



CYIENT

BRINGING ANALYTICS- DRIVEN PROCESS EXCELLENCE IN RAIL

How big data analytics optimized condition monitoring of rolling stock

Contents

Abstract	01
Introduction	01
Condition monitoring of wheelsets	02
Data analytics framework for condition monitoring of rolling stock	04
Case discussion: Rail wheel wear computation	04
Conclusion	07
About the Authors	07
References	07

Abstract

Modern rail and metro networks, typically include a combination of physical assets and information systems that need constant monitoring and management. The information systems, supported by a mix of legacy and modern platforms, are increasingly moving towards a virtual cloud infrastructure. The digitalization of the industry has accelerated these changes and now offers the potential to improve rail and metro asset availability and operational performance.

As a result of these efforts, terabytes of raw data are being generated, which comprise of mission-critical information, maintenance reports, and passenger-focused analytics. Today's challenge is the ability to access this data, analyze and interpret it, and derive actionable insights based on the information.

This paper discusses a use-case, wherein, the Cyient Insights team developed a condition monitoring system for rolling stock by analyzing big data and applying analytics approaches. The project involved the development of a robust analytical framework for monitoring and forecasting deterioration of wheel sets, and optimization of maintenance processes. The proposed framework incorporates varied data sources including trackside monitoring stations, operating conditions, and maintenance records. Near real-time monitoring of assets is enabled by the solution infrastructure.

An ensemble model of machine learning (ML) and statistical approaches were developed to predict the deterioration in wheelset profile. Prioritized alert generation aimed to aid operations and maintenance managers in better planning for asset management and customer services.

Introduction

Deterioration of rolling stock in general, and wheelsets in particular, is a natural outcome of regular operations. It is influenced disproportionately by extreme operating conditions, foreign substances, and obstacles. Deterioration damages the rolling stock, leading to reduced asset availability and accidents like a derailment. This makes it critical to address the issue proactively. Condition-based monitoring and reliability research have always been an important part of operating and maintaining rail assets, especially rolling stock.

The results obtained from historical monitoring data research have typically been used for changes in design and development of such assets, and also in planning corrective maintenance actions and replacement schedules of components. Existing literature in reliability research has extensively focused on the Failure Mode Effects and Criticality Analysis (FMECA). The results of this analysis are applied to product design and maintenance decisions.

Technological advances, over the last decade, have revolutionized how data can be gathered, analyzed, and visualized. The five Vs i.e. volume, velocity, variety, value, and veracity, commonly used to describe the aspects of big data, are derived from the META group's original 3Vs [Laney, D. 2001]. Monitoring and predictive solutions with cognitive and informational elements have been developed using advanced analytical approaches, to bring better insights for business managers. Recent developments in data acquisition (IoT and sensors), data storage and handling systems like big data technologies have enabled advanced logging systems for rail assets.

Modern data pipelines allow rail operators to collect data regarding an equipment's operation, maintenance, and health in real-time. The IT infrastructure of a typical rail operator comprises of large quantities of data, which are generated by the varied systems.

Examples include onboard sensors, wayside monitoring, operations and maintenance (scheduled and unscheduled) activities, social media, and customer relations management. The data can be textual, numeric, and even audio-visual. Technological innovations have made these data acquisition techniques and storage cheaper and faster.

The improved economics of data acquisition and storage, and the development of advanced analytical tools and methodologies have enabled the transformation of large quantities of disparate data into useful information using statistical, mathematical, and machine learning models. This aids transportation professionals and decision-makers in making accurate and timely decisions leading to the phenomenon of data monetization.

Condition monitoring of wheelsets

Condition monitoring of rolling stock is an active research area in the rail industry and academia. A few relevant papers and their contributions are listed below to indicate the diverse approaches being adopted by the railway industry.

