



HAL
open science

Reconstruction of missing retroreflexion data according to yearly inspection of markings

Maxime Redondin, Nadège Faul, Laurent Bouillaut, Dimitri Daucher

► To cite this version:

Maxime Redondin, Nadège Faul, Laurent Bouillaut, Dimitri Daucher. Reconstruction of missing retroreflexion data according to yearly inspection of markings. MMR 2017 - 10th International Conference on Mathematical Methods in Reliability, Jul 2017, Grenoble, France. 8p. hal-01815501

HAL Id: hal-01815501

<https://hal.science/hal-01815501v1>

Submitted on 14 Jun 2018

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.

RECONSTRUCTION OF MISSING RETROREFLEXION DATA ACCORDING TO YEARLY INSPECTION OF MARKINGS

MAXIME REDONDIN, NADEGE FAUL
Institut VEDECOM
77 rue des chantiers, Versailles, 78000, France
email: *firstname.name@vedecom.fr*

LAURENT BOUILLAUT, DIMITRI DAUCHER
UPE Paris EST – IFSTTAR/COSYS
Marne-la-Vallée, 77447, France
email: *firstname.name@ifstar.fr*

The quality and reliability of road infrastructure and its equipment play a major role in road safety. This is especially true if we are interested by autonomous cars traffic able to read road markings. This kind of vehicles needs an accurate maintenance strategy to guarantee a road with marking perceptible for a human eye or an autonomous car. To simplify the study of a road, a solution based on an Agglomerative Hierarchical Clustering segments a road according to the retroreflection level in time. If the follow-up of the maintenance for markings doesn't exist then a maintenance detector could estimate laying date. However, this strategy needs regular inspection data. Currently, French roads are irregularly inspected once a year and missing data appear. Three options are confronted: accept missing data, estimate missing data thanks to a linear interpolation or an original approach based also an Agglomerative Hierarchical Clustering. The last possibility evaluates the most reliable estimations for missing data. This approach is a first step to analyze the useful life of markings with a Weibull analysis. The broken centerline of the French National Road 4 is considered to illustrate our approach.

Keywords: Road markings, Retroreflection, Agglomerative Hierarchical Clustering, Linear interpolation, Missing data.

1. Introduction

The quality and reliability of road infrastructure and its equipment play a major role in road safety. This is especially true if we are interested by autonomous cars traffic. This kind of vehicles needs an accurate maintenance strategy to guarantee a road with markings perceptible for a human eye or an autonomous car. Markings are subjected to degradation over time, mainly due to the traffic, environmental conditions. Measured in millicandela by square meter and by luminance ($\text{mcd}/\text{m}^2/\text{lx}$), the retroreflection level is the standard visibility measurement. According to AFNOR rules [1], a minimum threshold of $150 \text{ mcd}/\text{m}^2/\text{lx}$ has to be respected by all constructors. Only devices able to measure dynamically, like the Ecodyn, is considered.

A performance-based approach of the maintenance is chosen to evaluate the economic impact of replacing a marking. Preliminary questions must be solved: how each marking was laid? What kind of marking was chosen? What is the

current maintenance strategy? Are there some strategic areas? The literature supposes some conditions: a test desk is considered [2], the study starts 6 months after the lay date [3] or the kind of marking is known [4]. In all cases, a follow-up of the maintenance campaigns for markings and regular inspection campaigns are complete.

Similar to [5], the study considered in this paper is based on a specific road: broken centerline of the French National Road 4 (NR4) managed by the DIR Est [6]. Unfortunately, there isn't any follow-up of the maintenance campaigns for markings. As a result, marking retroreflection data is incomplete, thus installment dates of markings must be estimated. Restricted to areas systematically inspected (~18% of the road), a solution based on an Agglomerative Hierarchical Clustering (AHC) [7] to cluster all marking according to the geographic position and the retroreflection evolution is proposed to simplify the road into five clusters [8]. Each cluster represents a specific part of the road. This part admits its own retroreflection evolution and past maintenance actions can be deduced thanks to a maintenance detector based on the previous clustering. To extend this approach to the whole NR4, two solutions are currently possible: ignore incomplete data because the AHC accepts them or complete data thanks to a linear interpolation [9]. The first solution cannot be exploited because the typography of each cluster are very uncertain. The second was proposed to some conditions: a test desk was considered, the study started seven days after the lay date and monthly inspections were organized. These hypothesis aren't verified in the NR4 case.

A reconstruction based also on an AHC on each inspection was therefore proposed to try to overcome this problem. After a brief presentation of each reconstruction algorithm, their efficiency will be compared.

