XII DBMC
International Conference on Durability of Building Materials and Components
PORTO - PORTUGAL, April 12th-15th, 2011

Survival Analysis Methodology for Service Live Prediction and Building Maintenance

Carles Serrat [1]
Vicenç Gibert [2]

ABSTRACT

This paper deals with, on one hand, the introduction of survival analysis techniques for being used in building maintenance and, on the other hand, the application of this methodology for analyzing a large building stock in order to obtain information for maintenance strategies and/or prevention policies. In particular, in this contribution the description of the time to the event when the event of interest is some damage (or some level of degradation or extent) on the building façade is the main goal to reach.

For the time being, building follow-up is based on inspections. However, data coming from building inspections are always not completed, but censored, due to the fact that, at each inspection time, the event of interest is already happened, or not yet. In order to solve this problematic, the existing methodology for fields like medicine, biology, industrial engineering or event history analysis is adapted, and routines in S-PLUS for a numerical and graphical systematic analysis are implemented. Estimates for non-parametric durability and hazard functions are derived.

The possibilities of the proposed methodology will be illustrated with its application to the building façades in Hospitalet de Llobregat, the second most important city in population in Catalonia (Spain), where more than 14.000 buildings have been inspected. The analysis of the results allows technicians to detect different zones and levels of intervention to be applied in the city.

KEYWORDS

Censored data, Durability, Maintenance, Nonparametric estimator, Survival analysis.

[1] Escola Politècnica Superior d'Edificació de Barcelona of Universitat Politècnica de Catalunya-BarcelonaTECH, Barcelona, SPAIN, carles.serrat@upc.edu

[2] Escola Politècnica Superior d'Edificació de Barcelona of Universitat Politècnica de Catalunya-BarcelonaTECH, Barcelona, SPAIN, vicenc.gibert@upc.edu

1 INTRODUCTION

The management of building maintenance is increasingly an important topic around the world, especially in recent decades, with the incorporation of tools to automate the tracking information and potential follow-up. From the social point of view it is illogical to leave without controlling the building stock. Moreover, if we take into account the high cost that the non-maintenance of buildings in appropriate good service conditions represents.

The primary objective of the study presented in this paper is that of establishing building maintenance policies based on maintenance criteria from which director plans are created that allow defining the which way maintenance actions are affected the built assets; no matter if the actions are at macro or micro urban level. This will allow establishing more consistent decisions in accordance with the findings of analytical reliable studies.

Throughout this process an important issue must not be overlooked: the methodology should enable generating a base of information and results on the historical evolution of the facades of buildings and their most relevant elements. This fact underlines the crucial role of future interventions and actions to be implemented with a scientific supported base. The sample size in this study was sufficiently large to contain all data needed to achieve the proposed targets to analyze the results and assess its impact without necessarily having to apply not provided corrective actions (what saves the corresponding extra costs that they represent).

2 STATE OF THE ART

In respect to buildings, the use of survival analysis techniques based on the statistical approach of a large sample size does not exist in the literature. First contributions dealing with building or civil maintenance issues appeared in the nineties and they are basically devoted to the prediction of the service life of materials of construction [Frohnsdorff 1996].

One of the first statistical approaches is the one proposed by Flourentzou *et al.* [2000]. They introduce MEDIC as a method for predicting residual service life and refurbishment investment budgets. Predictions are based on the combination of prior probability coming from the experience from a large number of previous investigations/refurbishments and the current state of the subject (building) under study. The goal is to obtain a conditional probability of changing (to worse) between states, given the current state. However a drawback of this methodology is the fact that the deterioration process is described by stratifying it in four states, what is impossible to determine if we do not know the (a priori) true distribution of the time to deterioration. The relevance of having information about the survival functions of buildings or building elements have already pointed out by Buergel-Goodwin *et al.* [2005].

