Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 14 (2016) 3751 – 3760

Transportation Research Procedia

www.elsevier.com/locate/procedia

6th Transport Research Arena April 18-21, 2016

Preliminary study on runway pavement friction decay using data mining

Mario De Luca [a], Francesco Abbondati [a,*], Monica Pirozzi [a], Daiva Žilionienė [b]

[a]*University of Naples "Federico II", Via Claudio 21, Naples 80125, Italy*
[b]*Vilnius Gediminas Technical University, Saulėtekio al. 11, Vilnius 10223, Lithuania*

Abstract

Surfaces of airport pavements are subject to the friction decay phenomenon. A recurrent problem for the runways is represented by the deposits of vulcanized rubber of aircraft tires. This happens mainly in the touch-down areas during landing operations, and the loss of grip compromises the safety of both take-off and landing operations. This study moves from the International Civil Aviation Organization and the Italian Civil Aviation Authority provisions concerning runway friction measurement and reporting to a better way to analyze friction data. Being data mining the computational process of discovering patterns in a large data sets, data mining techniques are very helpful to reach this target. Unsupervised and supervised classification methods to analyze friction data detected by Grip Tester Trailer were employed. First, *K-means* and *Subtractive Clustering* were applied to divide data into a certain number of clusters representing the different areas of consumption. In a second time two different *Classification and Regression Trees* models, *CART* and *GCHAID*, were employed to split the data points of the runway into nodes. At the end of the process scatterplots were built and better visualized through non-linear regressions. The decay curves obtained were of service to compare the results achieved using data mining techniques versus the International Civil Aviation Organization and the Italian Civil Aviation Authority provisions in order to find out the best way to analyze friction data. The final goals are to assure an optimum scheduling of the Airport Pavement Management System, as well as users safety.

Peer-review under responsibility of Road and Bridge Research Institute (IBDiM)

* Corresponding author. Tel.: +39-081-768-3370; fax: +39-081-768-3946.
E-mail address: francesco.abbondati@unina.it

Peer-review under responsibility of Road and Bridge Research Institute (IBDiM)
doi:10.1016/j.trpro.2016.05.460

Keywords: runway; safety; rubber deposits; friction assessment; grip number; decay curves; Airport Pavement Management System

1. Introduction

According to statistics, air transport accidents are far more uncommon than any other means of transport. However, as note Čokorilo et al. (2014) they can be very perilous and fatal at the same time when they occur. Take--off and landing are the most dangerous phases of the entire flight: the pilot may lose control of the airplane and get out the runway. It was estimated that 45% of the accidents happens during take-off and landing which duration is only the 2% of the entire flight.

Research management, future issues, special problems, new technologies and innovation opportunities are addressed in Dell'Acqua et al. (2011a, 2012b, 2013a, 2013b, 2015). Examples of working systems are provided, as well as guidelines for implementation in Žilionienė et al. (2013), Dell'Acqua (2011), and De Luca et al. (2012, 2013).

Surface conditions is one of the categories of runway accidents hazards, as by Oriola and Adekunle (2015). See also Dell'Acqua et al. (2011b, 2012a). The parameter used in this study to assess the characteristics of the upper layer of the runway pavement, in order to evaluate its safety and function ability by De Luca et al. (2014), is friction.

The International Civil Aviation Authority (*ICAO*) *Doc 9137-AN/898 Part 2, Airport Services Manual, Pavement Surface Conditions* by 2002 requires to periodically assess the friction value so as to constantly control the safety of the runway.

The Italian Civil Aviation Authority (*ENAC* − L'Ente Nazionale per l'Aviazione Civile) provisions in the matter of friction measuring and reporting, *APT l0A Criteri per la valutazione delle condizioni superficiali di una pista* by 2014, take inspiration from ICAO Guides substantially reflecting them.

Other organizations have issued regulations in the field of aviation safety, among which that by *English Civil Aviation Authority* (*CAA*) *CAP 683: The Assessment of Runway Surface Friction Characteristics* by 2010 is noteworthy.

Cluster analysis is an unsupervised learning process of grouping observations into classes or clusters so that the observations in the same class share more common features than to those in other classes. Various algorithms perform this process, *K-means* algorithm by Hartigan (1975) is one of the most popular. An efficient version of the algorithm is presented by Hartigan and Wong (1979).

Subtractive Clustering by Chiu (1994) is an algorithm based on fuzzy logic to better handle imprecise observations. As note Bataineh et al. (2011) it generates accurate models and produces consistent results.

