Return to Session

VALIDATING CABLE "DIAGNOSTIC TESTS"

Miroslav BEGOVIC, School of ECE - Georgia Institute of Technology, (USA), miroslav@ece.gatech.edu
Nigel HAMPTON, NEETRAC - Georgia Institute of Technology, (USA), nigel.hampton@neetrac.gatech.edu
Rick HARTLEIN, NEETRAC - Georgia Institute of Technology, (USA), rick.hartlein@neetrac.gatech.edu
Joshua PERKEL, School of ECE - Georgia Institute of Technology, (USA), jperkel@gatech.edu

ABSTRACT

Diagnostic Techniques are increasingly employed by utilities to manage their infrastructure assets. These are sophisticated techniques being applied to complicated and diverse real world networks. Consequently there are many concerns that these techniques a) are not accurate and b) damage the system by, at the very least, robbing other areas of vitally short resources. Thus there is a compelling need to develop and deploy simple and robust analytical techniques that can address these problems. These evaluation approaches would then identify the effective programmes such that support could be strengthened to these areas, whilst minimizing the resources deployed on approaches that are ineffective.

KEYWORDS

Cable, Accessory, Diagnostic Tests, Reliability

INTRODUCTION

Utilities the world over, and especially in North America, are facing a significant future challenge to maintain and renew their underground (cable) assets. These ageing assets (>20% of the presently installed cables are older than their design lives) are leading to ever increasing failures (Figure 1) whilst, at the same time, the power delivery requirements of some of these cables are increasing. Immediate replacement of these aged cables is not practical – the cost would be enormous and the resources required (manpower and materials) are simply not available. Thus asset management strategies are increasingly being used to help address the issue, such that the replacement of the ageing infrastructure is managed.

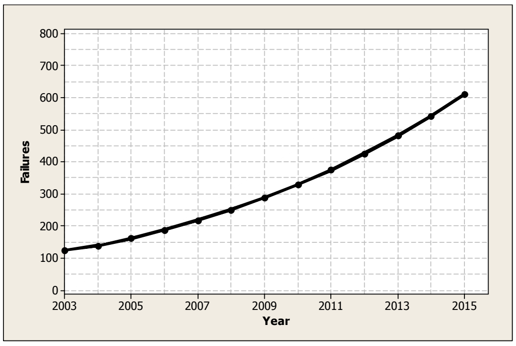

Figure 1: Example of increasing failure rates

A central component of the approach to asset management is the availability of appropriate information on underground assets. Although it is well known that old and unjacketed cables are the least reliable group, not <u>every</u> old or unjacketed cable is at *"death's door"*. Thus extra information is needed if a utility is to undertake *"smart maintenance",* that is, replacement of only those cables that will likely impact the near future reliability. This information is invaluable in helping to determine where maintenance and replacement funds should best be spent. Performance modeling supported by good quality and reliable diagnostic information can be a powerful tool for establishing a) the correct level of resources and b) the most effective way that they may be employed.

It is therefore clear that if we rely on diagnostic information to have an effective asset management programme, then we need to be certain that the information gathered is both relevant and accurate. We find it convenient to term this the Diagnostic Yield. In this area, most practical engineers recognize that results from diagnostic tests are not perfect (accuracy close to 100%). However, certain assurance is needed to ensure that the funds used to conduct diagnostic tests are well spent. They must deliver higher value compared to replacement and repair strategies based on chance selection. To this end, we have examined a number of ways to test and validate diagnostic information against the true system performance. As there are a large variety of diagnostic techniques at a utility's disposal, we have further concentrated on the methods that are 'technique independent" and applicable to all cable systems.

It is not the intention of this paper to dwell on the well known issues associated with either the diagnostic techniques themselves or their interpretation. Instead this paper focuses on a number of the methods we have developed to assess how well diagnostic information on cable systems relates to the performance of a specific system. Primarily this means comparing the predictions from the diagnostic information with real life both before and after the diagnosis. The paper will look at three main approaches:

- Direct Comparison - do the cables identified as "Bad" fail in service or, perhaps more importantly and rarely addressed, do the "Good" not fail?
- Performance Ranking - consideration of the whole continuum of performance (not just "Good" and "Bad") as measured by diagnostic data and correlation/validation with service experience.
- Diagnostic Outcome Maps – how the failures in service are affected by selection, testing & maintenance actions.

