PMF Series for Availability and Reliability Probabilistic Assessments

Jan B. Smith, PE, Zachry Holdings, Inc.

Key Words: availability, probability distribution, probability mass function, reliability, risk assessment

SUMMARY & CONCLUSIONS

PMF series (a contiguous family of probability mass functions) is a new method for probabilistic assessment of availability and reliability of both repairable and non-repairable systems. The same general methodology is applied to both availability and reliability. This allows availability assessment of complex systems with both repairable and non-repairable subsystems. Binominal and Weibull can be used in narrow circumstances, and in those cases results match those of PMF series, thus validating the more general PMF series. The method uses small quantities of system data that is always available for important operating systems, whereas Monte Carlo requires substantial data or assumptions.

The probability tables that are developed are analogous to the capacity outage tables of the electric power industry. However, PMF series is more broadly applicable as it also applies to single plants, the more common situation in manufacturing. Probability tables are embedded in specific user application software for efficient day-to-day business and operating decisions. These may include applications such as product inventory optimization and the absolute and relative risk difference between operating strategies.

This practical methodology using available data promises to greatly expand the use of probabilistic assessments. In the future, probabilistic risk may routinely be integrated into business and operating decisions. Practical methods have not heretofore been available for these decisions; therefore, decision makers are unaware that probabilistic risk can be used in their applications. Educating the decision makers in this technology is necessary.

The ability to measure and analyze risk that previously was unmeasured and unrecognized leads to surprise. Management and Reliability Engineers are generally unaware of the actual risk in meeting the production goals for their plants. When the risks are recognized, they are invariably larger than supposed. As a consequence, the reliability performance gap in any manufacturing company is perhaps significantly larger than single-valued measurement of mean availability, reliability and capacity would suggest.

1 INTRODUCTION

Conventional probabilistic analysis is not now widely used and its application for practical business and operational use is currently quite limited. A very significant exception is the electric power industry [ref 1]. Loss of load probability and loss of load expectation data are used for routine operating decisions as well as to evaluate the need for capital expansions. Why should probabilistic analysis be an integral part of the business in one industry, and yet be insignificant to most others? The key reason is that in the electric power industry the required data is readily available high level system data. In this particular case, mean availability (forced outage rate) and rated capacity of each unit of a system of units is all that is required to obtain probabilistic information.

The methods that serve special electric power industry needs so well are not normally applicable to other industries. Those methods provide no probabilistic information for a single power plant or a system of only a very few plants; although, this is generally the greatest need in industry.

Any method that is to play a significant role in the business decisions of an industry must use data that is readily available. What is already available for any important system is its availability performance data. PMF series uses this readily available data and only a small number of failures are required due to the use of data permutations, to be explained later.

PMF series extracts the likelihood, consequence and risk data that are embedded in system failure or availability data through a measurement process. The results are a dense series of unique probability mass functions that measure the probabilistic nature of the system. While results are analogous to electric power capacity outage tables, they are not limited to systems of multiple units. Single units and systems of single units in parallel and series are readily accommodated. The basic method of analysis applies to both reliability and availability and to both repairable and non-repairable systems and components.

The method is computationally heavy, but excellent results can be obtained on personal computers. Accuracy is increased with computation time. Results converge with theoretical in the special cases in which there is a theoretical method with which to compare. However, the usefulness of PMF series is not so much as a competitor with existing methods, but it provides excellent results where there is no alternative method.

The method was introduced in the 2002 RAMS symposium [ref 2]. At that time the method was not fully developed for reliability. That paper dealt primarily with availability and capacity of systems and the term PMF series was not used. PMF series is more descriptive of the total concept that applies to reliability and availability of repairable and non-repairable systems, equipment and components. Patents protect PMF series use for availability (not reliability) for other than academic and method evaluation purposes. Contact the author or ZHI Reliability and Asset Management Services for conditions and limitations of use, technical support, software, etc.

