CHAPTER 3

CORROSION MODELING

3.1 Corrosion under Insulation

CUI refers to external corrosion of piping and vessels fabricated from carbon manganese, low alloys and austenitic stainless steel that occurs underneath externally clad/jacketed insulation due to the penetration of water (European Federation of Corrosion, 2008). According to API 571 (API, 2003), CUI is described as “corrosion of piping, pressure vessels and structural components resulting from water trapped under insulation or fireproofing”.

By definition, corrosion is the loss of material as a result of chemical reaction between a metal or metal alloy and its surroundings (Jones, 1996). In other words, corrosion is an electrochemical process and with ferrous materials, the corrosion process continuously develops and forms iron oxide, which lacks the strength of the original metal component. It is changed back into a material similar to iron ore. The consequences are material weakening and subsequent loss of strength and stability in load bearing components.

Similarly, CUI is also considered as an electrochemical process that involves the transfer of electrically charged ions between the anode and cathode through the pore fluid of the insulation. The principles of the electrochemical corrosion for a basic corrosion cell require the same components as the electrolytic cell which includes the anode, the cathode and an electrolyte. In order for corrosion to occur, both anode and cathode must be connected in a manner that permits electron flow.

The electrochemical process of corrosion involves oxidation at the anode and reduction at the cathode as illustrated in Figure 3.1. The site where the base metal corrodes is called the anode. Metallic iron (Fe) from the steel oxidizes to produce ferrous ions and electrons are released according to Eq. (3.1) (Fontana, 1986).

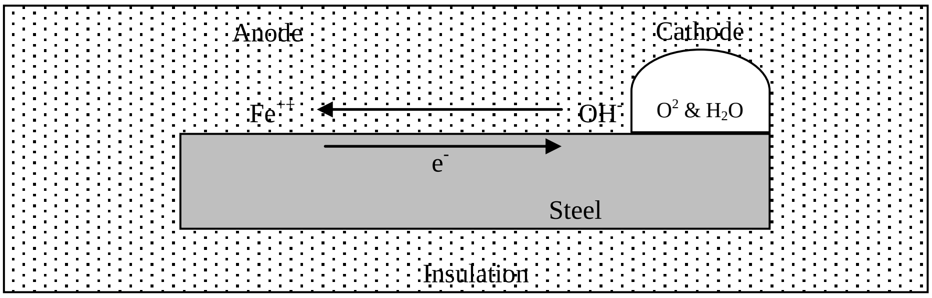

Figure 3.1: CUI mechanism - Corrosion cell in carbon steel covered by insulation

$$\text{Anodic reaction:} \quad Fe \leftrightarrow Fe^{2+} + 2e^- \qquad (3.1)$$

In order to maintain equilibrium of charges, an electrochemical reduction occurs at the cathode. In an acidic medium, the reaction taking place at the cathode is the reduction of hydrogen ions to hydrogen. However, insulation is highly alkaline (pH 7 to pH 11) and usually has a sufficient supply of oxygen and water to form hydroxyl ions, as displayed in Eq. (3.2):

$$\text{Cathodic reaction:} \quad O^2 + 2H_2O + 4e^- \leftrightarrow 4(OH)^- \qquad (3.2)$$

The current drives both the anodic and cathodic reactions to flow through a medium termed the electrolyte. The electrolyte conducts current primarily through ionic diffusion and must have specific minimum ion content and minimum water content to allow the flow of ions. In the case of CUI, the pore water in insulations acts as the electrolyte as a result of rain or even moisture condensed from the air. The combination of the anode and cathode processes results in the equations that transform the metallic iron (Fe) into hydroxides (rust) as shown in Eq. (3.3):

$$Fe + ½\, O_2 + H_2O + 2e^- \leftrightarrow Fe^{2+} + 2(OH)^- + 2e^- \qquad (3.3)$$

Eq. (3.3) can be simplified to Eq. (3.4) as follows

$$Fe + \tfrac{1}{2} O_2 + H_2O \leftrightarrow Fe^{2+} + 2(OH)^- \quad (3.4)$$

The Fe^{2+} cation combines with the hydroxyl ions $(OH)^-$ to form a fairly soluble ferrous hydroxide, $Fe(OH)_2$, which is rust that possesses a whitish appearance. The reaction is shown in Eq. (3.5). With sufficient oxygen, $Fe(OH)_2$ is further oxidized to form rust that has a reddish brown appearance.

$$Fe^{2+} + 2(OH)^- \leftrightarrow Fe(OH)_2 \quad (3.5)$$

For the transformation of metallic iron to rust to occur, all three of the following conditions must take place: (1) Iron must be available in a metallic state at the surface of steel; (2) During the anode process, oxygen, and moisture must be available; (3) During the cathode process, the electrical resistivity in the insulation must be low to facilitate electron to flow through the metal from anodic to cathodic areas.

3.1.1 Factors for Corrosion under Insulation

The triggering factor for CUI is always due to the presence of moisture. Three factors are necessary for CUI to occur:

1. Water

Water is the key point for corrosion to occur. Ordinarily, iron or steel corrodes in the presence of both oxygen and water, and corrosion does not take place in the absence of one of these factors (Schweitzer, 1989). Water normally contains dissolved oxygen (i.e. oxygen that is dissolved in water) and this dissolved oxygen may introduce corrosive environment. When the free oxygen dissolved in water is removed, the water is practically non-corrosive. If water is practically maintained neutral or slightly alkaline, it will also be non-corrosive to steel (Schweitzer, 1989). However, water that ingresses inside the insulation may contain chemical and acidic solution.

Normally, water can be introduced from two sources, external and internal. Water infiltrates from external sources such as rainfall, steam discharge, spray fire sprinkles or drift from cooling tower. External water enters an insulation system through breaks or damages of the insulation which can happen during insulation storage and/or installation, through ineffective waterproofing, through maintenance or through service lapses.

Even if the external sources are eliminated, water can still be introduced in the insulated system by the internal sources such as internal system leaks (e.g. water leak and steam tracing leak) or condensation. Condensation occurs when temperature of the metal surface is lower than the atmospheric dew point and causes poultice to trap in between metal and insulation as illustrated in Figure 3.2. It is created when the temperature and the dew point of the air have become the same, or nearly the same.

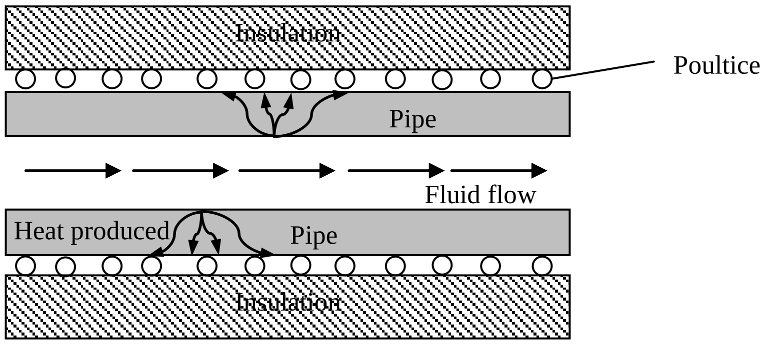

Figure 3.2: Illustration for water being introduced by internal sources in insulated systems

2. Chemical content of water

Chemical content of water plays an important factor for CUI to take place. Chlorides may be introduced by rainwater, plant and cooling tower atmospheres, misty sea (or road salt) environments or even portable water often used for fire fighting, deluge testing or wash downs. Besides, traditional thermal insulation materials contain chlorides (Corrosion under insulation, n.d.). If they are exposed to moisture, chlorides released may form a moisture layer on the pipeline surface, resulting in corrosion (i.e. pitting/stress corrosion cracking). Therefore, the quality of the materials used for the insulation has to be controlled in a way that these materials will not contain certain ‘‘acids’’ that can reduce the pH. “As the pH drops below 4, corrosion climbs dramatically. Such acidic corrosion is especially common with carbon steel. Consequently, quality assurance requirements often limit the pH of insulation to the neutral/alkaline range 7.0 to 11.7” (Corrosion, n.d).

3. Temperature

Operating temperature also contributes to CUI. According to API 581, equipment or piping systems operating in the temperature range between -12°C and 121°C are more susceptible to CUI, with temperature range of 49°C to 93°C being the most severe environment. API 581 and API 571 also provide several general guidelines as follows:

- Service temperatures between 0°C and 100°C allow water to exist as a liquid. Within this temperature range, the corrosion rate doubles for every 15°C and 20°C temperature increase. The maximum corrosion potential generally lies between these two extremes.
- As a general rule, plants located in areas with high annual rainfall or warmer, marine locations are more prone to CUI than plants located in cooler, drier, mid-continent locations.
- Regardless of the climate, units located near cooling towers and steam vents are highly susceptible to CUI.
- CUI is particularly aggressive where operating temperatures cause frequent or continuous condensation and re-evaporation of atmospheric moisture.
- Carbon steel systems that normally operate in-service above 121°C but are in intermittent service or are subjected to frequent outages.
- Cold service equipment consistently operating below the atmospheric dew point.
- Two temperature-corrosion conditions are of special note:
    - Cyclical temperatures which accelerate corrosion. For example, in regeneration process, the equipment or piping systems operate in cyclic operating temperature such as operating at 300°C and during normal condition it operates at ambient temperature; it is most likely that CUI will be triggered. Here, the warm temperature normally results in more rapid evaporation of

moisture and reduced corrosion rates. However, a surface that is covered with insulation will create an environment that holds in the moisture instead of allowing evaporation.

- The lack of temperature extremes during extended plant shutdowns, where water accumulates without freezing or evaporating.

3.2 Piping Systems Inspection Strategy

3.2.1 Inspection Strategy based on API 570

According to API 570, the inspection frequency for piping system shall be established and maintained using the following criteria:

a. Corrosion rate and remaining life calculations.
b. Piping service classification.
c. Applicable jurisdictional requirements.
d. Judgment of the inspector, the piping engineer, the piping engineer supervisor, or a corrosion specialist, based on operating conditions, previous inspection history, current inspection results, and conditions that may warrant supplemental inspections covered in Section 5.4.5 in API 570.

Inspection intervals for thickness measurements shall be scheduled based on the calculation of not more than half the remaining life determined from corrosion rates or at the maximum intervals suggested in Table 3.1 whichever is shorter. Table 3.2 describes the meaning of each piping class.

