

Enhanced Pipeline Risk Assessment

Part 1—Probability of Failure Assessments Rev 4

W. Kent Muhlbauer, PE

www.pipelinerrisk.com

This document presents new material that is to be incorporated into the book: Pipeline Risk Management Manual, 4th Edition, by W. Kent Muhlbauer, published by Gulf Publishing Co. This material should be viewed as a book excerpt. As a standalone document, it lacks some of the definitions and discussions that can be found in other chapters of that book. Philosophies of risk, data management, segmentation, uncertainty, and other variables impacting pipeline risk are among the topics into which this new material fits. The reader is encouraged to refer to the 3rd edition text for details and clarifications of concepts that are not fully developed in this document.

This material represents ideas and possible approaches to problem solving that may or may not be appropriate in certain situations and applications. The user of this material is strongly urged to exercise careful judgment in the use of all information presented here. Author provides no guarantee, expressed or implied, with regard to the general or specific application of this information. The user accepts all liability in any and all applications of any material taken from this document.

Copyright © 2008 by WKM Consultancy. All rights reserved.

Table of Contents

1. Introduction.....	3
2. Background.....	3
3. “Limitations” in Previous Approaches	5
4. Improvement Opportunity	7
4.1 Why Change Now?	7
4.2 Change Without Pain	7
5. Characteristics of the New PoF Algorithms	9
5.1 PoF Triad	9
5.2 PoF Assessment Steps.....	10
5.3 Model Features.....	10
5.4 Orders of Magnitude	12
5.5 Effective Zero	12
5.6 AND gates OR gates	13
6. Probability of Failure	15
6.1 Failure Mechanisms	15
6.2 From TTF to PoF	17
6.3 Role of Inspection and Testing	20
7. PoF: Exposure.....	22
7.1 Exposure as MPY (Pipe Wall Degradation Rate).....	22
7.2 Exposure as Events per Length-Time	26
8. PoF: Mitigation	29
8.1 Mitigation of Time-Dependent Failure Mechanisms.....	29
8.2 Mitigation of Time-Independent Failure Mechanisms	33
9. PoF: Resistance.....	34
9.1 Current Pipe Strength.....	35
9.2 Estimates of Effective Pipe Wall	36
9.3 Procedure to Estimate Resistance	36
10. Time Independent Failure Mechanisms.....	44
10.1 Third Party	44
10.2 Incorrect Operations.....	45
10.3 Geohazards.....	47
11. Final PoF.....	49
12. Calibration and Validation.....	51
12.1 Calibration vs. Validation	51
12.2 Use of Incident History	52
12.3 SME Validation	52
12.4 Diagnosing Disconnects Between Results and Reality	53

1. Introduction

Scoring or ranking type pipeline risk assessments have served the pipeline industry well for many years. However, risk assessments are being routinely used today in ways that were not common even a few years ago. For example, many operators are asking questions today such as:

- How to make full use of inspection data in a risk assessment
- How to generate results that directly suggest timing of integrity assessments
- How to quantify the risk reduction benefit of integrity assessment and other mitigations.
- Beyond the prioritization, how big is the risk? Is it actionable?
- How widespread is this risk issue?
- How can subjectivity be reduced?
- How to use past incident results in a risk assessment.

The new roles of risk assessments have prompted some changes to the way risk algorithms are being designed. The changes lead to more robust risk results that better reflect reality and, fortunately, are readily obtained from data used in previous assessments.

2. Background

Scoring systems as a means of analysis have been around for a long time. When knowledge is incomplete and a decision structure is needed to simultaneously consider many factors, scoring systems often appear. Boxing matches, figure skating, beauty contests, financial indices, credit evaluations, and even personality and relationship “tests” are but a few examples.

Many risk assessments were based on such scoring systems. They were often a simple summation of numbers assigned to conditions and activities that are expected to influence risks. Whenever more risk-increasing conditions are present with fewer risk-reducing activities, risk is relatively higher. As risky conditions decrease or are offset by more risk-reduction measures, risk is relatively lower.

The form of these simple assessments is normally some variation on:

$$\text{CondA} + \text{CondB} + \dots + \text{CondN} = \text{Relative Probability of Failure (or relative Consequence of Failure)}$$

Or sometimes:

$$(\text{CondA} \times \text{WeightA}) + (\text{CondB} \times \text{WeightB}) + \dots + (\text{CondN} \times \text{WeightN}) = \text{Probability of Failure}$$

Where

CondX represents some condition or factor believed to be related to risk, evaluated for a particular piece of pipeline.

WeightX represents the relative importance or weight placed on the corresponding condition or factor—more important variables have a greater impact on the perceived risk and are assigned a greater weight.

In the pipeline industry, relative risk scoring or ranking systems have been around for decades. Early published works from the late 1980's and early 1990's in scoring type risk assessments include:

- Dr. John Kiefner's work for AGA,
- Dr. Mike Kirkwood from British Gas,
- W. Kent Muhlbauer's first edition of The Pipeline Risk Management Manual, and
- Mike Gloven's work at Bass Trigon.

Such scoring systems for specific pipeline operators can be traced back even further, notably in the 1980's with gas distribution companies faced with repair-replace decisions involving problematic cast iron pipe.

Variations on this type of scoring assessment have now been in common use by pipeline operators for many years. The choices of categorization into failure mechanisms, scale direction (higher points = higher risk or vice versa) variables, and the math used to combine factors are some of the differences among these type models.

The scoring approach is often chosen for its intuitive nature, ease of application, and ability to incorporate a wide-variety of data types. These methodologies have served the industry well in the past. Prior to 2000, such models were used primarily by operators seeking more formal methods for resource allocation—how to best spend limited funds on pipeline maintenance, repair, and replacement. Risk assessment was not generally mandated and model results were seldom used for purposes beyond this resource allocation. There are of course some notable exceptions where some pipeline operators incorporated very rigorous risk assessments into their business practices, notably in Europe where such risk assessments were an offshoot of applications in other industries or already mandated by regulators.

The role of risk assessment in the U.S. expanded significantly in the early 2000's when the DOT, OPS—now, Pipeline and Hazardous Materials Safety Administration (PHMSA)—began mandating risk ranking of all jurisdictional gas and liquid pipelines that could affect a High Consequence Area (HCA). Identified HCA segments were then scheduled for integrity assessment and application of preventative and mitigative measures depending on the integrity threats present.

3. “Limitations” in Previous Approaches

The simple scoring assessment is still a useful screening and prioritization tool. However, these risk-ranking models were generally not intended for use in applications where outside parties were requesting more rigorous risk assessments. For example, risk assessment has now been used in project presentation and acceptance in public forums; legal disputes; setting design factors; addressing land use issues; etc, while previously, the assessment was typically used for internal decision support only.

Given their intended use, the earlier models did not really suffer from “limitations” since they met their design intent. They only now appear as limitations as the new uses are factored in. Those still using older scoring approaches recognize the limitations brought about by the original modeling compromises made. Some of the more significant compromises arising from the use of the simple scoring type assessments include:

- Without an anchor to absolute risk estimates, the assessment results are useful only in a rather small analysis space. The results offer little useful information regarding risk-related costs or appropriate responses to certain risk levels. Results expressed in relative numbers are useful for prioritizing and ranking but are limited in their ability to forecast real failure rates or costs of failure. They cannot be readily compared to other quantified risks to judge acceptability.
- Difficult to directly link to integrity re-assessment timing. Without additional analyses, the scores do not suggest appropriate timing of ILI, pressure testing, direct assessment, or other integrity verification efforts.
- Potential for masking of effects when simple expressions cannot simultaneously show influences of large single contributors and accumulation of lesser contributors. An unacceptably large threat—very high chance of failure from a certain failure mechanism—could be hidden in the overall failure potential if the contributions from other failure mechanisms are very low. This is because, in some scoring models, failure likelihood will only approach highest levels when all failure modes are coincident. A very high threat from only one or two mechanisms would only appear at levels up to their pre-set cap (weighting). In actuality, only one failure mode will often dominate the real probability of failure. Similarly, in the scoring systems, mitigation is generally high only when all available mitigations are simultaneously applied. The benefit of a single, very effective mitigation measure is lost when the maximum benefit from that measure is artificially capped.¹
- Some older models are unclear as to whether they are assessing damage potential versus failure potential. For instance, the likelihood of corrosion occurring versus the likelihood of pipeline failure from corrosion is a subtle but important distinction since damage does not always result in failure.
- Some previous approaches have limited modeling of interaction of variables (now required in some IMP regulations). Older risk models often did not adequately represent the contribution of a variable in the context of all other variables. Simple summations cannot properly integrate the interactions of some variables.
- Some models forced results to parallel previous leak history—maintaining a certain percentage or weighting for corrosion leaks, third party leaks, etc—even when such history might not be relevant for the pipeline being assessed.²

- Balancing or re-weighting was often required as models attempt to capture risk in terms that represent 100% of the threat or mitigation or other aspect. The appearance of new information or new mitigation techniques required re-balancing which in turn made comparison to previous risk assessments problematic.
- Some models can only use attribute values that are bracketed into a series of ranges. This creates a step change relationship between the data and risk scores. This approximation for the real relationship is sometimes problematic
- Some models allowed only addition, where other mathematical operations (multiply, divide, raise to a power, etc) would better parallel other engineering models and therefore better represent reality
- Simpler math does not allow orders of magnitude scales and such scales better represent real world risks. Incident frequencies and related probabilities can commonly range, for example, from nearly annually to less than 1 in ten million chance per year.

Notes:

1. *In general, the use of pre-set weightings or averaging of conditions can obscure higher probabilities of one or more failure mechanisms. The user of such models is usually cautioned to either examine enough lower level results (prior to averaging or application of weighting) to ensure this does not happen, or to migrate to an algorithm that will prevent the masking.*
2. *The assumption of a predictable distribution of future leaks predicated on past leak history might be realistic in certain cases, especially when a database with enough events is available and conditions and activities are constant. However, one can easily envision scenarios where, in some segments, a single failure mode should dominate the risk assessment and result in a very high probability of failure rather than only some percentage of the total. Even if the assumed distribution is valid in the aggregate, there may be many locations along a pipeline where the pre-set distribution is not representative of the particular mechanisms at work there.*

Users of the older scoring type risk assessments should recognize these potential difficulties in such methodologies. Serious practitioners always recognized these “limitations” and worked around them when more definitive applications were needed. However, when the limitations are coupled with the now-routine need to get more out of the risk assessments, the case for change becomes compelling.

4. Improvement Opportunity

4.1 Why Change Now?

While the previous generation of algorithms served the industry well, the technical compromises made can be troublesome or unacceptable in today's environment of increasing regulatory and public oversight. Risk assessments commonly become the centerpiece of any legal, regulatory, or public proceedings. This prompts the use of assessment techniques that more accurately model reality and also produce risk estimates that are anchored in absolute terms: “consequences per mile year,” for example. Fortunately, a new approach to algorithm design can do this while making use of all previously collected data and not increasing the costs of risk assessment. The advantages of the new algorithms are that they can overcome many of the previously noted limitations:

- More intuitive;
- Better models reality;
- Eliminates masking of significant effects;
- Makes more complete and more appropriate use of all available and relevant data ;
- Greatly enhances ability to demonstrate compliance with U.S. IMP regulations;
- Distinguishes between unmitigated exposure to a threat, mitigation effectiveness, and system resistance—this leads directly to better risk management decisions;
- Eliminates need for unrealistic and expensive re-weighting of variables for new technologies or other changes; and
- Flexibility to present results in either absolute (probabilistic) terms or relative terms, depending on the user's needs.

4.2 Change Without Pain

A migration from an older style risk assessment to the new approach is quite straightforward. An objective of the new approach is to retain the advantages of earlier approaches such as their simplicity and intuitively transparent nature, and still avoid overly-analytic techniques that often accompany more absolute quantifications of risk. In all risk analyses, the designer of the assessment model must strike a balance between complexity and utility—using enough information to capture all meaningful nuances (and satisfy data requirements of all regulatory oversight) but not information that adds little value to the analysis.

The new model described here uses the same data as previous approaches, but uses it in different ways. Weightings are not needed, but as with older models, valuations sometimes must still need to arise from engineering judgment and expert experience when “hard data” is not available. The new valuations are, however, more verifiable and defensible since they are grounded in absolute terms rather than relative. Some time and energy will still need to be invested into setting up the new assessment model with legitimate values for the systems being assessed. This investment is no greater than that needed to set up and maintain the older models.

In recent risk model upgrades, the time needed to convert older scoring type risk assessment algorithms into the new approach has averaged less than 40 man-hours. The new approach makes use of existing data to help with continuity and to keep costs of conversion low. The primary algorithm modifications consist of simple and straightforward changes to categorization of variables and the math used to combine them for calculating risk scores. The new algorithms are easily set up and executed in spreadsheets, desktop databases, GIS environments—basic (Structured Query Language) SQL handles all routines very readily. No special software is needed.

5. Characteristics of the New PoF Algorithms

5.1 PoF Triad

As in previous models, risk is the product of probability of failure (PoF) and consequence of failure (CoF). PoF normally harbors more opportunities for risk reduction than does CoF. Therefore, more attention to PoF is often warranted and the discussion of the new modeling approach begins there.

The suggested basis for this model is to examine each failure mechanism (threat) in three parts for:

- Exposure
- Mitigation effects
- Resistance to failure.