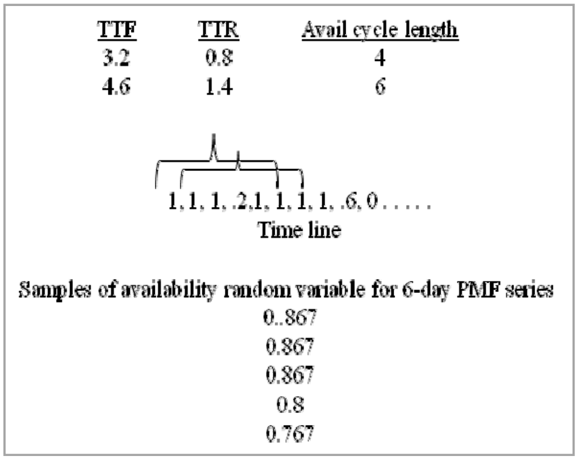

Figure 1
Illustration of two availability cycles of TTF and TTR forming a time line. Samples of the availability random variable are extracted, grouped into histograms and normalized to provide a probability mass function for every window length of interest. This may be 366 windows for forecasting up to one year in 1 day increments.

2 NEW DATA FORMAT

Preparing the data in a new format was presented in some detail in the 2002 paper, and is only summarized here. Time-to-failure (TTF) and time-to-restore (TTR) data are treated as an inseparable data set. Each set forms an available cycle composed of a series of numbers that range from 0 to 1. Each number represents the availability fraction for the smallest time interval of interest (STII). We will assume the STII is day for simplicity. These are laid end to end on a time line as seen in figure 1. Samples of the availability random variable for any time interval are obtained by incrementally advancing a window whose length is equal to the time interval of interest. In practice, this is a running average in a computer spreadsheet. Each position returns a sample of the available random variable for that window length. The time line is extended by considering permutations of the availability cycles. Permutations of the data are allowed if the data are independent. This is commonly the case for system data and can be confirmed with root cause analysis commonly conducted on important systems. Any dependence can be recognized and the data corrected accordingly. A typical small database contains numerous billions of available random variable values as explained and justified in the 2002 paper.

3 COMPARISON WITH BINOMINAL

The first distribution of the PMF series family of distributions can be validated for accuracy by considering a system of six plants (units) in parallel. The binominal distribution works in this special case and gives exact results. For convenience in using the binominal, the plant or units have the same mean availability and rated capacity. Figure 2 shows the formation of the system time line from those of the individual units. In figure 2, data permutations are used to develop unit time lines that are about one million days long. For unit 6, only a small sample of the 1.3 trillion data permutations (15!) are used while the smaller data base of unit 1 (5!*775=93,000 day time line) is repeated multiple times. The system time line is formed by the daily average.

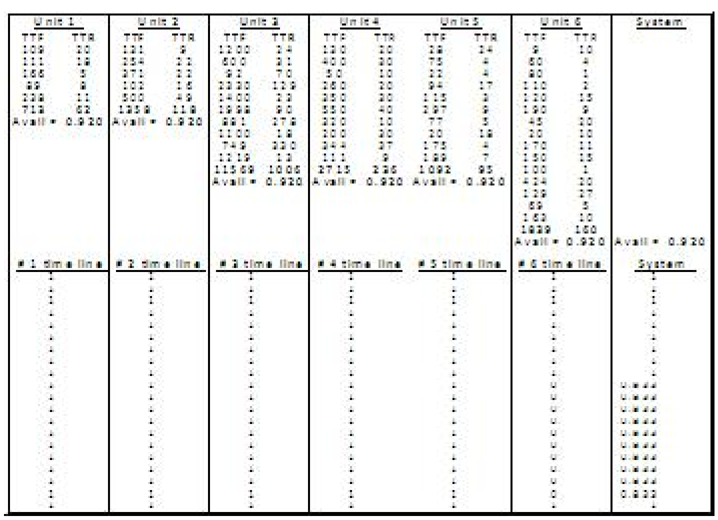

Figure 2
The time lines for six parallel units are developed that extend nearly 1 million days by use of data permutations. A system time line is formed on the right by determining system performance by day. Only the first 20 days are shown. This system time line is then used in PMF series according to figure 1 to determine the probability mass functions for each time window.