2. Reconstruction algorithms

2.1. "Do nothing" algorithm and mathematical formalism

An adapted mathematical formalism is chosen from now on. Let's suppose n markings. Each one is defined by $i = 1, \dots, n$. According to the current literature, the retroreflection level of a marking i at the inspection t is denoted $RL_t(i) > 0$. The retroreflection evolution between two inspections t and $t - 1$ is defined by:

$$\Delta RL_t(i) = RL_t(i) - RL_{t-1}(i) \quad (1)$$

If the marking i is not inspected at t , $RL_t(i) = NA$ and must be estimated. For all reconstruction algorithms, an estimated retroreflection level of a marking i at the inspection t is denoted $\widehat{RL}_t(i)$. In consequence, an estimated retroreflection evolution between two inspections t and $t - 1$ is $\widehat{\Delta RL}_t(i)$.

In France, all marking products must be tested during two years on a specific part of the NR2. The ASCQUER is charged to publish [10] a theoretical performance of each marking. The larger retroreflection level measured is 1140 mcd/m²/lx.

This value represents the maximum level of an estimator. The “do nothing” algorithm consists to accept all incomplete data i.e. $\widehat{RL}_t(i) = NA$.

2.2. Linear interpolation

The linear interpolation simply consists to model a line between two reference points. Applied to retroreflection data, this approach could estimate a missing data of a marking according to two reference measures [9]. Let's takes a marking i that admits at least one missing inspection at τ and let suppose the two near reference measures $t_a < t_b$. Then, the estimation is given by:

$$\widehat{RL}_\tau(i) = \begin{cases} RL_{t_a}(i) + \left(RL_{t_b}(i) - RL_{t_a}(i) \right) \frac{\tau - t_a}{t_b - t_a} & \text{if } t_a < \tau < t_b \\ RL_{\tau-1}(i) - \left(RL_{t_a+1}(i) - RL_{t_a}(i) \right) & \text{if } \tau < t_a \\ RL_{\tau+1}(i) + \left(RL_{t_b}(i) - RL_{t_b-1}(i) \right) & \text{if } \tau > t_b \end{cases} \quad (2)$$

If the retroreflection data at inspection $t_a < \tau < t_b$ is missing then $\widehat{RL}_\tau(i)$ is estimated by the basic linear interpolation formula. The two last cases estimate the retroreflection level of a marking i always inspected after the inspection t_a or always inspected before the inspection t_b . In the first case, if $t_a \neq 1$ and $t_a = \min\{t \in [1, T] / RL_t(i) \neq NA\}$ then $t_b = t_a + 1$. The inspection data $\tau < t_a$ is progressively estimated by (2). In the last case, if $t_b \neq T$ and $t_b = \max\{t \in [1, T] / RL_t(i) \neq NA\}$ then $t_a = t_b - 1$. The inspection data $\tau > t_b$ is progressively estimated by (2).

A drawback of this approach is that it can generate negative estimations but also bigger than 1140. If a marking admits an absurd estimator, then it cannot be considered in the study.

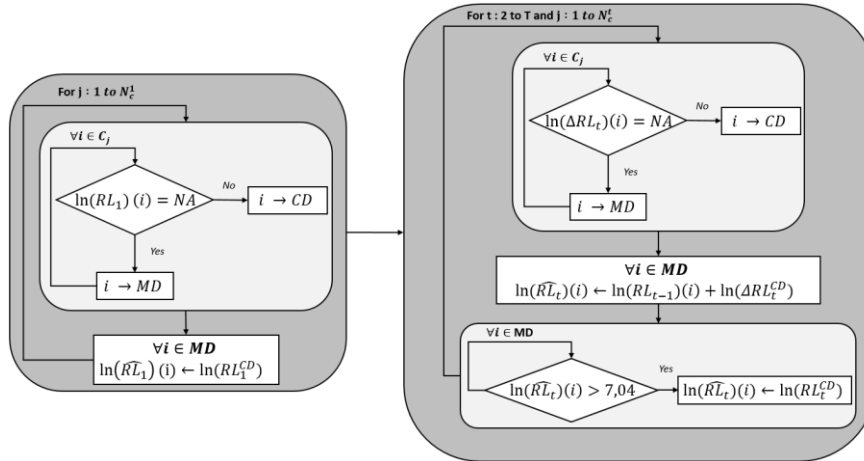


Figure 1. Presentation of the clustering approach for the reconstruction of retroreflective data

2.3. Clustering approach

The proposed clustering process, introduced in figure 1, is based on an AHC [9] and considers the retroreflection variation level between two inspections. The AHC consists to produce sequences of partitions of increasing heterogeneity between partitions into n clusters where each data is isolated and partition into one cluster which includes all the objects. The general form of the AHC is the following:

1. Each data is a cluster.
2. Distances between clusters are calculated.
3. The 2 closer clusters are merged into one.
4. Start again at step 2 until there is only one cluster containing all data.

