PREDICTION OF EACH ROAD DETERIORATION CONSIDERING TRAFFIC AND THE INTERACTION WITH OTHER SURFACE DETERIORATIONS, USING AUTOMATED LEARNING MACHINE TECHNICS

R. MUZZULINI[*], M. PAGOLA[*] AND O. GIOVANON[*]

[*] School of Engineering, IMAE Road Laboratory
National University of Rosario
Pellegrini 250, Rosario, Argentina
email: remuzzu@gmail.com, pagola@fceia.unr.edu.ar, ogiovanon@gmail.com

Key words: Road Deterioration Model Prediction, Learning Machine Technics

Abstract: For road user, the pavement is a surface that must allow the circulation of mixed traffic, in conditions of safety and comfort, under any climatic condition, for a long time. Once the pavement is in service, it begins to have deteriorations that can cause that the user modifies their behavior and start driving at a slower speed in order to maintain traffic safety conditions; this circumstance causes travel times increasing and therefore circulation costs increase.

There are different deteriorations to consider, roughness, road surface adherence, rutting, cracking and potholes. Periodic deterioration evaluation and prediction modelling allows that corrective actions can be anticipated, so that road quality does not fall below acceptability limits. To prioritize improvements and routine maintenance, it is necessary to develop adequate tools to predict the deterioration evolution, which can be incorporated into the pavement management systems used to prepare multi-year works and maintenance plans.

Periodic observations of surface deteriorations of sections in service located on routes of Littoral region of Argentina were used in the paper. It was possible to develop predictive models using Support Vector Machine Regression SVR and Random Forest Regression RFR; these are learning machine tools, which can be used to solve estimation problems of multidimensional functions. First a model to predict cracking was developed. When it was optimized, the model to predict rutting was realized. And at the end, the model to predict roughness was adjusted, using cracking and rutting models developed previously. Results indicate that SVR and RFR regression models have the capacity to perform training and prediction that help to develop road surface deterioration models.

1 INTRODUCTION

In this paper, "automated learning" techniques are presented to predict pavement surface deteriorations evolution; where analysed models for roughness, rutting and cracking.

Machine Learning is the subfield of computer science and a branch of artificial intelligence that aims to develop techniques that allow computers to learn.

The current analysis of the problem was focused on supervised learning, whose objective is to create a function capable of predicting the value corresponding to any valid input object, after having seen a series of examples. This makes predictions of unknown variables based on

behaviors or characteristics that have been seen in the data already stored.

What machine learning algorithm is the appropriate to use? The algorithm to be used depends on the size, quality and nature of the data. The fundamental idea of automated learning is to find patterns that can be generalized, in order to apply this generalization about cases that have not yet been observed, and make predictions. The objective of the regression is to minimize the error between the approximate function and the value of parameter. Many variables can lock some learning algorithms and cause the training time to be too long. To develop good regression models, is very important the selection of independent variables as an input of models.

Python was used as language for the development of test techniques, and Scikit-learn was used as software packages. These libraries also facilitate the evaluation, diagnosis and validation tasks since they provide several factory methods included to perform these tasks in a very simple way. [1]

It is expected that this tool can be implemented in road management systems, and as a transfer function in structural design programs.

2 REGRESSION MODELS USED

Two regression models were analyzed and compared: Support Vector Machine Regression and Random Forest Regression. [2,3,4,5]

2.1 Support Vector Machine Regression (SVR)

SVR objective is to make the prediction from a problem of geometric optimization, which can be written as a problem of quadratic convex optimization with linear constraints, in principle solvable by any non-linear optimization procedure.

The technique of support vectors is a universal tool for solving estimation problems of multidimensional functions. The aim is to select the hyperplane regressor that best fits the training data set, based on considering a distance margin ε, so that all the examples are in a band or tube around said hyperplane. The function is intended to be as close as possible to the points, that is, the formation of the band or tube around the true regression function.

To define the hyperplane, only data points that are more far than ε of the hyperplane are considered. These data points are the considered support vectors; they are identified with the strict possibility of the associated artificial variables, which quantify the error between the approximation and the real value of each data point of the training set. In Figure 1 the reduction in support vectors is shown when ε increase.