Classification and Regression Trees are called supervised learning algorithms, as the outcome variable of interest is previously known and supervises the process. The models are obtained recursively partitioning the data space and fitting a simple prediction model within each partition. As a result, the partitioning is represented graphically as a decision tree. The *Classification and Regression Trees* (*CART*) algorithm was popularized by Breiman et al. (1984). Considering an impurity index, it breaks up the initial data through a dichotomic splitting; only two nodes is generated from each element of the tree.

The *General CHi-Squared Automatic Interaction Detector (GCHAID*) algorithm is based on *CHi-Squared Automatic Interaction Detector* (*CHAID*), one of the oldest classification methods proposed originally by Kass (1980). *CHAID* algorithm provides a splitting condition that is either dichotomic or multiple through a Chi-square test; each element of the tree generates two or more nodes.

2. Techniques used in the study

Measurements data are not useful for an *Airport Pavement Management System* (*APMS*) if not properly analyzed to give back reliable results about the advancement of the phenomenon.

In fact, the main analytical tools for the representation of the decay phenomenon, are the *Grip Number* (*GN*) Time (*GN-Time*) decay curves, which require a value to be fixed as representative at each survey date.

To evaluate the best way to fix this value, different methods of elaboration that are divided into two groups were used:

- Analysis by Guides − in this group there are methods that are clearly stated or suggested by Organizations that deal with civil aviation safety. The two considered methods come from the Italian *ENAC APT-10A* and the English *CAA CAP 683;*
- Data mining analysis − in this group there are two unsupervised learning algorithms, *K-means* and *Subtractive Clustering*, and two supervised learning algorithms, *CART* and *GCHAID*. *K-means* algorithm has been associated with an auxiliary algorithm to find the optimal number of clusters. *Subtractive Clustering* algorithm does not need an additional procedure to give the optimal number of clusters. *CART* and *GCHAID* algorithms perform the task of choosing the optimal number of terminal nodes in the pruning step.

The system studied refers to a standard runway that is seen as composed of a series of points, placed on the ±3 and ±6 alignments, each of them having a specific *GN* value with a spacing 10 m.

The aim is to divide longitudinally the runway into parts that reflect the different areas of consumption at each survey date, in order to locate the one which presents greater degradation.

The mean value of *GN* of the area presenting greater degradation, is fixed as representative of the decay phenomenon of the whole runway and used to implement the decay curve.

2.1. Analysis by Guides: ENAC provisions

ENAC *APT-10A* clearly states necessary to provide friction surface information for each third of a runway. The one-third segments are called A, B and C. For the purpose of reporting information to aeronautical service units, segment A is always the one associated with the lower runway designation number. Assigning to each a friction value obtained by averaging the measurements collected in the part, the whole runway is characterized by the lower value.

The alternative criterion expressed by the *APT-10A*, leaves the analyst free to judge significant parts of only 100 m in order to locate the areas with functional deficit. But no rating scale on functional deficit is expressed, so the turn to the *CAA* provisions was did.

2.2. Analysis by Guides: CAA provisions

CAA CAP 683, on the other hand, suggests a method to locate the 100 m part and its friction value.

The *Continuous Friction Measurement Equipment* (*CFME*) collections are rearranged using what *CAA* calls the *Minimum 100 Meters Rolling Averages* method.

On a standard run, the Grip Tester Trailer provides values for each 10 meters increment along the run so that, over a distance of 100 m, an average is calculated. The first rolling average is the sum of the first 10 readings divided by 10 (RA1). The second rolling average is the sum of the readings 2 to 11 divided by 10 (RA2) and so on to the end of the run. A rolling average is best visualized as a 100 m long cursor passing over the surface of the runway. It can be moved to 10 different positions but still including the 10 m increment considered. After a value has been attributed to every 10 m increment of the run, the lowest of them is selected. Thus the process is repeated throughout the run in order to locate the minimum 100 meters rolling average at any 10 m segment on the run.

2.3. Data mining analysis: K-means algorithm

K-means algorithm is a crisp (or hard) clustering method, where the terms crisp/hard refer to the case of classic logic. It was published for the first time by Hartigan (1975), the aim is to divide the observations X_j= (*Chainage; GN*), into an assigned number of groups (k) minimizing the distance of each point from its cluster center (equation (1)):

$$V(U,C) = \sum_{i=1}^{k} \sum_{X_j \epsilon P_i} ||X_j - C_i||^2 \quad (1)$$

where C_i − the cluster centers, P_i − the clusters, U − a partition matrix of the objects j among the clusters i.

The drawback of the algorithm is that it asks to choose the number of clusters (k) to find.