Table 3.1: Recommended maximum inspection intervals for piping systems (API, 2001)

Type of circuit	Thickness measurements	Visual external
Class 1	5 years	5 years
Class 2	10 years	5 years
Class 3	10 years	10 years
Injection points	3 years	By piping class
Soil-to-air interfaces	Not applicable	By piping class

Table 3.2: Description for the piping class (API, 2001)

Class	Description	Examples
1	Services with the highest potential of resulting in an immediate emergency if a leak were to occur is in Class 1. Such an emergency may be safety or environmental in nature.	• Flammable services that may auto-refrigerate and lead to brittle fracture. • Pressurized services that may rapidly vaporize during release, creating vapors that may collect and form an explosive mixture, such as C2, C3, and C4 streams. • Hydrogen sulfide (greater than 3 percent weight) in a gaseous stream. • Anhydrous hydrogen chloride. • Hydrofluoric acid. • Piping over or adjacent to water and piping over public throughways.
2	Services not included in other classes are in Class 2. This classification includes the majority of unit process piping and selected off-site piping.	• On-site hydrocarbons that will slowly vaporize during release. • Hydrogen, fuel gas, and natural gas. • On-site strong acids and caustics.
3	Services that are flammable but do not significantly vaporize when they leak and are not located in high-activity areas are in Class 3. Services that are potentially harmful to human tissue but are located in remote areas may be included in this class.	• On-site hydrocarbons that will not significantly vaporize during release. • Distillate and product lines to and from storage and loading. • Off-site acids and caustics.

The remaining life shall be calculated as follow:

$$\text{Remaining life (in years)} = \frac{t_{\text{actual}} - t_{\text{required}}}{\text{corrosion rate}} \qquad (3.6)$$

where t_{actual} = the actual thickness measured at the time of inspection for a given location or component (in inches or mm), t_{required} = the required thickness at the same location or component as the t_{actual} measurement computed by the design formulas (e.g., pressure and structural) before corrosion allowance and manufacturer's tolerance are added (in inches or mm).

The long-term (LT) corrosion rate of piping circuits shall be calculated from the following formula:

$$\text{Corrosion rate (LT)} = \frac{t_{\text{initial}} - t_{\text{actual}}}{\text{years in service}} \tag{3.7}$$

The short term (ST) corrosion rate of piping shall be calculated from the following formula:

$$\text{Corrosion rate (ST)} = \frac{t_{\text{previous}} - t_{\text{actual}}}{\text{years in service}} \tag{3.8}$$

where t_{initial} = the thickness at the same location as t_{actual} measured at initial installation or at the commencement of a new corrosion rate environment (in inches or mm), t_{previous} = the thickness at the same location as t_{actual} measured during one or more previous inspections (in inches or mm).

The relationship between corrosion rate of insulated carbon steels with operating temperature and type of environment is also described by American Petroleum Institute in API 581. The type of environment is classified into three categories which are marine, temperate and arid based on the average rainfall. The marine area is defined as area having more than 1000 mm/yr of rainfall. For temperate area, the average rainfall is between 500 to 1000 mm/yr, whereas, the average rainfall for arid area is less than 500 mm/yr (Mokhtar & Che Ismail, 2008). Imperial unit used in API 581, which is mil/yr, has been converted to rounded metric unit, mm/yr, as shown in Table 3.3. According to API 581, "…the corrosion rate used in API 510 or API 570 calculation to determine the remaining life and the inspection frequency. In some cases, a measured rate of corrosion may not be available. The technical modules will provide default values, typically derives from published data or from experience with similar processes, to use until inspection results are available".

Table 3.3: CUI corrosion rate default matrix for carbon steel (API, 2000)

Temperature	Corrosion Rate (mm/yr)		
	Marine	**Temperate**	**Arid/Dry**
< -12°C	0	0	0
-12°C to 16°C	0.13	0.08	0.03
16°C to 49°C	0.05	0.03	0
49°C to 93°C	0.25	0.13	0.05
93°C to 121°C	0.05	0.03	0
> 121°C	0	0	0

External visual inspections for CUI should also be conducted at maximum intervals listed in Table 3.1 to evaluate the insulation condition such as insulation damage, missing jacketing/insulation, sealing derioration, bulging etc and shall be conducted on all piping systems susceptible to CUI. Piping systems that are known to have a remaining life of less than 10 years or that are inadequately protected against external corrosion need to be included for the NDE inspection recommended in Table 3.4.

Table 3.4: Recommended extent of CUI inspection following visual inspection (API, 2001)

Pipe class	Approximate amount of follow-up examination with NDE or insulation removal at areas with damage insulation	Approximate amount of CUI inspection by NDE at suspect area on piping systems within susceptible temperature ranges
1	75%	50%
2	50%	33%
3	25%	10%

Each piping system shall be monitored by taking thickness measurements at the thickness measurement locations (TMLs) which are the designated areas on piping systems where periodic inspections and thickness measurements are conducted (API, 2001). TMLs are specific areas along the piping circuit where inspections are to be made. The nature of the TML varies according to its location in the piping system. The selection of TMLs shall consider the potential for localized corrosion and service-specific corrosion. The followings are the specific types and areas of deterioration (API, 2001):

- Injection points
- Dead legs
- Corrosion under insulation (CUI)
- Soil-to-air (S/A) interfaces
- Service specific and localized corrosion
- Erosion and corrosion/erosion
- Environmental cracking
- Corrosion beneath linings and deposits
- Fatigue cracking
- Creep cracking
- Brittle fracture
- Freeze damage

Piping circuits with high potential consequences if failure should occur and those subject to higher corrosion rates or localized corrosion will normally have more TMLs and be monitored more frequently. TMLs should be distributed appropriately throughout each piping circuit. Figure 3.3 illustrates the typical TMLs within the injection point piping circuits

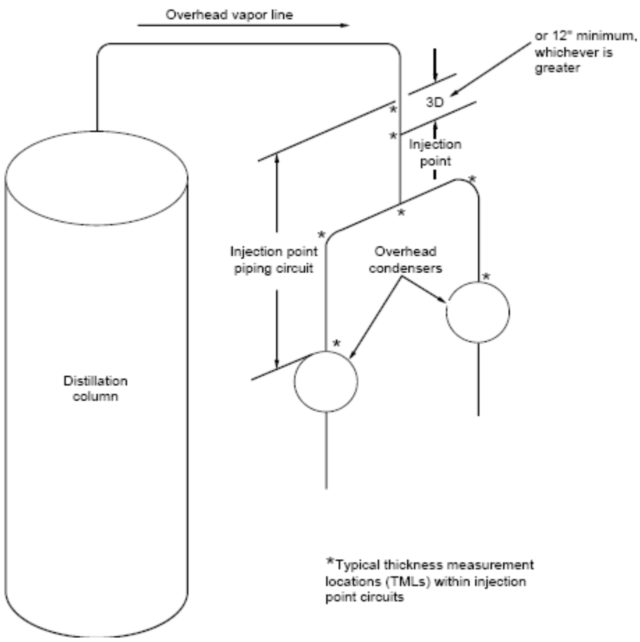

Figure 3.3: Typical injection point piping circuit (API, 2001)

The thickness at each TML can be measured by ultrasonic scanning, radiography or electromagnetic techniques. The thinnest reading or an average of several measurement readings taken within the area of a test point shall be recorded and used to calculate corrosion rates, hence assessing the remaining life. Where appropriate, thickness measurements should include measurements at each of the four quadrants on the pipe, with special attention to the inside and outside radius of elbows and tees where corrosion/erosion could increase corrosion rates.

In selecting or adjusting the number and locations of TMLs, the inspector should take into account the patterns of corrosion that would be expected and have been experienced in the process unit. In theory, a piping circuit subject to perfectly uniform corrosion could be adequately monitored with a single TML. In reality, corrosion is never truly uniform, so additional TMLs may be required. More TMLs should be selected for piping systems with any of the following characteristics:

- Higher potential for creating a safety or environmental emergency in the event of a leak.
- Higher expected or experienced corrosion rates.
- Higher potential for localized corrosion.

- More complexity in terms of fittings, branches, dead legs, injection points, and other similar items.
- Higher potential for CUI. TMLs should be established for areas with continuing CUI

3.2.2 Inspection Strategy based on Risk Assessment

Alternatively, external visual inspection intervals can be established by using RBI assessment conducted in accordance with API 581. In API 581, CUI is treated as a special case of external damage mechanism as well as a special concern due to CUI can cause failures in areas that are not normally of a primary concern to an inspection program. The next section will discuss what RBI is and how CUI is assessed in RBI methodology.

3.2.2.1 Risk-Based Inspection

For more than 100 years, it is believed that through inspection, potential failures can be detected and therefore, inspection has been regarded as an important activity in industry (Noori & Price, 2006). The traditional view of inspection is that by doing inspection, the probability of an unexpected failure should be reduced. The term 'inspection' refers to the planning, implementation and evaluation of examinations to determine the physical and metallurgical condition of equipment or a structure in terms of fitness-for-service (Wintle et al., 2001). Risk-based approach to inspection, or being well-known as RBI, is a method for using risk as a basis for managing an inspection program. The American Petroleum Institute, in the Base Resource Document API 581 (2000), defined RBI as a method for using risk as a basis for prioritizing and managing the efforts of an inspection program.

The concept of risk is used to target inspection and maintenance resources at areas of the plant where they can have the greatest effect in reducing risk, the occurrence and consequence of unplanned failures. The concept also aims to reduce the cost of unproductive inspections. Accordingly, RBI identifies 10% to 20% of items that cover 80% to 95% of the risk exposures of the equipment (Lee & Teo, 2001). RBI has been an industry standard for prioritizing inspection of static equipment, such as pressure vessels, tanks, heat exchanger, piping systems, relief valves and control valves. Although, the application of RBI is more on static equipment, it has also been applied to rotating equipment (Fujiyama et al., 2004). RBI provides many advantages, which include (1) an increase in plant availability, (2) a decrease in the number of failure occurrences, (3) a reduction in the level of risk due to failure, and (4) a reduction in the direct inspection cost of the plant (Khan et al., 2006).

In theory, risk is defined as the product of the probability of failure and its likely consequences as in Eq. (3.9):

$$\text{Risk} = \text{Probability of failure} \times \text{Consequence of failure} \tag{3.9}$$

Therefore, the main steps in RBI modeling are the estimation of the probability of failure and its consequences.