Industry experts, Olga Fink, Enrico Zio, and Ulrich Weidmann⁶, in their study of discrete-event diagnostic data, used a fuzzy classification approach with a combination of Echo-State Networks (ESNs) and a Restricted Boltzmann Machine (RBM), for predicting potential railway rolling stock system failures. The developed approach was then applied to railway door system with real data. They concluded that the combination of algorithms performed well in terms of prediction accuracy on the railway door system failure.

In another study on railway rolling stock failures³, the potential risks of unexpected failures occurring in rolling stock were identified, analyzed, and evaluated using failure mode, effects, and criticality. The most critical failure modes in the system on both reliability and economic criteria were reviewed while determining the levels of failure criticality and

possible methods for mitigation. The study concluded with the usability of the results for:

- a) the performance evaluation of existing maintenance practices and,
- b) planning a cost-effective, preventive maintenance program for different components of rolling stock.

The door sub-system is arguably the most actively analyzed among all the train sub-systems. Unipart Rail⁴ suggests that door fault accounts for over 30% of train failures and is the primary cause for delayed train schedules. The delay is due to safety reasons, whereby, the brake systems cannot be released on a train until the doors are completely shut and locked. The Unipart Rail case study describes how the Door Diagnostic Unit (DDU) developed by them, provides a cost-effective door condition monitoring system while enabling remote monitoring of train doors and processing the data in real-time. This is made possible with the use of complex, parallel processing such as digital signal processing and other methods.

Another research uses neural network models for detailed diagnostics once a fault has been detected in the train door operation¹. This method of detection and diagnosis was implemented in a distributed architecture, which would result in a low-cost industrial solution for monitoring the health of the assets in rapid transit systems. The research also concluded that it is feasible to integrate the results of the diagnosis process directly into the operator's legacy Maintenance Information System (MIS), thus producing a proactive maintenance regime.

Pereira P. and others², developed a proactive failure detection method, by using data from the train logging system to predict train doors breakdowns, before they happen. The researchers studied three methods for failure detection:

- 1) outlier detection,
- 2) novelty detection, and
- 3) supervised support vector machines (SVM).

The emphasis was given to reduce false alarms that could be provoked by the algorithms in case a passenger interrupts the movement of a door.

Studies indicate that a wide variety of analytical approaches such as time series modeling, anomaly detection, signal processing, and supervised and unsupervised ML techniques are used for condition monitoring, failure forecasting, alerts generation, and prioritized sequencing of operations and maintenance. Commercial solutions do exist to address some of the issues in the management of rolling stock assets—few of them are listed in the next section. However, the complexity and uniqueness of rail networks necessitate the customization and development of analytical workflows for successful implementations. The benefits include increased asset availability and reliability, efficient maintenance schedules, reduction in operational expenditure, inventory holdings, and higher scores on safety and service metrics.

Commercial solutions available in the field of rail asset monitoring and analysis

- MERMEC, through its products RAMSYS and TRACKWARE provides data planning, storage, processing, and custom reporting both in real time and offline.
- Müller-BBM Rail Technologies GmbH provides condition monitoring solutions such as Wheel Monitoring System (WMS), Wheel Roughness Measuring System, and Rail Roughness Measuring System to reduce the noise of vehicles and maintenance costs.
- Beena Vision has two main product lines for condition monitoring, train monitoring, and track monitoring systems. While the train monitoring product line has products such as WheelView, TreadView, BrakeView, and TruckView, the track monitoring product line has products like SurfView-Rail (rail and track imaging & inspection system) and TrackView-Profile (rail and track profile measurement system).
- KOLTECH specializes in the production of mechanical equipment and devices for rail vehicle maintenance and repair facilities worldwide. The company also offers integrated services for implementation of advanced product life cycle management (PLM) systems such as Wheel Lathes, Wheelset Measurement, and Monitoring Equipment.



Data analytics framework for condition monitoring of rolling stock

A typical infrastructure of a rail operator comprises of linear and non-linear assets. By defining these assets to taxonomies and ontologies, a big data storage system can be formed. This storage infrastructure is either setup at the client premises, or in a cloud-based third-party storage system such as Amazon Web Services (AWS) or Microsoft Azure. This enables gathering and storage of a wide variety of data originating from operations, maintenance, manufacturing, warehouses, geo-locating devices (GPS), people, IoT, sensors, social media, and other sources. This infrastructure is used as a foundation to create pipelines for data preprocessing and analysis.

