fault code	**Test 1**	**Test 2**
Normal	3075	17325
SVD00	420	1284
SVD50	376	1132
PVD00	420	642
PVD50	0	566

Table 4. Fault observations inside test data.

The observations were grouped into speed ranges and track sections to limit excitation variability. Fault observations were not uniformly distributed in all speed ranges. An example of the observation speed distributions for each fault in straight sections (Str1 (4) and Str2 (2)) in both tests is shown in Figure 2. Observations were group into different operating conditions ranges to minimize data imbalance.

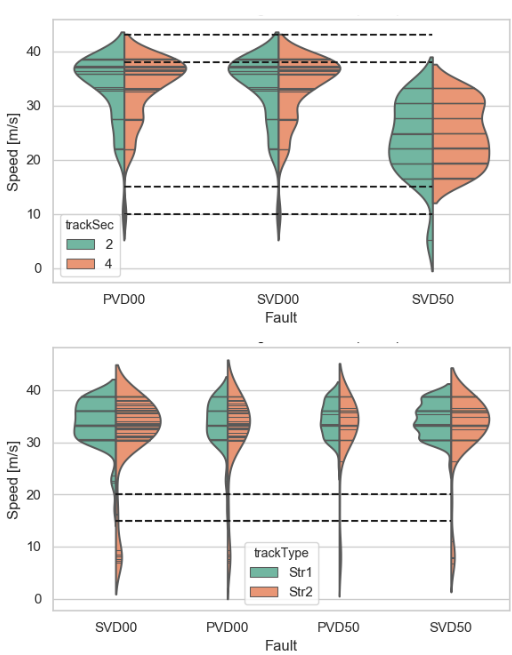


Figure 2. Train speed distributions of fault observations in straight track sections, in test 1 (top) and test 2 (bottom).

At speeds over 38 m/s it is not possible to generate a model for all faults in test 1. And in the same way, it is not possible to make a model with data from test 2 at speed under 25 m/s.

3.2 Excitation variability

The energy transferred between the track and the vehicle, is a combination of track and soil stiffness, track irregularities, train speed and vehicle dynamics. This interaction must be high enough to excite the modes that contain information about the status of the monitored component. Also, the acoustic signals captured by the on-board pressure sensors can be influenced by other environment conditions (e.g, presence of vegetation, passing through tunnels and bridges). PSDs of normal and faulty condition were analysed to find which frequencies contain information about the monitored component and to analyse the influence of the track quality in different sections. For comparative reasons, in PSD figures the red line corresponds to the average of blue PSDs (faulty condition), and black to the average of orange PSDs (normal condition). In Figure 3, in a narrow speed range, from 30 to 34 m/s, a clear difference in the spectrum between both straight sections (Str1 and Str2) can be seen at frequencies higher than 700 Hz. It can be induced that excitation over this frequency was higher in section Str1 than in Str2, as the acoustic energy radiated is higher. This will explain the better performance of the data-driven models in Str1 compared with the models created on section Str2. The acoustic radiated energy in the frequency range between 700 and 1100 Hz could contain rail acoustic emissions, which often dominates mid-frequency region between 400 and 2000 Hz (Zhang et al., 2018).

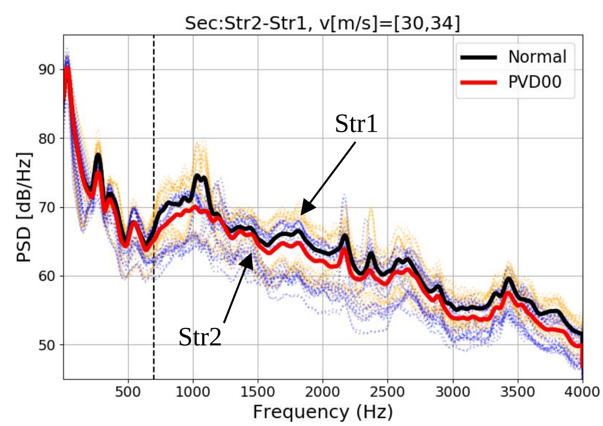


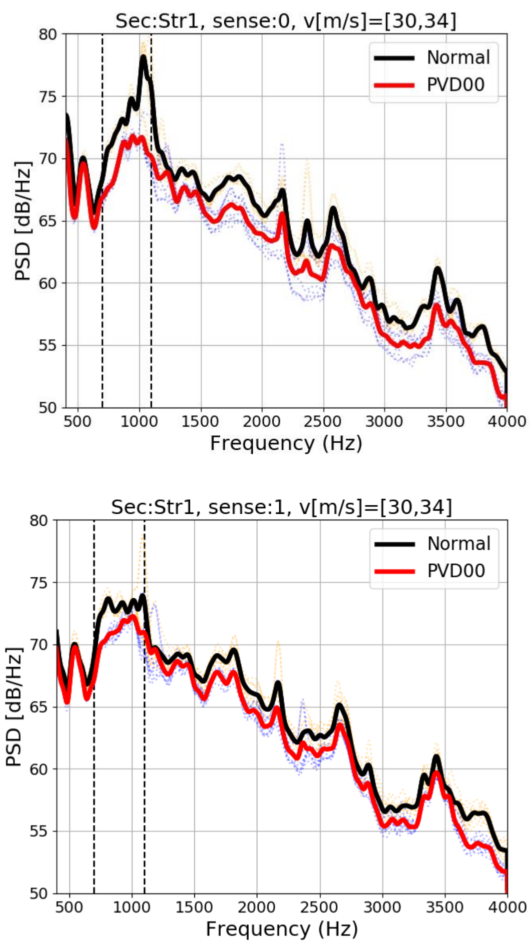
Figure 3. Fault PVD00 PSDs on straight sections (Str1 and Str2).