3.2.3 Principles of Failure Probability Assessment

According to Giribone and Valette (2004), the main driver for scheduling periodical inspection is the probability of failure. The probability of failure is defined as the mean frequency or rate with which the specified failure event would be expected to occur in a given period of operation, normally one year (Wintle et al., 2001). It is also defined in API 581 as the likelihood of a specific outcome, measured by the ratio of specific outcomes to the total number of possible outcomes (API, 2000). Estimating the probability of failure is an important input to RBI analysis in aiming to develop the inspection program. Giribone and Valette (2004) stated in estimating the failure probability, it is important to clearly identify the origin of the data and the expected outcome. They also outlined the situation of the various techniques in the spectrum of failure probability estimation. Below are several possibilities that can arise:

1. The expected outcome is clearly defined, and its probability of occurrence is firmly established on solid statistical grounds: e.g. failure rate of an electronic component in specified working conditions: in that case, the classical approach based on failure frequency is applicable.
2. The outcome is clearly defined but probabilities estimates can hardly rely on statistical data because situations under concern are not generic enough. This is typically the case of structural reliability as no two structures are really similar.
3. It can happen that such a procedure is not possible but that, fortunately, tests can be used. This is typically the case of medical diagnosis and industrial inspection. Several approaches may be used such as Bayesian approach.
4. It can also happen that no probability estimates whatever shaky or approximate is available. In that case, structured expert opinion elicitation can be used and integrated in a learning process (if any) with some benefit.

The above situations are typical situations encountered in risk analysis. Besides, other possibilities are of concern:

5. It can happen that no probability estimate is available whatsoever. This case (uncertainty) is in principle out of the scope of traditional risk analysis and is to be dealt with by some special techniques such as scenario analysis.
6. It can happen that probabilities are more or less known but that the outcome itself is only but poorly defined. A typical situation is for example when the condition of a piece of equipment is only specified by a rather loose statement such as ‘not so good’, etc. In this situation one must resort to the so-called ‘fuzzy logic’, in which the degree of belonging of an attribute to a given category is itself subject to the usual rules of probability.
7. Finally, the worst situation is when both probabilities and outcomes are problematic: a typical case is the controversial global climate change or the possible harmful effect of cellular phones on health. For this state, usually referred to as ‘ignorance’, none of the strategies previously examined is relevant. In those cases, the so-called precautionary principle can be of some help.

The schematic summary of the situation is shown in Table 3.5.

Table 3.5: Risk, fuzziness and ignorance (Giribone & Valette, 2004)

<table>
<tr><th colspan="2">Knowledge about outcomes →</th><th>Outcomes well defined</th><th>Outcomes poorly defined</th></tr>
<tr><td colspan="2">Knowledge about likelihood ↓</td><td></td><td></td></tr>
<tr><td colspan="2">Firm basis for probabilities</td><td>Risk (Frequentist approach)</td><td rowspan="3">Fuzziness</td></tr>
<tr><td rowspan="2">Shaky basis for probabilities but learning process</td><td>Quantitative use</td><td>Risk (Bayesian approach)</td></tr>
<tr><td>Qualitative use</td><td>Risk (Elicitation of expert opinions)</td></tr>
<tr><td colspan="2">No basis for probabilities</td><td>Uncertainty</td><td>Ignorance</td></tr>
</table>

To relate the above mentioned principles to CUI scenario, it is found that the expected outcome for CUI is clearly defined. However, its knowledge about likelihood is rather shaky since it cannot be firmly established on solid statistical grounds due to lack of failure data. Therefore, possibility 2 and 3 are the closest to describe the knowledge about CUI. In this study, the structural reliability analysis will be explored (possibility 2). For possibility 3 where a test can be used if a procedure is impossible, the degradation analysis will be investigated.

3.2.4 Failure Probability Approaches in RBI Methodology

There are different approaches, from qualitative to quantitative, to assess the failure probability in RBI methodology. In qualitative failure probability assessment, the probability of failure is primarily based on engineering judgments made by experts. The failure probability is described using terms such as very unlikely, unlikely, possible, probable or highly probable where subjective scores are assigned to different factors which are thought to influence the probability of failure. Criteria for the descriptive categories should be defined to ensure that this term will be used consistently.

In semi-quantitative failure probability assessment, the probability of failure should generally be more numerically based and detailed than the qualitative approach, but still contain a large element of engineering judgments. The common method is based on the guidelines by American Petroleum Institute, API 581 Risk-based Inspection Base Resource Document (API, 2000) which will be discussed in detail in the next section.

In a fully quantitative failure probability assessment, the approach is to statistically estimate the failure probability based on the actual data collected such as historical failure data and/or inspection data. Using this analytical approach, the numerical data are then analyzed using suitable models such as using a mathematical model, logic flow diagram and others.

3.2.5 Quantitative Failure Probability Assessment in API 581

It is important to understand the quantitative approach used to assess the failure probability using the standard API 581 because the expected results generated from the proposed models will be compared to those generated using the standard. The probability of failure analysis in API 581 is calculated using Eq. (3.10):

$$\text{Frequency}_{\text{adjusted}} = \text{Frequency}_{\text{generic}} \times \text{FE} \times \text{FM} \qquad (3.10)$$

where $\text{Frequency}_{\text{adjusted}}$ is the adjusted failure frequency, $\text{Frequency}_{\text{generic}}$ is the generic failure frequency, FE is the equipment modification factor and FM is the management systems evaluation factor.

It begins with a database of generic failure frequencies which is based on a compilation of available records of equipment failure history from a variety of sources. For equipment item that has not operated long enough to experience a failure, it is necessary to turn to larger equipment pool to find enough failures to provide a reasonable estimate on the true failure probability. This generic equipment pool is used to produce a generic failure frequency. Generic failure frequencies have been developed using records from all plants within a company or from various plants within an industry, from literature sources, past reports, and commercial data bases for each type of equipment and each diameter of piping. A generic database is presented in Table 3.6.

Table 3.6: Suggested generic failure frequencies for piping systems (API, 2000)

Piping Size	Leak frequency (per year for four hole sizes)			
	¼ in.	1 in.	4 in.	Rupture
Piping 0.75 in. diameter, per ft	1×10^{-5}			3×10^{-7}
Piping 1 in. diameter, per ft	5×10^{-6}			5×10^{-7}
Piping 2 in. diameter, per ft	3×10^{-6}			6×10^{-7}
Piping 4 in. diameter, per ft	9×10^{-7}	6×10^{-7}		7×10^{-8}
Piping 6 in. diameter, per ft	4×10^{-7}	4×10^{-7}		8×10^{-8}
Piping 8 in. diameter, per ft	3×10^{-7}	3×10^{-7}	8×10^{-8}	2×10^{-8}
Piping 10 in. diameter, per ft	2×10^{-7}	3×10^{-7}	8×10^{-8}	2×10^{-8}
Piping 12 in. diameter, per ft	1×10^{-7}	3×10^{-7}	3×10^{-8}	2×10^{-8}
Piping 16 in. diameter, per ft	1×10^{-7}	2×10^{-7}	2×10^{-8}	2×10^{-8}
Piping > 16 in. diameter, per ft	6×10^{-8}	2×10^{-7}	2×10^{-8}	1×10^{-8}

If enough data were available for given equipment item, true failure probabilities could be estimated from actual observed failures. However, if data is null, the RBI method recommends a generic failure frequency to be used to “jump start” the failure probability analysis. A data source should be chosen that represents plants or equipment similar to the equipment being modeled. For example, much high-quality generic data can be derived from nuclear power plant databases; however, the data may not be appropriate to be applied in refineries or petrochemical plants because of the differences in maintenance and inspection quality, and the nature of the service.

These generic frequencies are then modified by two factors, the equipment modification factor (FE) and the management systems evaluation factor (FM), to yield an adjusted failure frequency, as follows:

1. The equipment modification factors (FE) examines details on each equipment item and to the environment in which that item operates, in order to develop a modification factor unique to that piece of equipment. FE includes:
   - the technical module that examines materials of construction, the environment and inspection program
   - universal conditions that affect all equipment items at the facility
   - mechanical considerations that vary from item to item
   - process influences that can affect equipment integrity

2. The management systems evaluation factor (FM) adjusts for the influence of the facility's management system on the mechanical integrity of the plant. FM is used to describe direct impact that inspection, maintenance, process and safety personnel have on the equipment failure frequency. This adjustment is applied equally to all equipment items.

In API 581, Technical Modules is a systematic method used to assess the effect of specific failure mechanisms on the probability of failure (generic failure frequency). The Technical Modules serve four functions:

1. Screen the operation to identify the active damage mechanisms.
2. Establish a damage rate in the environment.
3. Quantify the effectiveness of the inspection program.
4. Calculate the modification factor to apply to the generic failure frequency

The Technical Module evaluates two categories of information which are (1) the deterioration rate of the equipment item's material of construction, resulting from its operating environment and (2) the effectiveness of the facility's inspection program to identify and monitor the operative damage mechanisms prior to failure. Inspection techniques required to detect and monitor one failure mechanism may be totally different from those needed for another mechanism. These differences are addressed by creating a separate Technical Module for each damage mechanism. The Technical Module for CUI follows the external damage technical module as represented in Appendix A.

3.3 Logistic Regression Model

Typically, for corrosion failure mode, the wall thickness data collected during inspection period are used to assess the probability of failure by analyzing the data statistically. However, the wall thickness data is not always available for statistical methods to be used. Typically what is usually available in CUI inspection reports is the result from inspection after insulation removal which is corrosion was found and treated, or corrosion was not seen as illustrated in Figure 3.4. These types of data are classified as binary responses with 0 and 1. Binary responses can be used to predict the probability of CUI occurrence by analyzing binary data using logistic regression model (Hosmer & Lemeshow, 1989).