The implications of the Diagnostic Yield upon the economic value models for Diagnostic Testing will also be discussed.

SAGE

The process of employing diagnostics to increase the efficiency of reliability improvement contains four elements that are summarized as:

• Selection – Choose the circuits for testing that will produce a high Diagnostic Yield. Typically this selection is based on age, failure rate, or other engineering judgment.
• Action – What actions will be performed as the result of certain diagnostic outcomes or interpretation? The actions are in two groups (Act or Not Act) and may include replacement, defer action, rejuvenation, and/or repair. These actions are chosen based on those that are most suitable for the system topology (conduit or direct buried) and most prevalent failure mechanisms (local or global defects).
• Generation – Diagnostic tests generate data that are well fitted to the type of maintenance actions and prevalent failure mechanisms.
• Evaluation – Are the methods employed for Selection, Action, and Generation, giving the expected results: lower rates of failure and increased times between failure? Can the diagnostic elements be improved?

Evaluation may appear to be a theoretical issue yet it has a profound practical influence. Diagnostic tests have costs in the range of a few thousand dollars per day, and these monies are most often taken from the overall maintenance / replacement budget. Thus an inappropriate or ineffective diagnostic will reduce the resources available by a few hundred meters of cable for each day of testing.

The focus of this paper is on the methods of Evaluation, both in terms of the accuracy of the individual diagnostic technologies (Generation) as well as the overall programme (Selection, Action, and Generation). It is important to understand that employing diagnostics effectively is a process that requires careful analysis and consideration before the first test is performed. The results of Evaluation depend heavily on how well this process has been conducted. This understanding / analysis begins with the data that are to be generated.

DATA

The analysis techniques that will be discussed in subsequent sections were developed considering the availability of data within US utilities in the period 2000 - 2006. Two types of data are needed: (1) diagnostic performance data and (2) service performance data. The level of detail contained within each of these data types is important as it limits the detail that may be obtained from any analysis to the coarsest level of the input data. In other words, if diagnostic data is available for each segment of a feeder circuit and performance data is only available for the feeder as a whole, then the analysis is limited to looking at the feeder as a whole. This requires that the diagnostic data be transformed in such a way that the condition of the entire feeder is extracted from the condition of its individual segments. The process may be performed using a weighted average based on the relative lengths of the segments. As a rule, this applies to any data (diagnostic & service) that is considered for analysis.

The following sections discuss the two data types needed for the validation techniques.

Diagnostic Performance Data

Diagnostic data are available in a plethora of formats and may be used in any form; but only provided it includes enough information to be able to distinguish between the circuits. This becomes an issue with interpreted results that provide a class membership (i.e. this is a Level 3); as many circuits may belong to the same class. We term these as “tied data”. The more detailed the measurement data, the less likely ties will arise. The preference, therefore, is have numerical measurements available but more qualitative information can be used. The difference between the quantitative and qualitative data is in the level of interpretation needed by the analyzer. It is also important to note that the diagnostic data must include a minimum of circuit data in order to combine the diagnostic data with the correct failure data. The following list, although not exhaustive, summarizes the minimum information needed:

- Circuit identification (i.e. segment, feeder, etc.)
- Date of test
- Cable type (i.e. XLPE, EPR, PILC, etc.)
- Type of diagnostic (i.e. PD, Tan δ, etc.)
- Numerical data
- Test protocol - Voltage Source (60 Hz., 0.1Hz., DC, etc.), Voltage magnitude, Test duration

Service Performance Data (Utility)

Service performance data needs to obey similar rules to that of the diagnostic data with an emphasis on the circuit information. In this case, the numerical data corresponds to the number of failures before and/or after the testing. The minimum information required is as follows:

- Circuit identification (typically by feeder)
- Date for each failure
- Circuit length

Circuit length becomes very important for service data as circuits typically do not experience more than a few failures; thus they are highly “tied”. Therefore, with some validation techniques it is more useful to look at failure rates, say failures per km or mile per year, rather than the total number of failures.

