UNCERTAINTIES, QUALITY OF INFORMATION AND EFFICIENCY IN MANAGEMENT OF SEWER ASSETS

S.M. Elachachi[1], D. Breysse[2] and E. Vasconcelos[3]

ABSTRACT

Decisions on sewer management have large, long-lasting consequences. These decisions have to be taken under uncertainty. Managers of sewer system are also faced with their infrastructure system ageing. A statistical and decision (using multicriteria analysis) model has been developed and used to improve strategies of IMR interventions (Inspection, Maintenance and Rehabilitation). For estimating the asset performances, an approach based on Relative Risk analysis was developed. It enables to sort out prioritary sewer sections before any decision, accounting for the specific requirements of the manager regarding the system itself or its urban or socio-economical environment. This contribution focuses on the indirect sensitivity of decisions (accuracy in final decision) to uncertainties in external inputs and/or imperfect knowledge, based on a definition of "the quality" of the given information. A set of parametric studies is performed to determine the effect of lack of knowledge or imperfect knowledge on risk estimation.

KEY WORDS

Management, sewer network, performance, inspection strategy, uncertainty

INTRODUCTION

The general requirements of the society for quality services considering the best economic conditions with the respect to the environment and the new constraints induced by the sustainable development, encourages the managers of technical systems to develop a coherent and integrated policy of management. The control of time effects (ageing) on the components of these inheritances is also a major economic stake insofar as the budgets of maintenance mobilize a significant fraction of the resources of the public or private managers. Decisions on sewer management have large, long-lasting consequences and these decisions have to be taken under uncertainty. The investment required to keep an acceptable performance of the national network in France is evaluated to one billion euros per year for a 250.000 km asset (Le Gauffre et al. 2004).

1 Professor, LM2SC, department of civil engineering, U.S.T.O, B.P 1505, Oran, Algeria, and C.D.G.A, University of Bordeaux 1, Av des Facultés, 33405, Talence, France, Phone 33-5-40006588, FAX 33-5-40003113, sm.elachachi@cdga.u-bordeaux1.fr

2 Professor, C.D.G.A, University of Bordeaux 1, Av des Facultés, 33405, Talence, France, Phone 33-5-40008840, FAX 33-5-40003113, d.breysse@cdga.u-bordeaux1.fr

3 Research Engineer (PhD), LMCU, Bd Carnot, Nord 59000, Lille, France, Phone 33-3-20213915, FAX 33-3-20216078, evasconcelos@cudl-lille.fr

Information about the structural, hydraulic and tightness conditions of sewer networks is uncertain due to the fact that this underground infrastructure system is buried and difficult to access. Therefore, estimating the sewer network's performances is problematic and the investments can involve considerable "risk" when decisions are taken under imperfect knowledge.

This contribution focuses on the sensitivity of decisions (confidence in final decision) to uncertainties in external inputs and/or imperfect knowledge, from a definition of "the quality" of any given information. Among the chief sources of imperfect knowledge are:

- The fact that the asset parameters themselves (diameter, depth, material,…) are only known on a part of the whole asset,
- The limited number of inspections (assessment) which gives a measure of the condition of the sewer's section,
- The fact that information is often decayed from a high-order quantitative scale (continuous) to a lower qualitative one (for example: light, medium, and heavy).

This contribution aims to show the effect of this "imperfect knowledge" on the performances and related to:

- "crude" knowledge (distance between the true behaviour and what is seen or known),
- "rebuilt" knowledge (distance between the true performance and that estimated while using the confusion matrices).

ORGANIZATIONAL SCHEME

Figure 1 shows through a systemic diagram the sequentiality and interdependence relations (Vasconcelos 2005) between the various aspects related to a coherent management and being able to be considered as optimized (by some aspects).

Information useful for the patrimonial management of a network, which one can obtain starting from a geographical information system (GIS), is stored in a data base. These data are classified in two groups:

- Parameters: characterizing the sewers' section (e.g.: section's diameter, laying time, material, depth...). They are invariant with time except when forced by external interventions,
- Anomalies: evolutionary data indicating the condition of a section (ex: defects or defaults like longitudinal or transverse fractures, total or partial obstruction, localised collapse...) and included in the form of statement histories.

