Proceedings of the Sixth International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 7, Paper 6.11
Civil-Comp Press, Edinburgh, United Kingdom, 2024
ISSN: 2753-3239, doi: 10.4203/ccc.7.6.11

Condition Based Maintenance for Railway Turnouts

D. Garcia Sanchez[1], G. Arteta[2], P. Pascual[3] and P. Infante[3]

[1]TECNALIA, Derio, Spain
[2]TECSA Empresa Constructora S.A. (TECSA), TECSA Empresa Constructora S.A. (TECSA), Bilbao, Spain
[3]Grupo de Investigación de Tecnología de la Construcción (GITECO) , University of Cantabria, Cantabria, Spain

Abstract

This article focuses on the specific study of special type A turnout. Today, this type of track apparatus is inspected by visual reconnaissance of the tracks and using specialized measuring equipment to detect irregularities in the rails such as wear or deformation. Both the visual recognition and the measurements made are recorded in a control form that is then evaluated in order to determine the necessary control action.

Thus, this article presents an algorithm based on data analysis that allows us to evolve towards a predictive maintenance model for special track segments.

It comprises the following main technical objectives: Analysis of the potential of data-driven anomaly detection methods, proposing a new approach that incorporates machine learning techniques through statistical pattern recognition. Diagnosis or evaluation of the condition of the track apparatus that allows the fault to be detected, identified, or located. Implementation of a valuable tool that allows the evolution of the maintenance strategy towards predictive maintenance management. Recommendation in terms of maintenance.

Keywords: railway turnout, condition based maintenance, principal component analysis, manual inspection, visual inspection, damage detection.

1 Introduction

In Spain, there are six types of standard turnouts depending on their length and the maximum permitted passage speed [1] and this paper focuses on Type A vehicles: maximum speed of 140 km/h and 30 km/h per deviation.

There is a set of parameters that allow the condition of this type of switchgear to be measured and evaluated. These parameters, which will later be used to implement a statistical pattern recognition algorithm, are defined in the ADIF NAV 7-3-8.2 standard [2].

2 State of the Art

According to references [3], models used to assess track conditions for diagnostic and prognostic purposes can be grouped into mechanical models and data-driven models.

Data-driven methods uncover sets of workable characteristics and decision criteria from observed data. These methods include statistical modelling [6... 10] and Machine Learning models [4] [11... 14]. The main difference between these two types lies in the main purpose of the analysis. Statistical models make inferences about the relationships between variables, while machine learning models focus on making accurate predictions. Both types can handle multivariate and high-dimensional data and extract hidden relationships between track state and measurement data. Overall, data-driven methods can help railway engineers better understand the condition of the railway track and make corresponding maintenance decisions. However, the performance of data-driven methods depends on the proper choice of data preprocessing and analysis models.

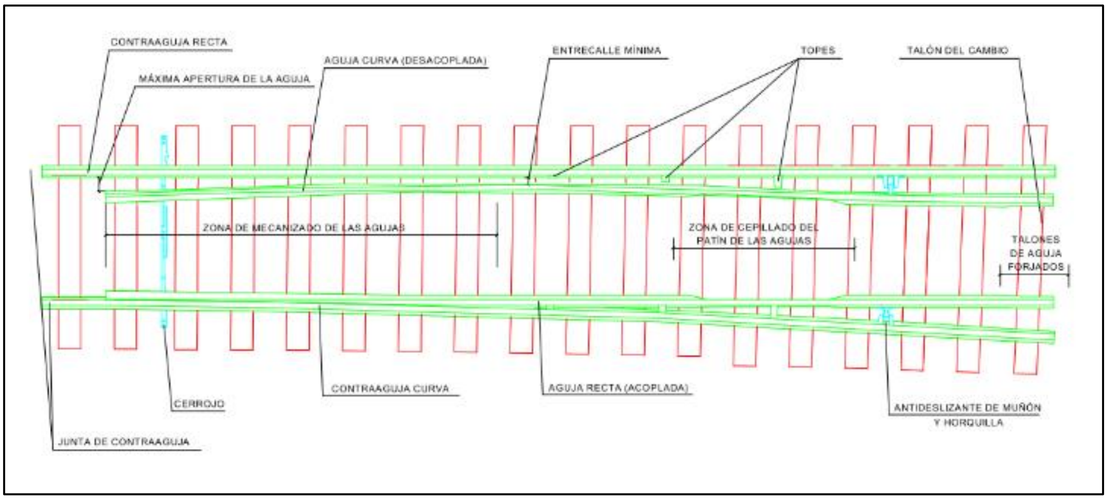

Figura 1. Track change zone within a railway turnout [5]

3 Case Study

3.1 Analysis of the initial data

This paper presents a simple tool that integrates functions for diagnosis, detection, and prediction of faults in track devices based on Machine Learning capable of recognizing hidden patterns among the variables that define the system.