For this reason it is associated with an auxiliary algorithm to find the optimal number of clusters, chosen to be similar to the one set to perform the pruning step both in *CART* and *GCHAID* algorithms, to obtain more comparable results. The general idea of the *V*-fold cross-validation is to divide the overall observations into a number (v) of folds. The same analysis is then successively applied to the observations belonging to the v-1 folds, and the results of this analysis are applied to the sample v to compute an index of predictive validity, i.e. how well the observations in sample v is assigned to homogeneous clusters using the current cluster solution computed from the v-1 learning sample. The number of folds used to perform the cross-validation is 10. The best number of clusters is given by the stabilization of a cluster cost function (calculated on the basis of average distance of cases from cluster centers).

2.4. Data mining analysis: Subtractive Clustering algorithm

Subtractive Clustering algorithm is a fuzzy clustering method that is belonging to the fuzzy logic, proposed by Chiu (1994). The algorithm is based on computing a potential value for each data point based on its distances to surrounding points and on the idea that each of them is a potential cluster center. A data with many data points nearby will have a high potential value. The measure with the highest potential value is chosen as the first cluster center. The key idea is that once the first cluster center is chosen, the potential of all data points is reduced according to their distance from the cluster center. Points near the first cluster center will have greatly reduced potential. The next cluster center is then placed at the data point with the highest remaining potential value. This procedure of acquiring new cluster center and reducing the potential of surrounding points repeats until the potential of all of them falls below a threshold.

This procedure is carried out by the equations (2) and (3):

$$P_i = \sum_{j=1}^{n} e^{\frac{-\left\|x_i - x_j\right\|^2}{\left(\frac{r_a}{2}\right)^2}}, \tag{2}$$

$$P_i = P_i - P_1^* \sum_{j=1}^{n} e^{\frac{-\|x_i - x_1\|^2}{\left(\frac{r_b}{2}\right)^2}} \tag{3}$$

Equation (2) represents the potential of each data point x_i, where the symbol $\|.\|$ denotes the Euclidean distance, and a positive constant r_a is the accept ratio. The accept ratio is effectively the radius defining a neighborhood; data points outside this radius have little influence on the potential. The highest potential is selected.

Then the potential of each data point is revised with the second formula that depends on the distance between the point considered and the cluster center previously determined, and a positive constant r_b which is the reject ratio. It reduces the potential of data near the first cluster center chosen in order to look for another one in a different location, far away from the first.

In addition there are criterion for stopping the clustering finding and to avoid marginal cluster centers, in which are involved the other two parameters, influence range and squash, mentioned in Table 1 with the previously defined.

Table 1. Constants used with Subtractive Algorithm.

Constant	Value
Influence range	0.33
Squash	1.25
Accept ratio (r_a)	0.50
Reject ratio (r_b)	0.15

2.5. Data mining analysis: CART algorithm

Classification and Regression Trees are hierarchical splitting methods for constructing prediction models from data. The models are obtained dividing the data space recursively into a certain number of nodes determining a series of logical if-then until a stopping rule is reached. Afterwards a pruning rule provides the best number of terminal nodes. As a result, the partitioning is represented graphically as a decision tree, with prediction error measured in terms of misclassification cost. The last nodes of the sequence, called *terminal nodes,* are the parts in which the original data set has been divided.

The *CART* algorithm was popularized by Brieman et al. (1984). It breaks up the initial data through a dichotomic splitting (only two nodes is generated at each split) considering the Gini index as an impurity index of the node (t) (equation (4)):

$$I(t) = \sum_{i \neq j} p\left(\frac{i}{t}\right) p\left(\frac{j}{t}\right) = 1 - \sum_j p^2\left(\frac{j}{t}\right) \tag{4}$$

where $p\left(\frac{j}{t}\right)$ − the frequency of the observation i in the class t. The threshold value for the split is chosen in correspondence of the minimum impurity.

As stopping rule have been set n = 22 as the minimum number of cases.

As last step, for pruning and selecting the right-sized tree (i.e. the optimum number of terminal nodes), the *V*-fold cross-validation (with v = 10) optimizes a cost-complexity function (5):

$$R_\alpha(T) = \sum_{t \epsilon \tilde{T}} R_\alpha(t) = R(T) + \alpha|\tilde{T}|, \tag{5}$$

where $R_\alpha(t)$ − the rate of misclassification of a node t, $R(T)$ − the rate of misclassification of the generic tree T, α − a positive real, $|\tilde{T}|$ − the cardinality of the terminal nodes.