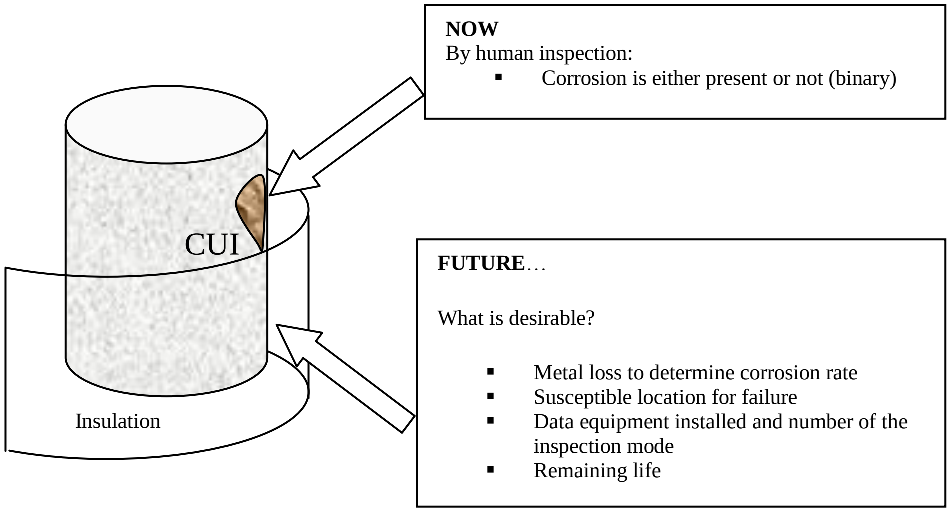

Figure 3.4: Data picture of CUI for logistic regression model

In statistics, logistic regression is used for prediction of the probability of occurrence of an event by fitting data to a logistic curve. It is a generalized linear model used for binomial regression. Like many forms of regression analysis, it makes use of several explanatory variables that may be either numerical or categorical.

Prior to engaging in a study of logistic regression modeling, it is important to understand that the goal of using logistic regression for data analysis is the same as that of any model-building technique used in statistics, that is, to find the best fitting and most parsimonious model. As in regression, a logistic regression model, sometimes called a logistic model or a logit model, describes a relationship between a response and a set of explanatory variables. A response is also known as a dependent variable or an outcome. Explanatory variables are also often referred to as covariates, independent variables or predictors.

The methods employed in an analysis using logistic regression follow the same general principles used in linear regression. However, there are two main differences in logistic regression compared to linear regression. The first difference is that in logistic regression, the response variables are categorical where the values are binary, with 0 or 1, rather than continuous. 0 is classified as a ''success'' and 1 is a ''failure''.

Another major difference between logistic and linear regression is that the response variable is assumed to follow a binomial distribution rather than a normal distribution for response variable in linear regression. For observations on a categorical variable with two categories, the binomial distribution applies to the sum of the outcomes when the following three conditions hold true (Agresti, 1990):

- For a fixed number of observation n, each falls into one of two categories.
- The probability of falling in each category, π for the first category and $(1-\pi)$ for the second category, is the same for every observation.
- The outcomes of successive observations are independent; that is, the category that occurs for one observation does not depend on the outcomes of other observations.

The binomial distribution for binary random variables specifies probabilities $P(Y=1)=\pi$ and $P(Y=0)=1-\pi$ for the two outcomes. The binomial probability mass function is

$$f(Y_i, \pi_i) = \pi_i^{Y_i}(1-\pi_i)^{(1-Y_i)} \qquad (3.11)$$

where Y_i = 0 or 1 and i is the observation number. When evaluating the piping systems subject to CUI, the interest is the present or absence of CUI. Therefore, the binary response is either CUI is present (Y = 1) or CUI is not present (Y = 0).

To better explain the concept of logistic regression, the logistic function that describes the mathematics behind this regression should be defined. The logistic function $f(z)$ is as follows:

$$f(z) = \frac{1}{1+e^{-z}} \qquad (3.12)$$

The logistic function $f(z)$ ranges between 0 and 1. Plots of $f(z)$ yield an S-shaped curve resembling the cumulative distribution plot for a random variable, as illustrated in Figure 3.5.

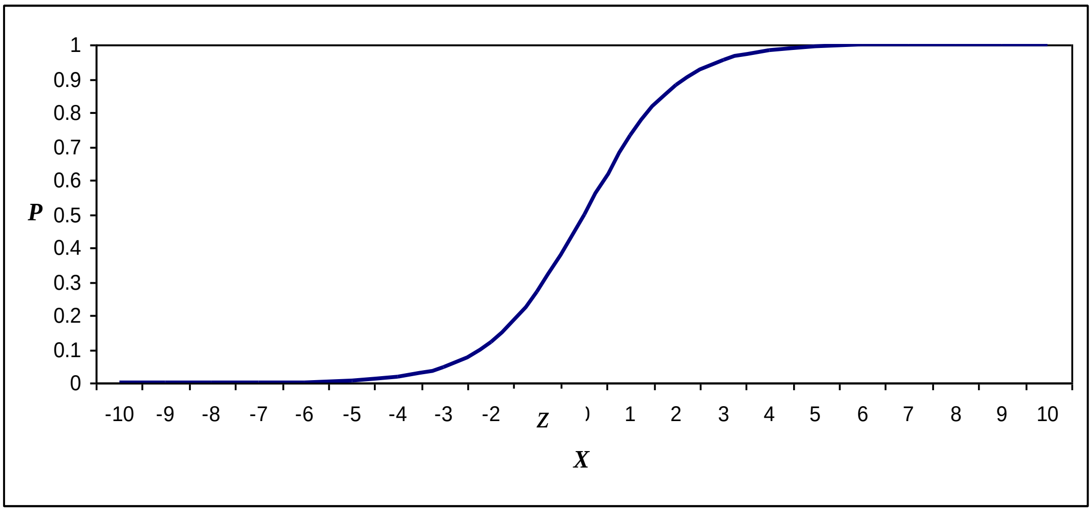

Figure 3.5: Logistic function

From the logistic function, the logistic regression model is obtained through the parameter z that can be written as the linear sum of the explanatory variables as follows:

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{3.13}$$

where $x_1, x_2, \dots x_n$ are defined as the independent variables of interest and $\beta_0, \beta_1, \beta_2, \dots \beta_n$ are the coefficient representing unknown parameters. Estimates of the parameters $\beta_0, \beta_1, \beta_2, \dots \beta_n$ are obtained using a mathematical technique called maximum likelihood. Refer to Appendix B for further explanation on maximum likelihood technique.

Newton–Raphson is an iterative method that is used to obtain parameter estimation for maximum likelihood. Basically, this concept of iteration is embedded with MATLAB software in order to solve the likelihood estimation. The Newton-Raphson concept will also be explained in Appendix B. For this concept, theoretically, it will choose initial estimates of the regression coefficients, such as $\beta_0 = 0$. At each iteration t, it will update the coefficient and this iteration will stop when the percentages of error decrease to the smallest value which approximately becomes zero.

A regression can simultaneously handle both quantitative and qualitative explanatory variables. In the logistic regression model, the response variable is a binary variable whereas the explanatory variables can be either quantitative or qualitative variables. The quantitative variables are pipe age and operating temperature and the qualitative variable is type of insulation.

The quantitative variable can be further classified as a continuous variable, one that takes any value within the limits of variable ranges. In this model, age or year of service is considered under continuous variable (i.e. 6, 10 and 15 years of services). The quantitative variable can also be considered as a categorical variable such as operating temperature. For example, the pipe having operating temperature 290°C can be categorized in group of pipes having operating temperature more than 121°C.

The qualitative variable such as types of insulation is also considered as categorical covariates. Here, dummy variable needs to be used in order to overcome the weakness of the categorical variable as it cannot be meaningfully interpreted in regression model. Dummy variables are artificial explanatory variables in a regression model whereby the dummy codes are a series of numbers assigned to indicate group. In dummy variable, it will be dichotomous variable as each variable is assumed one of two values, 0 or 1, indicating whether an observation falls in a particular group.

For dummy variable to be used, if there are K groups, one needs to have $K - 1$ dummy variables to represent K groups. Let say, if there are six operating temperature groups, one needs to have 5 dummy variables to represent the group which one of the groups will not be represented as dummy variable. It will be considered as a reference to which each of the group should be compared.

3.3.1 Wald Test

After generating the parameters $\beta_0, \beta_1, \beta_2, \ldots \beta_n$ (coefficient for each variable), it is necessary to test the statistical significance of each coefficient in the model. A Wald test calculates a *Z* statistic, which is:

$$z = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} \tag{3.14}$$

In other words, a Wald test is carried out by using the coefficient divided by its standard error (SE). The Wald tests are based on chi-square statistics that tests the null hypothesis that a particular variable has no significant effect given that the other variables are included in the model.

3.3.2 Backward Stepwise Elimination

Once a full logistic regression model is developed, the backward stepwise elimination procedure will be used to eliminate the explanatory variable with insignificant coefficient. The backward stepwise elimination procedure begins with a full model. Then, the variables which are found to be insignificant are eliminated from the model in an iterative process. The fit of the model is tested after the elimination of each variable to ensure that the model still adequately fits the data. When no more variables can be eliminated from the model, the analysis has been completed. The following steps describe the procedure:

Step 1: Assume the original model with all possible covariates is

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

Then, the following $r-1$ tests are carried out, $H_{0j}: \beta_{0j} = 0,\ \ j = 1,2,\dots,r-1$. The lowest partial F-test value F_l corresponding to $H_{0l}: \beta_l = 0$ or t-test value t_l is compared with the preselected significance values F_0 and t_0. One of two possible steps (step2a and step 2b) can be taken.

Step 2a: If $F_l < F_0$ or $t_l < t_0$, then x_l can be deleted and the new original model is

$$Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_{l-1} x_{l-1} + \beta_{l+1} x_{l+1+\cdots} + \beta_{r-1} x_{r-1} + \varepsilon$$

Go back to step 1.

Step 2b: If $F_l > F_0$ or $t_l > t_0$, the original model is the model we should choose.

Hypothetical example

Suppose the preselected significance level is α = 0.1. Thus, $F_0 = F_{1,8,0.9} = 3.14$

Step 1: The original model is

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$$

$F_3 = 0.018$ corresponding to $H_{03}: \beta_3 = 0$ is the smallest partial F value.

Step 2a: $F_l = 0.018 \leq F_0 = 3.14$

Thus, x_3 can be deleted. Go back to step 1.

Step 1: The new original model is

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_4 x_4 + \varepsilon$$

$F_4 = 1.86$ corresponding to $H_{04}: \beta_4 = 0$ is the smallest partial F value.

Step 2a: $F_4 = 1.86 < F_0 = 3.14$

Thus, x_4 can be deleted. Go back to step 1.

Step 1: The new original model is

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

$F_1 = 144$ corresponding to $H_{01}: \beta_1 = 0$ is the smallest partial F value.

Step 2b: $F_1 = 144 > F_0 = 3.14$. Thus,

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

is the selected model.

