Proceedings of 7th Transport Research Arena TRA 2018, April 16-19, 2018, Vienna, Austria

Automatic prediction of maintenance intervention types in transport linear infrastructures using machine learning

Antonio Reyes[a], Francisco J. Morales[a], Noelia Caceres[b], Luis M. Romero[a], Francisco G. Benitez[a]*, Joao Morgado[c], Emanuel Duarte[c], Teresa Martins[c]

[a]*Transportation Engineering, Faculty of Engineering, University of Seville, Camino de los Descubrimientos, s/n, Seville 41092, Spain*
[b]*Transportation Research Unit, AICIA, Seville, Spain*
[c]*Infrastruturas de Portugal, Almada, Portugal*

Abstract

This article presents a methodology to automate the prediction of maintenance intervention alerts in transport linear infrastructures. It combines current and predicted asset conditions with operational and historical maintenance data, to predict the needed tasks to avoid later severe degradation. By means of data analytics and machine learning models, a prioritised listing ranked on severity level corresponding to the alerts generated for all assets of the infrastructure is inferred. The scientific part presents: a discussion on relevant data to train machine learning algorithms in order to generate reliable predictions of the interventions to be carried out in further time scenarios, a schematic flow chart of the automatic learning procedure, and the self-learning rules from automatic learning from false positive/negatives. The empirical part describes a road network pilot case, the available historical data information, measurements, maintenance interventions, and a selected set of outcomes.

Keywords: Road; linear infrastructure; predictive maintenance; machine learning.

* Corresponding author. Tel.: +34-954-487-315
E-mail address: benitez@us.es

1. Introduction

Predictive maintenance for linear transport infrastructures has attracted considerable interest in the last two decades. The importance of pushing forward the development of reliable predicting tools, in order to maximise the availability of road and rail networks and to optimise the resources devoted to keep them serviceable, is of a paramount importance for society. Currently, from the managerial facet, several linear-asset management professional codes (e.g. Bentley Exor, Bentley Optram, HDM-4, INFOR EAM, PAVER, PMS Core) include applications providing some sort of predictive capabilities on maintenance alerts and operations, which facilitates the forecasting task; though they are still at the point of departure with a long way to the goal.
In the technical foundation sphere, a major leap took place in the years 1940-60 with the introduction of the concept of reliability and the use of statistics and optimisation techniques (Dekker, 1996), which gave rise to Reliability Centered Maintenance (Moubray, 1998), one of the most used tool nowdays. A second step forward had got with the use of computing data analytics for inferring data-driven models with the purpose of estimating the evolving condition of particular assets and components. Since the appearance of automated data-driven techniques (e.g. artificial intelligence, expert systems, statistical learning) in the 80s, the organisation in a scientific corpus (e.g. machine learning) in the 90s and 2000 and the recent popularisation (e.g. data mining), a huge interest has aroused in predicting the state of a road/rail infrastructure with the purpose to envisage the appropriate interventions and set maintenance planning programmes, according to the available resources and minimum impact to the infrastructure functioning.

The detection of maintenance alerts are, most customarily, based on inspecting the state/condition by means of visualising/auscultating/measuring explanatory features of the asset involved. The evolving of these features, estimated in either a quantitative or qualitative manner, using projection techniques or experience, and the late/immediate crosschecking with thresholds and limits (defined by technical standards prescribed by the corresponding infrastructure administration/regulator), has been the main tool to outlook a prospective malfunctioning. These thresholds are grounded on the accumulated knowledge acquired during a prolonged period of time in relation to the adequacy of the condition of the analysed assets, and they respond to a conservative envelope which guarantees the safety, integrity and right performance of the asset as a part of the system it works for. In order to improve the "predictive" capability of this procedure, new and additional explicative single and combined features have been considered and have been incorporated, during the past, to the listing of indexes to be monitored and measured. Even though, the large diversity of asset typologies and affecting factors make unlikely to envisage an ideal working case where any state/condition can be comprehensively explained using a fixed set of measurable features and, as further consequence, to get a reliable estimation of the evolution. This lack of determinism is intendedly overcome with the knowledge possessed by the maintenance team.

During decades, building either deterministic or probabilistic models based on aprioristic explicative features (i.e. mechanistic-empirical models) has been the tendency (e.g. Lytton, 1987; NCHRP, 2004; AASHTO, 2008; Mubaraki, 2010). At present, large efforts are invested in substituting that way of proceeding by data-driven modelling, using the very trendy wave of data mining and machine learning techniques (e.g. Quinlan 1986; Schwartz 1993; TRB 1999; Dick et al. 2003; Podofillini et al. 2006; Iqbal 2010; Podder 2010; Witten et al. 2011; Karlaftis et al. 2015; Plati et al. 2015), and laying on the increasing availability of data captured from auscultation/monitoring activities and campaigns. Thus, these intelligent (i.e. automatic and mining) techniques have promoted the concept of learning from data, facilitating the extraction of patterns and trends by "let the data speak by themselves". The use of these techniques, complemented with the historical information contained in the repository of work-orders can make a swift improvement of predictions.

