

Reliability Model for Underground Gas Pipelines

R.M. Cooke, E. Jager*, D. Lewandowski

Dept. Mathematics

Delft University of Technology

r.m.cooke@its.tudelft.nl

Abstract

A model is constructed for the failure frequency of underground pipelines per kilometer year, as a function of pipe and environmental characteristics. The parameters in the model were quantified, with uncertainty, using historical data and structured expert judgment. Fifteen experts from institutes in The Netherlands, the United Kingdom, Italy, France, Germany, Belgium, Denmark and Canada participated in the study.

Keywords

Underground pipeline, corrosion, third party interference, marked point process, uncertainty analysis, expert judgement, risk analysis.

1 Introduction

Many countries have invested extensively in underground gas pipelines in the 1960's and 1970's. These pipes are approaching the age at which problems

*Nederlands Gasunie

of corrosion are expected to appear with increasing frequency. These countries may be facing massive investment. The Netherlands for example has large gas reserves and is a major exporter of natural gas. The national gas company, Gasunie, is responsible for maintaining over 11,000 km of underground gas pipelines with a current replacement value of about 15 billion US dollars. Over the last forty years, however, several technologies have been introduced to protect pipes against corrosion and to detect and replace weak points. If, as is generally believed these efforts have had a positive effect, then predicting these effects is important for impending decisions regarding inspection and repair. This article describes a recent effort to upgrade the basis for decisions regarding inspection and replacement of underground pipelines¹.

Previous studies (see for example, Kiefner et al 1990) focused on developing ranking tools which provide qualitative indicators for prioritizing inspection and maintenance activities. Such tools perform well in some situations. In The Netherlands, however, qualitative ranking tools have not yielded sufficient discrimination to support inspection and maintenance decisions. The population of gas pipelines in The Netherlands is too homogeneous. Moreover, as the status of current pipes and knowledge of effectiveness of current technologies is uncertain, it was felt that uncertainty should be taken into account when deciding which pipelines to inspect and maintain.

We therefore desire a quantitative model of the uncertainty in the failure frequency of gas pipelines. This uncertainty is modeled as a function of observable pipeline and environmental characteristics. The following pipe and environmental characteristics were chosen to characterize a kilometer year of pipeline (Basalo 1992, Lukezich et al 1992, Chaker and Palmer 1989):

Pipe Characteristics

- pipe wall thickness
- pipe diameter

¹This article is based on Cooke and Jager (1998), but is expanded to account for new measurement data

- ground cover
- coating (bitumen or polyethylene)
- age of pipe (since last inspection)

Environmental Characteristics

- frequency of construction activity
- frequency of drainage, pile driving, deep plowing, placing dam walls,
- percent of pipe under water table
- percent of pipe exposed to fluctuating water table
- percent of pipe exposed to heavy root growth
- percent of pipe exposed to chemical contamination
- soil type (sand, clay, peat)
- pH value of soil
- resistivity of soil
- presence of cathodic protection
- number of rectifiers
- frequency of inspection of rectifiers
- presence of stray currents
- number of bond sites

Although extensive failure data is available, the data is not sufficient to quantify all parameters in the model. Indeed, the data yield significant estimates only when aggregated over large populations, whereas maintenance decisions must be taken with regard to specific pipe segments. Hence, the effects of combinations of pipe and environmental characteristics on the failure frequency is uncertain and is assessed with expert judgment. The expert judgment method is discussed in “Expert Judgment in the Uncertainty

Analysis of Dike Ring Failure Frequency” (Cooke and Slijkhuis, this volume) was applied (Cooke 1991). Fifteen experts participated in this study, from The Netherlands, Germany, Belgium, Denmark, The United Kingdom, Italy, France and Canada.

When values for the pipe and environmental characteristics are specified, the model yields an uncertainty distribution over the failure frequency per kilometer year. Thus the model provides answers to questions like:

- Given a 9 inch diameter pipe with $7mm$ wall laid in sandy soil in 1960 with bitumen coating etc, what is the probability that the failure frequency per year due to corrosion will exceed the yearly failure frequency due to 3rd party interference?
- Given a 9 inch pipe with $7mm$ walls laid in 1970 in sand, with heavy root growth, chemical contamination and fluctuating water table, how is the uncertainty in failure frequency affected by the type of coating?
- Given a clay soil with $pH = 4.3$, resistivity 4000 [Ohm cm] and a pipe exposed to fluctuating water table, which factors or combinations of factors are associated with high values of the free corrosion rate?

In carrying out this work we had to solve three problems:

- How should the failure frequency be modeled as a function of the above physical and environmental variables, so as to use existing data to the maximal extent?
- How should existing data be supplemented with structured expert judgment?
- How can information about complex interdependencies be communicated easily to decision makers?

In spite of the fact that the uncertainties in the failure frequency of gas pipelines are large, we can nonetheless obtain clear answers to questions like those formulated above.

The next section discusses the general issue of modeling uncertainty. Section 2 discusses the modelling of pipeline failures, sections 3 through 5 treat third party interference, environmental damage and corrosion. Section 6 presents recent data validating the model. Results are presented in section 7, and section 8 gathers conclusions.

2 Modeling Pipeline Failures

The failure of gas pipelines is a complex affair depending on physical processes, pipe characteristics, inspection and maintenance policies and actions of third parties. A great deal of historical material has been collected and a great deal is known about relevant physical processes. However, this knowledge is not sufficient to predict failure frequencies under all relevant circumstances. This is due to lack of knowledge of physical conditions and processes and lack of data. Hence, predictions of failure frequencies are associated with significant uncertainties, and management requires a defensible and traceable assessment of these uncertainties.

Expert judgment is used to quantify uncertainty. Experts are queried about the results of measurements or experiments which are possible in principle but not in practice. Since uncertainty concerns the results of possible observations, it is essential to distinguish failure *frequency* from failure *probability*. Frequency is an observable quantity with physical dimensions taking values between zero and infinity. Probability is a mathematical notion which may be interpreted objectively or subjectively. Experts are asked to state their subjective probability distributions over frequencies and relative frequencies. Under suitable assumptions, probabilities may be transformed into frequencies and vice versa. In this model the following transformations are employed. Let N denote the number of events occurring in one year in a 100-kilometer section of pipe. N is an uncertain quantity, and the uncertainty is described by a distribution over the non-negative integers. Let \mathcal{N} denote the expectation of N . If we assume that the occurrence of events along the pipe follows a Poisson distribution with respect to distance, then

$\mathcal{N}/100$ is the expected frequency of events in one kilometer of pipe. If $\mathcal{N}/100$ is much less than one, such that the probability of two events occurring in one kilometer in one year is very small, then $\mathcal{N}/100$ is approximately the probability of one event occurring in one kilometer in one year. $(1 - \mathcal{N}/100)$ is approximately the probability of no event occurring in one kilometer in one year, and the probability of no events in the entire 100 kilometers is approximately $(1 - \mathcal{N}/100)^{100}$.

The result becomes more accurate if we divide the 100 kilometers into smaller pieces. Using the fact that $\lim_{x \rightarrow \infty} (1 - \mathcal{N}/x)^x = e^{-\mathcal{N}}$, we find that the probability of no event in 100 kilometers in one year is $e^{-\mathcal{N}}$; the probability of no event in one kilometer in one year is $e^{-\mathcal{N}/100}$. The probability of at least one event in one kilometer is $1 - e^{-\mathcal{N}/100}$, and if $\mathcal{N}/100 \ll 1$, then this probability is approximately $\mathcal{N}/100$. To accord with more familiar usage, however, it is often convenient to suppress the distinction between small frequencies and probabilities.

