Myth busting – I don't have enough data (Part 2)

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

The first part of this discussion, which was published in the June edition of Pipelines International, hopefully dispelled some myths about low data availability. We discussed how reasoning is used to generate data and how many useful pieces of risk insight emerge from even simple pieces of knowledge.

We also contrasted a statistical approach to risk assessment with a physics-based approach. The point of that discussion is that we should always use reasoning in addition to whatever historical event frequencies we might have. What happened in the past rarely paints the complete picture – it is only a piece of the puzzle. Let's explore that a bit more here.

"I don't have enough data to quantify risk"

First, however, let's address the myth that low information suggests the use of a simple risk assessment – one that does not really quantify risk. Using a lesser risk-assessment process in an attempt to compensate for low information is an error. Pairing weak data with a weak model generates nothing useful. The proper approach is to begin with a full risk-assessment structure, make conservative assumptions where necessary, and then work on 'back-filling' the data that will ultimately drive the risk management.

So, we should use a robust risk assessment, regardless of the current data availability, for which there are two choices. Let's compare the statistics-based and the physics-based approaches in solving a typical risk-assessment problem: how often will a specific segment of pipeline experience failure from outside excavator force (third-party damage)?

Statistics-based approach

In this approach, we focus on historical event frequencies. Let's say that a slice of the national pipeline incident database shows that US transmission averages show 0.0003 reportable third-party damage incidents per mile per year. With some investigation, we can get averages for ranges of pipe diameters, product type, or other characteristics, should we believe that they are discriminating factors for third-party damages. We can assume that some of the historical 'unknown' causes of failure (a significant proportion of the data) were also third-party damage related. We can further assume that the entire population of

third-party failures is higher than the reportable-only count. At the end of this exercise, we have a decent estimate of a historical failure rate for an 'average' pipe segment.

Physics approach

In this approach, we focus on the physical phenomena that influence pipeline failure potential. We first make a series of estimates that show the individual contributions from exposure, mitigation, and resistance. For exposure, we ask "how often is there likely to be an excavator working near this pipeline?" We perhaps examine records in planning and permitting departments; take note of nearby utilities, ditches, waterways, or public works that require routine excavation maintenance; and tap into other sources of information. Then we estimate the role of mitigation measures as applied to this particular segment of pipe. We ask "What fraction of those excavators will have sufficient reach to damage the pipe?" (suggesting the benefit of cover depth), "What fraction of

excavators will halt their progress due to one-call system use, recognition from signs, markers, and briefings?" and "What fraction will halt their work due to intervention by pipeline patrol?"

Finally, we discriminate among the fraction of excavation scenarios with sufficient force potential to puncture the pipe, based on pipe characteristics and the types of force likely to be applied. This tells us the resistance – how often is there damage, but not failure? This discrimination between damage likelihood and failure likelihood is essential to our understanding.

All of these estimates can come from simple reasoning, at one extreme, and from literature searches, market analyses, database mining, finite-element analyses, and scenario analyses, at the other extreme. The level of effort should be proportional to the perceived contribution of the issue to the total risk picture.

Approach comparisons

Both of these approaches have merit and yield useful insight. However only the latter provides the location-specific insights we need to truly manage risk. The statistics-only approach yields an average value, suggesting how a population of pipeline segments may behave over time. There are huge differences among all the pipeline segments that go into a summary statistic, and therefore, we cannot base risk management on such a summary value. Risk, and hence 'risk management', ultimately occurs at very specific locations, whose risk may be vastly different from the population average. Stated even more emphatically, using averages will always result in missing the 'generally rare but critical at this location' evidence. For example, most pipelines are not threatened by landslides, but in the few locations where they are, this seemingly rare threat may well dominate the risk.

So, we use the physics-based approach to drive risk management. Using the statistics-based approach is very useful in calibrating risk estimates from populations of pipe segments – but more about that in a future article.

Another aspect of data availability

As we conclude this discussion of data availability, let's not dismiss the bona fide 'absence of key information' scenario. It is not uncommon for an operator to have inherited a system with a genuine lack of basic data. Perhaps a gathering or distribution system, assembled over decades and with very poor records, has been acquired. Even basic location and materials of construction data might be missing. This is frustrating for a prudent operator wanting to understand risk. He might also encounter resistance in moving resources towards improving the information status.

Information acquisition can be considered risk reduction, when uncertainty is modelled as increased risk. Therefore, a cost-benefit for the information collection efforts can be shown. This is of use in demonstrating the value of information collection.

Here are some steps which can, over time, remedy the absence-of-information situation using risk management techniques.

STEP 1

Formalise and centralise all available information – collect and digitise every scrap of paper in every filing cabinet and every piece of information in the minds of all the experienced personnel and all information that becomes available in the course of operations and management (O&M). This means building a robust database and establishing processes to make its upkeep a part of day-to-day O&M processes.

STEP 2

Next, perform a risk assessment using all of this information plus conservative defaults to fill in the knowledge gaps. This will produce risk estimates based on both actual risk and risk driven by the conservative defaults.

STEP 3

Finally, use these risk estimates to drive an information collection process. This might require that resources be initially spent specifically on filling knowledge gaps – conducting surveys, inspections, tests, etc., solely to gain the information that can replace the conservative defaults and thereby reduce the 'possible' risks.

In this approach, the risk assessment itself identifies the most critical information to collect. This is an efficient and defensible strategy to tackle the 'lack of data' issue. ●

> **Information acquisition can be considered risk reduction, when uncertainty is modelled as increased risk. Therefore, a cost-benefit for the information collection efforts can be shown. This is of use in demonstrating the value of information collection.**

FUTURE COLUMN TOPICS

- The troubles with weightings
- Damage vs failure – an important distinction
- Measuring damage potential – what is attacking? How effective are defences?
- Consequences of failure – ID the scenarios
- 'The perfect storm' chain of events
- How do I assess non-pipe components?
- Getting info from SMEs – facilitation!
- PL risk assessment – is it helping me?
- Monetisation of risk – a controversial common denominator
- How safe is 'safe enough'?