These three elements make up the Risk Triad, for evaluating PoF. They are generally defined as follows:

- Exposure = likelihood of force or failure mechanism reaching the pipe when no mitigation applied,
- Mitigation = actions that keep the force or failure mechanism off the pipe, and
- Resistance = the system's ability to resist a force or failure mechanism applied to the pipe.

An analogous naming convention is 'attack', 'defense', and 'survivability', respectively, for these three terms. The evaluation of these three elements for each pipeline segment results in a PoF estimate for that specific segment.

Probability of Damage—damage without immediate failure—also emerges from the triad. Using the first two terms without the third—exposure and mitigation, but not resistance—yields the probability of damage.

$$\text{Probability of Damage (PoD)} = f(\text{exposure, mitigation})$$

$$\text{Probability of Failure (PoF)} = f(\text{PoD, resistance})$$

Damage does not always result in immediate failure. Some damage may trigger or accelerate a time-dependent failure mechanism. Calculation of both PoD and PoF values creates an opportunity to gain better understanding of their respective risk contributions.

This three-part assessment also helps with model validation and most importantly, with risk management. Fully understanding the exposure level, independent of the mitigation and system's ability to resist the failure mechanism, puts the whole risk picture into clearer perspective. Then, the roles of mitigation and system vulnerability are both known independently and also in regards to how they interact with the exposure. Armed with these three aspects of risk, the manager is better able to direct resources appropriately.

5.2 PoF Assessment Steps

The overall steps for assessment of PoF under the new algorithms are as follows:

1. Exposure: Estimate exposure from each threat
 - a. Degradation rate from time-dependent failure mechanisms
 - b. Event rate from time-independent failure mechanisms
2. Mitigation: Estimate combined effect of all mitigations
 - a. Identify all mitigation measures
 - b. Rate effectiveness of each
3. Resistance:
 - a. Produce best estimate of current pipe strength by finding the “governing” information from:
 - i. Pipe wall implied by last pressure test
 - ii. Pipe wall implied by last inspection (including ILI, bell hole exam, etc)
 - iii. Pipe wall implied by assumption of leak-free at current NOP
 - iv. Possible pipe wall weaknesses
 - b. Estimate pipe’s resistance to failure from each threat
4. PoF: Calculate PoF from each threat
 - a. Risk Triad: combine Exposure, Mitigation, Resistance
 - b. TTF and then PoF for time-dependent failure mechanisms
 - c. PoF for time-independent failure mechanisms
 - d. Combine all PoF’s

5.3 Model Features

Other characteristics of this model distinguish it from previous risk assessment approaches and include the following.

1. Measurement Scales

More precise numerical scales are employed to fully capture the orders-of-magnitude differences between “high” risk and “low” risk. The new scales better capture reality and are more verifiable. Some exposures are measured on a scale spanning several of orders of magnitude—“this section of pipeline could be hit by excavation equipment 10 times a year, if not mitigated (annual hit rate = 10)” and “that section of pipeline would realistically not be hit in 1000 years (0.001 annual hit rate).”

The new approach also means measuring individual mitigation measures on the basis of how much exposure they can independently mitigate. For example, most would agree that “depth of cover”, when done as well as can be envisioned, can independently remove almost all threat of third party damage. Deep cover is theoretically a

characteristic that can mitigate 95-99% of the third party damage exposure. If buried deep enough, there is very little chance of third party damage, regardless of any other mitigative actions taken. This is because deeper burial places the pipe out of reach of most excavation equipment. “Public Education” on the other hand, is recognized as an important mitigation measure but most would agree that, independently, it cannot be as effective as depth of cover in preventing third party damages.

Improved valuation scales also means a more direct assessment of how many incidents can be avoided when the pipeline is more resistant or invulnerable to certain damages.

2. Variable Interactions

This model uses combinatorial math that captures both the influences of strong, single factors as well as the cumulative effects of lesser factors. For instance, 3 mitigation measures that are being done each with an effectiveness of 20% should yield a combined mitigation effect of about 49%. This would be equivalent to a combination of 3 measures rated as 40%, 10%, and 5% respectively, as is shown later.

In other cases, all aspects of a particular mitigation must simultaneously be in effect before any mitigation benefit is achieved. An example is patrol conducted with high frequency—multiple times per week—but low effectiveness (due to altitude of aircraft or overgrowth, etc) or a powerful ILI but with inadequate confirmatory investigations. So, if three aspects are each rated as 80% and each are essential to the performance of the mitigation, then the mitigation is actually only about 51%.

These examples illustrate the need for OR and AND “gates” as ways to more effectively combine variables. Their use eliminates the need for “importance-weightings” seen in many older models.

The new approach also provides for improved modeling of interactions: for instance, if some of the available pipe strength is used to resist a threat such as external force, less strength is available to resist certain other threats.

See section 5.5 for a further discussion of the math involved.

3. Meaningful Units

The new model supports direct production of absolute risk estimates. The model can be calibrated to express risk results in consistent, absolute terms: some consequence per some length of pipe in some time period such as “fatalities per mile year” or “\$ per km-decade”. Note that the definition of risk is embodied in these units: the frequency (probability) of some consequences.

Of course, this does not mean that such absolute terms must always be used. They can easily be converted into relative risk values when those simpler (and perhaps less emotional) units are preferable. The important thing is that absolute values are readily obtainable when needed.

5.4 Orders of Magnitude

Orders of magnitude scales are needed to appropriately characterize the range of failure probabilities. This is a departure from how most older scoring models approach risk quantification. It is a necessary aspect to properly mirror real-world effects and express risk estimates in absolute terms.

An order of magnitude is synonymous with a factor of 10 or “10 times” or “10X.” Two orders of magnitude means 100X, and so forth, so an order of magnitude is really the power to which ten is raised. A range of values from 10E2 to 10E-6 (10^2 to 10^{-6}) represents 8 orders of magnitude

This PoF model measures most mitigation effectiveness and resistance to failure in terms of simple percentages. However, contrary to conventional wisdom regarding significant digits, many digits must be carried with the percentages to capture the orders of magnitude. For example, mitigation deemed to reduce exposure by 1000 times—not an unusual scenario—carries a mitigation of 99.9%. If the unmitigated exposure is 0.1—the event was happening once every 10 years on average—it would be reduced to $0.1 \times (1 - 0.999) = 0.1 \times 0.001 = 1E-4$ by mitigation. The mitigation has reduced the event frequency by a factor of 1000—only one in a thousand of the events that would otherwise have occurred will occur under the influence of the mitigation.

Numbers for mitigated exposures will get very, very small whenever the starting point (unmitigated exposure) is small: 1000 times better than a “1 in a million” starting point is very small; 1000 times better than a “1 in a 100” starting point is not so small.

It might take some out of their comfort zone to begin working with numbers like this. If so, relative scales are easily created to be surrogates for the complex numbers. However, having access to the complex—and more correct—values at any time will add greatly to the risk model’s ability to support a wide range of applications.

Creating a correct range of orders of magnitude for a model is part of the tuning or calibration process.

5.5 Effective Zero

For some calculations, a lower limit or “effective zero” is needed to make the mathematical relationships perform properly. An effective zero is also a concept grounded in reality. Intelligent minds are never *absolutely* certain of anything. There is always some very slim possibility of almost anything. So, the effective zero can be seen as assigning a value to what we mean when we say “never.” For instance, to most, a chance of an event of around 1 in a trillion or perhaps 1×10^{-12} is the equivalent of saying “never.” This then would be the “effective zero” value to use in the risk assessment equations. There are some subtleties involved in selecting this value, as will become apparent when some risk values are generated. The value is also subject to change when a risk model is calibrated to produce results in absolute terms such as failures per mile-year.

5.6 AND gates OR gates

The probabilistic math used to combine variables to capture both the effects of single, large contributors as well as the accumulation of lesser contributors is termed “OR” & “AND” “gates.” This terminology is borrowed from flowchart techniques. Their use in pipeline risk assessment modeling represents a dramatic improvement over most older methods. This type of math better reflects reality since it uses probability theory of accumulating impacts to

- Avoid masking some influences;
- Captures single, large impacts as well as accumulation of lesser effects;
- Shows diminishing returns;
- Avoids the need to have pre-set, pre-balanced list of variables;
- Provides an easy way to add new variables; and
- Avoids the need for re-balancing when new info arrives.

OR Gates

OR gates imply independent events that can be added. The OR function calculates the probability that any of the input events will occur. If there are i input events each assigned with a probability of occurrence, P_i , then the probability that any of the i events occurring is:

$$P = 1 - [(1-P_1) * (1-P_2) * (1-P_3) * \dots * (1-P_i)]$$

OR Gate Example:

To estimate the probability of failure based on the individual probabilities of failure for stress corrosion cracking (SCC), external corrosion (EC) and internal corrosion (IC), the following formula can be used.

$$\begin{aligned} P_{\text{failure}} &= \text{OR}[P_{\text{SCC}}, P_{\text{EC}}, P_{\text{IC}}] = P_{\text{SCC}} \text{ OR } P_{\text{EC}} \text{ OR } P_{\text{IC}} \\ &= \text{OR} [1.05\text{E-}06, 7.99\text{E-}05, 3.08\text{E-}08] \quad (\text{using some sample values}) \\ &= 1 - [(1-1.05\text{E-}06)*(1-7.99\text{E-}05)*(1-3.08\text{E-}08)] \\ &= 8.10\text{E-}05 \end{aligned}$$

The OR gate is also used for calculating the overall mitigation effectiveness from several independent mitigation measures. This function captures the idea that probability (or mitigation effectiveness) rises due to the effect of either a single factor with a high influence or the accumulation of factors with lesser influences (or any combination).

$$\begin{aligned} \text{Mitigation \%} &= M_1 \text{ OR } M_2 \text{ OR } M_3 \dots \\ &= 1 - [(1-M_1) * (1-M_2) * (1-M_3) * \dots * (1-M_i)] \\ &= 1 - [(1-0.40) * (1-0.10) * (1-0.05)] \\ &= 49\% \end{aligned}$$

or examining this from a different perspective,

$$\begin{aligned} \text{Mitigation \%} &= 1 - [\text{remaining threat}] \\ &\text{Where} \end{aligned}$$

[remaining threat] = [(remnant from M₁) AND (remnant from M₂) AND (remnant from M₃)] ...

AND Gates

AND gates imply “dependent” measures that should be combined by multiplication. Any sub-variable can alone have a dramatic influence. This is captured by multiplying all sub-variables together. In measuring mitigation, when all things have to happen in concert in order to achieve the mitigation benefit, a multiplication is used—an AND gate instead of OR gate. This implies a dependent relationship rather than the independent relationship that is implied by the OR gate.

AND Gate Example:

The modeler is assessing a variable called “CP Effectiveness” (cathodic protection effectiveness) where confidence in all sub-variables is necessary in order to be confident of the CP Effectiveness—[good pipe-to-soil readings] AND [readings close to segment of interest] AND [readings are recent] AND [proper consideration of IR was done] AND [low chance of interference] AND [low chance of shielding] . . . etc. If any sub-variable is not satisfactory, then overall confidence in CP effectiveness is dramatically reduced. This is captured by multiplying the sub-variables.

When the modeler wishes the contribution from each variable to be slight, the range for each contributor is kept fairly tight. Note that four things done pretty well, say 80% effective each, result in a combined effectiveness of only ~40% (0.8 x 0.8 x 0.8 x 0.8) using straight multiplication.

6. Probability of Failure

The most compelling and useful definition of probability is “degree of belief.” When we speak of the probability of a pipeline failure, we are expressing our belief regarding the likelihood of an event occurring in a specified future period. Probability is most often expressed as a decimal ≤ 1.0 or a percentage $\leq 100\%$. Historical data, usually in the form of summary statistics, often partially establishes our degree of belief about future events. Such data is not, however, the only source of our probability estimates.

Probability is often expressed as a forecast of future events. In this application, the expression has the same units as a measured event frequency, i.e. events per time period. When event frequencies are very small, they are, for practical purposes, interchangeable with probabilities: 0.01 failures per year is essentially the same as a 1% probability of one or more failures per year, for purposes here. When event frequencies are larger, a mathematical relationship—reflecting an assumed underlying distribution—is used to convert them into probabilities, ensuring that probabilities are always between 0 and 100%.

The pipeline risk assessment model described here is designed to incorporate all conceivable failure mechanisms. It is then calibrated using appropriate historical incident rates, tempered by knowledge of changing conditions. This results in estimates of failure probabilities that match the judgments and intuition of those most knowledgeable about the pipelines.

6.1 Failure Mechanisms

This model recognizes that the two general types of failure mechanisms—time dependent and time independent—require slightly different calculation routines. Time dependent mechanisms of corrosion and fatigue can be initially measured in terms of how much damage they are causing over time. The initial “damage rate” measurement will then be used to calculate a time-to-failure (TTF) and then a probability of failure (PoF), perhaps in failures/mile/year. TTF and PoF can be estimated using common engineering and statistical relationships, either very complex (fracture mechanics, finite element analyses, rupture estimates, etc) or with simple approximations (% of Barlow-required pipe wall thickness, etc).

Time-independent failure modes are assumed to either cause immediate failure or create a defect that leads to a time-dependent failure mechanism. For failure mechanisms such as third party damage, weather, human error, and earth movement events, the PoF estimation process is a bit simpler. Constant failure rate or random failure rate events are assessed with a simple “frequency of occurrence” analysis. The estimated frequency of occurrence of each time-independent failure mechanism can be directly related to a failure probability—PoF—and then combined with the PoF’s from the time-dependent mechanisms.