2.1 Example of Modeling Approach

The notation in this section is similar to but a bit simpler than that used in the sequel.

Suppose we are interested in the frequency per kilometer year that a gas pipeline is hit (H) during third party actions at which an overseer from Gasunie has marked the lines (O). Third party actions are distinguished according to whether the digging is closed (CL ; drilling, pile driving, deep plowing, drainage, inserting dam walls, etc) and open (OP ; e.g. construction). Letting F denote frequency and P probability, we could write:

$$\begin{aligned} \text{Frequency}\{\text{Hit and Oversight present per km}\cdot\text{yr}\} &= F(H \cap O/\text{km}\cdot\text{yr}) \\ &= F(CL/\text{km}\cdot\text{yr})P(H \cap O|CL) + F(OP/\text{km}\cdot\text{yr})P(H \cap O|OP) \quad (1) \end{aligned}$$

This expression seems to give the functional dependence of $F(H \cap O)$ on $F(CL)$ and $F(OP)$, the frequencies of closed and open digs respectively. However, (1) assumes that the conditional probabilities of hitting with oversight given closed or open digs does not depend on the frequency of closed

and open digs. This may not be realistic; an area where the frequency of 3rd party digging is twice the population average may not experience twice as many incidents of hitting a pipe. One may anticipate that regions with more 3rd party activity, people are more aware of the risks of hitting underground pipelines and take appropriate precautions. This was indeed confirmed by the experts.

It is therefore illustrative to look at this dependence in another way. Think of $F(H \cap O)$ as a function of two continuous variables, FCL = frequency of closed digs per kilometer year, and FOP = frequency of open digs per kilometer year. Write the Taylor expansion about observed frequencies FCL_0 and FOP_0 . Retaining only the linear terms:

$$F(H \cap O/kmyr) = F(FCL, FOP) = F(FCL_0, FOP_0) + p_1(FCL - FCL_0) + p_2(FOP - FOP_0) + \dots \quad (2)$$

If $P(H \cap O|CL)$ and $P(H \cap O|OP)$ do not depend on FCL and FOP then (2) is approximately equivalent to (1). Indeed, put $p_1 = P(H \cap O|CL)$; $p_2 = P(H \cap O|OP)$, and note that $F(FCL_0, FOP_0) = p_1 FCL_0 + p_2 FOP_0$.

The Taylor approach conveniently expresses the dependence on FCL and FOP , in a manner familiar to physical scientists and engineers. Of course it can be extended to include higher order terms .

If we take the ‘zero-order term’ $F(FCL_0, FOP_0)$ equal to the total number of times gas lines are hit while an overseer has marked the lines, divided by the number of kilometer years in The Netherlands, then we can estimate this term from data. FCL_0 and FOP_0 are the overall frequencies of closed and open digs. p_1 and p_2 could be estimated from data if we could estimate $F(FCL, FOP)$ for other values of FCL and FOP , but there are not enough hittings in the data base to support this. As a result these terms must be assessed with expert judgment, yielding uncertainty distributions over p_1 and p_2 . Experts are queried over their subjective uncertainty regarding measurable quantities; thus they may be asked:

Taking account of the overall frequency $F(FCL_0, FOP_0)$ of hitting a pipe line while overseer has marked the lines, what are the 5, 50 and 95 percent

quantiles of your subjective probability distribution for:

The frequency of hitting a pipeline while overseer has marked the lines if frequency of closed digs increases from FCL_0 to FCL , and other factors remaining the same.

In answering this question the expert conditionalizes his uncertainty on everything he knows, in particular the overall frequency $F(FCL_0, FOP_0)$. We configure the elicitation such that the ‘zero order terms’ can be determined from historical data, whenever possible.

How do we use these distributions? Of course if we are only interested in the average situation in the Netherlands, then we needn’t use them at all, since this frequency is estimated from data. However, it is known that the frequency of third party activity (with and without oversight) varies significantly from region to region. If we wish to estimate the frequency of hitting with oversight where $FCL \neq FCL_0$ and $FOP \neq FOP_0$, then we substitute these values into (2), and obtain an uncertainty distribution for $F(H \cap O)$, conditional on the ‘zero-order’ estimate and conditional on the values FCL, FOP . This is pure expert subjective uncertainty. If we wish, we may also include uncertainty due to sampling fluctuations in the zero-order estimate.

2.2 Overall Modeling Approach

The failure probability of gas pipelines is modeled as the sum of a failure probability due to third party actions and a failure probability due to corrosion²:

$$\begin{aligned} P\{\text{failure of gas pipelines/ km}\cdot\text{yr}\} = \\ P\{\text{direct leak due to 3rd parties/ km}\cdot\text{yr}\} + \\ P\{\text{failure due to corrosion/ km}\cdot\text{yr}\} \end{aligned} \quad (3)$$

²The model does not include stress corrosion cracking or hydrogen induced cracking, as these have not manifested themselves in The Netherlands. Damage to pipelines during construction and installation is not modeled. Low probability scenarios like earthquake and flood have not been modeled, and ‘exotic’ scenarios like sabotage, war, malfeasance and the like are neglected.

Both terms on the right hand side will be expressed as functions of other uncertain quantities and parameters. The parameters will be assigned specific values in specific situations, the uncertain quantities are assigned subjective uncertainty distributions on the basis of expert assessments. This results in an uncertainty distribution over possible values of $P\{\text{failure of gas pipelines/km}\cdot\text{yr}\}$, conditional on the values of the known variables.

Failure due to corrosion requires damage to the pipe coating material, and (partial) failure of the cathodic and stray current protection systems. Damage to coating may come either from third parties or from the environment (Lukezich et al 1992). The overall model may be put in the form of a fault tree as shown in Figure 1.

3 Third Party Interference

The model described here enables the calculation of uncertainty distributions over the probability per kilometer year of various damage categories, with or without repair, resulting from third party interference. The underlying probability model is a so-called “marked point process”. For a given one kilometer section of pipe, third party activities (within 10 meters of the pipe) are represented as a Poisson process in time. Each dig-event is associated with a number of “marks”; i.e. random variables which assume values in each dig-event (Figure 2). The random variables and their possible values are shown in Table 1.

For each 1 km pipe section, the following picture emerges:

On the first dig the pipe was not hit, hence the damage was none ($D = n$) and no repair was carried out ($R = rn$). The second dig was an open dig with oversight; small line damage occurred, but was repaired. The third dig was closed without oversight, the line was hit and resulted in a direct leak.

By definition, repair was unable to prevent leak, hence $R = rn$.

The one-kilometer pipe section is described by a number of parameters, which are assumed to be constant along this section:

- t : pipe wall thickness

- gc : depth of ground cover
- $f = (fop, fcl)$: frequency of open and closed digs within 10m of pipe

The values of these parameters will influence the distributions of the random variables in Table 1. Hence, we regard these as random variables, and their influence on other random variables is described by conditionalization. In any one-kilometer section the values for these variables can be retrieved from Gasunie data, and the distributions of other variables can be conditionalized on these values. From a preliminary study (Geervliet 1994) it emerged that pipe diameter and coating type were not of influence on the probability of hitting a pipe.

The damage types indicated in Table 1 are defined more precisely as:

- dl : direct leak (puncture or rupture)
- ldl : line damage large (at least 1mm of pipe material removed, no leak)
- lds : line damage small (less than 1mm of pipe material removed)
- $cd3$: coating damage without line damage due to 3rd parties

Every time a gas pipeline is hit, we assume that one and only one of these damage categories is realized. Hence, $D = n$ if and only if $H = hn$. By definition, if $D = dl$, then repair prior to failure is impossible.