A different approach is found in Gaspar and de Brito [2003 & 2005]. The authors propose identification techniques for durability factors in order to predict service life of buildings or civil infrastructures. The same authors remark in the convenience of working with field data recollection techniques in order to assess the degradation level of a building element (in a facade, for example) in real life service conditions, in order to capture the effect on the deterioration process of the full complexity of the environmental context. In their work the estimation process at each time is based on an overall degradation level defined as a weighted average of specific degradation levels in several points in the moment of inspection, and performing a less squares cubic regression. It is obvious that this approach assumes that the time to event coincides with the inspection time; this is an unfeasible assumption. In fact, the resulting estimates for the deteriorating times are positively biased in an unknown quantity.

European guidelines for life time design and management of buildings can be found, among others, in the Project Cluster Lifetime by Sarja *et al.* [2005]. On the other hand, North American standards and advanced methodologies for predicting service life of building materials and components can be found, for instance, in the references of the Institute for Research in Construction (Montreal, Canada) [Lacasse 2008]. In all these cases, guidelines represent relevant information in order to define the severity of the injuries, the states in the initiation and progression of the deteriorating process and the transitions between them.

3 INSPECTION AND ANALYSIS METHODOLOGIES

In what follows the methodology that has been used for the inspection process is introduced as well as the survival analysis techniques used for the data analysis.

3.1 Inspection Methodology

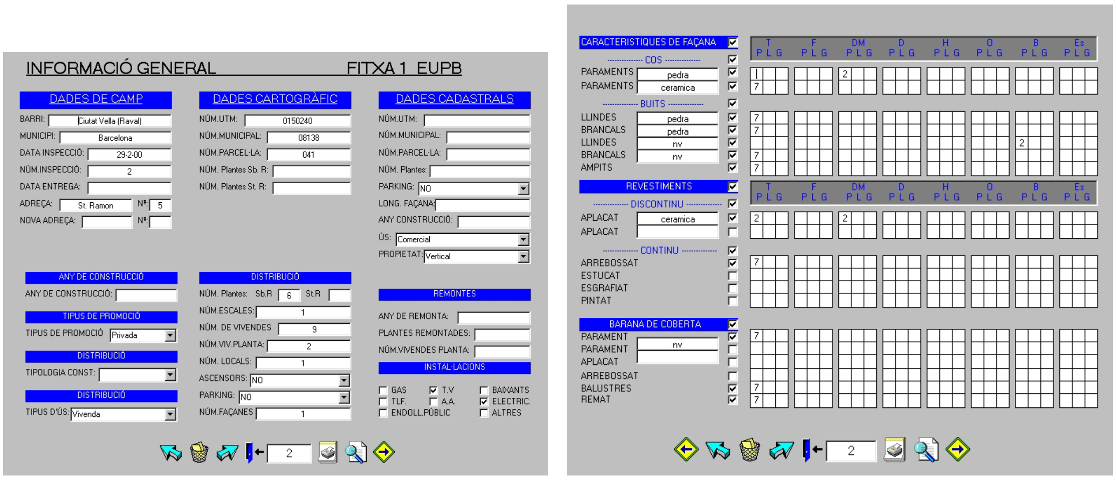
The systematic inspection of façade elements is aimed at detecting risk factors for façades in order to later evaluate their condition and likely time to repair and refurbishment. To develop an inspection methodology and obtain homogeneous results for further analysis, criteria based on a standardization of the most significant indicators were established. In this sense a classification of buildings, building elements that make up the façades and their materials was completed, as well as a cataloguing of defects that may affect the façades, the severity of the defects and, finally, the extent, as follows:

- The characterization of the buildings was carried out by the location of the building, year of construction, construction typology, and so on.
- The characterization of the constructive elements of the façades was done by subdividing the body into two parts: corps of the façade and corps outgoing. Within them the elements that the façade may contain have been identified. For example, in the case of the body of the façade we find cladding, voids, coatings and deck railings, and for example, within the voids we find lintels, jambs and parapets. Finally, we have defined the materials that can be used for each element.
- The definition of defect has been established to include the most common defects that can affect façade elements. In particular, the following eight defects have been considered: breach, fissure, material degradation, deformation, dampness, oxidation, desunity (incompatibility?) and detachment.
- The classification of the severity of defects has been established in terms of the action that is necessary to remove the risk that it may represent for people. After taking into account the term of the corresponding preventative action the following apply: 0 = “low severity”, 1 = “intermediate severity” and 2 = “high severity”.
- The definition of the extension of the defects has been established according to the proportion of the inspected element that it is affected by the defect: punctual (when the defect affects an element less than a 25%), local (affects element between 25% and 50%) and general (when the affected zone is greater than a 50% of the element).