A trade-off between spectrums separability, considering excitation and response, between normal and faulty status and number of available samples must be achieved to create robust data-driven models.

It can be seen in Figure 4 and Figure 5 that at around 1100 Hz there is a frequency range that could contains relevant information to classify fault PVD00. There is also a clear difference in track quality when changing traveling direction sense in both straight sections. Looking in both section in sense 0, a considerable difference in energy is observed between Str1 and Str2, see Figure 4a and Figure 5a.

If the vehicle structural modes are not damped due to a failure in the bogie suspension dampers, energy from the track vehicle interaction is converted into vehicle kinematic energy, which can be seen at low frequencies. Possible relative movements between rail and wheel induced by the vehicle dynamics can increase the energy injected into the rail increasing the acoustic energy radiated from the sleeper, rail, and wheel. This effect could be observable around rail acoustic modes frequencies, like the first pinned-pinned resonance at midspan in lateral (around 300 Hz to 600 Hz) and vertical direction (around 900 to 1100 Hz) (Zhao, Wang, & Xing, 2017).

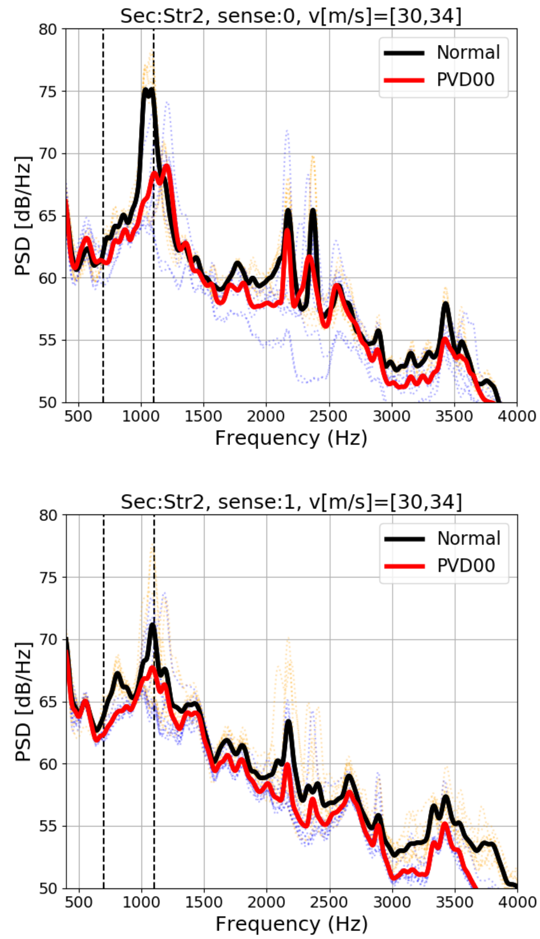
Further investigations (e.g. track settlement and surroundings as well as irregularities could be surveyed) would be needed to determine the true nature of these difference in both track sections.



a)

b)

Figure 4. Fault PVD00 PSDs on straight Str1 section in both travelling sense: a) sense 0; b) sense 1.



a)

b)

Figure 5. Fault PVD00 PSDs on straight Str2 section in both travelling sense: a) sense 0; b) sense 1.

The energy in the PSD observations from transitions and curved sections are in the same level at all frequencies as it can be seen in Figure 6. For this reason, transition sections are grouped together with curved sections as similar operating conditions to generate data-driven models.

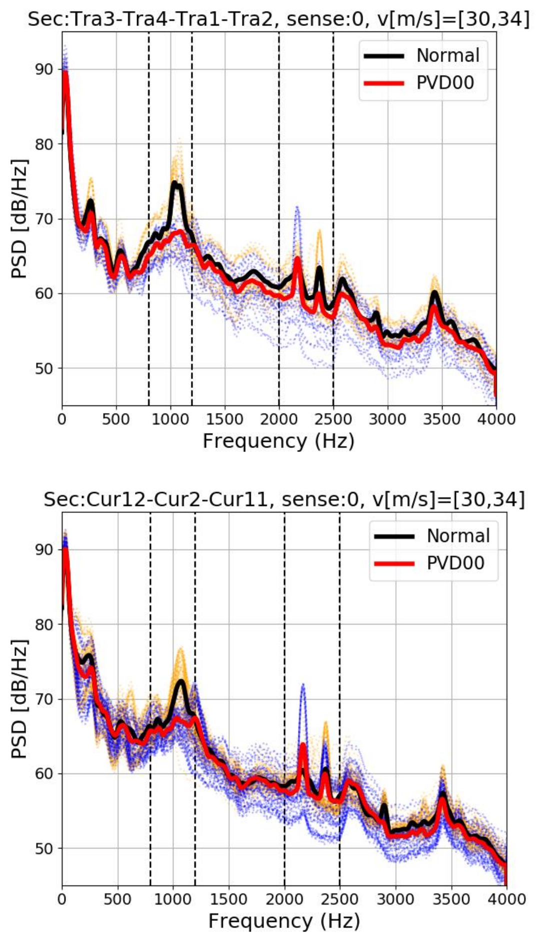


Figure 6. Fault PVD00 PSDs on transitions (top) and curves (bottom) sections at speed range 30 to 34 m/s.