The step 2 introduces a distance between two clusters. In the case of two data, the Euclidian distance d is considered. Else, the Ward distance (3) is chosen where X, Y denote clusters; x, y their average profiles and n_x, n_y their numbers. Let's consider N_C clusters, noted by C_j where $j = 1, \dots, N_C$. The number of markings classified in C_j and is denoted $\#C_j$.

$$d_W(X, Y) = \begin{cases} d(x, y) & \text{if } n_x = n_y \\ \left| \frac{1}{n_x} - \frac{1}{n_y} \right| & \text{if } n_x \neq n_y \\ d(x, y) & \text{else} \end{cases} \quad (3)$$

The whole monitoring data base is transformed by the natural logarithm. The process admits two steps: the initialization ($t=1$) and the iterations ($t>1$). Each cluster must admit at least one marking inspected:

$$N_C^t = \begin{cases} \max\{n \in \mathbb{N}^* / \forall j \in [1, n], \exists i \in C_j, \ln(RL_1)(i) \neq NA\} & \text{if } t = 1 \\ \max\{n \in \mathbb{N}^* / \forall j \in [1, n], \exists i \in C_j, \ln(\Delta RL_t)(i) \neq NA\} & \text{else} \end{cases} \quad (4)$$

The estimation is given by (5) $\forall t \in [1, T], \forall j \in [1, N_C^t], \forall i \in C_j$:

$$\ln(\overline{RL}_t)(i) = \begin{cases} \ln(RL_1^{CD}) & \text{if } t = 1 \\ \ln(RL_{t-1})(i) + \ln(\Delta RL_t^{CD}) & \text{else} \end{cases} \quad (5)$$

The initialization step rebuilds the first inspection. This inspection doesn't admit a variation of the retroreflection level. A data is a marking i described by its PR and $\ln(RL_1)(i)$. For each cluster C_j , all complete data are placed either in the data base CD (Complete Data) or in the data base MD (Missing Data). The \ln retroreflection mean level of CD (6) is the estimator of the \ln retroreflection level of each marking in MD at the first inspection.

$$\ln(RL_t^{CD}) = \frac{1}{\#CD} \sum_{i \in CD} \ln(RL_t)(i) \quad (6)$$

The iterative steps rebuilt next inspections. A data is a marking i described by its PR and $\Delta RL_t(i)$. Again, each cluster C_j are segmented by the data base the CD and MD. The retroreflection variation mean level of CD is used to estimate the \ln retroreflection level of a marking $i \in MD$:

$$\ln(\Delta RL_t^{CD}) = \frac{1}{\#CD} \sum_{i \in CD} \ln(\Delta RL_t(i)) \quad (7)$$

The logarithm transformation guarantees to not generate negative or null estimators. However, it is still possible to obtain the inequality (8) for a marking $i \in MD$. In this case, $\widehat{RL}_t(i)$ is replaced by $\ln(RL_t^{CD})$.

$$\exp(\ln(\widehat{RL}_t(i))) > 1140 \leftrightarrow \ln(\widehat{RL}_t(i)) > 7.04 \quad (8)$$

3. Comparison of algorithms

3.1. Reconstruction processes

Table 1. Monitoring of the reconstruction process realized by a linear interpolation

Level of reconstruction	Impossible	Yes	No	Reconstruction total	Observable total	Total
Numbers	106	754	181	935	777	1041
Frequency	10%	72%	17%	90%	75%	100%

The linear interpolation process abilities are detailed in table 1. 106 markings cannot be rebuilt because too few inspected. 754 markings benefits of at least one reconstruction. In fact, the monitoring of 935 markings, i.e. 90% of the road, is completed. However, 158 rebuilt markings present at least one absurd estimation. Finally, the linear interpolation can observe 777 markings, i.e. 75% of the NR4. For the comparison between reconstruction algorithms, this sample is considered from now.

