



**HAL**  
open science

## Impact of road infrastructure characteristics on road markings

Ali Tidjani, Maxime Redondin, Laurent Bouillaut, Dimitri Daucher

► **To cite this version:**

Ali Tidjani, Maxime Redondin, Laurent Bouillaut, Dimitri Daucher. Impact of road infrastructure characteristics on road markings. ESREL 2019, 29th European Safety and Reliability Conference, Sep 2019, Hanovre, Germany. 9p. hal-02968219

**HAL Id: hal-02968219**

**<https://hal.science/hal-02968219>**

Submitted on 15 Oct 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Impact of road infrastructure characteristics on road markings

Ali Tidjani and Maxime Redondin

VEDECOM Institute, 23 allée des marronniers, F-78000 Versailles, France.

E-mail : alitidjani680@gmail.com, maxime.redondin@vedecom.fr

Laurent Bouillaut<sup>1</sup> and Dimitri Daucher<sup>2</sup>

<sup>1</sup>Université Paris-Est, Grettia (IFSTTAR), F-7755 Marne-la-vallée, France

<sup>2</sup>Université Paris-Est, Lepsis (IFSTTAR), F-7755 Marne-la-vallée, France

E-mail : laurent.bouillaut@ifsttar.fr, dimitri.daucher@ifsttar.fr

**ABSTRACT:** Road markings channel traffic flows and guide road users visually, especially at night. Their quality and reliability are, therefore, important issues which must be closely followed to ensure the safety of road users and especially their perceptibility for the future autonomous cars. Thus, previous study has shown that it is possible to segment markings of the center line into maintenance strategic areas. However, this study didn't take into account all the characteristics of road infrastructures that could have an impact on the degradation of markings. Indeed, depending on the structure and complexity of all road infrastructures, these road markings could present a different degradation process over the time and in specific areas. So, the identification of these areas can help road managers to develop localized maintenance strategies. In this study, a Factor Analysis of Mixed Data (FAMD) combined with an Agglomerative Hierarchical Clustering (AHC) is proposed as the method to identify these areas or clusters in the case of the French National Road 4 (NR4).

**Keywords:** Road markings, FAMD, AHC, road infrastructure, surface signing, degradation.

## 1. Introduction

The surface signing regroups, all the marks on roads used to materialize the rules of use of the ways open to the public traffic. It is an important aid to driving, channeling traffic flows, specifying rules of roads, materializing parkings and guiding the user visually, especially at night. Recent papers from VEDECOM Institute (Reveilloud, Gruyer et Rahal, A lane marker estimation method for improving lane detection 2016) (Reveilloud, Gruyer and Rahal, A new multi-agent approach for lane detection and tracking 2016) show that a clear and visible horizontal signage is more important than a roadway in good condition for an autonomous vehicle.

The road markings are made perceptible at night thanks to their retroreflection luminance. Indeed, markings return the light emitted by the headlights of a vehicle towards its driver. In other words, retroreflection allows the visibility of road markings in night time.

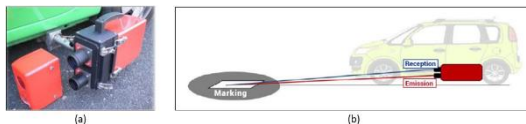


Fig. 1. (a) The retroreflectometer Ecodyn, (b) Measurement process (Nextroad 2019)

The retroreflection luminance is quantified by devices called retroreflectometers. Figure 1 presents the case of the Ecodyn from MLPC (Nextroad 2019). The device is designed to be

fixed on a vehicle and measured dynamically the retroreflection of road markings (Fig.1(a)). The measuring process is to attach the unit on the right or left side of the inspection vehicle. A white light source illuminates the marking at 6 m at the front of the camera and the amount of retro-reflected light is

However, road markings are subjected to a phenomenon of degradation. The figure 2 present it. A new marking is a white object and admits a strong retroreflection luminance ( $600 \text{ mcd/m}^2/\text{lx}$ ). Over time, the marking undergoes a decrease in their level of retro-reflection luminance. In the best case, it will become a grey object and admit a low retroreflection luminance ( $50 \text{ mcd/m}^2/\text{lx}$ ).