We wish to calculate uncertainty distributions over the probability of unrepaired damage:

$$P(D = x \cap R = rn | t, gc, f); \quad x \in \{cd3, lds, ldl, dl\}. \quad (4)$$

Letting \sum_{HOG} denote summation over the possible values of H , O , G , we have

$$\begin{aligned} P(D = x \cap R = rn | t, gc, f) = \\ \sum_{HOG} P(D = x \cap R = rn | H, O, G, t, gc, f) P(H, O, G | t, gc, f) = \\ \sum_{OG} P(D = x \cap R = rn | H = hy, O, G, t, gc, f) P(H = hy, O, G | t, gc, f) \quad (5) \end{aligned}$$

since third party damage can only occur if the pipe is hit.

For each of the four damage types, there are four conditional probabilities to consider, each conditional on four continuous valued random variables. To keep the model tractable it is necessary to identify plausible simplifying assumptions. These are listed and discussed below. The expressions “X depends only on Y” and “X influences only Y” mean that given Y, X is independent of every other random variable.

1. D depends only on (H, G, t)
2. R depends only on $(H, G, O, D \in \{cd3, lds, ldl\})$
3. gc influences only H
4. gc is independent of f
5. G is independent of f
6. (H, O, G) is independent of t given (gc, f) .

Assumptions 1, 3, 4 and 5 speak more or less for themselves. Assumption 2 says the following: If the pipe is hit and the damage is repairable ($D \neq dl$) then the probability of repair depends only on the type of dig and the presence of oversight; it does not depend on the type of repairable damage inflicted. Assumptions 1 and 2 entail that (D, R) depends only on $(H, G, O, t, D \in \{cd3, lds, ldl\})$.

To appreciate assumption 5, suppose the uncertainty over $f = (fop, fcl)$ is described by an uncertainty distribution, and consider the expression $P(fop, fcl | G = cl)$. Would knowing the type of dig in a given 3rd party event tell us anything about the frequencies of open and closed digs? It might. Suppose that either all digs were open or all digs were closed, each possibility having probability 1/2 initially. Now we learn that one dig was closed; conditional on this knowledge, only closed digs are possible. Barring extreme correlations between the uncertainty over values for fop and fcl , knowing $G = cl$ can tell us very little about the values of fcl and fop . Assumption 5 says that it tells us nothing at all.

To illustrate how these assumptions simplify the calculations, we consider the event ($D = cd3 \cap R = rn$); which we abbreviate as ($cd3 \cap rn$). We can show³:

$$P(cd3 \cap rn|t, gc, f) =$$

$$\sum_{OGP} (rn|hy, O, G)P(cd3|hy, G, t)P(O, hy|G, f)P(G)P(gc|hy)/P(gc) \quad (8)$$

Similar expressions hold for damage types lds and ldl . For dl , the term $P(rn|hy, O, G)$ equals one as repair is not possible in this case. The terms

$$P(cd3|hy, G, t), P(G|hy), P(G), P(gc|hy) \text{ and } p(gc)$$

can be estimated from data; the other terms are assessed (with uncertainty) using expert judgment. The uncertainty in the data estimates derives from sampling fluctuations and can be added later, although this will be small relative to uncertainty from expert judgment. The term

$$P(gc|hy)/P(gc) = P(H = hy|gc)/P(H = hy)$$

is called the ‘‘depth factor’’; it is estimated by dividing the proportion of hits at depth gc by the total proportion of pipe at depth gc . $P(G = cl|hy, cd3)$

³Using elementary probability manipulations and assumptions 1 and 2:

$$\begin{aligned} P(cd3 \cap rn|hy, O, G, t, gc, f)P(hy, O, G|t, gc, f) &= \\ P(cd3 \cap rn|hy, O, G, t)P(hy, O, G|t, gc, f) &= \\ P(cd3|rn, hy, O, G, t)P(rn|hy, O, G, t)P(hy, O, G|t, gc, f) &= \\ P(cd3|hy, G, t)P(rn|hy, O, G)P(hy, O, G|t, gc, f) & \end{aligned} \quad (6)$$

Reasoning similarly with assumptions 3, 4, 5 and 6, and using Bayes’ theorem:

$$\begin{aligned} P(hy, O, G|t, gc, f) &= P(hy, O, G|gc, f) = \\ P(gc, O, G|hy, f)P(hy|f)/P(gc|f) &= \\ P(gc|O, G, hy, f)P(O, G|hy, f)P(hy|f)/P(gc|f) &= \\ P(gc|hy)P(O, G|hy, f)P(hy|f)/P(gc) &= \\ P(gc|hy)P(O, hy|G, f)P(G|f)/P(gc) &= \\ P(gc|hy)P(O, hy|G, f)P(G)/P(gc) & \end{aligned} \quad (7)$$

is estimated as the percentage of coating damages from third parties caused by closed digs; $P(G = cl|hy)$ is the percentage of hits caused by closed digs, and $P(G)$ is the probability per kilometer year of a closed dig. This probability is estimated as the frequency of closed digs per kilometer year, if this frequency is much less than 1 (which it is).

The term $P(rn|hy, O, G)$ is assessed by experts directly when the terms $P(O, hy|G, f)$ are assessed using the Taylor approach described in section 3. There are no ruptures directly caused by third party activities in the Dutch data base. To assess the probability (with uncertainty) of ruptures due to third parties, experts assess, for two different wall thickness, the percentages of direct leaks which will be ruptures. Let RUP71 and RUP54 denote random variables whose distributions reflect the uncertainty in these percentages for thickness 7.1 and 5.4 mm respectively. We assume that RUP71 and RUP54 are comonotonic⁴. Putting

$$RUP54 = RUP71 + x(7.1 - 5.4)$$

we can solve for the the linear factor x , and for some other thickness t :

$$RUPt = RUP71 + x(7.1 - t)$$

gives an assessment of the uncertainty in the probability of rupture, given direct leak, for thickness t . This produces reasonable results for t near 7.1. For $t > 10\text{mm}$ it is generally agreed that rupture from third parties is not possible (Hopkins et al 1992).

4 Damage Due to Environment

In dealing with damage to coating due to environmental factors per kilometer year, we revert to the frequency notation, as this frequency can be larger than one. For both bitumen (*bit*) and polyethylene (*pe*) coatings, the probability of environmental damage depends on the pipe diameter (d), on the soil type (st) and on the percentage of the pipe exposed to fluctuating

⁴That is, their rank correlations are equal to one

water table (wt_f). Bitumen coating is also sensitive to the proportion of the one-kilometer length exposed to tree roots (rt) and chemical contamination (ch). The effects of these factors are captured with a first order Taylor expansion whose linear terms p_5, \dots, p_{10} are assessed by experts.

$$\begin{aligned} F(bit) &= Fo(bit) + p_5 \cdot (d - d_0) + p_6 \cdot wt_f + p_7 \cdot rt + p_8 ch + st \cdot bit \quad (9) \\ F(pe) &= Fo(pe) + p_9 \cdot (d - d_0) + p_{10} \cdot wt_f + st \cdot pe \quad (10) \end{aligned}$$

To determine the probability of at least one coating damage per kilometer, these frequencies are divided by 100 to determine the frequency per 10 meter section. As these frequencies will be much less than 1, and assuming that damage to different 10 meter sections are independent, we have:

$$P\{\text{at least 1 coating damage per km}\} = 1 - (1 - F/100) \quad (11)$$

On substituting (9) and (10) into (11), we obtain the probabilities per km year of bitumen and polyethylene coating damage per km year, due to environmental factors, notated $P(CDE_{bit})$ and $P(CDE_{pe})$.