2.6. Data mining analysis: GCHAID algorithm

The *GCHAID* algorithm is a version of the original *CHAID* proposed by Kass (1980), able to deal with quantitative dependent variables, like *GN*, instead of categorical only. It provides splitting conditions that is either dichotomic or multiple (each node of the tree generates two or more nodes) through a statistical *F*-test: if the test performed on predictor variables values in a node is significant (i.e. less than an alpha value) the algorithm will compute an adjusted *p*-value and will choose for the split the predictor variable with the smallest adjusted *p*-value; otherwise the node is terminal.

As stopping rule and pruning have been set the same as in *CART*.

3. Data collection

In order to perform the analysis, it is necessary to have a base of raw data regarding friction. A case study concerns the runway of the International Civil Airport of Lamezia Terme (Italy). It is equipped with a 4D class runway named RWY 10/28, length 2,416 m and wide 60 m. Horizontal landmarks border 2,200 m and 45 m. The measurements were made along the alignments +3 m, −3 m, +6 m, −6 m with respect to the center line, which are the most used areas according to the main landing gears of the aircraft traffic mix, and for the whole length of the runway.

Data were collected using a Grip Tester Trailer device, according to *Doc 9137-AN/898*, a CFME made up of a small carriage with a smooth surfaced tire and fixed skidding percentage, connected to a truck which enables the Grip Tester to carry along the runway. It is endowed with a self-wetting device that allows it to pour water on the pavement before the measurement tire passing. The measured parameter is the *GN*, the output is given each 10 m. Each measurement was conducted under dry conditions according to *Doc 9137-AN/898*. The final database consists of 60 alignments raw data, which refer to 15 surveys made in the years from the year 2009 to 2014. In November 2011 it was made a de-rubberizing maintenance operation by means of high-pressure water TrackJet machine.

A first analysis of raw data by means of scatterplots gives immediately an information about the behavior of the *GN* along the runway, and therefore about the different areas of consumption: its value is significantly low in the first and last sides (respectively side 10 and side 28) where the operations of take-off and landing are more frequent. In particular, the values are lower in the first part, side 10, which is more frequently used for landing.

It was observed that the two opposite tracks (+3 and −3; +6 and −6) representing the evolution of the *GN* along the runway has almost symmetrical shapes. Hence, to simplify the subsequent steps of elaboration the friction values of the two symmetrical alignments were averaged and from now on have been referred to them with ±3 and ±6 (Figure 1).

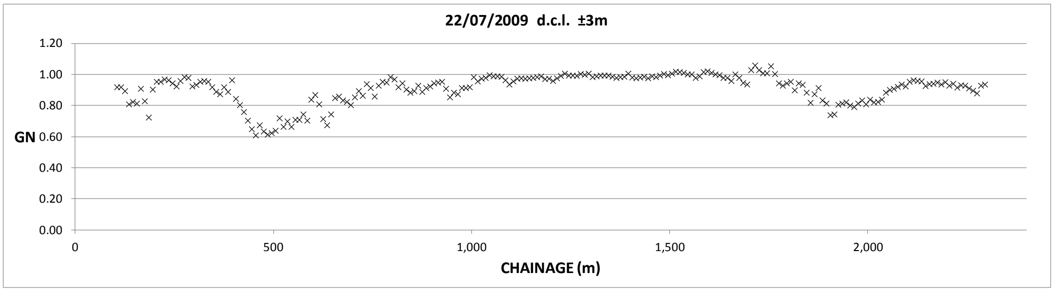

Fig. 1. Raw data survey July 22, 2009, distance from center line ±3.

4. Data analysis

The *ENAC* – one-third provision has always given the segment A, which matches with threshold 10, as the most interested by the decay phenomenon.

For the *CAA* provision, the whole runway at each survey and for each alignment (±3; ±6) has been characterized by the minimum rolling average, representing a part of 100 m changing location over time.

At the end of the two clustering analysis it was obtained different average number of clusters (Table 2). The *K-means* algorithm appears to have always determined more clusters than the *Subtractive Clustering*. At each survey date and for each alignment (±3; ±6) the whole runway is characterized by the cluster center with the lower *GN* value.

Table 2. Average number of clusters obtained.

Algorithm	Av. N° of clusters 2009–2011 ±3	Av. N° of clusters 2012–2014 ±3	Av. N° of clusters 2009–2011 ±6	Av. N° of clusters 2012–2014 ±6
K-means	6.57	6.12	6.42	5.75
Subtractive Clustering	3.85	3.25	4.00	3.50

Table 3 shows the average numbers of terminal nodes obtained with *CART* and *GCHAID* algorithms. *CART* appears to have always determined more nodes than *GCHAID*. At each survey date and for each alignment (±3; ±6) the whole runway is characterized by the node with the lower *GN* mean value.

Table 3. Average number of terminal nodes obtained.