Table 2. Monitoring of the reconstruction process realized by the clustering approach

Yearly inspection	2007	2008	2009	2010	2011	2012	2013	2014	2015
Numbers of incomplete data	419	428	174	23	472	386	156	193	11
Frequency against to the NR4	54%	55%	22%	3%	61%	50%	20%	25%	1%
Numbers of cluster used by the algorithm	16	10	31	677	13	7	32	29	723
Numbers of absurd estimation (> 1140 mcd/m ² /lx)	0	15	0	0	7	0	0	1	0

The clustering approach process is considered in table 2. Theoretically, this approach allows rebuilding all markings. The numbers of clusters depends on the distribution of incomplete data in the sample. However, it seems that few incomplete data needs many clusters (2015). We can note that only few absurd estimators exist and could be specifically studied. For example in 2008, all absurd estimations are ranged in the cluster 6 (18/34 markings inspected). The reconstruction process calculated for 15 markings:

$$\ln(\widehat{RL}_{2008})(i) = \ln(RL_{2007})(i) + \ln(\Delta RL_{2008}^{CD}) = 5 + 2.11 > 7.04 \quad (9)$$

Table 3. Comparison of estimators evaluated by a reconstruction algorithms

Markings disabled	15,1 in 2008	34 in 2010	59 in 2012	86 in 2014	98,2 in 2015	Average error
Reference values	148	66	177	364	248	
Linear interpolation	114	190	208	179	513	40,2
Clustering approach	158	80	282	359	182	11,6

Since 2007, five areas are systematically inspected. Five markings inspected in these areas are chosen and one inspection is disabled. The table 3 presents estimators calculated both by the linear interpolation and the clustering approach

and are compared to original values. It can be noticed that the clustering approach gives the smaller average error.

3.2. Monitoring of the retroreflection level

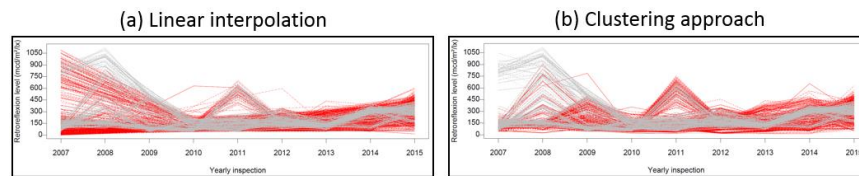


Figure 2. Comparison of retroreflection monitoring done by two different reconstruction algorithms. The figure 2 compares two monitoring of the retroreflection level between 2007 and 2015 of the NR4's skip center line: Fig 2a is rebuilt by the linear interpolation and Fig 2b by the clustering approach. Both indicated by red lines all markings with at least one estimated measure. The main difference is between 2007 and 2010 and concerns 224 markings placed between the PR 24 and 54.

Based on measures in 2009 and 2010, the linear interpolation proposes a linear degradation between 2007 and 2010. This information suggests a no-verifiable maintenance campaign in 2007 whereas the clustering approach completes an obvious maintenance campaign in 2008 (“grey pyramid”) and suggests a most important maintenance campaign in 2009 (“red pyramid”).

3.3. Temporal clustering

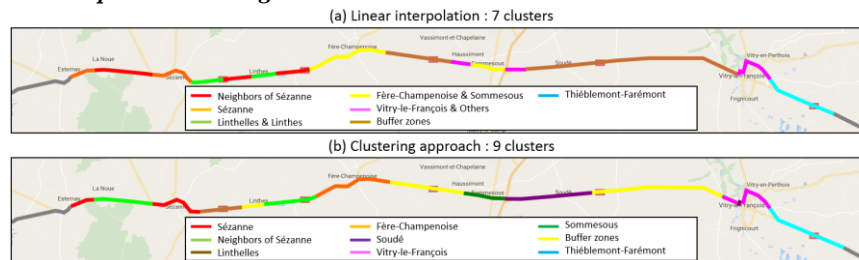


Figure 3. Comparison of temporal clustering done by two different reconstruction algorithms. According to a previous study, the AHC is able to segment the NR4 [8]. Each marking is defined by its PR and its nine inspection measures. The number of clusters is based on the semi-partial R^2 criterion which represents the loss of between-class inertial. It must be as small as possible. The “do nothing” algorithm estimates the admissible level of the semi-partial R^2 : 20% and eight clusters. For the same level, the AHC proposes seven clusters (Fig 3a) for the linear interpolation and nine clusters (Fig 3b) for the clustering approach. Seven consensus areas and two conflicts area can be extracted.

After Sézanne is one of two: 133 markings between the PR 24 and 38. The “do nothing” algorithm segments this area into five zones shared by three clusters. The linear interpolation proposes four zones shared by two clusters and the clustering approach is shared by three cluster. The last proposition is the most convenient.

The “do nothing” algorithm is the most segmented. On consensus areas, both the linear interpolation and the clustering approach are equivalent whereas, on conflict areas, the clustering approach is the more convenient. Finally, the clustering approach seems to propose the simplest segmentation of the NR4.