The anomalies, accessible via inspections (visual for the large diameters and CCTV for the nonvisitable pipes), make it possible to define and quantify the default indicators (DI). These default indicators allow to determine the deficiency indicators (DIi), in combination with the parameters defined in agreement with European regulation EN 13508-2 (EN 13508-2, 2003).

The crossing of the deficiency indicators and the vulnerability (V) of the medium at the position of the network elements, determines the Performance indicators (PIi) for a given section. By taking into account the priorities stated by the managers, the different (PIi) are balanced to obtain a performance grade for each section. This performance grade can be incorporated through statistical models to estimate the performance of the entire network.

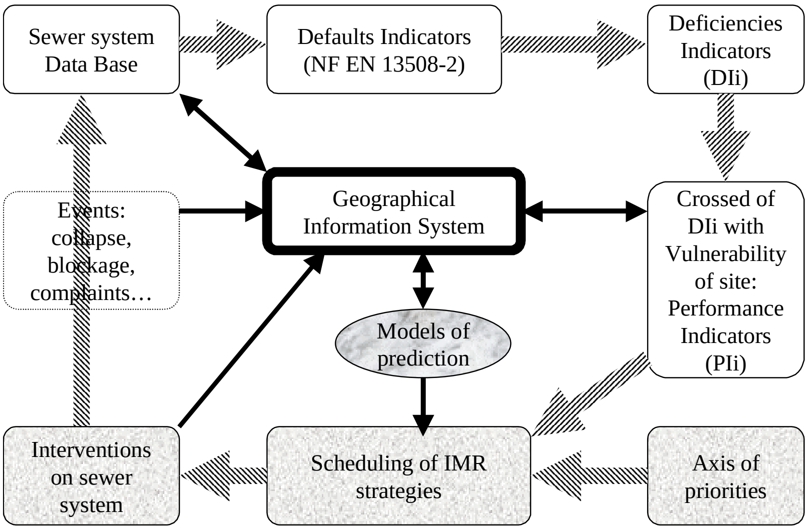


Figure 1: Organizational scheme (Vasconcelos 2005)

A statistical and decision (using multicriteria analysis) model has been developed and used to improve IMR strategies (Inspection, Maintenance and Rehabilitation), (Elachachi and Breysse, 2005). It is able to reproduce all the functionalities of a sewer network by integrating: (a) the ageing of the sewer system, (b) the practice of the manager regarding IMR operations (cleansing, performing CCTV inspections, gathering measurements on the system or on its efficiency, and rehabilitation...). The developed numerical support tool has been designed such as to show the manager how more efficient practices can provide him:

- a better knowledge of the actual performances at various scales, from the local scale (pipe) to the global scale (whole system), and on their evolution with time,
- a performance modelling, which enables to forecast future evolution and to take decisions with less uncertainty.

However, any future prediction requires a sufficiently and complete control of Data acquisition.

The quality of the information collected and to be fed to the Database is an essential task. Its exhaustiveness, representativeness, accuracy, reliability and homogeneity will condition the relevance of the predictive models, therefore the quality of the estimates of the future evolution of the network.

It is not easy to evaluate all the aspects of the quality of each information contained in a Database. That can be carried out by specifying some the circumstances of acquisition and which can be the cause (control, complaint, diagnosis), the means implemented (instrumentation), the person in charge for acquisition (experiment...) and also the context of acquisition (normal situation, crisis...).

Being given the expensive character (in time and money) of Data acquisition, control, storage in an exploitable form and Data updating, this acquisition has to be justified by operational objectives in coherence with the strategies followed by the manager.

RELATIVE RISK BASED ANALYSIS

For estimating the performances of the sewers, an approach based on Relative Risk analysis was developed.

The level of service of each section can not be checked each year on the entire of asset (budgetary limits), which results in having to estimate its deficiency conditions from available information (old inspections) and with the support of the evolution models.

While estimating the deficiency conditions, two cases arise:

- no inspection was ever carried out on the section,
- an inspection was already carried out on the section providing the deficiency condition observed {1 (best condition), 2, 3, 4 (worst condition)} and at a given inspection time. In this case, the condition observed in the past is used to consider the current condition.

Due to the fact that the inspection actions are generally limited and correspond to a small percentage of the network (for example for sewers networks, few percent per year in average France), the concept of relative risks (RR) for each modality is an interesting tool to estimate the risk of deficiency for any section.