Analytical models based on statistical, mathematical, and machine learning approaches are used to develop the advanced

analytics engines. Business Intelligence (BI) dashboards and simulation engines provide insights on asset conditions, suitable recommendation, and early warning alerts, while prioritized maintenance alerts provide a decision-support system for managers that enable informed and optimal decision-making.

Case discussion: Rail wheel wear computation

An analytical framework for condition monitoring of assets was developed for a rail operator handling both metro and rail transportation (passengers and goods) in a European market. The project uses the data obtained from a trackside wheel profile monitoring system designed to measure profiles and associated parameters on the wheelsets of a wagon as it passes over the system at normal track speed.

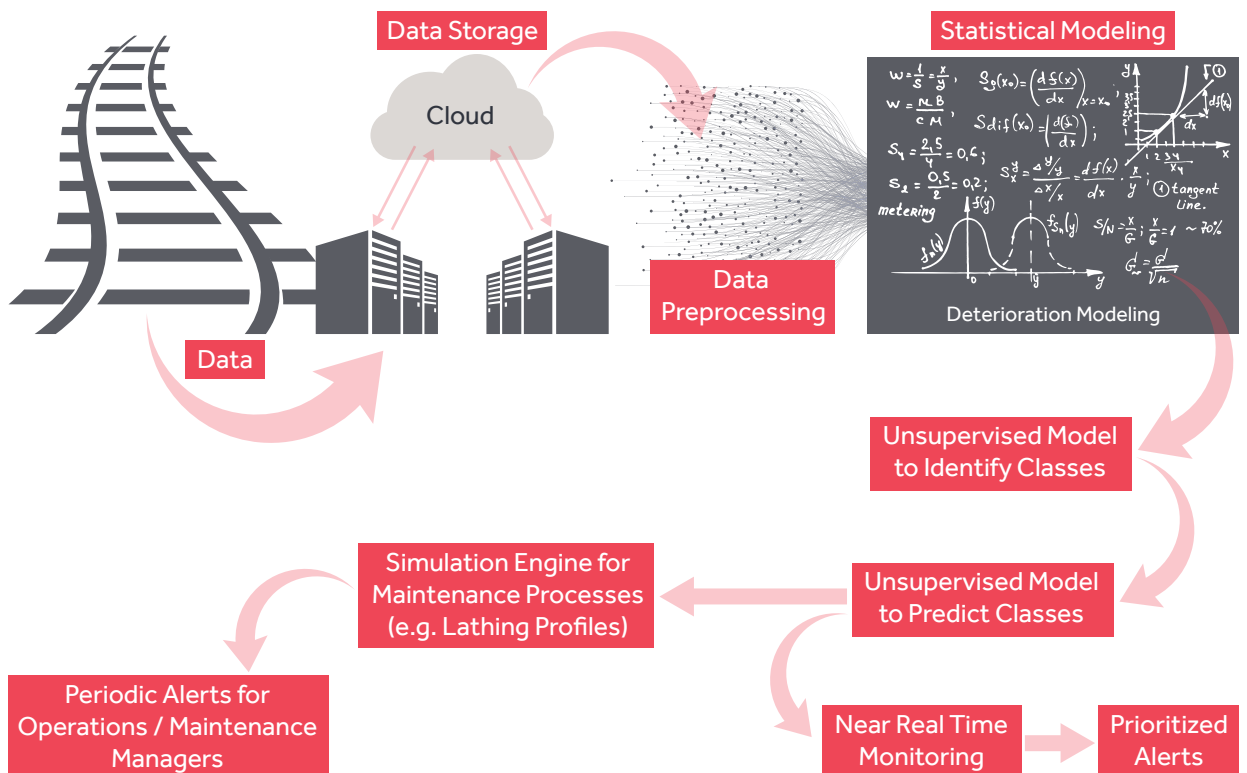


Fig. 1 | Data analytics enabled condition monitoring framework at Cyient

Data acquisition

A sophisticated high-speed, high-quality, laser-based digital imaging system kicks into operation when wheel sensors detect a train's arrival. RFID tags identify the train and the wheelset. The axle position of the wheelset is also noted. All information is tagged with accurate timestamps. Further, image processing algorithms are used to study the wheel profile data from the acquired images. The information that results from this processing includes: a) wheel profile, b) shear forces in rails, and c) load impact and notch size computations.

The processed data is stored in the cloud-based database in text format. The frequency of the data depends on the number of times a train passes through the measuring station, located along the key tracks across the rail network.

The figures below show samples of the system output:

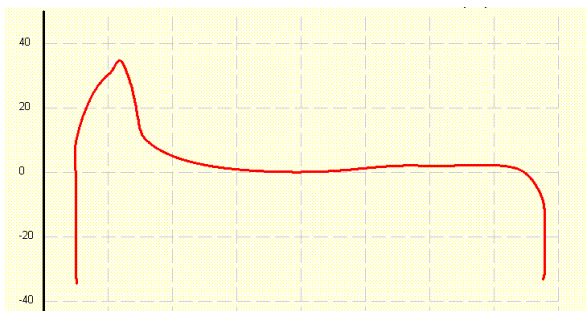


Fig. 2 | Sample wheel profile

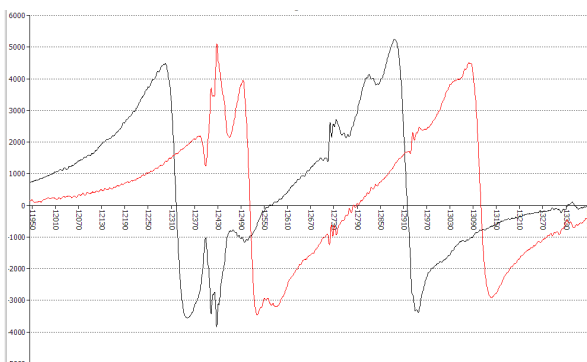


Fig. 3 | Shear forces inside rail track