The model that systematizes the inspection process allows encoding and entering data in a computer application for the assessment and the management of the results, with an interface similar to the field sheets used in data collection. Fig. 1 shows two screens available from the application.

By using this software application, several descriptive indicators can be obtained, referred to the inspected sample that the user wishes to consult for the analysis. For example, the software allows obtaining, among others, the following descriptors:

- List of maximum severities affecting the façades of the sample.
- Location of elements affected by defects of a certain severity.
- Number of detected defects by specific type of construction element, for a particular choice of type of defect and severity.



a) b)

Figure 1. Screens that allow a) the introduction of parameters of the inspected buildings and b) definition of materials for façade building elements and severity and extent of each type of defect.

3.2 Analysis Methodology

In this subsection, basic definitions of durability, the censorship mechanism issue and the non-parametric approach are introduced that were subsequently used for statistical analysis of data. Among many possible references, the primary reference on reliability and survival analysis used in this study was the work by Klein and Moeschberger [1997].

3.2.1 Fundamental Concepts in Survival Analysis

Three basic concepts in durability analysis (survival in the bio-medical field, reliability in the industrial-technologic field or event history analysis in the social field) are the failure, the durability function and the hazard function, that are defined as follows.

Definition 1: *failure* is an event of interest in which we are concerned about.

According to the International Organization for Standardization (ISO), quality is the set of properties and characteristics of a product that enable it to satisfy the requirements for whom it has been defined. In this sense, let T be the time from the beginning of the follow-up (time zero) until the failure (the event of interest) happens. The time T is a random variable and of interest in its distribution, such that in order to compute probabilities of occurrence under its density $f(t)$ or to obtain statistics like, for instance, the quantiles of the distribution. This will allow us to estimate the time until a proportion of damaged buildings in the population or, in the reverse sense, the proportion of damaged buildings at certain time for a particular defect. In the maintenance setup, time zero will mean the date that the building is built and by failure it is understood the successive grades of severity defects or successive grades of the extent of defects.

Definition 2: The *durability function* for the random variable T, $R(t)$, is the complement to one of the distribution function of T:

$$R(t) = 1 - F(t) = 1 - P(T \le t) = P(T > t) \tag{1}$$

Definition 3: The *mean hazard rate* in the interval $[t_1, t_2]$ is defined as

$$h(t_1, t_2) = \frac{R(t_1) - R(t_2)}{(t_2 - t_1) R(t_1)}, \tag{2}$$

where the durability, $R(t_1)$, is the proportion of survivors at the moment t_1 and $\left[R(t_1)-R(t_2)\right]/R(t_1)$ represents the proportion of elements that, have not yet failed at time t_1, but fail in the interval of time $[t_1,t_2]$. When t_2 tends to t_1 the *(instantaneous) hazard function* at time t_1 is given as: $h(t_1)=\lim_{t_2\to t_1} h(t_1,t_2)$. It can be shown that $h(t)=f(t)/R(t)$ and, in this sense, the hazard function measures the propensity to failure of the elements or statistical subjects as a function of time. If the cumulative hazard function, $H(t)$, is defined as $H(t)=\int_0^t h(s)\,\mathrm{ds}$, then the relationship between the durability function and the cumulative hazard function is:

$$R(t)=e^{-H(t)} \tag{3}$$

3.2.2 Censorship Mechanism

If *T* is the true time until failure occurs and C_1 denotes the time of an inspection event, then at time C_1 there are two possibilities: that failure has already occurred and thus $T<C_1$, or; failure has yet to occur and $T>C_1$. So, as a consequence, for each object of interest on which an inspection is performed the exact time at which *T*, is observed is not likely possible, hence only a part of the information is obtained upon inspection.