This study has been developed for the specific study of type A switchgear. These devices are examined through visual inspections of the tracks and using specialized measuring equipment to detect possible irregularities in the rails, such as wear or deformation. All the information collected during the visual reconnaissance and measurements of the variables that determine the structural condition of the track apparatus is recorded on a control sheet.

From this control form, a first list can be extracted with all the parameters or variables that describe the state of health of the track apparatus. This first list has been structured in table format and contains the forty-one variables extracted from the control sheet along with a brief description of what each of them measures, an operating threshold, and theoretical values.

It is important to note that the range of operation contained in the description of each variable will be fundamental in the generation of synthetic data, to implement values that represent a healthy state of the track apparatus and with which the model can be trained.

The objective of this phase is to define a data vector that collects the values of each variable, in such a way that each data vector represents a completed inspection, thus generating datasets or datasets with which to work in the modelling phase. But to do this, it will be necessary to study the optimal dimension of this vector, that is, it will be necessary to carry out an analysis of the variables and filter those that are essential and that provide the maximum information. In this way, we seek to define the system to be modelled in the most simplified way possible so that the implemented model is optimal and provides results with high precision.

Tipo A (v<30km/h), [1668 mm]	LEYENDA: Actuación programada Actuación inmediata	VALORES LÍMITES MÍNIMOS: (1) 1.626 mm Valor ETI (2) 70 mm Valor ETI Cota de seguridad Valor ETI	ESPACIO RESERVADO PARA LAS EMPRESAS:
TIPO APARATO: (Seleccionar) Tipo A (v>30km/h)[1668 mm]	UBICACIÓN: Irun	Nº APARATO:T7	LINEA:
FECHA: 10/10/2021			

CAMBIO (*)		1er Cerrojo		2º Cerrojo		3er Cerrojo		4º Cerrojo	
APERTURA EN BIELA (**)		V direc	V desv	V direc	V desv	V direc	V desv	V direc	V desv
	Real								
Nominal:165 mm; Tol.: [-15, 15] mm	Corregido								
ENCERROJAMIENTO (**)		V direc	V desv	V direc	V desv	V direc	V desv	V direc	V desv
	Real								
Nominal:45 mm; Tol.: [-10, 10] mm	Corregido								

ANCHOS Y PERALTES DE VÍA DIRECTA					ANCHOS Y PERALTES DE VÍA DESVIADA				
ANCHO [-2, 10] mm Progr [-3, 15] mm Inmediata			PERALTE [-5, 5] mm Progr [-10, 10] mm Inmediata		ANCHO [-2, 10] mm Progr [-3, 15] mm Inmediata			PERALTE [-5, 5] mm Progr [-10, 10] mm Inmediata	
TRAVIESA	TEORICO	REAL	TEORICO	REAL	TRAVIESA	TEORICO	REAL	TEORICO	REAL
-15	1668	1661,0	0						
-10	1668	1665,0	0						
-5	1668	1668,0	0						
JCA	1668	1674,0	0						
5	1668	1668,0	0		5	1668	1682,0	0	
10	1668	1661,0	0		10	1668	1682,0	0	
15	1668	1658,0	0		15	1668	1684,0	0	
20	1668	1668,0	0		20	1668	1682,0	0	
25	1668	1668,0	0		25	1668	1680,0	0	
30	1668	1667,0	0		30	1668	1678,0	0	
35	1668	1667,0	0		35	1668	1675,0	0	
40	1668	1669,0	0		40	1668	1673,0	0	
45	1668	1668,0	0		45	1668	1665,0	0	
	1668		0			1668		0	
	1668		0			1668		0	
	1668		0			1668		0	
	1668		0			1668		0	
	1668		0			1668		0	
	1668		0			1668		0	
	1668		0			1668		0	