The first of the PMF series distributions, developed from a window length of 1 day, is compared with the binominal distribution in Table 1. The results from the two methods are nearly identical and closer accuracy is accomplished with additional computation time. There is no alternative method against which to compare all the remaining probability mass functions, but these are formed with the same methodology as the first distribution.

Units out of service	PMF series	Binominal
0	0.6068651	0.6063550
1	0.3159091	0.3163591
2	0.0682507	0.0687737
3	0.0083529	0.0079738
4	0.0006090	0.0005200
5	0.0000131	0.0000181
6	0.0000000	0.0000003

Table 1

Values of binominal and PMF series showing probability of x units out of service. The empirical PMF series gives essentially the same values, but is not limited to systems of several plants. It is also used on single plants and systems of a few plants.

4 OTHER AVAILABILITY DATA TYPES

Availability cycles of TTF and TTR data and their permutations are appropriate for systems such as power and petrochemical plants. For many other systems failures cannot be identified in terms of TTF and TTR. In chemical plants for example, there may be reduction in maximum production for multiple reasons on almost a daily basis. These are for a variety of reasons and can be process and operational problems as well as equipment related issues. They may result from human error and management system deficiencies, or may simply be related to normal variation in chemistry. Specific failures cannot generally be identified and they are most certainly not independent. Therefore, permutations cannot be used.

A wind plant is an example of a plant in which TTF and TTR data are not defined. The performance of the plant is related to wind velocity. Historical data and/or wind forecast could be used for PMF series. But wind speed during one hour is more likely to be low or high, depending upon the velocity in the preceding hour. Hour to hour data are not independent, so the wind plant is analyzed without data permutations with velocity normalized between 0 and 1.

In a manufacturing plant, one common way of obtaining system availability data is based on daily production quantities. All losses from expected production are accounted for and classified according to type of loss. Those losses from forced outages allow a ratio that ranges from 0 to 1 to be determined for that day. This is the availability for that day and is one of the sequence of numbers that makes up the time line of figure 1.

A chemical plant example is seen in Figure 3. The ratio of actual production to the sum of actual production plus unavailability losses is determined for each day. Data for 14 months, seen in figure 3, are laid end to end multiple times to extend the time line. This is done without using data permutations. The time line extension reduces error from the uneven weighting of data that occurs when the window of figure 1 starts and stops on the time line. This error is driven to trivial levels by making the time line long.

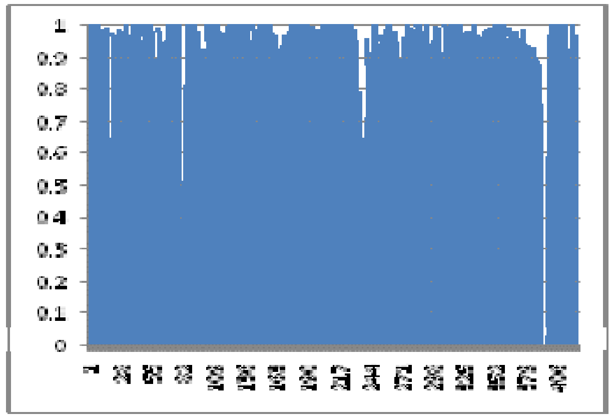

Figure 3

The availability by day for a 14 month period from a chemical plant. Obtained by measurement of production quantities and production loss classification.

Another type of system performance data useful for PMF series is based on downtime. The PMF series for a cement ball mill system, figure 4, uses data from a downtime collection and classification system over a one year period.

