# RELIABILITY INDEXES FOR WATER SUPPLY SYSTEMS

*indicadores de Fiabilidade de redes de abastecimento de água*

***João Pedro Poças Martins***
*Assistente*
*Secção de Construções Civis*
*Grupo GEQUALTEC*
*jppm@fe.up.pt*

***Joaquim Poças Martins***
*Professor Associado*
*Secção de Hidráulica, Recursos Hídricos e Ambiente*
*Águas do Porto, EM*
*jpm@fe.up.pt*

## Resumo

As questões relacionadas com a fiabilidade são fundamentais para a gestão de redes de abastecimento de água. Na ausência de indicadores com aceitação universal, a fiabilidade tem sido medida com base em indicadores simples, que avaliam o excesso de capacidade de componentes chave da rede. Esta abordagem não permite avaliar os diferentes tipos de falhas que podem ocorrer e não descreve a heterogeneidade da rede. É importante distinguir as falhas rotineiras das falhas catastróficas. O primeiro tipo de falhas pode ser avaliado com base em critérios probabilísticos. O segundo deve ser abordado tendo em conta critérios de redundância aplicados aos componentes chave da rede. Neste texto são avaliados os indicadores de fiabilidade mais utilizados e são propostos dois indicadores alternativos. Os indicadores descritos têm sido testados numa rede com cerca de 150 000 clientes. Foram desenvolvidos de modo a fornecerem indicações úteis a clientes, entidades gestoras e entidades reguladoras.
**Palavras-chave:** fiabilidade, indicadores, rede, abastecimento, água

## Abstract

Reliability issues are a major concern for the design and operation of modern water supply systems. In the absence of universally accepted indexes, the reliability of this water supply systems is usually measured using simple indexes, which evaluate the excess capacity in a network's key components. This approach does not distinguish between different possible failure modes and does not reveal the existence of different performance levels across the network. Routine and catastrophic situations must be dealt with separately. The former may be addressed using probabilistic analysis. The latter are better addressed through the duplication of key components in the network. In this paper, the most popular reliability indexes are reviewed and two alternative indexes are proposed. The indexes described in this paper have been tested on a water company serving 150 000 customers. These indexes aim to provide meaningful information to customers, management and regulators.
**Keywords**: reliability, indexes, water, supply, network

## 1 Introduction

Reliability issues are a major concern for the design and operation of modern water supply systems. The reliability of a public water system can be measured by the likelihood that its customers will be supplied in acceptable conditions. These conditions can be defined in terms of water quality, limit pressures (maximum and minimum), etc. [1].

The criteria used to measure reliability can be divided in the following categories:

- System capacity – Measured by the relation between the system's actual (or future) capacity and the predicted reference year annual consumption.
- Redundancy – These criteria measure the dependence of the system on a set of critical components.
- Probabilistic criteria – Measured by the probability of occurrence of a water supply failure for the network as a whole, for one of its regions or for an individual customer. The calculation of this probability requires a considerable amount of information and a larger amount of processing power than required by other kinds of indexes.
- Customers' opinions – These criteria are often found to be of limited use. The majority of customers do not present formal complaints as a result of supply failures. When polls are carried out, meaningful correlations between customers' opinions and recorded supply failures are often inexistent [2].

The indexes described in this article have been developed with the following purposes:

- To allow the detection and the analysis of heterogeneous behavior across the network – Persistent problems in certain regions of a network are unacceptable. Network-wide indexes such as the Security if Supply Index proposed by the Office of Water Services [3] do not identify this kind of failure.
- To measure reliability from the customer's point of view – Value is better identified by the customer. The customer's point of view (how likely is the customer to be affected by supply failures during a certain period of time) should, therefore, be included in the analysis, besides broader scoped indexes.
- To consider different failure modes – Routine and catastrophic events must be distinguished. A water supply system that is found to perform reliably under normal conditions, for example, might be very venerable under severe conditions (droughts, earthquakes, pollution, etc.).

## 2 Routine and extreme events – the manager's point of view

Pipe bursts and power shortages are examples of routine events. These events are quite predictable because it is easy for water companies to gather data concerning previous occurrences in their operational records.

Extreme events include earthquakes, severe floods or droughts, pollution, war, etc. Since the probability of occurrence of this kind of events is very low, useful data based on previous events is usually unavailable.

The magnitudes of failures associated with each of these kinds of events are obviously different. As a result of these events, supply failures of different durations are expected to occur. Besides the (real) impact on the quality of service, the possible impact of the event itself and its influence on the customer's perception of the supply failure should also be considered. As noted, causes of supply failure may range from a simple pipe burst to the occurrence of a natural catastrophe.

