

# Probabilistic Model of Cumulative Damage in Pipelines Using Markov Chains

Francisco Casanova-del-Angel, Esteban Flores-Méndez, Karina Guadalupe Cortes-Yah

Sección de Estudios de Posgrado e Investigación. Escuela Superior de Ingeniería y Arquitectura, Unidad Profesional ALM, Instituto Politécnico Nacional, Ciudad de México, Mexico  
Email: fcasanova@ipn.mx, fmendeze@ipn.mx, kcontesy@ipn.mx

**How to cite this paper:** Casanova-del-Angel, F., Flores-Méndez, E. and Cortes-Yah, K.G. (2020) Probabilistic Model of Cumulative Damage in Pipelines Using Markov Chains. *Journal of Applied Mathematics and Physics*, 8, 620-642.  
<https://doi.org/10.4236/jamp.2020.84048>

**Received:** February 11, 2020

**Accepted:** March 28, 2020

**Published:** April 1, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

This paper presents a probabilistic model of cumulative damage based on Markov chains theory to model propagation of internal corrosion depth localized in a hydrocarbons transport pipeline. The damage accumulation mechanism is unit jump type, depending on the state. It uses a shock model based on Bernoulli trials and probabilities to remain in the same state or the next one. Data are adjusted to Lognormal distribution and proven with a Kolmogórov-Smirnov test. The vector obtained from multiplying the initial state vector with the transition matrix was developed and the system of equations to find each transition probability with a single inspection report was solved. In order to calculate propagation of internal corrosion after inspection, an exponential equation was proposed and a parameter was adjusted to the data. Time to expected failure was obtained by adding the time expected in each damage state. Each time step was adjusted to real time.

## Keywords

Localized Corrosion, Cumulative Damage, Failure, Pipelines, Markov Chains

## 1. Introduction

Throughout Mexico, there are approximately 54,000 kilometers of terrestrial pipeline in operation, transporting crude oil, natural gas, sour gas, sweet gas, gasoline, diesel, and other refined products. In addition, there are 2000 kilometers of underwater pipeline. Almost half of pipeline has been operating for more than 30 years and, in spite of continual maintenance, inspection and patrolling works, problems caused by corrosion, damage by third parties, operation and design mistakes that may lead to system failure are present. Failures of these pipeline systems constitute significant financial losses and, in the worst scenario,

loss of human lives and damage to the environment. Failure may be due to several causes and may occur in a small area of the pipeline; however, to identify significant factors contributing to failure in the pipeline is possible. Accidents due to corrosion in pipelines transporting hydrocarbon are the most frequent.

Corrosion is damage to materials caused by ions transportation and due to potential difference in the material itself or between the material and its surrounding environment. It is an irreversible interfacial reaction, generally of a metal, with its surrounding environment which damages or modifies its properties. Corrosion causes loss in mechanical resistance properties of the material, resulting in changes of structures and components geometry; therefore the material loses the function for which it was intended, causing: 1) direct losses due to change in corroded tubular structures; 2) indirect losses resulting from loss of production due to temporary suspension of productive systems and facilities and pollution of goods generated; 3) loss of wellbeing and human lives; 4) efficiency loss; and 5) increase of exploitation costs.

Corrosion is of major interest since the loss of metal invariably implies a reduction of the structural integrity of the pipe and an increase in the risk of failure. The two most significant factors involved in the corrosion process are the type of material and the environment. The environment includes conditions impacting on the internal and external walls of the pipe. Since most pipelines undergo various environmental conditions, the assessment must allow proper sectioning or consideration of each type of environment in each given segment. In addition, corrosion may be divided into uniform and localized corrosion. Uniform corrosion is that which develops at the same speed on all of the surface of the material, while localized corrosion refers to an accelerated attack on a specific superficial position, generally due to separation of anodic and cathodic areas, and is caused by inequalities of the structure or composition of the corroding material or differences in the environment. Pitting corrosion is a special type of localized corrosion, defined as a very localized attack on passive metals, causing very narrow and deep cavities.

The study of corrosion has been mainly based on deterministic approaches, particularly on the electrochemical corrosion theory; however, localized corrosion cannot be explained without statistical and stochastic points of view due to high dispersion of laboratory and field data (see ref. [1] and ref. [2]). Assessment of the integrity and reliability of pipelines has been studied in various ways in recent years and risk and reliability analyses have been applied to determine failure probabilities.

A formal analysis is carried out using reliability tests in the risk of failure assessment. Thus, from a structural point of view, reliability or probability of success is the probability of a structure, an element or structural element satisfying the limit and service conditions for which it was intended. In this manner, the risk of failure assessment on ducts based on mechanical or structural reliability methods is a better approach in assessing safety, providing an analytical tool which allows the assessment of risk with more reasonable and verifiable methods

that may always be improved. Reliability methods provide elements to design a structural system and are also used to know the current conditions of a system. This is very important when making decisions regarding maintenance, repair or substitution of structural elements in the system, in order to take the structural system back to the required safety level specified.

Cumulative damage is the irreversible accumulation of damage throughout the life of the element, which eventually results in degradation and failure. A stochastic process is a mathematical model of a dynamic process which evolution in time is governed by probabilistic laws (see ref. [3]). In probability theory, there were no general procedures or particular schemes to solve this type of problems. This brought a new general theory of random processes to study these problems. Any knowledge about the states of a system will lead us to study the type of stochastic processes called process without further effect or, by analogy, Markov chains or processes. Andrei Andreyevich Markov was a Russian mathematician known for his works on theory of numbers and probabilities theory. The model proposed assumes a finite and discrete state of time using Markov processes, which assumes that probability of what will happen in a given instance will depend exclusively on the immediate previous past.

J.L. Bogdanoff and F. Kozin started their study of Markov's cumulative damage models in 1978. In the beginning, there was doubt about its usefulness in engineering and its practical value. However, this negative point of view changed, since its potential application field in the cumulative damage phenomenon is vast. These models have been capable of describing and successfully analyzing life data of various phenomena, including fatigue, growth of cracks due to fatigue, corrosion, and change of properties of the material, among others. In addition, these models have provided profound knowledge about many interesting problems in engineering (see ref. [3]). Models of phenomena may be deterministic or probabilistic. The interest of above-mentioned researchers is focused on the evolution in time under cyclic uses, damage due to fatigue, crack lengths, and loss of material. A probabilistic model is proper for cumulative damage phenomena in which uncertainties to estimate the mean life may not be ignored. Localized corrosion is known for showing high variation in measurable parameters such as corrosion rate, maximum pitting depth, and time to perforation, among others. Variation in results is derived from the influence of heterogeneities on the metal surface throughout pitting development and variations in the corroding environment throughout time. All these facts suggest that randomness is an inherent and unavoidable feature of pitting corrosion throughout time; therefore, stochastic models are a better choice for describing pitting corrosion than deterministic models are (see ref. [1]). Provan and Rodríguez used a non-homogeneous Markov process to model growth of pitting depth without taking into account the generation process. This model aims at describing the growth of pitting as a function of time of exposure, considering a discrete space of states of possible pitting depths. They compared the estimated results with experimental data reported for aluminum and directed their own experiences in

pitting corrosion towards stainless steels (see ref. [4]). H.P. Hong considered pitting corrosion as a combination of two stochastic processes: he modeled the generation of the pitting process as a Poisson process and the propagation of depth as a Markov process (see ref. [5]). The probability distribution of pitting corrosion depth and probability of time to failure are obtained by combining both processes. To use the method, examples based on experimental data were developed. In this work, the transition intensity used to describe growth of a specific depth in the homogeneous Markov process depends on the total number of states used and does not define the number of optimal states to describe a given set of experimental data (see ref. [5]). Valor and his research team developed a stochastic model to simulate localized corrosion. Localized corrosion was modeled as a non-homogeneous Poisson process in which the time of induction to pitting initiation was simulated as carrying out a Weibull process. To simulate growth of pitting they used a non-homogeneous Markov process. The model proposed is validated using data from published experiments on localized corrosion (see ref. [6]). A stochastic prediction model of damage evolution due to corrosion in active sites, applicable under professional practice conditions, is presented by J.L. Alamilla and E. Sosa (see ref. [7]). It uses integrity and structural reliability analyses. Damage of a material and its evolution are determined by the damage state in a given time instant and a damage occurrence rate. To achieve this, the probability density function of the pitting corrosion damage depth of the system is estimated. Its application depends on the quantity of inspection reports available. Two scenarios are presented: when there is one inspection report and when there are two inspection reports. In the latter scenario, there is variation regarding whether defects may be identified in the two inspections or not (see ref. [7]). In this work, localized corrosion due to internal pitting is studied, since depending on a lower number of factors than external corrosion, simpler mathematical models may be developed (see ref. [8]). Corrosion analyzed in this work is focused on *loss of pipe metal*. In addition, a probabilistic model of cumulative damage due to corrosion is carried out. Cumulative damage reduces reliability as time increases. Keeping high reliability levels, carrying out inspections, repairs, partial replacements, and keeping proper operation conditions may increase the cost of the system's life cycle. This study is useful because it creates a tool that enables the prediction of the corrosion state in a given instant of time, in order that resources for inspection, maintenance and repair programs may be optimized.