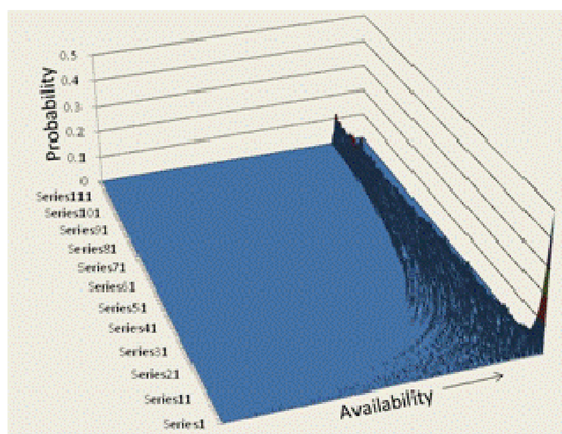

Figure 4

Cement ball mill system PMF series from downtime data. Individual failures are not identified in terms of TTF and TTR. Only downtime by classification was used. Each of the series is a unique availability probability distribution that is also a distribution of fraction of maximum capacity. The distributions are related to units of production by a multiplier that is plant rate.

Here a number ranging from 0 through 1 is calculated for each day from the uptime to total time ratio for the day. Downtime not related to available is excluded, such as inventory control downtime.

5 PROBABILITY AND RISK ASSESSMENT

The location and magnitude of each probability mass is known for all distributions. Any availability goal can be assessed for any and all distributions. For a process plant, the goal may be determined from the following:

- product shipping schedule or production plan
- product inventory available for use
- planned outage schedule
- plant rate

The goal may currently be a single-valued goal, such as production required to meet next quarter sales forecast. PMF series requires the above inputs be developed by STII (day for most manufacturing). Each day will have a different cumulative goal derived from the above information. The cumulative availability needed to match the cumulative production requirement will be irregular over time and evaluation may be required for each day. The look-ahead time may typically extend as long as one year. This would mean evaluating all 365 of the distributions because risk may rise and fall from one day to another over the course of the one year look-ahead. The look-ahead time could be shorter, such as the length of a peak sales season.

Data such as that in figure 4 is intersected with the required availability for each of the PMF series distributions. Probability, consequence and risk calculations are then made in the manner explained in the 2002 paper. When the probability table is embedded in user interface software, calculations are efficient. Immediate display of risk over time for any scenario allows risk to be managed to optimum levels.

Figures 5 and 6 are probability of product shortfall and risk of shortfall measured in units of production, respectively, for each of 120 look-ahead days for the cement ball mill system of figure 4. Using the probability table, risk is easily evaluated with changing inventory, planned outage schedule

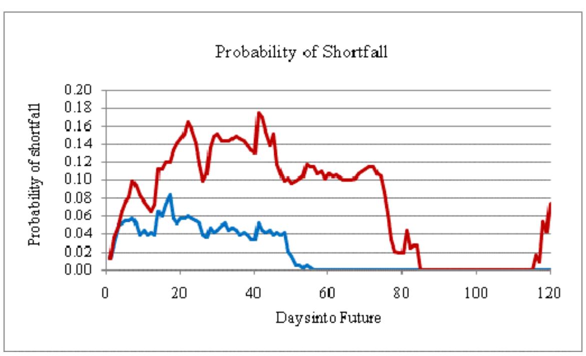

Figure 5
Probability of product shortfall for 120 look-ahead days for the cement plant of figure 4. Two different operating strategies are assessed. Probability of shortfall is a function of business situation as well as availability variation.