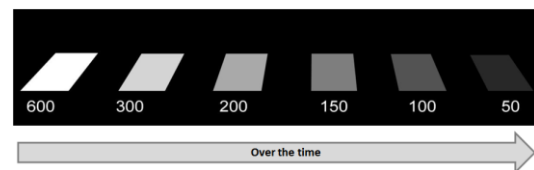


Fig. 2. Road markings degradation over the time

Available inspection devices stipulate that the retroreflection luminance of markings is the only measure used for evaluating marking degradation. The NF EN 1436 rules (AFNOR 2009) indicates that a new marking must admit a minimum of  $150 \text{ mcd/m}^2/\text{lx}$ . As this degradation of markings is an important issue, it is necessary for road network managers to take care of the

markings to avoid any unfortunate consequences. Also, it must be added that the development of a system maintenance strategy requires a first analysis study which can allow identifying maintenance areas of road markings. In addition, the literature of the last thirty years presents mainly models of degradation based on regression methods. For example, (Lu 1995) proposed an exponential model as a function of age of markings, (Abboud and Bowman 2002), developed different exponential models as a function of the Annual Average Daily Traffic (AADT) and the age of markings, (Sitzabee, Hummer and Rasdorf 2009) proposed a multilinear model as a function of time, the initial retroreflection, the AADT, the lateral locations of markings and marking color. However, this approach remains relatively difficult to apply to the reality of the field.

The thesis report of Maxime Redondin (M. Redondin, 2018) presents different maintenance strategies for road markings according to their retroreflection luminance. This thesis shows that it is possible to segment, at the scale of a road, the markings of the center line in strategic areas of maintenance (M. Redondin, et al. 2017) and analyzed the life cycles of the markings of the center line through a Weibull analysis (M. Redondin, L. Bouillaut, et al. 2018). However, this thesis didn't consider the different characteristics of a road infrastructure that could have an impact on the degradation of markings.

The ambition of this study is to segment the road infrastructure according to different characteristics such the degradation of the pavement, the road surface signing or traffic lanes configuration. This approach produces an identification of strategic areas of maintenance based on the road infrastructure. The general objective of this study is the evaluation of the impact of the characteristics of road infrastructure on the quality of road markings. In other words, the central question that will be dealt with throughout this study is to know to what level road infrastructure can impact the strategic areas of road markings maintenance.

In this study, a Factor Analysis of Mixed Data (FAMD) (Escofier and Pagès 2008) (Pagès 2004) combined with an Agglomerative Hierarchical Clustering (AHC) (Tufféry 2011) is proposed as the method to identify strategic areas of road markings maintenance.

Then, three markings lines are identified: the center line, the emergency one and the edge one. The annual inspections of these three marking lines are provided by CEREMA and a monitoring of the bitumen is provided by DIR EST. Also, some data bases that describe the major characteristics of road infrastructures such as the types of roadways were collected on the Ministère

de l'Ecologie et de la Transition Energétique (*the French Ministry of ecological and solidarity transition*) website. A fusional algorithm has been elaborated to gather all these data bases into a big one before applying the FAMD and AHC.

As an illustration of this approach, a 100km section of the French National Road 4 (NR4) is considered.

After the introduction of the classification algorithm, this paper will focus on the fact that the surface signing is more segmented according to both characteristics and parameters of the considered road infrastructure. Another interesting point is that the areas which weren't classified in the previous study (M. Redondin, et al. 2017) have been classified in new clusters in this case.

Finally, this study is a first step to introduce the modeling and optimization of road infrastructure maintenance for the autonomous car.

## 2. Collection of the used data

This section presents the three data sources. To simply, each variable is defined by a key word. The study is based on different inspections realized between 2007 and 2015.

The first data source is provided by CEREMA EST. The surface signing is annually inspected each September since 2007. The chosen retroreflectometer is an Ecodyn. Three lines are identified: the broken center line, the emergency one and the median strip one. However, the FAMD considers only the center and the emergency line because those of the median strip line is not available all over the NR4. To simply: C20XY, E20XY, M20XY represent respectively the retroreflection luminance evaluated during the 20XY campaign of the Center, Emergency and Median strip line.

The second data source is provided by DIR EST. The pavement is yearly inspected. The different types of bitumen identified are: Pure Thin Bitumen (PTB), Modified Thin Bitumen (MTB), Very Thin Bitumen (VTB), Very Thin Bitumen Modified (VTBM), Semi Granular Bitumen (SGB), Semi Granular Modified Bitumen (SGMB) and High Module Coated (HMC). There was no replacement of bitumen between 2007 and 2015. The age (Age) of pavement is also provided.

The final data source is provided by *the French Ministry of ecological and solidarity transition*. The ministry launched an open data initiative in 2017 as part of the Open Data in a goal of transparency in the light of several sectors (Ministère de l'Ecologie et de la Transition Energétique 2018). The objectives of the government in making the data available to the public are to inform decision makers and citizens

in a transparent way about the state of the heritage and the use of public funds, to provide data to the digital sector companies who develop user services and strengthen the role of the state as a service provider in certain areas. The Ministry feeds and publishes a dozen sub-bases.

Each Point of Reference (PR) is localised in Lambert coordinates 93. The distance in meter of each consecutive PR is deduced.