Since routine and extreme events have distinct causes, their effects (both real and perceived) are different and the amount of data available for the estimation of their probability of occurrence are so dissimilar that each of these kinds of events should be analyzed separately.

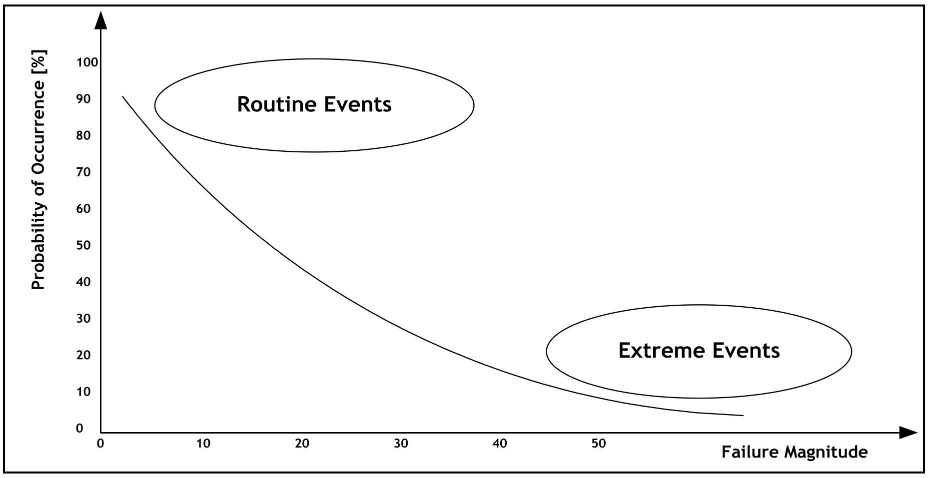


Fig. 1 – Perfil de fiabilidade. Eventos rotineiros e eventos extremos.

The following examples show that different networks might be well prepared for one kind of events but they may be particularly vulnerable to the other:

- Network *N* was built recently. The company that manages this network strives to control leakage so pipe bursts are detected and repaired quickly. Customers, therefore, experience few service interruptions each year. The duration of these interruptions is usually short. On the other hand, the system relies on a single water treatment plant which renders the network vulnerable to extreme events with a low probability of occurrence.
- Network *O* hás developed throughout decades, along with the city it serves, absorbing smaller existing networks along this growth process. The network includes, therefore, several alternative water sources. Although the old age of a large part of its components contributes to the occurrence of regular, short service interruptions, the existence of redundant water sources makes it more resistant to the effects of extreme events.

From the manager's point of view, the network's vulnerability to different kinds of events should be analyzed in order to decide how and how much should be spent in order to increase its overall reliability.

Fig. 2 shows the distinction between normal operation costs and the costs required to increase the reliability of a system regarding routine and extreme events. The costs related to extreme events include those necessary to build redundant components for all the major parts of the system. This is roughly a fixed cost. The reliability of a network regarding routine events can be increased almost indefinitely as a result of increasing the investment in leakage detection and pipe burst reparation, on one hand, and by adding redundant components to the distribution network, in particular by building loops in the network, on the other. Obviously, there is an upper limit to the amount of money that should be spent on increasing the reliability of network regarding routine events.

The manager must, therefore, decide if he should concentrate the company's resources on the prevention of short, recurring service interruptions (to which the customers are particularly sensitive) or if he should strive to mitigate the consequences of a hypothetical extreme event that the customers are unlikely to ever experience.