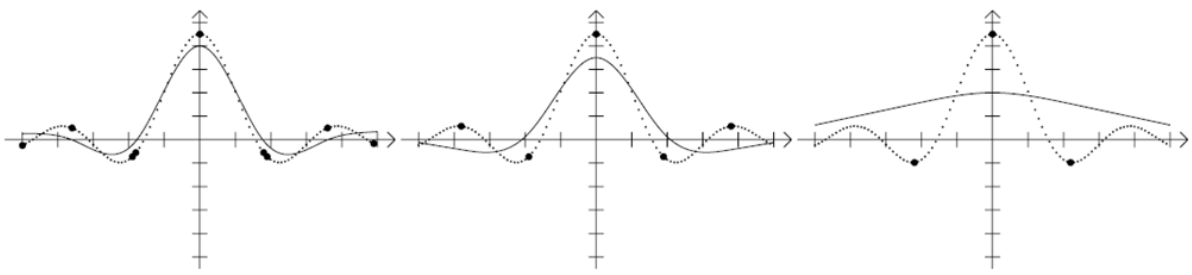

Figure 1. Left to right: regression (solid line), data points (small dots) and SVs (big dots) for an approximation with ε = 0.1, 0.2 and 0.5. Note the decrease in the number of SVs. [6]

2.2 Random Forest Regression (RFR)

They are a combination of predictive trees in such a way that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The generalization error for forests converges to a limit as the number of trees in the forest becomes large.

RFR are a set of different regression trees and are used for multiple non-linear regressions, where each sheet contains a distribution for the continuous output variable(s). The objective of these methods is to inject the algorithm with just randomness to maximize the independence of the trees while maintaining a reasonable precision. The results turn out to be insensitive to the number of features selected to divide each node. In general, when selecting one or two characteristics, optimal results are obtained.

Random forests are effective in eliminating noise in the input data of the model. Given a long list of input variables and a reduce data set, it is very likely that any predictive model will discover false relationships between those inputs and the target variable chosen.

Because RFR constructs many trees using a subset of input variables and their values, it contains underlying decision trees that omit the noise-generating variable. In the end, when it is time to generate a prediction, there is a vote among all the underlying trees and the majority prediction value wins.

3 PAVEMENT SURFACE DETERIORATIONS

Since a road is built deterioration process begins, and evolves causing reduction in road quality, and increase maintenance and user costs.

Once put in service and influenced by the climate and traffic, the pavement will have a loss of quality and deterioration will appear. To reduce these deteriorations, certain periodic maintenance and rehabilitation activities must be carried out on time, in order to reduce the impact that different failures may have on the structure, in order to optimize the available resources and to avoid reconstruction.

For this, pavement management techniques and pavement road evaluation methodologies are required, which consist in the evaluation of superficial and structural deteriorations level that affect the quality of the service provided to the users, using defined analysis methodologies and appropriate equipment.

There are different deterioration parameters to evaluate road surface quality provided to users: roughness, rutting, cracks, potholes, raveling, bleeding, etc. [7,8]

3.1 Roughness

Roughness evaluates the deviations of road longitudinal profile with respect to a flat surface, that affect in a very important way vehicle dynamics, quality of circulation, dynamic effect of the loads and the drainage. Users perceive roughness as vibratory movements that affect the comfort of circulation.

Roughness value is expressed in IRI (International Roughness Index, defined in 1982 by the World Bank), which is an index obtained by a mathematical simulation of the passage of a virtual vehicle, circulating on road profile at a speed of 80 km/h.

IRI is a continuous type variable, where the range goes from zero for ideally flat surfaces (runways of airports can provide values lower than unity), with values between 1 and 2 m/km for new pavements, and values higher than 6 m/km for deteriorated surfaces.

Data used in the present work are measurements made with dynamic response equipment, according to class 3 World Bank.

3.2 Rutting

Rutting is the loss of the road transversal profile, with depressions located in correspondence with the wheel path; this is the area where most of heavy vehicles circulate.

In Argentina, rutting is measured as the depression under a 1.20 m straight edge. Data used in this work have been obtained from manual measurement made by trained people, at 1 km intervals.

3.3 Cracking

The presence of cracking in asphalt surface layer is an indicator that the material has consumed its fatigue capacity. This indicator is only visualized by the user when the surface presents a high degree of deterioration, when there is loss of material. When cracks are in early stages of development, it is an indicator that layer material is failing, but the user does not perceive them, because it does not affect circulation comfort.

