Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 71 (2023) 292–299

Transportation Research Procedia

www.elsevier.com/locate/procedia

XV Conference on Transport Engineering, CIT2023

International Roughness Index (IRI) prediction models for freeways

Heriberto Pérez-Acebo[a]*, Miren Isasa[b], Itziar Gurrutxaga[b], Harkaitz García[b], Aimar Insausti[b]

[a]*University of the Basque Country UPV/EHU, Pº Rafael Moreno "Pitxitxi", 2, 48013 Bilbao, Spain*
[b]*University of the Basque Country UPV/EHU, Plaza Europa, 1, 20018 Donostia-San Sebastián, Spain*

Abstract

In Pavement Management Systems (PMS), pavement performance models, or pavement deterioration (or evolution) models are regarded as a key element because they are able to forecast the future condition of the pavement based on available data. Hence, once pavement performance is predicted for next years, the optimal moment and treatment can be planned to be conducted, maximizing the existing limited budget for road maintenance and rehabilitation (M&R). There is a wide variety of characteristics that are assessed in a pavement, and additionally, there are various indices to measure those characteristics too. However, it can be said that there is a property, pavement roughness, measured by the International Roughness Index (IRI), which is the most widely employed index worldwide. Most of the road administrations around the world measure the roughness by means of IRI.
The Regional Government of Gipuzkoa (RGA) manages the entire road network in the province of Gipuzkoa, except from the municipal roads. Using the IRI data, traffic and pavement structure information of the A-636 freeway of Gipuzkoa, the aim of this paper is to develop some IRI prediction models for freeways in Gipuzkoa, adjusted to the climate characteristics of the province. Results showed that accurate models can be created if adequate variables are included, such as the pavement type, achieving a determination coefficient of $R^2 = 0.827$. This fact underlines the importance of recording as much information as possible, especially pavement structural section, in the PMS.

Keywords: International Roughness Index; prediction model; Pavement Management System; Gipuzkoa; network level, pavement deterioration models; IRI

* Corresponding author. Tel.: +34-94-601-7820.
E-mail address: heriberto.perez@ehu.eus

2352-1465
10.1016/j.trpro.2023.11.087

1. Introduction

As the total cost for pavement maintenance and rehabilitation (M&R) activities is generally higher than the available budget, highway administrations must to optimize assigned fund by means of Pavement Management Systems (PMS). They can be defined as a "*set of tools or methods that assist decision makers in finding the optimum strategies for providing, evaluating, and maintaining pavements in a serviceable condition over a period of time*" (AASHTO, 2012). PMS depends on some key elements: data collection of present pavement condition, prediction of future pavement condition using accurate performance model, and finally, network and project-level plans for M&R taking into consideration existing traffic and local circumstances, and available financial resources.

There is a wide variety of characteristics that can be evaluated in a pavement. They can be classified in the following groups: pavement distress measurements, surface characteristics (including longitudinal profile and roughness, and surface texture and friction), sub-surface characteristics, and structural evaluation (AASHTO, 2012). Moreover, there are no universal methodology to collect pavement condition data and each road agency follows its own procedure. Additionally, there are various indices to measure those characteristics too. Surface characteristics represent a small part of the pavement section but they are essential because they measure the contact point between vehicles and pavement, implying a great influence on travel safety and comfort. Pavement roughness, or longitudinal profile unevenness, is a widely measured characteristic by road agencies; it is related to ride comfort and safety and provides information about users' cost (travel time and operating cost). Many devices were available in the 80s, reporting results in various indices. With the aim of correlating them, the World Bank funded a correlation experiment in 1982. When processing the data, it was saw that almost all the roughness-measuring devices of the experiment could produce measures on the same scale if that scale was selected accordingly (Sayers et al. 1986a). As a result, the International Roughness Index (IRI) was developed, including a guideline and a computer code to calculate it (Sayers et al., 1986b). The IRI is defined in the algorithm proposed by Sayers (1995), and represents the accumulated suspension stroke of a vehicle; divide by the distance travelled during the same time period, expressed in mm/m or m/km. Since then, due to its transportability and stability over time, the IRI has become a real international index for measuring roughness.