As an example of failure mechanism categorization, ASME B31.8 Appendix S nomenclature identifies time-dependent threats as External Corrosion (EC), Internal Corrosion (IC), Stress Corrosion Cracking (SCC), and fatigue. Time independent threats are Third Party damage (TP), Incorrect Operations (IO), weather and other outside forces such as earth movement events (WOF) and Equipment failures (EQ, classified as “random”). Also noted are potential manufacturing (MFG) and construction (CON) issues as contributors to failure potential.

While the ASME reference categorizes these last as ‘random’ failure mechanisms, they are really potential points of weakness, not failure mechanisms. In the modeling approach recommended here, all possible weaknesses are best captured as elements of resistance—the ability of a pipeline segment to resist failure when exposed to a threat. Equipment failure can often also be included as part of the other mechanisms, where valves, flanges, separators, etc are treated the same as pieces of pipe but with different strengths and abilities to resist failure. Large rotating equipment (pumps, compressors) and other pieces may warrant independent assessment.

The model described here supports any logical categorization of threats or failure mechanisms. The following table summarizes one categorization scheme.

Table 6.1-1 Failure Mechanism Categories

Failure Mechanism	Mechanism Type	Probability Model Structure
Third Party; geohazards; human error; sabotage; theft	Time-independent	(failure rate) = [unmitigated event frequency] x (1 – [exposure reduction]) Where [exposure reduction] = f (mitigation effectiveness, resistance)
Ext corrosion; Int corrosion; Fatigue, SCC	Time-dependent	(failure rate in year one) = $1 - \text{EXP}(-1 / \text{TTF})$ or other user-defined relationship, where $\text{TTF} = (\text{available pipe wall}) / [(\text{wall loss rate}) \times (1 - \text{mitigation effectiveness})]$

Under the assumption that most forecasted failure rates for pipelines will be very small, this document will often substitute “probability of failure” for “failure rate.” So, the two basic equations used are modified from the table above and become:

$$\text{PoF}_{\text{time-indep}} = [\text{unmitigated event frequency}] \times (1 - [\text{exposure reduction}])$$

where
[exposure reduction] = f (mitigation, resistance)

$$\text{PoF}_{\text{time-dep}} = f(\text{TTF})$$

where
TTF = “time to failure”
= (available pipe wall) / [(wall loss rate) x (1 – mitigation)]

And then:

$$\text{PoF} = f(\text{PoF}_{\text{time-indep}}, \text{PoF}_{\text{time-dep}})$$

Terms and concepts underlying these equations are discussed in the following sections.

6.2 From TTF to PoF

The relationship between an estimated TTF and the probability of failure in the next year—year one—can be complex and warrants special discussion. The PoF is normally calculated as the chance of one or more failures in a given time period. In the case of time-dependent failure mechanisms, TTF estimates are first produced. The calculated probability assumes that at least one point in the segment is experiencing the estimated degradation rate and no point is experiencing a more aggressive degradation rate.

The TTF estimate is expressed in time units and is calculated by using the estimated pipe wall degradation rate and the theoretical pipe wall thickness and strength, as was shown above. In order to combine the TTF with PoF from all other failure mechanisms, it is necessary to express the time-dependent failure potential as PoF. This requires a conversion of TTF to PoF. It is initially tempting to use the reciprocal of this time-to-failure number as a leak rate—failures per time period. For instance, 5 years to failure implies a failure rate of once every five years perhaps leading to the assumption of 0.20 failures per year. However, a logical examination of the TTF estimate shows that it is not really predicting a uniform failure rate. The estimate is actually predicting a failure rate of ~0 for 4+ years and then a nearly 100% chance of failure in the fifth year. Nonetheless, use of a uniform failure rate is conservative and helps overcome potential difficulties in expressing degradation rate in probabilistic terms. This is discussed later.

An exponential relationship can be used to show the relationship between PoF in year one and failure rate. Using the conservative relationship of [failure frequency] = 1/TTF, a possible relationship to use at least in the early stages of the risk assessment is:

$$\text{PoF} = 1 - \text{EXP}(-1 / \text{TTF})$$

where

PoF = probability of failure in year one

TTF = time to failure

This relationship ensures that PoF never exceeds 1.0 (100%). As noted, this does not really reflect the belief that PoF's are very low in the first years and reach high levels only in the very last years of the TTF period. The use of a factor in the denominator will shift the curve so that PoF values are more representative of this belief. A Poisson relationship or Weibull function can also better show this, as can a relationship of the form $\text{PoF} = 1 / (\text{fctr} \times \text{TTF}^2)$ with a logic trap to prevent PoF from exceeding 100%. The relationship that best reflects real world PoF for a particular assessment is difficult if not impossible to determine. Therefore, the recommendation is to choose a relationship that seems to best represent the peculiarities of the particular assessment, chiefly the uncertainty surrounding key variables and confidence of results. The relationship can then be modified as the model is tuned or calibrated towards what is believed to be a representative failure distribution.

The relationship between TTF and PoF includes segment length as a consideration. PoF logically increases as segment length increases since a longer length logically means more opportunity for active failure mechanisms, more uncertainty about variables, and more opportunities for deviation from estimated degradation rates. This is discussed more fully in a later section.

6.2.1 Using Two Relationships for Year One PoF Estimations

Using the basic relationship described above can result in excessive conservatism. In more advanced applications of risk assessment, a more complex relationship between PoF and TTF might be warranted. By adding an extreme value analysis to the basic TTF analysis, early year TTF's can be dismissed in certain scenarios.

For example, a new pipeline has little chance of corrosion leak in the early years, even when aggressive corrosion rates are possible. Therefore, even if a worst case TTF is 5 years, the new pipeline enjoys a very low PoF in year one. Use of the simple $PoF = 1/TTF$ does not show this. It yields a 20% chance of failure in year one, until the extreme value analysis demonstrates that this is over-conservative.

The extreme value analysis requires the creation of a variable called TTF99. TTF99 is the min plausible TTF—a value that is lower than any actual value will be, 99-99.9% of the time—the subject matter expert (SME) is 99% confident that the TTF cannot be worse than this value, even considering a highly improbable coincidence of very unlikely factors.

Establishing this extreme value can be done by taking the best pipe wall thickness estimate and degrading that by the highest plausible unmitigated corrosion rate. Alternatively, statistical methods can be used to establish the 99% confidence level, when data is available.

Using both TTF and TTF99 creates four scenarios, each with its own relationship to PoF. These scenarios involving TTF (best estimate of current time to failure) and TTF99 (lowest plausible TTF) are examined to arrive at an estimate of PoF:

The scenarios are summarized as follows, assuming the time of interest is 1 year—a year one PoF is sought:

If TTF99 less than 1 year AND TTF less than 1 year, then $PoF = 99\%$

If TTF99 less than 1 year AND TTF greater than 1 year, then use constant failure rate, basically the reciprocal of the TTF, to estimate PoF

If TTF99 greater than 1 year AND TTF greater than 1 year, then 'use more conservative of a less conservative relationship (such as $\text{lognormal}(TTF99)$) vs. an assumed constant failure rate ($1/TTF$)

Scenario 1. If it is plausible to have a year one failure AND the best estimate of TTF is also less than one year.

If TTF99 is less than 1 year, indicating that failure in year one has $>1\%$ probability of occurrence, and TTF is also < 1 year, then failure during year one is probable and PoF is assigned 99%. Pipeline segments are conservatively assigned this value when little information is available and a very short TTF cannot be ruled out.

Scenario 2. If it is plausible to have a year one failure AND the best estimate of TTF is greater than one year.

If TTF99 is < 1 year and $TTF > 1$ year, then PoF needs to reflect the probabilistic mpy embedded in the TTF estimate. Probabilistic mpy means that, for instance, a 10%

chance of a 100mpy degradation rate, will be treated as 10 mpy. To ensure that the PoF estimate captures the small chance of a 100mpy rate actually occurring next year, a constant and conservative failure probability— $PoF = 1/TTF$ —is associated with the 10mpy. Pipeline segments will fall into this analysis category when very short TTF cannot be ruled out but the most probable TTF values exceed one year.

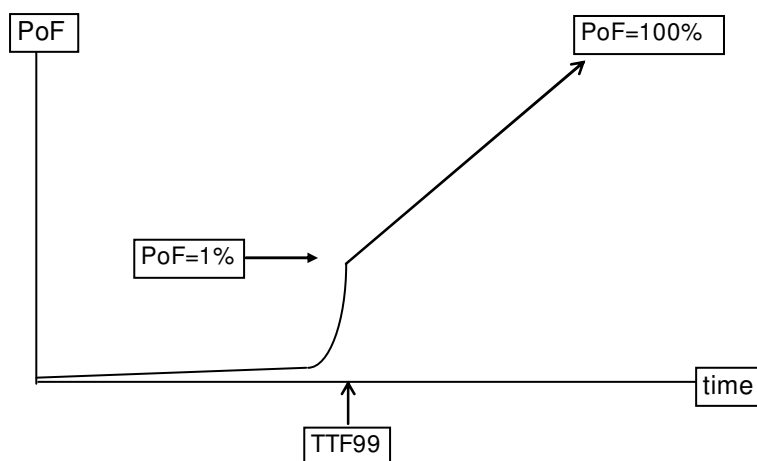
Scenario 3. If it is not plausible to have a year one failure, even using extreme values

If $TTF_{99} > 1$ year then TTF_{99} , rather than the actual TTF governs PoF. The relationship between TTF_{99} and PoF can be assumed to be lognormal or Weibull or some other distribution, with parameters selected from actual data or from judgments as to distribution shapes that are reflective of the degradation mechanism being modeled. Very low year one PoF's will emerge. A new pipeline with a relatively slow plausible degradation rate, will have a PoF governed by this analysis.

Scenario 4. TTF is very high

When TTF is very high, it overrides TTF_{99} for PoF. This is again logical. Even if TTF_{99} is close to one--PoF approaching 100%--TTF might indicate that the segment's actual TTF (best estimate) is so far from this low probability event, that it should govern the final PoF estimate. A pipeline segment with very high confidence in both current pipe wall and a low possible degradation rate will have a high TTF. Even if a short TTF is theoretically possible—as shown by TTF_{99} —a confidence in the estimated TTF will govern. Such high confidence is often obtained via repeated, robust inspections.

Scenarios 3 & 4 are possible only when $TTF_{99} > 1$ year—virtually no chance of failure in year one. Then the worst case between scenario 3 and scenario 4 governs. See the figure below where PoF is on the vertical axis and time is on the horizontal axis.



6.3 Role of Inspection and Testing

Inspection and testing plays a significant role in pipeline failure potential. An operator will often have several pieces of information, differing in age and accuracy, for a single pipeline location. For example, he will often know the date and pressure level of the last pressure test, results and accuracy of the latest ILI, and perhaps also non-destructive evaluation (NDE) measurements taken from bell hole examinations over the years. The challenge is to integrate all of these and extract the most information from the data.

The model makes a distinction between inspection results that produce measurements of damage and those that produce measurements of damage rate. The former shows where a previously active damage rate has resulted in damage such as metal loss or cracks. The latter provides an estimate of currently active damage rates—a metal loss rate or crack growth rate. Both measurements are derived from inspections of various types, but they are used differently in the model. The use of this information in the model should mirror the deliberations of a SME. In other words, the model is designed to arrive at the same conclusions that would emerge from an SME studying the same information.

Damages are quantified as either metal loss—depth and/or volume of metal loss from internal and external corrosion—or cracking—depth and shape characteristics of cracks. Damages are obtained from any or all of the following sources:

- Measurements from ILI;
- Measurements from NDE (including simple visual investigations); and/or
- Other (such as inferential results drawn from corrosion coupons, probes, analyses of pigging effluents, and other sources).

Actual measurements of damage should override damage estimates that were derived from indirect data. In the absence of actual measurements, estimates of damage are produced by taking the most recent measurement and then applying the estimated damage rate for the period since the measurement was taken.

Damage *rates* for internal corrosion, external corrosion, and cracking are used to estimate the remaining service life and PoF of a pipeline segment. As with damage estimates, each of these damage mechanisms can have rate estimates based on any or all of the following:

- Calculations based on ILI measurements
- Calculations based on NDE measurements
- Inferential information from pressure testing
- Other (such as from corrosion coupons, probes, analyses of pigging effluents, and other sources)

This damage rate is applied to the segment from the time of the last damage measurement until the time of the risk assessment. It is also used to forecast future conditions.

Each measurement—for either damage or damage rate—carries its own accuracy, age, and confidence level. Accuracy involves the potential error band of the tool itself, but also potential for operator-, procedural-, and other inaccuracies. Confidence level should be based

not only on the accuracy and age of the measurement but also the location relevance. In some instances, but not all, a nearby measurement may provide some information for the pipeline segment being evaluated. All of these aspects must be considered when making conclusions about current conditions. For instance, a highly accurate NDE taken 3 years ago might yield as much information about current conditions as a less accurate ILI performed one year ago. By conservatively considering all the aspects that make up a SME's confidence in a measurement, the measurements are made to be equivalent. Once made equivalent, the model can legitimately use the smallest damage rate.