In Argentina, this parameter is measured assigning deterioration degrees according to evolution pattern, because of fatigue, increasing from 0 to 10. Grade 0, no cracks. Grade 2, are fine and isolated cracks located in correspondence with wheel path. Grade 4, are branched cracks with tendency to form meshes. Grade 6, are interconnected cracks forming blocks. Grade 8, are high severity level cracks, alligator cracking. Grade 10 are generalized cracks, pieces may move and can be lost.

Data used in the present work have been obtained from visual survey performed by trained personnel walking along the road, analyzing 20 meters at 1 km intervals. After that, measurements have been expressed as total cracking, using correlations between methodologies.

4 ROAD NETWORK DETERIORATION MODELING

4.1 Independent variables selection

Independent variables, predictors in the future, are independent in the sense that they are external and measurable variables.

The choice of these variables is as important as the choice of the target variable, since it determines modeling success. Most of the time invested in the development of models is used precisely in the analysis and selection of independent variables set.

In the present analysis, data used are: year of measurement, deflection, annual ESAL, rutting, percent of total cracking, and roughness. The objective parameter is roughness, and intermediate models were developed for rutting and cracking.

4.1.1 Traffic

It is an important variable in road design, because truck loads are responsible of deterioration. It is necessary to know the number and type of vehicles that will pass, as well as load intensity and axle configurations. In the present study, traffic intensity and loads are expressed in annual equivalent single axle of 80 KN (ESAL), for each year of service.

4.1.2 Rutting

Rutting is the changes in road transversal profile, with respect to the original profile. These variations of the transversal profile are found in the tracks and appear because of permanent deformation caused by heavy traffic. This deformation has a fundamental impact on the functional conditions (for the user) and structural conditions (for the engineer) of the existing pavement. This causes water accumulation in wheel tracks, loss of control of the vehicle, driving insecurity, lack of comfort, and increasing accidents risk.

4.1.3 Cracking

Cracking is the discontinuity in surface layer material. This deterioration reduces layer resistance, and after them more deteriorations appears, like raveling, potholes and roughness. They also provide a potential water access to the lower layers, which accelerates the deterioration evolution.

These variables are known to have significant impact on roughness. The prediction of rutting and cracking as intermediate models, improve roughness prediction.

4.2 Data used

Data were obtained from different sections located on in service roads. Measurements were made by Road Administration staff and shared with the University.

Data are from homogeneous sections, with same structure and traffic. Sections are located in different geographic location from Argentina Littoral Region. From each section the following information were known: traffic, structure, layer materials, maintenance applied, and periodic measurements results: rutting, cracking and roughness.

For the sections analyzed in this study 325 data points were collected from 59 homogeneous sections. Benkelman beam deflection was not measured every year, and was assumed constant during years without data. Roughness was measures annually. Table 1 shows and example of section data for sections 51 and 55.

4.3 Experimental models analysis

In this point progressive analysis tested with SVR and RFR are described, using data for model training and validation as described previously.

First a model to predict cracking was developed. When it was optimized, the model to predict rutting was done. And at the end, the model to predict roughness was adjusted, using cracking and rutting models developed previously. Standard deviation of prediction error was used to compare both techniques.

To minimize errors, data were filtered using following criteria:

- When roughness increased more than 0.6 m/Km IRI from one year to the other, it was assumed that surface maintenance was conducted and it was not documented. Then this section

data was divided into two sections. One section up to roughness increment, and another section after that.

- In roughness data, when difference between results of two consecutive years was more than 0.3 m/km and less than 0.6 m/Km, a filter was applied. The strange data was removed and a new data was obtained using a grade 3 polynomial regression between data from previous an after year. This situation was noted in roads with high deterioration level.
- For years without deterioration data, like year 2012 of section 51, data were created as extrapolation of previous data using a grade 3 polynomial regression between data from previous an after year. In this period, traffic was increased using 2% increment rate.
- Deterioration data were forced to have an increment between years. Then, if cracking, rutting or roughness from one year to the other were decreasing, they were force to increase.
- To train with Support Vector Machine Regression, the input vector was normalized over all data (less roughness), that is, the characteristics were readjusted so that they have the properties of a standard normal distribution with $\mu = 0$ and $\sigma = 1$. In Python, we used the preprocessing library that provides a quick and easy way to perform this operation in a single data set.