Pavement performance or deterioration models are included in PMS as a key element due to their ability to predict future pavement condition. There are many types of models, which can be grouped according to various criteria. For example, Uddin (2006) classified the models into deterministic (based on regression analysis), probabilistic (including Markov and Bayesian models), and Artificial Neural Networks (ANN) models. Similarly, the Pavement Management Guide (AASHTO, 2012) grouped them into deterministic, probabilistic, Bayesian, and subjective models. At present ANN models are attracting researchers' attention and many models are being developed. However, deterministic and probabilistic models are generally recognized as the basic groups (Abaza, 2017, Pérez-Acebo et al. 2018, 2020a).

Deterministic models are developed when large historical pavement condition information is available and statistically significant deterioration correlations can be established. In these models, a correlation between the predicted index (dependent variable) and various influencing factors (independent variables) is calculated by means of regression analyses, generally, the least-square regression approach. On the other hand, probabilistic models provide the probabilistic distribution of the considered index (independent variable), and not an exact value. Normally, Markov chains are employed, with Transition Probability Matrices. Moreover, ANN models are employed more and more due to their capacity to solve problems that are difficult with traditional methods, after a training period. However, some researchers regard them as a "*black box*" because the exact relationship between the dependent variable and the independent variables is unknown (Sollazo et al. 2017, Pérez-Acebo et al. 2020a).

The aim of this paper is to develop an IRI deterioration model for the A-636 freeway, in the province of Gipuzkoa, in the Basque Country introducing the factors that are statistically significant on its evolution. The freeway has some segments with flexible pavement and others with semi-rigid pavement.

2. IRI performance models

There are three main types of pavements. A Portland Cement Concrete (PCC) at the top layer characterizes rigid pavements. In flexible pavements, a top surface asphalt layer (or layers), which supports directly the traffic loads, and beneath it, an untreated base layer is placed. The usual material for the untreated base is unbound aggregates. As traffic volumes are constantly increasing, higher stresses and more frequently act over the pavement, causing permanent deformation in flexible pavement. For improving the behavior of base layers, treated materials are employed, such as gravel and cement or soil-cement. Pavements with treated bases are called semi-rigid pavements, the third type, representing an intermediate type between the classic ones, rigid and flexible.

For IRI evolution prediction, many models have been developed for flexible pavements, with all the model types commented in the previous section. Some of those models are summarized in Abdelaziz et al. (2020), Perez-Acebo et al. (2020b, 2021). As it can be seen, the main factors introduced in the models to predict the future IRI are traffic volumes, initial IRI value, pavement age, structural data (existing materials and thicknesses of the layers), defects (cracking, rutting, potholes, etc.), subgrade characteristics, and climate factors. With regard to semi-rigid pavements, very few models are proposed, especially for network-level applications. The Mechanistic-Empirical Pavement Design Guide (MEPDG) in its 3rd version (AASHTO, 2020) has proposed a model for semi-rigid pavement, which include many variables, including many defect values and subgrade characteristics. Pérez-Acebo et al. 2021 proposed a model for semi-rigid pavement in single carriageways. The A-636 freeway that is used for IRI performance modeling in this study includes flexible and semi-rigid pavement segments, so presented models can be used for both types.

3. Used data and methodology

Gipuzkoa is one of the three provinces that composes the autonomous region of the Basque Country, in Spain, together with Biscay and Araba/Álava. The region has a special legislative status and each of the regional governments of these provinces has the competence about all the roads (freeways and highways) of the territory, even important freeways connecting with other regions or countries. Hence, the Regional Government of Gipuzkoa (RGA) (Diputación Foral de Gipuzkoa) is the owner and manages the entire interurban road network of the province, except for the municipal roads, with a total length of more than 1200 km (Pérez-Acebo 2018). The RGA has developed its own Pavement Management System, where the usual inputs are registered: inventory data, with all the road segments identified and their main characteristics; traffic history data; environmental data; pavement condition data (by means of IRI, in the case of pavement roughness); and for new projects, the pavement structure is available, obtained from as-built drawings.