We set as an objective to establish a methodology based on a probabilistic damage model and Poisson and Markov stochastic processes that allow describing and predicting the state of corrosion in terrestrial pipelines for transport of hydrocarbons in a given period of time.

For such purpose, we proposed to statistically analyze the data on pipeline inspection; to define the limit states of variable *depth of corrosion defect*; to determine the starting and transition probability of the model; to assess corrosion of pipelines transporting hydrocarbons in several instants of time; and to calcu-

late expected time to failure.

## 2. Stationary Stochastic Model of Cumulative Damage

In our model of cumulative damage in pipelines using Markov chains, it is assumed that: 1) seriousness of the service cycle is repetitive and constant (that is, it is considered that anything that happens in a service cycle occurs in every other cycle); 2) damage states are discrete and called  $1, 2, \dots, j, \dots, b$  (failure); 3) cumulative damage in a service cycle depends only on the service cycle and the damage state at the beginning of the cycle; and 4) damage may only be increased in a service cycle, from the state occupied at the beginning of the cycle, to the state of a higher unit.

Time  $x$  is made discrete by the first assumption, for that reason  $x = 0, 1, 2, \dots$ . Permissible damage states have been made discrete in the second assumption, therefore, the model has finite (discrete) damage and time states. As for the third assumption, the model does not mention anything about quantitative details of what happens within a service cycle. This third assumption is Markov's, which is damage accumulation in a service cycle, only depending on the corresponding service cycle and the damage state in which the cycle started. In the model, to know how the starting damage was achieved is irrelevant. When state  $b$  is reached, withdrawal or failure occurs. Thus, state  $b$  is an absorbent state and  $\rho_b = 1$ . The fourth assumption implies that it must be transitory since, once damage moves from a state to the next one, it never goes back to the immediately previous state.

This version of the cumulative damage model is a stationary Markov process, with discrete time and finite states. The cumulative damage mechanism is of the unit jump type, depending on the state.

Let  $D_0$  be the random variable defined as the state occupied by damage in time  $x = 0$ . The starting probability distribution  $p_0$  over damage states in  $x = 0$  are specified by vector line  $(1 \times b)$ :  $p_0 = \{\pi_1, \pi_2, \dots, \pi_{b-1}, 0\}$ , where  $P\{D_0 = j\} = \pi_j \geq 0$  y  $\sum_{j=1}^{b-1} \pi_j = 1$ . It should be noted that in no case does it start at failure state, since  $\pi_b = 0$ .  $\pi$ 's creates a mass function with  $D_0$  probability.

The constant probability  $P$  transition matrix is associated if each service cycle has constant seriousness; in addition, based on the fourth assumption, it is only possible to pass from a state to a higher state, therefore:

$$P = \begin{bmatrix} p_1 & q_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & p_2 & q_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & p_3 & q_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & p_{b-1} & q_{b-1} \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \tag{1}$$

where  $1 > p_j > 0$  and  $p_j + q_j = 1$ .  $p_j$  is the probability to remain in state  $j$  for a step;  $q_j$  is the probability that, in a step, damage goes from state  $j$  to state  $j + 1$ . It

may be observed that for (1) all damage states  $1, \dots, b-1$  are transitory and state  $b$  is absorbent.

Let  $D_x$  be the damage state occupied in time  $x$ , and let:  
 $P\{D_x = j\} = p_x(j) \forall j = 1, \dots, b$ , be probability that damage occupies state  $j$  at time  $x$ . Since  $p_x(j) \geq 0$  and  $\sum_{j=1}^b p_x(j) = 1$ ,  $p_x(j)$  creates a probability mass function in time  $x$  in damage states  $1, \dots, b$ . Now, using vector line  $(1 \times b)$ :  
 $p_x = \{p_x(1), \dots, p_x(b)\}$  to represent this probability mass function and, based on Markov's chain theory:

$$p_x = p_0 P^x = p_{x-1} P \quad \forall x = 0, 1, 2, \dots \tag{2}$$

where  $P_0 = I$  and  $I$  is the identity matrix  $(b \times b)$ .

A graphic view of Equation (2) is that of a set of vertical bars representing  $p_x(j)$ , where  $x$  is on the time axis at regular intervals. When  $x$  increases, the probability mass is displaced from lower to higher states and, gradually, deviated to state  $b$ . In the last case, all masses accumulate in state  $b$ . Its final structure depends on  $p_0$  and  $P$ . It should be noted that  $p_x(b)$  is a distribution function accumulated in time  $x$ , since the limit when  $x \rightarrow \infty$  is  $p_x(b) = 1$ , as all masses in the last case come to state  $b$  and stay there. Probability  $p_x(j)$  is the probability that  $D_x$  is in state  $j$  in time  $x$ .

Given  $p_0$  and  $P$ , in (2)  $p_x = \{p_x(1), \dots, p_x(b)\}$  is calculated and it is known that the accumulated distribution function of time  $W_b$  to absorption in state  $b$  (with  $p_b = 1$ ) is  $p_x(b)$ .  $W_b$  becomes time to failure in the cumulative damage model. Cumulative distribution function of  $W_b$  is given by:

$$F_w(x; b) = P\{W_b \leq x\} = p_x(b) \quad \forall x = 0, 1, 2, \dots \tag{3}$$

$t$  reliability function  $\mathcal{F}_w(x; b)$  es:  $\mathcal{F}_w(x; b) = 1 - F_w(x; b) \forall x = 0, 1, 2, \dots$  Failure rate function  $h_w(x; b)$  is:  $h_w(x; b) = 1 - \frac{F_w(x; b)}{\mathcal{F}_w(x-1; b)} \forall x = 0, 1, 2, \dots$  The mean and variance of  $W_b$  are, respectively:

$$E\{W_b\} = \sum_0^\infty \mathcal{F}_w(x; b) \quad \text{and} \quad \text{var}(W_b) = 2 \sum_0^\infty x \mathcal{F}_w(x; b) + E\{W_b\} - [E\{W_b\}]^2$$

The probability that damage  $D_x$  is in state  $j$  in time  $x$  is determined by  $p_x(j)$ :  
 $P\{D_x = j\} = p_x(j) \forall j = 1, \dots, b$  and the cumulative distribution function of  $D_x$  is:  $F_D(j; x) = P\{D_x \leq j\} = \sum_{k=1}^j p_x(k) \forall j = 1, \dots, b$ . Its mean and variance are given by:

$$E\{D_x\} = \sum_{j=1}^b j p_x(j) \quad \text{and} \quad \text{var}(D_x) = \sum_{j=1}^b p_x(j) - [E\{D_x\}]^2$$

Thus, Equation (2) gives results required to determine basic probabilistic data for cumulative damage in which failure occurs upon arriving at state  $b$ . This equation is easy to program using standard routines to multiply matrices. If advantage of  $P$ 's shape is used, calculations may be efficiently managed.

For this case,  $\rho_b = 1$  and  $\pi_1 = 1$ , cumulative damage process starts in state 1 with  $x = 0$ . Random variable  $W_{1,b}$ , is time to failure in state  $b$  since it started in state 1 in  $x = 0$ . Geometric transformation of cumulative distribution function  $F_w(x; 1, b)$  de  $W_{1,b}$  is:

$$\psi_W(z;1,b) = \frac{q_1 \cdots q_{b-1} z^{b-1}}{(1-p_1 z)(1-p_2 z) \cdots (1-p_{b-1} z)(1-z)} \tag{4}$$

and the transformation of the probability mass function  $p_W(x;1,b)$  is:

$$\phi_W(z;1,b) = (1-z)\psi_W(z;1,b) \tag{5}$$

Inversion of these two transformations is direct. If we assume that:  $1 > p_1 > p_2 > \cdots > p_{b-1} > 0$ , then:

$$F_W(x;1,b) = 1 - g_1(1,b)p_1^x + g_2(1,b)p_2^x - \cdots + (-1)^{b-1} g_{b-1}(1,b)p_{b-1}^x \tag{6}$$

where:

$$\begin{aligned} g_1(1,b) &= \frac{q_2 q_3 \cdots q_{b-1}}{(p_1 - p_2)(p_1 - p_3) \cdots (p_1 - p_{b-1})} \\ &\quad \dots \\ g_j(1,b) &= \frac{q_1 \cdots q_{j-1} q_{j+1} \cdots q_{b-1}}{(p_1 - p_j)(p_2 - p_j) \cdots (p_{j-1} - p_j)(p_j - p_{j+1}) \cdots (p_j - p_{b-1})} \\ &\quad \dots \\ g_{b-1}(1,b) &= \frac{q_1 q_2 \cdots q_{b-2}}{(p_1 - p_{b-1})(p_2 - p_{b-1}) \cdots (p_{b-2} - p_{b-1})} \end{aligned} \tag{7}$$

It should be noted that:

$$F_W(x;1,b) = 0 \quad \forall x = 0, 1, \dots, b-2 \tag{8}$$

since a unit jump model may not reach state  $b$  from 1 in less than  $(b-1)$  n-steps. Therefore,  $g_j(1,b)$  satisfy certain relationships implied in (8). Equation (6) and Equation (7) are valid for  $1 > p_1 > p_2 > \cdots > p_{b-1} > 0$ .