4 RESULTS AND DISCUSSION

Performance classification of data driven modes from both test campaigns are presented separately to show the influence of acoustic sensor location. Different monitoring parameters have been studied: track section type, faulty wheelset location, sensor near leading of trailing edge, signal sub segmentation, number of features used for classification. Logistic regression was used as a classifier in all the models generated.

4.1 Microphones located on a central bogie-position

Fault classification performance with microphones located on a central bogie-position is less influenced by aerodynamics noise, achieving a more robust classification.

NGini and its standard deviation for fault SVD00 using microphones on a central bogie-position in different sections, straight and curved at different speed are shown in Figure 7. Additionally, the imbalance between normal and faulty observation in all generated models is also shown at the bottom in Figure 7.

Higher classification performance is achieved in curved track sections than in straight track sections when observations are grouped into similar speed ranges. Reducing excitation variability and sensor redundancy increased classification performance. The model bias is reduced at higher speeds, where more samples are available. Models for the complete speed range show better performance when using the microphone near the trailing edge in straight track sections. The classification performance is higher when the ratio between normal and faulty is closer to 50%. The speed range between 36 and 39 m/s shows higher performance in straight and curved sections.

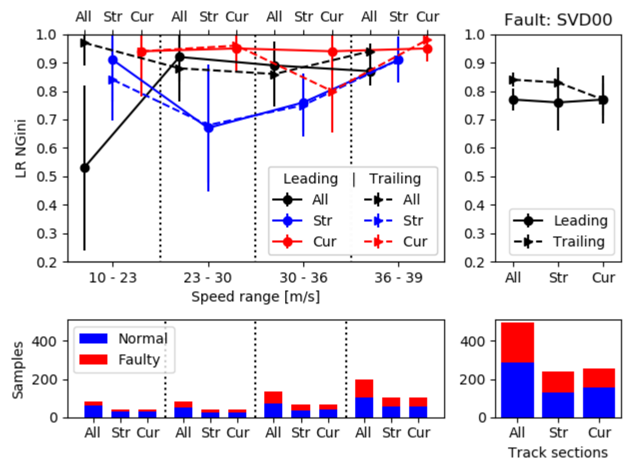


Figure 7. Fault SVD00 classification normalized Gini coefficient with logistic regression model in all sections, straight and curved at different speed ranges (top left), overall speeds using sensor located near a leading and trailing wheelset (top right).

Model performance parameter COF for fault SVD00 with logistic regression is shown in Table 5. Not all the models generated for fault SVD00 can be considered acceptable following the criterion previously defined (COF>1.3), as the model was not capable to classify the component condition with enough accuracy and low variance.

v [m/s]	COF (near leading wheelset sensor)			COF (near trailing wheelset sensor)		
	All	Straight	Curved	All	Straight	Curved
10-23	0.8	2.2	2.1	3.5	2.0	2.7
23-30	3.3	1.2	3.3	2.5	1.3	3.3
30-36	4.4	2.4	3.8	2.5	2.3	1.9
36-39	4.3	3.4	4.8	6.0	3.7	5.4
10-39	4.3	2.5	3.6	5.6	4.1	2.9

Table 5. Fault SVD00 classification performance at different speeds and track sections. Model is considered acceptable if COF>1.3.

The cross-validated accuracy and its standard deviation corresponding to the NGini showed in Figure 7 when all in all sections are considered and using the sensor near the leading wheelset is 92±2%, while using the near trailing wheelset sensor is 95±3%. The exceptions are: for all sections in the leading wheelset sensor at speeds between 10 and 23

m/s with 78±12%, in straight tracks between 23 to 30 m/s with 76±5% and in the same operating conditions using the near the trailing wheelset sensor is 85±12%.
NGini for fault SVD50 in several wheelsets at different speed ranges considering all track section types is shown in Figure 8. The classification can vary considerably depending on its location in the train. Regardless, Ngini is close to 0.7 at middle speeds when considering samples from all track sections and all speed ranges.

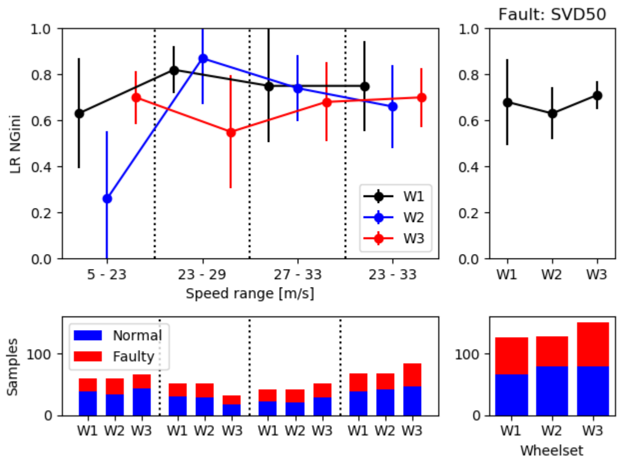


Figure 8. Fault SVD50 classification normalized Gini coefficient with logistic regression model for all sections in three wheelsets at different speed ranges (top left) and overall speeds (top right).

COF for fault SVD50 is considerable smaller than SVD00. As fault SVD50 is less severe than SVD00, more energy must be introduced into the system to detect it. This hypothesis will explain the low classification performance at low speeds, and how it grows with increasing speed. Models show overfitting at low speeds, in which many features were required to classify between faulty and normal component status.