For developing an IRI prediction model for freeways, the A-636 was selected. It is an approximately 20 km-long freeway connecting the cities of Bergara and Beasain, through the south of the province and it was constructed in 8 different segments during the last 20 years. In the entire length, two lanes were constructed in the carriageway for each direction. Initially, the first segments, one near Bergara and the other near Beasain were constructed, and the last segment, which eliminates the mounting pass of Deskarga, was finished in 2019. For each segment, constructed from one link to the next one, the complete pavement structure is registered, with the employed materials and their layer thicknesses. The exact month when the segment was opened to traffic is also recorded. Finally, the traffic volumes in each segment are also available at the documents that the RGA published yearly (GFA/DFG, 2022). For pavement condition data, IRI data from 2018 and 2021 were also obtained. Consequently, considering the available information, it was decided to develop a deterministic IRI performance model. Various variables could be employed to try to forecast future pavement condition, and it can be observed if they are statistically significant. Hence, the following variables were prepared.

- *IRI*. It is the dependent variable, the one that is wanted to be predicted. IRI was measured in the right lane, which is the one that is supposed to get damaged first because heavy vehicles mainly circulate on the right lane. IRI data were collected in the right lane in each of the wheel paths, left and right. The average value of each of the wheel path was calculated for segments with the same traffic volumes and pavement section. As segments were constructed from one link to the next one, using the same pavement structure, traffic volumes and pavement structures are the same for each of the 8 segments. As a result, there are 4 IRI values in each data collection year for each segment, left and right wheel path of the lane in one direction and another two average IRI values for the lane in the other direction. It is expressed in m/km.

- Real Age (*R.Age*): it is the real age of the pavement. It is the difference between the year of the data collection (2021 or 2018) and the year when the segment was opened to traffic. It is expressed in years, in decimal form, i.e., if the data collection was conducted in July of 2018, it is used 2018.5. Similarly for the year when the segment was opened to traffic.
- Annual Average Daily Traffic (*AADT*). It is the Annual Average Daily Traffic of the segment in the year of the IRI data collection (2018 or 2021). It represents the AADT in both direction and is expressed in vehicles/day
- Annual Average Daily Traffic of Heavy Vehicles ($AADT_{hv}$). It is the AADT of heavy vehicles. In Spain, a vehicle weighting more than 3500 kg is considered as a heavy vehicle (MFOM, 2003). It counts both directions, expressed in heavy vehicles/day.
- Total Vehicles (*TotVeh*). It is the total number of vehicles that passed through a segment since the segment was opened. It takes into account the AADT of each year and when the data collection was conducted and when the segment was opened to traffic. It is expressed in thousands of vehicles, considering both directions.
- Total Heavy Vehicles (*TotHVeh*). Similarly, is the total number of heavy vehicles that circulated on the segment since it was opened to traffic until the IRI data collection. Once again, it is expressed in thousands of heavy vehicles, considering both directions.
- Total thickness of bituminous layers (*BitTot*). It is the total thickness of the bituminous layers in the pavement section, expressed in cm. It considers all the bituminous layers.
- Pavement type (*PavType*). It is a categorical variable, not numerical, which considers the pavement type: flexible or semi-rigid. There are not rigid pavements in Gipuzkoa.

With these variables or transformed variables, a Multiple Linear Regression (MLR) is developed, using all the numerical variables (all except for *PavType*), and with all the variables, a General Linear Multiple (GLM) regression model is created. The MLR model analysis assumes some hypothesis that must be verified after the model is created: a linear relationship between the dependent and the independent variables, the independence of observations, the homoscedasticity, errors must be normally distributed; the variance of errors must be equal across all levels; and their little or no multi-colinearity in the data. The GLM regression model is a general form of linear regression and it can include numerical and categorical variables.