RELATIVE RISK OF A DEFICIENCY

The relative risk is a number, without unit, ranging between 0 and the infinity. More (RR) it is distant from 1 (superior or inferior), more association exists between the occurred deficiency and the presence of the studied factor (modality) is strong. The interpretation which one can make of the concept of relative risk is that, for any pair of (deficiency, factor):

If (RR) = 1, that means that there is no detection of excess of risk. There is no relation shown between the deficiency and the exposure to the studied factor;

If (RR) > 1, that means that there is an excess of risk. There is thus a relation between the exposure to the studied factor the occurrence of deficiency. The factor can be regarded as factor of risk. One concludes by affirming that if a section is exposed, the risk to observe a dysfunction is (RR) times higher than if it were not exposed.

If (RR) < 1, that means that there is a lower risk of deficiency. This modality can be regarded as a protective factor.

The (RR) values are obtained by a statistical analysis of correlations between factors and occurrences of deficiencies. This statistical analysis is performed "real time", each new information updating the previously estimated values.

These (RR) values are used in a second step to provide estimates of actual condition from that observed in the past.

The quality of the estimate depends on the quality of available information (frequency of the inspections) and of the quality of the models used for the estimate (model of RR). The models of (RR) can be more accurate or of comparatively less accuracy for mainly two reasons:

- The sample size (number of occurrences of deficiencies) is small,
- The parameter value (factor) is false, due to the imperfect knowledge of the asset itself.

The following principle is used to obtain the relative risk (RR). Over one given year, the number of events $N^{events}_{observed}(\Delta)$ on a network Δ is stored, the statistical distribution Φi by modality i for a given parameter being known. Assuming the number of unacceptable events awaited for each modality is (1):

$$N^{events}_{awaited\ i}(\Delta) = N^{events}_{observed\ i}(\Delta) * \Phi_i \tag{1}$$

The Relative Risk (RR) of the modality i in the network Δ is determined by:

$$RR_i(\Delta) = \frac{N^{events}_{observed\ i}(\Delta)}{N^{events}_{awaited\ i}(\Delta)} \tag{2}$$

$RR_i(\Delta)$ is calculated every year. To facilitate stability and convergence of the numerical process, the $RR_i(\Delta)$, for modality i, at the year t, is updated according to:

$$\overline{RR}^{t}_{i}(\Delta) = \frac{M * RR^{t-1}_{i}(\Delta) + RR^{t}_{i}(\Delta)}{M+1} \tag{3}$$

with $\overline{RR}^{t}_{i}(\Delta)$ being the relative risk updated for year t, $RR^{t-1}_{i}(\Delta)$ the relative risk known for year (t–1), $RR^{t}_{i}(\Delta)$ the relative risk calculated for year t, and M-a coefficient taking account of the quantity of the generated information. The coefficient M evolves with the number of years of management.

Figure 2 shows the RR for two deficiencies (stopping and mechanical collapse) for four chosen parameters (Diameter, Depth, Slope and Material). For example, one can see that for stopping deficiency, the RR due to slope is the highest (RR=1.64) when the slope is weak and it is lower for medium for larges slopes. On the other hand, the Depth seems to have no effect (RR near to one for all the modalities).

RELATIVE RISK OF A SEWER SECTION

On a given section identified at year t by all its modalities m for all parameters p, the global relative risk $\Re_{PI}$ can be computed.

$$\Re_{PI} = \prod_{pm=1}^{pm=p} RR_{pm} \tag{4}$$

Where (PI) is the studied performance indicator, pm the urban parameter modality and (RR) the relative risk associated with the modality of the parameter on a given section.