Inversion of (5) gives the probability mass function  $p_W(x;1,b)$  of  $W_{1,b}$ ; and it is found using the technique used to obtain (6):

$$p_W(1;1,b) = q_1 g_1(1,b) p_1^{x-1} - \cdots + (-1)^{b-1} q_{b-1} g_{b-1}(1,b) p_{b-1}^{x-1} \tag{9}$$

This equation may also be obtained noticing that:

$$p_W(x;1,b) = F_W(x;1,b) - F_W(x-1;1,b) \tag{10}$$

$p_W(x;1,b)$  is equal to zero for  $x = 0, 1, \dots, b-2$ .

The mean and the first central moments of  $W_{1,b}$  are obtained from the characteristic function of (5), using the differentiation technique. These are:

$$\begin{aligned} E\{W_{1,b}\} &= \sum_1^{b-1} (1+r_j) \\ \text{var}(W_{1,b}) &= \sum_1^{b-1} r_j (1+r_j) \\ \mu_3\{W_{1,b}\} &= \sum_1^{b-1} r_j (1+r_j)(1+2r_j) \\ \mu_4\{W_{1,b}\} &= 3[\text{var}(W_{1,b})]^2 + \sum_1^{b-1} r_j (1+r_j)(1+2r_j)(1+3r_j) + \sum_1^{b-1} r_j^2 (1+r_j) \end{aligned}$$

where

$$r_j = \frac{p_j}{q_j} = \frac{1}{q_j} - 1 \quad \text{and} \quad p_j = \frac{r_j}{1 + r_j} \tag{11}$$

The reliability function  $\mathcal{F}_W(x;1,b)$  corresponding to the cumulative distribution function of (6) is:

$$\mathcal{F}_W(x;1,b) = g_1(1,b)p_1^x - g_2(1,b)p_2^x + \dots + (-1)^{b-1} g_{b-1}(1,b)p_{b-1}^x \tag{12}$$

The risk function  $h_w(x;1,b)$  is obtained using (9), (10) and (12), these being:

$$h_w(x;1,b) = 1 - \frac{g_1(1,b)p_1^x - \dots - (-1)^{b-1} g_{b-1}(1,b)p_{b-1}^x}{g_1(1,b)p_1^{x-1} - \dots - (-1)^{b-1} g_{b-1}(1,b)p_{b-1}^{x-1}} \tag{13}$$

$W_{1,b}$  is time required to reach state  $b$  since it started with state 1 at time  $x = 0$ . Given that:

$$W_{1,b} = T_1 + \dots + T_{b-1} \tag{14}$$

where  $T_j$ 's are independent and, upon adding  $T_j$ 's characteristic function, we obtain:

$$\phi_j(u) = \frac{q_j e^{iu}}{1 - p_j e^{iu}} \tag{15}$$

we have that  $W_{1,b}$ 's characteristic function is:

$$\phi_w(u;1,b) = \prod_1^{b-1} \frac{q_j e^{iu}}{1 - p_j e^{iu}} \tag{16}$$

If we substitute  $e^{iu}$  with  $z$ , we obtain the geometric transformation of the probability mass function of  $W_{1,b}$  given in (17)

$$\phi_w(z;1,b) = \frac{q_1 \dots q_{b-1} z^{b-1}}{(1 - p_1 z)(1 - p_2 z) \dots (1 - p_{b-1} z)} \tag{17}$$

This formula is the same as the one obtained in (5) using (4).

Let us assume now that  $p'_1 p'_2 \dots p'_{b-1}$  is a permutation of  $p_1, p_2, \dots, p_{b-1}$ , then (4) becomes:

$$\Psi_w(z;1,b) = \frac{q'_1 q'_2 \dots q'_{b-1} z^{b-1}}{(1 - p'_1 z)(1 - p'_2 z) \dots (1 - p'_{b-1} z)(1 - z)} \tag{18}$$

This equation may be inverted to find  $F_w(x;1,b)$  and obtain exactly the same equation as that defined in (6); for this reason, it may be said that there is  $(b-1)!$  permutations of  $p_j$ . Thus, there are  $(b-1)!$  models with the same cumulative distribution function  $F_w(x;1,b)$  of  $W_{1,b}$ . In order to know how these models are different from each other, the notion of sample function mean must be introduced.

Time used in state  $j$  is random variable  $T_j$ , whose characteristic function is given in (15). It is found that this is the expected time value used in each state, and it is given by:

$$E\{T_j\} = 1 + r_j \tag{19}$$



In the succession of the mean time used in each state, the adjacent points are linked with line segments; their segmented curve is the mean of the sample function, since it shows behavior of the sample function mean. Actually, the family of sample functions creates a propagated cloud coming from state 1 at  $x=0$ . Mean simple function is the position of the mean of this cloud in various states. Time in which the sample function mean reaches  $b$  is, in case of the unit jump,  $E\{W_{1,b}\}$ .

Now, it is possible to know the difference between models. While all  $(b-1)!$  models have the same  $F_w(x;1,b)$ , each has a different sample function mean. This is a very important point, since it explains that the cumulative distribution function  $F_w(x;1,b)$  of  $W_{1,b}$  does not define the evolution of the cumulative damage process.

Let us consider the set of cumulative distribution functions  $F_w(x;1,j)$  for  $j=2,\dots,b$ . This set depends on  $r_j$ 's order. If the set is specified, this defines, in a unique manner, the evolution of the cumulative damage process even when in each  $F_w(x;1,b)$  set it is the same.

The above observations apply to the  $\pi_1=1$  case. Let us assume that  $\pi_1,\pi_2,\dots,\pi_k$  are different from zero or one, where  $k < b-1$ . Therefore,  $p_1,\dots,p_k$  may not be reordered without changing  $F_w(x;b)$ , since  $F_w(x;j,b)$  depends on  $p_j$  order for  $j=1,\dots,k$ . However,  $p_{k+1},\dots,p_{b-1}$  may be reordered without altering  $F_w(x;b)$ . Thus, unless all  $\pi_j$  for  $j=1,\dots,b-1$  are different from zero,  $F_w(x;b)$  does not specify in a unique manner the cumulative damage process of the model.

The most important points of this simple stationary unit jump model with probability transition matrix given by (1) are:

- starting damage distribution may be incorporated into the model;
- variable seriousness in service cycles may be incorporated into the model, changing transition  $P$  matrices in time;
- explicit shapes for  $W_b$ 's moments may be obtained;
- $r_j$ 's order does not change the cumulative distribution function of  $W_{1,b}$ ; and;
- knowing the cumulative distribution function  $F_w(x;1,b)$  does not determine the evolution of  $D_x$ .

### 3. Application of the Model with Inspection Data

Below some observations are shown on the bases of the cumulative damage model:

- time is discrete  $x=0,1,2,3,\dots$  Each  $x$  represents an actual time interval;
- damage states are finite and discrete, called  $1,2,\dots,j,\dots,b$  (failure). Each state is represented by certain thickness of the pipeline's wall;
- damage states are finite and discrete, called  $1,2,\dots,j,\dots,b$  (failure). Each state is represented by certain thickness of the pipe's wall;
- the starting damage value may be random, but in this work it is considered that the pipeline has no damage at the beginning, and because of that the

probability to be in state 1 at time  $x = 0$  is one, that is, the starting damage vector is  $p_0 = \{1, 0, \dots, 0\}$ ;

- the state in which failure occurs is given by  $p_b = 1$ , that is, failure occurs in the last state, state  $b$ , not before that, and;
- cumulative damage in an  $x$  time step is constant and does not decrease.

Data of internal pitting corrosion depths in a 155,500.34 m long, 5.6 mm thick diesel-transporting pipeline were obtained. Inspection reported 96 corrosion pits. Upon carrying out a statistic analysis, their histogram was obtained, where each class interval of the histogram is considered a damage state of the model.

Probability of each interval is obtained from dividing the number of defects of each interval into the total number of defects, therefore obtaining probability vector  $p_x = \{p_x(1), \dots, p_x(7)\}$ . Probabilities obtained from the probability density function are shown in **Table 1**.

In order obtain transition probabilities, equations resulting from developing Equation (2) for times  $x = 1, 2, 3, 4, 5, 6, 7$  are used. The starting probability vector  $p_0 = \{1, 0, \dots, 0\}$  is multiplied times the  $P$  transition matrix described in (1), raised to the power of  $x$ . Upon equaling probabilities of **Table 1** with equations involving  $p_1, p_2, p_3, p_4, p_5, p_6$ , and  $p_7$  we have:

$$p_x(1) = p_1^7 = 0.1042$$

$$p_x(2) = p_1 \left( p_1 \left( p_1 \left( p_1 \left( p_1 p_1 + q_1 p_2 \right) + q_1 p_2^2 \right) + q_1 p_2^3 \right) + q_1 p_2^4 \right) + q_1 p_2^5 + q_1 p_2^6 = 0.3542$$

upon solving above equations, the values of **Table 2** were obtained.

As can be observed, in this case it is not possible to find all the values of  $p_j$  with the histogram, since its shape leads to numerical problems.

A solution to this problem is to calculate the cumulative distribution function of data, to make an adjustment and obtain probabilities of each interval. **Figure 1** shows cumulative distribution obtained from data.