A new integrity inspection can “reset the clock” by overriding older and/or less accurate data. This is the value of inspection or testing: obtaining new information that leads to a better estimate of current pipe integrity. As an added ‘bonus’ in this approach to integrating various inspection techniques, a cost-benefit to various inspections can be easily created. This allows comparisons of highly-accurate, but potentially more expensive, inspection options with lower-accuracy but inexpensive options. In the case of indirect measurements such as certain ILI techniques, the follow-up protocols must also be considered in assessing the accuracy of the inspection.

See Section 9 for examples of direct integration of inspection and testing data into a risk assessment.

7. PoF: Exposure

“Exposure” is the name given to this model’s measure of the level of threat to which the pipeline segment is exposed, if absolutely no mitigation measures were present. Exposure is the attack on pipeline integrity. It can be thought of as a measure of how active a failure mechanism is in the pipeline segment’s environment. Each failure mechanism contributes some threat exposure to each pipeline segment. Exposure is measured differently for the two different categories of failure mechanisms:

Pipe Wall Degradation Rate for time-dependent mechanisms

- External corrosion,
- Internal corrosion,
- Cracking,
- SCC (and all other forms of environmentally-enhanced cracking).

Events per length-time (mile-year, for instance) for time independent / random mechanism

- third party
- incorrect operations
- weather
- land movements (geohazards)
- equipment failures
- theft/sabotage

7.1 Exposure as MPY (Pipe Wall Degradation Rate)

For time-dependent threats, each segment of pipeline can have varying degrees of exposure to each of the time-dependent mechanisms such as:

- External corrosion,
- Internal corrosion,
- Fatigue cracking, and/or
- SCC.

These exposures can potentially cause a damage rate—a rate of pipe wall loss or crack growth through the pipe wall. The exposure from each threat, whether producing a corrosion or cracking damage rate, is measured in mils per year where 1 mil = 1/1000th of an inch or in mm. Note that an important distinction is being made between *damage* and *damage rate* (see section 6.3). This discussion centers on damage rate. Damage—the amount of pipe wall lost due to previously active damage rates—is the focus of the discussion on “effective pipe wall.”

In the case of corrosion, the pipe's reaction to its environment establishes a rate for material loss by corrosion. Although metal loss is actually a loss of mass and is perhaps best characterized by a loss of volume, using a one-dimensional measure—depth of metal loss—conservatively assumes “narrow and deep” corrosion versus “broad and shallow.” It is, after all, the loss of effective wall thickness that is of primary importance in judging impending loss of integrity for time-dependent failure mechanisms. MPY or mm/year are the metrics commonly used by corrosion control experts and can refer to pitting—a more aggressive and localized damage—or general corrosion, to characterize metal loss. In some cases, considerations of volume or weight loss instead of thickness loss might be warranted—note the difference in depth associated with a 1 lb/year metal loss when a pitting mechanism is involved versus a generalized surface corrosion.

This approach to measuring exposure is readily applied to pipe corroding or deteriorating in any environment: unburied pipe (atmospheric exposure), pipe atop supports, in snow banks, in flowing vs stagnant water, in marine splash zones, wrapped in insulation, and many others.

MPY can also be used to measure crack growth rates if some simplifying assumptions are used. Crack propagation is a very complex mechanism. To fully estimate cracking potential, concepts of fracture mechanics must be applied, including possible presence of defects, characteristics of defects, stress levels including stress concentrators, metallurgy, etc. As one important variable—sometimes as the only measure of exposure—fatigue cycles are usually measured, both in terms of their magnitude and frequency. The two general types of fatigue loadings commonly seen are large magnitude low frequency cycles as is typically seen in internal pressure fluctuations and smaller magnitude but more frequent cycles typically seen in traffic or temperature loading scenarios. All scenarios involving all combinations of frequency and magnitude should be identified. Most will be directly additive. In other cases, OR gate math applied to all simultaneous causes ensures that any scenario can independently drive fatigue and also show the cumulative effect of several lesser exposures.

SCC can be considered a special form of degradation involving both cracking and corrosion. Since aggressive corrosion can actually slow SCC crack-growth rates, the interplay of cracking and corrosion phenomena can be difficult to model. Recent literature has identified factors that seem to be present in most instances of SCC. These factors can be used to estimate an exposure level, expressed in units of mpy, and this exposure can be added to internal corrosion, external corrosion, and fatigue crack-growth, for an overall exposure level.

Other forms of environmental cracking, blistering, or other damages, should also be considered in the exposure estimates for time-dependent mechanisms.

As a modeling convenience, mpy and mils lost assumes a consistent rate of damage. This is normally not the case in reality. Both corrosion and cracking are known to exhibit a relatively wide range of rates even when all conditions appear to be unchanging. Allowances for more aggressive, shorter duration damage rates might be warranted for certain modeling applications.

Unless additional details are incorporated into the model, the potential mpy rate applies to every square centimeter of a pipe segment—the degradation could be occurring everywhere uniformly and simultaneously. See also discussion of length effects below.

7.1.1 Measurements vs Estimates

At times, an operator will have direct measurements of damage rates derived from coupons, probes, or recurring inspections (see Section 6.3). When available, these measured damage rates can be made equivalent to estimated damage rates by assigning confidence levels to each. Confidence in a damage rate should be based on many considerations such as:

- Proximity of measurement to location being assessed
- Age and accuracy of measurement
- Relevance of measurement to future damage rates at location of interest
- Accuracy of estimates (often governed by accuracy of inputs to estimating model)

An uncertainty or confidence should be assigned to any mpy estimate or measurement, especially when derived from inferential sources. A measured mpy is often inferred from some evidence such as a monitoring point some distance from the point of interest or by assuming a time period over which a measured damage might have occurred. An estimated mpy is often based on assumed characteristics of the electrolyte. Uncertainty can be thought of as a probability adjustment, reflecting the greater opportunity for damaging conditions with increasing uncertainty.

All measured and estimated damage rates should be adjusted by confidence levels. For example, having 50% confidence that a measured 2 mpy corrosion rate represents the best estimate of future corrosion rates, may warrant the use of 4 mpy for future corrosion rates. The next section further discusses this ‘probability of mpy’ concept.

Having made all information equivalent, measurements and estimates can then be compared and the more optimistic can be used. Using the more conservative will not allow inferior information to be replaced by better information. Ensuring equivalency is the key to being able to select the more optimistic value.

7.1.2 Potential Damage Rate = Mitigated Exposure

Exposure to time-dependent mechanisms is measured as an unmitigated damage rate, usually mpy. It is the rate of material loss that would occur on a completely unprotected pipe in the environment of interest. After applying mitigation (discussed in a later section), the assessment changes slightly to now estimate the *probability* of a certain damage rate occurring.

The unmitigated exposure—from soil as an electrolyte, in the case of buried pipeline—is assumed to be fairly uniform at a given location (not always an accurate assumption). Whether or not the exposure is actually causing pipe damage depends on mitigation effectiveness. Where there is no mitigation, the damage rate is obviously 100% of the unmitigated rate. Where mitigation reduces the likelihood of the damage rate occurring, the damage rate is modeled to be less. The premise is that severity of potential damage can be captured via mitigated damage rate: a 10% chance of 100 mpy damage rate is modeled as a 10 mpy damage rate. The usefulness of this simplification and possible ramifications to the assumption are discussed here. This discussion is a slight variation on the previous discussion of TTF to PoF in section 6.2. The user should satisfy himself that a probability-weighted mpy is a valid modeling approach for his needs.

The rationalization that a probability of a baseline damage rate can be modeled as a constant damage rate equivalent to the probability multiplied by the baseline damage rate, can be examined through an example. Suppose that there is an estimate of a 10% chance of a 50 mpy annual damage rate. Further, suppose that a loss of 50 mils results in failure. Does a 10% chance of 50 mpy produce essentially the same failure rate as a 5 mpy damage rate?

In the first expression, 10% chance of 50 mpy, there is a 10% chance of a failure after 1 year which implies failure every 10 years or 0.1 failures per year. In the second expression, assuming a constant 5 mpy damage rate, a failure-causing loss of 50 mils requires 50 mils/5 mpy = 10 years suggesting failure after 10 years, perhaps leading to an assumption of 0.10 failures per year. Both calculations appear to produce the same TTF and failure frequency using such simplifying calculations.

In reality though, conceptual complications arise. In the case of a constant 5 mpy, there would be an extremely low probability of failure in the first years since damages would not have accumulated to the point where failure could occur. As a matter of fact, failure is forecast only in year 10. However, in the case of 10% chance of 50 mpy, there is a real chance of failure in early years. If the 50 mpy does manifest in year one or year two, then failure would occur in those years.

An advantage of using a probability-weighted mpy is that it simultaneously captures both scenarios of

- More probable, lower damage rates leading to accumulated (and finally fatal) damage over many years; and
- Low-probability, aggressive damage rates leading to rapid damages and failure in early years.

The strong tie between mpy and TTF means that manipulation of mpy must not compromise the estimation of TTF. The complications with using a probabilistic mpy are avoided and the simplifying rationalization can be used if a conservative translation between the TTF and failure rate (or PoF) is used. As noted in a previous section, when the PoF is modeled as being essentially inversely proportional to TTF (failure rate = 1/TTF), the probability-adjusted, constant mpy will always be a conservative PoF.

When done properly, the TTF emerging from the mpy damage rate estimate helps ensure that both scenarios are preserved. Mathematically, the formula (failure rate) = 1/TTF, ensures that the 'low probability of a high damage rate' scenario causing a failure in an early year is not lost by focusing only on smaller damage accumulation over many years.

7.1.3 Length of Pipe Implications

Time-independent failure mechanisms have a clear connection to length and time of exposure—how many times an event happens over a certain length of pipe in a certain period of time. Time-dependent failure mechanisms have a clear connection to time since exposure is measured in damage rate—material lost over a certain period of time. However, the connection between damage rate and pipe length may not be immediately apparent.

A probabilistic damage rate logically contains a relationship to the exposed area. A larger surface area presents a greater opportunity for an aggressive exposure and/or a failure of

some mitigation measure. Therefore, damage rates used in this model are captured in length- and time-normalized units such as mpy/mile, although the ‘ per mile’ portion is often omitted until the PoF is calculated.

For example, the number of active corrosion locations on a pipe segment is based on the probability of coating holiday per square foot of coated area and the probability of CP coverage per square foot of pipe surface area. A unit-length of one mile is used and pipe surface areas are based on that unit-length. SME’s base their estimates of holidays and CP coverage gaps on this unit-length of one mile. Hence, the number of active corrosion points—and the probability of one or more active corrosion points—is on a per-mile basis. The probabilistic mpy is based on the probability of an active corrosion location somewhere on the pipe segment, so, the mpy damage rate is also on a per-mile basis. Mathematically, the probability cannot exceed one, but the number of corrosion locations can continue to increase with each additional mile of pipe included in the assessment.

The role of pipe surface area introduces the length effect for most time-dependent failure mechanisms. In the case of external corrosion potential on buried pipe, influence of surface area is most clearly seen when mitigation is considered. Since the coating and CP defect rates are both based on surface-area and are therefore length-sensitive (per-mile), the mitigated exposure estimate is also length-sensitive.

For internal corrosion, the link to length may not be as apparent. Corrosion rate estimates, are not normally expressed with a unit-length component but logically embody length-related issues such as uncertainty that changes with distance from the evidence and the role of surface area in corrosion rate and mitigation effectiveness. Coupon data, pig effluent analyses, and pipe wall inspection are common sources of corrosion rate estimates. The effectiveness of common mitigations such as maintenance pigging and chemical treatments must be estimated in order to assess the corrosion potential. Surface area of pipe interior plays varying roles. Many corrosion mechanisms are focused on only a portion of the pipe interior (such as 6 O’clock position) while the wetted surface area may play a large role in chemical treatment effectiveness.

To make the length connection more clear, it is useful to have SME’s prepare their estimates for a unit-length of pipe. This is illustrated in the discussion of mitigation for external corrosion.

7.2 Exposure as Events per Length-Time

For time-independent threats, exposure should also be quantified independently of any mitigation. Since historical data and typical pipeline experience does not include mitigation-free scenarios, this type of analysis may seem unusual. However, quantifying threats in this manner provides a better understanding of the exposure and helps in tuning the model to actual experience.

The concept of measuring a threat as if there was absolutely no mitigation applied normally requires a bit of “imagineering.” For example, in the case of third party damage, one must envision the pipeline in a completely unmarked ROW (actually indistinguishable as a ROW), with no one-call system in place, no public education whatsoever, and buried with only a few

millimeters of cover—just barely out of sight. Then, a “hit rate” is estimated—how often would such a pipe be struck by agricultural equipment, homeowner activity, new construction, etc.? This exercise is actually very illuminating in that it forces one to recognize the inherent threat exposure without the often taken-for-granted role of mitigation.

A range of possibilities can be useful in setting boundaries for assigning exposure levels to specific situations. A process for estimating a range of exposure levels is

- Envisioning the worst case scenario for a completely unprotected, specific length of pipe and extrapolating (or interpolating) that scenario as if it applied uniformly over a mile of pipe and
- Envisioning the best-case scenario and extrapolating (or interpolating) that scenario as if it applied uniformly over a mile of pipe.