Fault SVD50 COF (All sections)			
v [m/s]	**FWS 1**	**FWS 2**	**FWS 3**
5 – 23	**1.0**	**0.4**	**2.0**
23 – 29	**2.5**	**1.6**	**0.8**
27 – 33	**1.2**	**1.8**	**1.4**
23 – 33	**1.4**	**1.4**	**1.9**
5 – 33	**1.4**	**1.9**	**3.1**

Table 6. Fault SVD50 classification performance on different wheelset at different speeds considering observation from all sections. FWS: Faulty wheelset.

The cross-validated accuracy and its standard deviation corresponding to the NGini showed in Figure 8 using the sensor near wheelset W1 is 83±8%, while using the sensor near wheelset W2 is 85±7% and on wheelset W3 is 72±7%. The exceptions are: on W1 in the speed range between 5 and 23 m/s which is just 79±10%, and between 27 and 33 m/s which is 82±11%; on W2 between 5 and 23 m/s which is just 70±14%; and on W3 between 23 and 29 m/s which is just 75±11%.
NGini of fault PVD00 considering all sections is shown in Figure 9. These curves show how sensor redundancy can increase robustness and classification performance up to 20%. NGini in fault PVD00 remains close to 0.8 for speed ranges over 23 m/s, and best speed range interval is again the one with more available observation samples, 36 to 38 m/s.

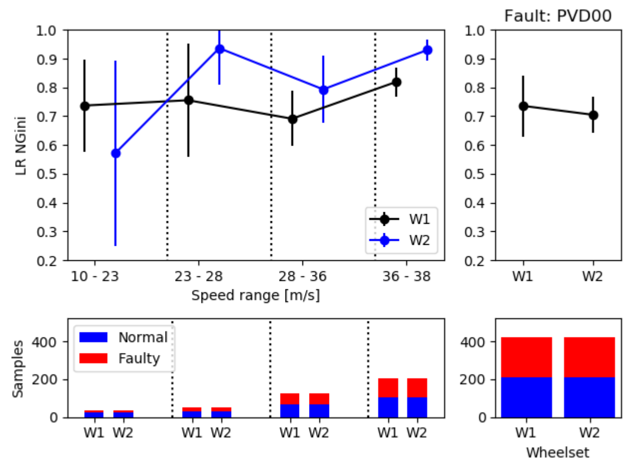


Figure 9. Fault PVD00 classification normalized Gini coefficient for logistic regression model for all sections in two wheelsets at different speed ranges (top left) and overall speeds (top right).

COF for fault PVD00 increases with the amount of data available, in this case at higher speed ranges, see Table 7. A drop in classification is observed in speed range between 28 and 36 m/s. A possible explanation could be an overlapping with other noise sources that could mask relevant fault information in the used frequencies band features.

Fault PVD00 COF (All sections)		
v [m/s]	**FWS 1**	**FWS 2**
10 – 23	**1.6**	**0.7**
23 – 28	**1.5**	**2.5**
28 – 36	**2.3**	**2.3**
36 – 38	**4.0**	**5.2**
10 – 39	**2.3**	**3.1**

Table 7. Fault PVD00 classification performance in different wheelsets at different speeds considering observation from all sections.

The cross-validated accuracy and its standard deviation corresponding to the NGini showed in Figure 9 using the sensor near wheelset W1 is 85±4%, while using the sensor near wheelset W2 is 92±2%. The exceptions on W2 in the speed range between 5 and 23 m/s which is just 77±10%.

4.2 Microphones exposed bogie-position to aerodynamic noise

Fault classification performance is more challenging when microphones exposed to aerodynamic noise. Advanced signal processing methods were required to extract useful features to achieve an acceptable fault classification.

Time length of each observation varies within track sections and train speed between 20 and 40 seconds. The observations were split into 10 second subsegments (SubSeg) as a first step to increase available data. This did not increase NGini in all cases but increased the model performance parameter COF for fault SVD00 and PVD00. The probability to introduce the required excitation energy is higher with longer time signals. This effect is not observed for faults SVD50 and PVD50. This sub segmentation (SubSeg) increased NGini and COF on fault SVD50 in straight section Str1, but did not achieve similar results with PVD50, see Table 8.

A NGini reduction in fault SVD50 is obtained when including also Str2 compared with the case considering just straight track section Str1. Reducing excitation variability grouping observation into similar track sections types improved classification, see COF of fault SVD00 in Table 8.

COF	Sub Seg	Fault	FWS	LR NGini	1*std	Track
2.8	no	SVD00	6	**0.78**	**0.09**	Cur+Tra
1.9	yes	SVD00	6	**0.80**	**0.15**	Str1
3.7	no	SVD00	11	**0.82**	**0.06**	Cur+Tra
2.4	yes	SVD00	11	**0.58**	**0.07**	Cur+Tra
3.8	no	SVD00	11	**0.87**	**0.06**	Cur
3.1	no	SVD00	11	**0.77**	**0.07**	Tra
8.0	no	SVD00	11	**0.99**	**0.01**	Str1
4.6	yes	SVD00	11	**0.98**	**0.05**	Str1
2.3	yes	SVD00	11	**0.71**	**0.10**	Str1
4.9	yes	SVD50	11	**0.92**	**0.04**	Str1
2.4	no	SVD50	11	**0.77**	**0.11**	Str1
1.4	no	SVD50	11	**0.53**	**0.14**	Str1+Str2
3.5	yes	PVD00	6	**0.63**	**0.04**	All
3.8	yes	PVD00	6	**0.69**	**0.04**	Cur+Tra
2.7	no	PVD00	6	**0.65**	**0.07**	Cur+Tra
4.3	yes	PVD00	6	**0.98**	**0.06**	Str1
6.0	no	PVD00	6	**0.99**	**0.03**	Str1
3.9	no	PVD00	11	**0.80**	**0.05**	Cur+Tra
2.4	no	PVD00	11	**0.72**	**0.10**	Cur
5.4	no	PVD00	11	**0.99**	**0.04**	Str1
8.1	yes	PVD00	11	**0.99**	**0.01**	Str1
2.6	yes	PVD50	6	**0.80**	**0.10**	Str1
5.1	no	PVD50	6	**0.96**	**0.04**	Str1
2.9	no	PVD50	6	**0.84**	**0.09**	Str1