Objectives of the project

Train wheels undergo wear (e.g. the loss of material from the surface, deformation of flange curve) due to regular operations. Wear (deterioration) refers to the changes in the shape of wheel and rail surfaces which further influences rolling contact fatigue. The main business objective of this project was to develop an analytics-enabled framework for monitoring the deterioration of the wheelsets. A secondary objective was to develop a model to predict the rate of deterioration. The final requirement was to come up with recommendations on the optimal scheduling of lathing operations on the wheel to maximize the wheel's expected lifetime:

Methodology

The analysis component of the framework included the following steps:

- Computation of wear loss of wheel profile in operation
- Development of solution for near real-time monitoring of wheel condition
- Development of a simulation engine to identify optimal lathing schedules, based on the historical values of computed wear loss

AUC (area under the curve) was used as a proxy for the wheel material lost in operations. If the curve function $F(X)$ is smooth (neither sharp corners nor sharp turning points) and continuous, a curve-fitting and a numerical integration method might need to be applied to compute the area under a curve. This means the definite integral of function $F(X)$ is limited to the area under a curve, which can be represented as:

$$AUC = \int_a^b F(X) dX$$

Where $[a, b]$ is the finite X domain.

If a curve cannot be described as a smooth and continuous function $F(X)$, as in our case, the "Trapezoidal Rule" will be used to compute the area of each small trapezoid. The AUC computing method becomes relatively simple when we use the trapezoidal rule:

- The total sum of those individual areas equals the total area under the curve (i.e., total AUC)
- Here X and Y are discrete, and the area is calculated on a baseline value (either $Y=0$ or $Y=$ some other predefined value).