This communication presents the methodologies, approaches and models (developed under the H2020 project INFRALERT) for triggering alerts associated to assets of road/rail linear infrastructures needed of maintenance interventions, be corrective or predictive. INFRALERT aims at the development of models and ICT tools to optimise the performance of existing linear land transport infrastructure. The concept and scope of INFRALERT is aimed at developing an expert based information system to support and automate infrastructure management from measurement to maintenance using a modular approach. For this purpose, it includes a data management system (the Data Farm) and a set of toolkits (asset condition, alert management and the RAMS & LCC) and a decision support tool which receives the results of these toolkits and optimises maintenance interventions. This paper presents the methodology framework and results obtained from the alert management toolkit. The estimated alerts are assessed according to the information provided by a decision making tool based on the forecast of the state evolution reflected by physical explanatory features, relevant to the condition of the assets of interest, and the historical interventions database. The output of the said tool will tag each estimated alert with a level of severity and will rank all alerts in a hierarchical listing of interventions and their associated probabilities of occurrence. The final purpose is to provide a procedure for managing all active and predicted interventions, optimizing maintenance operations.

2. Relevant information for a predicting approach

In the approach proposed, alerts are inferred by correlating the estimated values of the explanatory features (X_n) of the asset state, in a requested scenario of interest for the Maintenance Managerial Body (MMB), with the information stored in the historical maintenance work-order repository. Two different main data sources are involved. The first one is the measurements carried out in the linear infrastructure in which the values of the relevant features are included; the second one is the historical maintenance database which stores, at least: a) each intervention type conducted, and b) the identification of the measurements prior to the intervention. This makes the triggering of alerts based not just on comparing their estimated value of explanatory features with their pre-set thresholds, but also using the non-explicit information hidden in these data sources, which may explain the needed intervention carried out in past cases. This repository may also contain recorded information regarding the subjective assessment of the asset condition inspected by the MMB team just before the intervention; in particular, the subjective assessment of each single explanatory feature (SX_n), each combined explanatory feature (CX), and a global valuation of the asset state (G).

Besides, there are others endogenous and exogenous characteristics/variables affecting the asset state (condition) evolution. These characteristics complete the set of explicative features to be taken into account in the predictive intervention approach; this is for instance the traffic flow and road category, in a road network case; or the axle load and freezing index in the rail study case.

3. Procedural approach

The overall framework proposed is sketched in Fig. 1 where different modules (embodying techniques, methodologies, algorithms and models) and their interactions, inputs and outputs are shown. The approach aims at estimating and prioritising maintenance alerts, and predicting the required maintenance interventions. In particular, two kinds of alerts are predicted and involved in the proposed methodology: those triggered by the deviation of the predicted condition of an asset from standards (e.g. European Standard, Road Administration Standard, MMB); and those inferred by correlating recorded information from previous interventions.

Module AM1 is responsible for generating pre-alerts based on limits, from the point of view of those features that overcome their associated thresholds, using as inputs the forecasted values of the explanatory features of the asset. In particular, the goal of this module is to compare the value of each forecasted feature with the corresponding threshold (based on design/quality/safety parameters) to determine/quantify the asset state. As result, the following outcomes are provided: i) Pre-alerts (warnings) indicating that a specific feature exceeds its prescribed threshold, and ii) technical severity levels (TSL) of the estimated pre-alerts. The TSL is an objective value used to prioritise the pre-alerts according, for instance, based on a distance criterion between the value of the features and the thresholds. Module AM2 is responsible for predicting alerts based on work-orders (WO), from the point of view of whether a maintenance action is required (Yes/No); it also estimates the most probable maintenance interventions to be conducted. To achieve this, the module embodies two different functional submodules. The first one (AM21) is specifically devoted to triggering alerts regarding the need of maintenance (Yes/No) and their corresponding level of global technical severity (GTSL) in terms of all forecasted features considered as a whole. The GTSL is derived as a function of all TSLs of the corresponding explanatory features $X=\{X_1, X_2, \ldots, X_n\}$, which are previously normalised (in order to refer all values to the same scale) and weighted by pre-set values α_i to each individual feature X_i (subject to a constraint $\alpha_1 + \alpha_2 + \ldots + \alpha_n = 1$). Here, the alerts are triggered by the estimator contained in the first block (Alert Estimator) which has been previously trained with the adequate information (see section 2) by a machine learning (ML) processing. The learning methodology consists of an automatic classification in a binary variable (1-0: Yes/No). A set of four automatic binary classification models (i.e. DT-Decision Tree, ANN-Artificial Neural Network, KNN-K Nearest Neighbourhood, SVM-Support Vector Machine) was considered.