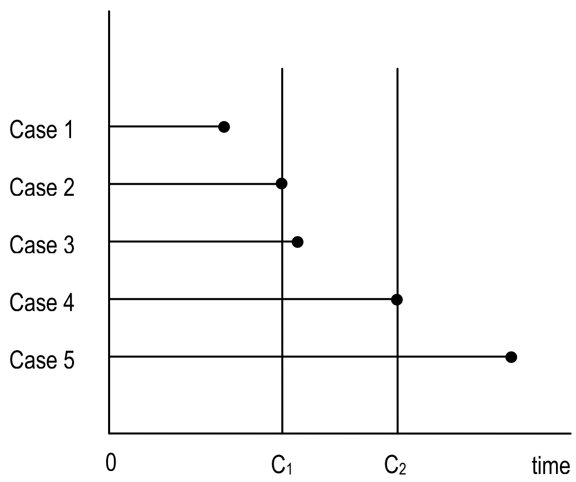


Figure 2. Illustration of the censorship mechanism: a) At time C_1, case 1 is left censored, case 2 is exact observation and cases 3, 4 and 5 are right censored observations. b) With a second inspection time C_2, case 1 is left censored, cases 2 and 4 are exact observations, case 3 is interval censored and case 5 is right censored.

In fact, if $T<C_1$, we only know that the true value of T is in the interval $(0,C_1)$, in this case the variable T is *left censored* and denote the observation by C_1^-. Otherwise, if $T>C_1$, $P\left(T\in(C_1,\infty)\right)=1$ and that T is *right censored* and denote the observation by C_1^+. This is usual approach taken to categorise failure data in "durability" studies, and in this instance all data are either left or right censored. This is referred to as *current status data*. Note that if there is a second inspection time, C_2, with $C_1<C_2$, then the observed data would be left, right or interval censored as illustrated in Fig. 2. . In the current study only one inspection time was considered and so only left or right censored observations are dealt with.

Concerning to the censorship mechanism it is important to assume a non-informative hypothesis, which is that the inspection time *C* is independent on the random variable of interest *T*. In general, when *C* is defined by design the non-informative hypothesis can be assumed.

3.2.3 A Non-Parametric Approach

Since there are no references on the distributions of failure times for the different defects and elements it follows that the durability function and the hazard function should be estimated non-parametrically. That is, estimates were only based on the data and no hypothetical (and non-testable) distributions were assumed for the unknown density f of *T*. Kaplan and Meier [1958] proposed a non-parametric estimator, called product limit estimator, for the durability function in the case of right censored data. Turnbull [1976] extended the Kaplan-Meier estimator to the case of interval censored data. The Turnbull's estimator is an iterative algorithm that maximizes the non-parametric likelihood function

$$L = \prod_{i \in O} \left(F(o_i) - F(o_i^-) \right) \prod_{i \in R} \left(1 - F(r_i) \right) \prod_{i \in L} F(l_i) \prod_{i \in I} \left(F(r_i) - F(l_i) \right), \tag{4}$$

where *O*, *R*, *L* and *I* are, respectively, the subsets of exact, right-censored, left-censored and interval-censored observations. Turnbull's algorithm identifies the intervals where probabilistic mass can be estimated and the resulting durability and hazard functions derived.

3.2.4 Implementation of the Analysis Methodology

We have developed all the durability analysis methodology; a library containing routines and procedures for obtaining estimates for failure probabilities, durability and hazard functions and summary statistics tables, have been implemented in S-PLUS® (Insightful®).

4 ILLUSTRATION

4.1 Scope of the Project

The project and methodology presented in the previous section have been applied in several cities and locations, not only nationally in Spain but also internationally. Table 1 shows the location of inspections and the number of inspected façades.

Table 1. Scope of the project to date

Country	*City*	*Location*	*Number of façades inspected*
Spain	Barcelona	Ciutat Vella	2631
		L'Eixample	2736
	L'Hospitalet de LLobregat	Whole city	13193
France	Toulouse	Old Town	80
México	México DF	Casco Antiguo	525
Chile	Valparaíso	Zona de los Cerros	396
	Santiago de Chile	Casco Antiguo	1403

As an illustration of these techniques some results are presented of the software application to the inspection of L'Hospitalet de Llobregat (hereafter L'Hospitalet) in the metropolitan area of Barcelona.

