Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 120 (2023) 702–707

www.elsevier.com/locate/procedia

56th CIRP Conference on Manufacturing Systems, CIRP CMS '23, South Africa

Leveraging digitilisation and machine learning for improved railway operations and maintenance

M. Bezuidenhout[a], J. L. Jooste[a]*, D. Lucke[b], C. J. Fourie[a]

[a]Department of Industrial Engineering, Stellenbosch University, 145 Banghoek Rd, Stellenbosch 7613, South Africa
[b]ESB Business School, Reutlingen University, Alteburgstraße 150, Reutlingen 72762, Germany

* Corresponding author. Tel.: +27218084234. E-mail address: wyhan@sun.ac.za

Abstract

The efficient and safe movement of goods and people require reliable railway systems. Quality assurance of manufactured and assembled systems and correct maintenance of such systems are required to keep rolling stock in good operational condition. Quality assurance and maintenance in the railway industry can be costly and time-consuming, but the expansive growth of data due to smart sensors and monitoring technologies makes it possible to leverage the potential of machine learning to reduce cost and labour. Improved reliability and safety, and reduced costs are benefits that the use of "Big Data" and machine learning techniques can realise. However, despite these potential benefits for manufacturers, rail operators, and passengers, the rail industry is still labelled for its lack of innovation, while in most other industries, data is regarded as a strategic asset for competitive advantage.
This paper demonstrates how machine learning and data analysis can be used to benefit railway industry manufacturers and operators when applied to rolling stock data. It also illustrates the lost opportunity in the rail industry for not applying data-driven solutions to their full potential. The paper also discusses the current applications of machine learning in the railway industry and provides the requirements for the implementation of machine learning techniques. Machine learning is applied to pantograph data of a South African railway operator's rolling stock. Classification – a machine learning technique – is used to identify and categorise events within the dataset to discover whether pantograph bounce occurs due to faulty sensors, faulty pantographs, or defective infrastructure. In this paper it is demonstrated how machine learning can benefit rail manufacturers and operators to improve manufacturing and assembly processes, as well as maintenance practices. It is concluded that railways should treat data similarly to other railway assets, with suitable management and governance practices.

Keywords: Machine Learning; Predictive Maintenance, Railway; Smart Sensors

1. Introduction

The rise in digitalisation across various industry sectors and public entities provides the opportunity to create new value and increase competitiveness. In the railway sector this is especially true, with benefits including improved efficiency, lower operating costs and providing better service to customers [1]. Furthermore, rail companies and infrastructure managers can leverage digitalisation to enhance many segments of their business, such as manufacturing, operations and maintenance practices [1].

Advancements in predictive technology, data mining and machine learning have enhanced preventive and condition-based maintenance practices to either supplement or replace previous generation maintenance systems, which often come with difficulties and inefficiencies [2]. The purpose of a maintenance programme is to ensure that physical assets are available and continue to accomplish their intended purpose with the least number of resources spent [3]. To perform such

10.1016/j.procir.2023.09.062

maintenance practices, however, it is first needed to collect the appropriate data.

Condition monitoring technology has made it possible to track various functions and components on rolling stock through onboard sensors. The intended purpose of this data is then to predict defects and breakdowns or to retroactively analyse the course of events that lead to a failure and find the root cause [1]. This means that rolling stock conditions can be predicted based on past and current system conditions.

Big Data can be described by its five V's. These are volume, variety, velocity, veracity and value [4]. The scale and complexity of these datasets make it difficult to analyse with traditional data-processing techniques, but machine learning offers a solution.

Machine learning applications are already applied in similar industries and expensive physical assets, such as aircraft, heavy earth-moving equipment, power plants and chemical plants [5]. Machine learning makes it possible to uncover accurate degradation patterns, develop prediction models, aid in decision making and generate better maintenance plans [6, 7].

In the field of railways, the problem is that a large amount of data is gathered, but the data is not always transformed into valuable information. In other words, although the amount of data increases exponentially, the rail industry still lacks automated solutions and barely uses machine learning to solve problems [6, 4].