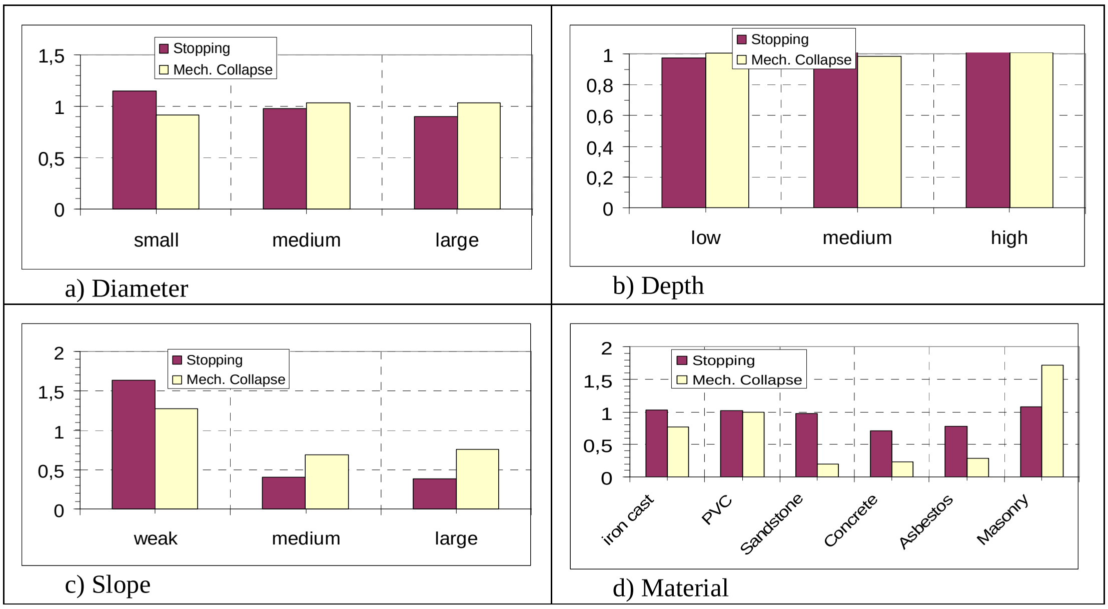


Figure 2: Relative Risks of two deficiencies corresponding to four parameters.

For instance, the Figure 3 presents the temporal evolution of $\Re_{PI}$ over ten years of a non maintained sewer and for two performance indicators (stopping, mechanical collapse) considering three different sewer sections. One can see that the RR of the section #28 is 2.6 times higher than that the section #29 for mechanical collapse (3.2 versus 1.2).

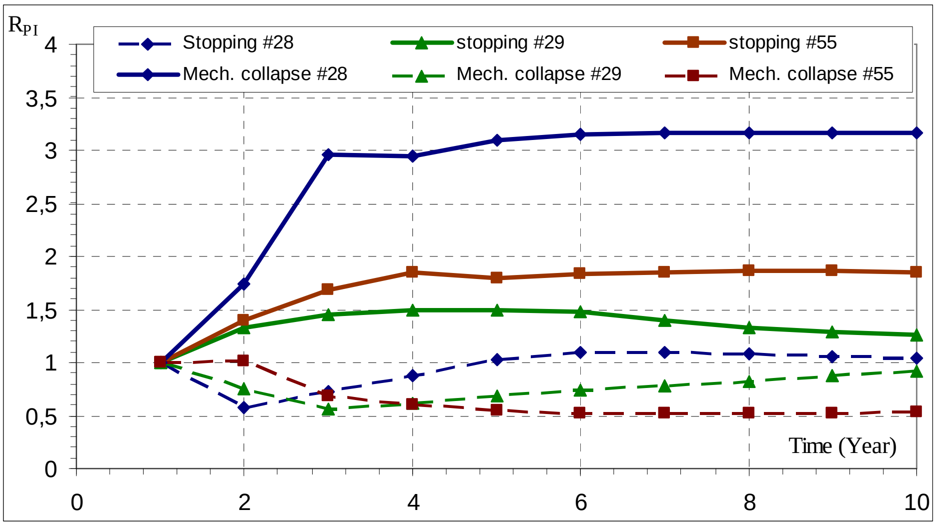


Figure 3: Temporal evolution of $\Re_{PI}$ for two performance indicators for three sewer sections (#28, 29 and 55, non maintained sewers).

EFFECTS OF THE "IMPERFECT KNOWLEDGE"

UNCERTAINTY AND "IMPERFECT KNOWLEDGE"

The condition rating that follows a sewer assessment is used to objectively determine the current condition of sewers. A rating system that minimizes subjective evaluation and is repeatable can be effectively used to predict future condition. It is acknowledged that it is impractical to develop a sophisticated condition rating system if the deterioration process of a sewer structure is not fully understood, as it is the case when new methodologies or materials are involved.

CASE STUDY: VIRTUAL DATA BASE

A virtual Database has been constructed. The generation of its components (sections) is based on the probabilistic distributions of a real but incomplete Database of the town of Lilles in France (Vasconcelos, 2005). It contains 2400 sections, which correspond to 120 km.