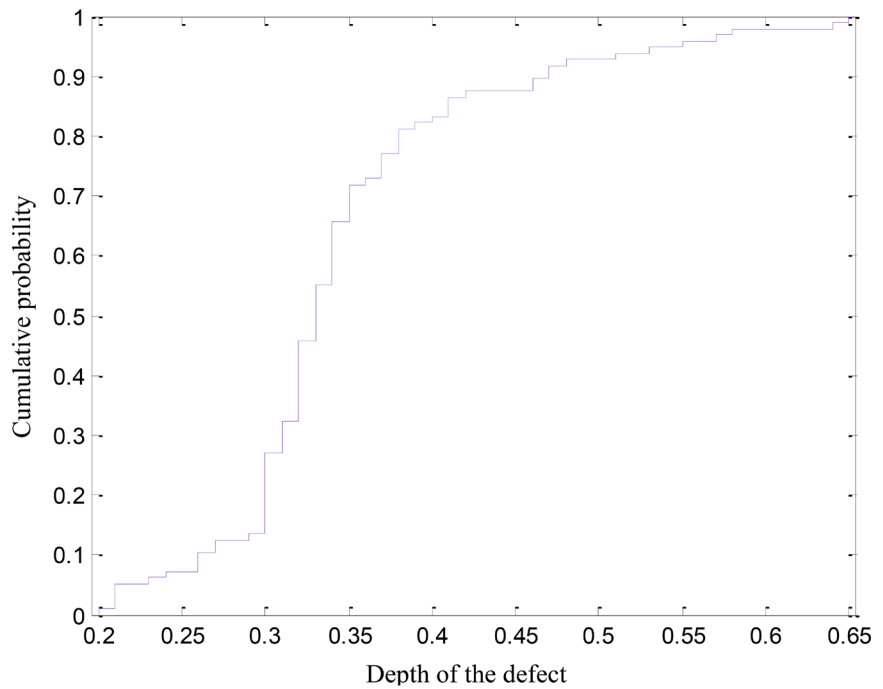
Several adjustments were made to data with various probability functions. **Figure 2** shows the various adjustments that were made. Adjustments with Lognormal and Birnbaum-Saunders distributions were found to allow solving numerical problems (see Birnbaum and Saunders, ref. [9]). **Figure 3** and **Figure 4** show a comparison of histograms with both adjustments. The Birnbaum-Saunders distribution, also known as the fatigue life distribution, is a probability distribution

**Table 1.** Probabilities of each state at time  $x = 7$ .

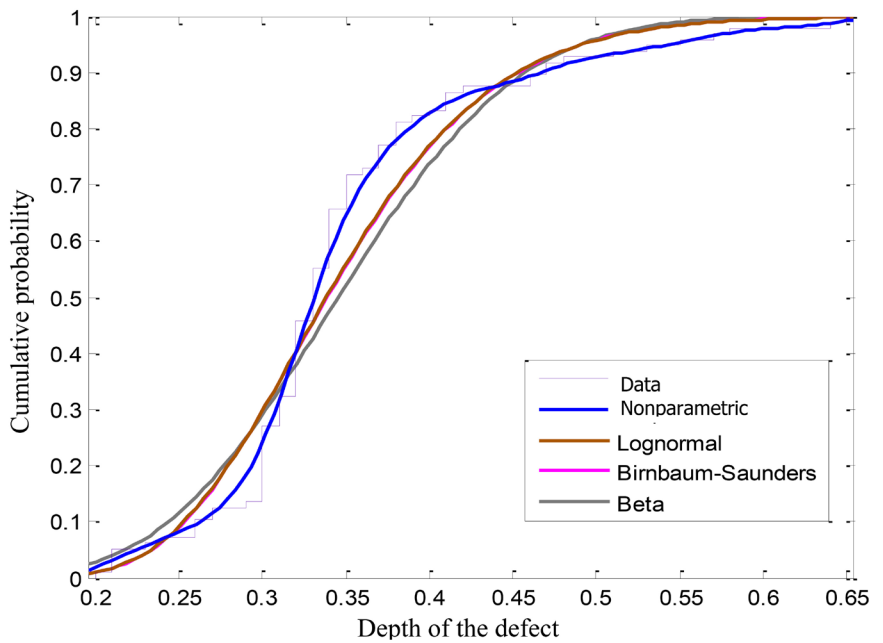
$p_x(1)$	$p_x(2)$	$p_x(3)$	$p_x(4)$	$p_x(5)$	$p_x(6)$	$p_x(7)$
0.1042	0.3542	0.3646	0.0521	0.0625	0.0417	0.0208

**Table 2.** Transition probabilities for  $P$  matrix at time  $x = 7$ .

$P_1$	$P_2$	$P_3$	$P_4$	$P_5$	$P_6$	$P_7$
0.6859	0.7450	0.7930	0.0000	0.1950	0.7040	-



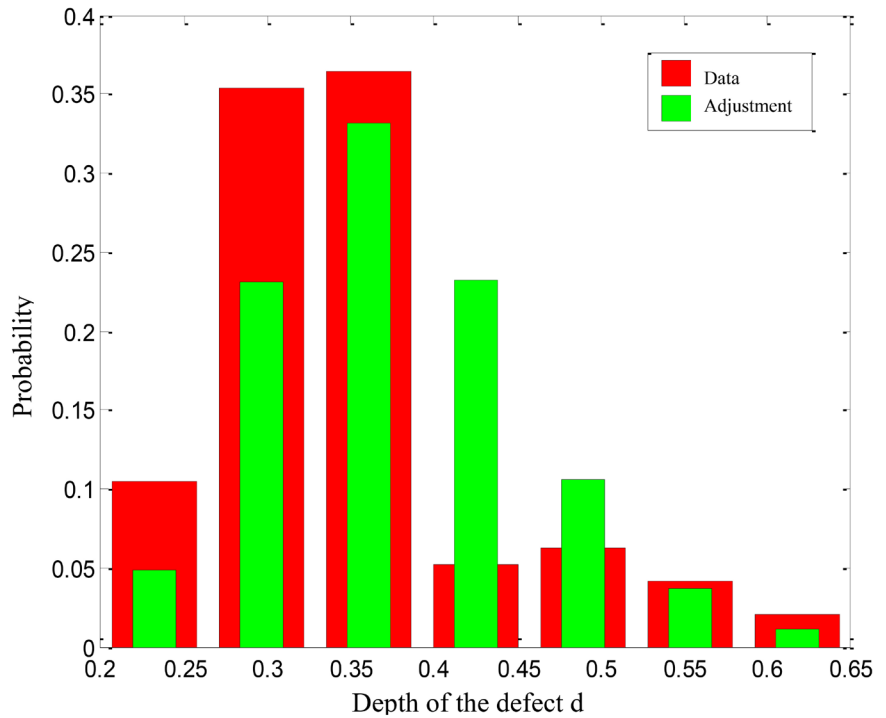
**Figure 1.** Empirical cumulative probability distribution function.



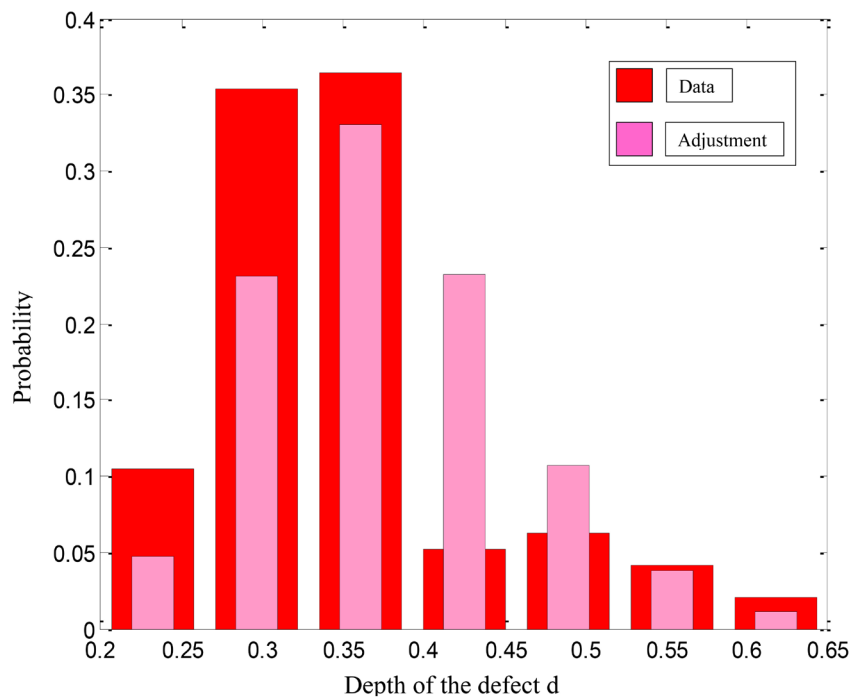
**Figure 2.** Adjustment to the cumulative probability distribution function curve.

extensively used in reliability applications to model failure times (see Birnbaum and Saunders, ref. [9]).