3.3.3 Kruskal Wallis Test

Kruskal–Wallis test is employed to check the sensitivity of the model. In statistics, the Kruskal–Wallis test is a non-parametric method for testing equality of population medians among 3 or more groups. Since it is a non-parametric method, the Kruskal-Wallis test does not assume a normal population. However, the test does assume an identically-shaped and scaled distribution for each group, except for any difference in medians. The test procedure is as follows (Ariaratnam et al., 2001):

1. Rank all the N observations for all the j number of the data sets (i.e. N is the total samples).

2. Let r_{ij} be the rank of all observations, y_{ij} (i = 1, 2, . . . , n and j = 1, 2, . . . , t) where i is the number of sample in one data set and j is the number of the data sets.
3. Let r_j be the sum of the rank of each data set.
4. Compute the test statistic Kruskal-Wallis using the following equation:

$$KW = \frac{12}{N(N+1)} \sum_{j=1}^{3} \frac{r_j^2}{n} - 3(N+1) \qquad (3.15)$$

   where N = total samples and n = sample size for data set/group j.
5. Under the null hypothesis H_0, Kruskal-Wallis follows approximately the chi-square distribution with $t - 1$ degrees of freedom.
6. Reject H_0 at α-level (i.e., 5% for this test) if $KW > \chi^2_{\alpha,t-1}$.

3.4 Degradation Analysis

In a situation where the mean time between failures (MTBF) data are scarce, degradation analysis is useful for the analysis of failure time distributions in reliability studies. The key to the analysis is the perceived link between the degradation measurements and the failure time. By assuming a stochastic model for the degradation, the lifetime distribution is implied. For equipment or piping systems subject to corrosion, the wall thickness is typically measured at the thickness measurement location (TMLs) during each inspection period as the collection of wall thickness data has become necessary in assessing the failure probability for systems subject corrosion using them in degradation analysis.

3.4.1 Degradation models

Degradation models vary markedly across the fields of reliability modeling. A level of degradation at which a failure is said to have occurred needs to be defined first. The use of the term *failure* in this context is defined as when the wall thickness reaches the minimum wall thickness allowed. Here, the failure can be defined as when the corrosion defect reaches a pre-specified threshold value. For piping system, the pipe failure is characterized when the pipe wall thickness reaches the minimum wall thickness specified. In assessing the reliability for systems subject to corrosion, the actual failure data are replaced by these "time-to-failure" data.

The extrapolation typically can be done using the one of the following models:

Linear model : $d(t) = at + b$ (3.16)

Exponential model : $d(t) = be^{at}$ (3.17)

Power model : $d(t) = bt^{a}$ (3.18)

Logarithmic model : $d(t) = a\ln(t) + b$ (3.19)

where $d(t)$ represents the wall thickness at time t , t represents time and a and b are model parameters to be solved for. Once the model parameters have been estimated, time-to-failure t_i can be extrapolated, which corresponds to the defined level of failure $d(t_i)$. The computed t_i at each TMLs can now be used to as the mean time-to-failure data for the life data analysis. Figure 3.6 shows the illustration.

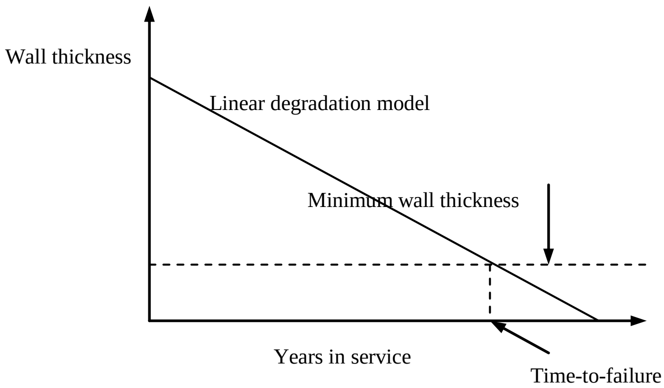

Figure 3.6: Illustration of time-to-failure using linear degradation model

For this study, a linear model is assumed for CUI degradation. Therefore, a represents the corrosion rate and b represents the pipe nominal thickness. Linear model is always becomes preferable in estimating the rate of corrosion degradation (refer to Eq. (3.7) and Eq. (3.8) which are based on the standard API 570) especially for long term prediction. Noor et al. (2007) and Yahaya, et al., (2009) also employed a linear model to assess the corrosion rate.

3.4.2 Lifetime Distribution

Life data analysis is one of the well-known engineering tools for analyzing failure data and becomes the tool of choice for many reliability engineers. In life data analysis, the mean life is determined by analyzing time-to-failure data. For the case of the piping reliability, the time-to-failure data are extrapolated using an appropriate degradation model.

There are several distribution models that have been successfully applied to failure data such as Weibull, exponential, lognormal and many other distributions. Figure 3.7 illustrates the concept of a distribution model for failure data.

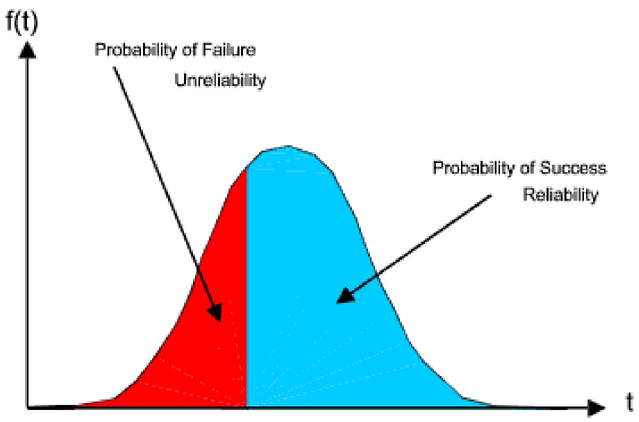

Figure 3.7: The concept of a distribution model for failure data

Weibull analysis is a powerful tool for a reliability assessment that can be used to classify failures and to model failure behavior. The very common forms of the Weibull distribution are:

Probability density function, $f(t)$: $$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}} \quad (3.20)$$

Cumulative density function, $F(t)$: $$F(t) = 1 - e^{\left[-\left(\frac{t}{\eta}\right)^{\beta}\right]} \quad (3.21)$$

where t is the time-to-failure. The distribution is characterized by two parameters, the scale parameter η and the shape parameter β. The value of the parameter β identifies the mode of failure. For example, $\beta < 1$ means infant mortality (decreasing failure rate), $\beta = 1$ indicates random failure (constant failure rate) and $\beta > 1$ describes wear-out failure (increasing failure rate). The scale parameter η is defined as the life at which 63.2% of units will fail.

Once the Weibull parameters are known, then it is easier to estimate the failure rate defined as the frequency with which a system or component fails, expressed as failures per unit time. For example, the failure rate may be expressed as 2×10^{-4} failures/year. It is often denoted by the Greek letter λ (lambda) and the failure rate function is

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1-F(t)} \tag{3.22}$$

If the failures exhibit a constant failure rate, then the time-to-failure data fits an exponential distribution which is described as follows.

$$F(t) = 1 - e^{-\lambda t} \tag{3.23}$$

The mean time between failures of exponential distribution is

$$\text{MTBF} = \frac{1}{\lambda} \tag{3.24}$$

Another common distribution is a lognormal distribution. The general forms of the lognormal distribution are:

Probability density function, $f(t)$: $$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(t)-\mu)^2}{2\sigma^2}\right] \tag{3.25}$$

Cumulative density function, $F(t)$: $$F(t) = \Phi\left(\frac{\ln(t)-\mu}{\sigma}\right) \tag{3.26}$$

where t is the time-to-failure, μ is the mean, σ is the standard deviation and $\Phi(\cdot)$ is the is the standard normal cumulative density function.

3.5 Structural Reliability Analysis

One of the probabilistic engineering approaches for supporting the structural integrity analysis of structures and components under service loads is structural reliability analysis. This probabilistic approach is capable in incorporating the uncertainties in the reliability assessment. Uncertainties in the reliability assessment may be divided in several categories but only a few will be highlighted as follows:

1. *Modeling uncertainty*: Model uncertainty is the uncertainty related to imperfect knowledge or idealizations of the mathematical models used or uncertainty related to the choice of probability distribution types for stochastic variables.
2. *Physical uncertainty*: Physical uncertainty is related to the natural randomness of a quantity, for example the uncertainty in the yield strength of the material due to production variability and the variability of the physical dimensions of the structural components.
3. *Statistical uncertainty*: Statistical uncertainty is the uncertainty due to limited sample sizes of observed quantities. Statistical estimators (e.g. sample mean and coefficient of variation) can be typically estimated from available data and then used to suggest an appropriate probability density function and associated parameters. Generally, the observations of the variable are not represented perfectly and as a result there may be bias in the data as recorded. Moreover, different sample data sets usually produce different statistical estimators. Statistical uncertainty can be incorporated in a reliability analysis by repeating the analysis using different values of the parameters to indicate sensitivity.

In assessing the reliability of structures, the following steps are typically followed:

1. To identify the limit-state function that defines the failure,
2. To assess the variability that lies in the input variables of the limit-state function.

3.5.1 Limit State Function

Structural reliability analysis requires a limit state function, which defines failure or safe performance. Also known as a failure function or a performance function, the limit states could relate to strength failure, serviceability failure, or anything else that describes unsatisfactory performance. The limit state function, $g(x)$ is defined as

$g(x_1, \dots, x_n) > 0 \rightarrow$ *Safe*

$g(x_1, \dots, x_n) = 0 \rightarrow$ *Failure surface*

$g(x_1, \dots, x_n) < 0 \rightarrow$ *Failure*

where $x_1, \dots, x_n$ are the basic random variables and n is the number of random variables. The basic random variables are the variables employed in a structural analysis such as dimensions, materials, loads and material strengths. Often it is common for the limit state function to be the resistance minus the load as follows

$$g(x) = R - S \tag{3.27}$$

where R is the resistance and S is the load.

In structural reliability analysis, $x_1, \dots, x_n$ are considered to be random and statistically distributed with known distribution type and distribution parameters. The failure probability p_f can then be calculated as follows with the failure domain F:

$$p_f = \int_F f_1(x_1) \dots f_n(x_n) dx_1 \dots dx_n \tag{3.28}$$

where $f_n(x_n)$ represents the probability densities of the respective basic variables x_n, which for the sake of simplicity are assumed to be stochastically independent i.e. any (Melchers, 1999). This assumption means that the distributions for the random variables are independent.