Table 8. Faults classification performance at speed range 30-40 m/s in different track sections for sensors exposed to aerodynamic noise. NF: number of features used; %F: percentage of faulty cases.

Fault PVD50 was just possible to be successfully classified in straight section Str1. Fault SVD50 was classified successfully in straight and transitions. Fault PVD00 and SVD00 were classified in all sections, but meanwhile models COF for these two faults reached values around 8 in section Str1, it stayed around 5 for PVD50 and SVD50.

Model decision functions of all fault classifications in straight track Str1 are shown in Figure 10 and Figure 11. Under the mentioned operating conditions, first Mel frequency cepstral coefficient skewness (skew) and impulse factor (im), together with the PSD frequency band around 1105.2 Hz and spectral bandwidth (sbw) are useful features to classify between normal and faulty status for fault PVD00 and PVD50.

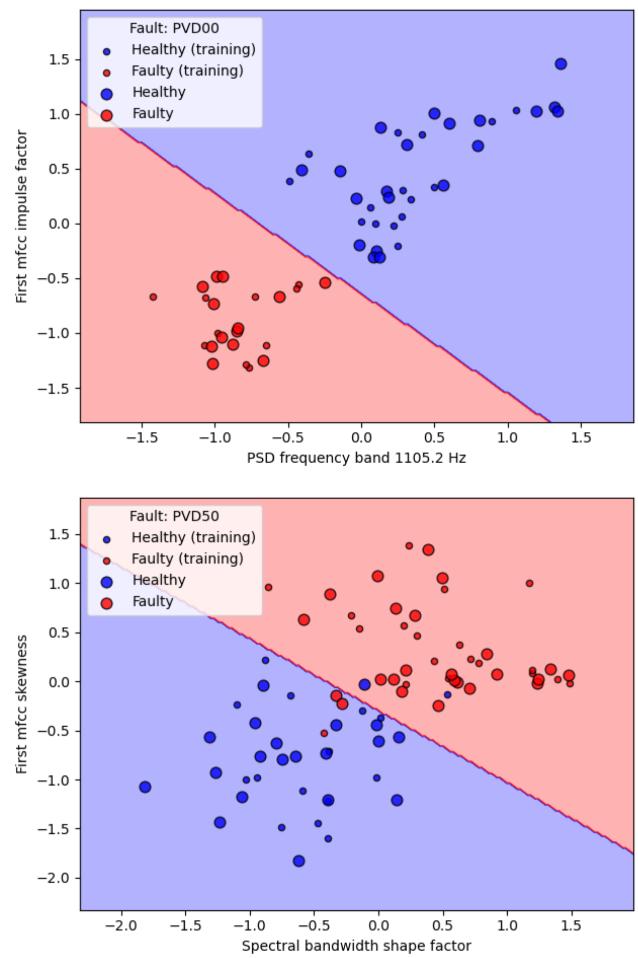


Figure 10. Decision function of primary suspension faults on straight track section Str1 using logistic regression, PVD00 (top) and PVD50 (bottom).

In the case of faults SVD00 and SVD50, the PSD frequency band around 562.8 Hz, the first and seventh Mel frequency cepstral coefficient achieved a considerable classification performance.

A small improvement could be achieved using a nonlinear classifier, in the case of 50% damping efficiency reduction faults (SVD50 and PVD50), but the present work was focused on finding proper features for classification rather than model selection and hyperparameter tuning of the best machine learning algorithm.