Table 1: Example of section data available

Sec.	Year	Benk Beam Deflection (0.001 mm)	ESAL Annual (10^6)	All Crack (%)	Rutting (mm)	IRI (m/Km)
…….	…….	…….	…….	…….	…….	…….
51	2007	660	1.50	0	0	1.54
51	2008	660	1.55	0	0	1.54
51	2009	660	1.67	0	3	1.96
51	2010	660	1.46	1	5	1.96
51	2011	660	1.66	0	6	2.56
51	2013	660	1.70	3	8	2.16
51	2015	660	1.77	1	10	2.56
51	2016	660	1.74	0	10	2.95
…….	…….	…….	…….	…….	…….	…….
55	2010	590	1.50	31	4	2.16
55	2011	590	1.49	31	4	2.95
55	2012	590	1.55	33	3	2.95
55	2014	590	1.75	37	5	2.95
55	2017	590	1.85	33	7	2.95

4.4 Training models with n-1 year's data

Training was done for three parameters: first cracking, after that rutting containing cracking, and at last roughness containing rutting and cracking; in the same way that is recommended in Highway Development and Management software (HDM). Results indicate that SVR and RFR regression models have the capacity of perform training and prediction. [9]

4.4.1 Cracking prediction

First, the model was trained with n-1 year's data to adjust both regression models. Cracking percentage of previous year was taken as input data, and cracking percentage of the present year was the prediction objective. Year, deflection, ESAL annual were data entry too.

In Figure 2 actual cracking percentage versus SVR and RFR prediction is shown. Prediction errors, expressed in percentage, are 4.41 for SRV model and 4.19 for RFR.

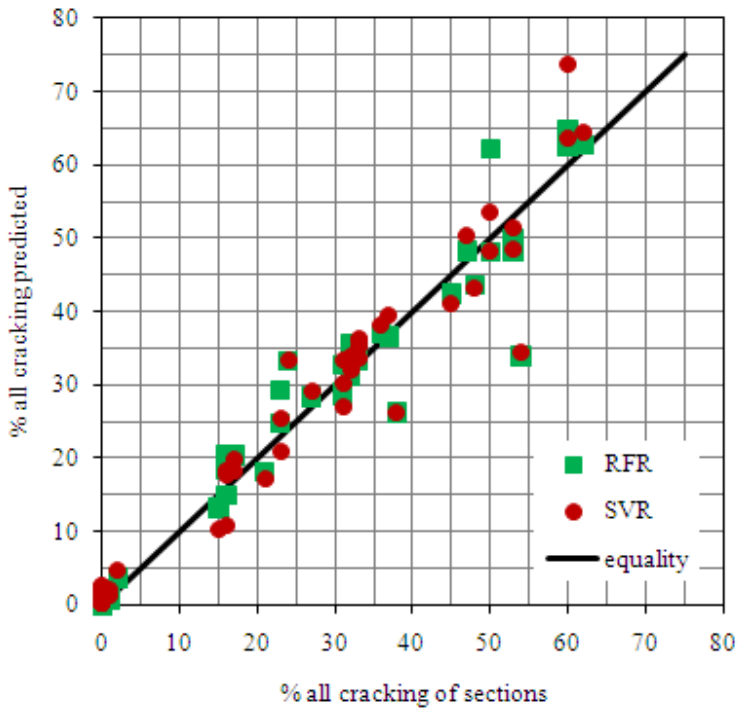

Figure 2: Actual cracking percentage versus SVR and RFR prediction

4.4.2 Rutting prediction

As second step, rutting prediction model was adjusted. It was done by training the models with n-1 year's data. Cracking of previous year was obtained using cracking prediction model data. Rutting of previous year was taken as data, and rutting of present year was the objective.

In Figure 3 measured rutting versus SVR and RFR prediction is shown. Prediction errors, expressed in mm, are 1.13 for SRV model and 1.60 for RFR.

4.4.3 Roughness prediction

In a similar way than for cracking and rutting, roughness prediction was done. Roughness measurement of previous year, cracking and rutting predicted were used as data, and roughness of the present year was the objective.

In Figure 4 roughness measured versus SVR and RFR prediction is shown. Prediction errors, expressed IRI (m/Km), are 0.20 for SRV model and 0.10 for RFR.