VALIDATION TECHNIQUES

The following sections describe the validation techniques we have developed or considered.

Direct Comparison

Direct comparison is the method that has been generally employed by workers to evaluate the effectiveness of diagnostic testing. It compares the results from diagnostic testing with the outcomes in the field by looking to see if the areas identified as “Bad” by the diagnostic actually failed within a reasonable time following the testing. This method of assessment is very onerous and typically produces overly conservative results. This effect occurs for a number of reasons:

• Method ignores cables that diagnostic shows as “Good”.
• Requires that diagnostic has the ability to clearly separate cables into “Good” and “Bad” groups with no overlap. In our experience this condition is never fulfilled using the features and tools employed by most diagnostic tests.

- Anything less than 100% accuracy gives "Bad" results.

Notwithstanding the above, one thing is certain with a Direct Comparison: if this method shows that things are, in fact, "Good" then the reality is that the diagnostic has done an excellent job of identifying the "Bad" components. On the other hand, if Direct Comparison indicates the diagnostic was not effective then other methods should be employed to evaluate the diagnostic's performance as this method tends to exaggerate any imperfections.

Performance Ranking

This technique has been developed in our group as a means of evaluating the effectiveness of diagnostic testing by comparing the diagnostic data with real world performance. This comparison provides measures, quantitative & semi quantitative, of the accuracy of the diagnostic. Performance Ranking is the only technique that looks at the entire spectrum of data all the way from the best to the worst. It is not focused on the "Bad" segments as in the case of Direct Comparison. In addition, it may be used with any diagnostic test was well as with data provided in any form more detailed than a simple pass/fail. Example diagnostics include (but are not limited to) partial discharge (offline and online) and Tan δ. This facility is essential due to the multiplicity of data (diagnostic and service) formats.

Table 1: Illustrative Performance and Diagnostic (*A, B & C*) for 5 circuits.

Circuit	**I 200m**	**II 800m**	**III 150m**	**IV 200m**	**V 100m**
Failures in Service	1	2	1	0	1
Performance Rank	**4**	**1**	**3**	**5**	**2**
Diagnostic *A* – *level based*	0	5	4	2	2
Diagnostic *A* Rank	**5**	**1**	**2**	**4**	**3**
Diagnostic *B* - *value based*	5	20	5	10	22
Diagnostic *B* Rank	**5**	**2**	**4**	**3**	**1**
Diagnostic *C* - % to be replaced	2%	15%	0%	15%	20%
Diagnostic *C* Rank	**4**	**3**	**5**	**2**	**1**

Performance Ranking is completed by generating two distinct ranks (a number representing 1st, 2nd, 3rd, etc), Performance Rank and Diagnostic Rank, for each tested segment. Each of these ranks is a number that gives the relative performance of each circuit as it compares to all other circuits in the group. There cannot be duplicate ranks within either rank type. Furthermore, all circuits must be assigned both a Performance and Diagnostic rank for plotting. In other words, if a test group consists of 10 segments then there will be at most one #1, one #2, one #3, etc., for the Performance Rank and then the same would hold true for the Diagnostic Rank.

The basic procedure can be summarized as follows:

1. Determine the "Diagnostic Rank" using the available diagnostic data and the circuit information.
2. Determine the "Performance Rank" using the available failure and circuit information.
3. Plot Diagnostic Rank versus Performance Rank.
4. Determine the best line with a statistical method.

The concept of ranking the circuits is quite simple. However, with test groups containing more than a few feeders it is very likely that there will be cases where the ranking criteria produce ties (Table 1). As one of the requirements of this technique is to assign a single rank to each circuit, breaking these ties becomes critical. Several methods have been developed to address this issue for both ranks. Each will be discussed in detail in the following sections in conjunction with the steps outlined above.