This paper aims to address how Big Data and machine learning can improve the rail industry and why emerging machine learning applications are not being used more widely currently. The advantages and current applications of machine learning in railway are discussed, along with the challenges in its application. Finally, the paper will also reduce the gap between machine learning's potential and its applications by describing a case study on predictive maintenance with pantograph data of a South African rail operator.

2. Machine Learning

Machine learning is the scientific study of algorithms, statistical models and computational models to make predictions on new observations [8, 9]. Learning algorithms aim to perform a specific task by learning from underlying patterns in data without needing explicit instructions from human professionals [9, 10]. Machine learning can therefore be considered a subset of artificial intelligence.

There are two main subsets of machine learning, namely supervised and unsupervised learning algorithms. Supervised learning is used to predict future variables based on past data. An example of this type of machine learning is predicting the future state of assets. On the other hand, unsupervised learning is used to discover the inherent structure present in unlabeled data. Identifying outliers or critical incidents is an example of unsupervised learning [6, 10].

2.1 Maintenance in railway

Machine learning not only satisfies the need to process large datasets, but it can also support predictive maintenance in the railway industry. Predictive maintenance is a type of preventive maintenance, as its main goal is to prevent failures [2]. Unplanned failures are reduced by using indicators to predict the time of failure and proactively replace or repair components so that it does not deteriorate below a desired level [2]. Other forms of preventive maintenance are planned according to either a schedule or the operation time of assets. The manufacturer's recommendations are the primary source used to establish the maintenance frequency in these cases [11].

Traditionally, corrective maintenance, which is a more reactive approach, was performed when a failure already occurred. It includes repairs and replacements of components in order to restore the system [2]. This type of maintenance has a high downtime as it is often unplanned and can be costly [11].

Predictive maintenance can therefore reduce railway maintenance costs, improve reliability and increase asset availability [12]. The failure data gathered for predictive maintenance also extends beyond just maintenance, as it can be used for root-cause analyses and the redesign of components [12].

In order to implement predictive maintenance, three types of decisions are required [3]. These are which parameters to monitor; what is the inspection frequency; and what is the warning limit (the trigger). Once these questions are answered, the relevant data can be gathered.

2.2 Data collection

Rapid advancements in digitalisation and its accompanying technologies have enabled the exponential increase in data acquisition for railway assets. The affordability of such technology has also drastically improved, as expensive physical network infrastructure is not required.

Condition monitoring is the method or process to acquire or record certain parameters in order to support maintenance or operating decisions relating to machines, equipment, or in this case, rolling stock [13]. Condition-monitoring technologies therefore sense, measure and record data from physical assets either periodically or continuously. This data can then be analysed to predict future trends. Other benefits of condition-monitoring technology include limiting the probability of failures, identifying the root causes of failures, reduced maintenance costs and monitoring assets in real time [13].

Enormous amounts of data can be collected in ways that was previously not possible. This includes collecting data through GPS, on-board sensors, video cameras, fiber-optic cables and hand-held devices [4, 14]. These low-cost electronic devices have the added benefit of monitoring assets when rail operators experience cost and personnel constraints. Accurate and real-time physical measurements can also be taken on heat, sound, speed, voltage, stress and vibrations [12].

Data acquisition from condition-monitoring technology involves placing sensors and devices on the applicable components and infrastructure. These sensors and devices transmit data to data servers, where the data is then stored. Data can then be accessed by data science and rolling stock experts or by the train operator.

Condition monitoring is already prevalent in other industries with expensive physical assets. For example, Bombardier is a Canadian manufacturer of business jets and public transport vehicles and uses a maintenance technology called Orbita. Orbita anticipates faults by taking into account the environment and conditions of the assets in use [15]. French National Railways (SNCF) also uses remote diagnosis to optimise maintenance and prevent failures, which have resulted in the number of breakdowns being cut in half on those electric multiple units (EMUs) [16, 17]. Another company called Embraer, also developed a cloud-based technology called IKON, which automatically collects data on assets in the aeronautical industry [18]. Although new data is collected which was previously either too expensive, impractical or not possible, the value only lies in processing and analysing it.