5 Failure Due to Corrosion

5.1 Modelling Corrosion Induced Failures

The modeling of failure due to corrosion is more complicated than that of failure due to third parties. The probability of failure due to corrosion depends on many factors as listed in Table 2.

The model described here uses only pit corrosion. Given these factors, the corrosion rate is assumed constant in time (Camitz et al 1989).

For a pipeline to fail due to corrosion, two lines of defense must be breached. First the coating must be damaged, and second, depending on location, the cathodic or stray current protection system must not function as intended. Coating damage may be caused either by third party actions or by environmental factors. These protection systems have been in place since 1970.

We first elaborate the model for pipelines installed after 1970.

Assuming that the coating has been breached, pit corrosion will reduce the pipe wall thickness until a critical value is reached, at which point the pipe fails. This critical wall thickness, that is the thickness at which failure occurs, is expressed as a fraction x of the original wall thickness minus the pipe material removed during the damage event. x depends on the pressure of the gas in the pipe line, and on the geometry of the pipe damage, and this relationship has been established by experiment. In this model, x is introduced as a parameter whose value depends only on the damage type, thus we distinguish x_C , x_S and x_L for (only) coating damage, small and large pipe damage respectively. Coating damage is either caused by 3rd parties (*cd3*) or by the environment (*cde*).

A length of pipe can be inspected for corrosion, and if corrosion is found, the pipe is uncovered and repaired. Hence, after such inspection the pipe is as good as new. The effective birthday (*eb*) of a pipe section is the calendar year of the last inspection.

Given a corrosion rate (CR) and a damage type, we define the effective life of a pipe section as the time required for the corrosion to reduce wall thickness to the critical wall thickness. Letting t denote the original pipe wall thickness, i.e. $C, S, L, t_C = 0, t_S = 0.5mm, t_L = 2mm$:

$$EL(CR, i) = x_i(t - t_i)/CR. \quad (12)$$

In this equation, CR is uncertain and x_i, t_i are parameters with uncertain indices. $EL(CR, i)$ is the time a pipe survives given corrosion rate CR after sustaining damage type $i, i \in \{C, S, L\}$.

Suppose we are interested in the event ‘first failure of a one-kilometer length of gas pipeline occurs in calendar year y ’. For each given value of CR , there are three years, $y_C(CR)$, $y_S(CR)$ and $y_L(CR)$ such that damage type i in year y_i , somewhere on this one-kilometer length of pipe, causes failure in year y . y_i is called the critical year for damage type i . The situation is pictured in Figure 3.

Referring to Figure 3, we see that failure due to damage type C is impossible; the pipe isn’t old enough in year y . If small damage (S) occurs in year y_S ,

and not before, and if large damage has not occurred before y_L , then the pipe fails in y due to small damage in y_S . The probability of this is

$$(1 - P_S)^{Y_S - eb} P_S (1 - P_L)^{Y_L - eb}. \quad (13)$$

where we write $P_S = P(lds)$, $P_L = P(ldl)$, $P_C = P(cd3) + P(cde) - P(cd3 \cap cde)$. However, if y is 'next year' then we already know that the pipeline has not already failed due to corrosion from small or large pipeline damage. Hence, we should conditionalize on the event 'no small damage before y_S and no large damage before y_L '. In this case, the probability of failure in year y due to small damage is simply P_S , and the probability of failure due to corrosion is, neglecting higher order terms, $P_S + P_L$ (all of these probabilities are conditional on CR).

If y is in the future, and we conditionalize on our knowledge that no failure has occurred up to now, with $T = y - \text{now}$, $T < y_S - eb$, then the probability of failure due to corrosion between in year y is (again, conditional on CR and neglecting small terms):

$$(1 - P_S)^T (1 - P_L)^T (P_S + P_L). \quad (14)$$

In general, let

$$q_i(eb, y, \text{now}, CR) = \text{MIN}\{(y_i(CR) - eb), (y - 1 - \text{now})\} \quad (15)$$

denote the number of years between y and now in which a failure due to damage type i could have caused failure between now and year $y - 1$, conditional on CR , and let

$$1_i = \begin{cases} 1, & \text{if } y_i(CR) > eb \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

then (sum and product are over $i \in \{C, S, L\}$);

$$P_{f|cr}(CR, eb, y, \text{now}, P_C, P_S, P_L, t, x_C, x_S, x_L) = \prod (1 - p_i)^{q_i} \sum 1_i p_i \quad (17)$$

is approximately the probability of failure in year y due to corrosion, given CR . $P_{f|cr}$ is an uncertain quantity since the arguments written in capital letters represent uncertain quantities.

5.2 Pit Corrosion Rate

The free rate of pit corrosion CR_f [mm/yr] is modeled to depend on the soil type (clay, sand, peat), the soil resistance (r), the acidity (pH) and the proportion of pipeline under the water table, above the water table and fluctuating under and above the water table (wt_u, wt_a, wt_f). CR_f is the rate of corrosion which would obtain if the cathodic protection were not present. Using a zero-order corrosion rate with arguments $r_0, pH_0, wt_{u0} = wt_{f0} = 0, wt_{a0} = 1$; we apply the linear approximation (supported by experiment):

$$CR_f = CR_{fo} + p_{11}(r - r_o) + p_{12}(pH - pH_o) + p_{13}wt_f + p_{14}wt_u. \quad (18)$$

The linear terms p_{11}, \dots, p_{14} are assessed with expert judgment. All of the terms in (18) depend on soil type.

We distinguish three states of the cathodic protection system:

- CP_f : wholly non-functional, $CR = CR_f$
- CP_p : partially functional (pipe-soil potential outside prescribed range),
 $CR = CR_p$
- CP_{ok} : wholly functional as prescribed; $CR \sim 0$.

Before the cathodic protection system was installed in 1970, only state CP_f was available. CR_p is determined via expert judgment as a fraction of CR_f . $P(CP_i)$ is the fraction of one-kilometer pipe length for which cathodic protection is in state i , $i \in \{CP_f, CP_p, CP_{ok}\}$. Since the factors affecting the cathodic protection do not change from year to year, we assume that the states CP_f and CP_p affect the same portions of pipe each year.

Stray currents can induce corrosion against which cathodic protection is ineffective. In 1970 a protection system of bonds was installed to drain off strong stray currents in locations where these are known to occur. Each bond is inspected once a month hence if a bond has failed the stray current corrosion rate CR_{st} has been operative on the average for one half month. In the neighborhood of a bond, the corrosion rate before 1970 due to stray currents is CR_{st} , and after 1970 it is assumed to be $CR_{st}/24$. If bs is the

proportion of a one-kilometer length of pipe in the neighborhood of a bond sites and $P(SP)$ is the probability that the stray current protection system fails at one site, then $bs \cdot P(SP)$ is the probability that CR_{st} (before 1970) or $CR_{st}/24$ (after 1970) obtains, given that damage has occurred somewhere in the pipe section.

Unconditionalizing equation (18) on CR , we obtain the probability of failure per kilometer year due to corrosion for pipe installed after 1970:

$$P_{cor>70} =$$

$$P_{f|cr}(CR_f)P(CP_f) + P_{f|cr}(CR_p)P(CP_p) + P_{f|cr}(CR_{st}/24)bsP(SP) \quad (19)$$

This is an uncertain quantity whose distribution is the uncertainty distribution for the failure frequency for a one-kilometer length of gas pipeline with specified pipe and environment parameter values.