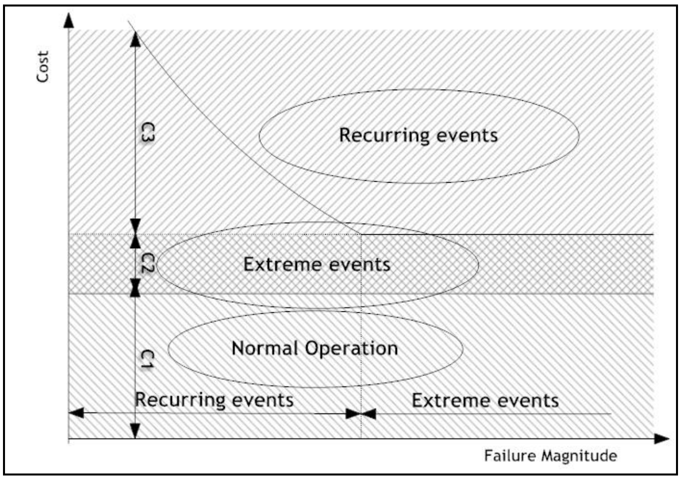


Fig. 2 – Kinds of events and respective effect mitigating costs

## 2 ProbIn – Probabilistic index

The most rational approach to the assessment of the reliability of a network is the adoption of a probabilistic index. Given the adequate information, the probability of the occurrence of a service interruption can be calculated for any of the network's nodes. Unfortunately, the variables required to calculate the reliability functions for the network's individual components are often unknown or the information that is needed to determine the value of these variables is not adequately organized. A number of simplifying assumptions should, therefore, be adopted in order to analyze the network. [1]. Besides a large volume of data, considerable processing power is also required to perform the calculations, especially for networks with a large number of components or with many loops.

Despite these disadvantages, it is believed that it is possible to gather the information that is necessary to characterize the reliability of a network's components, in particular those that are part of the distribution network. A major part of the effort required to organize and analyze component reliability data should be performed locally, by the water companies, Previous study [4] shows that the durability of pipes, for example, depends largely on the quality of the workmanship used to assemble them so studies about the reliability of these components are not universally valid. Probabilistic indexes are, therefore, a valid solution for the analysis of water supply networks, in particular distribution networks, although they require water companies to gather and analyze data.

The proposed probabilistic index is calculated by means of an algorithm that is also used to obtain the POCINDEX composite index, also presented in this work

### 2.1. Description of the algorithm

#### 2.1.1 Input

The following data is necessary for the calculation of the índex:

- Network configuration (using an EPANET *.INP file format).
- Variables for the calculation of the network's individual components' reliability functions (pipe materials, the existence of generators at pumping stations, etc.).
- *Source* and *sink* nodes for the network or for the region of the network to be studied:
  - The *sink* will usually represent an individual customer.
  - The source will represent a treatment plant, a water source, etc.

- If the network includes several sources, a fictitious node should be added and should be connected to all of the sources. For calculation purposes, this node will assume the role of the source.

### 2.1.2 Output

The algorithm's main output is the probability of occurrence of any kind of failure that will result in a service interruption. An additional output is a list of critical components for a pair of source-sink nodes (Fig. 3). A failure in a critical components results in a service interruption due to the absence of redundant components.

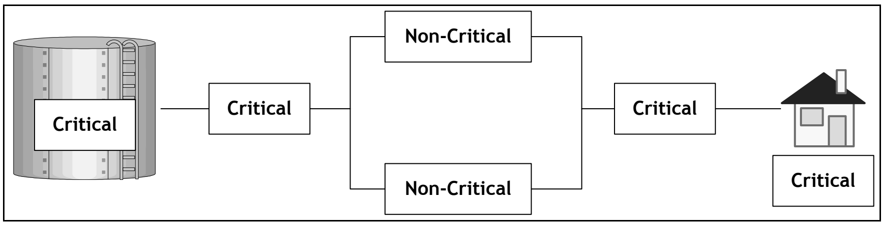


Fig. 3 – Critical components in a network – example.

### 2.1.3 Algorithm description

The algorithm is divided in four main steps:

1. Identification of all possible paths from the source to the sink.
2. Elimination of all non-functional paths, from a hydraulic point of view.
3. Calculation of the probability of success associated with each individual path.
4. Calculation of the joint probability of success of the set of paths.

The first step is performed by a *breadth-first search* (BFS) algorithm [5]: Beginning at the source ($N_0$), the links to its neighboring nodes ($N_1$) are considered, then to the following nodes ($N_2$) and so forth until a dead end (a node that is only connected to nodes that have already been visited by the algorithm), or the sink is reached. The analysis of networks containing a large number of nodes and loops will require considerable processing power. The most complex configuration for a network is a complete graph (Fig. 4), where each node is connected to every other node in the network.