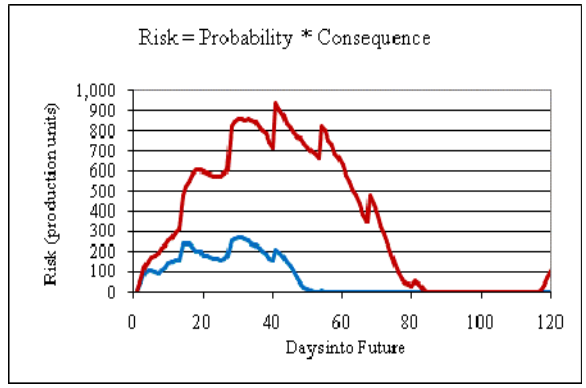

Figure 6
Risk assessment associated with figure 5. Risk is probability multiplied by consequence. Risk for two operating strategies is seen. Risk is measured in production units above. Risk is measured in dollars when the economic loss resulting from a unit of production shortfall is applied.

and new shipping requirements, allowing continual risk evaluation and management.

Realistic operating strategies are seldom risk free, so it is often useful to see the difference in risk between two strategies. Figures 5 and 6 show the difference between two strategies. Volatility of risk and probability can be due to variation in business circumstances, as well as plant availability.

6 RELIABILITY OF REPAIRABLE SYSTEMS

Obtaining the PMF series for repairable systems is a simple modification to the method. The availability cycle is terminated with a single zero that indicates failure. The analysis, including use of permutations, is otherwise the same.

Figure 7 is a top view of a 3-D plot of the reliability PMF series analysis for a power plant. Reliability is on the horizontal axis. The number of failures, time, reliability and probability of occurrence is easily obtained. As with

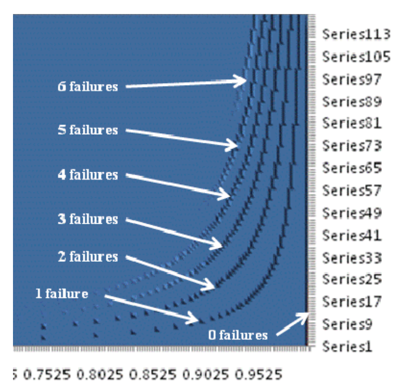

Figure 7
Reliability evaluation of a power plant as an example of a repairable system. Each series is a unique reliability distribution. The 3-D data can be sliced in a number of ways for reliability, number of failures, time and probability.

availability, this is an essentially exact measurement of the variability within the database of 30 forced outages over four years.

7 NON-REPAIRABLE RELIABILITY

Time-to-first-failure (TTFF) analysis requires parallel time lines formed by each failure. Each failure forms the equivalent of an availability cycle of figure 1, but is terminated with a single zero that reflects failure. In the case

#1	#2	....	#200	Reliability
1	1	....	1	1
1	1	....	1	1
1	1	....	1	1
1	1	....	1	1
1	1	....	1	1
0	1	....	1	0.995
	1	....	1	0.995
	1	....	1	0.995
	0	....	1	0.99

Table 2

Failure data #1 (TTFF=6), #2 (TTFF=9) and #200 (TTFF=120) are shown to show how the reliability of this non-repairable data is formed. Only the first 9 days are shown for illustration. Failures #3 through #199 are not shown for simplicity.

of data suspension, the failure cycle is terminated with 1. The numbers 1 and 0 may also be considered as success and failure, respectively. The time line is only one failure long – one cycle long. The reliability fraction at each time is formed in a manner similar to that of figure 2. Table 2 demonstrates how the reliability fraction is calculated for every STII and the time line formed.