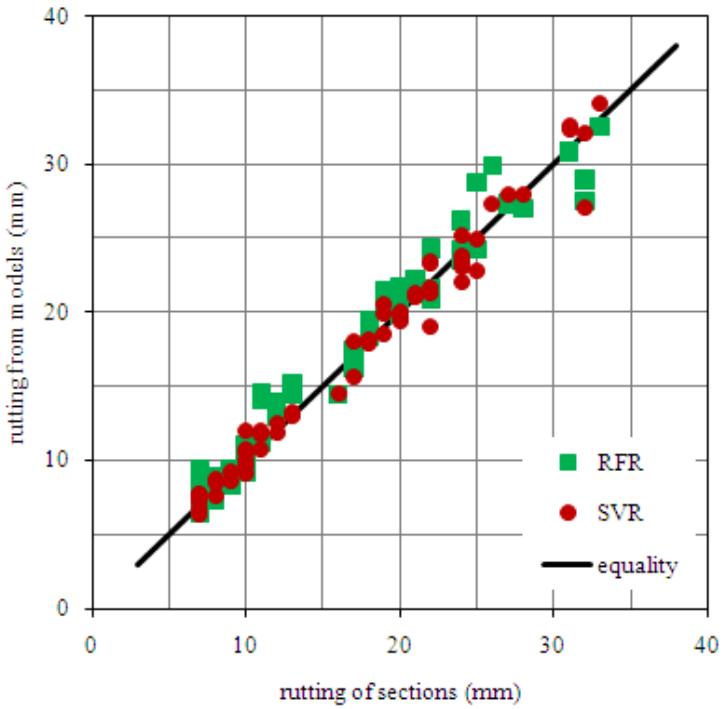

Figure 3: Rutting measured versus SVR and RFR prediction

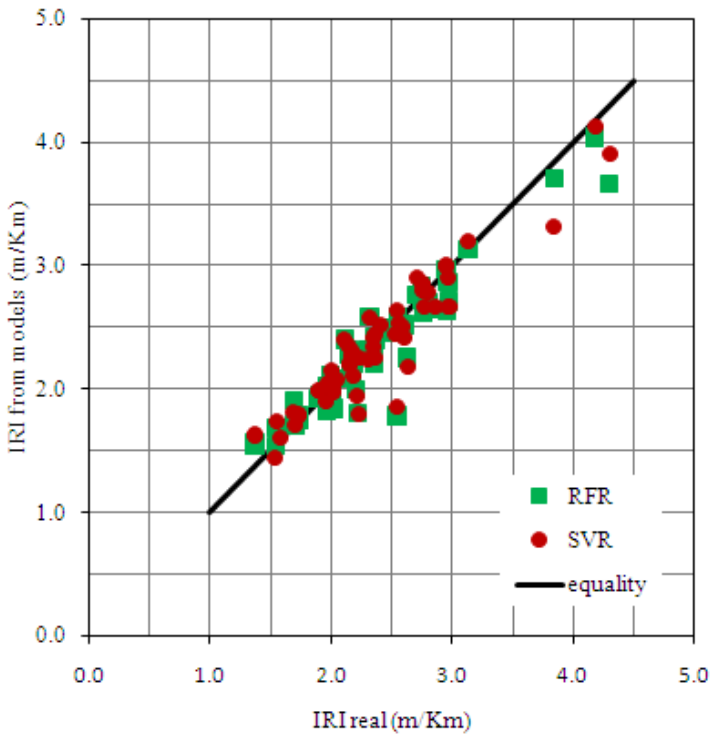

Figure 4: Roughness measured versus SVR and RFR prediction

4.4.4 Prediction analysis using developed models

Up to now, data from n-1 year were used for models training, and the ability of models to predict each deterioration for the last year was analyzed. The errors obtained using SVR and RFR were similar. Roughness prediction errors expressed in IRI (m/Km), were 0.20 for SRV model and 0.10 for RFR.

Now is time to use data and models to validate prediction for many years. Beginning with the first data, models were used to predict year by year roughness evolution, using the three models of SVR and RFR.

For this analysis, data were taken as follows: year was incremented in one year, deflection was considered as a constant, traffic increment rate was 2%, and cracking and rutting were obtained from adjusted models.

Figure 5 shows roughness evolution for one section, as an example.