Figure 4 explains the process of computation of wear loss:

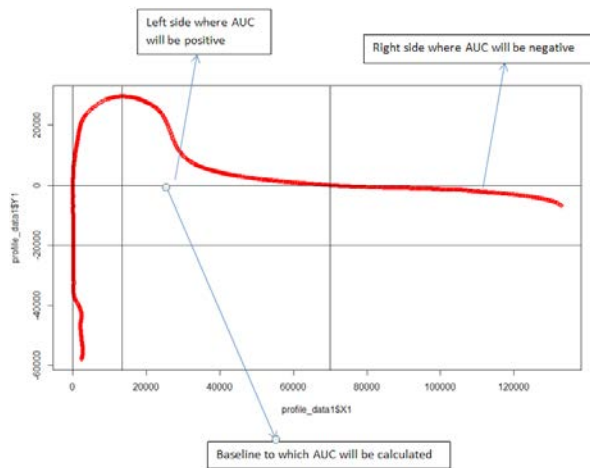


Fig. 4 | Computation of AUC

The steps involved in the computation of AUC are:

- For each profile starting from first available data from measuring station, the AUC is computed using a baseline (so the computation will be the area between profile curve, the red line as shown in the above figure 4, and the baseline, such as at $Y=0$).
- Sum all the absolute values to arrive at the final value of the area.
- Similar computations are then done for the rest of the available profile historical data for each of the wheels.
- The difference in area, in time, will be equivalent to wear loss in operations, until lathing operation takes place. This is shown in figure 5 below. The blue line represents

the first profile of wheel and red shows the second profile. The difference between the curves is shown in green.

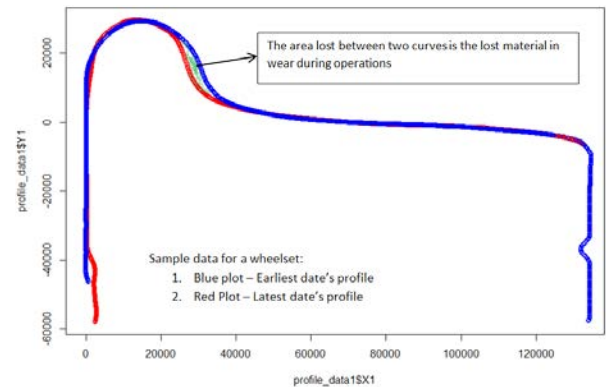


Fig. 5 | Computed AUC

- Similarly, the computations are performed for wheel profiles immediately after first or subsequent lathing operations.

Following this, a monitoring system was developed using ensemble models to enable near real-time assessment of wheel deterioration (wear). The models developed incorporated various data sources such as weather conditions, shear forces (to compute impact loading profiles on the rails), repair, maintenance and replacement history, as well as inventory data. This system generates alerts to the operations and the maintenance managers of the fleet. A wheelset with high wear loss or a high loss in the subsequent lathing operation is prioritized. This enables managers to schedule the various tasks such as allocations of assets for operations and maintenance.

The simulation engine enables the what-if analysis by the management to study the trade-off between maintenance frequency, costs and asset availability, and overall lifetime utility of the assets.

Results

Analysis validated the hypothesis that there is a high correlation between the operational material loss in a wheel and the material loss in the subsequent lathing. A robust framework

was developed to monitor the condition of the rolling stock asset (wheelset) and manage its maintenance to enhance its lifetime availability. The robustness of the framework pertains to data-related issues such as missing or inaccurate data from the monitoring stations and asset provenance. The analytical framework is validated on historical data and ready for implementation in the production environment.

Conclusion

Big data offers new possibilities in collecting and processing data for monitoring the condition of heavily used assets in rolling stock. Advanced analytics techniques facilitate the mining of this data for information that enables intelligent decision making. This paper presents a framework for "Data analytics-enabled condition monitoring for rolling stock." A case study is also presented wherein disparate data from operational, maintenance, weather, inventory, and other sources were analyzed in an integrated fashion. Wheelset condition was monitored on a near real-time basis, and actionable alerts were generated for the maintenance and operations teams. Additional insights include improvements in data acquisition and processing and changes in maintenance processes. A downstream benefit of this includes a reduction in inventory holding that will result from the improved asset availability.