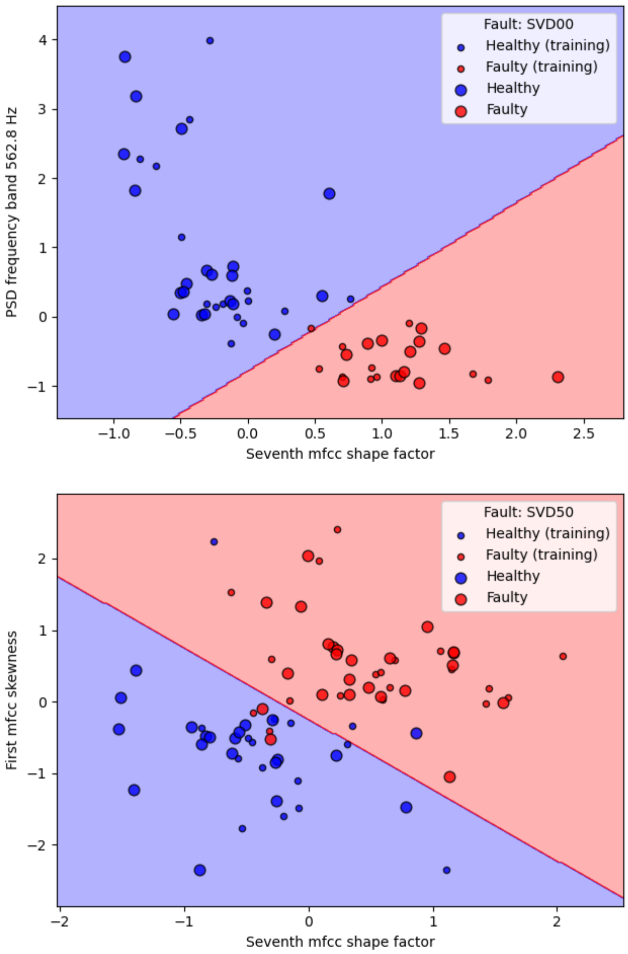
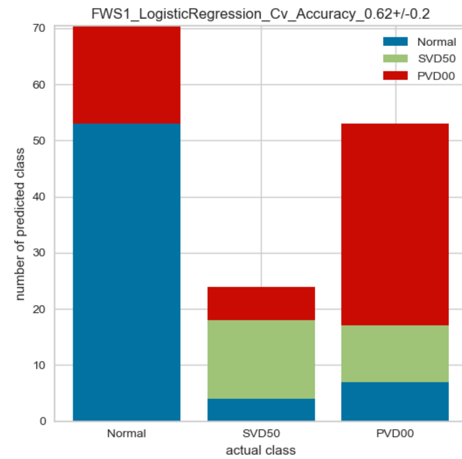


Figure 11. Decision function of secondary suspension faults on straight track section Str1 using logistic regression, SVD00 (top) and SVD50 (bottom).

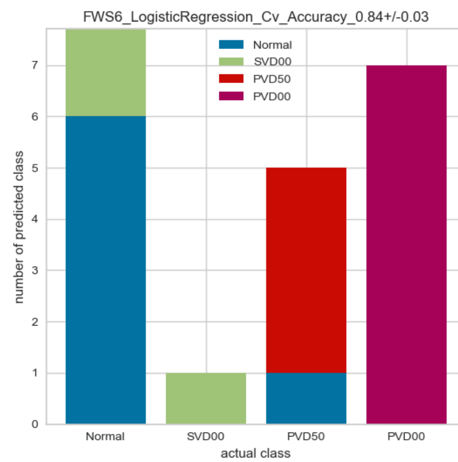
The cross-validated accuracy of the classifications for the same operating condition as the one shown in Figure 10 and Figure 11 is: 98±3% for PVD00, 92±6% for PVD50, 98±3% for SVD00 and is 90±8% for SVD50.

Figure 12 shows an intuitive representation of the confusion matrix of the multiclass classification models for both tests in two different wheelsets, in which the wrong classified observations are stacked one on top of the other. Just one existing fault at a time was considered. If multiple faults occur simultaneously, masking between important frequencies could happen (Lacey, 2008). The classification using microphones located near a leading and trailing wheelset (Figure 12a), in which just PSD energy bands were used in the initial feature pool, shows lower performance. SVD50 is just misclassified with PVD00 in both wheelsets, while PVD00 is misclassified also with normal condition.

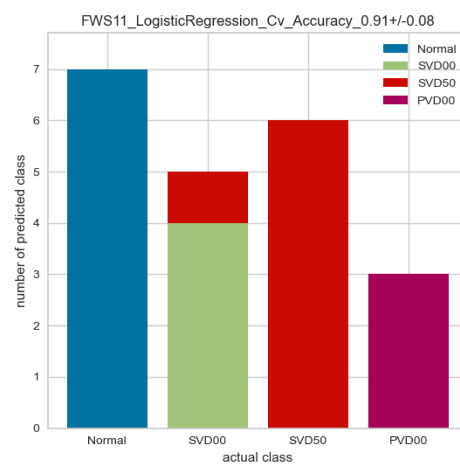
The amount of data available from signals sensor on an exposed bogie position in straight track Str1 (Figure 12b and Figure 12c) is considerably reduced. Depending on the wheelset, misclassification can vary considerably.



a)



b)



c)

Figure 12. Multiple fault classification models confusion bars for first tests in all section (a) and second tests in straight track Str1 (b and c).

Normal condition and PVD00 are mostly well classified, while SVD00 and SVD50 show higher misclassification.

Multiple fault classification models showed the capacity to be used as fault model selector for specific operating conditions.

5 CONCLUSIONS

The approach presented in this paper for suspension damper fault detection and classification with data-driven models using on-board acoustic sensors shows promising results.

Normalized Gini coefficient NGini (where Gini coefficient represents twice the area under the receiver operating characteristic curve minus one, normalized with the area of a perfect classifier, ground truth) has been used as classifier scorer. Data-driven models are considered as good with NGini values over 0.7.

With microphones located sensor near the leading or trailing wheelset, removed dampers in vertical suspension were classified considering all section types and speed ranges with a NGini of 0.75±0.1 (PVD00) and 0.77±0.04 (SVD00), while classification of secondary vertical 50% damper efficiency reduction fault achieved a NGini of 0.71±0.06 (SVD50). The cross-validated balanced accuracies of the classifications for the same operating condition are 92±2% (PVD00), 95±3% (SVD00) and 85±7% (SVD50).