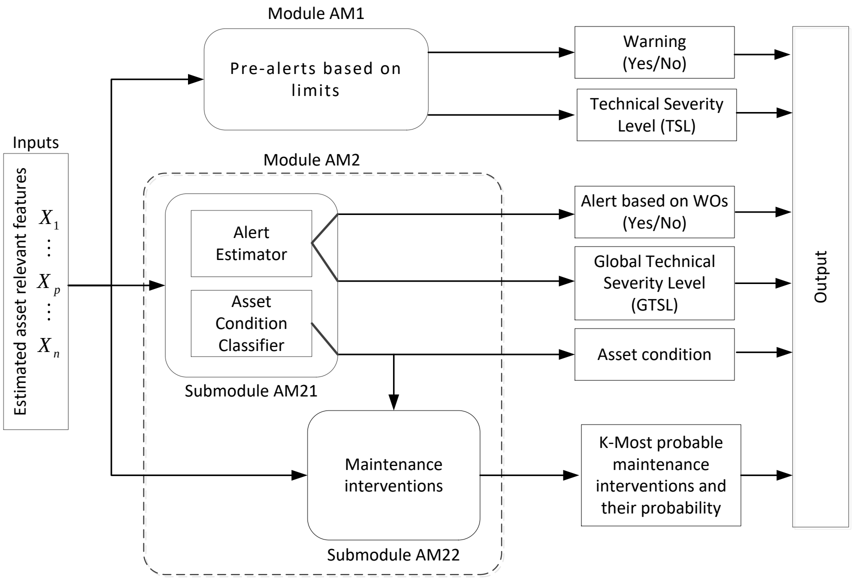

Fig. 1 Methodology framework.

The model which yields the best predictive capabilities was chosen though all four outputs show similar accuracies. Submodule AM21 also provides an optional output using a second block, Asset Condition Classifier, which "learns" from the MMB know-how, with the final purpose of predicting a subjective evaluation of the asset condition (from the set of forecasted features) without the intervention of the MMB. The second submodule (AM22) aims at determining the set of k-most probable maintenance interventions that have to be conducted, as well as their corresponding probabilities of occurrence, via a learning procedure based on historical intervention database. In order to estimate the most probable intervention type on a specific asset, according to the know-how contained in the historical database, a process has to be launched by correlating the estimated values of the relevant explanatory features ($\hat{X}$) and the forecasted subjective state/condition ($S\hat{X}, C\hat{X}, \hat{G}$) provided by the second block (Asset Condition Classifier) of submodule AM21; this process makes use of similar asset samples reported in the historical repository; the intervention type predictions are based on an unsupervised ML scheme using clustering and k-neighborhood techniques. As result, Module AM2 provides: i) triggered alerts identifying those assets where maintenance is required; ii) global technical severity level (GTSL) of those implied assets; iii) K-Most probable predicted interventions associated to each asset and triggered alert; and iv) the probabilities of occurrence of the most likely interventions.

4. The self-learning rules from automatic learning

Since the approach is based on machine learning techniques, a set of self-learning rules has been defined for improving the predictive capabilities. This section focuses on this point and the methodology used for discerning "false positives" and "false negatives" predictions, according to the following rules:

- A false positive arises when an estimate (from models) indicates a given condition has been reached, when it has not; it is commonly regarded as a "false alarm". Then, an erroneous positive case has been assumed.
- A false negative appears when an estimate (from models) indicates that no alert has been detected (i.e. the predicted asset state is right), while it was later detected by a corrective maintenance intervention; therefore, erroneously no failure/fault was forecasted.

Before getting into the various cases which may arise, Fig. 2 sketches a description of the way-of-proceeding to be followed by the maintenance team when a triggered alert is communicated; it also includes the pieces of information necessary to be recorded in the database. This may have either a "corrective" or "predictive" cause; the case of preventive intervention is not taken into account herein as it follows a predefined plan laid on specific rules. According to this, all recorded data have associated a timestamp field to identify the instant when any decision/action was made (when the alert is regarded as either "not-attended", "attended" or "intervened"). The false cases may arise when wrong estimates are due to: i) wrong forecasting of a feature, ii) wrong prediction for requesting maintenance (Yes/No), iii) wrong prediction for maintenance type. A description of the cases is presented in next paragraphs.