4.2 About L'Hospitalet de Llobregat and the Dataset

L'Hospitalet has 266,973 inhabitants (Institute of Statistics of the city, 31 December, 2008). In Catalonia, it is the city having the second largest population and is among the twenty most populated cities in Spain. The municipality covers an area of 12.5 square kilometres between the towns of Barcelona, Esplugues, Cornellà and El Prat. Administratively L'Hospitalet is divided in 7 districts and 12 sectors.

4.3 Univariate and Multivariate Analyses by Severity and by Extent

Univariate analyses were first performed on failure times for all the defects according to the severity and extent criteria. Thereafter, multivariate analyses were performed. In this paper only the univariate analysis is presented.

With the aim of illustrating the capabilities of the software system and the merit of our proposal, a small sample of graphical and numerical results taken from a severity analysis is presented in this section. The durability function for jambs for which defects are characterised by the presence of fissures along the jamb is given in Fig. 3 and Table 2. Among other conclusions, one can determine that after 10 years, 5.2% of jambs will suffer a punctual fissure of low severity and 1.6% and 0.1% of jambs punctual fissures of intermediate and high severity, respectively.

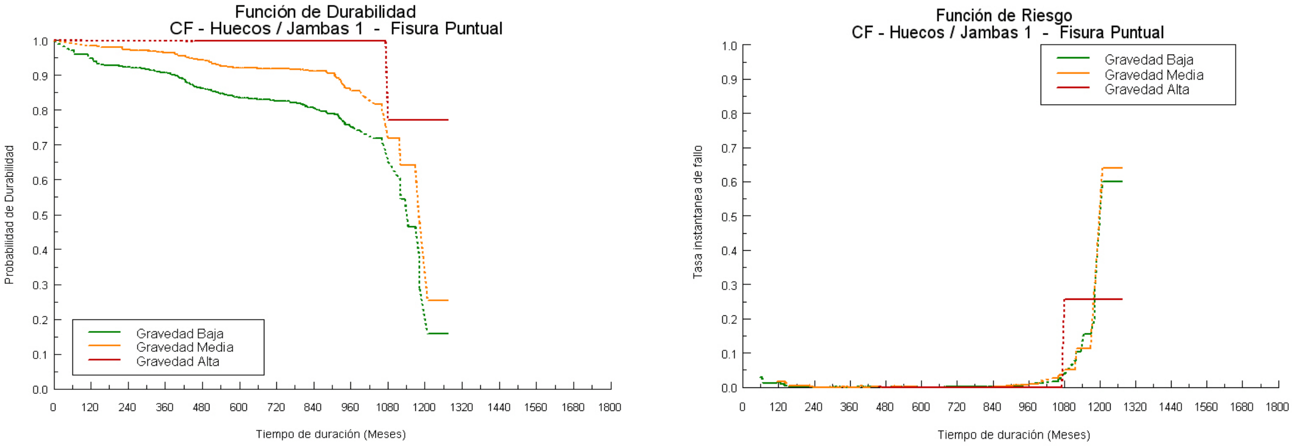


Figure 3. Durability and hazard functions for punctual fissure on the jambs (by severity level)

Table 2. Numerical results of the graphical analysis in Fig 3.