In the case where a pipe has more than one corrosion defect and assume that these defects are designated at location 1, 2, 3,..., n and the corresponding failure probabilities are by $p_1, p_2, p_3, \dots, p_n$, respectively. If it is assumed that the individual failures are mutually independent (which implies that the failure of any one of them is sufficient to cause pipe failure) then the failure probability of the pipe can be approximated from:

$$p_{f(\text{pipe failure})} = 1 - (1 - p_1)(1 - p_2)(1 - p_3) \dots (1 - p_n) \tag{3.29}$$

3.5.2 Reliability Analysis Methodologies

There are many random variables in the reliability assessment of corroded piping systems such as the corrosion rate, pipe thickness and pipe diameter. To simulate the failure probability, the appropriate probabilistic model must be employed so that it can take account the influential variables. Many probabilistic analyses are generic and therefore, applicable for assessing the probability of failure of piping systems subject to CUI. Several methods to measure the reliability of a structure are as follows (Tong, 2001):

1. *Strength-load interference method*: Using this basic structural reliability method, the uncertain parameters are modeled by one characteristic value. The problem considers only one load effect S resisted by one resistance R which are described by a known probability density function, $f_S(\sigma)$ and $f_R(\sigma)$, respectively. The structural element will be considered to have failed if its resistance R is less than the load effect S acting on it. The probability of failure p_f of the structural element can be stated in any of the following ways:

$$p_f = P(R \leq S) \tag{3.30}$$

$$= P(R - S \leq 0) \tag{3.31}$$

$$= P\left(\frac{R}{S} \leq 1\right) \tag{3.32}$$

The probability of failure becomes:

$$p_f = P(R - S \leq 0) = \int_F \int f_{RS}(r, s)\, drds \tag{3.33}$$

where $f_{RS}(r, s)$ is the joint probability density function of $f_R(r)$ and $f_S(s)$ and F describes the failure domain. When R and S are independent, then $f_{RS}(r, s) = f_R(r) f_S(s)$. Eq. (3.33) becomes

$$p_f = P(R - S \leq 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\geq r} f_R(r) f_S(s) drds \tag{3.34}$$

This method of failure probability analysis is one of the oldest methods in structural reliability analysis and it continues to be popular due to its simplicity and its ease of use. The major disadvantage of this method is the assumption that load and resistance are statistically independent which may not be valid for some problems when there is a correlation between load and resistance (Tong, 2001).

2. *Second-Moment methods*: The Second-Moment' methods are approximation methods which are popular due to their inherent simplicity. In these methods, the probability density functions of each variable are simplified by representing them only by their first two moments (i.e. mean and variance). In this case, the probability density function is described by the normal distribution function. The higher moments, such as skewness and flatness of the distribution, are ignored. Then, the next useful step is to transform these variables to their standardized normal distribution with zero mean ($\mu = 0$) and unit variance ($\sigma^2 = 1$). This transformation and approximation of the random variables via standard normal distribution completely simplifies the integration procedures in determining the failure probability, and hence, all the useful properties of the normal distribution can be utilized.

   Typically when involved with complex problems, the limit state function is often non-linear. However, the limit state function can be linearized to allow further simplification and the technique is known as the 'first-order' method. Thus, the first limit state and second moment random variables approximation techniques are brought together to give the First-Order Second Moment (FOSM) reliability method and it is the basis of this reliability method.

   FOSM technique yields the exact probability of failure when the random variables are normally distributed and when the limit-state is linear. However, it is not always the case in most reliability problems where the distribution of some of the variables is not normal. Moreover, the failure probability determined using FOSM, is commonly taken as the nominal failure probability rather than the 'true' probability of failure (Melchers, 1999).

   To overcome these disadvantages, one of the extensions of the second moment and transformation methods is First-Order Reliability method (FORM). FORM is well equipped to handle this type of problem and hence will be employed in this study. Details of FORM will be discussed in the next section.

3. *Monte Carlo simulation (MCS):* MCS involves 'sampling' at 'random' to simulate a large number of experiments and to observe the results. MCS involves sampling each random variable x_i randomly to give a sample value $\hat{x}_i$. Then, the limit state function $g(\hat{x}_i) = 0$ is checked. If the limit state is violated, the structure or structural element has failed. The experiment is repeated many times, each time with a randomly chosen sample value. If N trials are conducted, MCS estimates the probability of failure by

$$p_f = \frac{N_H}{N} \tag{3.35}$$

where N_H is the total number of trials where failure has occurred and N is the total number of trials conducted.

MCS exhibits the following characteristics: it can be applied to many practical problems, allowing the direct consideration of any type of probability distribution for the random variables; it is able to compute the probability of failure with the desired precision; it is easy to implement. However, despite the advantages it presents, the use of this method is not widespread in structural reliability because it is not efficient when compared to second-moment methods.

This is due to MCS requires a great number of trials for the analysis in order to evaluate the probability of failure of a structure with a prescribed precision. The number of trials depends on the order of magnitude of that probability. As the values of the probability of failure associated to the ultimate limit states vary normally between 10^{-4} and 10^{-6}, the number of analyses to be performed for ensuring a 95% likelihood that the actual probability be within 5% of the computed one must be at least 1.6×10^7 to 1.6×10^9, according to (Shooman, 1968).

3.5.3 First-Order Reliability Method

FORM has been applied in the reliability analysis of stress corrosion cracking (Cizelj et al., 1994), creep and fatigue failure (Mao, 2000). FORM is a process which can be used to determine the probability of a failure given the probability distribution of the basic random variables and the limit state function. For this approach, the variables need to be characterized only by their means and variance. In the case where the variance is hardly determined due to lack of data, the coefficient of variation (COV) is commonly applied. COV is a normalized measure of dispersion of a probability distribution and is defined as the ratio of standard deviation σ to the mean μ as shown in Eq. (3.36). For simplicity, it will be assumed, that these variables are mutually independent.

$$\mathrm{COV} = \frac{\sigma}{\mu} \tag{3.36}$$

FORM uses a first order approximation of the limit state function and therefore, the calculated probability of failure is also approximate. FORM method is based on the Hasofer-Lind reliability index where this method involves linearising the limit state function at the design point and then determining the value of the reliability index, β which satisfies the failure function, corresponding to the linearized limit state function (Low et al., 2004). The Hasofer–Lind reliability index and FORM have been well-documented in Ditlevsen (1981), Shinozuka (1983), Ang and Tang (1984), Madsen et al. (1986), and Tichy (1993), for example. The continuing interest in FORM is also evident in Melchers (1999), Haldar and Mahadevan (1999), Nowak and Collins (2000).

The Hasofer-Lind reliability index, β can be described as the minimum distance, in standard deviation units, between the mean value point to the limit state function as shown in Figure 3.8. β is defined by

$$\beta = \min_{g=0} \sqrt{\left\{\frac{x_i - m_i}{\sigma_i}\right\}^T [R]^{-1} \left\{\frac{x_i - m_i}{\sigma_i}\right\}} \quad \text{for } i = 1, \ldots, N \tag{3.37}$$

where x is a representing the set of random variable values, m is the mean values, σ is the standard deviation and $[R]$ is the correlation matrix of the variables. It should be noted here that β is a measure of factor of safety, a term describing the structural capacity of a system beyond the applied loads or actual loads.

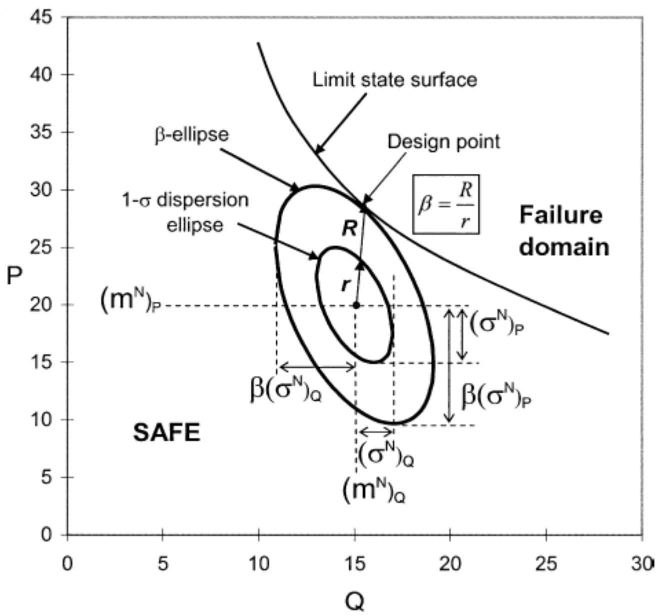

Figure 3.8: Design point and equivalent normal dispersion ellipsoids illustrated in the plane (Low & Tang, 2004)

The method to determine β is based on an iterative technique and the iteration is continued until a desired convergence is obtained. Since the determination of β is an iterative process to find the minimum value of a matrix calculation subject to the constraint that the values result in a system failure, common Solver routines found in several software packages (e.g. Excel) can easily arrive at the solution.

Once β is found, the pipe failure probability corresponding to the corrosion defect can be determined from:

$$p_f = \text{Prob}(g(x) < 0) = \Phi(-\beta) \tag{3.38}$$

where $\Phi(-\beta)$ is a distribution function of a variable and this function is assumed to be distributed normally with zero mean and unit standard deviation.

FORM also allows non-normal distributed functions to be incorporated into this technique. When the basic random variables are non-normal, the approach known as the Rackwitz–Fiessler equivalent normal transformation will be used to transform the non-normal distribution to normal distribution (Rackwitz & Fiessler, 1978). The following two functions are used to do the transformation:

Equivalent normal standard deviation: $$\sigma^N = \frac{\phi\{\Phi^{-1}[F(x)]\}}{f(x)} \tag{3.39}$$

Equivalent normal mean: $$m^N = x - \sigma^N \times \Phi^{-1}[F(x)] \tag{3.40}$$

where x is the original non-normal variate, $\Phi^{-1}[.]$ is the inverse of the cumulative density probability of a standard normal distribution, $F(x)$ is the original non-normal cumulative density probability evaluated at x, $\phi\{.\}$ is the probability density function of the standard normal distribution, and $f(x)$ is the original non-normal probability density ordinates at x. For details of Rackwitz–Fiessler equivalent normal transformation, refer to Rackwitz & Fiessler (1978).

The advantage of the FORM technique over the numerical integration and Monte-Carlo simulation is its simplicity in calculating or approximating the probability of failure. Having many random variables describing the structural reliability problem will create a problem for integration methods. However, this is not so critical for FORM technique.