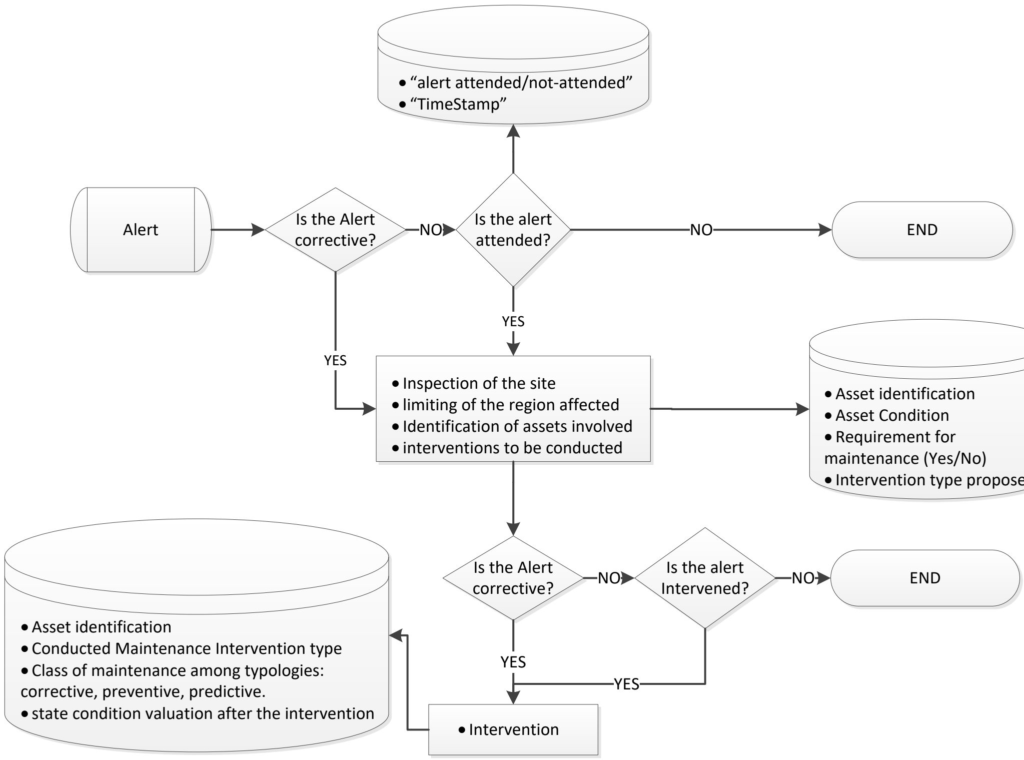

Fig. 2 Triggered-alert procedure diagram flow.

4.1. Wrong forecasting of a feature

Two possible cases, false positive and false negative, are analysed in separated manner.

a) The false positive case takes place when the alert is triggered by the expected value of a single feature, a prior combined feature, or a multiplicity of features, due to a wrong forecasting of the features values. The triggered-alert procedure, launched by the MMB, ends by assessing all assets implied in the alert, and by recording information on the condition of those assets. The new information will enrich the database corresponding to the condition of the assets involved. The machine learning procedure will learn from the enriched database information, improving the success rate according to the quality of the captured information in further prediction runs.
b) The false negative case arises when the expected value of the features (single, combined or a multiplicity of features) is below the value-to-be and also falls below the TSL (Technical Severity Level) threshold. In this case the triggered-alert procedure is not activated and no alert is triggered. Eventually a corrective warning may be brought up and the triggered-alert procedure activated, which will record the pieces of information identified in Fig. 2.

4.2. Wrong prediction for requesting maintenance

Similar to the previous paragraph, two possible cases, false positive and false negative, may arise:

a) False positive. This takes place when Submodule AM21 predicts the need of maintenance, and an alert is triggered. This case follows a similar pattern than case a) of section 4.1. The ML procedure will be improved according to the database enrichment.

b) False negative. This takes place when Submodule AM21 does not detect the need of maintenance and no alert is triggered. This case follows a similar pattern than case b) of section 4.1.

4.3. Wrong prediction for maintenance type

In this case, the requesting for maintenance is supposed to be correctly estimated as positive, otherwise the procedure goes back to section 4.2. Only the general "false type" case prediction is possible:

a) False type. The alert has been triggered due to a requesting for maintenance (Submodule AM21), and an estimated maintenance intervention type is provided. The Triggered-Alert Procedure is activated by the MMB, and the assets implied are assessed by the maintenance intervention team. The team detects that the maintenance type estimated does not corresponds to the predicted type; it means a wrong maintenance type prediction by submodule AM22 has taken place. The Triggered-Alert Procedure is continued, acting according the unit actions specified in the procedure. The information captured by this procedure (reflecting all pieces presented in Fig. 2) is recorded in the database, which enriches the stored information. Later estimations by the ML algorithms will be benefited by the new enriched database.