The Lognormal distribution was chosen to present the probability distribution function of corrosion pitting. At first sight, differences may be observed between the histograms and the probability distribution functions; however, probability has a theory to know if the set of data resulting from sampling a random variable



**Figure 3.** Comparison between measured probabilities (heavy) and with Lognormal adjustment.

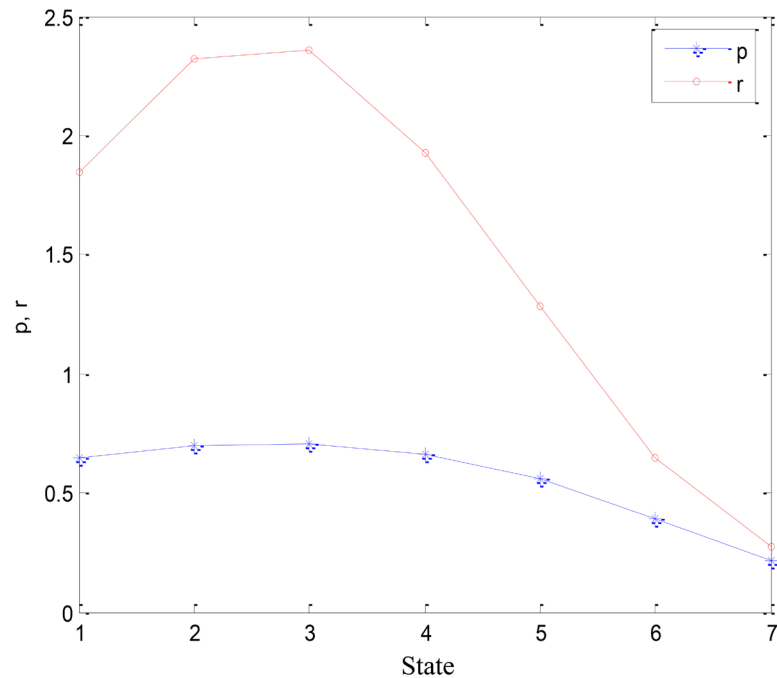


**Figure 4.** Comparison between measured (heavy) probabilities and with Birnbaum-Saunders (light) adjustment.

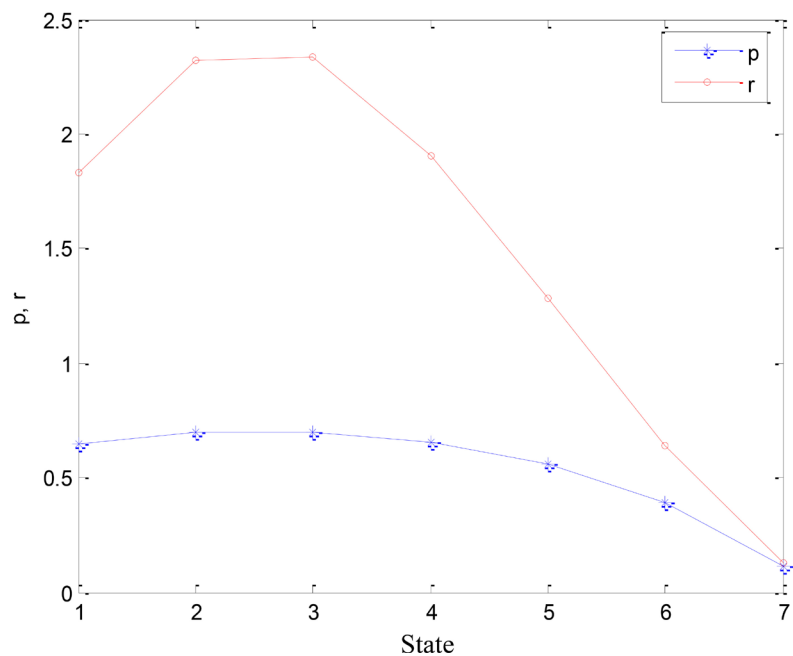
may be represented with a certain distribution function. Based on the above, it is important to know if data actually adjust to the distribution proposed with an

acceptable reliability level. To this end, a Kolmogorov-Smirnov hypothesis test was carried out. The adjustment proposed passed this test with a 1% significance level, and so it is possible to use the Lognormal distribution for the following calculations.

**Figure 5** and **Figure 6** show the evolution of transition  $p_j$  probabilities and the  $r_j$  relationship with both adjustments.



**Figure 5.**  $p$  and  $r$  variation (Lognormal).



**Figure 6.**  $p$  and  $r$  variation (Birnbaum-Saunders distribution).

The values of the probability to be in each state at time  $x = 7$  from the histogram adjusted with Lognormal function are shown in **Table 3**.

Again, the procedure to find transition probabilities is carried out. Equations resulting from multiplying the starting state times the transition matrix raised to the power of  $x = 7$  is vector

$$p_x = \{p_x(1), p_x(2), p_x(3), p_x(4), p_x(5), p_x(6), p_x(7)\}$$

Upon equating each  $p_x(j)$  to probability numerical values and solving equations, it is possible to find  $p_j$ . **Table 4** shows such values. **Table 5** shows  $r_j$  values, remembering that  $r_j = \frac{p_j}{q_j}$ .

It can be observed that, with the Lognormal adjustment, it is possible to find transition probabilities up to  $x = 7$ . Until now, necessary data has been obtained up to inspection time. Probabilities to fill the transition matrix and  $r$  ratio have been obtained. This data is essential to know the time to failure expected, since the next step is to observe and assess the trends of probabilities found to extrapolate data until failure. A way to use data obtained up to inspection time to obtain the probabilities of transition to failure is by using an equation that is a function of a parameter and adjusted to data. In this way, probabilities after inspection may be known.

Taking into account the shape of  $q_j$  evolution for various states, the following equation is proposed:

$$q_j = 1 - (1 - Y_q) e^{-\lambda(j - I_q)^2} \quad \forall j = 1, \dots, 7$$

where  $Y_q$  is the lowest value of  $q$ ,  $I_q$  is the state of the lowest value of  $q$ , and  $\lambda$  is a parameter adjusted to data.

Since the thickness of the pipe under study is 5.6 mm and the class intervals of the histogram are 0.0643, it may be deduced that the number of  $b$  states to failure must be:

**Table 3.** Probabilities of each state at time  $x = 7$ .

$p_x(1)$	$p_x(2)$	$p_x(3)$	$p_x(4)$	$p_x(5)$	$p_x(6)$	$p_x(7)$
0.0481	0.2309	0.3311	0.2319	0.1058	0.0372	0.0111

**Table 4.** Transition probabilities for  $P$  matrix at time  $x = 7$ .

$P_1$	$P_2$	$P_3$	$P_4$	$P_5$	$P_6$	$P_7$
0.6483	0.6990	0.7020	0.6580	0.5620	0.3940	0.2130

**Table 5.**  $r_j = \frac{p_j}{q_j}$  ratio at time  $x = 7$ .

$r_1$	$r_2$	$r_3$	$r_4$	$r_5$	$r_6$	$r_7$
1.8431	2.3223	2.3557	1.9240	1.2831	0.6502	0.2706

$$b = \frac{t}{\Delta x} = \frac{5.6}{0.0643} = 87.111$$

where  $t$  is the thickness of pipe equal to 5.6 mm and  $\Delta x$  is the class interval equal to 0.0643 mm.

With data known up to the inspection time,  $\lambda$ 's value is found by calculating the mean square error.  $\lambda$ 's value minimizing error is  $\lambda = 0.055$ . Knowing this value, it is possible to calculate  $q_j$  up to failure. **Figure 7** shows a comparison of  $q_j$  measured from data and that is calculated with  $\lambda$ 's function. It may be seen that there is good adjustment of the function. In this way,  $q_j$  has been calculated up to failure. **Figure 8** shows comparison of  $r_j$  variation calculated as a function of  $\lambda$  and the one calculated from data.

With the above adjustments, it is possible to obtain all the transition probabilities of passing from one state to another one, considering all states, up to failure.

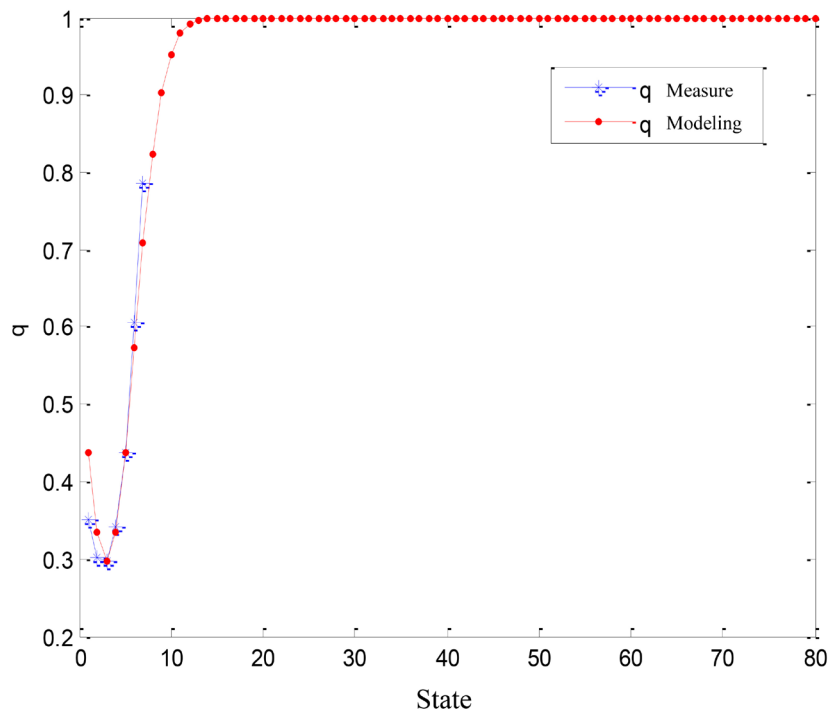
Now that we have the transition matrix, the cumulative distribution function of  $D_x$  is calculated.