The traffic lane between two PRs are described (see figure 3):

- single lane: 2 ways superior to 7m ( $2(>7)$ ),
- single lane: 2 ways inferior to 7m ( $2(<7)$ ),
- single lane: 3 ways superior to 10m ( $3(>10)$ ),
- single lane: 3 ways inferior to 10m ( $3(<10)$ ),
- single lane: 4 ways (4),
- separate lane  $2 \times 2$  ways ( $2 \times 2$ ),
- separate lane  $2 \times 1$  way ( $2 \times 1$ ).

Crossings on a given PR: interchange with a highway (Interchange), crossing with another road (Crossing) and split. In other case, there is no specific observation (No).

The general degradation level of the roadway: Excellent (Exc), Acceptable (Acc), Bad, Poor.

Other information between two PRs: the Annual Average Daily Traffic (AADT) and the width of roads (Width).

The databases are updated year after year and allow a follow-up in time. After the collection of these databases, a fusional algorithm has been elaborated to gather them into a big one that will serve as the support to the methodological framework.

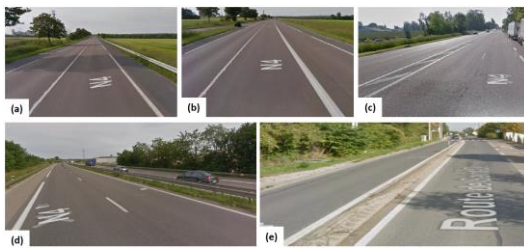


Fig. 3. (a) Single floor 2 ways, (b) single floor 3 ways, (c) single floor 4 ways, (d) separate floor  $2 \times 2$  ways, (e) separate floor  $2 \times 1$  way.

This final database defined each PR by a specific road infrastructure, a degradation level of pavement and a monitoring of the surface signing. This description is done by qualitative and quantitative variables. The FAMD proposes a data management and identifies different correlation between variables. The final clustering identified different PR which admit both a similar description and a similar monitoring of markings.

### 3. Methodological framework of the study

#### 3.1 Factorial Analysis of Mixed Data

The Factorial Analysis of Mixed Data (FAMD) is dedicated to the analysis of data containing both qualitative and quantitative variables. In practice, the FAMD behaves globally as a Principal Component Analysis (PCA), transforming the qualitative data into a Complete Disjunctive Table (CDT). In fact, a CDT makes it possible to transform qualitative data into a new table indicating whether individuals possess the property of the qualitative variable in question. The CDT's weight is fixed by (Pagès 2004).

PCA is an exploratory technique that summarizes the information contained in a set of data into several synthetic variables which are linear combinations of the original variables based on data discrimination.

To produce a summary of information in the sense of the PCA is to establish a similarity between individuals, to search for groups of homogeneous individuals, to highlight a typology of individuals. As for the variables, it is to highlight the links between them, by means of synthetic variables and to highlight a typology of variables. The PCA seeks in a general way to establish links between these two typologies.

One of the main objectives of this method is to reduce the size of the set of original variables or, in other words, the representation of the maximum amount of information contained in the starting variables using a reduced number of new variables, correlated between them. Indeed, a small set of non-correlated variables is easier to understand and use in different analyses than a larger one.

Thus, the first steps of the FAMD under the formalism of the PCA correspond to upstream data processing. As a first step, the data table is split in two: the quantitative data on the one hand and the qualitative data on the other. In accordance with this formalism, the first table is normalized. The qualitative part is transformed into a CDT, each column of which is weighted by the inverse of the square root of the standard deviation. Finally, it is just about merging the two tables. The resulting table is interpretable by the PCA.

After these steps, the following ones correspond to two different but complementary points of view: the matrix of individuals and the matrix of correlations. The first describes each individual according to the descriptive variables and the second describes the descriptive variables according to the individuals. On each matrix, the eigenvalues are calculated and the eigenvectors are deduced.

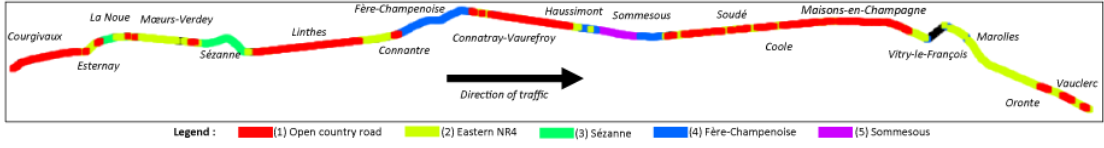


Fig. 4. The NR4 broken center line segmented into 5 replacement areas

Then, geometrical interpretations of the matrices of the individuals and correlations are carried out: the geometric axes are formed by the eigenvectors, the hyperplane of the individuals and the hypersphere of the correlations are represented. Descriptive variables are always enclosed in the hypersphere which radius is 1 (Pagès 2004).