5. Case study

5.1. Description of pilot road case

The pilot case selected is a meshed road network in the central region of Portugal, managed by *Infraestruturas de Portugal*, totalling 620 km; it includes several road categories, as principal itineraries, supplementary itineraries, national roads, regional roads and other roads; the road categories are classified on the basis of features such as travel speed, traffic volume, traffic mix and strategic importance. Regarding traffic levels, the chosen demo case presents heterogeneity among the chosen sub-networks (i.e. itineraries), between 2,500 and 10,000 vehicles per day, with an average of 9% of heavy vehicles. This network is divided in 1,241 sections of 500 meters each.

Regarding the measurements data source, the records chosen corresponds to the period 2012-2016, when the auscultation retrieving system was laser-based. The raw data consists of a very rich database with a large number of features whose values are referenced to the position where the measurements were captured. The relevant features selected were the International Roughness Index (IRI), the Rutting (RUT) and the surface area with Crocodile Cracking (CT). Data was captured every 10 m of road and 100% of the net was monitored. These three explicative features are used for building the ML models. The evolution of the features constitutes a piece of information very relevant for the whole predictive maintenance intervention stage, and it conforms part of the inputs once the ML models are calibrated. The second data source, historical maintenance database, consists of a series of records since 1933. However, as this database has to be correlated with measurements, only the period between 2012 and 2016 was used. These records mainly contain the maintenance type and the position where the intervention was carried out.

Needless to say that a filtering pre-process has to be conducted to the recorded data for extracting the relevant information free of inconsistencies (e.g. unidentified/undetermined geometrical location of measurements, unclear description of the intervention carried out, assets affected, et-cetera), which may reduce a non-negligible percentage the final number of valid records. The number of filtered records in the historical interventions database is limited according to the network extension of the pilot case and the time period. In order to circumvent this drawback, a multiplicity of simulated data-sets derived from the statistics distributions of the available empirical real data was generated in order to select the most appropriate machine learning model.

5.2. Results

This section presents a selection of the results obtained after applying the presented methodologies to a set of real measurements on roads not included in the training process. That is, the model is trained using a specific partition of asset data (training set). The testing set is completely excluded from the training process and is used for independent assessment of the final models. A toolkit that includes the modules described in Fig. 1 has been implemented within the H2020 INFRALERT project. Fig. 3 shows a combined colour and scale-symbol theme map that exposes the results obtained in the road case. It can be seen the conditions, indicated with GTSL value by the line thickness, of all road sections of the pilot case in 2015 and the most recommended maintenance intervention proposed by the methodologies (line colour).

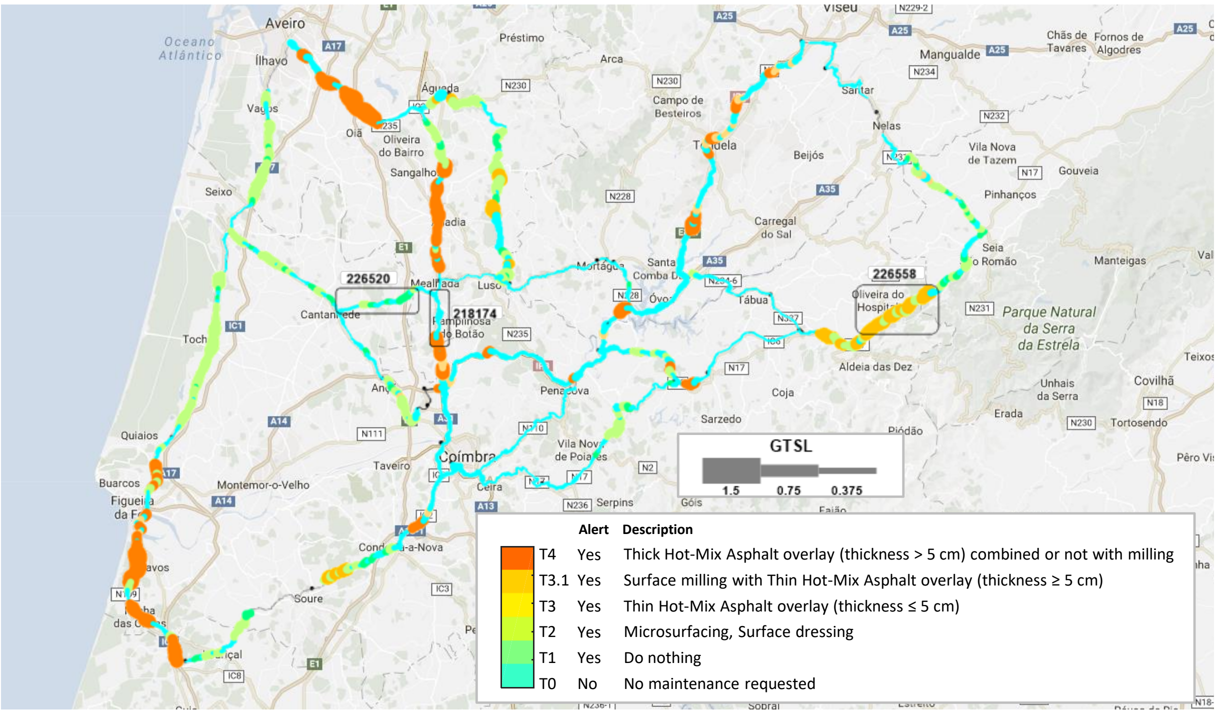

Fig. 3 Results obtained in the pilot case in 2015.