4. Results and discussion

Initially, the correlation between each independent numerical variable and IRI was calculated by means of the Pearson coefficient and its significance (Table 1).

Table 1. Correlations between the dependent variable (IRI) and the independent variables (Pearson coefficient) and significance of correlation

Independent variables	Correlation with IRI (Pearson coefficient, R)	Significance of the correlation (bilateral)
R.Age	0.602	< 0.001
AADT	0.302	0.021
AADThv	0.197	0.139
TotVeh	0.579	< 0.001
TotHVeh	0.602	< 0.001
BitTot	0.441	< 0.001

The MLR models were created and tested with the IBM SPSS v28 software, using the Step by Step and Forward functions. Models were accepted if they were globally significant (a p-value of the Fisher-Snedecor test below 0.05) and all the introduced variables were significant (the coefficient of the variables are different from zero with 95% significance). Several models were tested. As seen in Table 1, *AADT* and $AADT_{hv}$ had lower correlations and generally were discarded from the models. In some, not all the variables were significant. In other, although all the variables were significant, there was an important problem of collinearity. In fact, the age of the pavement and the total number of vehicles (or heavy vehicles) that circulated on the segment are highly correlated. The model with the higher determination coefficient (R^2) and verifying all the hypothesis of a MLR model (commented in section 3) is shown in Equation 1:

$$IRI = 0.888 + 6.824 \cdot 10^{-9} \cdot TotHVeh^2 + 0.029 \cdot TotBit \tag{1}$$

Where *IRI*, *TotHVeh* and *TotBit* are as defined in Section 3. The equation has a determination coefficient of R^2 = 0.470, and an adjusted of R^2_{adj} = 0.450. The complete statistical analysis of the model is shown in Table 2, Table 3 and Figure 1.

Table 2. Analysis of the Variance of the model of Equation (1)

Source	Sum of Squares	d.o.f.	Mean Squares	F Value	p-value	Durbin-Watson	RMSE	R
Model	10.528	2	5.264	24.348	< 0.001	0.822	0.355	0.685
Error	11.891	55	0.216				R2	Adj R2
Corrected total	22.418	57					0.470	0.450
Variable	Parameter estimates	Standard Error	t-value	p-value	95% confidence level		Tolerance	VIF
Intercept	0.888	0.256	3.472	0.001	0.375	1.401		
$TotHVeh^2$	6.82*10-9	< 0.001	5.338	< 0.001	0.000	0.000	0.887	1.128
TotBit	0.029	0.012	2.436	0.018	0.005	0.053	0.887	1.128

Table 3. Colinearity diagnosis of Equation (1).

Dimension	Eigenvalue	Correlation Index	Proportions of the Variance		
			Intercept	$TotHVeh^2$	TotBit
1	2.596	1.000	0.01	0.05	0.01
2	0.377	2.625	0.03	0.89	0.01
3	0.028	9.695	0.97	0.06	0.98