Then successive Hilbert projections to a two-dimensional Cartesian plane are made to represent the map of individuals and the circle of correlations. The map usually has point clouds. Here, each cloud represents individuals who tend to share the same descriptive variables. The circle presents the main discriminating variables: the closer the variable is to the circle boundary, the more discriminating it is. Moreover, the more variables are distant, the more they are opposed. In the opposite, the closer two variables are, the more they are associated.

Each eigenvector contributes to the total variance of the ACP model. The principal components are the vectors that contribute the most.

### 3.2 Agglomerative Hierarchical Clustering

The Agglomerative Hierarchical Clustering (AHC) is an iterative classification method whose principle consists in producing sequences of nested partitions of increasing heterogeneity (between the partition in  $N$  classes where each object is isolated, and the partition in 1 class which groups all the objects) (Tufféry 2011).

The AHC is usable as soon as there is a notion of distance: in a space of individuals or in a space of variables. It is necessary to have defined the distance of two objects, which is generally natural, and the distance of two classes, which leaves more possibilities. It is precisely this distance that constitutes the criterion.

The algorithm of the AHC is established according to the following scheme:

1. The initial classes are the objects;
2. The distances between classes are calculated;
3. The two closest classes are merged and replaced by one;
4. Resume from 2. until you have only one class, which contains all the objects.

It should be emphasized at this level that the dissimilarity criterion used in the studies here is Ward's distance. This choice is justified by the fact that Ward's distance minimizes the loss of information because of data aggregation. This distance is defined by Eq.1 as:

$$d_W(X, Y) = \frac{d(x, y)^2}{\frac{1}{n_x} + \frac{1}{n_y}} \quad (1)$$

## 4. Application to the French National Road 4

### 4.1 The French National Road 4

The National Road 4 (NR4), located in the department of Marne, is French National Road that connects Paris and Strasbourg. It is the basic example that will serve as a support in this study. The previous study (M. Redondin, et al. 2017) identified five strategic areas on this road as shown in Figure 4 below. This figure also shows that the streets of Vitry le François are inaccessible.

The five clusters of the figure 4 are such as:

- The cluster 1 corresponds to open country roads.
- The cluster 2 corresponds mainly to the region of Vitry-le-François.
- The cluster 3 corresponds to the bypass of Sézanne.
- The cluster 4 corresponds to the bypass of Fère-Champenoise.
- The cluster 5 corresponds to a local highway near of Sommesous.

However, these clusters of the last study are obtained without considering the road infrastructure characteristics (M. Redondin, 2018).

### 4.2 Factorial Analysis of Mixed Data on NR4

After applying the PCA to the reconstructed quantitative data and qualitative data transformed into a CDT (through an algorithm of the FAMD developed on the software R), the proportions of the total variance explained by the first eleven main components are as presented by the Figure 5. It shows that the first component alone accounts for 14.7% of the total variance and the three first combined explain 33.2% of the variance.

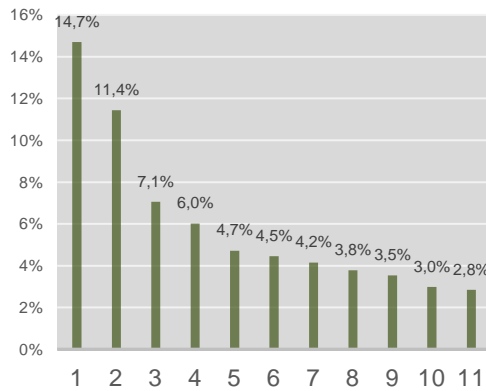


Fig. 5. Proportions of the variance explained by the factorial axes.

There are several criteria for choosing the number of components in a given study. The most common criterion is the Kaiser one. It is such that: an eigenvalue greater than 1 indicates that the principal component (PC) concerned, represents more variance compared to a single original variable, when the data are standardized. This is usually used as the threshold from which the main components are kept. Note that this is only applied when the data is normalized. In the present study, this criterion leads to the selection of 8 axes, explaining 59.9% of the total information contained in the data.

Another method for determining the number of principal components is the Karlis-Saporta-Spinakis criterion (KSS). The KSS is defined by the equation (2). Indeed, this criterion is one of the most reliable for determining the number of main axes because it takes into account a ratio between the number of variables and the number of individuals. This selection principle consists in retaining the components whose eigenvalues are greater than KSS with:  $p$  the number of variables and  $n$  the one of the individuals. This is, indeed, a robust variant of Kaiser's principle (Karlis, Saporta et Spinakis 2003).

$$KSS = 1 + 2 \sqrt{\frac{n - 1}{p - 1}} \quad (2)$$

The application of the Karlis-Saporta-Spinakis principle selects the first nine main components with which, results a space having a large proportion of the total variance.