3.5.4 Bootstrap Method

One of the random variables that are required to estimate the probability of failure for corroded piping systems using the structural reliability analysis is the corrosion rate. Not much attention has been focused by researchers for developing models for CUI corrosion rates. There is no empirical and semi-empirical model which has been developed to determine CUI corrosion rate found in the literature that is based on field data (e.g. inspection data). Moreover, no published data, neither experimental data nor field data, have been found on the mean and variance of CUI corrosion rate as well as its statistical distribution.

The corrosion rates data collected during inspection period (after insulation is being removed) are available, even though, the quantity of data is limited. Using this data, fortunately alternative approaches exist and the bootstrap re-sampling approach is an alternative method that can be employed when the data is limited (Chernick, 2008). Bootstrap sampling allows the empirical cumulative distribution function for CUI corrosion rates to be obtained from random sampling results generated from the measured CUI corrosion rate. Introduced in 1979, the bootstrap method is a very general re-sampling procedure for estimating the distributions of statistics based on independent observations. The bootstrap method is shown to be successful in many situations. For example, this method was applied to predict the time of first repair and subsequent rehabilitation of concrete bridge decks subject to chloride-induced corrosion (Kirkpatrick et al., 2002).

For the mathematical description of the bootstrap re-sampling procedure, the following nomenclature is used: hats $(\hat{.})$ denotes estimates, asterisks $(.^*)$ denotes quantities related to bootstrap samples, N is the sample size, K is the number of bootstrap simulations (Riesch-Opperman et al., 2007).

Consider the problem of estimating CUI corrosion rate by the bootstrap re-sampling method. Suppose the original sample of corrosion rate consists of $x_1, \dots, x_N$ of independent data points, from which a statistic of interest $s(x_1, \dots, x_N)$ is computed. $x_1, \dots, x_N$ are independent and identically distributed (iid) samples from a k-dimensional population distribution F. A schematic representation of the bootstrap re-sampling procedure is shown in Figure 3.9.

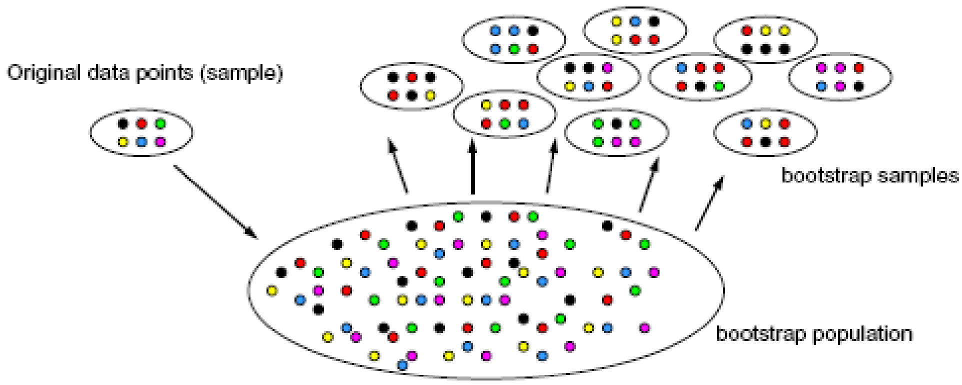

Figure 3.9: Schematic representation of the bootstrap resampling procedure (Riesch-Oppermann et al., 2007)

A bootstrap sample $x^* = (x_1^*, \dots, x_N^*)$ is obtained by randomly sampling, N times, with replacement, from the original data points $x_1, \dots, x_N$, i.e. each data point carries equal probability weight for sampling. If this procedure is repeated K times, we obtain a large number K of independent bootstrap samples $x^{*1}, \dots, x^{*K}$, each of size N. As indicated in Figure 3.9, this procedure can be regarded as artificially inventing a so-called bootstrap population by replicating the original sample a large number of times.

From each bootstrap sample x^{*k}, a bootstrap replication which is a statistic of interest, $s(x^{*k})$, is calculated. After ordering according to increasing values of $s(.)$, an empirical distribution function $\hat{c}(s)$ is obtained from which quantiles of the distribution of s can be obtained.

In the present application, the statistic of interest is the mean and coefficient of variance (COV) of CUI corrosion rate, CR from the set of bootstrap samples of CUI corrosion rates. The bootstrap replications for the mean and variance lead to an empirical distribution function $\hat{G}(CR)$ reflecting the scatter of CUI corrosion rates.

The $1 - 2\alpha$ percentile interval for corrosion rate is defined by the α- and $(1 - \alpha)$-quantiles of $\hat{G}$. From K independent bootstrap samples, the percentile confidence intervals by taking the $K \times \alpha$th value in the ordered list of the K bootstrap replications of the estimated CR as the lower limit and the $K \times (1 - \alpha)$th value in the ordered list of the K bootstrap replications of the estimated CR as the upper limit of the confidence interval. Typically, K = 200 to 1000 bootstrap samples are appropriate to obtain stable confidence interval (DiCiccio & Epron, 1996; Riesch-Opperman & Walter, 2001; Riesch-Opperman et al., 2007). Selecting 5% and 95% quantiles from the CR distribution, a 90% confidence interval for CR is obtained.

3.6 Markov Model

Markov model is a well-established, powerful mathematical tool for reliability modeling. Markov model is a discrete random process having Markov property. A discrete random process means a system that can be in various states and can change randomly in discrete steps. The Markov property states that the probability distribution for the system at the next step only depends on the current state of the system and is independent of past states. In other words, the description of the present state fully captures all the information that could influence the future evolution of the process. In Markov model, the probability of moving into a state at time $t + 1$ only depends upon the state at time t.

Being a stochastic process means that all state transitions are determined by random chance where at each step, the system may change its state from the current state to another state (or remain in the same state) according to a probability distribution. The changes of state are called transitions, and the probabilities associated with various state-changes are called transition probabilities. Transition probabilities indicate the likelihood that the condition of the system will change from one state to another state in a unit time. A typical transition probabilities matrix **P** has the following general structure where the sum of probabilities in each row must be 1.

$$\mathrm{P} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0j} & \cdots & p_{0m} \\ p_{10} & p_{11} & \cdots & p_{1j} & \cdots & p_{1m} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{i0} & p_{i1} & \cdots & p_{ij} & \cdots & p_{im} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{n0} & p_{n1} & \cdots & p_{nj} & \cdots & p_{mm} \end{bmatrix} \tag{3.41}$$

where p_{ij} = transition probability from state i to state j; $i = 0,1,2 \dots m$ and $j = 0,1,2 \dots m$. Based on the Chapman-Kolmogorov equation, the probability of the system moving from State i to State j after n periods (n transitions), that is the n-step transition probability matrix, $P^{(n)}$, can be obtained by multiplying the matrix P by itself n times (Ross, 2000). Thus:

$$P^{(n)} = P^n \tag{3.42}$$

Let the initial state vector $Q^{(0)}$ be the probability that the Markov chain is in state i at time 0. Then, the state vector $Q^{(n)}$ which is the probability that the chain in state j after n transitions, can be expressed as

$$Q^{(n)} = Q^{(0)} P^{(n)} \tag{3.43}$$

where $Q^{(0)} = [q_1, q_2, \dots, q_m]$ = condition vector at stage n; q_i = probability of being in State i at time 0.

Markov model described above is a discrete-time Markov chains. Another type of Markov chain model is a continuous-time Markov chain. Normally, the degradation process can be modeled as a continuous-time Markov chain and will be described as follows.

3.6.1 Continuous-Time Markov chains

In continuous-time Markov chain, there are no smallest time steps and hence one-step transition probability matrices are no longer valid. If one is at state $i \in I$ at time t, then one can ask for the probability of being in a different state $j \in I$ at time $t + h$,

$$f(h) = P\{X(t + h) = j | X(t) = i\} \tag{3.44}$$

Consider h is a very small value where $h > 0$. For $h = 0$, $f(0) = 0$. Assume that f is differentiable at 0, the information about f at the origin is its derivative $f^{'}(0)$

$$f^{'}(0) = q_{ij} = \lim_{h\to 0}\frac{P\{X(t+h)=j|X(t)=i\}}{h} = \lim_{h\to 0}\frac{f(h)-f(0)}{h} \tag{3.45}$$

and can be written as

$$P\{X(t+h) = j|X(t) = i\} = q_{ij}h + o(h) \tag{3.46}$$

$o(h)$ is a abbreviation for any function with the property that $\lim_{h\to 0}\frac{o(h)}{h} = 0$, the function is of smaller order than h. This notation denotes the actual function o where o might be different one in each line, but there is no need to invent a new name for each occurence.

The Markov property states that if one knows the state $X(t)$ then all additional information about X at times prior to t is irrelevant for the future: for all $j \neq i$, $t_0 < t_1 <\ldots< t_n < t$ and $x_0, \ldots, x_n \in I$,

$$P\{X(t+h) = j|X(t) = i, X(t_i) = x_i \forall i\} = P\{X(t+h) = j|X(t) = i\} = q_{ij}h + o(h) \tag{3.47}$$

In other words, given that one is in state i at time t, it does not matter how one get there and it does not matter what t is, the probability that in time h one is in state j is some number $P_h(i,j)$ depending only on i, j and h.

Let a matrix $Q = (q_{ij}: i, j \in I)$ which contains all the information about the transitions of the Markov chain X. This matrix is called the Q-matrix of the Markov chain. The properties of Q-matrix are as follows:

- All off-diagonal entries $q_{ij}, i \neq j$, are positive.
- All diagonal entries q_{ii} are negative.
- The sum over the entries in each row is zero.

q_{ij} is the transition rate at which one tries to enter state j when one is at state i, sometimes called as the jump intensity from state i to state j. Since the Markov chain has the memoryless property, the only continuous random variables which have the memoryless property are exponential random variables. Hence it can be deduced that in continuous-time Markov chain, the amount of time one waits at each state has an exponential distribution.

For example, consider a continuous time Markov chain having three states S = {0, 1, 2} as shown in Figure 3.10.

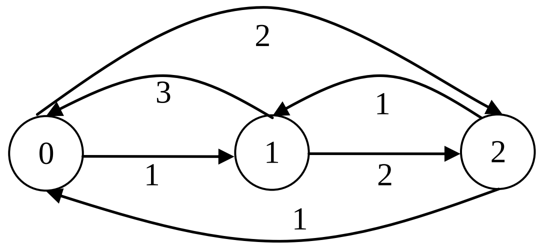

Figure 3.10: Example of a three-state continuous time Markov model

having the transition rates

$$Q = \begin{bmatrix} -3 & 1 & 2 \\ 3 & -5 & 2 \\ 1 & 1 & -2 \end{bmatrix}$$

The corresponding rates at which the chain leaves the states are $q_{00} = 3$, $q_{11} = 5$, and $q_{22} = 2$. This means that if one is currently in state 0, one will wait, on average, 1/3 time units until the next transition; if one is currently in state 1, one will wait, on average, 1/5 time units until the next transition; if one is currently in state 2, one will wait, on average, 1/2 time units until the next transition.