2.3 Applications and benefits of machine learning

Depending on the type of problem and the available data, various specialised algorithms can be used. Generally, classification is used to determine to which predetermined categories an event belongs and is, therefore, a subset of supervised learning. Regression is another subset of supervised learning and is used to predict future outcomes. Clustering falls under unsupervised learning and is used to group events intuitively when data is not labelled. Anomaly detection is also an unsupervised learning method and aims to find outliers or abnormal behaviour in the data [19].

Machine learning applications are relevant to the manufacturing, operations, and the maintenance of assets. The use of machine learning in manufacturing can optimise the production of assets, whereas its use in maintenance can assist in assets continuing to function as intended over time.

Machine learning has been applied extensively in the manufacturing industry both within and outside of the railway sector. Numerous manufacturing parameters can be continuously monitored at various stages of the production process. This can include characteristics relating to equipment, production lines, human resources operating the lines, materials, and the environment, such as temperatures [20]. One such example is the use of support vector machines, which forms part of supervised machine learning, for the monitoring of the health of tools and machines [21]. As the technology becomes cheaper and more accessible, digitalisation and asset monitoring in rail improves. Several machine-learning methods have already been studied and implemented in select railways.

Chalouhi et al. [22] developed a machine-learning model to detect anomalies at each train passage of a railway bridge in Northern Italy to detect damage in railway bridges. In Canada, a classification-based method is used to detect rail defects by using acceleration data [23]. In Greece, classification trees and regression trees are used to predict the diagnosis of a train fleet and create strategic decision support to help with maintenance planning [7]. A new strategy has also been introduced by Asad et al. [24] to track travelers and profile age groups in order to address disease outbreaks. It is therefore possible to recommend designated safe track routes, certain train carriages, stations, and platforms for staff and the public.

One of the most important elements required for the continuous and uninterrupted operation of railway systems is rail tracks. Due to the mechanical forces of trains and environmental conditions, rail tracks degrade and can incure great costs for rail operators or influence the safety and riding quality for passengers [10]. Artificial Neural Networks (ANN), Grey Models (GM) and Support Vector Machines (SVM) are all machine-learning methods used to predict the degredation of railway tracks [25]. In Japan specifically, a track condition monitoring system has been studied successfully to identify faults in tracks automatically based on car-body vibrations [14].

Research has also been done in the area of predictive maintenance specifically. Hongfei et al. [5] built failure prediction models by using several approaches, such as machine learning, correlation analysis and time series analysis. Bukhsh et al. [6] also trained two machine-learning classifiers to predict maintenance needs and treatments. These classifiers can be used as an add-in tool for decision support and maintenance planning.

Although the mentioned applications of machine learning proved to be beneficial for railway operators, few machine learning solutions in this sector have been reported in the literature when compared to other industries [6].

From these examples it is clear that machine learning can optimise the production process, improve performance and maintain machine health. Machine-learning applications have also proven to be advantageous on the operations side, such as reducing service interruptions, increasing network velocity, and saving enormous costs for rail operators. There will always be a trade-off between the maintenance schedule, cost, and an operator’s capacity to perform maintenance, but machine learning can optimise this trade-off.

2.4 Challenges in applying machine learning in railway

The collection of immense amounts of data in the railway sector through condition monitoring technology is unproductive if it will not be used to gain valuable information.

It is important to note that railway industries can differ significantly in terms of technology and infrastructure from country to country, or occasionally even within a single country. This also means that their adoption of machine learning into manufacturing, operations or maintenance decisions may vary significantly. However, there are also common machine learning challenges, irrespective of the diversity of technology and infrastructure.

As discussed in the previous section, many applications of machine learning and the use of data for predictive maintenance is already applied in railway, however, the implementation of these types of innovative solutions are scarce. Although there is a need for innovation, the failure to take advantage of data can be attributed to several challenges.

Barriers to innovation include a resistant culture in railway, outdated procurement processes, and fragmented data [26]. One example of such a challenge is that the same procurement process is used for both low-risk technology and safety-critical technology, which makes small changes more difficult than it needs to be.

There are also challenges that relate to machine-learning techniques, especially when it needs to be used for predictive maintenance. One such challenge is that data is collected from multiple sources. This makes combining information difficult, as different information from various sensors do not have the exact same measurements for similar events [5]. Najafabadi et al. [27] also point out several obstacles, such as unstructured data formats, lack of data quality, unlabeled data and multi-source data. Data ownership and communications have also fallen short in the railway industry [12]. There is therefore no single accountable entity to oversee data management.

