Statistical Signature Analysis: Modeling Complex $\lambda_D(t)$ from Proof Test Data and the Effects on Computing PFDavg

by

Julia V. Bukowski, Ph. D.

for

exida.com L.L.C.

November 30, 2008

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Previously we analyzed several pressure relief valve (PRV) proof test data sets using quantal response analysis (QRA) and obtained estimates for λ_D, the constant failure rate for the "stuck-shut" failure mode, in the range of [10^{-8}, 10^{-7}] failures/hour. We also found that in that range of values, λ_D alone explained only a small fraction of the observed failures; the remaining failures were modeled by probability of initial failure (PIF) in the range of 1% - 1.6%.

Using Statistical Signature Analysis (SSA), we now re-analyze two of the data sets, this time considering more complex models of $\lambda_D(t)$ to explain the observed failures and to compute average probability of failure on demand (PFDavg). Our findings are:

- Any analysis of proof test data, based solely on operating times until proof test and results of proof test (pass/fail), must necessarily admit multiple explanations for the failure mechanisms underlying the test results, i.e., one data set will be statistically consistent with more than one $\lambda_D(t)$.
- Since PFDavg serves to measure risk reduction, it should include all dangerous failures regardless of underlying cause. By using more complex models for $\lambda_D(t)$ than are currently considered, e.g., models which can include constant λ_D, both zero or non-zero PIF, and models for infant mortality and/or failure rate growth, it is possible for $\lambda_D(t)$ to account for all observed failures regardless of underlying cause.
- Each different explanation of failures observed, i.e., each different $\lambda_D(t)$ model produced from a single data set, potentially produces a different value for PFDavg for any given proof test interval, T_P. Thus, one data set can support multiple values of PFDavg for any fixed T_P. However, *for the data analyzed to date* and for a fixed value of λ_D, these difference are small for T_P > 4.5 years. Generally, in these cases, PFDavg > 0.01. Hence, once T_P reaches about 4.5 years, the data analyzed support Safety Integrity Level (SIL) 1 behavior for all of the complex models of $\lambda_D(t)$ explored thus far. For values of T_P < 2 years, *some* complex models for $\lambda_D(t)$ support SIL 2 behavior. But we can not tell by currently available analysis methods which models best represent the actual failure mechanisms responsible for the proof test data observed. Thus, in general, current data supports SIL 1 behavior for PRV.

1. *INTRODUCTION*

In order to assign a SIL to equipment in low demand applications, we must be able to compute PFDavg. To compute PFDavg, we must first have a model for $\lambda_D(t)$, the failure rate of the equipment in the dangerous failure mode. A dangerous failure occurs when equipment designed for prevention or mitigation of an unsafe condition cannot properly respond to the unsafe condition, i.e., the equipment fails on demand. For example, consider a PRV, which, in normal operation, is closed. Should it fail in the "stuck-shut" mode, it would be in a state of dangerous failure as it would be unable to respond to an overpressure event if one occurred.

Once we have determined a model for $\lambda_D(t)$, we can compute PFDavg which is defined as

$$PFDavg = (1/T_P) \int_0^{T_P} PFD(t)\,dt = (1/T_P) \int_0^{T_P} \{1 - R(0) \exp[-\int_0^{t} \lambda_D(s)\,ds]\}\,dt. \quad (1)$$

Current practice for computing PFDavg is based on three assumptions: $R(0) = 1$, i.e., that each PRV is working at $t = 0$, $\lambda_D(t)$ is adequately characterized by a constant, λ_D, and proof testing results in perfect restoration. Based on these assumptions, PFDavg is usually approximated as

$$PFDavg \approx \lambda_D * T_P/2. \quad (2)$$

Does currently available evidence support the appropriateness of these assumptions? No. For PRV, the generally accepted range of values for λ_D is 10^{-8} to 10^{-7} failures/hour. (For some types of valves, some in the European safety community advocate even smaller values for λ_D, e.g., [1].) However, even if we allow λ_D to be as large as 10^{-7}, we find that λ_D accounts for only a small fraction of the total number of stuck-shut failures observed in available PRV proof test data sets [2].

While retaining λ_D as a constant, in [2, 3] Bukowski and Goble relaxed the assumption that $R(0) = 1$ to $R(0) \leq 1$ by introducing PIF, defined as $PIF = 1 - R(0)$, and showed that the remaining observed failures not accounted for by λ_D can be explained by values of PIF ranging from 1% - 1.6%. With this modification, PFDavg can be approximated as [3]

$$PFDavg \approx PIF + (1 - PIF) * \lambda_D * T_P /2. \quad (3)$$

Note that (3) reduces to (2) if PIF = 0, i.e., if $R(0) = 1$, so (3) includes (2) as a special case. In [3] it was also shown that even if λ_D is as large as 10^{-7}, PIF dominates PFDavg. This approach treats all failures not accounted for by λ_D as initial failures and provides a conservative (though somewhat controversial) approach to the assessment of PFDavg.