The speed range from 30 to 36 m/s seems to be the most appropriate for classification of fault SVD00 in straight section, while in curves the range 36 to 39 m/s shows higher performance. The speed ranges from 23 to 28 m/s and 36 to 38 m/s show higher classification performance for fault PVD00.

Using microphones exposed to aerodynamic noise, all faults were classified successfully with just 2 features in the speed range of 30 to 40 m/s. Removed dampers in vertical suspension were classified in straight track Str1 with a NGini of 0.99±0.03 (PVD00) and 0.99±0.01 (SVD00), while classification of 50% damper efficiency reduction fault achieved a NGini of 0.84±0.09 (PVD50) and 0.77±0.11 (SVD50). The respective cross-validated balanced accuracies of the classifications for the same operating condition are 98±3% (PVD00), 92±6% (PVD50), 98±3% (SVD00) and 90±8% (SVD50).

Multiclass classification models achieved acceptable classification performance when faults were grouped under similar required excitation, they can be used for model selection of specific fault detection depending on the train operating conditions.

ACKNOWLEDGEMENTS

The publication was written at VIRTUAL VEHICLE Research GmbH in Graz and partially funded by the COMET K2 – Competence Centers for Excellent Technologies Programme of the Federal Ministry for Transport, Innovation and Technology (bmvit), the Federal Ministry for Digital, Business and Enterprise (bmdw), the Austrian Research Promotion Agency (FFG), the Province of Styria and the Styrian Business Promotion Agency (SFG).

We would furthermore like to express our gratitude to our supporting industrial partner Siemens Mobility Austria GmbH.

REFERENCES

Ali, S. M., Hui, K. H., Hee, L. M., & Leong, M. S. (2018). Automated valve fault detection based on acoustic emission parameters and support vector machine. *Alexandria Engineering Journal*, *57*(1), 491–498. https://doi.org/10.1016/j.aej.2016.12.010

Amini, A. (2016). *Online condition monitoring of railway wheelsets*. University of Birmingham. Retrieved from http://etheses.bham.ac.uk/6957/4/Amini16PhD.pdf

Amini, A., Entezami, M., & Papaelias, M. (2016). Onboard detection of railway axle bearing defects using envelope analysis of high frequency acoustic emission signals. *Case Studies in Nondestructive Testing and Evaluation*, *6*, 8–16. https://doi.org/10.1016/j.csndt.2016.06.002

Anastasopoulos, A., Bollas, K., Papasalouros, D., & Kourousis, D. (2010). Acoustic emission on-line inspection of rail wheels. In *EWGAE*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.462.610&rep=rep1&type=pdf

Atamuradov, V., Medjaher, K., Dersin, P., Lamoureux, B., & Zerhouni, N. (2017). Prognostics and Health Management for Maintenance Practitioners -Review, Implementation and Tools Evaluation. *International Journal of Prognostics and Health Management*, 2153–2648. Retrieved from https://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2016/ijphm_17_060.pdf

Bernal, E., Spiryagin, M., & Cole, C. (2019). Onboard Condition Monitoring Sensors, Systems and Techniques for Freight Railway Vehicles: A Review. *IEEE Sensors Journal*, *19*(1), 4–24. https://doi.org/10.1109/JSEN.2018.2875160

Bozzone, M., Pennestrì, E., & Salvini, P. (2011). Dynamic analysis of a bogie for hunting detection through a simplified wheel-rail contact model. *Multibody System Dynamics*, *25*(4), 429–460. https://doi.org/10.1007/s11044-010-9233-8

Chong, S. Y., Lee, J. R., & Shin, H. J. (2010). A review of health and operation monitoring technologies for trains. *Smart Structures and Systems*, *6*(9), 1079–1105. https://doi.org/10.12989/sss.2010.6.9.1079

Contributors, W. (2019). Gini coefficient. Retrieved from https://en.wikipedia.org/w/index.php?title=Gini_coefficient&oldid=928797425

Elamin, F. (2013). *Fault Detection and Diagnosis in Heavy Duty Diesel Engines Using Acoustic Emission*.

Gasparetto, L., Alfi, S., & Bruni, S. (2013). Data-driven condition-based monitoring of high-speed railway bogies. *International Journal of Rail Transportation*, *1*(1–2), 42–56. https://doi.org/10.1080/23248378.2013.790137

Gong, C. S. A., Lee, H. C., Chuang, Y. C., Li, T. H., Su, C. H. S., Huang, L. H., … Chang, C. H. (2018). Design and implementation of acoustic sensing system for online early fault detection in industrial fans. *Journal of Sensors*, *2018*. https://doi.org/10.1155/2018/4105208

Höjer, M., Bergseth, E., Olofsson, U., Nilsson, R., & Lyu, Y. (2016). A Noise Related Track Maintenance Tool for Severe Wear Detection of Wheel-Rail Contact. *Proceedings of the Third International Conference on Railway Technology: Research, Development and Maintenance*, (April), 1–17.

Jesussek, M., & Ellermann, K. (2015). Fault detection and isolation for a railway vehicle by evaluating estimation residuals. *Procedia IUTAM*, *13*, 14–23. https://doi.org/10.1016/j.piutam.2015.01.004

Lacey, S. J. (2008). An Overview of Bearing Vibration Analysis. *Maintenance and Asset Management*, *23*(6), 32–42.