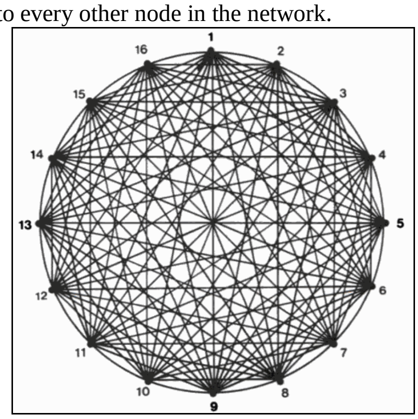


Fig. 4 – Complete graph ($K_n$) for a set of $n$ nodes, with $n$=16

A complete graph containing $n$ nodes has $n_a$ edges (in this context, $n_a$ pipes):

$$n_a = \frac{n(n-1)}{2} \qquad (1)$$

A complete graph with *n* nodes contains $n_p$ paths:

$$n_p = 1 + \sum_{i=1}^{n-2} \left( \prod_{j=i}^{n-2} j \right) \qquad (2)$$

The efficiency of the algorithm, in extreme cases, will therefore be:

$$O\left((n-2)!\right) \qquad (3)$$

This result should be viewed as a clearly pessimistic one. In fact, the configuration of real water supply networks is never even close to a complete graph. As a note, the proposed algorithm analyzes real networks (in particular, the city of Oporto's distribution network) in only a few seconds (the results depend, of course, on the source and sink nodes chosen by the user) so the algorithm is adequate for the dimension of admissible problems.

The second step of the algorithm can be performed by two different methods:

1. Simplified analysis – In each individual path, the heights of its nodes are compared to the height of the source. If a node is found that has a height that is greater than the height of the source in a path that contains no pumps, the path is eliminated. This process is performed very quickly, but it may result in the undue consideration of non-functional paths, from a hydraulic point of view, in the following stages of the algorithm.
2. Detailed analysis – Each path is exported to EPANET as an independent network and it is analyzed to detect eventual errors and/or failures to supply the sink. This process is usually slow.

The first method is usually the preferred one. Even if a path is unduly considered as valid, it will in fact be "approximately valid". This means that the path might become valid by performing slight changes to the network. This problem might occur as the result of using, for instance, pumps with insufficient elevation to link source and sink under the conditions defined by the user.

The third step is still a work in process. As mentioned earlier, the data that is necessary to characterize the reliability of a network's individual components is not always available. Even if the reliability of a certain class of components has been studied, the conclusions of the studies may not be universally applicable [6-8]. Water companies should perform their own evaluation of the reliability of components (of pipes, in particular) according to the operational data gathered by them. This task is critical for reliable results.

In the algorithm's final step, the paths are combined in order to evaluate the reliability of the service. Once it is divided in individual paths, a system may be considered successful if any of these paths are. The reliability of each path should, therefore, be calculated. The reliability of the system is calculated as a union of its *n* paths:

$$R_S = P\left(P_1 \cup P_2 \cup \ldots \cup P_n\right) \qquad (4)$$

$R_S$ is the reliability of the system and $P(P_i)$ is the probability of success of path *i*. The probability of success of a set of two paths, for instance, can be calculated using the following expression:

$$P\left(P_1 \cup P_2\right) = P\left(P_1\right) + P\left(P_2\right) - P\left(P_1 \cap P_2\right) \qquad (5)$$

Since paths are serial successions of components, the probability of success of path *i*, $P(P_i)$, is the joint probability of success of its *m* components ($C_1, \ldots, C_m$):

$$P_i = P\left(C_1 \cap C_2 \cap \ldots \cap C_m\right) \qquad (6)$$

This evaluation process for complex systems is called a path-tracing method.
This algorithm has been implemented in a T-SQL application. A Visual Basic application has also been developed to import data from EPANET files.