VIA DIRECTA		VIA DESVIADA	
PASO LIBRE DE RUEDA EN EL CAMBIO: < 1618 mm			
REAL	CORREGIDO	REAL	CORREGIDO
ENTRECALLE MINIMA: ≥ 58 mm; Tol.: < 58 mm Progr, < 55 mm Inmed			
REAL	CORREGIDO	REAL	CORREGIDO
64		88	
(1) COTA DE PROTECCION: 1628 mm; Tol.: [-2, 3] mm Prog. [-2, 5] mm Inmediata			
REAL	CORREGIDO	REAL	CORREGIDO
1626,0		1625,0	
ENTRECALLE CARRIL - CONTRACARRIL: 40 mm Tol.: [-2, 5] mm			
REAL	CORREGIDO	REAL	CORREGIDO
41		42	
PASO DE RUEDA LIBRE EN LAS PUNTAS DEL CRUZAMIENTO: < 1590 mm			
REAL	CORREGIDO	REAL	CORREGIDO
1586		1573	
PASO DE RUEDA LIBRE EN LA ENTRADA DE PATA LIEBRE: <1620 mm			
REAL	CORREGIDO	REAL	CORREGIDO
1591		1592	
PASO DE RUEDA LIBRE EN LA ENTRADA DE CONTRACARRIL: <1620 mm			
REAL	CORREGIDO	REAL	CORREGIDO
1610		1597	
ANCHURA DE LA GARGANTA DE GUÍA: >38 mm			
REAL	CORREGIDO	REAL	CORREGIDO
61		61	
PROFUNDIDAD DE LA GARGANTA DE GUÍA: >40 mm			
REAL	CORREGIDO	REAL	CORREGIDO
43		43	
(2) SOBREELEVACIÓN DE CONTRACARRILES: 20 mm Tol.: [-5, 5] mm			
REAL	CORREGIDO	REAL	CORREGIDO
1		1	

Figure 2. Data set including measures from manual inspection.

Next, a study of each variable was carried out, analysing the characteristics, functionality and information provided by each of them. To carry out this study, ADIF Technical Instructions were used, which describe the characteristics of each of the parameters that measure the condition of the track apparatus, as well as the methodology followed for the inspection of each of them.[1][15]

This study revealed the existence of two relationships between the six most relevant parameters detailed below: wheel clearance at change, track width (wheel clearance), minimum spacing, track width (protection level), protection height, lane spacing and counter-rail.

After carrying out an in-depth study and analysing the characteristics and functions of each of the variables, a first screening was carried out of some variables that, according to the criteria adopted in this project, did not provide useful information for the implementation of the model that is sought to be developed. Specifically, fourteen variables were discarded, all of them were track gauge measurements at multiple points that have been considered irrelevant to the case study of this project.

Later, two new parameters were added to the table of variables, "Wheel clearance on the change on the direct track" and "Wheel clearance on the deviated track". These two parameters appeared on the initial control sheet but had not been considered because there was no data on them. However, after the study and documentation carried out on the parameters measured in the inspection of switchgear, it was concluded that these two parameters provided broadly relevant information on the condition of type A turnouts.

Consequently, the table of variables was readjusted to include twenty-nine variables that defined a data vector of twenty-nine components.

Even though the vector is too large to make data processing and modelling difficult, we moved on to the next phase in order to start programming the algorithm and perform the first tests with the resulting data vector.

3.2 Methodology

Principal Component Analysis (PCA) is an unsupervised machine learning technique widely used in various areas, such as data analysis, pattern recognition, computer vision, and bioinformatics, among others. Its main purpose is to analyse and reduce the dimensionality of multivariate datasets. To do this, it tries to find a new reduced set of variables, called principal components, that capture most of the variability present in the original data.

PCA is based on the idea that multidimensional data often contains redundancy or correlation between variables. It seeks to transform the original data into a new coordinate system in which the variables must meet the condition of non-correlation with each other.

As is known, principal components are obtained as linear combinations of the original variables, where the first principal component captures the largest possible variance in the data, the second principal component captures the next largest variance, and so on (Figure 3).

The principal component analysis process consists of the following steps:

- Data standardization: A transformation is applied to the data so that all variables have mean zero and standard deviation one.
- Calculation of the covariance matrix: The covariance matrix or correlation of the standardized data is calculated, which shows the covariance or correlation relationships between all the variables.

- Principal component calculation: The eigenvectors (also called eigenvectors) and eigenvalues (also called eigenvalues) of the covariance matrix are calculated.
- Principal component selection: A set number of principal components are selected to retain, usually based on the cumulative explained variance.
- Data Transformation: The original data is projected into the new principal component space. To do this, the standardized data array is multiplied by the eigenvector array corresponding to the selected principal components. The result of this projection is the data transformed into the new coordinate system.