To compare PMF series TTFF analysis with that using Weibull analysis, 200 failures were generated from a Weibull distribution of eta = 60 and beta = 2.5. The 200 failures were processed with the PMF series methodology and the assumption free results are plotted against the Weibull for

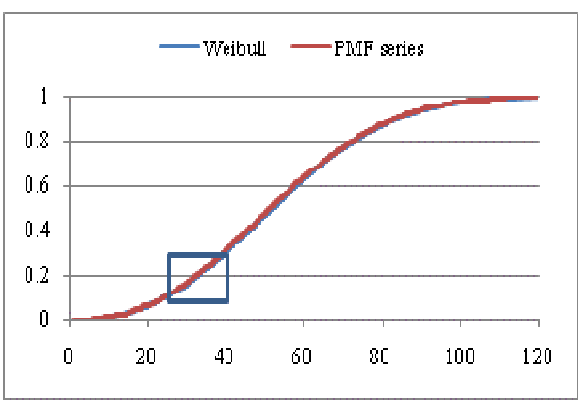

Figure 8

Cumulative probability of failure for PMF series overlays the Weibull. Insert is seen in figure 9.

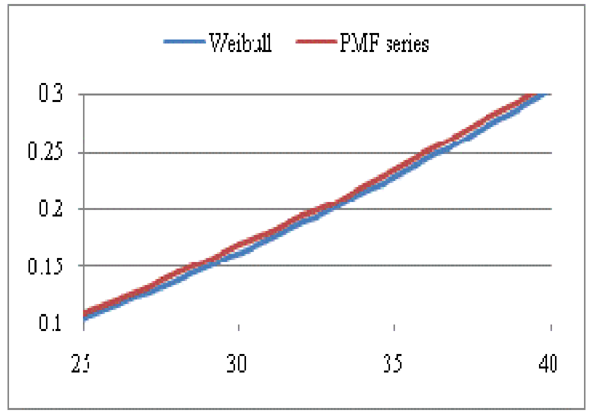

Figure 9

PMF series is empirical and discrete and can never exactly equal the continuous distribution that generated the failure data. However, unlike Weibull, there were no assumptions made in the PMF series analysis.

comparison. The results are essentially identical as seen in figures 8 and 9.

The significance of closely matching Weibull is:

- PMF series is distribution free. It can be used in cases where the assumption of identically distributed data cannot be defended.
- Because of commonality of the methodology, PMF series TTFF analysis can be combined with availability, as discussed in the next section.

8 COMBINING REPAIRABLE AND NON-REPAIRABLE

The system of measurement is the same for repairable and non-repairable systems. This allows complex systems containing repairable components or systems to be combined with non-repairable components and subsystems. Such a system is seen conceptually in Figure 10.

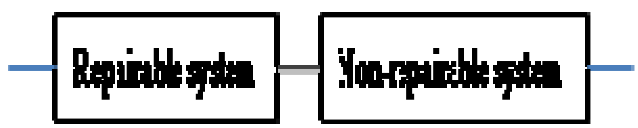

Figure 10

Illustrating a complex system with both repairable and non-repairable elements. Availability PMF series is obtainable for the overall system as a result of the measurement system being common to repairable and non-repairable systems and to availability and reliability.

The probability table of the cement plant from figure 4 is combined with the PMF series time-to-first-failure of figure 8 to obtain the PMF series availability for the complex system, figure 11. This is done by multiplying each probability mass of each series with reliability at that series and adding a probability mass at zero availability equal to unreliability.

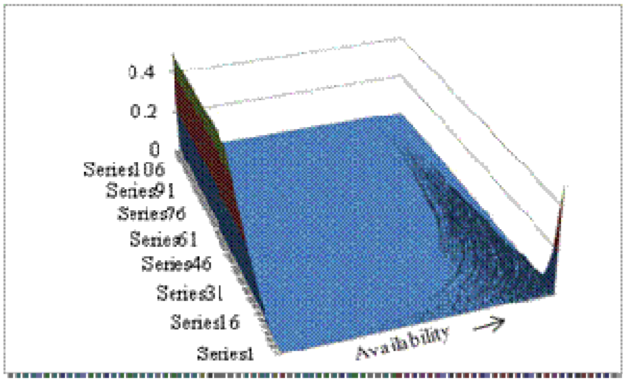

Figure 11
Cement plant availability of figure 4 and Weibull TTFF data of figure 8 combined for availability of the figure 10 total system. Probability scale is limited to 0.5.