Lanza di Scalea, F., Zhu, X., Capriotti, M., Liang, A. Y., Mariani, S., & Sternini, S. (2017). Passive Extraction of Dynamic Transfer Function From Arbitrary Ambient Excitations: Application to High-Speed Rail Inspection From Wheel-Generated Waves. *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems*, *1*(1), 011005. https://doi.org/10.1115/1.4037517

Li, C., Luo, S., Cole, C., Spiryagin, M., & Sun, Y. (2017). A signal-based fault detection and classification method for heavy haul wagons. *Vehicle System Dynamics*, *55*(12), 1807–1822. https://doi.org/10.1080/00423114.2017.1334929

Li, P., & Goodall, R. (2004). Model-Based Condition Monitoring for Railway Vehicle Systems. In *Control*. Bath.

Li, P., Goodall, R., Weston, P., Seng Ling, C., Goodman, C., & Roberts, C. (2007). Estimation of railway vehicle suspension parameters for condition monitoring. *Control Engineering Practice*, *15*(1), 43–55. https://doi.org/10.1016/j.conengprac.2006.02.021

Liu, X. Y., Alfi, S., & Bruni, S. (2016). An efficient recursive least square-based condition monitoring approach for a rail vehicle suspension system. *Vehicle System Dynamics*, *54*(6), 814–830. https://doi.org/10.1080/00423114.2016.1164869

Luber, B., Müller, G., Sorribes-Palmer, F., Pietsch, L., & Six, K. (2018). Influence of wheel profile condition on wheel-rail contact loading and vehicle dynamics – a stochastic approach. In *4th International Conference on Railway Technology: Research, Development and Maintenance*. Barcelona.

Mei, T. X., & Ding, X. J. (2009). Condition monitoring of rail vehicle suspensions based on changes in system dynamic interactions. *Vehicle System Dynamics*, *47*(9), 1167–1181. https://doi.org/10.1080/00423110802553087

Ngigi, R. W., Pislaru, C., Ball, A., & Gu, F. (2012). Modern techniques for condition monitoring of railway vehicle dynamics. In *Journal of Physics: Conference Series* (Vol. 364, p. 12016). https://doi.org/10.1088/1742-6596/364/1/012016

Nielsen, J. C. O., & Johansson, A. (2000). Out-of-round railway wheels-a literature survey. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, *214*(2), 79–91. https://doi.org/10.1243/0954409001531351

Pollock, A. A. (2018, August 1). Acoustic Emission Inspection. (A. Ahmad & L. J. Bond, Eds.), *Nondestructive Evaluation of Materials*. ASM International. https://doi.org/10.31399/asm.hb.v17.a0006454

Shahidi, P., Maraini, D., & Hopkins, B. (2016). Railcar Diagnostics Using Minimal-Redundancy Maximum-Relevance Feature Selection and Support Vector Machine Classification. *International Journal of Prognostics and Health Management*, *7*, 2153–2648.

Shahidi, P., Maraini, D., Hopkins, B., & Seidel, A. (2014). Estimation of Bogie Performance Criteria Through On-Board Condition Monitoring. In *Annual Conference Of The Prognostics And Health Management Society*. Retrieved from https://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2014/phmc_14_066.pdf

Shahidi, P., Maraini, D., Hopkins, B., & Seidel, A. (2015). Railcar Bogie Performance Monitoring using Mutual Information and Support Vector Machines. *Annual Conference of the Prognostics and Health Management Society*, 1–10.

Tidriri, K., Chatti, N., Verron, S., & Tiplica, T. (2016). Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges. *Annual Reviews in Control*, *42*. https://doi.org/10.1016/j.arcontrol.2016.09.008

Tsunashima, H., Mori, H., Tsunashima, H., & Mori, H. (2010). Condition monitoring of railway vehicle suspension using adaptive multiple model approach. *Control Automation and Systems (ICCAS), 2010 International Conference On*, *3*(1), 584–589. https://doi.org/10.1299/jmtl.3.243

Ward, C. P., Goodall, R. M., Dixon, R., & Charles, G. (2013). Condition Monitoring of Rail Vehicle Bogies. *UKACC International Conference on Control*, 1178–1183. https://doi.org/10.1049/ic.2010.0447

Ward, C. P., Weston, P. F., Stewart, E. J. C., Li, H., Goodall, R. M., Roberts, C., … Dixon, R. (2010). Condition monitoring opportunities using vehicle-based sensors. In *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* (Vol. 225, pp. 202–218).

https://doi.org/10.1243/09544097JRRT406

Zhang, X., Thompson, D., Ryue, J., Jeong, H., Squicciarini, G., & Stangl, M. (2018). The sound radiation of a railway rail fitted with acoustic shielding. In *25th International Congress on Sound and Vibration 2018, ICSV 2018 : Hiroshima Calling* (Vol. 7). International Institute of Acoustics and Vibration (IIAV). Retrieved from https://eprints.soton.ac.uk/430205/

Zhao, C., Wang, P., & Xing, M. (2017). Research on the Matching of Fastener Stiffness Based on Wheel-Rail Contact Mechanism for Prevention of Rail Corrugation. *Mathematical Problems in Engineering*, *2017*. https://doi.org/https://doi.org/10.1155/2017/6748160

Zoljic-Beglerovic, S., Golkani, M. A., Steinberger, M., & Horn, M. (2018). Robust Parameter Identification for Railway Suspension Systems. In *15th International Workshop on Variable Structure Systems and Sliding Mode Control*.

Zoljic-Beglerovic, S., Stettinger, G., Luber, B., & Horn, M. (2018). Railway Suspension System Fault Diagnosis using Cubature Kalman Filter Techniques. In *10th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes*.