Number at risk	% of censoring		Estimated % of injuries				Years until a % of cumulative injuries				Cumulative % of injuries across the years			
			Initial		Final									
	Left	Right	Years	%	Years	%	10%	25%	50%	75%	10	25	50	100
2960	12.4	87.6	4.7	2.8	106.1	84.3	32.6	80.7	95.1	99.2	5.2	8.3	16.4	79.8
2960	5.3	94.7	9.5	1.6	106.1	74.7	75.7	89.5	98.4	--	1.6	3.0	7.9	66.9
2960	0.1	99.9	37.8	0.2	106.1	22.8	89.7	--	--	--	0.1	0.2	0.2	22.8

5 DISCUSSION AND FUTURE WORK

We have introduced a proposal for the use of "survival techniques" for building maintenance. As a probabilistic technique it allows estimating the failure distributions for all defects which may be of interest to a maintenance manager. Moreover, those characteristics which play (or not) a significant role in the product deterioration process can be tested. All this knowledge now allows an understanding of how the elements of the building envelope system functions, as well as how failures should be anticipated in service conditions as a consequence of aging, and especially how if indeed failures are anticipated, what preventive actions can be taken in-service. As well, it permits a viewpoint on what decisions need be taken at the design stage to help reduce anticipated defects and thus ensure the long-term performance (durability) of the building envelope.

The analysis platform that was developed and implemented becomes the basis for a continuous analysis in the future that will allow not only the design of maintenance strategies and prevention policies, these being based on economic and safety criteria, but also permit on-line decisions to be made.

We would like to provide information on two aspects for future development on which we are now working. Firstly, development of programming to achieve a more accurate inspection schedule. This will allow improvement of the quality of the data by reducing the proportion of censoring required or replacing left or right censoring by interval censoring, the latter method being the preferred approach. Secondly, the interest of adjusting the resulting distributions to the covariates of interest in the sense we mentioned in Section 4.3, when we referred to the multivariate analysis. In many cases, the failure distribution will be sensitive to building conditions. It is the belief of the authors that both these development items will make the building maintenance strategy more efficient.

ACKNOWLEDGMENTS

This work has been partially supported by the School of Building Construction of Barcelona (EPSEB-UPC), by the Building Laboratory at EPSEB and by the Spanish Ministry of Science and Innovation under grant MTM2008-06747-C02-01. Special thanks are due to the City Hall and to the inhabitants of L'Hospitalet de Llobregat for their collaboration in data collection, to the authors of several Graduation Final Projects for their contribution to this work, and to the members of the IEMAE for their valuable comments and suggestions.

REFERENCES

Buergel-Goodwin E., Ferrara C. & Bradley PE. 2005, “Survival functions of buildings and building elements”, in Proceedings of the 10DBMC International Conference On Durability of Building Materials and Elements. Lyon, France.

Flourentzou F., Brandt C. & Wetzel C. 2000, “MEDIC, A method to predicting residual service life and refurbishment investment budgets”, *Energy and Building* **31,** 167-170.

Frohnsdorff GJC. 1996, “Predicting the Service Lives of Materials of Construction”, in Proceedings of the Fourth Materials Engineering Conference. Chong KP. (ed), 38–53, American Society of Civil Engineers (ASCE), Washington, DC, USA.

Gaspar PL. & de Brito J. 2005, “Assessment of the overall degradation level of an element, based on field data” in Proceedings of the 10DBMC International Conference On Durability of Building Materials and Elements. Lyon, France.

Gaspar PL. & de Brito J. 2003, “Service Life Prediction: Identifying Independent Durability Factors”, in Proceedings of the Second International Symposium for Integrated Lift-time Engineering of Buildings and Civil Infrastructures. 319–324, ILCDES Kuopio, Helsinki, Finland.

Kaplan EL. & Meier P. 1958, “Nonparametric estimation from incomplete observations”, *Journal of the American Statistical Association*, **53**, 457-481.

Klein JP. & Moeschberger ML. 1997, *Survival Analysis. Techniques for Censored and Truncated Data.* Springer, New York, USA.

Lacasse MA. 2008, “Advances in service life prediction - an overview of durability and methods of service life prediction for non-structural building components” in Proceedings of the Annual Australasian Corrosion Association Conference. Wellington, New Zealand.

Sarja A., Bamforth, Caccavelli and Chevalier JL. 2005, *European Guide for life time design and management of civil infrastructures and buildings*, Project Cluster Lifetime

Turnbull BW. 1976, “The empirical distribution function with arbitrarily grouped, censored and truncated data”, *Journal of the Royal Statistical Society, Series B,* **38** (3), 290-295.