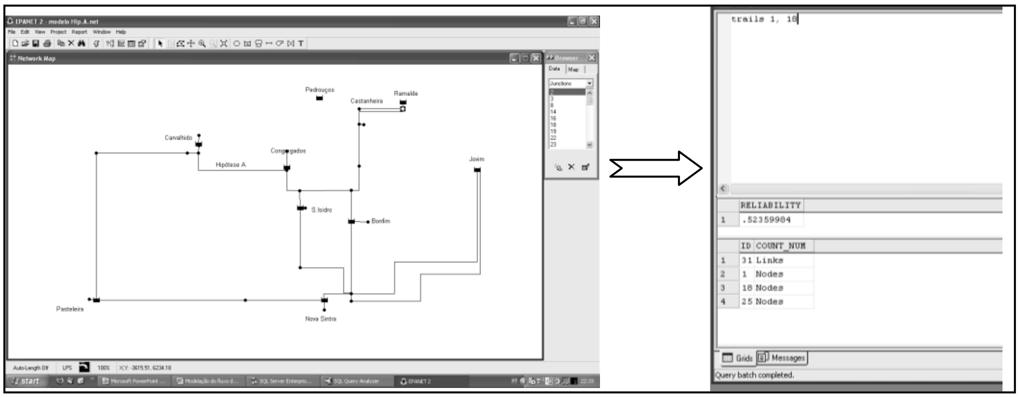


Fig. 5 – Software applications developed to implement the proposed algorithm

### 2.2. Probabilistic indexes. Conclusions.

Although probabilistic indexes appear to be a rational approach to the evaluation of the reliability of water supply systems, the scarceness of useful information and the inconsistencies detected in data collected from different networks hinders their adoption in real cases. Even so, it is possible to gather enough data to characterize the reliability of network components locally, making these indexes useful as tools for the evaluation of the reliability of networks regarding routine events.
The existence of redundant components is a necessary condition for favorable results to be achieved concerning the reliability of networks under extreme events. The reliability of networks under routine events is increased considerably when customers are supplied from domestic water tanks

## 3. POCINDEX

POCINDEX is a composite index that was developed in order to avoid some of the disadvantages that were found in other reliability measures. This index was developed considering not only the data that is currently gathered by water companies (in particular, by Águas do Porto, EM) but also other data that can be obtained presently.

### 3.1. Description of the algorithm

POCINDEX is divided in two separate indexes. Each of these indexes is composed by information about the past behavior of the system and redundancy criteria.
Information about past behavior is important for reliability assessment although it should be used with caution. The most popular reliability index for urban water networks, IWA's QS13 [9, 10] is based on this kind of information:

$$QS13 = \frac{NI \times DI}{NT \times HY} \qquad (7)$$

- NI – Number of customers that were affected by service interruptions during the report period.
- DI – Duration of the service interruption, measured in hours.

- NT – Total number of customers.
- NY – Duration of the report period, measured in hours.

Electrical networks also use information about past behavior as a measure of reliability. The IEEE 1366 standard includes several indexes of this nature [11]. The $CEMI_n$ index (*Customers Experiencing Multiple Interruptions*), for instance, displays some network vulnerabilities that are hidden by indexes that measure a system as a whole:

$$CEMI_n = \frac{CN_{(k>n)}}{N_T} \tag{8}$$

- $CN_{(k>n)}$ – Total number of customers that experienced more than *n* sustained service interruptions (with a duration of one minute or more).
- $N_T$ – Total number of customers.

Several states in the US are considering the usage of this index, not only as a measure of reliability, but also as a criterion for the application of fines that would be paid by companies to customers affected by the insufficient performance of the supply network [11, 12]. The application of a similar index to water supply networks is believed to be beneficial since the distribution of failures across a network tends to be very heterogeneous.

A disadvantage of this kind of indexes is that, in general, reliable and auditable information about the duration of service interruptions and the number of customers affected by them is not collected be water companies. It is important, therefore, not only to develop useful indexes, but also to make these indexes viable and auditable by suggesting ways to calculate them from information collected through reliable processes.

The times of the occurrence of failures and of the respective reestablishment of service should be recorded. The consistency of this information can be verified against the operational data gathered by the companies, especially if ERP systems (*Enterprise Resource Planning*) are used to support human resource management.

The number of customers affected can be obtained by several different methods:

1. The simplest method, but also the least reliable and least auditable one, is performed *in situ* by the work team responsible for the reestablishment of the supply (in the cases where the interruption is caused by a pipe burst). This information should be added to the work hours tally sheet that is used in many companies to record the number of work hours necessary to perform the repair which in turn is used to calculate the cost of the repair.
2. Given the location of the failure, the valves that should be closed during the repair can be identified and the number of customers affected can be determined. This method obviously requires a reliable Geographical Information System and a customer database, but it can be automated, making it auditable. This process is likely to be the most suitable for most water companies, although it might require the development of adequate information systems.
3. The last alternative is the least realistic in the short term, although its adoption is feasible with today's technology. It requires AMR equipment (Automatic Meter Reading) to be installed upstream of all water meters so that network data can be measured and transmitted in real time.

The Customer POCINDEX index is calculated as follows:

$$CRI = \left(NI, ST, R_Q\right) / \left(R_{loop}, R_{source}, R_{trunk}\right) \tag{9}$$

- *NI* – Number of service interruptions experienced by the customer during the report period (the previous year).
- *ST* – Storage available for the customer, measured in days, assuming na average day water consumption along the network.