For pipelines whose effective birthday is before 1970,

$$x_i(t - t_i) - (y - 1970) \cdot CR \quad (20)$$

is the thickness of pipe wall, under damage type i , exposed to corrosion at the rate obtaining before protection systems were installed. Let

$$1_{i,CR} = \begin{cases} 1, & \text{if } y - x_i(t - t_i)/CR > 1970 \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

If $1_{i,CR} = 1$, then $y_i > 1970$; if $y_i < 1970$ then we must account for the absence of protection systems. We compute the effective life as follows:

$$\begin{aligned} EL(CR_f, i) &= x_i(t - t_i)/CR_f; \\ EL(CR_p, i) &= 1_{i,CR_p}x_i(t - t_i)/CR_p \\ &\quad + (1 - 1_{i,CR_p})(x_i(t - t_i) - (y - 1970)CR_p)/CR_f; \\ EL(CR_{st}, i) &= 1_{i,CR_{st}}x_i(t - t_i)24/CR_{st} \\ &\quad + (1 - 1_{i,CR_{st}})(x_i(t - t_i) - (y - 1970)CR_{st}/24)/CR_{st} \end{aligned} \quad (22)$$

$P_{cor<70}$ is obtained by using (22) instead of (12) in (17).

6 Validation

This model was originally developed in 1996 (see Cooke and Jager, 1998). In the last few years the Dutch gas company has launched a program of “intelligent pig runs”. An intelligent pig is a device that can be sent through a large diameter pipe to measure corrosion defects. These pig runs are quite accurate but also quite expensive. The gas company is interested in using these runs to calibrate the failure model, so that the model can be used to support the selection of pipes to be pigged in the future.

At present data from two pig runs are available. For each run, the data consists of a list of defects, their position on the pipe, and their depth. Run A covered 66 km of a gas pipeline with bitumen coating laid in sand in 1966, with average diameter of 12.45 inches at an average depth of 1.76 m. There were 65 incidents in which the removal of pipe material was at least 10 percent of the wall thickness.

Run B covered 84 km of a gas pipeline with bitumen coating laid in sand in 1965, with average diameter of 11.45 inches at an average depth of 1.87 m. There were 92 incidents in which the removal of pipe material was at least 10 percent of the wall thickness. A part of the data is included in this document (Figure 4).

The Laplace test was applied to each data set separately to test the hypothesis that the spatial inter-arrival times came from an exponential distribution, against the hypothesis that the data come from a non-homogeneous Poisson distribution. The Pipeline A data indicates a statistically significant spatial trend since 30% of observed corrosion appear in first 3km of the pipe (the pipe is 65 km long). This spatial clustering could not be explained and is not used in the further analysis. There was no significant ‘spatial trend’ in the Pipeline B data.

A non-parametric Kolmogorov-Smirnov test was used to test the hypothesis that the spatial inter-arrival intervals for events removing at least 10 percent of pipe wall material came from the same distribution. The hypothesis was not rejected at the 5 percent level. Hence, no significant difference was found

between the two data sets.

The model predicts the frequency of leak due to corrosion per kilometer year. As there have been no leaks, the model must be adapted to predict the frequency of corrosion events removing specified percentages of pipe wall material. This is done by manipulating the critical fraction (xc) of pipe wall material which must be removed in order to cause a leak due to coating damage. By setting $xc = 10$ percent the 'failure frequency' output by the model corresponds to the frequency of corrosion events removing at least 10 percent of the pipe wall material. By setting xc successively equal to 10, 15, 20 . . . 40 percent we obtain 7 uncertainty distributions for the frequency per kilometer year of removing at least xc percent of the pipe wall material. For each value of xc , we retrieve the number of events removing at least xc percent of pipe wall material from the data. Dividing this number by the number of kilometer years, we obtain the empirical frequency per kilometer year of removing at least xc percent of pipe wall material. We then compare these empirical frequencies with the appropriate uncertainty distribution for these frequencies from the model. The results are shown in Figures 5 and 6. We see that the model places the observed frequencies well within the central mass of the respective uncertainty distributions⁵.

7 Results for Ranking

7.1 Case-Wise Comparisons

Two types of results can be obtained with the model. First, we can perform case-wise comparisons. By specifying parameter values for two or more types of kilometer-year sections of pipelines, the uncertainty distributions for the frequency of failure can be compared. In Figure 7 compares three cases, namely

- bitumen coated pipe laid in 1975 in sand

⁵The exact time at which cathodic protection was installed could not be retrieved at this writing, but is estimated to be 1970

- bitumen coated pipe laid in 1975 in clay
- polyethylene coated pipe laid in 1975 in sand

Percentiles of the subjective uncertainty distribution are shown horizontally; the logarithm of the failure frequency per kilometer year is plotted vertically (the absolute values are proprietary). Other parameters are the same in all cases, and those describing the frequency of 3rd party intervention are chosen in accord with the generic values retrievable from the Dutch data. Each graph plots the frequency per kilometer year of failure against the percentiles of the uncertainty distribution for the failure frequency. Each graph shows three curves, a curve corresponding to failure due to corrosion (*corlk*), a curve corresponding to failure due to 3rd party interference (*3leak*), and a curve corresponding to the sum of these two (*leak*).

Because of the choice of 3rd party interference frequencies, the curves for failure due to 3rd party interference are constant - this means that there is no uncertainty regarding this failure frequency. The curves for failure due to corrosion are not constant. In the first graph (bitumen in sand), we see that the 66th percentile of the uncertainty distribution for failure due to corrosion corresponds to the same failure frequency as 3rd party intervention. In other words, there is a 0.66 probability that the frequency of failure due to corrosion will be lower than that due to 3rd party intervention.

In the second graph (bitumen in clay) we find a probability of 0.85 that the frequency of failure due to corrosion will be lower than that due to 3rd party intervention. In the third graph (polyethylene in sand), we find a probability of 0.77 that the frequency of failure due to corrosion will be lower than that due to 3rd party intervention. In the second graph, note that the failure frequency curve for corrosion drops off more rapidly than in the third graph. This means that very low values are more likely for bitumen coated pipe in clay than for polyethylene coated pipe in sand.

Comparisons of this nature can only be made on the basis of fully specified cases. We cannot conclude, for example, that “bitumen in clay is about the same as polyethylene in sand”. The comparisons in Figure 7 depend

on the values of all other environmental and pipe variables. Thus, changing the amount of root growth, the soil resistivity the age and thickness of the pipe, or any of the other parameters, might produce very different pictures. Finally, we note that the failure frequency due to corrosion is highly uncertain. Nevertheless, clear comparisons may be made by taking this uncertainty into account.

7.2 Importance in Specific Case

As mentioned in the introduction, we use Monte Carlo simulation to compute the uncertainty distribution of the failure frequency for given pipe and environmental characteristics. When we focus on a particular kilometer of pipe, i.e. a particular set of values for all the parameters in the model; we may ask “which factors are important for the failure frequency in this specific case?” Since this failure frequency is uncertain, we are really asking “which factors are important for the uncertainty in failure frequency in this case?”

To gain insight into this type of question a new graphic exploratory tool has been developed, termed “cobweb plots”⁶. These plots enable the user to gain insight into complex relations between interdependent uncertain quantities.

We illustrate by considering the uncertainty in failure due to corrosion in a bitumen coated pipe laying in sand for five years without cathodic protection.