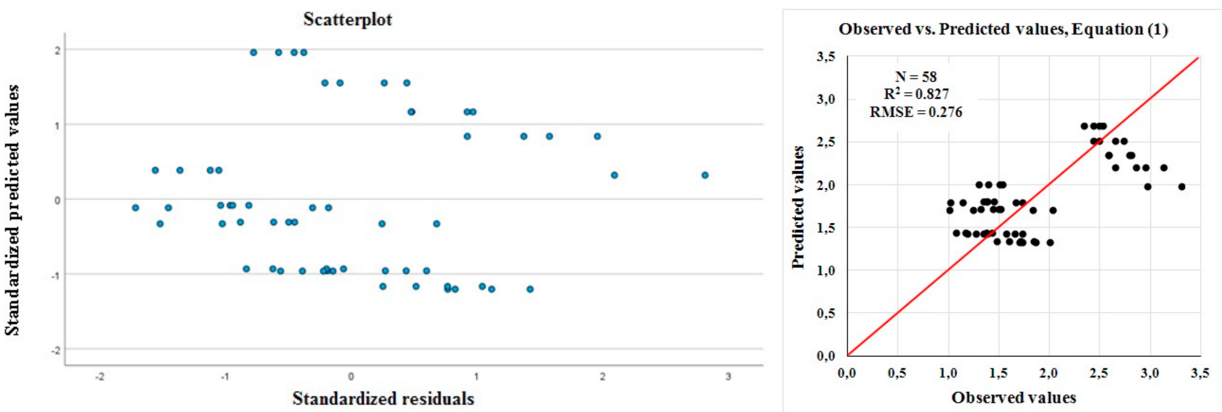

Fig. 1. Analysis of Equation (1), (a) scatter plot of the standardized residuals vs. standardized predicted values; (b) observed values vs. predicted values.

In addition, some GLM models were tested. Once again, it is aimed to obtain the model with the highest R^2 with all the variables statistically significant and verifying all the hypotheses. Equation 2 was selected because it had a high R^2 and all the variables were statistically significant, even the coefficients for the categories of the categorical variables.

$$IRI = 3.68 + A \cdot TotBit + 2.2 \cdot 10^{-10} \cdot TotVeh^2 + B \cdot R.Age^2 - 6.8 \cdot 10^{-9} \cdot TotHVeh^2 \quad (2)$$

Where *TotBit*, *TotVeh*, and *TotHVeh* are as defined in section 3, and A and B are two coefficients depending on the pavement type, flexible and semi-rigid, according to Table 4. Statistical analysis of the model proposed in Equation 2 is shown in Tables 4 and 5, and in Figure 2. The determination coefficient of the model is $R^2 = 0.827$. As seen, the inclusion of the pavement type is an important factor, which is able of improving considerably the model accuracy. Hence, it must be concluded that flexible and semi-rigid pavements have a complete different behavior. It was probed in Perez-Acebo et al. (2020b, 2021) that models for flexible pavements cannot be applied in semi-rigid pavements. Moreover, in a Pavement Management System it is important to include as much information as possible.

Table 4. Values of coefficients A and B depending on the pavement type.

Pavement type	A	B
Flexible	0.032	-0.011
Semi-rigid	-0.117	-0.007

Table 5. Test of Between-Subjects effects of the model of Equation (2).

Origin	Type III sum of squares	d.o.f.	Mean Square	F	Significance	Observed Power
Corrected model	18.534	6	3.089	40.555	<0.001	1.000
Intercept	5.640	1	5.640	74.047	<0.001	1.000
PavType*TotBit	4.659	2	2.330	30.585	<0.001	1.000
$TotVeh^2$	2.645	1	2.645	34.728	<0.001	1.000
PavTyp*$R.Age^2$	2.839	2	1.420	18.640	<0.001	1.000
$TotHVeh^2$	0.397	1	0.397	5.211	0.027	0.610
Error	3.885	51	0.076			
Total	221.712	58				
Corrected total	22.418	57				

Table 6. Parameter estimates for the model of Equation (2).

Parameter	B	Standard Error	T	Significance	95% confidence level		Observed Power
Intercept	3.678	0.427	8.605	<0.001	2.820	4.536	1.00
[PavType=FLEX]*TotBit	0.032	0.017	1.926	0.06	-0.001	0.066	0.47
[PavType=SEMIR]*TotBit	-0.117	0.024	-4.779	<0.001	-0.166	-0.068	0.99
$TotVeh^2$	2.2*10-10	3.7*10-11	5.893	<0.001	1.5*10-10	3*10-10	1.00
[PavType=FLEX]*$R.Age^2$	-0.011	0.002	-6.103	<0.001	-0.014	-0.007	1.00
[PavType=SEMIR]*$R.Age^2$	-0.007	0.001	-5.123	<0.001	-0.009	-0.004	0.99
$TotHVeh^2$	-6.8*10-9	3*10-9	-2.283	0.027	-1*10-8	-8*10-10	0.61