7.2.1 Examples for Third Party Damage Potential

Example worst case scenario: 2500 ft of pipe with 1" cover, no signs, no information available to excavators, located in an active farming and construction zone with potential for line strikes every week. Assessor assigns a value of 50 hits per year for 2500' = 100 hits/mi-yr

Example best case scenario: 10 miles of pipe in controlled, uninhabited desert, no utilities, area with limited access. Assessor assigns 1 hits on 10 miles in 100 years = 0.001 hits/mi-yr

Expressing beliefs in numbers requires a change in mindset for some. It is common to hear that something will “never” happen or that steps must be taken to ensure a certain event is “practically impossible.” Do such terms suggest “once every 100 years”? Once every million years? Or even lower frequencies? Assigning numbers to these qualitative terms removes the emotionalism and makes the term real. In order to help anchor all exposure estimates, a guidance chart can be used:

Table 7.2.1-1. Exposure Levels

Failures/yr	Years to Fail	Approximate Rule Thumb
1,000,000	0.000001	Continuous failures
100,000	0.00001	fails ~10 times per hour
10,000	0.0001	fails ~1 times per hour
1,000	0.001	fails ~3 times per day
100	0.01	fails ~2 times per week
10	0.1	fails ~1 times per month
1	1	fails ~1 times per year
0.1	10	fails ~1 per 10 years
0.01	100	fails ~1 per 100 years
0.001	1,000	fails ~1 per 1000 years
0.0001	10,000	fails ~1 per 10,000 years
0.00001	100,000	fails ~1 per 100,000 years
0.000001	1,000,000	One in a million chance of failure
0.000000001	1,000,000,000	Effectively, it never fails

It is sometimes difficult to imagine the lower ends of the exposure scale. Values implying frequencies like once every 100,000 years or 10,000,000 years are not normally mentioned in the pipeline industry. The reality, however, is that these are real and valid numbers for many stretches of pipeline. A 10 mile stretch of pipeline with one exposure (hit or near miss) in 80 years implies a frequency of 0.00125 incidents per mile-year (1 incident/(10 miles x 80 years) = about once every 800 years for each mile of pipe). If there were no exposures in 80 years—and many lines do exist for decades with no exposures—then one could reasonably assign a frequency of 1/10,000 or higher. When there is little or no historical data, a comparable situation and/or judgment can be used.

With practice, several distinct advantages of the approach become apparent:

- Estimates can often be validated over time through comparison to actual failure rates on similar pipelines.
- Estimate values from several causes are directly additive. For example, many external force threats such as falling objects, landslide, subsidence, etc, each with their own frequency of occurrence can be added together for an overall exposure level.
- Estimates are in a form that consider segment-length effects and supports PoF estimates in absolute terms (failures per mile-year) when such units are desired.
- Avoids need to standardize qualitative measures such as “high,” “medium,” and “low.” Experience has shown us that such standardizations often still leave much room for interpretation and also tend to erode over time and when different assessors become involved.
- Can directly incorporate pertinent company and industry historical data.
- This approach forces subject matter experts (SME) to provide more considered values. It is more difficult to present a number such as 1 hit every 2 years, compared to a qualitative labels such as “high.”

Many geohazards are already commonly expressed in units that are directly linked to event frequency. Recurrence intervals for earthquakes, floods, and other events can be used to establish exposure.

8. PoF: Mitigation

Risk reduction occurs either through reducing the exposure to the threat—mitigation—or reducing the failure likelihood in the face of mitigated exposures—resistance. An underlying premise in assigning values, is that mitigation and resistance can work together to offset much of the exposure, thereby reducing the PoF. In this section, mitigation is discussed. Recall that mitigation is synonymous with ‘defense’ in this application. Exposure is the ‘attacking’ mechanism and mitigation is defending against the attack. If the defense fails, then damage occurs.

Probability of Damage (PoD) = (exposure) x (1-mitigation)

In this model, a percentage is assigned to a mitigation measure that reflects its possible effectiveness in risk reduction. For example, a value of 90% indicates that that measure would independently reduce the failure potential by 90%—only one out of every ten exposures would reach the pipe, based solely on the protection offered by this measure. A mitigation range for each measure is set by the best-case amount of mitigation the variable can independently contribute. So, the “best” possible level of mitigation is an estimate of how effective the measure would be if it was done as well as can be envisioned. A very robust mitigation can theoretically reduce the exposure level to a very low level—sometimes independently eliminating most of threat. Recall the earlier discussion of orders of magnitude and the need to preserve enough significant digits to show these orders of magnitude when PoD is calculated.

In order to capture the belief that mitigation effects can be dominated by either strong independent measures or by accumulation of lesser measures, OR gate math is used, as previously discussed. The OR gates also capture the notion of diminishing returns—overall mitigation effectiveness is less and less improved as more and more is applied.

As was pointed out, this means that very small exposures turn into very, very small PoD’s under the influence of strong mitigation.

8.1 Mitigation of Time-Dependent Failure Mechanisms

For time-dependent mechanisms such as corrosion and fatigue, mitigation is a direct reduction in exposure. Recall that exposure for these mechanisms is measured as a degradation rate in units of mpy or mm per year. Since the previously-estimated exposure is a constant or baseline threat, it is the mitigation estimate that introduces the probabilistic component required for PoF estimates. For instance, unmitigated exposure might be estimated to be 12 mpy, but the effects of mitigation lead to a conclusion that there is only a 20% chance of that 12 mpy damage actually occurring on the pipeline segment. This is equivalent to a 2.4 mpy damage rate, as was shown in the previous discussion.

8.1.1 External Corrosion

Common mitigation measures for external corrosion include coating and application of cathodic protection (CP). These two are usually employed in parallel and provide redundant protection. Since each can independently prevent or reduce corrosion, an OR gate is

appropriate in assessing the combined effect. Some practitioners rate these measures as equally effective, in theory at least.

The modeling approach quantifies external corrosion potential as the probability of one or more active corrosion points on the pipeline segment. This probability is based on an estimate of the frequency of active corrosion locations, derived from estimates of coating holiday rates plus the efficiency with which CP prevents those holidays from experiencing corrosion.

8.1.1.1 Coating Effectiveness

Coating is designed primarily to provide a barrier to the electrolyte. Discounting its role in supporting CP, coating effectiveness is appropriately assessed in terms of its barrier effectiveness or defect rate. Sometimes the effectiveness of coating can be estimated from the current demand of the CP system. This is often expressed as a % bare value.

In the absence of good coating defect rate information for a particular pipeline, the rate can be inferred by using some general corrosion failure rates. This is illustrated by a example of a very coarse quantification of coating condition.

Example: Converting Qualitative Assessments of Coating Condition into Quantitative Terms Based roughly on US DOT failure statistics, first assume that the overall failure rate of a subject pipeline is 0.001 failures per mile-year. Next, assume that 30% of these failures are due to corrosion. Finally, assume that for each corrosion failure, there are 100 coating failures that are protected by CP (so, only one out of every 100 coating defects leads to a failure). Using several representative pipe diameters, this leads to coating failure rates on the order of 1 per million square feet of coated pipe. With several key assumptions, this suggests that an “average” coating has a failure rate on the order of 10E-6 per square foot per year. Perhaps poor coatings are several orders of magnitude worse and superior coatings are several orders of magnitude better. A scale from these very rough assumptions can be generated and is illustrated below

Table 8.1.1.1-1 Sample Linkage of Coating Descriptors to Coating Defect Rates

Coating Evaluation	Assumed % Effectiveness of Coating	Defect Rate Implied by % Effectiveness	Score	Estimated Defect Rate from Calculation (per sq ft per year)
excellent	0.9999999	1E-07	0.9	5.01E-07
good	0.99999	1E-05	0.8	2.51E-06
fair	0.99	0.01	0.5	0.000316
poor	0.9	0.1	0.2	0.039811
absent	0	1	-1	1

By observation, a holiday rate is assigned to each qualitative descriptor to be used in a simplifying formula. This links the descriptor—perhaps carried over from a previous risk assessment—to a defect rate implied by that descriptor. This is obviously a very coarse assessment and should be replaced by better knowledge of the specific pipeline being evaluated.

To better visualize the implications of this simple relationship, and perhaps to help SME’s calibrate terms, consider the following ‘defect rate estimates’ for a sample pipe diameter of 12”. For various lengths of the 12” pipe, the probability of a coating defect is estimated. This can then be used to help validate and tune the coating assessment protocols since records and/or SME’s can often relate actual experiences with a particular coating to such defect rates.

Table 8.1.1.1-2 Visualizing Coating Defect Rates

Coating Score	Defect Rate per sq ft per year	Probability of Defect in Segment, per year for varying lengths of 12” pipe, (L = ft of length)				
		L = 1	L = 10	L = 100	L = 1000	L = 5280
excellent	5.0E-07	0.00%	0.00%	0.02%	0.16%	0.83%
good	2.5E-06	0.00%	0.02%	0.18%	1.75%	8.91%
fair	3.2E-04	2.46%	22.1%	91.8%	99%	100%
poor	4.0E-02	11.8%	71.4%	100%	100%	100%
absent	1.0E+07	100%	100%	100%	100%	100%

In the above table, a mile of “good” coating has about a 9% chance of having at least one defect. A ‘fair’ coating under this system is almost certain to have at least one defect every 1000 ft. These results might seem reasonable for a specific pipeline’s coating. They show that the probability of a coating defect is proportional to both the quality of coating and the length of the segment (length as a surrogate for surface area of the segment). If the results are not consistent with expert judgment—perhaps ratings for “fair” are too severe, for instance—then the modeler can simply modify the equation that relates coating score to defect rate.

Of course, this model is using many assumptions that might not be reasonable for many pipelines. In addition to the highly arguable initial assumptions, many complications of reality are ignored, including:

- Coatings fail in many different ways;
- The meaning of coating “failure” (shielding vs increased conductance vs. holiday etc); and
- The widely varying failure rates of pipelines (DOT data includes all ages, types, conditions, environments, etc, of coatings).

Nonetheless, these valuations capture the perceived relationship between coating quality and surface area in estimating probability of coating damage or defect. Note that in this application, the probability of a defect diminishes rapidly with diminishing segment length. As segments are combined to show PoF along longer stretches of the pipeline, the small defect counts must be preserved (and not rounded to 0.0). The modeler should be cautious that, through length-reduction and rounding, the probabilities are not lost.

8.1.1.2 CP Effectiveness

CP effectiveness is measured by a variety of factors including various pipe-to-soil voltage and other over-line surveys, survey protocols (on, off, de-polarized, etc) interpretation of survey results, rectifier inspections, interference potential, and others. The interactions of CP variables can suggest OR gate math in some cases and AND gate math in others. This is fully discussed elsewhere in the text.

For purposes of this model, CP effectiveness is expressed as a reliability rate in protecting steel beneath coating defects. The factors used to judge CP effectiveness therefore should produce an estimate of how many coating defects have been compensated by CP or, phrased another way, “How many of the coating defect areas will experience the mpy exposure rate because the CP has failed to protect the pipe?” The specifics of estimating the CP effectiveness rate are discussed elsewhere.

8.1.1.3 Combining Coating and CP

The following example illustrates the process of estimating external corrosion mitigation effectiveness. With an assumed coating defect rate of 0.005 (5 defects per 1,000 square feet of coating), and 17, 623 ft² of coated pipe surface area, the expectation is for a coating defect count of 88 per linear mile of pipeline. If CP effectiveness is judged to fully address 999 out of every 1000 coating defects, then the frequency of corroding defects is 0.088 per mile or one episode of corrosion every $1/0.088 = 11.3$ miles or ~9% chance of one or more corroding locations along a mile of pipeline. Under the premise that, mathematically, a 9% chance of a corrosion rate occurring anywhere is the same as corrosion occurring everywhere at 9% of that rate, then the mitigation effectiveness is $1 - 0.09 = 91\%$ and the *mitigated* external corrosion rate is modeled to be 91% less than the *unmitigated* corrosion rate.

The key variables and calculations results from the above example are shown in the table below.

Table 8.1.1.3-1 Example of Coating-CP Mitigation Effectiveness

Attribute	Value
Estimated coating defect rate per 1,000 ft ²	0.005
Total ft ² of pipe surface area	17,623
Resultant number of defects per mile of pipeline	88.1
Estimated number of CP gaps (failures to protect) per coating defect	0.001
Resultant number of unprotected coating defects per mile	0.0881
Final mitigation effectiveness against external corrosion	91.2%

Recall that the degradation rates in this model are actually probabilistic estimates of damage rates—the unmitigated damage rate and the probability of that damage rate being experienced by the pipe. Viewing mitigation as an “all or nothing” defense against external corrosion, the model seeks an estimate of the number of active corrosion points in a given length of pipe. Each of those points is assumed to be experiencing the unmitigated corrosion rate. The number of points is converted into a probability of one or more active corrosion points in the

segment. This probability times the exposure (unmitigated corrosion rate) yields the mpy to be used to calculate TTF and then PoF.

As another example of the relationship between exposure and mitigation (coating defect rate and CP effectiveness), consider the analysis shown in Table 8.1.1.3-2.