Algorithm	Av. N° of nodes 2009–2011 ±3	Av. N° of nodes 2012–2014 ±3	Av. N° of nodes 2009–2011 ±6	Av. N° of nodes 2012–2014 ±6
CART	6.85	7.62	7.57	6.62
CHAID	3.42	3.87	3.57	3.50

Scatterplots were built using the results by Guides and by data mining techniques. The raw data collected by Grip Tester Trailer provide, for each of the 15 survey date and for each alignment ±3 and ±6, six values of *GN*. So 12 decay curves were built. For each decay curve a non-linear regression was made to better represent the physical phenomenon, cumulative type. The quadratic polynomial function has been found to be the best fit for all the curves, generating determination coefficients, R^2, higher than any other non-linear or linear function.

A general shape shown by all the curves is due to the maintenance operation made at the end of the year 2011: at the de-rubberizing date there is a sharp increase of the *GN* values that divides each curve into two separated branches, one before and one after the de-rubberizing maintenance operation. The different slope of the two branches of each curve as well as the different start and end points, are the result of the different methods.

Comparing the curves obtained with the different methods, have been assigned different degrees of reliability to the obtained results. The consideration made are in the following.

- *ENAC* decay curves – It is known, this is the method mostly used in practice for scheduling maintenance operations on the runways. The main disadvantage is that the GN values are quite higher than that observed in other procedures, due to averaging measurements on each third of the runway. In such a way areas with surface texture more preserved because more distant from the landing area, make the average higher (Figure 2).

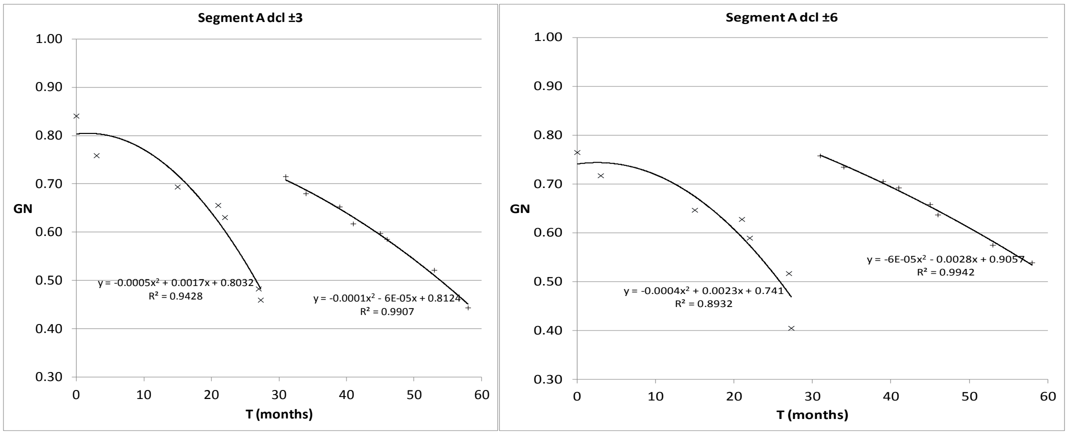

Fig. 2. *ENAC* decay curves, distance from center line ±3 and ±6.

- *CAA* decay curves – The CAA elaboration procedure shows points with lower GN value, i.e. it reveals maintenance necessity more early than ENAC – one-third. The ±6 curve shows that it has been yield very differently from the ±3: the second branch of the ±6 curve has a different slope quite reversed and shifted very upward (Figure 3).

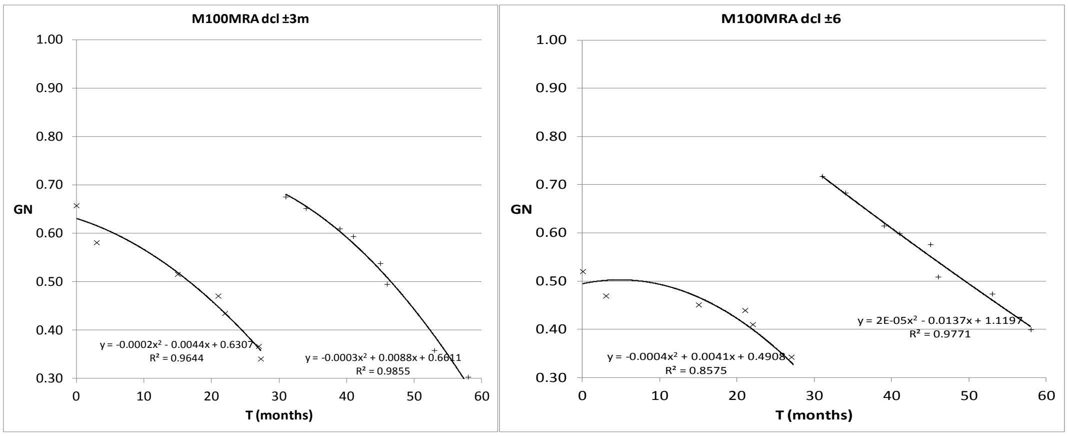

Fig. 3. *CAA* decay curves, distance from center line ±3 and ±6.