$$F_D(j;x) = P\{D_x \leq j\} = \sum_{k=1}^j p_x(k) \quad \forall j = 1, \dots, b$$

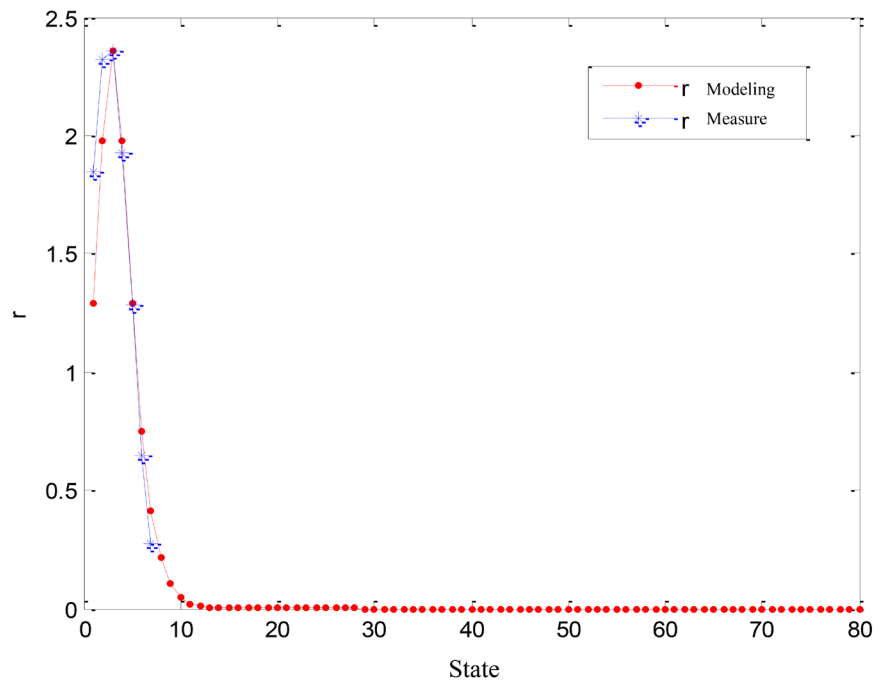
and the mean and variance of  $D_x$  are given by:

$$E\{D_x\} = \sum_{j=1}^b j p_x(j) \quad \text{y} \quad \text{var}(D_x) = \sum_{j=1}^b j^2 p_j(j) - [E\{D_x\}]^2$$

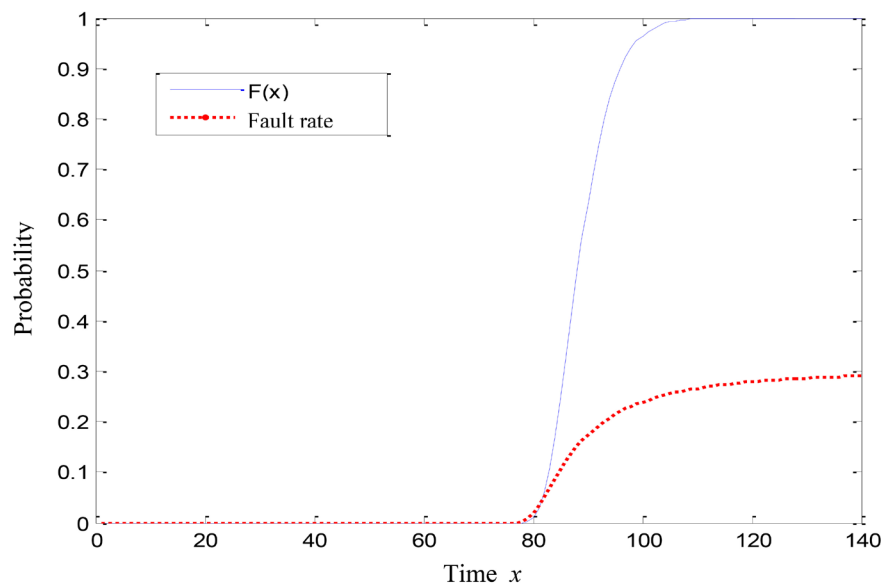
**Figure 9** shows the cumulative damage distribution function with actual data of the pipe under study, where it can be seen that it is complied with, that is to say:



**Figure 7.** Comparison of  $q_j$  calculated from data (measured) and  $\lambda$ 's function (modeled).



**Figure 8.** Comparison of  $r_i$  calculated as a function of  $\lambda$  (modeled) and calculated from data (measured).



**Figure 9.** Cumulative distribution function  $F_w(x;1,b)$  and  $h_w(x;1,b)$  of the pipe under study.

$$F_w(x;1,b) = 0 \quad \forall x = 0, 1, \dots, b-2$$

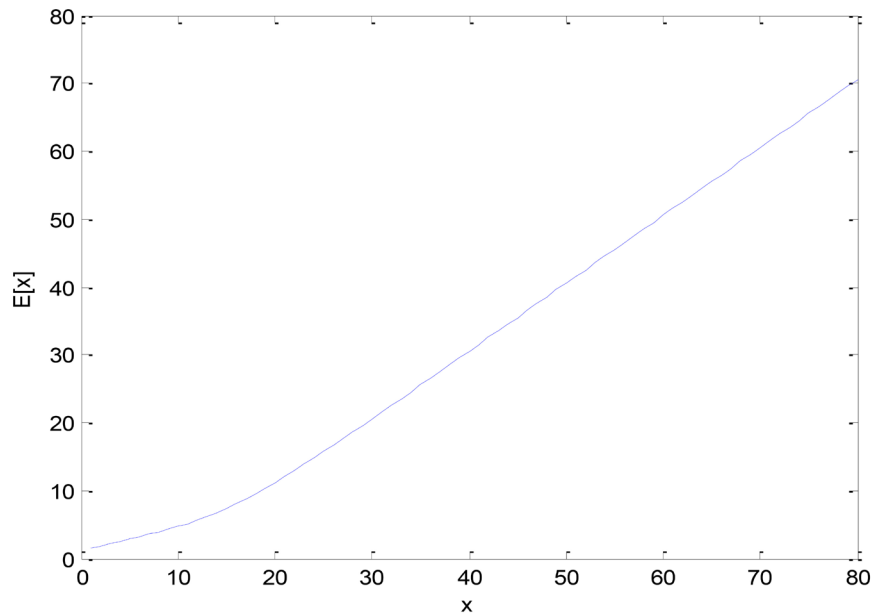
since a unit jump model may not reach state  $b$  from 1 in less than  $b - 1$  step. Based on the above, the distribution curve starts to be different from zero in state 79. It may be seen that it quickly rises and that probability of time to failure being lower or equal to 120 is one, which means, at that time, the system has



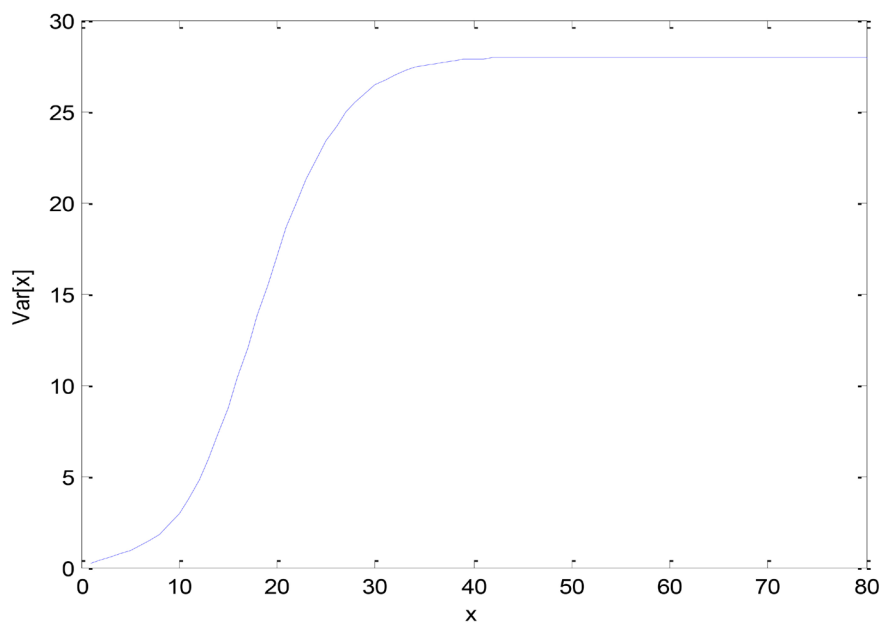
failed.