Table 8.1.1.3-2 Corrosion at Unprotected Points

Coating Condition	Coating Defect Type	Prob of Defect Type per sq ft	Is CP fully eff?	Prob of CP protecting pipe	Scenario Prob	Resultant MPY
excellent	none	99.9%	Y	0.9	89.9%	0
			N	0.1	0.01%	0
	hole	0.1%	Y	0.9	0.09%	0
			N	0.1	0.01%	16
	shielding	0.0%	Y	0.1	0.0%	0
			N	0.9	0.0%	16
Final probability of 16 mpy damage rate per sq ft					0.01%	16
Final probability of 0 mpy damage rate per sq ft					99.99%	0

In this example, the unmitigated exposure rate is taken to be 16 mpy. So, unprotected metal is assumed to be corroding at that rate. The column “Scenario Prob” shows the estimated probability of the corrosion rate shown in the “MPY” column. Note that there are only two possibilities: corrosion at full rate (no mitigation) or no corrosion (fully mitigated). Adding up all events that result in 16 mpy shows that there is a 0.01% of 16 mpy. There is a 99.99% chance of 0 mpy corrosion. This simple analysis shows that the “excellent” coating—very large probability of no defects—coupled with 90% effective CP, results in a 0.01% chance of 16 mpy damage rate per square foot of pipe. The resulting ‘probability of active corrosion points’ is then used for TTF estimates, leading to estimates of PoF for external corrosion.

8.1.2 Other Time-Dependent Mitigations

See discussions in text.

8.2 Mitigation of Time-Independent Failure Mechanisms

For time-independent mechanisms such as third party damage, incorrect operations, and geohazards, mitigation is modeled as a reduction in exposure. Recall that for these mechanisms, exposure is measured in events per length of pipe, per time period and will usually span several orders of magnitude. Therefore, mitigation will also span orders of magnitude. For instance, a very effective mitigation might reduce event frequency from 10 times per year to once every 100 years.

Discussion of mitigation for specific time-independent failure mechanisms is shown in a subsequent section of this document.

9. PoF: Resistance

Resistance, as the second component of threat reduction—along with mitigation—allows ready distinction between the damage potential and the failure potential. Resistance is simply the ability to resist failure in the presence of the failure mechanism. For time-dependent mechanisms, it is a measure of available strength, including:

- Wall thickness,
- Wall thickness “used up” for known loadings,
- Possible weaknesses in the wall, and
- Material strength including toughness.

For time-independent mechanisms, resistance includes the above factors plus considerations for external loadings:

- Buckling resistance,
- Puncture resistance,
- Diameter to wall thickness (D/t) ratio, and
- Geometry.

This is where the model considers most construction and manufacture issues involving longitudinal seams, girth welds, appurtenances, and metallurgy, as discussed in a later section.

Inspection and testing provide essential input into estimations of a pipeline’s ability to resist failure. Section 6.3 discusses how integrity assessment data of varying ages and accuracies can be integrated. This section continues that discussion and suggests an overall methodology to assessing resistance.

The key concepts of the new approach to modeling this aspect of PoF are summarized as follows:

- All wall thickness measurements available for a location on a pipeline are made equivalent by adjusting each for age and accuracy.
- Using age and accuracy of all measurements allows integration of many pieces of data collected over many years—a frequently encountered situation.
- Once equivalent, the more optimistic—not the more conservative—measurement can be used as the best estimate of pipe wall today.
- Using the more optimistic value ensures that inferior information is overridden by better information when available. (Ensuring equivalency is the key to being able to select the more optimistic value).
- In the absence of useful measurements, the wall thickness implied by current NOP is used.
- The risk-reducing value of all types of inspection and testing is immediately shown by this approach.

The power of this approach is that it can immediately and transparently integrate all of the following information:

- Pipe specification;
- Last measured wall thickness;
- Age of last measured wall thickness;
- Wall thickness "measured" (implied) by last pressure test;
- Age of last pressure test;
- Detection capabilities of last inspection (ILI, etc), including data analyses and confirmatory digs;
- Age of last inspection
- Estimated metal loss mpy since last measurement;
- Estimated cracking mpy since last measurement;
- Maximum depth of a defect remaining after last inspection;
- Maximum depth of a defect surviving at last pressure test
- Maximum depth of a defect surviving at normal operating pressure (NOP) or last known pressure peak;
- Penalties for possible manufacturing/construction weaknesses

This includes information gathered from every excavation and direct examination, as well as indirect inspection and test. This method also allows inclusion of non-pipe components into the same assessment.

Once a defensible estimate of current pipe wall is made, remaining life or failure potential estimates can be made. These estimates can be based on leak criteria, rupture criteria, or both.

9.1 Current Pipe Strength

An evaluation of pipe strength is critical to risk assessment and plays a large role in evaluating failure probability from all mechanisms, but especially the time-dependent mechanisms of corrosion and fatigue.

Current pipe strength and ability to resist failure incorporates pipe specifications, current operating conditions, recent inspection or assessment results, unknown pipe properties such as toughness and seam condition, as well as known or suspected stress concentrators and special external loading scenarios. This model captures these factors in a variable called "effective pipe wall."

Aspects of structural reliability analysis (SRA) are implicit in this approach since probability of defects is being overlaid with stresses or loads. A very robust SRA will use probability distributions to fully characterize the loads and resistances-to-loads, while this simplified approach uses point estimates. Simplifications employed here allow more direct calculations and avoid the need for iterative analyses (such as Monte Carlo type simulations) sometimes seen in the more robust SRA calculations.

Measured pipe wall thickness could be used directly to calculate remaining strength (available wall) if we have confidence that

- The measurement captures all defects that currently exist.

- There are no imperfections/weaknesses in the steel.
- There are no unintended stresses that are “using up” some strength.

Realistically, all measurements have limitations and many pipelines will have some age-of-manufacture issues as well as other issues that make us question the true available pipe strength, regardless of what a wall thickness measurement suggests. Issues include low freq ERW seam, inclusions, laminations, low toughness, girth weld processes, weakenings from other threat exposures, etc. Effective pipe wall captures such uncertainty about true pipe strength by reducing the estimated pipe wall thickness in proportion to uncertainty about possible wall weaknesses.

9.2 Estimates of Effective Pipe Wall

This is a more complex aspect of the risk evaluation—necessarily because the use of available and anticipated information must be done in several iterative steps. The challenge lies in determining what information tells us the most about the pipe condition today. A coarse measurement of wall thickness, taken recently, can be more valuable than a very precise measurement taken long ago. Every piece of direct or indirect data might yield the most useful knowledge.

In simultaneously considering all available information, the model not only produces the best estimate of pipe wall thickness, it is also able to much more accurately respond to queries regarding the benefit of performing new pressure tests or new ILI. The benefit is readily apparent, as are suggested re-assessment intervals. All data and all assumptions about exposure and mitigation are easily viewed and changed to facilitate model tuning and/or what-ifs.

The analysis begins with what is known about the pipe wall. In general, an owner will usually know at least:

1. That the pipe has integrity at its current pressure and stress condition (NOP).
2. The wall thickness that was last measured (visual, UT, ILI, implied by pressure test, etc or default to nominal design).

The beginning point of the analysis is these two factors. In addition, the owner (in the US) is now also normally required by regulation to estimate the potential for damages to the pipe since the last inspection. That estimated damage rate is used to calculate an effective wall thickness at any time after the last measurement was taken. A new integrity verification inspection or test will 're-set the clock'.

9.3 Procedure to Estimate Resistance

The steps required in the model's time-dependent failure mechanism analysis may initially appear to be daunting. However, they are actually very intuitive and leave a trail of very useful information:

1. NOP-based wall: Produce an estimated wall thickness, based on leak-free operation at current NOP (if that assumption is defensible). This may include an estimate of the deepest non-leaking defect that could be present at this pressure.
2. Pressure test based wall: Calculate a pipe wall thickness inferred by the most recent pressure test. This might be the original post-construction test. It can also be a recent higher-than-normal pressure to which the segment has been exposed.
3. Inspection based wall: Calculate an estimated wall thickness based on the most recent inspection. This is normally ILI, but can also be bell hole exams where reliable and comprehensive wall thickness measurements were taken. The accuracy of the inspection for all types of possible defects should be a part of this estimate. It will sometimes be prudent to model different defect types since detection sensitivities vary (see Table 9.3.3).
4. Exposure: Produce an estimate of steel metal-loss / crack-growth in the absence of any mitigation. This includes at least external corrosion, internal corrosion, and cracking and should reflect, for instance in the case of external corrosion, the corrosion rate of the pipe if it was buried uncoated and unprotected in the segment's environment.
5. Mitigation: Evaluate the effectiveness of current mitigation measures. This effectiveness will be used to directly offset (reduce) the steel deterioration rate that would otherwise occur.
6. Estimated pipe wall: Calculate an estimated pipe wall. This is the larger of the pipe wall thickness estimates based respectively on:
 - NOP (and largest surviving defect),
 - Last pressure test minus possible metal-loss / crack-growth since, and
 - Last inspection minus possible metal-loss / crack-growth since.

Additional estimates might be warranted when metal-loss and cracking scenarios are evaluated separately.

7. Effective pipe wall: Assign a penalty to reduce pipe strength whenever there is the possibility of manufacture or construction issues that would reduce the pipe's strength. This is an adjustment factor used to move from an estimated pipe wall to an effective pipe wall.
8. Resistance: Calculate the available pipe wall by comparing the effective pipe wall with the wall thickness needed to contain NOP.

It is recognized that this modeling approach makes several simplifying assumptions that do not fully account for the complex relationships between anomaly sizes, types, and configurations with leak potential, rupture potential, and fracture mechanics theories. In addition, metal loss and cracking phenomena have been shown to progress in non-linear fashion—sometimes alternating between rapid progression and complete stability. A constant deterioration rate is used only as a modeling convenience in the absence of more robust predictive capabilities. It should be noted that remaining strength calculations and TTF estimates should not be taken as precise values but rather as relative measures that characterize overall system behavior but may be significantly inaccurate for isolated scenarios.

Nonetheless, after accounting for uncertainty and application of appropriate safety factors, the TTF values directly support integrity management in a way that previous approaches could not. A re-assessment interval is readily apparent from these calculations. Integrity assessment schedules can be directly linked to calculations that fully integrate all pertinent data.

More details of these overall steps used to estimate resistance to failure follow.

9.3.1 NOP-Based Wall

This is the entry point into the effective pipe wall estimate. The wall thickness implied by leak-free operation at normal operating pressure (NOP) or a recent high pressure can be calculated by simply using a hoop stress calculation with NOP to infer a minimum wall thickness. Since defects can be present and not be causing failure, a value for “max depth of defect surviving NOP” can also be assumed. The depth of defect that can survive at any pressure is a function also of the defect’s overall geometry. Since countless defect geometries are possible, assumptions are required.

With assumptions, a wall thickness based solely on operating leak-free at NOP, $pipe_wall_NOP$, can be inferred using the Barlow formula for stress in the extreme fiber of a cylinder under internal pressure as follows:

$$pipe_wall_NOP = ([NOP]*[Diameter]) / (2*[SMYS]*1000)$$

This simple analysis can account for defects that are present but are small enough that they do not impact effective pipe strength by adjusting for an assumed population of possible defects. There is some precedent in using 80% to 90% of the Barlow-calculated wall thickness to allow for non-critical defects that might soon grow critical. The analysis could be made even more robust by incorporating a table or chart of defect types and sizes that could be present even though the pipe has integrity at NOP. An appropriate value can be selected knowing, for example, that a pressure test at 100% SMYS on 16", 0.312, X52 pipe could leave anomalies that range from 90% deep 0.6" long to 20% deep, 12" long. All combinations of geometries having deeper and/or longer dimensions would fail. Curves showing failure envelopes can be developed for any pipe.

Of course, the estimate of $pipe_wall_NOP$ pre-supposes that the portion of pipe being evaluated is indeed not leaking and is exposed to the assumed NOP.

9.3.2 Pressure-Test Based Wall Thickness Estimate

To conclude that there is greater wall thickness than implied by NOP, additional information must support that premise. The above analysis can be repeated using the most recent test pressure instead of the NOP. Then, using an estimated deterioration rate (discussed below) from the time of test until today, another pipe wall estimate is produced. This can be compared to the estimate from the NOP. If this conservatively estimated deterioration since the test suggests a larger wall thickness than implied by the Barlow-NOP calculation, then this value can be used instead of the minimum wall needed for NOP per Barlow.

9.3.3 Inspection-Based Wall Thickness Estimate

The calculation next uses the last actual measurement taken, including the uncertainty surrounding the measurement and the age of the measurement. This measured pipe wall value will override the other wall thickness estimates (from NOP and since the last pressure test), if the measured value shows with confidence that even more pipe wall is available. The capability of the measurement tool and the validation process is important. Better knowledge, obtained by either greater detection capability of all possible defects and/or a more aggressive validation program, reduces uncertainty in the measurement. As with the estimate based on pressure test, this estimate should include possible degradation of pipe wall since the measurement.

If an integrity assessment, including accuracy considerations, indicates “no anomaly,” there could nonetheless be an anomaly present that is just below the detection capability of the assessment. The modeler can use an assessment of integrity inspection capability (IIC) to adjust all measured or inferred wall thicknesses. The adjustment should be based on the largest surviving defect after the most recent inspection. It can also somewhat consider the severity of the defect—how much might it contribute to likelihood of failure. For instance, a detected lamination is normally not a significant threat to integrity unless it is very severe or also has the potential for blistering or crack initiation, both of which are very rare.

A complication in evaluating IIC is that several defect types must be considered. IIC is not consistent among inspection tools and defect types, so some generalizations are needed. Examples of defect types include metal loss (internal or external corrosion), axial cracks, circumferential cracks, narrow axial corrosion, long seam imperfections, SCC, dents, buckles, laminations, inclusions. Inspection or assessment techniques often focus on one or two of these with limited detection capabilities for the others. Since most ILI assessments provide unequal information on cracking versus metal loss, a two-part calculation is required in the TTF assessment. This is illustrated in Example 1 below.