- $R_Q$ – Relation between a customer's maximum water consumption flow and the flow made available by the public network under unfavorable conditions, as described in expression(10).
- $R_{loop}$ – Binary value – Does the distribution network's configuration allow supply through two or more different paths?
- $R_{source}$ – Binary value – Can the customer be supplied from two or more different water sources?
- $R_{trunk}$ – Binary value – Can the customer be supplied through two or more different trunk mains?

$$R_Q = \frac{Q_T}{Q_D} \qquad (10)$$

- $Q_T$ – Sum of the instant flows of all of the devices in the customer's domestic water system. The calculation of this parameter may be performed automatically if the domestic water system's designs are submitted as an information model (BIM) that follows a standard format.
- $Q_T$ – Maximum flow that can be delivered to the customer, regardless of the characteristics of his domestic water network, considering low pressure in the public network.

$$Q_D = 0.15\pi d^2 \sqrt{2gp/\gamma} \qquad (11)$$

- $d$ – Diameter of the pipe upstream of the water meter.
- $p/\gamma$ – Piezometric level immediately upstream of the customer's water meter.

All of these variables should be obtainable automatically from the water company's operational databases. The binary values can be obtained using the breath-first search algorithm described earlier, used to calculate the ProbIn index.

The Customer POCINDEX index is calculated as follows:

$$SRI = \left(\%1I, \%2I, \% > 2I\right) / \left(\%1S, \%1R, \% < 2ST\right) / \left(\%XD, \%XW, \%XM\right) \qquad (12)$$

- *%1I, %2I, %>2I* – Relative number of customers (expressed a percentage) that experienced one, two or more than two service interruptions during the report period, respectively.
- *%1S, %1R, %<2ST* – Relative number of customers (expressed a percentage) served by a single source, a single trunk main or with less than two days of storage available, considering average day consumptions.
- *%XD, %XW, %XM* – Excess storage capacity of the system considering average day, average week or average month consumptions.

Similarly to the Customer POCINDEX, the System POCINDEX's variables should be automatically obtainable from the water companies' operational databases.

## 4 Conclusions

When assessing the reliability of water supply networks, routine and extreme events should be considered separately. A system might be reliable under one kind of event but particularly vulnerable under the other. Usually, a network's behavior regarding reliability is not homogeneous. Some regions will perform better than others. Indexes that are applied to a network as a whole, which treat reliability as a one-dimensional problem and that hide the existence of customers affected by frequent service interruptions do not characterize the reliability of a water supply network adequately by themselves.

Two separate indexes have been presented: Probin (a probabilistic index) and POCINDEX (a composite index based on redundancy criteria and information about past behavior). These indexes have been developed so that they can be determined automatically for any node in a network. A systematic approach should be followed when collecting data and when calculating the indexes to guarantee the reliability and the auditability of the results. Automated calculation of this kind of indexes may require water companies to develop or to improve existing information systems.

**References**

[1] Harberg, R. J.; *Planning and managing reliable urban water systems* Denver, 1997.
[2] Roberts, D. , Davis, X.; *Customer-Oriented Reliability*, Transmission & Distribution World, December 1 2003.
[3] OFWAT; *Security of supply, leakage and the efficient use of water. 2002–2003 report* 2003.
[4] Hurley, R.; *Exclusive Result Findings Comparing The Reliability Of Plastic Vs. Metal Pipe Materials*, Global Water Leakage Summit, January 25-26 2007.
[5] Black, P. E.; *Depth-first search* Dictionary of Algorithms and Data Structures, 2008.
[6] O'Day, D. K., Weiss, R., Chiavari, S., Blair, D.; *Water Main Evaluation for Rehabilitation/Replacement* Denver, Colorado, 1986.
[7] Røstum, J.; *Statistical modelling of pipe failures in water networks*, PhD Thesis, Norwegian University of Science and Technology, 2000.
[8] Boxall, J. B., O'Hagan, A., Pooladsaz, S., Saul, A. J., Unwin, D. M.; *Estimation of burst rates in water distribution mains*, Water Management, June 2007.
[9] Alegre, H.; *Performance indicators for water supply services* London, 2000.
[10] Duarte, P., Alegre, H. , Baptista, J. M.; *Indicadores de desempenho para serviços de abastecimento de água* Lisbon, 2004.
[11] Warren, C. A.; *Overview of 1366-2001 the Full Use Guide on Electric Power Distribution Reliability Indices*, Power Engineering Society Summer Meeting, 2002 IEEE, July 25 2002.
[12] Matavalam, R. K. R.; *Power Distribuition Reliability as a Function of Weather* MsC Thesis, University of Florida, 2004.