Fig. 4 shows the predictions on GTSL and Alerts for section 218174, segmented in subsections of 500 meters. As it was previously explained, the toolkit provides GTSL values based on the values of IRI, RUT and CT, taking into account the distance to the reference thresholds. In parallel, the toolkit generates the associated alerts for each subsection (500 m) of the road, revealing the sections where maintenance is needed by means of triggering alerts. For validation purposes, the last block contains the actual intervention carried out for each analysed year. This fact is close related to an inadequate asset condition. As it can be seen in year 2012, focusing on the values of GTSL, some subsections have acquired an inadequate asset condition (triggering the corresponding alerts). As no maintenance is executed these inadequate levels remain in a similar (or even worse) order of magnitude in the next year. Here, it is necessary to clarify that the values of GTSL for 2013 are based on measurements coming from a campaign previously performed to the maintenance intervention exhibited in the last block in Fig. 4 (intervention by year). Based on the new measurements for the following years (2014-2016), the toolkit detects this maintenance in 2013 since there are no alerts and the GTSL reflects that the asset conditions have been restored to adequate levels. Regarding the intervention necessary to attend an alert, the toolkit provides an objective solution, by indicating only those sections where maintenance is needed. However, the intervention is finally decided by the MMB taking into account other information as a whole, such as budget constraints, traffic interruption priorities, working teams travel times, and various interrelations among maintenance activities. In fact, these interrelations among maintenance activities motivate that some subsections not exhibiting alerts (with no need of intervention), are finally maintained to preserve the continuity of the asset condition levels. A last issue worth to be highlighted focuses on unattended alerts. As it can be seen in the last two subsections on the analysed road, alerts are predicted; however, they were not attended and the toolkit remains generating alerts over there. Together with alerts and GTSL, the toolkit also provides the prediction of the K most probable maintenance interventions that can be applied, according to historical interventions carried out in similar conditions, and their occurrence probability. This is just only a recommendation to be used by the MMB. As with the alert case, for which other information is taken into account as a whole, the MMB finally will decide the intervention to be (or not) carried out to solve the alert.

Road id			Global Technical Severity Level (GTSL) by year					Estimated alert by year					Intervention by year			
idSection	netClass	Section	GTSL_2012	GTSL_2013	GTSL_2014	GTSL_2015	GTSL_2016	2012	2013	2014	2015	2016	2012	2013	2014	2015
218174	2521	0	0,358	0,353	0,285	0,307	0,398	--	--	--	--	--		Maintened		
218174	2521	0,5	0,367	0,645	0,287	0,32	0,342	--	--	--	--	--		Maintened		
218174	2521	1	0,596	0,688	0,318	0,436	0,429	--	Alert	--	--	--		Maintened		
218174	2521	1,5	0,737	0,653	0,448	0,656	0,601	Alert	Alert	--	--	--		Maintened		
218174	2521	2	0,477	0,475	0,279	0,37	0,445	--	--	--	--	--		Maintened		
218174	2521	2,5	1,144	1,052	0,187	0,212	0,227	Alert	Alert	--	--	--		Maintened		
218174	2521	3	1,078	1,247	0,204	0,228	0,231	Alert	Alert	--	--	--		Maintened		
218174	2521	3,5	0,801	1,051	0,268	0,346	0,407	Alert	Alert	--	--	--		Maintened		
218174	2521	4	0,563	0,757	0,194	0,21	0,231	Alert	Alert	--	--	--		Maintened		
218174	2521	4,5	0,338	0,585	0,157	0,174	0,187	--	Alert	--	--	--		Maintened		
218174	2521	5	0,498	0,532	0,195	0,219	0,293	--	--	--	--	--		Maintened		
218174	2521	5,5	0,969	0,56	0,169	0,188	0,2	Alert	--	--	--	--		Maintened		
218174	2521	6	0,505	0,328	0,219	0,233	0,332	--	--	--	--	--		Maintened		
218174	2521	6,5	0,377	0,288	0,208	0,224	0,232	--	--	--	--	--		Maintened		
218174	2521	7	0,779	0,642	0,248	0,263	0,277	Alert	Alert	--	--	--		Maintened		
218174	2521	7,5	0,481	0,925	0,718	1,044	1,056	--	Alert	Alert	Alert	Alert				
218174	2521	8	0,313	0,979	0,804	1,024	0,862	--	Alert	Alert	Alert	Alert				

Fig. 4 Results on GTSL and Alerts for the section 218174, segmented in subsections of 500 meters, in the time period 2012-2016.