A matrix can be set up capturing the beliefs about IIC. For example, see Table 9.3-1, below.

Table 9.3-1. IIC Benefit Matrix

Inspection Type	Validation (Pig-Digs) Protocol	Defect Type, max surviving defect depth as % of wall thickness							Max Surviving Defect	
		External Corrosion	Internal Corrosion	Axial Crack	Circum Crack	Dent/Gouge	Ovality/Buckling	Laminations	Metal Loss	Crack
MFL high resolution ILI	Aggressive	5	5	100	10	20	50	50	5	100
	Routine	10	10	100	50	50	50	50	10	100
	Min	15	15	100	100	50	50	50	15	100
MFL std resolution ILI	Aggressive	10	10	100	50	50	50	50	10	100
	Routine	15	15	100	100	50	50	50	15	100
	Min	20	20	100	100	50	50	50	20	100
Ultrasound ILI		5	5	100	100	20	20	5	5	100
TFI ILI		20	20	5	10	50	50	50	20	5
EMAT ILI		50	50	10	10	50	50	10	50	10
Ultrasound shear wave crack tool ILI		50	50	10	10	50	50	10	50	10
Caliper, sizing, gauging, inertial ILI		100	100	100	100	5	5	100	100	100
Pressure test		10	10	10	10	80	80	80	10	10

Other inspection techniques such as various forms of NDE UT (guided wave, transverse wave, etc), X-ray , etc can be added to this comparative matrix. This matrix is a simplification and is offered only as an example. It must be modified to reflect full consideration of all pertinent variables, before it is used to adjust real measurements.

Values shown represent defect sizes (depths normally), expressed as percentage of wall thickness, that might remain after the assessment. A value of 100 means that the assessment technique has no detection capabilities for that defect type—a near 100% thru-wall defect could remain after the inspection. The last 2 columns aggregate the various defects into two categories and assign an IIC to each category based on the capabilities for the specific defects. As an example of the use of this matrix, consider a pipeline that has been evaluated with a High Resolution MFL tool with a routine validation protocol. The corresponding maximum surviving defects for this assessment are 10% of wall for metal loss and 100% of wall for axial cracks. So, no information regarding crack presence is obtained.

9.3.4 Estimated Pipe Wall

The estimated pipe wall reflects the best estimate of how much metal is present and available to resist failure. This model assumes that a defect just below detection threshold exists, with “just below detection threshold” estimated using a hoop stress calculation assuming a relatively long length. Either cracking or metal loss may dominate the calculation, depending upon the estimated aggressiveness of each and the date/type of assessments performed. For many risk assessments, the two phenomena are best tracked independently.

Remaining wall thickness, or maximum surviving defect sizing, can be estimated using some simple relationships like the Barlow equation specified in US pipeline regulations. This has limitations since it does not accurately capture the effects of defect size (depth versus length and width are important) or type (cracking phenomena are not captured by the Barlow relationship). When increased accuracy is required, metal loss sizing routines such as RSTRENG and ASME B31.8G or fracture mechanics relationships can be substituted. It is recommended that the more robust calculations be used when data is available since the Barlow will produce overly conservative results. For example, in a 72% design factor pipeline, with a 12.5% wall thickness manufacturing tolerance, loss of only 15% wall would predict failure. Ignoring the manufacturing tolerance is often suggested in order to reduce the over-conservatism when Barlow is used (and this is consistent since ASME recommendations are to use nominal wall value in Barlow calculations).

9.3.5 Effective Pipe Wall

An estimated pipe wall thickness has now been created. The effective pipe wall calculation begins with this value and adjusts or penalizes it for anything that implies a reduced strength in that metal. A potential weakness is modeled as being equivalent to reduced wall thickness.

Possible manufacturing/construction weaknesses are identified for each pipeline segment. Typical age-of-manufacture/construction issues include

- Increased longitudinal seam susceptibilities (low freq ERW, for instance);
- Hard spots;
- Laminations;
- Low toughness;
- Girth weld weaknesses;
- Miter joints;
- Wrinkle bends;
- Stress concentrators;
- Sub-standard appurtenances; and
- Any other possible weaknesses.

The amount of weakness actually produced by these factors is often very situation-specific. Generalizations are used to avoid the sophisticated finite element analyses that would be required to fully model all of the possibilities. Some generalizations are available from

industry standards and even regulations. Note the seam factor used in U.S. regulations for pipeline design—this is an example of an adjustment value.

The effective pipe wall estimate can now be used for available wall calculation (time-dependent mechanisms) and in external force resistance models (third party pipe wall puncture resistance and landslide buckling resistance, for example). However, nominal pipe wall is often used in certain external force variables such as D/t and “geometry” factor since their influences are very coarse. Using effective pipe wall everywhere can also lead to troublesome to circular calculations, so some simplifications are often needed.

In this type of analysis, higher-grade (stronger) steels tend to have a higher (worse) PoF compared to lower strength steels. This is true because the mpy deterioration applies equally to all strengths of steel. So, heavier wall steel has the longest TTF, regardless of strength. If two wall thicknesses are equal, the one with the lower strength will have a longer TTF because it begins with a thicker wall under the “leak free at NOP” initial premise—i.e., it takes more wall thickness of a lower strength steel to contain the operating pressure.

When pipe grade is unknown, the often-recommended default of 24,000 psig is not conservative when calculating remaining wall thickness. Since the mpy deteriorates high strength steel as readily as low strength, using a higher SMYS default results in lower remaining wall and quicker TTF—a more conservative assessment overall.

The following examples illustrate some derivations of the effective pipe wall estimates.

Example 1

A 15 year old, non-leaking pipeline segment has a nominal wall thickness of 0.320” after accounting for the manufacturing tolerance of 12.5%. The Barlow hoop stress calculation using NOP shows that a minimum of 0.210” is required to contain normal operating pressure. (Considering also the max defect depth that could survive at this pressure, assuming a long corrosion defect, would bring the minimum wall thickness down substantially.) The conservatively estimated deterioration rate in this segment is 10 mpy from a combination of 8 mpy metal loss and 2 mpy cracking. A calculation can be done to estimate today’s wall thickness based degradation since original installation. Since it has been 15 years since construction, the calculated pipe wall is $0.320'' - (15 \text{ years}) \times (10 \text{ mpy}) \times (1 \text{ inch} / 1000 \text{ mils}) = 0.170''$. The minimum wall implied by NOP is higher, so the current estimate of pipe wall thickness is 0.210”.

A high resolution MFL ILI tool with routine confirmation excavations (follow-up) is subsequently used to assess integrity. This technique is assumed to have no capabilities to detect longitudinal crack-like indications and +/-10% accuracy of metal loss anomalies,. The assessment measures a minimum wall in this segment of 0.300”. So, ILI-estimated wall thickness for metal loss is $0.300 \times 90\% = 0.270''$. For cracking, the available wall could actually be 0.00” since the integrity assessment is assumed to have no detection capabilities. We can now adjust the estimated wall with cracking to be (wall after metal loss) – (mils potentially lost by cracking) = $0.270 - (2 \text{ mpy} \times 15 \text{ years}) = 0.240''$. Since there is not a measured value (sensitive to presence of cracks) to override this estimate, then it shall become the value for pipe wall estimated based on possible cracking.

These values along with the corresponding damage rates are used to set re-assessment intervals for cracking and metal loss mechanisms respectively.

For an overall pipe wall estimate, we should use the crack-adjusted value of 0.240". This is then adjusted for possible metal weaknesses to get effective pipe wall which can be used to estimate the ability of the pipe to withstand other threats such as external forces.

Without the ILI, the pipe wall would have been assumed to be 0.210". So, the ILI improved the risk picture by removing some uncertainty. This was done by a direct metal loss measurement and an adjustment for possible cracking. The ILI information also may prompt a revision of the deterioration rate, further reducing the conservatism brought on by uncertainty.

Example 2

Same scenario as above except that a 1.25 x MOP pressure test is the chosen integrity assessment technique. This technique is modeled to have a capability to find all defect types to the extent that they fail at the test pressure. The Barlow calculation using the test pressure indicates a minimum effective wall thickness of 0.263". (Note that a 1.5 x MOP test would have led to a 0.315" wall.) So, 0.263" is the value for pipe wall thickness estimate to be used in obtaining effective pipe wall.

This example assumes that a more robust inspection is achieved via pressure test, so risk is reduced more than in the previous example where a defect-specific ILI was used. That assumption will not always be valid.

In both of these examples, extreme cases of defects left behind after integrity assessment are possible. The user might wish to include additional conservatism by allowing for the possibilities of such defects. Note that both examples concluded much less pipe wall than nominal wall thickness. This will often be sufficient conservatism.

9.3.6 Resistance (Available Pipe Wall)

The difference between the available pipe wall thickness and the thickness required for anticipated loads (internal pressure, external loads) is the thickness of metal that can be lost before failure occurs (using "leak" criteria; "rupture" criteria might predict failure with less wall loss, depending on defect size assumptions). This estimated "extra" wall thickness represents a safety margin—failure potential is reduced as this increases since the TTF will be increased. This "available wall" can be also used in subsequent estimates of resistance to other failure mechanisms such as external forces. Similarly, the available wall estimate can be reduced on the basis of other results from the risk assessment. For instance, when external forces are "using up" more pipe strength, this reduces strength available to withstand other failure mechanisms.

Again, some significant simplifying assumptions underlie this value and should be carefully considered by the modeler.

10. Time Independent Failure Mechanisms

A general discussion of PoF assessment for time-independent failure mechanisms is offered here. The calculations leading to PoF for time-independent failure mechanisms is more straightforward than for time-dependent mechanisms. As previously noted, the underlying form of PoF estimation is as follows:

$$\text{PoF} = [\text{unmitigated event frequency}] \times (1 - [\text{Threat Reduction}])$$

Where
[Threat Reduction] = $f(\text{mitigation effectiveness, resistance})$

Threats modelled as mostly random in nature, third party, theft, sabotage, incorrect operations, geohazards, etc, are sensitive to segment length since the threat is assumed to be uniformly distributed across the entire segment. This results in a failure rate per length per time period (such as PoF / mile / year) which is then multiplied by the segment length to get a failure probability for the segment. A direct summation of failure probabilities is acceptable when numerical values are very small. Otherwise, a probabilistic summation, similar to the way in which all threats are combined, is used.

One of the keys to the new approach in risk assessment is to capture the orders of magnitude spans between risk levels. Older scoring systems did not normally provide for this.

The best possible value for each mitigation variable is determined based on that variable's perceived ability to independently mitigate the threat. The mitigation captures the possible span—orders of magnitude—of exposure reduction.

Discussion and notes regarding some assessments for specific failure mechanisms follow for the remainder of Section 10 of this document. This material begins the transition of specific threat analyses from the 3rd edition of Pipeline Risk Management to the 4th edition. The patterns shown in these examples can be applied to any other time-independent failure mechanism.

10.1 Third Party

Third party damage is modeled as a time-independent failure mechanism. It can be assumed that any third party damage that does not result in an immediate failure, initiates a time-dependent mechanism such as corrosion or fatigue.

10.1.1 Exposure

Exposure is the estimated events/per mile-year from excavation activity and certain other external forces. Unless considered elsewhere in the model, impacts should include:

- Excavation—farm equipment, construction equipment, dredging, boring, and piles.
- Traffic—vehicles, rail, marine, air.
- Falling objects—trees, utility poles, buildings, meteors, etc (anything that could fall onto the ROW).

External impacts—landslides, rock falls, etc—are normally considered in the Geohazard assessment.

Recall that all exposures are evaluated in the absence of mitigation. For example, the unmitigated exposure from falling trees might be estimated to be on the order of several times per year—perhaps coinciding with severe storm frequency. It is only after adding mitigation—notably depth of cover—that the threat appears as small as most intuitively believe it is.

See also examples in Section 7.

10.1.2 Mitigation

Mitigation = f (cover, patrol, one-call, damage prevention program, ROW condition, signs/markers, etc.). Some comments on measuring effectiveness of some specific mitigation measures follow:

- One-call effectiveness is generally an AND gate between sub-variables such as system type, notification requirement, and response. The AND gate is applicable since all sub-variables together represent the effectiveness of the mitigation.
- The mitigation of patrol is normally an AND gate between patrol type and frequency. Patrol type implies an effectiveness and includes combinations of different types—ground-air, for example.
- External protection is typically an OR gate between cover, warning mesh/tape, exterior protection since each measure can act independently to reduce the PoF.
- Casing is a mitigation if it is thought of as something added to a pipeline system. If it is considered part of the pipeline system, then it is a resistance. Either categorization can be used since either will have the same impact on PoF.

Maximum effectiveness of each mitigation measure should represent the modeler's belief about how much the failure potential is reduced by that measure independent of any other measure. For example, casing pipe or extreme depth of cover should probably warrant 95-99% reduction in exposure levels.

10.1.2 Resistance

Resistance = f (pipe wall thickness, pipe geometry, pipe strength, stress level, manufacturing and construction issues). The pipe wall thickness and material toughness can be used to assess puncture resistance. The geometry, diameter and wall thickness, can measure resistance to buckling and bending. Since internal pressure induces longitudinal stress in the pipe, a higher internal pressure can indicate reduced resistance to external forces.