The variable *corlk* or “leak due to corrosion” is potentially influenced by the following variables:

- crf: free corrosion rate
- crp: corrosion rate under partial functioning of cathodic protection
- crse: corrosion rate from stray currents
- ps: frequency of small unrepaired pipeline damage

⁶Wegman (1990) introduced a similar technique, though without conditionalization.

- *pl*: frequency of large unrepaired pipeline damage
- *pc3*: frequency of coating damage from 3rd parties
- *pcen*: frequency of coating damage from environment

The uncertainty distribution for *corlk* is built up by considering a large number of “scenario’s”, where each scenario is made by sampling values from all input variables. In each scenario, unique values are assigned to all the above variables. We are interested in how the high and low values of *corlk* co-vary with high and low values of the above variables.

Cobweb plots allow the user to explore this co-variation. Suppose we plot all the values of the above variables on parallel vertical lines, with high values at the top and low values at the bottom. Each individual scenario assigns exactly one value to each variable; if we connect these values we get a jagged line intersecting each variable-line in one point. Suppose we plot jagged lines for each of 200 scenario’s; the result will suggest a cobweb. It may be difficult to follow the individual lines; it is therefore convenient to “filter” or “conditionalize” on sets of lines. For example, we might conditionalize on all lines passing through high values of *corlk* and see where these lines intersect the other variables.

The first cobweb plot (Figure 8) shows lines for 500 scenario’s. The second cobweb plot (Figure 9) conditionalizes on high values of *corlk*: we see that these are associated with high values of *crf* and with high values of *pcen*. *crp* and *crse* are not affected by this conditionalization; by assumption, there is no cathodic protection in this case. The third cobweb plot (Figure 10) conditionalizes on low values of *corlk*. We see that these are strongly associated with low values of *crf* and with *crse* but not associated with other variables. We may conclude that damage from the environment is important for high values of failure frequency due to corrosion, but not for low values. This sort of behavior occurs quite often; the variables that are associated with high values of some target variable are not the same as the variables associated with low values of the target variable.

8 Conclusions

We collect a number of conclusions.

- A ranking tool has been developed which uses failure data and structured judgment.
- The tool characterizes pipe sections according to some 20 pipe and environmental characteristics
- The tool predicts failure frequencies per kilometer year, and gives uncertainty bounds
- These predictions allow distinctions to be made between pipe sections with different characteristics, and these distinctions are not swamped by uncertainties, despite the fact that the uncertainties are large
- For most pipes, the risk due to corrosion is significantly less than the risk due to 3rd party interference
- The depth of ground cover is of significant influence for the frequency and severity of 3rd party damage.

Exercises

1. The expected frequency of failure in a pipe is 2 per kilometer year. Derive the probability of failure per kilometer year using the assumptions in section 2.
2. Equations (1) and (2) give two expressions for the frequency of hitting a pipe with oversight. Whereas (1) simply gives a number, (2) gives this number as a function of the variables FCL and FOP . Write out the proof of equivalence sketched in the text and discuss its assumptions. In particular, suppose $P(H \cap O | CL)$ is a linear function of FCL , and that $P(H \cap O | OP)$ is a linear function of FOP . Rewrite (2) using these linear functions.
3. Write out in detail the derivation of equation 6, justifying all the steps.

4. Write out in detail the derivation of equation 7, justifying all the steps.
5. Suppose the effective life due to coating damage is equal to the effective life due to small line damage. Derive the relation between the critical thickness for coating damage, the critical thickness for small line damage and the pipe diameter.
6. Consider a cobweb plot with two independent uniform variables U, V . Draw a vertical line half way between the lines corresponding to the two variables and consider the density of crossings on this half way line. Show that this density is triangular (hint, show that lines crossing on the half-way line correspond to realizations $(u_1, v_1), (u_2, v_2)$ satisfying $u_1 + v_1 = u_2 + v_2$. The density is then proportional to the length of the line $U + V = cns$ in the unit square.)

Tables and Figures

Variable	Meaning	Values	Interpretation
H	Pipe hit?	hy, hn	hityes, hitno
O	Overseer notified?	oy, on	overtightyes, oversightno
R	Repair carried out?	ry, rn	repairyes, repairno
D	Damage?	$n, cd3, lds, ldl, dl$	none, coating damage, small line damage, large line damage, direct leak
G	Dig type?	op, cl	open dig, closed dig

Table 1: Marks for 3rd party digs

chemical contamination of soil	pipe diameter	oversight
soil type (clay, sand, peat)	pipe inspection	repair
soil resistance	stray currents	acidity
pipe thickness	tree roots	pipe age
water table	third party actions	

Table 2: Factors influencing failure due to corrosion

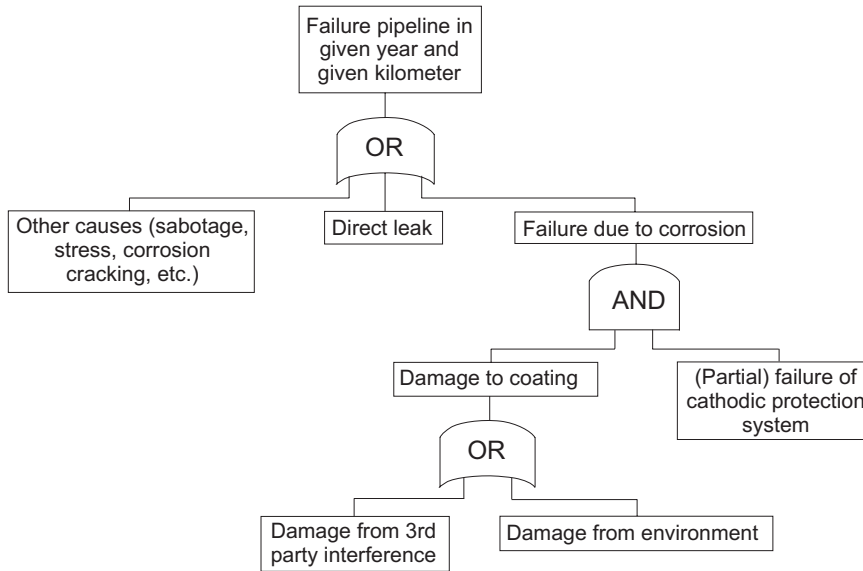


Figure 1: Fault tree for gas pipeline failure

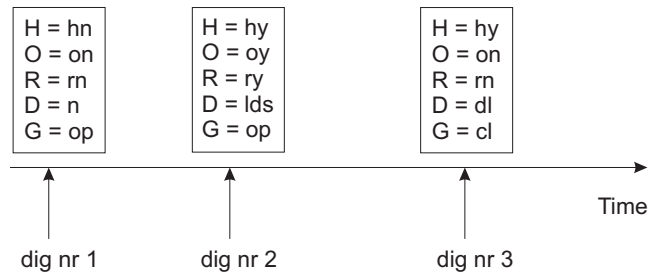


Figure 2: Digs as marked point process

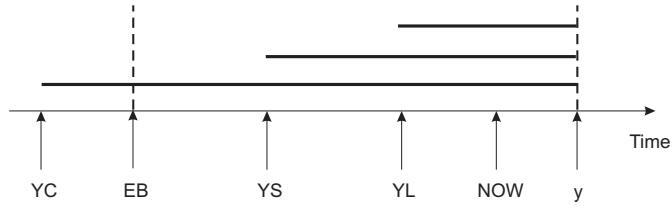


Figure 3: Effective lives and critical years for three damage types for fixed corrosion rate