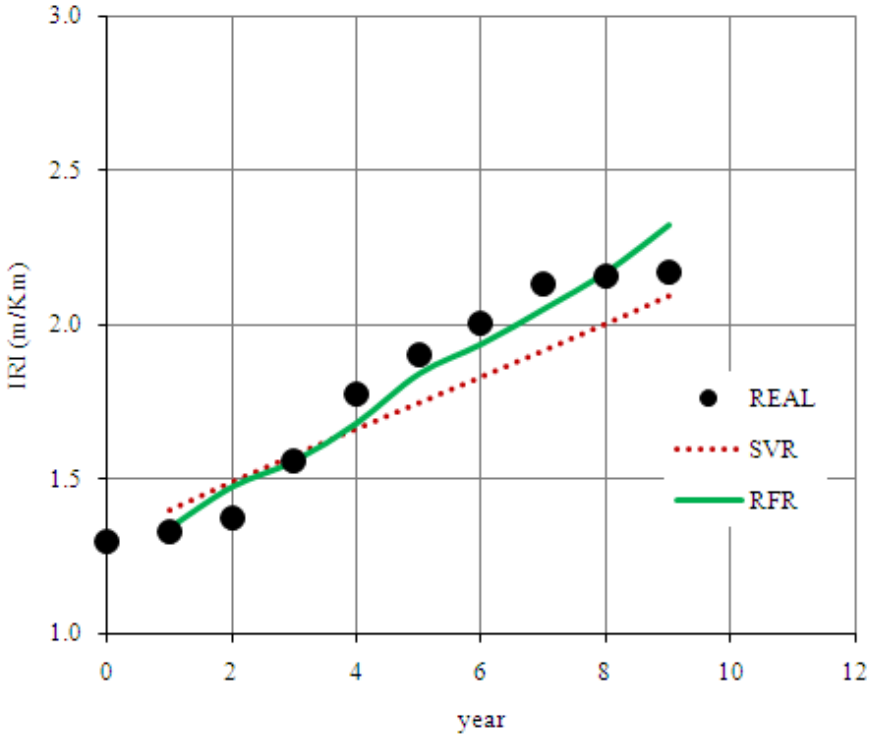

Figure 5: Roughness prediction using SVR and RFR regression models

Results show that both methodologies are capable to be trained and predict deteriorations.

The present study was our first intention to use SVR and RFR regression models to training and prediction deteriorations. Results look good and we are going to continue our researches using another group of sections data, for other regions of Argentina.

5 CONCLUSIONS

For road routine maintenance planning, it is necessary to have adequate tools/models to predict the evolution of deterioration. These models can be incorporated into pavement management systems to prepare the multi-year works and maintenance plans.

Other use of models is to use as transfer functions in mechanistic pavement design software as, to predict deterioration like cracking, rutting and roughness, as MEPDG guide. [10]

The present study was our first intention to use SVR and RFR regression models to training and prediction deteriorations. Results look good and we are going to continue our researches using another group of sections data, for other regions of Argentina.

REFERENCES

[1] Lopez Briega, R. 2015. *Machine Learning con Python.* http://relopezbriega.github.io/blog/2015/10/10/machine-learning-con-python/.

[2] Cutler, A. 2013. *Trees and Random Forests*. Mathematics and Stadistics. Utah State University.

[3] Drucker, C., Burges, J., Kaufman, L., Smola, A. and Vapnik, V. 1996. *Support Vector Regression Machines*. Bell Labs and Monmouth University Department of Electronic Engineering West Long Branch.

[4] Goddard J., Cobos Silva S., Perez Salvador B. and Gutierrez Andrade M. 2000. *Un algoritmo para el entrenamiento de máquinas de vector soporte para regresión*. Revista de Matemática: Teoría y Aplicaciones. issn: 1409-2433.

[5] Guareño, J. 2016. *Support vector regression: Propiedades y aplicaciones. Master'sthesis*. Facultad de Matemáticas, Departamento de Estadística e Investigación Operativa de la Universidad de Sevilla.

[6] Alex J. Smola and Bernhard Schölkopf. 2003. *A Tutorial on Support Vector Regression.* RSISE, Australian National University, Canberra 0200, Australia.

[7] Pagola, M. and Giovanon, O. 2016. *Curso Evaluación de Calzada*. FCEIA UNR. Argentina.

[8] Mohamed Jaafar, S., Ahlan, M and Uddin, W. 2015. *Modeling of pavement roughness performance using LTTP database for southern region in the U.S.* Bituminous Mixtures & pavements VI

[9] PIARC. 2005. HDM-4. *Highway Development and Management.*

[10] AASHTO. 2008. *Mechanistic-Empirical Pavement Design Guide.*