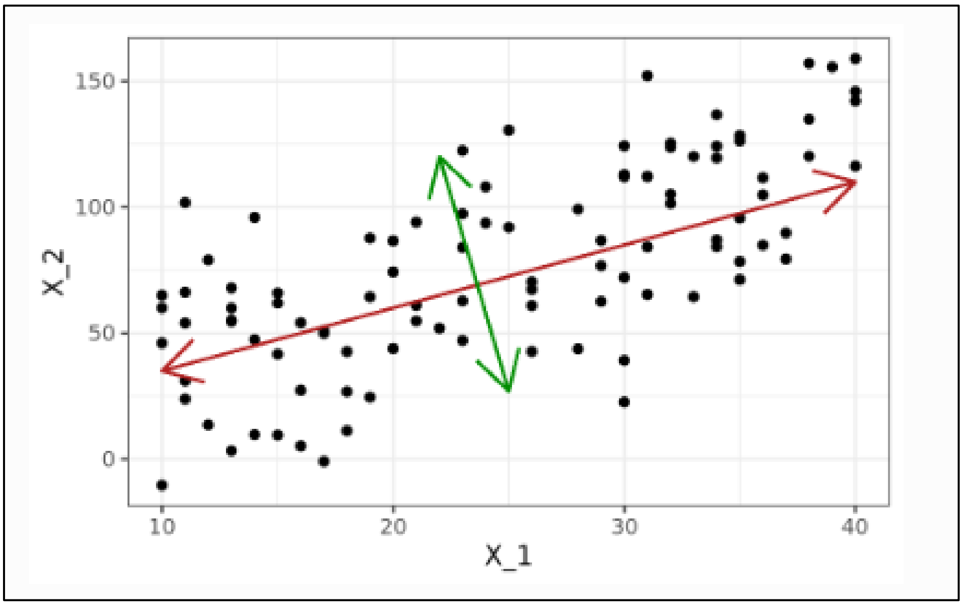

Figure 3. Example of a New Principal Component Coordinate System

3.3 Validation

An example program has been developed in Python where one hundred synthetic data samples are generated for the twenty-nine variables that make up the data vector.

For the sake of visibility and efficiency in programming, the synthetic data matrix was programmed in such a way that the rows were the variables of the data vectors, and the columns were the various synthetic samples.

The "Data Frame" function of the "pandas" library transforms the array-type structures generated with the "NumPy" library into data frames.

The line of research was to implement a model that would detect correlations when the data showed anomalous situations that foreshadowed that a failure was going to occur or that there was a defect in the structure.

Taking this approach into account, a study was carried out to test the influence of the number of samples in a dataset on the accuracy of the PCA technique.

The study consisted of implementing an algorithm that generated datasets composed of six variables and different sample sizes (10,100,1000 and 10000), applying the PCA technique to them and iterating this process a considerable number of times, with the aim of projecting the variation in the proportion of accumulated variance yielded by the PCA for each sample size. Specifically, the cumulative proportion of the third main component was compared as it provided quite enlightening results. The results of the study can be seen graphically in Figure 4.