It may be observed that the failure function starts from zero, at the same time as the distribution function, since the pipeline is considered to have no damage at the beginning, and it raises until becoming constant at a 0.3 value. Based on the above, it may be said that the probability that pipeline fails, if it has worked in a given period, is low. **Figure 10** and **Figure 11** show the mean and variance of the pipe under study, respectively.

The mean shows the damage state value expected for each step of time  $x$ . In



**Figure 10.** Mean of  $D_x$  of the pipe under study.



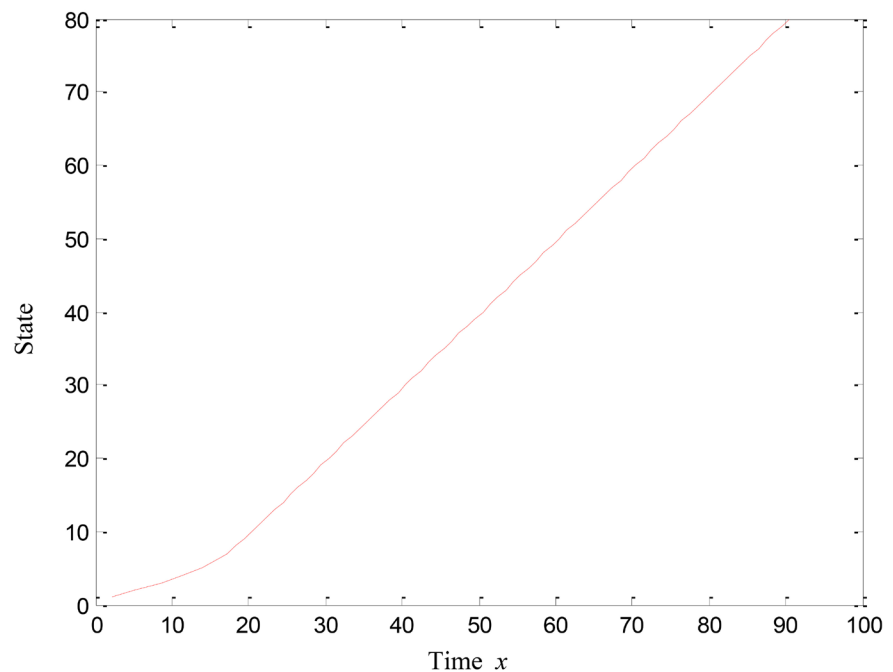
**Figure 11.** Variance of  $D_x$  of the duct under study.

the first stage, variation is not linear, but from state 20, variation of the mean is constant. Variance gives us the average of squares of deviations regarding the mean. In this case, it increases until it becomes constant in a value lower than 30, which shows that there is not a high variation and that the model is reliable. To know time to failure of the pipe, time used in state  $j$  is the random  $T_j$  variable, it being possible to apply Equation (19) with the  $r_j$  until failure, in order to know the mean time used in each state, as well as its succession. **Figure 12** shows the curve representing the mean of the sample function  $E\{T_j\} = 1 + r_j$ .

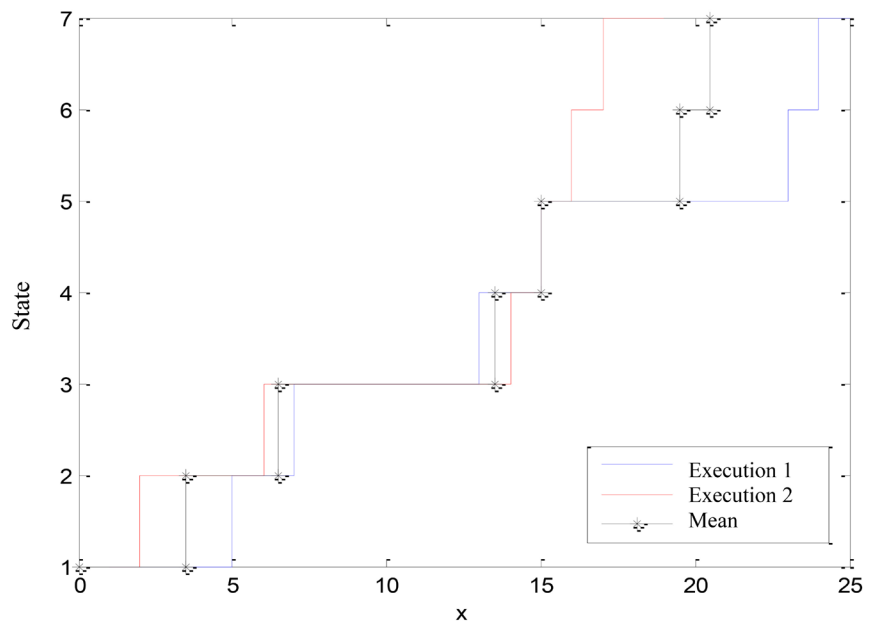
In order to prove and have an idea of how the theory works, simulations were carried out using Monte Carlo simulations. That is, a uniform distribution was used to obtain random times in each state, following a geometric distribution. The procedure includes simulating sampling functions and obtaining their mean to compare it with the theory's.

**Figure 13** shows simulation of two sampling functions and their mean up to state 7. It should be noted that the mean is, in fact, in the middle of both realizations. Continuing in this way, **Figure 14** shows five realizations, considering the seven damage states.

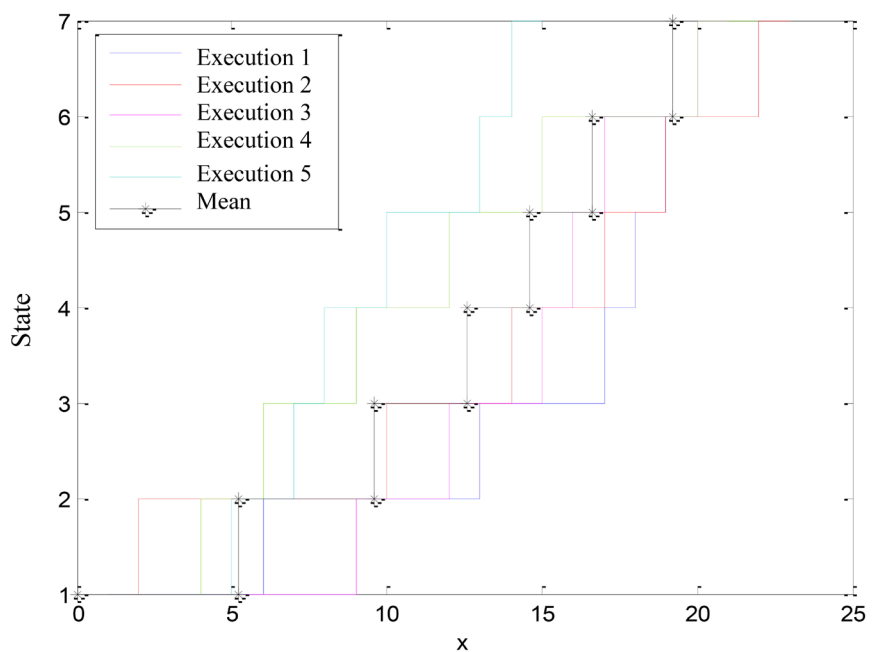
**Figure 15** shows 1000 sampling functions simulated and their mean in black. In addition, time expected was obtained with Equation (19), shown in red. It was observed that the difference between the mean obtained from realizations and the one obtained from data decreases as the number of simulations increases, because of this it can be said that the method is consistent. **Figure 16** shows the comparison of the time mean in each state using (19) and the Monte Carlo model using 1000 sampling functions and considering the 80 damage states. **Figure 16**



**Figure 12.** Evolution of  $E\{T_j\} = 1 + r_j$  calculated in function of  $\lambda$  (modeled).

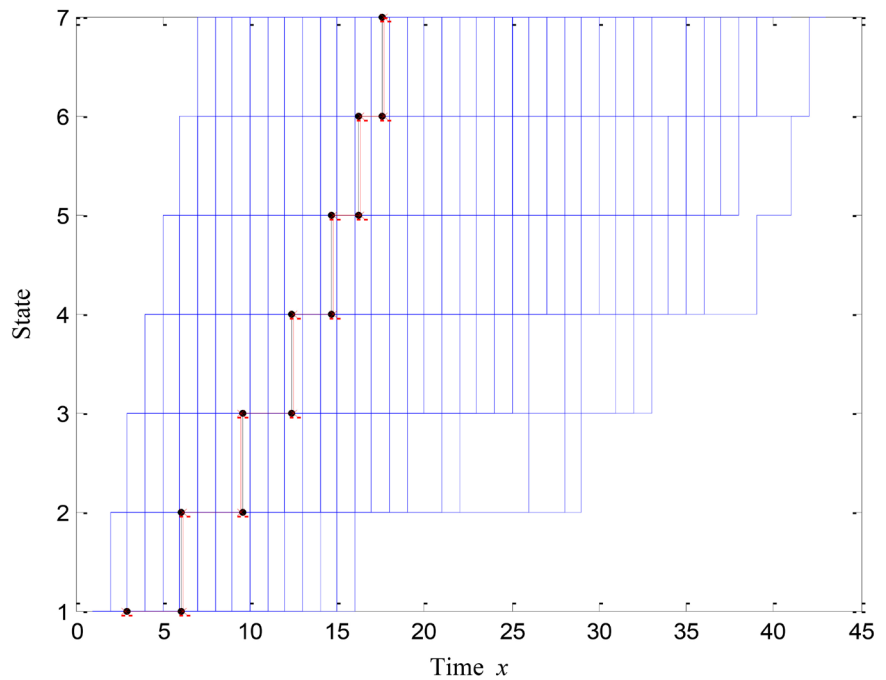


**Figure 13.** Two sampling functions simulated with Monte Carlo and its mean, considering the seven damage states measured.