The interpretation of the factorial axes is done sequentially, for each axis and each point cloud, by looking at the contributions to the formation of the axes. The variables contributing most to the formation of a given axis are those whose coordinates on this axis are close to 1 in absolute value.

The first Component opposes single floor 2 ways with acceptable condition, without crossing, and thin coated, modified bitumen to separate floor 2 × 2 ways that are in excellent condition.

The second component opposes single floor 2 ways with an acceptable condition, without crossing, and lined with pure thin asphalt with separate floor 2 × 2 ways that are in poor condition and crossing a portion with a difference in level.

The third one opposes single floor 2 ways with poor condition, crossing a portion with a difference in level and semi-grained bitumen pavement to separate floor 2 × 2 ways that are in poor condition and without crossing.

The remaining components can be described in a lighter way. Component 4 opposes markings whose pavement condition is acceptable to sections admitting emergency stop strip. Component 5 opposes single floor 4 ways in state mediocre with separate floor 2 × 2 ways crossing an elevation gain. Component 6 opposes single floor 4 ways in poor condition and made of pure paved bitumen very thin bitumen modified. Component 7 opposes separate floor 2 × 1 way in poor condition made of pure semi-grained bitumen to the rest. Component 8 is between pavements in an acceptable condition having an exchanger and made of very thin bitumen pure with separate floor 2 × 1 way in bad state and without crossing. The last component selected opposes the single floor 3 ways without any particular intersection with single floor 4 ways in acceptable condition with interchange or crossing and made of pure semi-grained bitumen.

### 4.3 Correlation between different characteristic of the NR4

After the main component's selection procedure, the variables of the study can be projected on the two first main components.

The Figure 6 is the circle of correlation of the variables. It shows the relationships between all variables in this study. It can be interpreted as follows:

- Positively correlated variables are grouped together;
- Negatively correlated variables are positioned on opposite sides of the chart origin (opposite quadrants);
- The distance between the variables and the origin measures the quality of representation of the variables. Variables that are far from origin are well represented by the PCA.

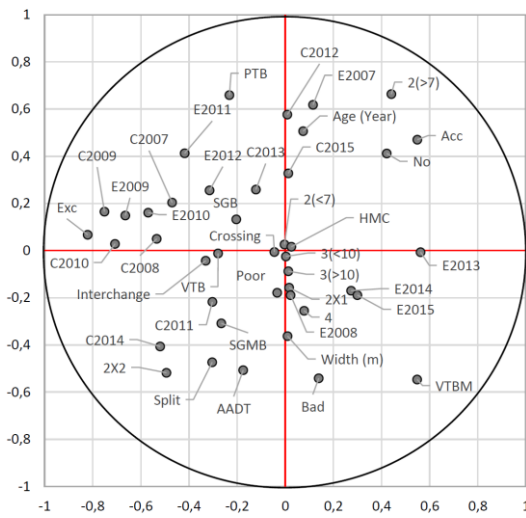


Fig. 6. Circle of correlations associated to the NR4

In this circle of correlations, the variables  $2(<7)$ , HMC, Crossing and  $3(<10)$ , for example, are positively correlated. This means that markings that are situated in the same areas have the same properties. However, markings that are correlated with the variable  $2 \times 2$  and those which are correlated with the variable  $2(>7)$  have opposite properties. The markings of center line C2009, C2007 and C20013 are positively correlated with the markings of emergency line E2011, E2012, E2010 and E2009 while they are negatively correlated to E2014, E2015, E2013 and E2008.

#### 4.4 Map of individuals and clustering

To segment the markings lines after the PCA, the next step is about to choose the number of clusters through the Agglomerative Hierarchical Clustering.

One solution would be to represent the jumps of inertia of the dendrogram which will make it possible to retain the optimal number of classes.

Indeed, depending on the height of the dendrogram section, the number of classes differs. In the first place, an analysis of the form of the dendrogram may give an indication of the number of classes to be retained. It is the jumps of inertia that allow this analysis.

Thus, after analyzing the inertial jumps obtained from the inertia graph, the dendrogram can be split into 2, 4, 7, 9 or 14 classes. To choose the optimal number among the five candidates, it is necessary to conduct a case-by-case study and choose the best one.

The selection criterion adapted for this study is the one of the qualities of the representation of individuals and variables according to the number of clusters selected. According to this criterion, individuals who tend to group together consider

that they belong to the same cluster. And so, the best representation is that which allows a good distinction of different classes.