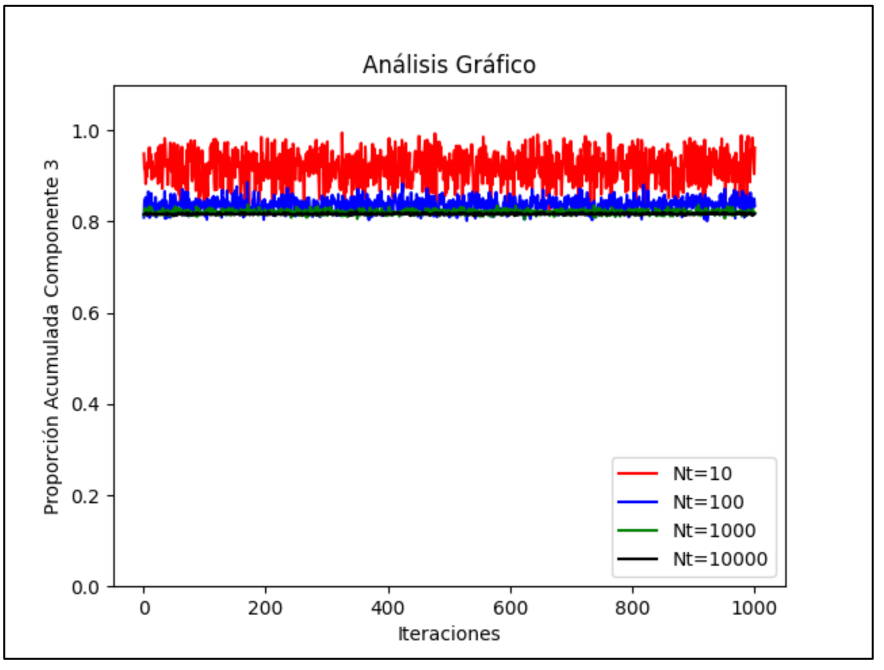

Figure 4. Influence of Sample Size on PCA Accuracy

Figure 5 shows a comparison between a dataset containing healthy values and a dataset with the variable "track gauge at minimum spacing measurement point" synthetically modified with the growth function described above. An iterative process of 1000 iterations are carried out with the aim of obtaining conclusive results.

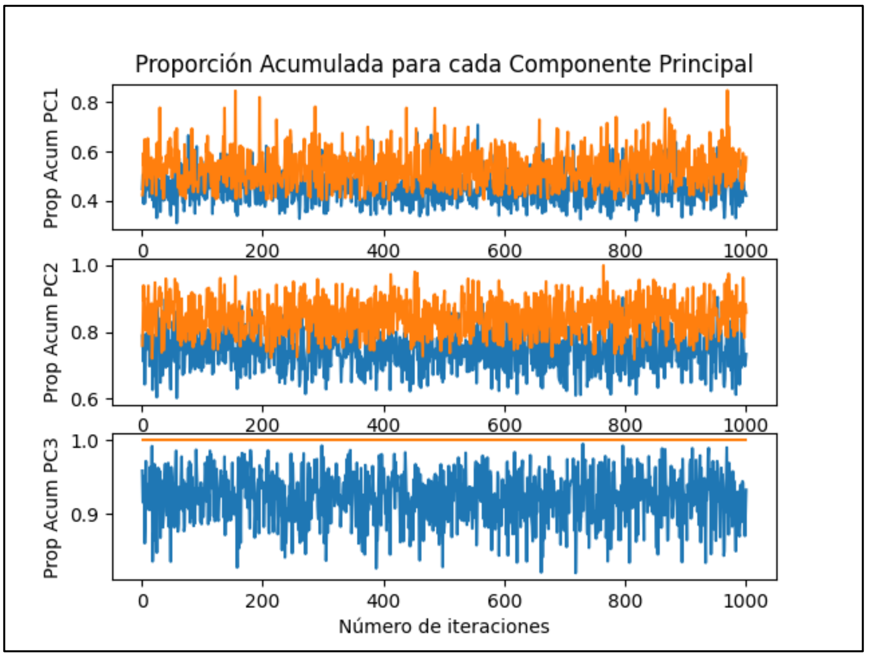

Figure 5. Validation test of the "track gauge at the minimum spacing measurement point"

Just as no conclusions can be drawn for the first two main components, the cumulative proportion of the third component reflects a significant distinction between the dataset generated with healthy data and the dataset modified with the synthetic growth function. This distinction of results between the dataset in good health and the dataset that simulates a synthetic anomaly validates the implemented fault detection and prediction algorithm.

The algorithm is also capable of detecting if there has been a specific failure in the needle, a fundamental element of the track apparatus subjected to high mechanical stress due to dynamic loads, lateral forces, impacts and vibrations. This type of fault is difficult to detect because the wear produced on the needle is invisible in the measurement of the "track width", since it would increase the "free passage of the wheel in the gearbox" but at the same time decrease the "minimum spacing" in the same proportion. This is due to the geometric conditions of the structure.

Finally, it should be noted that the algorithm is also capable of detecting a simultaneous failure in two different areas of the device. This feature gives the algorithm greater detection capacity, making it more robust and efficient.

4 Conclusions

This paper has presented a tool for detecting anomalies that can reflect a failure state in the structure or lead to a sudden failure. This tool is an algorithm implemented with a statistical pattern recognition technique such as PCA, which is currently widely used in the field of Machine Learning with the aim of improving operational efficiency and helping in decision-making in any field or sector.

The developed algorithm has great potential for application, since it allows to improve the maintenance plan of railway systems, whose current maintenance strategy consists of a procedure for the application of preventive tasks. The improvement would mean an evolution towards a predictive maintenance plan that would reduce the number of unplanned downtimes, improve safety and reduce operating costs.