The set of parameters, defaults, deficiencies, vulnerability and performances taken into account is larger than that used in this paper and describes all major aspects of the asset time evolution.

FOLLOWED APPROACH

In order to measure the effect of the imperfect knowledge of the asset, we have proceeded as follows:

- From the original database, calculation of the relative risk $\overline{RR}_i^t$ for two performances (stopping and mechanical collapse) according to each parameter's modality. These risks, after numerical convergence, are considered as a "true" reference,
- Calculation of the relative risk $\Re_{PI}$ for each section for the two performances. This relative risk is called RR_actual,
- Decay of the information regarding a given parameter (four are alternatively considered: depth, material, slope and diameter) for some fraction of the asset (10%, 20%, 40% and 60%),
- Recalculation of new relative risks $\Re_{PI}$ for all sections with "false" data regarding some of them. This relative risk is called RR_estimated.

We will study now how the quality of R _estimated will decrease when the information is progressively decayed.

The figure 4 shows RR_estimated vs RR_actual for the two performances assuming that the slope value is unknown for 60% of sections. The figure is sufficiently explicit but unfortunately unexploitable if one wishes to quantify the effect of this ignorance on the model. To do this, it is interesting to introduce what is known as the confusion matrix (Baillargeon and Périgny, 2002).

It concerns a table of contingency confronting estimated classes (columns) and real classes (lines) for all the Database. On the diagonal, one finds, therefore, well classified values and out of the diagonal badly classified elements. The sum of values on a line gives the exact number of elements of the category.

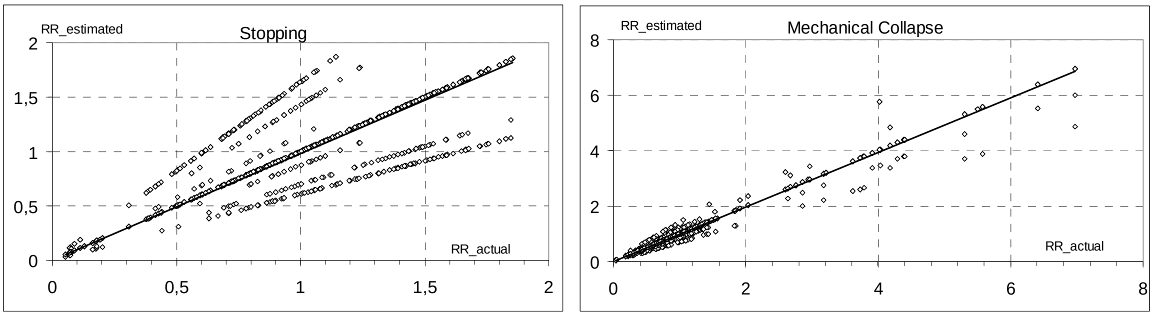


Figure 4: RR_estimated vs RR_actual for the two performances (ignorance of 60% on the slope parameter).

Tables 1 and 2 show respectively, the corresponding confusion matrices for stopping and mechanical collapse deficiencies.

One can define 3 measurements of error:

- The Global Error Rate GER: it is the ratio of badly classified elements to the number of elements of the Database.
- The A priori Error Rate AER: it is the rate of error for each category (8^{th} column),
- The Posteriori Error Rate PER: it is the rate of bad rankings related to the good classifications for a given class (8^{th} line).

The AER is always smaller than the PER for the two deficiencies. The AER and PER of the stopping deficiency are larger than those of the mechanical collapse deficiency for this example, (AER included between 0.8% and 69.2% while the PER is included between 1% and 45%). The larger AER does not correspond necessarily to the larger PER.

Table 1: Confusion matrix for different classes of RR (for Stopping deficiency, 60% of ignorance for the slope parameter)

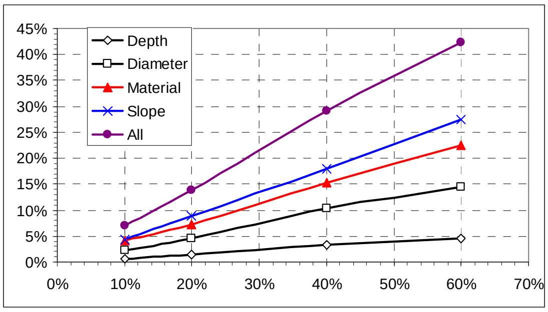
	Class 1 [0 – 0.4]	Class 2 [0.4 – 0.8]	Class 3 [0.8 – 1.2]	Class 4 [1.2 – 1.6]	Class 5 >=1.6		AER
Class 1	147	6	0	4	11	168	12.50%
Class 2	40	307	4	0	137	488	37.09%
Class 3	107	1	620	0	160	888	30.18%
Class 4	152	0	1	524	11	688	23.84%
Class 5	9	17	0	0	142	168	15.48%
	455	331	625	528	461	2400	
PER	67.69%	7.25%	0.80%	0.76%	69.20%		