After confronting the five cases, it is the 9-class score that stands out with the best quality representation. This partition in 9 classes is represented by the dendrogram of the figure 7 below.

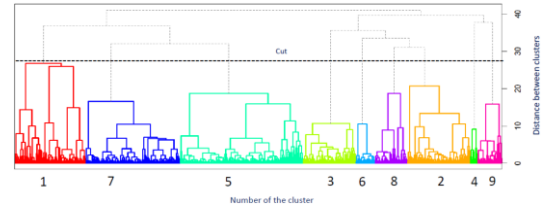


Fig. 7. Dendrogram associated to the segmented map of individuals

On the other hand, the projection of the individuals, and their 9 selected clusters, on the first factorial plane is also presented in figure 8. This figure makes it possible to better visualize the different groups of markings. It also allows them to associate the different characteristics given by the variables of the road infrastructure. Indeed, its reading of the map of markings is done by referring to the circle of correlations of the variables (figure 6).

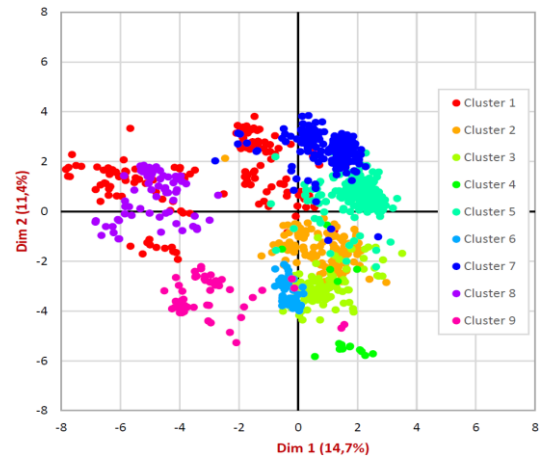


Fig. 8. Map of individuals associated to the NR4

This Chart of Figure 8 is known as the map of individuals (here markings). Its reading is facilitated by the circle of correlations (see figure 6) translating as a compass. According to the four quadrants, each marking, can be described by the corresponding variables in the same quadrant.

The chart (figure 8) particularly shows that the cluster 1 in red color is particularly split in two. And another observation can be made on the cluster 9 in pink. This cluster is the one of areas of the NR4: it is isolated.

## Impact of road infrastructure characteristics on road markings 7

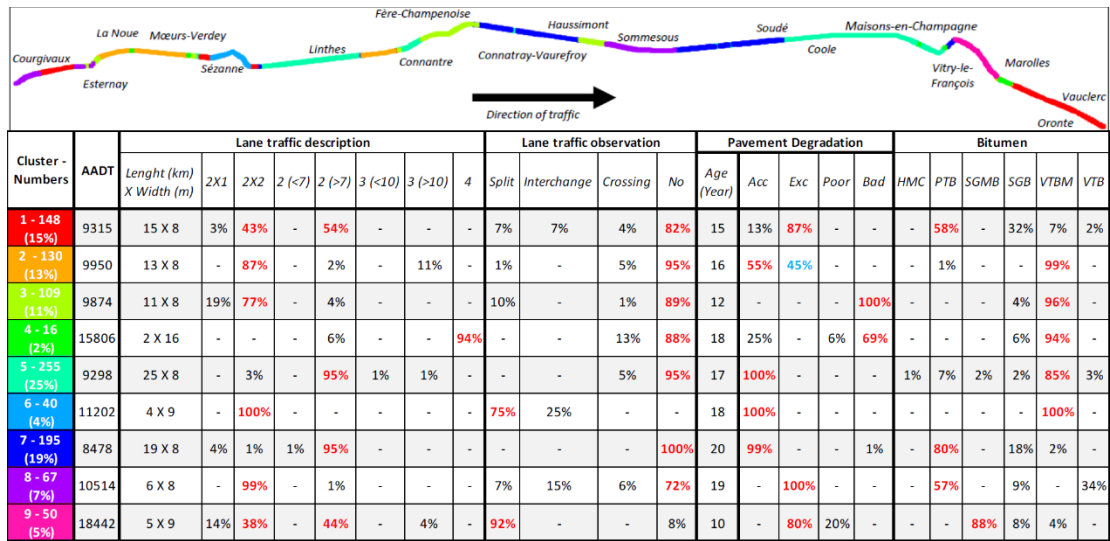


Fig. 9. The location of clusters on the RN4

### 4.5 Cluster projected on the NR4

In this classification in nine clusters, the cluster 1 (see figure 9) in red color is particularly split in two. This is due to the fact that the markings of this cluster are laid mainly on two types of roads: separate floor  $2 \times 2$  ways and single floor 2 ways.