3.4. Maintenance detector

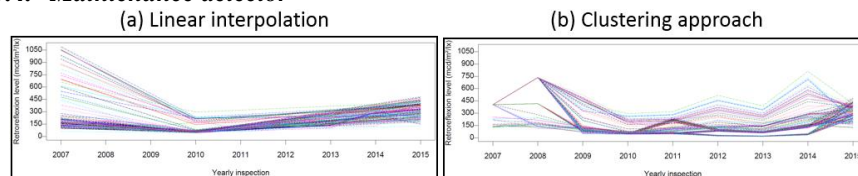


Figure 4. Comparison of retroreflection monitoring in After Sézanne area done by two different reconstruction algorithms: 133 markings

Finally, a maintenance detector, introduced in [8], was proposed to estimate, for a given cluster, using the previous temporal clustering and a given time, which marking has been replaced. Also based on the variation of retroreflection, this is the reason why a maintenance action in 2007 wasn’t detectable. In this paper, estimated maintenance dates are reported to consensus and conflicts areas. Due to missing data, the “do nothing” algorithm isn’t reliable.

In fact, the maintenance detector evaluates how many markings are replaced for a given area and a given time. The comparison of detectors shows two main divergences. First, the main difference between monitoring (Fig 2) is represented by the detector in the 2008 inspection. The second one is again “After Sézanne”. Indeed, the linear interpolation (Fig 4a) proposes a linear evolution in two phases. Between 2007 and 2010, markings have lost 200 mcd/m²/lx on average and since 2010 they linearly gain +50 mcd/m²/lx on average per year. This second evolution has no engineer explication and the detector approves, by default, at least one maintenance by year; that is quite unrealistic. The clustering approach (Fig 4b) proposes five local maintenance campaigns (2008, 2011, 2012, 2014 and 2015). This strategy and life cycles are more credible according to the current literature. The clustering approach seems therefore to be better than the linear approach.

4. Conclusion

The clustering approach proposed in this paper seems to be quite more adaptable to the NR4 situation since it allows to rebuild the whole NR4. However, if an inspection concerns few parts of the road then the algorithm produces few estimators for many missing data. Moreover, the clustering approach produces estimators with the smaller average error between estimators and reference values. This work is a first step to analyze the useful life of marking with a Weibull method.

5. Acknowledgments

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

6. References

- [1] AFNOR, *NF EN 1436+A1 - Road marking materials - Road marking performance for road users*, Saint Denis La Plaine (France): AFNOR Editions, 2009 .
- [2] G. Bahar, M. Masliah, T. Erwin, E. Tan and E. Hauer, «Pavement marking materials and markers: Real-world relationship between retroreflectivity and safety over time.,» *NCHRP web-only Document 92 - Transportation Research Board*, 2006.
- [3] V. Karwa and E. Donnell, «Predicting Pavement Marking Retroreflectivity Using Artificial Neural Networks Exploratory Analysis,» *Journal of Transportation Engineering*, vol. 137, n°12, pp. 91-103, 2011.
- [4] W. Sarasua, D. Clarke and W. Davis, «Evaluation of interstate pavement marking retroreflectivity,» Rep. No. FHWA-SC-03-01, Clemson Univ., Clemson, S.C., 2003.
- [5] J. Migletz, J. Graham, D. Harwood and K. Bauer, «Service Life of Durable Pavement Markings,» *Transportation Research Record*, n°11749, pp. 13-21, 2001.
- [6] Direction Interdépartementale des Routes de l'Est, «La signalisation horizontale - Politique d'entretien et de renouvellement et recommandations pour les concepteurs routiers,» Nancy, 2008.
- [7] S. Tufféry, *Data Mining and Statistics for Decision Making*, John Wiley & Sons, 2011.
- [8] M. Redondin, L. Bouillaut, D. Daucher and N. Faul, «Temporal clustering of retroreflective marking,» *European Safety and Reliability Conference, Portoroz, Slovenia, June 2017*.
- [9] S. Sathyanarayanan, V. Shankar and E. Donnell, "Pavement Marking Retroreflectivity Inspection Data - A Weibull Analysis," *Transportation Research Record : Journal of the Transportation Research Board*, vol. 08, no. 2055, p. 63–70, 2008.
- [10] ASCQUER, «PMA_NF2: Produits de marquage de chaussées - NF2,». Available: http://ascquer.fr/produit-certifie/NF/12/PMA_NF2/Produits-de-marquage-de-chaussees---NF2.
- [11] Y. Zhang and D. Wu, «Development of methodologies to predict service lives of pavement marking materials,» *Journal of the Transport Research*, vol. 45, n° 13, pp. 5-18, 2006.