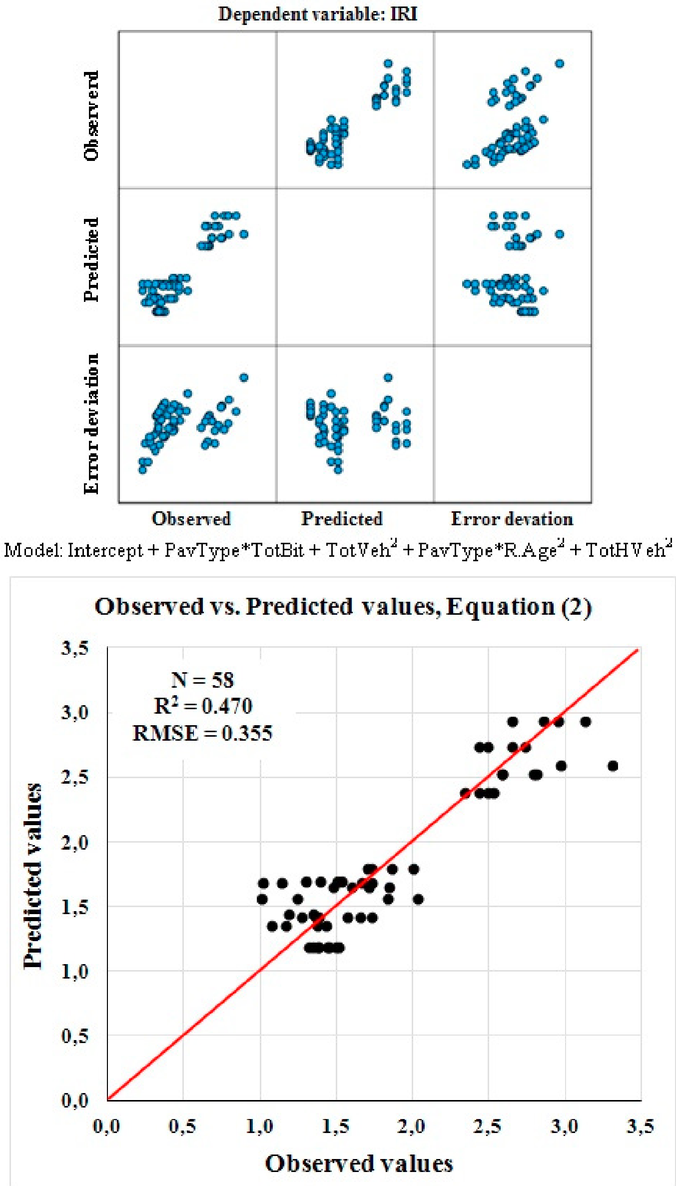

Fig. 2. Statistical analysis of Equation (2), (a) plot of residuals (standardized, observed values and predicted values, (b) observed values vs. predicted values.

With the inclusion of the pavement type as a variable, the model is almost twice more accurate (from R^2 = 0.470 to R^2 = 0.827). Consequently, road administrations must keep the as-built drawings of new projects because they contain important information to be recorded in the PMS.

5. Conclusions

In this study an IRI performance model was developed for freeways in the province of Gipuzkoa (Spain). The A-636 freeway was employed and a deterministic model was selected because there were many data available. Various variables were prepared as potential factors to be included in the model: the real age of the pavement; the Annual Average Daily Traffic (AADT) and the AADT of heavy vehicles of the year of the IRI data collection; the total number of vehicles and heavy vehicles that has circulated through the segment since it was opened to traffic; and the total thickness of bituminous materials. The best model obtained an R^2 = 0.470 using only two variables, the thickness of bituminous materials and the total number of heavy vehicles passed through the segment. If more variables were introduced, there were problems of colinearity. Additionally, a categorical variable was defined: the pavement type, with the aim of considering separately the flexible and semi-rigid pavement. Including this variable a GLM model was developed and the model accuracy was considerably improved, until R^2 = 0.827.