The lack of public benchmark datasets also causes shortcomings in the machine learning research in railway, as models cannot be compared or improved upon [10]. Another challenge is that machine learning solutions are not easily interpretable by operators. The solutions from these models are also not easily explained, making it untrustworthy to the end user [5, 10]. Although digitalisation and machine learning has made it possible to innovate and improve railway operations, these barriers have made train service performance fall short of expectations.

3. Case Study

A South-African railway operator has recently put new state-of-the-art commuter trains into operation. These trains are fitted with advanced condition-monitoring technology. Although an immense amount of valuable data is collected on various components and subsystems on the trains, the use of this data is limited. Many of the same challenges mentioned in Section 2.4 were also encountered during the course of the study. A vast amount of events is collected from a large amount of sensors and stored in the same database. Even when a user is knowledgeable about components that are monitored, the data is complex and easily misinterpreted. This then leads to either incorrect maintenance activities to be suggested or interpretations to be mistrusted. This results in missed opportunities for reducing costs and decreasing asset downtime. This provides reason to research how this data can be used to gain insights that are not only correctly interpreted but can also be easily understood and adopted.

3.1 Problem definition

A case study was carried out on the pantograph data of the new fleet to demonstrate how data analysis and machine learning can transform this data into valuable information that can be used to make better maintenance decisions.

The pantograph is a linkage assembly that is placed on the roof of a train and connects to the overhead current-supply system. Uninterrupted current supply is essential to the train's operation, but it can happen that the contact between the pantograph and overhead wire is momentarily broken. This phenomenon is known as pantograph bounce [28].

Pantograph bounce can be caused by many factors, but the broad categories these fall into is that there is either a faulty pantograph or a problem with the railway infrastructure and overhead traction equipment.

3.2 Challenges in current operations

The current solution to pantograph bounce is to only assess GPS data of the whole fleet. This lead maintenance personnel to come to the conclusion that infrastructure along the whole route was faulty and in need of maintenance. This conclusion was incorrect, because if there is a faulty pantograph on a train, there would be pantograph bounce recorded along the whole route. If it was indeed the infrastructure that needed maintenance, pantograph bounce would have only occurred at certain coordinates. In essence, maintenance personnel were not able to distinguish between what caused pantograph bounce and where maintenance was needed.

Machine learning applications are also not applied due to a resistance to change within current operations and a lack of data ownership. Silos of data is also being collected at a scale that has become unmanageable in some cases, especially since data processing is needed. Data also needs to be analysed in depth before valuable conclusions can be drawn, meaning that the automation of data has not been applied yet to help with decision making.

3.3 Data collection

Sensors are placed on each of the train cars, and as a result, it is possible to identify exactly where on the train the current supply was interrupted. The placement of these sensors are shown in Figure 1, where the four motorised coaches are denoted as M1, M2, M3 and M4 and the two trailer cars are denoted as TC1 and TC2. Two pantographs are placed on the roof of the trainset. The first pantograph feeds current to the M1 and M4 coaches and the second pantograph feeds current to the M2 and M3 coaches. The three causes of pantograph bounce can be classified according to the following scenarios:

- Only one of either the M1, M2, M3 or M4 sensors records pantograph bounce. This means that the sensor that recorded the event is most likely faulty.
- Either the M1 and M4 or M2 and M3 sensors record pantograph bounce. This means that a specific pantograph is faulty.
- Any three out of four sensors or all four sensors record pantograph bounce. This means that the problem lies with the railway infrastructure and overhead traction equipment.

It is important to note that these scenarios can only be classified for events that occur on the same train, at the same time and at the same location.

3.4 Machine learning application

Classification is an appropriate machine-learning solution to predicting what caused pantograph bounce. Classification is a form of supervised machine learning that analyses the relationships between features from past data and uses it to predict the labels of new data. The causes of pantograph bounce can be defined, meaning that the exact classes the model needs to categorise data into are already known. Classification can therefore be used to predict the labels or classes for pantograph bounce based on past data.