Performance Rank

The Performance Rank is based on the failure data from either before or after testing. It is determined by comparing the failure rates (annual or cumulative) for all tested components with one another and then ranking from worst (highest failure rate) to best (lowest failure rate). In the case of cables, the ranking is most commonly done considering multiple segments together as one "feeder" as this is the extent of detail available in the failure records at most utilities. This is not an issue for other devices such as transformers, breakers, or poles. However, it is important to note that the ranking approach is able to cope with whatever detail is available in the data.

Diagnostic Rank

The Diagnostic Rank is far more complicated to determine than the Performance Rank as different diagnostic techniques provide their assessments in different ways. The data may be quantitative measures of the degradation that has occurred in the device or may simply be qualitative such as "Good", "Bad", or "Okay". Furthermore, this data may be as specific as by component or may include several components at once. Whatever the level of detail may be, it is necessary to evaluate the diagnostic data in the same groupings as the performance data. An example would be, two cable segments may have been tested separately, however, they must be considered as one because the failure records do not distinguish between them.

Listed below are some examples of available cable diagnostic data that has been successfully analyzed using Performance Ranking, note that this list is in no way exhaustive:

- Recommended sections for replacement – C in Table 1.
- Partial discharge magnitude and count – B in Table 1.
- Tan δ values – B in Table 1
- Severity – A in Table 1

It must be emphasized that the only requirement for diagnostic data is that it be capable of providing some level of distinction between different circuits.

Ranking Tie Breaks

As mentioned above, it is common to see situations (Table 1) where ties can arise, especially in the case of the Performance Rank. These cases can be solved in a number of ways depending on the type of component. For example, in the case of cables, most ties may be dealt with by normalizing by the amount of cable. Other components such as breakers and transformers may be handled by considering age or even the number of exposures to fault currents. Whatever the device, the key is to choose a characteristic that will include sufficient variability within the component population. It is also possible, however, that multiple characteristics will be needed in order to break all the ties. Therefore, we are able to rank any dataset, however, the level of interpretation differs depending on the number of characteristics we need in order to break the ties.

The following hierarchy was developed for cables based on the circuit information that is typically available at utilities:

- Measurement data (Diagnostic Rank).
- Circuit length: Average per unit length. Also, longer circuits should be more prone to failure so give higher rank to longer circuits – Table 1.
- Number of accessories: More accessories lead to more opportunities for failure so give higher rank to circuits with more accessories.
- Age: Older circuits receive a higher rank as these are logically more prone failure.
- Construction: Primarily, insulation type, however, this should also include jacketing, whether the cable was direct-buried or installed in conduit, and type of neutral.

Analyzing the Ranks

Once the two ranks have been computed they may be analyzed either graphically (qualitatively) or statistically (quantitatively). In the former case, a plot of Diagnostic Rank versus Performance Rank is generated. A sample of such a plot is shown in Figure 2 which uses real diagnostic tests using different approaches on a cable system.

The accuracy of the diagnostic test is directly related to how far from the dashed line the dots are in Figure 2. The interpretation of Figure 2 is as follows: Circuits in the lower left corner, we consider by convention to be the worst performers (highest failure rate and classified as "Bad" by the diagnostic test) while the upper right corner contains the best performers (low failure rate and classified as "Good" by the diagnostic test).

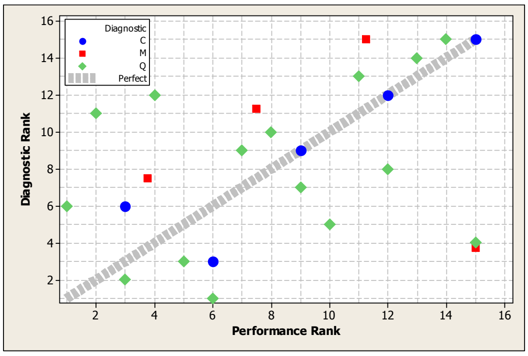

Figure 2: Sample Performance Ranking plot. Dots represent individual segments whereas the dashed line represents what would be a perfect correlation between Performance and Diagnostic ranks.