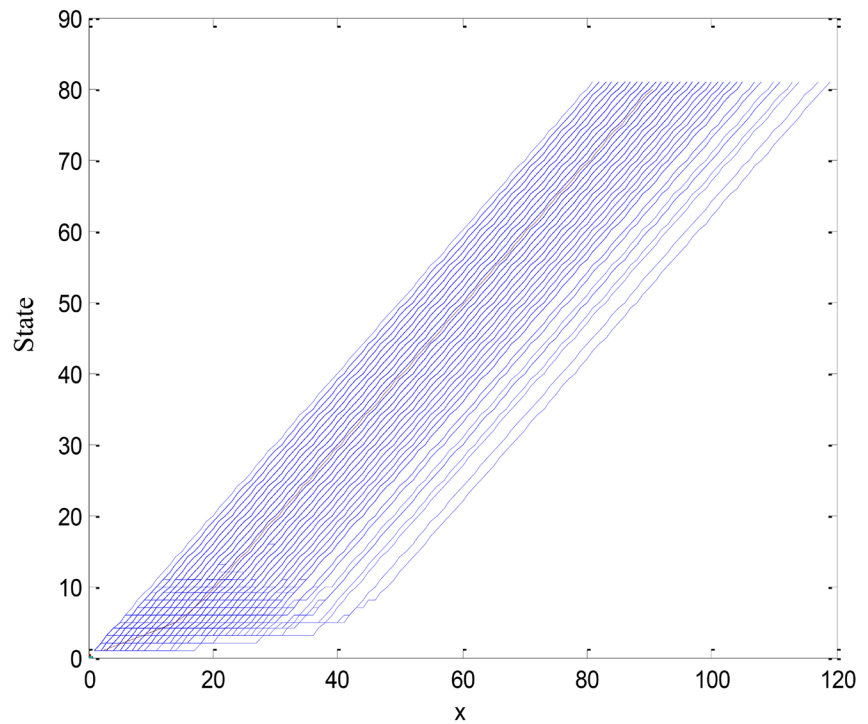


**Figure 14.** Five sampling functions simulated with Monte Carlo and its mean, considering the seven damage states measured.

shows that the mean lines are virtually the same, so it can be said that simulations and the theory give similar results. In this way, time in which the mean of the sample function reaches  $b$  may be obtained, in this case  $E\{W_{1,b}\}$ , that is, time to failure expected. In this way, for instance, to reach state 20, mean time expected is  $E\{W_{1,20}\} = 30$  time steps, and to reach failure  $E\{W_{1,b}\} = 90$  time steps.



**Figure 15.** Comparison of  $E\{T_j\} = 1 + r_j$  (red) with the mean of a model of M (black) using 1000 sampling functions (blue) considering the seven damage states measured.



**Figure 16.** Comparison of  $E\{T_j\} = 1 + r_j$  (red) with the mean of a Monte Carlo model (black) using 1000 sampling functions (blue) considering the 80 damage states measured.

Until now, the time steps to failure of the pipeline are known. It should be noted that throughout calculation and analysis, time steps  $x = 1, 2, 3, \dots$  have

been considered; however, it should also be considered that each time step used until now has an equivalence in actual time.

It should be remembered that  $x = 7$  was considered to realize equations  $p_x = p_0 P^x$ ; however, this number of states was arbitrary and may vary in accordance with certain considerations, depending on each data set. Therefore, it is better to obtain it directly from data. The mean and its state are calculated. Then, the time value expected to that state is also calculated. The mean of this data set is 0.3767 and is in the third state. Upon calculating the state of the mean on the vertical axis of **Figure 15** and projecting it on the horizontal axis, it is found that time elapsed to reach that state is  $x = 12.5$ . Relating actual time of four years, in which the inspection was carried out with theoretical time  $x = 12.5$ , it is obtained that a time step is equivalent to 0.32 years of actual time. With this data, it is possible to know the actual time to failure. In accordance with the analysis, the mean theoretical time to failure is  $E\{W_{1,b}\} = 90$  time steps, and in actual time it is  $90 \times 0.32 = 28.8$  years. In this manner, it is possible to know the mean time to failure of any pipeline having an inspection report, using Markov chains theory.

#### 4. Conclusions

The spread of corrosion shown in **Figure 2** represents the expected evolution of pitting corrosion in the pipe. Note that, in its first stage, the corrosion rate is low; it then increases, before becoming constant until failure. Its explanation is due to there being an average of time that it stays in each state, and since the system starts from there being no damage at the beginning; it can be said that the nucleation time is implicit. Nucleation is the time at which the corrosion cavity begins to form on a microscopic level. In light of the fact that it is an average and it is assumed that the system starts without damage; it takes some time for all of the cavities to be generated. After, there is another stage in which it remains constant. This is because once the cavities have been generated and the film formed, the corrosion rate remains constant, ref. [10].

According to the phenomenon of localized corrosion—as described in the literature, ref. [2]—the behaviour of the corrosion rate is due to the fact that, at an early stage, the part of the metal attacked is completely exposed to a high concentration of corrosive species, while later, a layer of oxides is formed with passive properties between the metal and the corrosive environment, thus serving as a protective barrier against corrosion, ref. [8]. This shows that the corrosion process in a pipe is almost stable if there are no sudden or radical changes that affect the kinetic process, ref. [7].

The model presented here, represents the evolution in time of damage caused by internal corrosion pitting in a hydrocarbon transport pipeline. It allows us to know, in detail, the spread of the corrosion depth at different moments from the time from the data of an inspection report. It is also possible to obtain a transition matrix from a single inspection report, as well as the fact that adjusting the

data to a normal log probability distribution avoided numerical problems. That said, it is essential to perform a Kolmogorov-Smirnov test at a 1% level of significance, indicating whether the normal Log distribution can be used with an acceptable level of reliability.

The comparison of the time used in each damage state, according to the theory of accumulated damage and with the Monte Carlo simulations, provides similar values, so it can be affirmed that the model is consistent.

With the model presented it is possible to calculate the expected time to the failure of any pipe that has an inspection report. This can be very useful for the industry to schedule inspections and maintenance.

## Acknowledgements

This article and its corresponding research were carried out thanks to the resources of the research projects IPN SIP-20120585. The collaboration of the Flores Méndez and Cortes Yah is appreciated, and the authors thank the reviewers for carefully reading the paper and for their constructive comments and suggestions which have improved the paper.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Shibata, T. (1996) W.R. Whitney Award Lecture: Statistical and Stochastic Approaches to Localized Corrosion. *Corrosion Science*, **52**, 813-830. <https://doi.org/10.5006/1.3292074>
- [2] Frankel, G.S. (1998) Pitting Corrosion of Metals. A Review of the Critical Factors. *Journal of the Electrochemical Society*, **145**, 2186-2198. <https://doi.org/10.5006/1.3292074>
- [3] Bogdanoff, J.L. and Kozin, F. (1985) Probabilistic Models of Cumulative Damage. John Wiley and Sons, USA.
- [4] Provan J.W. and Rodríguez, E.S. (1989) Development of a Markov Description of Pitting Corrosion. *Corrosion*, **45**, 178-192. <https://doi.org/10.5006/1.3292074>
- [5] Hong, H.P. (1999) Application of the Stochastic Process to Pitting Corrosion. *Corrosion Science*, **55**, 10-16. <https://doi.org/10.5006/1.3283958>
- [6] Valor, A., Caleyó, F., Alfonso, L., Rivas, D. and Hallen, J.M. (2007) Stochastic Modeling of Pitting Corrosion: A New Model for Initiation and Growth of Multiple Corrosion Pits. *Corrosion Science*, **49**, 559-579. <https://doi.org/10.5006/1.3283958>
- [7] Alamilla, J.L. and Sosa, E. (2008) Stochastic Modelling of Corrosion Damage Propagation in Active Sites from Field Inspection Data. *Corrosion Science*, **50**, 1811-1819. <https://doi.org/10.1016/j.corsci.2008.03.005>
- [8] Engelhardt, G. and Macdonald, D. (2004) Unification of the Deterministic and Statistical Approaches for Predicting Localized Corrosion Damage. I. Theoretical Foundation. *Corrosion Science*, **46**, 2755-2780. <https://doi.org/10.1016/j.corsci.2004.03.014>

- [9] Birnbaum, Z.W. and Saunders, S.C. (1969) A New Family of Life Distributions. *Journal of Applied Probability*, **6**, 319-327. <https://doi.org/10.2307/3212003>
- [10] Casanova del Angel, F. and Toquiantzi Butrón, R. (2008) Corrosion Phases of Structural Shapes Exposed to the Atmosphere. *Corrosion Science*, **50**, 2288-2295. <https://doi.org/10.1016/j.corsci.2008.05.015>