EVENT_NO	EVENT_NAME	CATEGORY	distance [m]	LIST_CLOCK	position [hour]	%ML	length [mm]	width [mm]	wall thickness [mm]
D - 1	Defect	External General	97.579	3:30	3:30	23	41	142	12.86
D - 2	Defect	External General	97.595	8:10	8:10	31	41	91	12.86
D - 3	Defect	External Circ Groove	97.666	7:10	7:10	18	20	145	12.86
D - 4	Defect	External Circ Groove	114.636	6:50	6:50	14	25	79	12.86
D - 57	Defect	External Pit	799.246	6:30	6:30	16	25	41	12.86

Figure 4: The Gasunie data

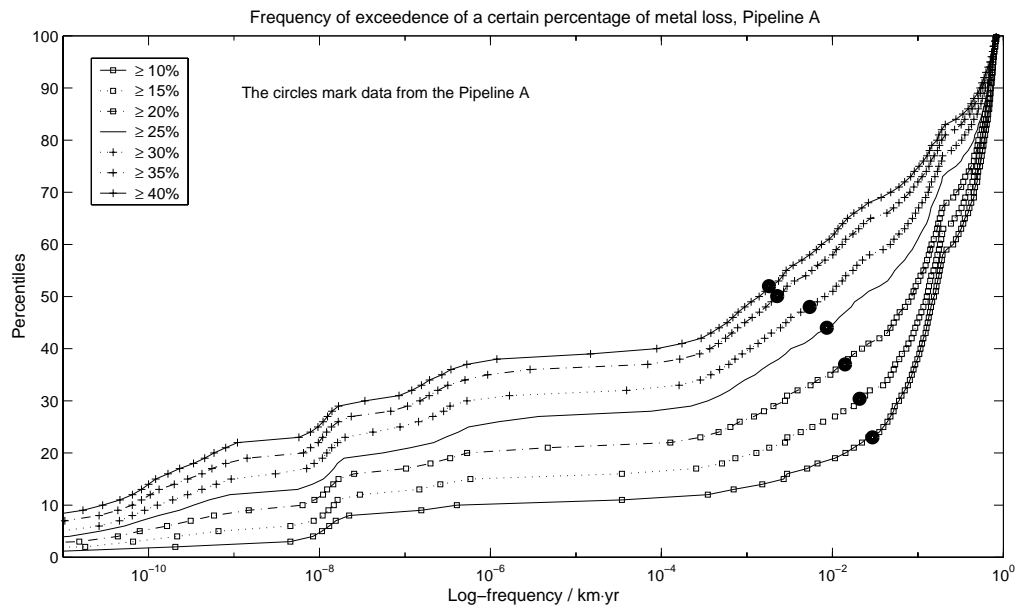


Figure 5: Uncertainty distributions and observed exceedence frequencies, pipe A

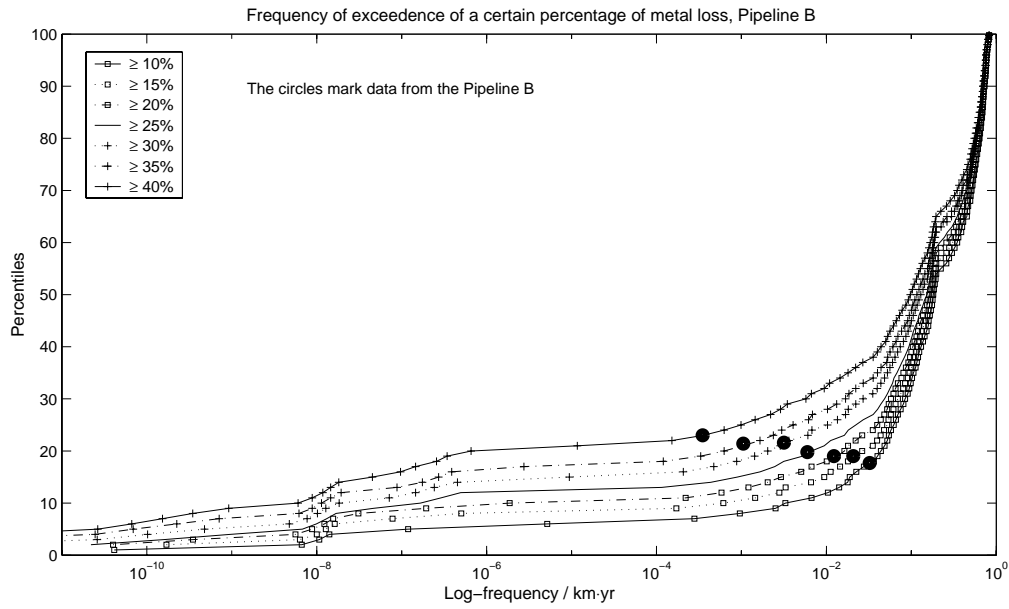


Figure 6: Uncertainty distributions and observed exceedence frequencies, pipe B

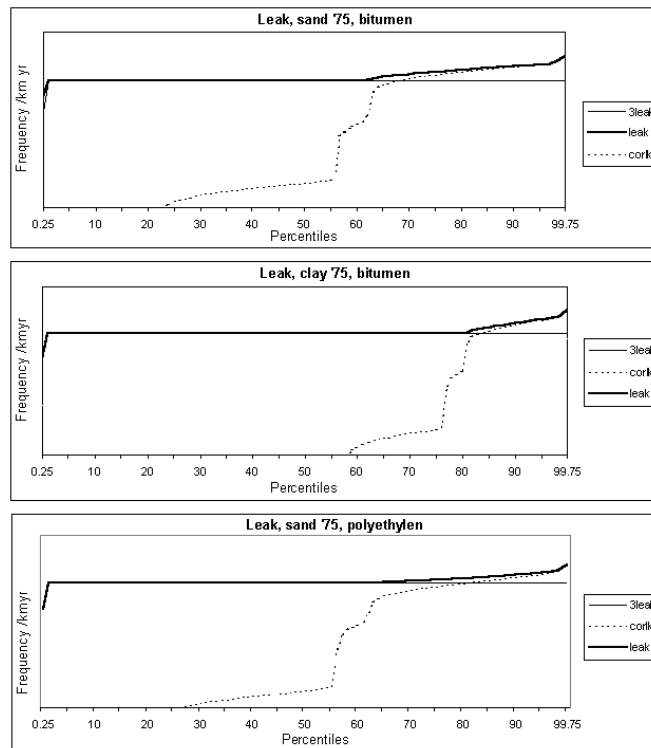


Figure 7: Uncertainty distributions for three cases

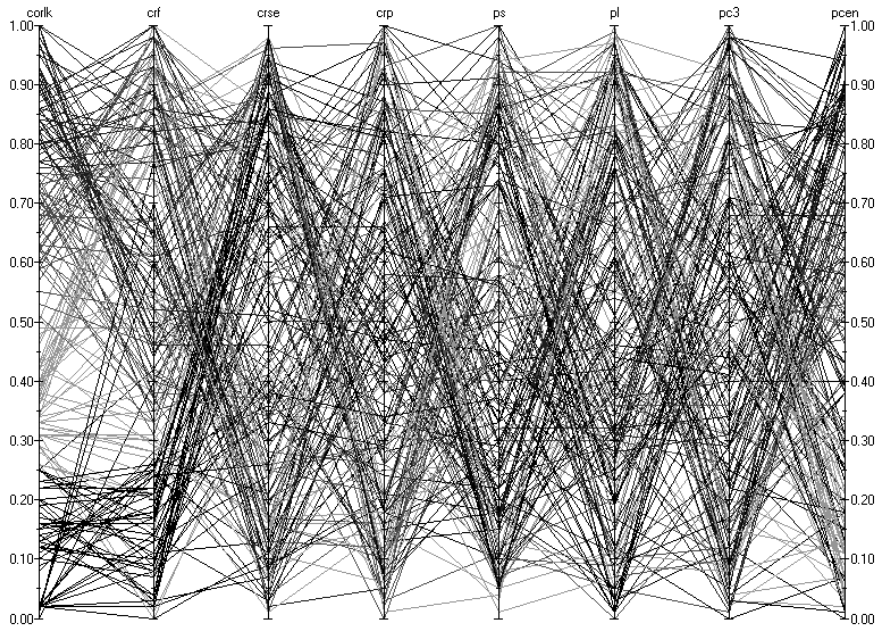


Figure 8: Unconditional cobweb plot

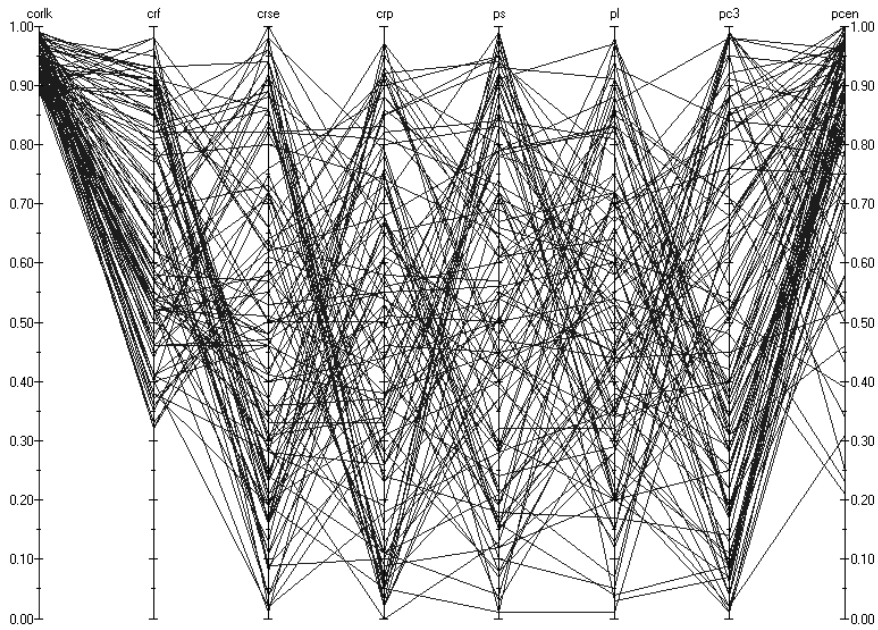


Figure 9: Cobweb plot conditional on high values of *corlk*

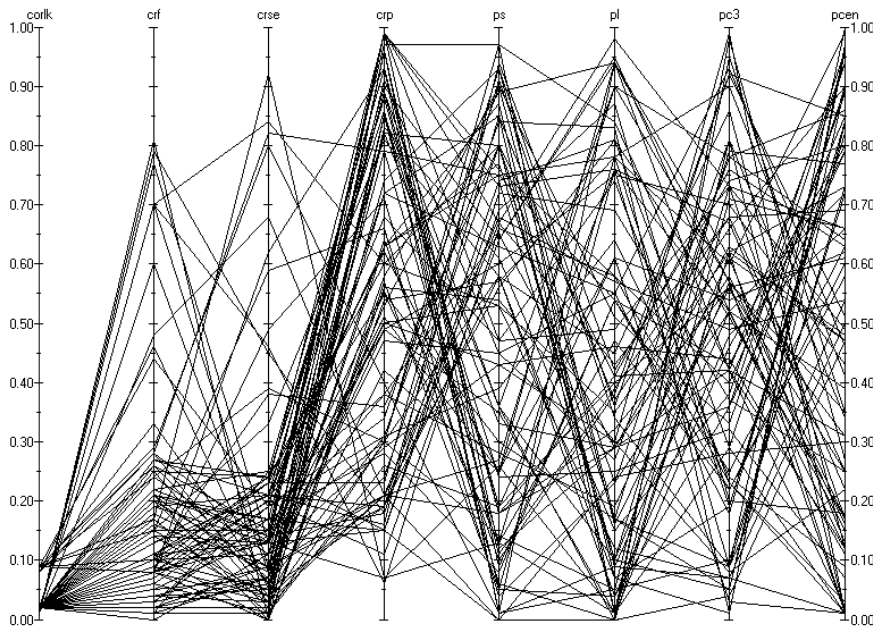


Figure 10: Cobweb plot conditional on low values of *corlk*

References

- Sept 28 (1990). *Methods for Prioritizing Pipeline Maintenance and Rehabilitation* by American Gas Association, Battelle Corporation.
- Basalo, C. (1992). *Water and gas mains corrosion, degradation and protection*, Ellis Horwood, New York.
- Bushman, J.B., and Mehalick, T.E. (1989). *Statistical analysis of soil characteristics to predict mean time to corrosion failure of underground metallic structures* in Chaker V., and Palmer, J.D., *Effects of Soil Characteristics on Corrosion*, Report No. STP 1013, American Society for Testing and Materials, Philadelphia.
- Camitz, G., and Vinka, T.G. (1989). *Corrosion of steel and metal-coated steel in Swedish soils - effects of soil parameters* in Chaker V., and Palmer, J.D., *Effects of Soil Characteristics on Corrosion*, Report No. STP 1013, American Society for Testing and Materials, Philadelphia.