A three-state Markov model for general wall thinning proposed by Fleming (2004) will be employed in this study, as shown in Figure 3.11. The state transition rates are denoted by ϕ, λ and ω. He defined the three states as follows:

- State 1: No detectable damage
- State 2: Detectable flaw (wall thinning)
- State 3: Failure (leak or rupture)

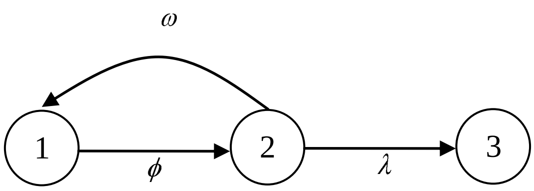

Figure 3.11: Three state continuous time Markov model (Fleming, 2004)

3.6.2 Solution to Differential Equations

The associated differential equations for the three-state Markov model can be written as

$$\frac{dP_1}{dt} = -\phi P_1 + \omega P_2 \tag{3.48}$$

$$\frac{dP_2}{dt} = \phi P_1 - (\lambda + \omega) P_2 \tag{3.49}$$

$$\frac{dP_3}{dt} = \lambda P_2 \tag{3.50}$$

The solution for Eqs. (3.48) - (3.50) can be obtained using Laplace transforms or other suitable technique, such as the eigenvalue method, as so long as the boundary conditions are specified. It is assumed as the piping are inspected to verify they are free of detectable flaws at the beginning of commercial operation, the appropriate boundary conditions are

$$P_1(t = 0) = 1$$

$$P_2(t = 0) = P_3(t = 0) = 0$$

Thus, the time dependent solutions for the state probabilities are given by

$$P_1(t) = \frac{1}{r_1 - r_2}\left[(r_1 + \phi)e^{r_2 t} - (r_2 + \phi)e^{r_1 t}\right] \tag{3.51}$$

$$P_2(t) = \frac{\phi}{r_1 - r_2}\left[e^{r_1 t} - e^{r_2 t}\right] \tag{3.52}$$

$$P_3(t) = 1 - \frac{1}{r_1 - r_2}(r_1 e^{r_2 t} - r_2 e^{r_1 t}) \tag{3.53}$$

where the term r_1 and r_2 are defined as

$$r_1 = \frac{-(\phi+\lambda+\omega)+\sqrt{(\phi+\lambda+\omega)^2-4\phi\lambda}}{2} \tag{3.54}$$

and

$$r_2 = \frac{-(\phi+\lambda+\omega)-\sqrt{(\phi+\lambda+\omega)^2-4\phi\lambda}}{2} \tag{3.55}$$

The solution using Laplace transform is provided in Appendix C. Once such solutions were obtained the calculations could be performed very easily via spreadsheet.

3.6.3 Method to Estimate the Transition Rate

An important prerequisite to ensure the successful application of the Markov-chain model is the availability of reasonably good estimates of the transition rates. Ideally, the transition rate is created based on comprehensive analyses of historical data of the system condition. Fleming (2004) used inspections data and service data for the selected failure in order to estimate the transition rate. However, the method proposed by Fleming (2004) is applicable in the case of CUI due to none or very few CUI failure data recorded. Therefore, direct estimation on the parameter is impossible.

To estimate the transition rates ϕ and λ, first-order reliability method (FORM) was used where the method was adopted by Vinod et al. (2003). FORM requires an appropriate limit-state function and the limit-state functions were developed based on the illustration in Figure 3.12 where the states were defined as in Table 3.7. Refer to Section 3.5.3 on how FORM is used. The strategies for estimating Markov model parameters can be observed in Table 3.8.

To determine the different transition rates ϕ and λ, the limit state functions, based on strength and resistance. The first limit state function is formulated to estimate the transition rate ϕ represents the transition rate from State 1, in which corrosion defect is less than 0.125 of the pipe nominal thickness $t_{nominal}$, to State 2, in which the corrosion defect is t_1. 0.125 of the pipe nominal thickness $t_{nominal}$ was assumed to be the undetectable defect in this study based on the value proposed by Vinod et al. (2003). This value can be considered as it has been found that for radiography, the minimum variation of the thickness that this technique can be detected is 1 % to 2% of the sample thickness (Bardal & Drugli, 2004). The limit state function can be defined as:

$$\text{LSF1: } g(x) = t_1 - (0.125t_{nominal} + CR \times T) \tag{3.56}$$

where $t_1 = (t_{nominal} - t_{min}) - \left(\frac{t_{nominal} - t_{min}}{2}\right)$ (in mm), CR is the corrosion rate (in mm) and T is the time of inspection which usually 10 years.

The second limit state function is used to estimate the transition rate λ represents transition rate from state 2, which has already crossed the detectable range t_1, to the state 3, in which the wall thickness is beyond the minimum wall thickness allowedt_2. The LSF for this case would be

$$\text{LSF2: } g(x) = t_2 - [t_1 + CR \times T] \tag{3.57}$$

where $t_2 = t_{nominal} - t_{min}$ (in mm).

Note that, in this model, the failure stage 3 does not specify the actual leak, but represents a stage where the corrosion defect reaches the minimum wall thickness.

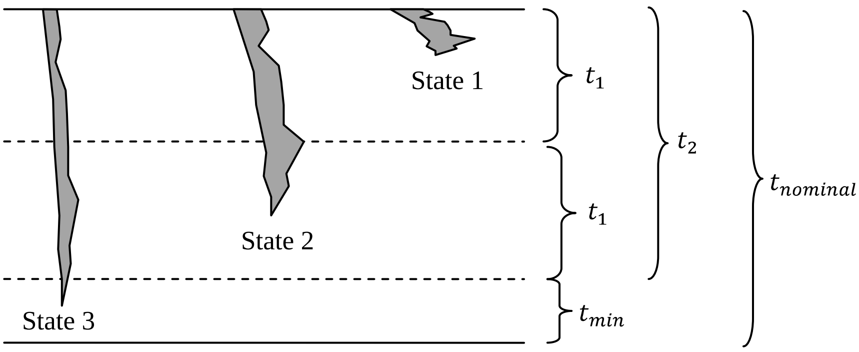

Figure 3.12: Illustration for the limit-state functions used to estimate the transition rates

Table 3.7: Description for the three-state Markov model

State	Definition
1	d less than t_1
2	d is between t_1 and t_2
3	d is more than t_2

Note: d is the depth of corrosion, $t_1 = (t_{nominal} - t_{min}) - \left(\frac{t_{nominal} - t_{min}}{2}\right)$, $t_2 = t_{nominal} - t_{min}$

Table 3.8: Strategies to estimate Markov model parameters

Symbol	Parameter definition	Strategy for estimation
$\boldsymbol{\phi}$	Occurrence rate of a flaw	Estimate using FORM with limit state function $g(x) = t_1 - (0.125t_{nominal} + CR \times T)$
$\boldsymbol{\lambda}$	Occurrence rate of a leak from a flaw state	Estimate using FORM with limit state function $g(x) = t_2 - [t_1 + CR \times T]$
$\boldsymbol{\omega}$	Inspection and repair rate of a flaw state	Model of Eq. (3.49) and estimates of P_I, P_{FD}, T_{FI}, T_R
$\boldsymbol{P_I}$	Probability per inspection interval that the pipe element will be inspected	Estimate based on specific inspection strategy; usually done separate for ASME Sec XI and RISI inspection programs
$\boldsymbol{P_{FD}}$	Probability per inspection that an existing flaw will be detected	Estimate based on NDE reliability performance data and difficulty of inspection for particular element
$\boldsymbol{T_{FI}}$	Flaw inspection interval, mean time between in-service inspections	Normally 10 years for ASME Section piping systems
$\boldsymbol{T_R}$	Mean time to repair the piping element given detection of a critical flaw or leak	Estimate of time to tag out, isolate, prepare, repair, leak test and tag in-service; if to be conditioned for at power, can be no longer than technical specification limit for operating with element tagged out of service

The repair rates ω are estimated using the model given in Eq. (3.58),

$$\omega = \frac{P_I P_{FD}}{T_I + T_R} \tag{3.58}$$

where

P_I = the probability that piping element with a flaw will be inspected per inspection interval. The value will be 1 if it is in the inspection program or else it will be 0 (Vinod et al., 2003).

P_{FD} = the probability that a flaw will be detected given this element is inspected. This parameter is related to the reliability of Non Destructive Examination (NDE) inspection which is often referred to as Probability of Detection. For most NDE, its values are between 0.84 and 0.95 (Vinod et al., 2003).

T_I= the mean time between inspections for defects (For piping system, the inspection interval proposed by API 570 is either 5 or 10 years depending on the piping class. Refer to Table 3.1)

T_R = the mean time to repair once detected (For this study, it is assumed to be 14 days based on the input from plant experts).

3.6.4 Hazard Rate for Markov Model

One way to determine the failure rate or hazard rate is to determine the system reliability function for the model. Next is to derive the hazard rate as a function of the reliability function according to the definition of the hazard rate as explained below. Using this concept, the reliability function for Markov model $r(t)$ is given by

$$r(t) = 1 - P_3(t) = P_1(t) + P_2(t) \tag{3.59}$$

From this, the hazard rate $h(t)$ is the time dependent frequency of pipe leak and can defined for pipe leak as

$$h(t) = -\frac{1}{r(t)}\frac{dr(t)}{dt} = \frac{1}{1-P_3(t)}\frac{dP_3(t)}{dt} \tag{3.60}$$

or

$$h(t) = \frac{r_1 r_2\left(e^{r_1 t} - e^{r_2 t}\right)}{r_1 e^{r_2 t} - r_2 e^{r_1 t}} \tag{3.61}$$

3.7 Concluding Remarks

This chapter discusses the mechanism of CUI, state-of-art of CUI inspection strategy and techniques used to assess the probability of failure. The basic theories for the four models namely logistic regression, degradation analysis, structural reliability analysis and continuous-time Markov model are also discussed in detailed.