It appears that the markings of the cluster 1 are posed on separate floor  $2 \times 2$  ways and single floor 2 ways in excellent condition. On these roads, the presence of a little elevation, exchanger and plan is noted and more than half (58%) are coated with pure thin bitumen. The markings on the center line deteriorated considerably in 2010, two years after being laid.

In cluster 2, 55% of markings have acceptable pavement condition and 45% are in excellent condition. The roadways are almost totally (99%) paved with very thin modified bitumen and are separate floor  $2 \times 2$  ways. In this cluster, there is the presence of plan and markings have evolved rather stable over time.

Cluster 3 includes only the markings whose pavements are in bad condition and it is essentially separate floor  $2 \times 2$  ways coated with very thin modified bitumen. The presence of altitude difference is quite important (10%).

A significant number of Cluster 4 markings have poor road conditions (69%). These include markings on single floor 4 ways coated with very thin modified bitumen. The crossing with a Plan is quite remarkable (13%). A rather peculiar evolution of the level of retro-reflection of the edge line, which had a significant increase between 2007 and 2008 then a considerable decline before stabilizing, is also noticed.

Cluster 5 markings have an acceptable pavement condition and are essentially laid on single floor 2 ways. This is the only cluster that contains a single floor 3 ways portion which size is little than 10.5 m.

Cluster 6, on the other hand, is the one of separate floor  $2 \times 2$  ways paved with very thin modified bitumen and in acceptable condition. There is a significant difference in altitude (75%) and heat exchanger (25%). In this cluster, the retro-reflection luminance level of the center line shows a remarkable peak in 2011 and for the emergency one it is in 2008 that a peak appears.

Cluster 7 groups together mainly single floor 2 ways coated with pure thin bitumen and in acceptable condition. It is the only cluster in which there is no crossing and its bitumen is the oldest of all and has the lowest daily traffic rate.

The group in which all the portions of roads are in excellent condition is the cluster 8. It contains a separate floor  $2 \times 2$  ways. Apart from the presence of elevation and plan in this cluster, a fairly large percentage of exchanger is noticed.

Finally, cluster 9 displays a significant percentage of elevation gain (92%). Single floor 2 ways and separate floor  $2 \times 2$  ways covered with modified semi-grained bitumen and in acceptable condition are encountered. The three types of markings in this cluster show a similar evolution.

In addition, a monitoring of the different lines of markings can be carried out according to the different years of the study.

### 4.6 Markings lines monitoring

From now, the median strip line is segmented according to the proposed clustering. As shown in the following figure 10, the follow-up of markings is deepened by observing each line of markings within the various clusters obtained.



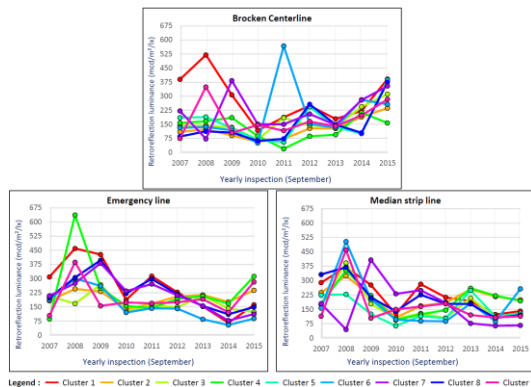


Fig. 10. Markings lines monitoring

The evolution of road markings is remarkable in some clusters. For example, in the cluster 1, the markings of the center line deteriorated considerably in 2010, two years after being laid. In cluster 4, a rather peculiar evolution of the level of retro-reflection of the edge line markings, which had a significant increase between 2007 and 2008 then a considerable decline before stabilizing, is also noticed. In cluster 6, the retro-reflection level of the center line markings has a remarkable peak in 2011 and for the emergency line markings it is in 2008 that a peak appears. In the cluster 9, the three types of markings lines show a similar evolution.

### 5. Comparison with previous study

To obtain strategic zones for the maintenance of road markings, Redondin (M. Redondin, *Approches de classifications à partir de données fortement censurées pour l'analyse de fiabilité et la définition de stratégies de maintenance - Application aux marquages routiers dans un contexte de véhicules autonomes 2018*) obtained 5 clusters through an AHC applied to the markings of the center line.

Obtaining groups of markings makes it possible to develop localized maintenance strategies by analyzing in particular the properties of the markings of these different groups. However, the AHC applied to center line markings, without considering certain infrastructure characteristics, does not give results covering the entire study portion of the RN4. In fact, the black area around Vitry le François (see figure 4), for example, is part of the unclassified zones. Also, it must be added that the previous study (M. Redondin, et al. 2017) is limited to the center line of markings in its AHC.