- Cooke, R. (1991). *Experts in Uncertainty*, Oxford University Press, Oxford.
- Cooke, R. M. and Jager, E. (1998). *A probabilistic model for the failure*

- frequency of underground gas pipelines* in *Risk Analysis* vol. 18, No. 4, pp 511-527, Society for Risk Analysis.
- Cooke, R.M. (1995). *UNICORN, Methods and Code for Uncertainty Analysis*, AEA Technology, The SRD Association, Thomson House, Risely, Warrington, Cheshire, WA3 6AT, United Kingdom.
- Chaker V. and Palmer J.D. (eds) (1989). *Effects of Soil Characteristics on Corrosion*, Report No. STP 1013, American Society for Testing and Materials, Philadelphia.
- Geervliet, S. (1994). *Modelling van de Faalkans van Ondergronds Transportleidingen* in *Report for two year post graduate program, department of Mathematics and Informatics*, performed under contract with the Netherlands Gasunie, Delft University of Technology, Delft.
- Hopkins, P., Corder, I., and Corbin, P. (1992). *The resistance of gas transmission pipelines to mechanical damage*, European Pipeline Research Group.
- Kiefner, J.F., Vieth, P.H. and Feder, P.I. (1990). *Methods for Prioritizing Pipeline Maintenance and Rehabilitation*, American Gas Association, Washington.
- Lukezich, S.J., Hancock, J.R. and Yen B.C. (1992). *State-of-the-art for the use of anti-corrosion coatings on buried pipelines in the natural gas industry*, topical report 06-3699, Gas Research Institute.
- Mayer, G.R., van Dyke, H.J. and Myrick, C.E. (Sept. 21 1987). *Risk analysis determines priorities among pipeline-replacement projects*, Oil and Gas Journal.
- Palmer, J.D. (1989). *Environmental characteristics controlling the soil corrosion of ferrous piping* in Chaker V., and Palmer, J.D., *Effects of Soil Characteristics on Corrosion*, Report No. STP 1013, American Society for Testing and Materials, Philadelphia.
- Wegman, E.J. (1990). *Hyperdimensional data analysis using parallel coordinates*, Journal of the American Statistical Analysis, vol. 90, No. 411.