Troubles with weightings

By W.Kent Muhlbauer, WKM Consultancy, LLC, Houston, TX, USA

There is currently a great disparity in approaches and level of rigour applied to risk assessment by pipeline operators largely due to the absence of complete standards or guidelines covering this complex topic. The disparity leads to inconsistent and problematic risk management, as was discussed in a previous column.

Most operators desire sound and useful risk assessment to support their decision-making. Weaknesses in an operator's risk-assessment practice are almost entirely due to insufficient guidance. This column strives to improve this situation by challenging past practice as well as discussing proper methods for pipeline risk assessment.

Focusing this time on our past missteps, the use of 'weightings' should be a target of critical review in any risk-assessment practice. Weightings have been used in some older risk assessments to give more importance to certain factors. They were usually based on a factor's perceived importance in the majority of historical pipeline-failure scenarios. For instance, the potential for AC-induced corrosion is usually very low for many kilometres of pipeline, so assigning a low numerical weighting appeared appropriate for that phenomenon. This was intended to show that AC-induced corrosion is a rare threat.

Used in this way, weightings steer risk-assessment results toward predetermined outcomes. Implicit in this use is the assumption of a predictable distribution of future incidents and, most often, an accompanying assumption that the future distribution will exactly track the past distribution. This practice introduces a bias that will almost always lead to very wrong conclusions for some pipeline segments.

The first problem with the use of weightings is finding a representative basis for the weightings. Weightings were usually based on historical incident statistics – '20 per cent of pipeline failures from external corrosion'; '30 per cent from third-party damage'; etc. These statistics were usually derived from experience with many kilometres of pipeline over many years of operation. However, different sets of pipeline kilometre-years show different experience. Which past experience best represents the pipeline being assessed? What about changes in maintenance, inspection, and operation over time? Shouldn't those influence which data sets are most representative to future expectations?

It is difficult, if not impossible, to know which set of historical population behaviour best represents the future behaviour of the segments undergoing the current risk assessment. If weightings are based on, for example, average country-wide history, the non-average behaviour of many kilometres of pipeline is discounted. Using national statistics means including many pipelines with vastly different characteristics from the system being assessed.

If the weightings are based on a specific operator's experience, then (hopefully) only a very limited amount of data is available. Statistics using small data sets are always problematic. Furthermore, a specific pipeline's accident experience will probably change with the operator's changing risk-management focus. When an operator experiences many corrosion failures, he will presumably take actions to specifically reduce the potential for corrosion occurring and, over time, a different mechanism should then become the chief failure cause. So, the weightings would need to change periodically and would always lag behind actual experience, therefore having no predictive contribution to risk management.

W. Kent Muhlbauer.

The bigger issue with the use of weightings is the underlying assumption that the past behaviour of a large population will reliably predict the future of an individual. Even if an assumed distribution is valid for the long term population behaviour, there will be many locations along a pipeline where the pre-set distribution is not representative of the particular mechanisms at work there. In fact, the weightings can fully obscure the true threat. The weighted modelling of risk may fail to highlight the most important threats when certain numerical values are kept artificially low, making them virtually unnoticeable.

Use of weightings as a significant source of inappropriate bias in risk assessment is readily demonstrated. One can easily envisage numerous 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.

Consider threats such as landslides, erosion, or subsidence, classed as failure mechanisms called geohazards. An assumed distribution of all failure mechanisms will almost certainly assign a very low weighting to this class since most pipelines are not significantly threatened by the phenomenon and, hence, incidents are rare. For example, to match a historical record that shows 30 per cent of pipeline incidents are caused by corrosion and 2 per cent by geohazards, weightings might have been used to make corrosion point totals 15 times higher than geohazard point totals (assuming more points means higher risk) in an older scoring methodology.

One can easily envisage numerous 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.

But a geohazard phenomenon is a much localised and very significant threat for some pipelines, and may dominate all other threats in some segments. Assigning a 2 per cent weighting masks the reality that, perhaps, 90 per cent of the failure probability on this segment is due to geohazards. So, while the assumed distribution may be valid on average, there will be locations along some pipelines where the pre-set distribution is very wrong. It would not at all be representative of the dominant failure mechanism at work there. The weightings will often completely mask the real threat at such locations.

This is a classic difficulty in moving between the behaviour of statistical populations and individual behaviour. The former is often a reliable predictor – hence the success of insurance actuarial analyses –but the latter is not.

In addition to masking location-specific failure potential, use of weightings can force only the higher-weighted threats to be perceived 'drivers' of risk, at all points along all pipelines. This is rarely realistic. Risk management can become driven solely by the pre-set weightings rather than actual data and conditions along the pipelines. Forcing risk-assessment results to resemble a predetermined incident history will almost certainly create errors.

Since weightings can obscure the real risks and interfere with risk management, their use should be discontinued. Using actual measurements of risk factors avoids the incentive to apply artificial weightings (see the previous column on the need for measurements). Therefore, migration away from older scoring or indexing approaches to a modern risk-assessment methodology will automatically avoid the misstep of weightings.

Kent Muhlbaner contributes a column to each edition of *Pipelines International* tackling specifics of pipeline risk in bite-sized portions to make this challenging subject more approachable.

FUTURE COLUMN TOPICS

- Consequences of failure–ID the scenarios
- The "Perfect Storm" chain of events
- How do I handle non-pipe assets?
- Getting info from SME's – facilitation!
- Monetisation of risks – a useful common denominator
- How safe is "safe enough"?
- Damage vs failure – an important distinction
- Q&A on risk assessment evolution