10.2 Incorrect Operations

The time-independent failure mechanism of human error is measured as “incorrect operations.” Specifically, the potential for failure directly precipitated by a “real time” human error, is measured here. Examples of such failures include improper operation of valves causing overpressure or disabling of control or safety devices. Such failures usually

require that a sequence of unlikely events, including failure of highly reliable safety devices, all occur.

As a modeling convenience and due to the normally consistent aspects of human error reduction across all failure mechanisms, the role of possible human error in all other failure mechanisms is often also assessed in one location in the risk analysis. This includes the potential for error in design and maintenance activities related to safety systems, corrosion control, third party damage prevention, and others. Results of this analysis are used to adjust mitigation effectiveness estimates. When human error potential is higher, mitigation effectiveness is conservatively assumed to be lower. For example, when procedures or training are found to be inadequate, then effectiveness of corrosion control methods might be more suspect; when instrument calibration and maintenance records are missing, effectiveness of safety devices is questionable.

10.2.1 Exposure

For exposure estimates, abnormal, unintended, inappropriate actions that could lead to pipeline failure are “events.” Exposure is measured as frequency of “events” . Measures employed to avoid an incident are mitigation. Ability of the system to resist a failure when exposed to an incident is resistance. So, stress level is resistance as well as exposure and is appropriately included in both aspects of the analysis.

The unmitigated exposure level for this mechanism should be based on a completely untrained workforce, with no procedures in place, no records of design or maintenance, no SCADA benefits, etc. As with some other exposure estimates, such an unmitigated scenario may require some imagination on the part of the assessors. As a modeling convenience, some integrity-threatening events such as surge and thermal overpressure are included in this threat category.

Exposure level should include an assessment of all pressure sources that can overpressure the pipeline segment of interest. Sources of potential overpressure typically include source pressure, thermal overpressure, and surges. All of these are modeled as real time human error exposures. Safety devices are ignored at this point in the analysis. Each source is assigned an event frequency, based on how often the overpressure event is theoretically possible. When the threat is continuous, a pre-set value can be assigned. An example is a pipeline segment downstream of a pressure-reducing regulator that protects the pipe from high upstream pressure (remember, no benefit from the safety device is credited yet). When this high pressure source can overpressure the segment, the exposure is continuous. Surge potential can also be considered in this part of the model.

10.2.2 Mitigation

Mitigation measures typically thought to reduce failure potential include:

- Safety systems,
- Training,
- Procedures,
- Proactive surveying,
- Maintenance practices,

- Materials handling,
- Quality assurance,
- Hazard Identification, and
- Others.

Some of the less obvious mitigation measures are briefly discussed below.

- ✓ [Surveys] is a mitigation variable that shows how much proactive information collection, digestion, and reaction to new information is being done. It overlaps aspects of surveys employed in other threat mechanisms (CIS, aerial patrol, depth cover, etc) but additional “credit” given here as evidence of overall corporate philosophy of proactively addressing possible exposures.
- ✓ [Maintenance practices] indicates a sensitivity to keeping things in high working order. It should be an AND gated variable combined with variables such as one measuring effectiveness of safety devices, since the latter requires the former in order to realize its full capability.
- ✓ [Materials] captures the company’s processes to ensure correct materials are used. This includes material selection and control as replacements/additions to the system are made.
- ✓ [QA] applies to quality control checks in design, construction, operations, and maintenance. The ability of such measures to reduce exposure can be assessed.
- ✓ [HazID] captures programs that identify and prompt appropriate actions to avoid human errors.

10.2.3 Resistance

The segment’s resistance to human error caused failures can be modeled as a function of:

- System safety margins,
- Stress level (% SMYS),
- Time-to-overpressure,
- Etc.

10.3 Geohazards

The potential for damages or failure from geologic or hydraulic forces, although a relatively rare threat for most pipelines, can be the main risk driver for certain segments and a challenge for risk assessment.

10.3.1 Exposure

One way to measure this exposure threats is to sum the contributions from each of three geohazard categories:

$$\text{Exposure} = [\text{geotech}] + [\text{hydrotech}] + [\text{seismic}]$$

Where:

Geotech = [landslide probability] * [landslide severity]

Hydrotech = [erosion]+[subsidence]+[buoyancy]+[flood-bank erosion]+[flood-undercut]+[debris loadings]

Seismic = [fault] + [liquefaction]

Fault = expected failure rate due to fault actions

Liquefaction = [peak ground acceleration (PGA)] * [soil suscept]

This general failure mechanism category includes mechanisms of two specific types: those that produce constant forces and those that produce random events. The constant forces can be modeled as continuously “using up” available pipe strength, thereby reducing resistance to other failure mechanisms. Of high priority would be the identification of coincident application of such geohazards with pipe weaknesses or higher exposures to other failure mechanisms. The forces generating random events are usually better modeled as non-continuous.

The process for assigning PoF values to these phenomena should include the use of historical incident rates and published recurrence interval data whenever available.

10.3.2 Mitigation

Mitigation measures are often phenomena-specific if not situation-specific and might require special handling in the assessment. Mitigation measures typically thought to reduce failure potential include:

- Strain gauges,
- Barriers,
- Soil removal,
- Erosion control structures,
- Drain control, and
- Etc.

10.3.3 Resistance

Resistance can be assessed in a fashion similar to third party (refer to Section 7.7). Resistance measures typically thought to reduce failure potential include:

- Supports,
- Anchors, and
- Pipe designs.

11. Final PoF

All failure probabilities and risk valuations should be normalized spatially and temporally. Units such as failures/mile-year are convenient for all mechanisms except equipment and certain materials/construction issues. In those cases, a segmentation strategy or assumptions regarding anomalies per unit length can be used to normalize.

The relationship between leak frequency and failure probability is often assumed to be exponential. The exponential relationship fits many observed rare-event phenomena and is frequently used in statistical analysis.

The probability of no events can be calculated from:

$$P(X)EVENT = [(f * t)^X / X !] * \exp (- f * t)$$

Where:

P(X)EVENT = probability of exactly X events

f = the average spill frequency for a segment of interest, events /year

t = the time period for which the probability is sought, years

X = the number of events for which the probability is sought, in the pipeline segment of interest.

The probability for one or more events is evaluated as follows:

$$\begin{aligned} P(\text{probability of one or more})EVENT &= 1 - \text{Probability of no spills} \\ &= 1 - P(X)EVENT \\ &\text{where } X = 0. \end{aligned}$$

At very small event frequencies, the probability values are equal to the event rates. So, the two can be used interchangeably until the event rates become higher.

In the risk assessment, a probability of failure is calculated for each pipeline segment for each threat. Under the assumption that each failure mechanism is basically independent, these probabilities are combined through an OR gate equation to give an overall failure probability for the segment. The segment probabilities are combined to give an overall PoF.

$$PoF = f(PoF_{\text{time-indep}}, PoF_{\text{time-dep}})$$

PoF values associated with each failure mechanism are combined using the widely accepted premise in probability theory that the “chance of one or more failures by any cause” is equal to 1 minus “the chance of surviving cause A” times “the chance of surviving cause B” times, etc. So, for a model that has categorized threats into third party, TTF, theft/sabotage, incorrect operations, and geohazard, the relationship would be:

$$PoF_{\text{overall}} = 1 - [(1 - PoF_{\text{thdpty}}) \times (1 - PoF_{\text{TTF}}) \times (1 - PoF_{\text{theftsab}}) \times (1 - PoF_{\text{incops}}) \times (1 - PoF_{\text{geohazard}})]$$

Where P_X = Failure Probability associated with failure mechanism X (Prob of one or more failures/ (mile*yr) or other appropriate units)

A simple summation of failure probabilities is acceptable when numerical values are very small.

While the assumption of independence is made for purposes of probabilistic math, dependences can also be modeled. For example, the effective pipe wall calculated in the TTF routines can be used in the resistance calculations for external forces. Similarly, the effects of external loadings can influence the “available wall” calculations in the TTF routines.

12. Calibration and Validation

Users should be vigilant against becoming too confident in using any risk model output without initial and periodic ‘reality checks’. Especially when risk output is expressed in numbers that appear to be very precise, it is easy to fall into what has been termed an “illusion of knowledge.” Regardless of the extent of the modeling rigor employed, assumptions and simplifications are still needed in any analysis. The uncertainty surrounding a risk assessment cannot be eliminated and a model without some simplifications is not justifiable in such a high uncertainty environment. The very nature of extremely rare events makes planning difficult. For example, an event might be very precisely measured to have a recurrence interval of 88 years. This is very useful information when the event is compared to many other events with say, intervals of 2 years and 250 years. However, the once-every-88 year event could occur next year, in year 24, in year 67, or even multiple times in any year. Decision-making should recognize this at all times.

12.1 Calibration vs. Validation

For some applications of pipeline risk assessment, especially in the early stages, relative risk values are the only values that will be required. Relative values can often adequately support prioritization and ranking protocols. The need for calibration—tuning model output so that it mirrors actual event frequencies—might be unnecessary in initial stages. In that case, only validation—ensuring consistent and believable output from the model—is required.

Prior to the need for PoF results expressed in absolute terms—failures per mile-year, for instance—the PoF values can be stripped of their time period implication and be used as relative numbers. A 2.3% PoF does not mean a 2.3% annual probability of failure until the risk assessment has been calibrated—it only means a 2.3% chance of failure over some time period. This might be one year or one hundred years. Until the calibration is done, the 2.3% value can be used as a relative measure of PoF.

Experience has shown that risk management permeates so many aspects of the organization that a good risk model’s role will eventually be expanded. As its output becomes more familiar, new users and new applications arise. Ultimately most assessments will be asked to anchor their output in absolute if not monetary terms. When this happens, the need for both validation and calibration arises.

A risk assessment model producing estimates of absolute risks—events or costs per pipeline length and per time period—is most useful when it is calibrated or “tuned” to produce results that are consistent with beliefs about the real failure probabilities. Such beliefs are normally based on historical experience, tempered by knowledge of changing factors. The process of calibrating risk assessment results begins with establishing plausible future leak rates based on relevant historical experience. These rates become ‘targets’ for risk assessment outputs, with the belief that large populations of pipeline segments, over long periods of time, would have their overall failure estimates approach these targets. The risk assessment model is then adjusted so that its outputs do indeed approximate the target values.

12.2 Use of Incident History

Incident history is one of the important pieces of evidence to consider when both validating and calibrating risk assessment results. This includes all incidences of measured metal loss, crack-like indications, damages found, anomalies detected, plus actual failures. In most cases, analyses of all previous repairs will be useful and relevant.

An incident impacts our degree of belief about future failure potential in proportion to its relevance as a predictor. Some will directly impact exposure estimates. Even if it has little or no direct relevance as a predictor, the related investigation would certainly yield information useful in effective pipe wall calculations.

A mechanism must exist to remove the incident's "penalty" in the risk assessment when there is no longer any relevance. An example would be where an ineffective coating is the root cause of a corrosion incident and that coating is subsequently replaced. Another example is a high incidence of third party damages or near-misses associated with some land use that has since changed.

All PoF estimates can be calibrated by using relevant historical failure rates when available. This generally involves the following steps:

- Perform detailed analysis of historical leak data.
- Evaluate data in the context of similar pipelines (similar environments and O&M practices) in other companies.
- Determine relevance of each incident to all segments of pipeline.
- Use relevant data to calibrate or tune the algorithms so that absolute risk levels—expressed in annualized costs, for example—can be produced.

Failures outside of the segment of interest might or might not be relevant so historical data should be used only in context of engineering judgment and experience.

12.3 SME Validation

Similar to the use of a benchmark for model validation, a carefully structured interview with SME's can also identify model weaknesses (and also often be a learning experience for SME's). If an SME reaches a risk conclusion that is different from the risk assessment results, a drill down into both the model and the SME's basis of belief should be done. Any disconnect between the two represents either a model error or an inappropriate conclusion by the SME. Either can be readily corrected. The key is to identify where the model and the SME first diverge in their assumptions and/or conclusions.

A good objective of risk assessment should be to make the risk assessment model house the collective knowledge of the organization—anything that anyone knows about a pipeline's condition or environment, or any new knowledge of how risk variables actually behave and interact, can and should be captured into the analysis protocol.

12.4 Diagnosing Disconnects Between Results and Reality

If model results are not consistent with a benchmark believed to closely represent performance of the system, any of several things might be happening:

- Benchmark is not representative of the assessed segments,
- Exposure estimates are too high or too low,
- Mitigation effectiveness is judged too high or too low, and
- Resistance to failure is judged too high or too low.

The distinction between PoF and probability of damage (but not failure) can be useful in diagnosing where the model is not reflecting perceived reality. If damages are predicted but not occurring, then the exposure is overestimated and/or the mitigation is underestimated.

Mitigation measures have several aspects that can be tuned. The orders of magnitude range established for measuring mitigation is critical to the result, as is the maximum benefit from each mitigation, and the currently judged effectiveness of each. More and more research is becoming available and can often be used directly in judging the effectiveness of a mitigation measure.

A trial and error procedure might be required to balance all these aspects so the model produces credible results for all inputs.

Point Estimates

Even though the more robust algorithms discussed here use almost all pertinent information, they are still normally set up to receive and produce point estimates only. In reality, many variables will vary over time as well as along a pipeline. To better model reality, the changes in many parameters like pressure, soil resistivity, wall thicknesses, etc should be captured by creating a distribution of the variations over time or space. Such distributions can also at least partially quantify the uncertainty surrounding all measurements. The range of possibilities for all pertinent variables must be understood and accounted for in producing the risk estimates.