Other case derived from results is analysed in Fig. 5. In this case, it is straightforward to see the GTSL significantly decreases since 2015 but no maintenance intervention is reported in the historical data. After checking this point with the MMB, a non-reported intervention was actually detected. This was discovered by looking at the videos of inspections over the road. In this regard, it is necessary to clarify that sometimes missing/erroneous pieces of information are present in the historical maintenance repository and not all maintenance/work done on each piece of road are properly reflected. Even though the major interventions are mandatory to be registered, although these are less numerous than small corrective operations as major maintenance plans are conducted in intervals of 5-10 years; smaller maintenance are also executed on roads, normally in few meters of the road, and frequently only on one of the lanes. In general, minor maintenance is executed by contractors under 3-year contracts for all aspects of road maintenance (not only pavement). This type of maintenance is seldom registered in the historical data, except if they are of a more significant extent.

Road id			Global Technical Severity Level (GTSL) by year					Estimated alert by year					Intervention by year			
idSection	netClass	Section	GTSL_2012	GTSL_2013	GTSL_2014	GTSL_2015	GTSL_2016	2012	2013	2014	2015	2016	2012	2013	2014	2015
226520	2522	0	0,492	0,484	0,674	0,4	0,489	--	--	Alert	--	--				
226520	2522	0,5	0,55	0,517	0,294	0,268	0,26	Alert	Alert	--	--	--				
226520	2522	1	1,102	1,1	0,984	0,275	0,281	Alert	Alert	Alert	--	--				
226520	2522	1,5	1,101	1,109	0,921	0,282	0,299	Alert	Alert	Alert	--	--				
226520	2522	2	1,085	1,092	1,01	0,264	0,259	Alert	Alert	Alert	--	--				
226520	2522	2,5	0,65	0,71	0,607	0,299	0,281	Alert	Alert	Alert	--	--				
226520	2522	3	0,571	0,423	0,51	0,689	0,62	Alert	--	Alert	Alert	Alert				
226520	2522	3,5	1,172	1,057	1,185	0,59	0,625	Alert	Alert	Alert	Alert	Alert				
226520	2522	4	0,757	0,61	0,821	0,531	0,708	Alert	Alert	Alert	Alert	Alert				
226520	2522	4,5	0,263	0,249	0,45	0,334	0,596	--	--	--	--	Alert				
226520	2522	5	0,332	0,307	0,451	0,386	0,411	--	--	--	--	--				
226520	2522	5,5	0,249	0,252	0,24	0,313	0,306	--	--	--	--	--				
226520	2522	6	0,353	0,269	0,447	0,563	0,488	--	--	--	Alert	Alert				
226520	2522	6,5	0,373	0,255	0,391	0,43	0,268	--	--	--	--	--				
226520	2522	7	0,304	0,293	0,317	0,431	0,331	--	--	--	--	--				
226520	2522	7,5	0,343	0,372	0,486	0,705	0,705	--	--	--	Alert	Alert				
226520	2522	8	0,428	0,389	0,364	0,682	0,811	--	--	--	Alert	Alert				
226520	2522	8,5	0,813	0,465	0,701	0,736	0,272	Alert	--	Alert	Alert	--				
226520	2522	9	0,215	0,281	0,261	0,391	0,188	--	--	--	--	--				

Fig. 5 Results on GTSL and Alerts for the section 226520, segmented in subsections of 500 meters, in the time period 2012-2016.

There may be other situations (Fig. 6) where the GTSL estimates a road bad condition, such the case of roads with low traffic flow (less important road), the situation when the IRI level is rather high but accompanied by low cracking (CT), or when the IRI is low and CT high. In all these cases the sections are under low-priority maintenance based on diverse factors (e.g. low traffic, budget restrictions, other prominences).

Road id			Global Technical Severity Level (GTSL) by year					Estimated alert by year					Intervention by year			
idSection	netClass	Section	GTSL_2012	GTSL_2013	GTSL_2014	GTSL_2015	GTSL_2016	2012	2013	2014	2015	2016	2012	2013	2014	2015
226558	2522	0	1,228	0,991	0,967	0,649	1,1	Alert	Alert	Alert	--	Alert				
226558	2522	0,5	1,145	1,161	1,067	0,945	0,864	Alert	Alert	Alert	Alert	Alert				
226558	2522	1	1,194	1,301	1,226	1,165	1,361	Alert	Alert	Alert	Alert	Alert				
226558	2522	1,5	1,22	1,27	1,252	1,217	1,276	Alert	Alert	Alert	Alert	Alert				
226558	2522	2	0,625	1,113	1,066	1,103	1,154	Alert	Alert	Alert	Alert	Alert				
226558	2522	2,5	0,991	1,261	1,147	1,182	1,292	Alert	Alert	Alert	Alert	Alert				

Fig. 6 Results on GTSL and Alerts for the section 226558, segmented in subsections of 500 meters, in the time period 2012-2016.