The first noticeable difference between the two figures (figure 4 and figure 9) concerns the unclassified areas in the case of center line markings only. On the other hand, taking into account all the characteristics of the road infrastructure has made it possible to classify

these unclassified areas. Also, it must be added that when the center line markings are analyzed alone, the groups obtained have a high degree of overlap in certain areas. This is not the case in the study of all road infrastructures where the different clusters are relatively well distinguished.

These differences in the number of clusters and areas of overlap between clusters bear witness to the impact of road infrastructure characteristics in the classification of horizontal road markings. This is a strong point of what has been accomplished in this study and is also more relevant than the case of the study of the center line only.

### 6. Conclusion

Road markings play an important role in the guidance of vehicles in general and future autonomous vehicles. This is especially true if we are interested to all road infrastructures that must be maintained to guarantee maximum safety for vehicular traffic.

The first results of this study show that considering all the characteristics of infrastructures makes it possible to segment road markings more effectively for strategic areas of maintenance. Indeed, the nine areas obtained in this study are better distinguishable (with less overlap) than those obtained without taking into account all the properties of road infrastructures. Another highlight of this work is that there are no more unclassified zones than what happens when we only consider horizontal road markings.

Moreover, this study made it possible to highlight a classification of all the road infrastructures contained in the data collected through the FAMD combined with an AHC methodology. This classification, compared to that of the center line markings, made it possible to detect the impact of road infrastructure properties on horizontal signage. Therefore, it would be interesting to analyze the life cycles of the markings of the center line, the emergency one and the edge one through a Weibull analysis.

In the end, it would be interesting to extend the present study by a panel type econometrics modeling to evaluate the level of impact of road infrastructure features on horizontal signaling.

### Acknowledgement

VEDECOM and IFSTTAR thank the DIR Est and the CEREMA for inspection data on the National Road 4.

## References

- Abboud, Nasser, and Brian Bowman. "Cost and longevity based scheduling of paint and thermoplastic striping." *Transportation Research Record : Journal of the Transportation Research Board* 1794 (2002): 55-62.
- AFNOR. "NF EN 1436+A1 - Road marking materials - Road marking performance for road users." Saint Denis La Plaine (France): AFNOR Editions, 2009 .
- Escofier, Brigitte, and Jérôme Pagès. *Analyses factorielles simples et multiples (Objectifs, méthodes et interprétation)*. Dunod, 2008.
- Karlis, D., G. Saporta, and A. Spinakis. "A simple rule for the selection of Principal Components." *Communication in Statistics - Theory and Application* 03, no. 32 (2003): 643-666.
- Lu, JJ. *Performance of traffic markings in cold regions*. University of Alaska Fairbanks - Report No. INE/TRC 95.03. 1995.
- Ministère de l'Écologie et de la Transition Énergétique. *l'ouverture des données du réseau routier national*. 2018. <https://www.ecologique-solidaire.gouv.fr/louverture-des-donnees-du-reseau-routier-national#e3>.
- Nextroad. *ECODYN mlpc® - Signing visibility*. 2019. <http://www.nextroad.com/> (accessed In migrated on Nextroad Website).
- Pagès, Jérôme. "Analyse Factorielle de Données Mixtes." *Revue de statistique appliquée* 52, no. 4 (2004): 93-111.
- Redondin, M., L. Bouillaut, D. Daucher, and N. Faul. "Temporal clustering of retroreflective marking." *European Safety and Reliability Conference* , 2017.
- Redondin, Maxime. *Approches de classifications à partir de données fortement censurées pour l'analyse de fiabilité et la définition de stratégies de maintenance - Application aux marquages routiers dans un contexte de véhicules autonomes*. Université de Paris-Est: Thèse de doctorat, 2018.
- Redondin, Maxime, Laurent Bouillaut, Allou Samè, Dimitri Duahcer, and Nadège Faul. "Alternative Weibull analysis for road markings : an EM approach." *ESREL 2018*. Trondheim (Norway), 2018.
- Revilloud, Marc, Dominique Gruyer, and Mohamed-Cherif Rahal. "A lane marker estimation method for improving lane detection." *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on IEEE*, 2016: 289-295.
- Revilloud, Marc, Dominique Gruyer, and Mohamed-Cherif Rahal. "A new multi-agent approach for lane detection and tracking." *Robotics and Automation (ICRA) 2016 IEEE International Conference* (2016): 3147-3153.
- Sitzabee, William E., Joseph E. Hummer, and William Rasdorf. "Pavement marking degradation modeling and analysis." *Journal of Infrastructure Systems* 15, no. 3 (2009): 190-199.
- Tufféry, S. *Data Mining and Statistics for Decision Making*. John Wiley & Sons, 2011.