The dashed line can also be thought of as the perfect correlation between the performance and diagnostic ranks. Therefore, the obvious statistical approach is to examine the Pearson Correlation Coefficient [3], [4] between the two ranks as shown in (Equation 1).

$$r_{DP} = \frac{n\sum D_i P_i - \sum D_i \sum P_i}{\sqrt{n\sum D_i^2 - \left(\sum D_i\right)^2}\sqrt{n\sum P_i^2 - \left(\sum P_i\right)^2}} \qquad \textbf{(1)}$$

Where r_{DP} = Pearson correlation coefficient, n = number of samples, $D_i = i^{th}$ Diagnostic Rank, $P_i = i^{th}$ Performance Rank

The value of the correlation coefficient is in the range [-1,1] where one corresponds to a perfect correlation, zero to no correlation, and negative one to an inverse correlation. Comparison of these coefficients enables the computation of the difference in accuracies between different diagnostic technologies.

Table 2 shows the Pearson correlation coefficients for the data shown in Figure 2. The respective significance levels indicate that only the results from Diagnostic C are not likely to occur randomly (with probability <0.05).

Table 2: Pearson correlation coefficients for three diagnostic techniques.

Diagnostic Technique	r_{DP}	Level of Significance
C	**0.900**	**<0.05**
M	**-0.200**	**>0.1**
Q	**0.321**	**>0.1**

Figure 2 deals with service data using 3 diagnostics, the example described in Figure 3, looks at field-aged XLPE samples in which Tan δ was measured at 0.1Hz. using a sinusoidal VLF source at U_0 and then taken to failure (breakdown) using a VLF (0.1 Hz) step ramp protocol. Figure 3 shows the breakdown voltage versus Tan δ value for each of the samples. Note that the samples shown in blue failed during 60 Hz. Tan δ measurement.

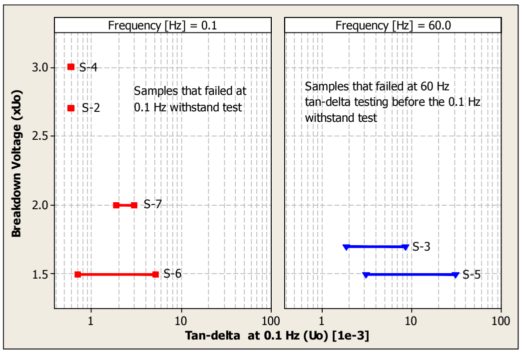

Figure 3: VLF breakdown strength versus average Tan δ value for service aged cables measured in the laboratory.

Direct inspection of Figure 3 is not a straightforward way to understand the manner in which the Tan δ value predicts the breakdown strengths of the samples. However, if the data are analyzed using Performance Ranking, the picture becomes much clearer. Table 3 shows the ranking criteria that were employed to rank this group of samples. Figure 4 shows the resulting Performance Ranking plot for this data.

Table 3: Criteria used to rank the samples.

Rank	Initial Criterion	Tie Break Criterion
Diagnostic	Average Tan δ (high Tan δ ⇒ low rank)	Variance of Tan δ
Performance	Breakdown Voltage (low BV ⇒ low rank)	Time on test

Return to Session

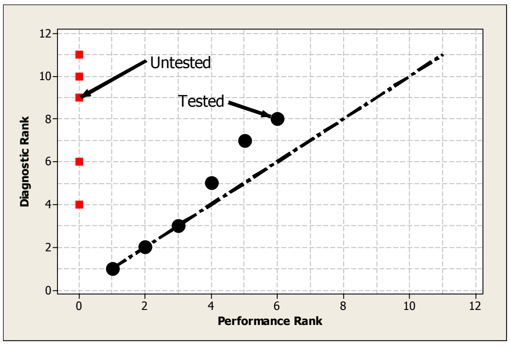

Figure 4: Performance Ranking plot for Tan δ and breakdown strength for data shown in Figure 3.

Computation of the Pearson Correlation coefficient for the six tested samples yields a value of 0.991 which is significant at the 0.001 level. Therefore, these results would only occur randomly with a probability of less than 0.1%.

Diagnostic Outcome Mapping (DOM)

DOM uses failure and testing data to evaluate whether or not changes in failure rate are coincident with diagnostic activities (as well as the activity called for by the test results). DOM uses a graphical representation for analysis. Experience with many analyses has shown that the well established Reliability Growth Model (often referred to as Crow-AMSAA technique specified in IEC 61164 [2]) is very suitable. Crow-AMSAA is a plotting technique that plots cumulative failures versus time on log-log scales.