This algorithm is a verified and validated pilot tool (synthetic data) for specific specifications. If real data on the variables treated in this project are known, an exhaustive study of these variables could be carried out and the pilot tool could be adjusted so that it would be able to detect possible trends or hidden correlations between the different parameters, well in advance to be able to carry out predictive maintenance.

Despite the fact that the algorithm has been designed for a specific type of special track segment, with certain characteristics and parameters, the methodology followed in the development of this project is applicable to any railway infrastructure. This gives the work carried out great versatility and opens up a range of new lines of research.

Acknowledgments

The authors are grateful for the financial support provided by LIAISON (HORIZON-CL5-2022-D6-02-06), by BIDEAI (Program HAZITEK 2021 and Core funding ZL-2021/00724 supported by the Basque Government) and 4RAIL (PROYECTOS EN COLABORACIÓN PÚBLICO-PRIVADA 2021 Proyecto CPP2021/008374 NextGenerationEU).

References

[1] Dirección Técnica (RENFE) and Dirección de Mantenimiento de Infraestructura (RENFE), “NAV 3-6-0.1: DESVÍOS. Características de los tipos y modelos.,” Jul. 1992.

[2] Grupo de trabajo GT-208, “NAV 7-3-8.2: INSPECCIÓN DE APARATOS DE VÍAS,” Feb. 2022.

[3] A. K. S. Jardine, D. Lin, and D. Banjevic, “A review on machinery diagnostics and prognostics implementing condition-based maintenance,” Mechanical Systems and Signal Processing, vol. 20, no. 7. 2006. doi: 10.1016/j.ymssp.2005.09.012.

[4] A. Falamarzi, S. Moridpour, and M. Nazem, “A review of rail track degradation prediction models,” Australian Journal of Civil Engineering, vol. 17, no. 2. 2019. doi: 10.1080/14488353.2019.1667710.

[5] “Aparatos de vía: los desvíos ferroviarios,” https://masqueingenieria.com/blog/aparatos-de-via-los-desvios/, 2016.

[6] J. Sadeghi and H. Askarinejad, “Development of improved railway track degradation models,” Structure and Infrastructure Engineering, vol. 6, no. 6, 2010, doi: 10.1080/15732470801902436.

[7] M. Audley and J. D. Andrews, “The effects of tamping on railway track geometry degradation,” Proc Inst Mech Eng F J Rail Rapid Transit, vol. 227, no. 4, 2013, doi: 10.1177/0954409713480439.

[8] C. Jia, W. Xu, F. Wang, and H. Wang, “Track irregularity time series analysis and trend forecasting,” Discrete Dyn Nat Soc, vol. 2012, 2012, doi: 10.1155/2012/387857.

[9] H. F. Lam, J. H. Yang, Q. Hu, and C. T. Ng, “Railway ballast damage detection by Markov chain Monte Carlo-based Bayesian method,” Struct Health Monit, vol. 17, no. 3, 2018, doi: 10.1177/1475921717717106.

[10] L. Bai, R. Liu, Q. Sun, F. Wang, and P. Xu, “Markov-based model for the prediction of railway track irregularities,” Proc Inst Mech Eng F J Rail Rapid Transit, vol. 229, no. 2, 2015, doi: 10.1177/0954409713503460.

[11] H. Guler, “Prediction of railway track geometry deterioration using artificial neural networks: A case study for Turkish state railways,” Structure and Infrastructure Engineering, vol. 10, no. 5, 2014, doi: 10.1080/15732479.2012.757791.

[12] Y. Jiang, H. Wang, G. Tian, Q. Yi, J. Zhao, and K. Zhen, “Fast classification for rail defect depths using a hybrid intelligent method,” Optik (Stuttg), vol. 180, 2019, doi: 10.1016/j.ijleo.2018.11.053.

[13] A. Falamarzi, S. Moridpour, and M. Nazem, “Development of a tram track degradation prediction model based on the acceleration data,” Structure and Infrastructure Engineering, vol. 15, no. 10, 2019, doi: 10.1080/15732479.2019.1615963.

[14] C. W. Tan, G. I. Webb, F. Petitjean, and P. Reichl, “Machine learning approaches for tamping effectiveness prediction,” 2017 International Heavy Haul Association Conference (IHHA, no. September, 2017.

[15] Dirección Técnica (RENFE) and Dirección de Mantenimiento de Infraestructura (RENFE), “NAV 7-5-3.1: CONSERVACIÓN DE LA VÍA. Mantenimiento de desvíos y otros aparatos de vía,” 1995.