About the Authors

Manoj Kumar

Manoj Kumar is an analytics professional and a subject matter expert for mining with over 13 years of work experience. He has more than five years of hands-on experience in developing business insights using advanced machine learning techniques and has been a successful contributor to multiple projects. He has also managed delivery for teams spread across geographies while being adept at

identifying and defining analytics solution to diverse, complex business problems across industry verticals.

Abraham KP

Abraham KP is a Principal Data Scientist with Cyient Insights. He has over 13 years' of experience in advanced analytics modeling and business consulting. His experience includes Manufacturing, Quality Management, Process Control, Supply Chain Management, and Inventory Optimization. He has worked with clients in Rail, Heavy Engineering, Medical Electronics, Automobile, Oil and Gas, and Finance domains.

References

- ¹ Lehasab N, et al., "Industrial fault diagnosis: Pneumatic train door case study," Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, Vol 216, Issue 3, 2002.
- ² Pereira P., Ribeiro R.P., Gama J., "Failure Prediction – An Application in the Railway Industry." In: Džeroski S., Panov P., Kocev D., Todorovski L. (eds) Discovery Science. DS 2014. Lecture Notes in Computer Science, Vol 8777. Springer, Cham
- ³ Dinmohammadi, F., Alkali, B., Shafiee, M. et al. "Risk Evaluation of Railway Rolling Stock Failures Using FMECA Technique: A Case Study of Passenger Door System," Urban Rail Transit (2016), Volume 2, Issue 3, pp 128–145. doi:10.1007/s40864-016-0043-z
- ⁴ Door Diagnostics - Product Innovation & Technology - Case Study, Available at <http://www.unipartrail.com/case-study---door-diagnostics.html>
- ⁵ Laney, Doug " 3-D Data Management: Controlling Data Volume, Velocity, and Variety", Application Delivery Strategies (File No. 949), Pub: Gartner-META Group, Feb 2001.
- ⁶ O. Fink, E. Zio, and U. Weidmann, "Fuzzy Classification With Restricted Boltzman Machines and Echo-State Networks for Predicting Potential Railway Door System Failures," in IEEE Transactions on Reliability, vol. 64, no. 3, pp. 861-868, Sept. 2015.

About Cyient

Cyient (Estd: 1991, NSE: CYIENT) provides engineering, manufacturing, geospatial, networks, and operations management services to global industry leaders. We leverage the power of digital technology and advanced analytics capabilities, along with domain knowledge and technical expertise, to solve complex business problems. As a Design, Build, and Maintain partner, we take solution ownership across the value chain to help our clients focus on their core, innovate, and stay ahead of the curve.

Relationships lie at the heart of how we work. With more than 15,000 employees in 22 countries, we partner with clients to operate as part of their extended team, in ways that best suit their organization's culture and requirements. Our industry focus spans aerospace and defense, medical, telecommunications, rail transportation, semiconductor, utilities, industrial, energy and natural resources.

For more information, please visit www.cyient.com

Contact Us

North America Headquarters

Cyient, Inc.
99 East River Drive
5th Floor
East Hartford, CT 06108
USA
T: +1 860 528 5430
F: +1 860 528 5873

Europe, Middle East, and Africa Headquarters

Cyient Europe Ltd.
The Space Holborn
235 High Holborn
London WC1V 7LE
UK
T: +44 20 7404 0640
F: +44 20 7404 0664

Asia Pacific Headquarters

Cyient Limited
Level 1, 350 Collins Street
Melbourne, Victoria, 3000
Australia
T: +61 3 8605 4815
F: +61 3 8601 1180

Global Headquarters

Cyient Limited
Plot No. 11
Software Units Layout
Infocity, Madhapur
Hyderabad - 500081
India
T: +91 40 6764 1000
F: +91 40 2311 0352