The instantaneous failure rate is determined by computing the slope, or gradient, of the curve at any particular point. A decreasing gradient indicates the failure rate is decreasing while an increasing gradient corresponds to an increasing failure rate. By adding the testing events to the same representation, it is possible to determine the effect that the magnitude of the test programme (diagnostic testing plus required action) has had on the reliability.

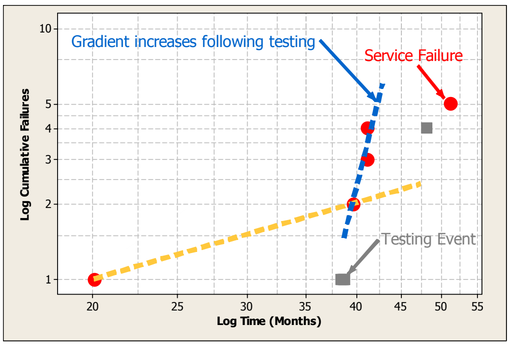

Figure 5: Sample outcome map in which the failure rate increases following a test. Dots represent service failures and squares represent test activities.

Figure 5 shows one of the possible outcomes that can occur for a particular cable circuit (accessories excluded from this analysis). An increasing failure rate is represented by an upward slope. The higher the gradient the larger the failure rate. Thus, in Figure 5 it is seen that testing at 38 months, after a quiescent period of 18 months, leads to 4 failures within 4 months. The alteration in gradients on the CROW-AMSAA plot shows this change in failure rates. In this case the testing event at 38 months would seem, on this singular circuit, to degrade the reliability.

It is also possible and quite common to see constant rates and decreasing failure rates following testing events for individual circuits and then even for an entire system. Figure 6 shows an example of an overall system outcome map since the initiation of a diagnostic testing and action programme. This approach extends the analysis in Figure 5 from a single circuit to a larger at risk population.

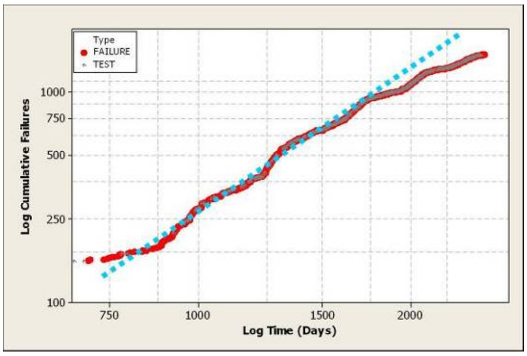

Figure 6: Sample system outcome map that shows reduced failure rate.

In Figure 6 it is clear that there has been little change during the first three years of the programme; however, the programme yields a reduced failure rate, seen by the depression of the data from the straight line, during the last three years. This observation may be investigated further by examination of the annual gradients for the curve in Figure 6. Comparison with the amount of testing and action accomplished within a particular year provides the results shown in Figure 7.

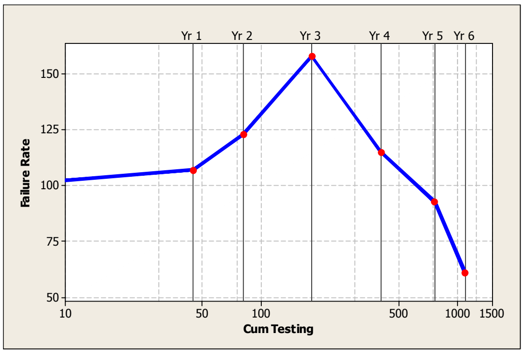

Figure 7: Failure rate (Rate at start of the programme = 100) versus cumulative tests for an overall testing and action programme.

Figure 7 clearly shows the annual gradients increase out to year 3. We presume that this represents the continuation of the pre programme trend. The failure rates then began to decrease after year three. That year corresponds to the point at which the number of tests and actions exceeded a critical level, somewhere around 175 actions, where the programme could bring the failure rate under control. This analysis illustrates the usefulness of DOM in showing real benefits of a diagnostic testing and action programme. It also shows that for a programme to make a real impact on the system reliability it must reach a minimum level of activity. This level will differ from system to system and will require a long term commitment on the part of the utility if successful results are to be achieved.