9 BUSINESS IMPACTS ON PROBABILITY MASS

The probability and risk information available through PMF series unveils issues that have not been measurable. At the same time this new technology provides an opportunity to address these issues by incorporating unusual but fundamental considerations in our risk assessments that heretofore were impractical, if not impossible.

The unit valuation of a plant's product (and therefore the valuation of its availability) is normally considered constant. But as figure 12 concept shows, the unit value on the low side of a probability distribution should generally have a higher valuation. The value is greater during scarcity. When commodity producers operate on the low side of the probability distribution, they sometimes purchase product from their competitors in order meet their customer obligations. Their economic loss is greater than simply not selling the product. They incur it because losing the customer would be an even greater loss. The valuation on the high side of the distribution often should have a lower valuation.

When probability masses are adjusted to reflect these variable valuations, new distributions are formed. These new business probability mass functions have a mean that is lower than before. To illustrate, three chemical plants producing the same commodity product were evaluated with push sales margins substantially below those of pull sales. Pull sales are long term commitments with reasonable profit margins. Push sales are those sold on the spot market and carry much lower profit margins. Push sales volumes have no long term commitment and are normally surplus volumes such as when plant availability and capacity have exceeded expectation. For these three plants, push sales margins were applied to probability masses and their reflected volumes above the plant mean and pull sales margins to those below mean. The "business mean availability" ranged from 1.3 to 3.8 percentage points lower than the mean before business adjustment. This significantly lower effective mean is a generally unrecognized cost of unreliability.

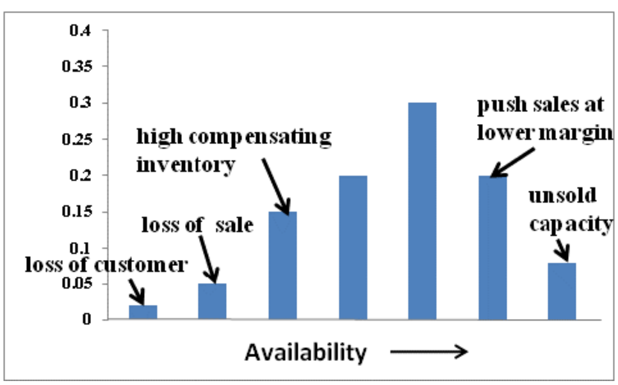

Figure 12
Weight of the economic impact varies along the distribution. Each probability mass can be assigned a different economic weight, and then normalized to provide a business probability distribution.

There are many situations that produce an additional cost of unreliability that have not been measured, and may even be unrecognized. As probabilistic analysis becomes more common, there will be greater recognition that the cost of unreliability is much higher than presumed. This should produce an increasing realization of the justification for even higher reliability.

REFERENCES

1. R Billington, R. N. Allan, Reliability Evaluation of Power Systems, 2nd edition, New York, Plenum Press, 1996.
2. J. B. Smith, W. B. Smith, "Probabilistic Assessment of Availability from System Performance Data," Proc. Ann. Reliability & Maintainability Symposium, (Jan.) 2002.

BIOGRAPHY

Jan B. Smith, P.E.
Zachry Holdings, Inc.
1499 Underwood Road
La Porte, TX 77571

E-mail: janb.smith@cebridge.net

Jan B. Smith, P.E., BSME, is a Registered Professional Engineer and serves as senior reliability consultant to ZHI's Reliability and Asset Management Services group. He has worked continually in Reliability Engineering for over 40 years. His experience includes reliability-centered maintenance, root cause failure analysis; finite element analysis; statistical and probabilistic analysis; vibration analysis; and capacity and availability forecasting. He has chaired national and regional conferences on plant reliability, authored several technical papers on the subject, and has organized and lectured at seminars on reliability issues. He has three patents on probabilistic assessments methods for availability, system and plant capacity.