6. Conclusions and prospective research lines

In this paper, a road network was studied from the maintenance interventions predictive approach. A methodology on several supervised and unsupervised machine learning techniques have substantiated the optimum choice of the best predictive models based on historical intervention work-orders, asset features and measurement auscultations. The main predicted outcomes are: a) the estimated intervention type for each road section and the probability of occurrence, b) a sorted out listing of estimated alerts according to the technical severity level. Each prediction set is referred to scenarios identified by its time-stamp. The results evidence that the methodology framework provides good predictive capabilities.

The methodologies and results presented herein are far from being exhaustive and conclusive, and several parallel lines of research are open: a) sensitivity to the quality of intervention description in the historical repository regarding the intervention timestamp, b) importance of the detailed/undetailed description of the asset state condition previous to the intervention, among others.

The work presented constitutes a step forward in generating a smart decision support tool to derive intervention plans based on alert forecasting generation and the optimal selection of activities regarding the most critical interventions to be carried out, which are the single bricks to build the final purpose of covering the full range of planning maintenance at operational, tactical and strategic level, under an expert intelligent framework.

Acknowledgements

The research has received funding from European Union's Horizon 2020 Research and Innovation Programme (grant agreement n° 636496). Some of the authors express their gratitude to the Spanish Ministry of Economy and Competitiveness for the partial subsidy granted under the national R&D program (TRA2015-65503) and the Torres Quevedo Programme (PTQ-13-06428). The content reflects only the authors' view and it is stated that the EU and IP are not liable for any use that may be made of the information contained therein.

7. References

AASHTO, 2008. *Mechanistic-Empirical Pavement Design Guide*. A manual of Practice. American MEPDG-1. Association of State Highway and Transportation Officials.

Bentley Exor. http://www.bentley.com/en-US/Products/Exor/.

Bentley Optram. http://www.bentley.com/Optram.

Dekker R., 1996. Applications of maintenance optimization models: a review and analysis, *Reliability Engineering & System Safety*, 51(3), 229-240.

Dick C.T., Barkan C.P.L., Chapman E.R., Stehly, M.P., 2003. Multivariate statistical model for predicting occurrence and location of broken rails. *J. Transportation Research Board*, 1825, 48-55.

HDM-4. http://www.hdmglobal.com/.

INFOR EAM. http://www.infor.com/solutions/eam.

Iqbal, Z., 2010. *A Study of AI techniques for Railheads, Vegetation, Switches & Crossings*. Master Thesis, E38440. Högskolan Dalarna.

Karlaftis A.G., Badr A., 2015. Predicting asphalt pavement crack initiation following rehabilitation treatments. *Transportation Research C*, 55, 510-517.

Lytton, R., 1987. Concept of Pavement Performance Prediction and Modelling. *Proceedings 2nd North American Conference on Managing Pavement*, Toronto, Canada.

Moubray J. *Introduction to relibility-centered-maintenance*. USA: TWI Press Inc.,1998.

Mubaraki M, 2010. *Predicting deterioration for the Saudi Arabia Urban Road Network*. PhD thesis, University of Nottingham.

NCHRP, 2004. *Guide for Mechanistic–Empirical Design of New and Rehabilitated Pavement Structures*. Final Report. National Cooperative Highway Research Program. TRB.

PAVER. http://www.paver.colostate.edu/.

Plati C., Georgiou P, Papavasiliou V., 2015. Simulating pavement structural condition using artificial neural networks. *Structure and Infrastructure Engineering*, 12(9), 1127-1136.

PMS Core. https://www.heller-ig.de/index.php?id=171.

Podder T., 2010. *Analysis & study of AI techniques for automatic condition monitoring of railway track infrastructure*. Master Thesis, E3845D. Högskolan Dalarna.

Podofillini L., Zio E., Vatn J., 2006. Risk-informed optimisation of railway tracks inspection and maintenance procedures. *Reliability Engineering and System Safety* 91(1), 20–35.

Quinlan J.R., 1986. *Induction of decision trees*. Machine Learning 1(1),81–106.

Schwartz C.W., 1993. Infrastructure condition forecasting using neural networks.*Proceedings: Infrastructure Planning and Management*, Denver, CO.

TRB A2K05(3), 1999. Use of artificial neural networks in geomechanical and pavement systems. Transportation Research Circular E-C012, *Transportation Research Board*, NRC. Retrieved from http://onlinepubs.trb.org/onlinepubs/circulars/ec012.pdf.

Witten I.H., Frank E., Hall M.A., 2011. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann Publishers.