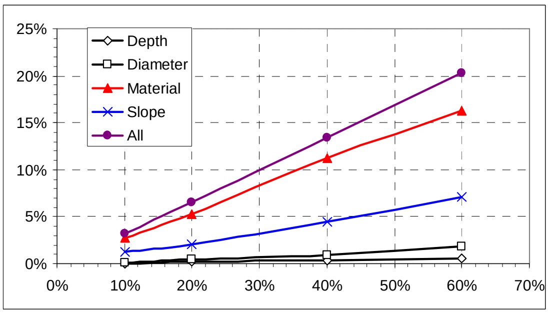
Table 2: Confusion matrix for different classes of RR (for Mechanical Collapse deficiency, 60% of ignorance for the slope parameter)

	Class 1 [0 – 1.4]	Class 2 [1.4 – 2.8]	Class 3 [2.8 – 4.2]	Class 4 [4.2 – 5.6]	Class 5 [5.6 - 7]		AER
Class 1	1962	94	0	0	0	2056	4;57%
Class 2	20	143	8	5	0	176	18.75%
Class 3	0	19	73	7	5	104	29.81%
Class 4	0	5	4	39	0	48	18.75%
Class 5	0	0	4	0	12	16	25.00%
	1982	261	89	51	17	2400	
PER	1.01%	45.21%	17.98%	23.53%	29.41%		

Figure 5 indicates the GER when each one of the parameters is alternatively considered as "unknown" and gathered. For this the information alteration included between 10% and 60%. The results of simulation where all four parameters have been taken as "unknown" has been added together. We can observe that all the parameters have not the same weight for each performance.

For instance, The GER is equal to 27.5% for the stopping performance and 7.13% for the mechanical collapse performance. It is interesting to see that the network plays the role of a filter insofar as the GER is attenuated for all the parameters effects (for example, for the parameter diameter and the stopping performance we have 2.21%, 4.58%, 10.3% and 14.5% for respectively 10, 20, 40 and 60%). The Material and the slope seems to be, through the GER, the dominant parameters, thus any effort to improve the data quality have to focus on these parameters.





a) b)

Figure 5: Global Error Rate vs imperfect knowledge of the Database [a) Stopping deficiency, b) Mech. Collapse deficiency].

CONCLUSIONS

Decisions on sewer management have to be taken under uncertainty and imperfect or partial knowledge. Managers of sewer system are faced with their infrastructure system ageing and the needed investment to keep an acceptable performance is a complex task.

A statistical and decision (using multicriteria analysis) model has been developed and used to improve strategies of IMR interventions.

A better knowledge of present performances at various scales and their evolution with time, from the local (pipe) to the global (entire network) scales, helps forecast future evolution and enables to to take decisions with less uncertainty.

A Relative Risk based analysis, for estimating the performances of the sewer's sections, was developed.

The sensitivity analysis of estimated risks (confidence in the final decision) and uncertainties in external inputs has direct consequences to sort out prioritary efforts to improve the quality of the Database.

REFERENCES

Baillargeon , G., Périgny , M.J. (2002), "Probabilities and statistics: with applications", SMG editors.

Elachachi, S.M, Breysse, D. (2005), " Study on the installation of operational tools of technical patrimony management: application to sewers systems" (in french), *Technical Report n°2005-28*, CDGA, University of Bordeaux I.

EN 13508-2 (2003), "Condition of drain and sewer systems outside buildings, Part 2: Visual inspection coding system", European Norms.

Le Gauffre, P., Joannis, C., Breysse, D., Gibello, C., Desmulliez, J.J. (2004), "Management of urban sewers systems: Methodological guide" (in French), Lavoisier.

Vasconcelos, E. (2005), "*Development of assistance tools with management of the nonworth visiting sewers system*" (in french), Doctoral Thesis, University of Bordeaux I.