- *GCHAID* and *Subtractive Clustering* decay curves – Both of them generating less cluster centers than CART and K-means, they show opposite results in the ±3 curves. GCHAID second branch shifted downward and has a lower slope than the first branch; besides, the two branches end with higher values than CART and K-means, so to result not precautionary. Subtractive Clustering second branch shifted upward and has a little higher slope than the first branch; besides, the two branches end with a little higher values than CART and K-means (Figure 4).

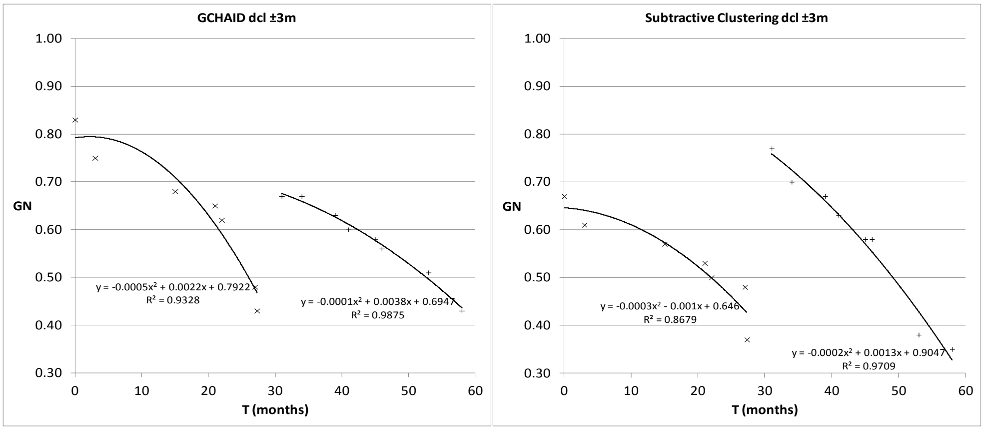

Fig. 4. *GCHAID* and *Subtractive Clustering* decay curves, distance from center line ±3.

- *CART* and *K-means* decay curves – They show very similar results both in ±3 and ±6 curves. Their branches have regular and comparable behavior in the slopes as well as in the start and end points. Besides, they always end with precautionary values (Figure 5).

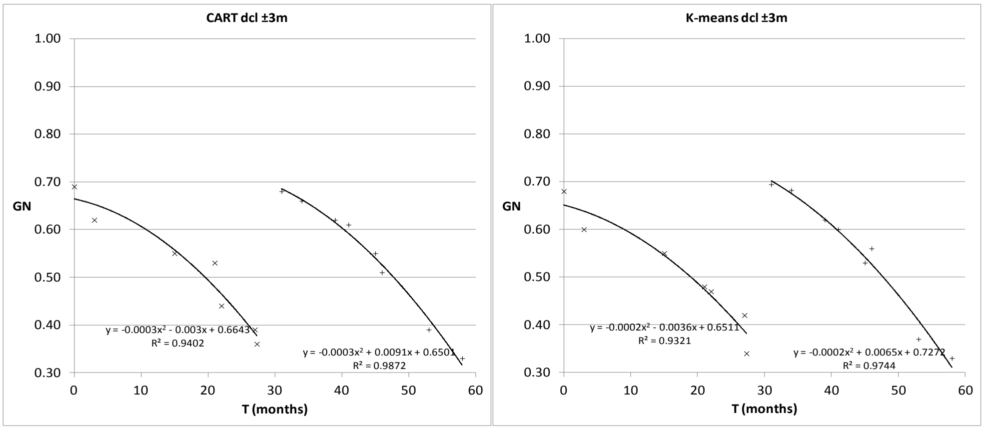

Fig. 5. *CART* and *K-means* decay curves, distance from center line ±3.

A comparison was made between the variance distributions of the values coming out from the two methods (Figure 6) to evaluate the precision of the partitions made by the two algorithms. The area swept by the ±3 *K-means* variance distribution is 99.7% than that by ±3 *CART*. Conversely, the area swept by ±6 *CART* is 90.0% than that by ±6 *K-means*. In both cases the percentages are very close to each other and suggest that both methods are equally capable of well representing the decay phenomenon.