ECONOMIC & RELIABILITY IMPLICATIONS

The detailed economics of a diagnostic testing and action programme involve numerous interdependencies that are ultimately beyond the scope of this paper. However, when analysing the cost of maintenance, based on the results of diagnostics, it becomes very clear that two rarely considered issues are extremely important. These issues are diagnostic accuracy and system quality disbursement (i.e. how much of the system is in need of maintenance). Information (failure rates, replacement costs, outage costs, diagnostic costs) from US utilities enables these issues to be illustrated in scenarios summarized in Figure 8.

The analyses show that the costs increase with the ratio of defective / non defective elements. The savings will accrue from the difference between the costs and the reference case. Any reference case will depend upon the specific details of the location and the Failure Tolerance of customers and regulator. Figure 8 also shows that as the diagnostic accuracy decreases, the cost ranges increase dramatically regardless of the system composition. Furthermore, the cost ranges for systems with low percentages of “Bad” components are far more sensitive to the accuracy of the diagnostic than those with higher ratios of “Bad” components. Therefore, the value that can be obtained from diagnostic testing is very much related to the accuracy of the diagnostic as well as the system to which it is applied. Of equal importance is the fact that the significant cost range will mean that year on year the costs will be indeterminate and variable. This will cause problems with budget planning.

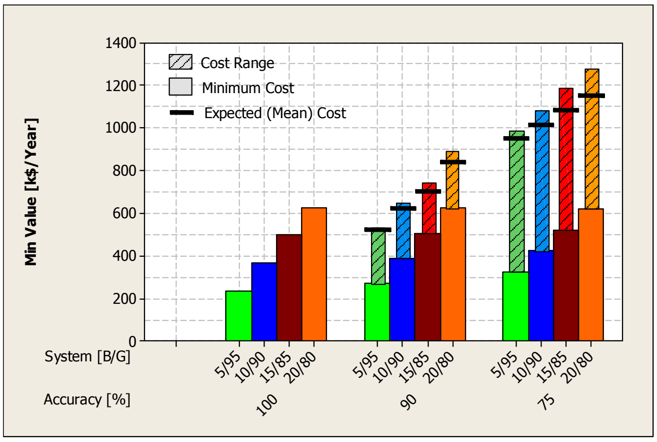

Figure 8: Effect of diagnostic accuracy and system composition on the possible annual maintenance costs. Solid bars indicate the minimum cost for diagnostics with accuracies less than 100% while striped bars show the cost ranges.

CONCLUSIONS / FUTURE WORK

This paper describes two primary techniques for evaluating the accuracy and effectiveness of diagnostic techniques and utility programmes for employing them. It is clear that the accuracy of a diagnostic will greatly influence the value that a utility can hope to obtain. Furthermore, utilities must be aware that this value also depends heavily on their understanding of their systems as well as their abilities to Select and Act in ways that take advantage of what diagnostics can offer. This is why Evaluation is so important within a whole Asset Optimisation programme.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the useful discussions with many of the engineers involved within the Cable Diagnostic Focused Initiative.

The work reported here was supported by a large number of utilities in North America and the US Department of Energy under award number DE-FC02-04CH11237

REFERENCES

[1] IEEE Power Engineering Society, “IEEE Guide for Field Testing and Evaluation of the Insulation of Shielded Power Cable Systems,” IEEE Std. 400-2001.

[2] International Electrotechnical Commission, “Reliability Growth – Statistical Test and Estimation Methods,“ IEC 61164, Second Edition 2004-03.

[3] E.W. Weisstein, "Correlation Coefficient." From *MathWorld*, Wolfram Web Resource. http://mathworld.wolfram.com/CorrelationCoefficient.html

[4] G.E.P. Box, W.G. Hunter, J.S. Hunter, 1978, *Statistics for Experimenters*, John Wiley & Sons, Inc., New York, NY, USA. 60-91.