Several features were also analysed individually, such as external temperature, speed, distance travelled, longitudinal and latitudinal coordinates, and the date of pantograph-bounce occurrence. Figure 2 shows the logic used to classify the cause

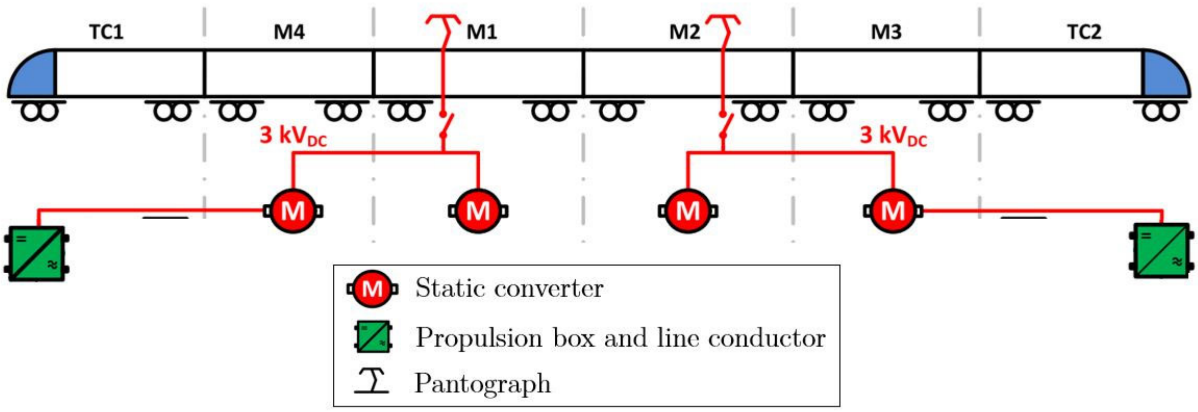

Figure 1: The placement of pantographs on a commuter train [29].

of pantograph bounce. Training data was then used to train a k-Nearest Neighbors (KNN) model, as well as a classification tree.

3.5 Results and impact

In this case study, the KNN model outperformed the classification tree. The KNN model was able to predict the cause of pantograph bounce accurately 85.37% of the time.

The case study poses multiple benefits to the rail operator. Firstly, it demonstrates that the data gathered can indeed be used to gain valuable insights to make better maintenance decisions. Furthermore, specific problems can be identified that was not possible before.

If pantograph bounce occurred due to a faulty sensor, that sensor and train can be identified and inspected if the problem persists. The same is true for a faulty pantograph that caused pantograph bounce, as the specific pantograph on a specific train can be identified with ease.

Exact coordinates of infrastructure faults can also be identified if pantograph bounce was classified as being due to faulty infrastructure. If there was a problem with the railway infrastructure, several trains in the fleets that travelled through those locations would have experienced pantograph bounce. This makes classifying the cause of pantograph bounce so important, as it would be clear to the rail operator that there is not a problem with any of the trains within the fleet, but that the problem can be fixed at the identified coordinates.

It is therefore possible to pinpoint exactly what needs maintenance and therefore save time, money and effort.

4. Recommendations

There is a strong need for innovation in the rail industry, especially in South Africa. Even though some rail operators have tried to implement modern solutions, the challenges mentioned in Section 2.3 still create a barrier. To overcome these challenges, the following is recommended:

- Funding intended to stimulate innovation in the rail industry should be advertised better and administered more effectively [26].
- Datasets should be made publicly available to encourage and increase research. Machine learning research in the rail industry is somewhat limited so granting access to academia could improve existing methods.
- Proper data governance is needed to ensure data availability, usability, and consistency. Data governance also includes the rules and policies that ensures quality data and that it is used for critical decision making.
- The proper data architecture is needed in order to sort, arrange and access data efficiently and effectively. Essential data should be accessible by the appropriate departments. Data should also be stored in a standardised format.
- Data ownership is lacking. Having a department or entities accountable for the efficient and accurate storage and processing of data, especially as the amount of data being generated increases would be beneficial.