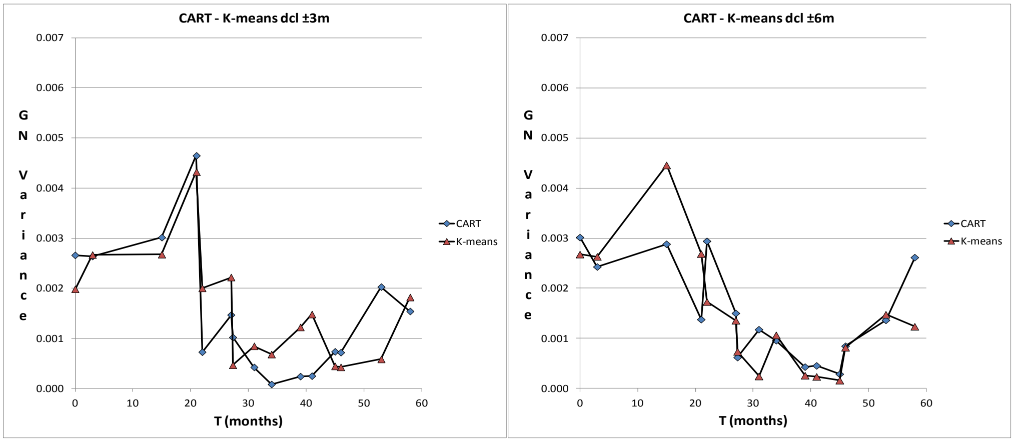

Fig. 6. Values variances.

Friction levels represented by the decay curves must be compared with the following friction levels provided by Guide *ENAC*:

- Design Objective Level (DOL) − 0.74
- Maintenance Planning Level (MPL) − 0.53
- Minimum Friction Level (MFL) − 0.43

A comparison has been made also between the Guide friction levels and the *ENAC* and *CART* decay curves obtained (Figure 7) highlighting in the charts the MPL and MFL levels.

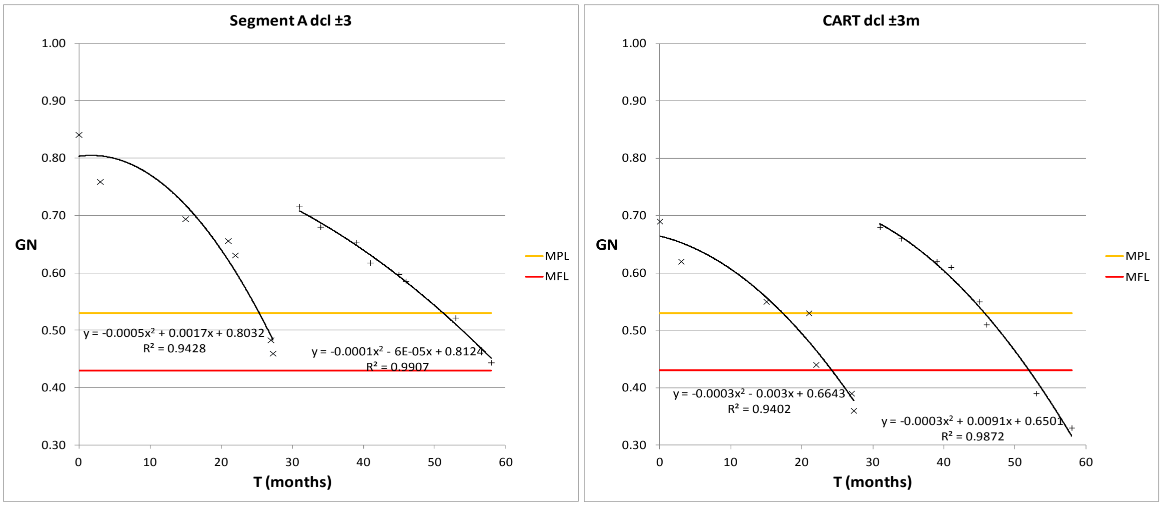

Fig. 7. ENAC and CART decay curves with highlighted MPL and MFL, distance from center line ±3.

It is well evident how the first method is not precautionary as the second. *CART* decay curve warns the maintenance need with and advance of 5−6 months at least: two points for each branch go below the line of the Minimum Friction Level.

5. Conclusions

Measurements carried out with high performance devices like *Continuous Friction Measurement Equipment* provide a large base of data, from which to extract information to properly characterize the physical phenomenon under investigation.

Modeling the appropriate decay laws is an important success factor for an *Airport Pavement Management System* in ensuring maximum safety and cost saving as possible.