Consequently, it was shown that flexible and semi-rigid pavements have completely different behavior and they must be considered separately, with different models or include a specific variable for distinguishing them. Additionally, this research also underlines the importance of recording as much information as possible in the Pavement Management Systems. Road administrations should registered all the available data for developing better performance models and forecast future pavement condition more accurately, which will result on better decision about maintenance and rehabilitation activities.

Funding

This research was funded by the Gipuzkoako Foru Aldundia / Diputación Foral de Gipuzkoa by means of the project "*Construcción y movilidad inteligentes y sostenibles en Gipuzkoa / Gipuzkoan eraikuntza eta mugikortasuna adimentsu eta jasangarriak*" of the research program·"Etorkizuna Eraikiz" under grant P10.

References

AASHTO, 2012. Pavement Management Guide, 2nd ed. American Association of State Highway and Transportation Officials; Washington, DC, USA.

AASHTO, 2020. Mechanistic-Empirical Pavement Design Guide: A Manual of Practice, 3rd Edition. American Association of State Highway and Transportation Officials, Washington, DC, USA.

Abaza, K.A., 2016. Simplified staged-homogeneous Markov model for flexible pavement performance prediction. Road Materials and Pavement Design, 17.2, 365-381.

Abdelaziz, N., Abd El-Hakim, R.T., El-Badawy, S.M., Afify, H.A., 2020. International Roughness Index prediction model for flexible pavements. International Journal of Pavement Engineering, 21, 88-99.

GFA/DFG (Gipuzkoako Foru Aldundia / Diputación Foral de Gipuzkoa), 2022. Información aforos en las carreteras de Gipuzkoa. Recopilación hasta 2021. Departamento de Infraestructuras Viarias, Donostia-San Sebastián, Spain.

MFOM (Ministerio de Fomento) 2003. FOM/3640/2003, de 28 de noviembre, por la que se aprueba la Norma 6.1-IC "Secciones de firme", de la Instrucción de Carreteras. Ministerio de Fomento, Madrid, Spain.

Pérez-Acebo, H., 2018. Carreteras. Volumen I: Red viaria y tráfico. Servicio Editorial de la Universidad del País Vasco UPV/EHU, Bilbao, Spain.

Pérez-Acebo, H., Bejan, S., Gonzalo-Orden, H., 2018. Transition probability matrices for flexible pavement deterioration models with half-year cycle time. International Journal of Civil Engineering, 16, 1045-1056.

Pérez-Acebo, H., Gonzalo-Orden, H., Findley, D.J., Rojí, E., 2020a. A skid resistance prediction model for an entire road network. Construction and Building Materials, 262, 120041.

Pérez-Acebo, H., Linares-Unamunzaga, A., Rojí, E., Gonzalo-Orden, H., 2020b. IRI performance models for flexible pavements in two-lane roads until first maintenance and/or rehabilitation work. Coatings, 10.2, 97.

Pérez-Acebo, H., Gonzalo-Orden, H., Findley, D.J., Rojí E. 2021. Modeling the international roughness index performance on semi-rigid pavements in single carriageway roads. Construction and Building Materials, 272, 121665.

Sayers, M.W., Gillespie, T.D., Queiroz, C., 1986a. International experiment to establish correlations and standard calibration methods for road roughness measurements. World Bank Technical Paper 45. The World Bank, Washington, DC, USA.

Sayers, M.W., Gillespie, T.D., Paterson, W.D., 1986b. Guidelnes for the conduct and calibration of road roughness measurements. World Bank Technical Paper 46, The World Bank, Washington, DC, USA.

Sayers, M.W., 1995. On the calculation of IRI from longitudinal road profile. Transportation Research Record, 1501, 1-12.

Sollazo, G., Fwa, T.F., Bosurgi, G., 2017. ANN models to correlate roughness and structural parameters in asphalt pavements. Construction and Building Materials, 134, 684-693.

Uddin, W., 2006. Pavement Management Systems, in "The Handbook of Highway Engineering". In: Fwa, T.F. (Ed). Taylor & Francis, Boca Raton, FL, USA.