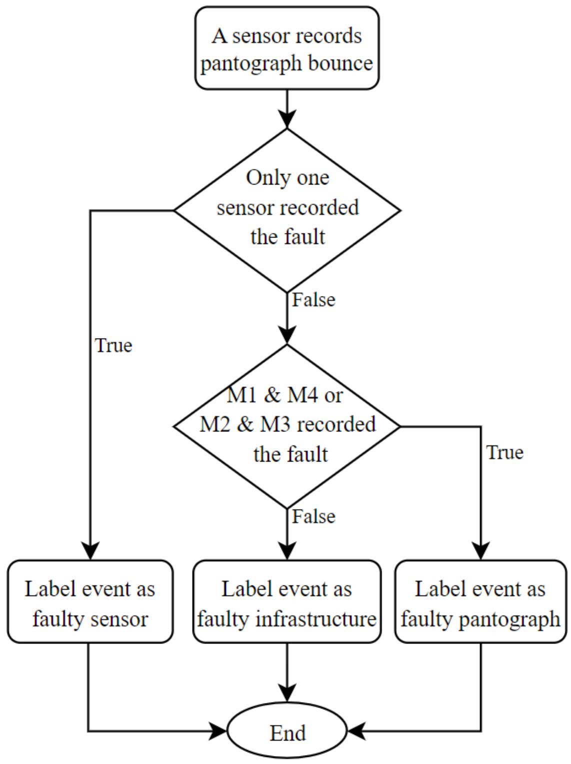

Figure 2: Classification of pantograph bounce events.

- Focus on solutions to be interpretable. There is a trade-off between smart predictive maintenance solutions and human-interpretable results [5].

5. Conclusion

The rail industry's greatest asset is no longer rolling stock, but rather data. The ability to capture, manage, and store vast amounts of data have created new opportunities for the railway industry. Data can now be used for the management, operation, and maintenance of railway assets.

This paper investigated the current applications of machine learning in railways and why it is not being used more widely. A case study was presented on the classification of pantograph bounce in order to make better maintenance decisions. The case study demonstrated that the value does not only lie in collecting data through emerging technologies but processing it and applying readily available machine-learning techniques to it.

Machine learning can improve many aspects of the rail industry when used correctly and combined with proper data management. This will in turn result in more reliable railway service, that operates more safely and is more cost-effective for rail operators.

Acknowledgements

This paper was written as a collaboration between the Industrial Engineering Department of Stellenbosch University, South Africa, and the ESB Business School of Hochschule Reutlingen, Germany. The authors would like to thank the Passenger Rail Agency of South Africa (PRASA) for providing data and funding for this research.

References

[1] D. Scordamaglia, "Digitalisation in railway transport: A lever to improve rail competitiveness," European Parliamentary Research Service (EPRS), 2019.

[2] A. Prajapati, J. Bechtel and G. Subramaniam, "Condition based maintenance: a survey," *Journal of Quality in Maintenance Engineering,* vol. 18, no. 4, pp. 384-400, 2012.

[3] A. H. Tsang, "Condition ‐ based maintenance: tools and decision making," *Journal of Quality in Maintenance Engineering,* vol. 1, no. 3, pp. 3-17, 1995.

[4] N. Attoh-Okine, "Big data challenges in railway engineering," in *2014 IEEE International Conference on Big Data*, Washington, 2014.

[5] L. Hongfei, D. Parikh, Q. He, B. Qian, Z. Li, D. Fang and A. Hampapur, "Improving rail network velocity: A machine learning approach to predictive maintenance," *Transportation Research Part C: Emerging Technologies,* vol. 45, pp. 17-26, 2014.

[6] Z. A. Bukhsh, A. Saeed and I. Stipanovic, "A machine learning approach for maintenance prediction of railway assets," in *Transport Research Arena TRA 2018*, Vienna, Austria, 2018.

[7] I. Kalathas and M. Papoutsidakis, "Predictive maintenance using machine learning and data mining: A pioneer method implemented to greek railways," *Designs,* vol. 5, no. 1, pp. 5-23, 2021.

[8] Z.-H. Zhou, Machine Learning, Singapore: Springer Nature, 2021, pp. 1-24.