On the basis of the preliminary empirical study performed, the method of analysis of the measurements is discriminative for the representation of the decay laws.

The current Guide's provisions in matter of friction measuring and reporting appear not univocally encoded in an analysis method of collected measurements, so to allow the analyst to investigate and deal with an appropriate one.

The comparison made shows that the data mining techniques based on *Clustering* and *Classification/Regression Trees* models gives very reliable and precise results. In fact their algorithm leave the data to declare themselves, resulting more flexible in assessing the different consumption areas in terms of extension and intensity of the phenomenon.

This stimulates the scientific community to further research in the field of assessment of the pavement surface characteristics.

References

Bataineh, K.M., Naji, M., Saqer, M., 2011. A Comparison Study between Various Fuzzy Clustering Algorithms. Jordan Journal of Mechanical and Industrial Engineering 5(4), 335−343.

Breiman, L., Friedman, J., Stone, C. J., Olshen, R.A., 1984. Classification and Regression Trees. Chapman and Hall/CRC, First edition, Boca Raton, London, New York, Washington D.C. pp. 368.

Chiu, S.L., 1994. Fuzzy Model Identification Based on Cluster Estimation. Journal of Intelligent and Fuzzy Systems 2, 267−278.

Čokorilo, O., De Luca, M., Dell'Acqua, G., 2014. Aircraft Safety Analysis Using Clustering Algorithms. Journal of Risk Research 17(10), 1325−1340.

De Luca, M., Dell'Acqua, G., 2012. Freeway safety management: case studies in Italy. Transport 27, 320–326.

De Luca, M., Dell'Acqua, G., 2013. Calibrating the passenger car equivalent on Italian two line highways: a case study. Transport published online 16 Oct 2013, 1–8.

De Luca, M.; Dell'Acqua, G., 2014. Runway Surface Friction Characteristics Assessment for Lamezia Terme Airfield Pavement Management System. Journal of Air Transport Management 34, 1−5.

Dell'Acqua, G., 2011. Reducing Traffic Injuries Resulting from Excess Speed Low-Cost Gateway Treatments in Italy. Transportation Research Record: Journal of the Transportation Research Board 2203, 94–99.

Dell'Acqua, G., 2015. Modeling Driver Behavior by Using the Speed Environment for Two-Lane Rural Roads. Transportation Research Record: Journal of the Transportation Research Board 2472, 83–90.

Dell'Acqua, G., Busiello, M., Russo, F., 2013b. Safety Data Analysis to Evaluate Highway Alignment Consistency. Transportation Research Record: Journal of the Transportation Research Board 2349, 121–128.

Dell'Acqua, G., De Luca, M., Lamberti, R., 2011b. Indirect skid resistance measurement for pourus asphalt pavement management. Transportation Research Record: Journal of the Transportation Research Board 2205, 147–154.

Dell'Acqua, G., De Luca, M., Russo, F., 2012a. Procedure for Making Paving Decisions with Cluster and Multicriteria Analysis. Transportation Research Record: Journal of the Transportation Research Board 2282, 57−66.

Dell'Acqua, G., De Luca, M., Russo, F., Lamberti, R., 2012b. Mix Design with Low Bearing Capacity Materials. The Baltic Journal of Road and Bridge Engineering 7(3), 204−211.

Dell'Acqua, G., Russo, F., 2011a. Road Performance Evaluation Using Geometric Consistency and Pavement Distress Data. Transportation Research Record: Journal of the Transportation Research Board 2203, 194−202.

Dell'Acqua, G., Russo, F., Biancardo, S.A., 2013a. Risk-type density diagrams by crash type on two-lane rural roads. Journal Of Risk Research 16, 1297–1314.

Hartigan, J.A., 1975. Clustering Algorithms. Probability & Mathematical Statistics, John Wiley & Sons Inc.

Hartigan, J.A., Wong, M. A., 1979. Algorithm AS 136: a K-Means Clustering Algorithm. Applied Statistics 28(1), 100−108.

Kass, G.V., 1980. An Exploratory Technique for Investigating Large Quantities of Categorical Data. Applied Statistics 29(2), 119−127.

Oriola, A.O., Adekunle, A.K., 2015. Assessment of Runway Accident Hazards in Nigeria Aviation Sector. International Journal for Traffic and Transport Engineering 5(2), 82–92.

Žilionienė, D., De Luca, M., Dell'Acqua, G., 2013. Evaluation of Climatic Factors Based on the Mechanistic-Empirical Pavement Design Guide. The Baltic Journal of Road and Bridge Engineering 8(3), 158−165.