[9] R. D. Sarode and C. A. Khanzode, "Advantages and disadvantages of artificial intelligence and machine learning: A literature review," *International Journal of Library & Information Science (IJLIS),* vol. 9, no. 1, pp. 30-36, 2020.

[10] M. C. Nakhaee, D. Hiemstra, M. Stoelinga and M. van Noort, "The recent applications of machine learning in rail track maintenance: A survey," in *International Conference on Reliability, Safety, and Security of Railway Systems*, Cham, 2019.

[11] S. Selcuk, "Predictive maintenance, its implementation and latest trends," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture,* vol. 231, no. 9, pp. 1670-1679, 2017.

[12] S. Davies, J. Daniels, T. Li, R. Grill and J. Baier, "Can the railway industry utilize data to its full potential?," WSP, London, 2019.

[13] B. Rao, Handbook of condition monitoring, Oxford: Elsevier, 1996.

[14] H. Tsunashima, "Condition monitoring of railway tracks from car-body vibration using a machine learning technique," *Applied Sciences,* vol. 9, no. 19, p. 2734, 2019.

[15] H. Sarih, P. Tchangani, K. Medjaher and E. Pere, "Data preparation and preprocessing for broadcast systems monitoring in PHM framework," in *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, 2019.

[16] D. Scordamaglia, "Digitalisation in railway transport: A lever to improve rail competitiveness," European Parliamentary Research Service, 2019.

[17] D. Burroughs, "Digital transformation improves SNCF's maintenance systems," International Railway Journal (IRJ), 25 June 2020. [Online]. Available: https://www.railjournal.com/in_depth/digital-transformation-improves-sncfs-maintenance-systems/. [Accessed 4 4 2022].

[18] Embraer, "Embraer Launches Predictive Maintenance System IKON with Big Data and Analytics Based on Amazon Web Services Cloud," Embraer, 19 6 2019. [Online]. Available: https://embraer.com/global/en/news?slug=1206610-embraer-launches-predictive-maintenance-system-ikon-with-big-data-and-analytics-based-on-amazon-web-services-cloud. [Accessed 4 4 2022].

[19] S. S. Skiena, The data science design manual, New York: Springer, 2017.

[20] D. Birant and A. Dogan, "Machine learning and data mining in manufacturing," *Expert Systems with Applications,* vol. 166, 2021.

[21] T. Wuest, K. D. Thoben, D. Weimer and C. Irgens, "Machine learning in manufacturing: advantages, challenges, and applications," *Production & Manufacturing Research,* vol. 4, no. 1, pp. 23-45, 2016.

[22] E. K. Chalouhi, I. Gonzalez, C. Gentile and R. Karoumi, "Damage detection in railway bridges using machine learning: application to a historic structure," *Procedia engineering,* vol. 199, pp. 1931-1936, 2017.

[23] C. Yang, Y. Sun, C. Ladubec and Y. Liu, "Developing machine learning-based models for railway inspection," *Applied Sciences,* vol. 11, no. 13, pp. 1-15, 2021.

[24] S. M. Asad, K. Dashtipour, S. Hussain, Q. H. Abbasi and M. A. Imran, "Travelers-tracing and mobility profiling using machine learning in railway systems," in *2020 International Conference on UK-China Emerging Technologies (UCET)*, Glasgow, 2020.

[25] Y. Liao, L. Han, H. Wang and H. Zhang, "Prediction models for railway track geometry degradation using machine learning methods: A review," *Sensors,* vol. 22, no. 19, pp. 1-26, 2022.

[26] K. Barrow, "Overcoming barriers to innovation: what can rail learn from developers?," International Railway Journal (IRJ), 18 8 2019. [Online]. Available: https://www.railjournal.com/in_depth/overcoming-barriers-to-innovation-what-can-rail-learn-from-developers/. [Accessed 15 10 2022].

[27] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *Journal of big data,* vol. 2, no. 1, pp. 1-21, 2015.

[28] T. Dahlberg, "Catenary, pantograph, and their interaction," *Vehicle System Dynamics International Journal of Vehicle Mechanics and Mobility,* vol. 44, no. 8, pp. 591-593, 2006.

[29] B. Thierry, "Operate the EMU trainset: Learner guide," Gibela, Johannesburg, 2015.