In [4, 5], using a new method of statistical analysis, Bukowski and Goble showed that more than one combination of λ_D and PIF were statistically consistent with the proof test data and that different combinations lead to different PFDavg assessments. They did not, however, explain why these different combinations existed and why proof test data supported more than one explanation for the underlying failure mechanisms responsible for the observed failures. Understanding this is important to being able to deduce which of the many possible combinations of λ_D and PIF best explains the observed proof test data. Lastly, they suggested that more complicated $\lambda_D(t)$ be considered when trying to explain the observed failures.

This report:

- argues that, logically, the calculation of PFDavg (from which SIL is assigned) must account for all dangerous failures regardless of underlying cause.;
- explains how $\lambda_D(t)$ together with PIF can model all dangerous failures regardless of underlying cause.;
- underscores why seeking to identify $\lambda_D(t)$ and PIF based on proof test data alone necessarily results in multiple $\lambda_D(t)$ and PIF combinations which provide equally good explanations for the observed proof test data;
- compares and contrasts two methods for the statistical analysis of proof test data;
- revisits two proof test data sets previously analyzed by QRA and re-analyzes the data using the new SSA method now including both constant and declining failure rates in $\lambda_D(t)$ as well as both zero and non-zero values of PIF;
- demonstrates that the different $\lambda_D(t)$ and PIF combinations resulting from SSA produce different PFDavg for the same proof test data;
- stresses the importance of appropriate data collection, analysis, and sharing of results to reduce costs while maintaining or improving safety.

2. *NOTATION*

CCPS	Center for Chemical Process Safety
CPI	chemical process industries
FMEDA	failure mode, effects, and diagnostic analysis
FITS	10^{-9} failures/hr
HPI	hydrocarbon process industries
PERD	Process Equipment Reliability Database
PFDavg	probability of failure on demand averaged over $[0, T_P]$
PFD(t)	probability of failure on demand as a function of time
PIF	probability of initial failure; 1 - R(0)
PRV	pressure relief valve(s)
QRA	quantal response analysis
R(0)	initial reliability; 1 - PIF; may be less than 1
R(t)	reliability as a function of time
SIL	safety integrity level(s)

SSA	statistical signature analysis
T_{IM}	end of the period of infant mortality failure rate
T_P	length of a particular PRV time in-service until proof test
λ_D	constant failure rate of the dangerous ("stuck-shut") failure mode
$\lambda_D(t)$	failure rate of the dangerous ("stuck-shut") failure mode as a function of time; may be a constant

3. *USING $\lambda_D(t)$ TO MODEL ALL DANGEROUS FAILURES*

Some safety analysts believe that it is necessary to include only certain failures resulting from some particular types of failure mechanisms in the computation of PFDavg. Often they use the PFDavg approximation given in (2) to justify including only failures attributable to the constant failure rate, λ_D. We believe it is self-evident that, from a safety standpoint, a dangerous failure is a dangerous failure regardless of cause and must be accounted for in PFDavg. After all, if a PRV does not respond to an overpressure event, the safety consequences exist regardless of why the dangerous failure occurred. Therefore, the question is not which dangerous failures to include in PFDavg but how to include them.

Consider the case where extensive *time-to-failure* data exists for a type of safety equipment. (This is not the case for the "stuck shut" mode in PRV but we will consider that complication subsequently.) Say we have 1000 equipment units and for each unit we know the entire state history over an interval of 0 to 5 years. Further assume that in the 5 year period exactly 7 units (0.7%) fail and the remainder are successful for the entire 5 year period. Figures 1 - 3 show three possible sets of time lines for 10 such units: 3 of the 993 successful units and all 7 of the failed units. The solid lines indicate time intervals when the unit was in a state of success and the F lines indicate time intervals when the unit was in a state of failure. If a unit is successful over the entire 5 year interval, its solid line ends in the letter S. The time-to-failure patterns are quite different in each of the figures. Figure 1 shows 5 of 7 failures clustered in the first year. Figure 2 shows the 7 failures spread over the entire 5 year interval. Figure 3 shows 6 of 7 failures clustered in the last year. Clearly, the three figures show different failure patterns that would be represented by three different types of $\lambda_D(t)$ which we could estimate given the time-to-failure data.

For example, Figure 1 would likely be consistent with a declining failure rate; Figure 2, with a constant failure rate; and Figure 3, with an increasing failure rate. These three examples are deliberately simple, each exhibiting a single failure rate characteristic. In general, time-to-failure data which included many units over an extended observation period might exhibit one, two, or all three types of failure rates. The presence of all three types of failure rates gives rise to the traditional "bath tub curve" model for $\lambda_D(t)$.

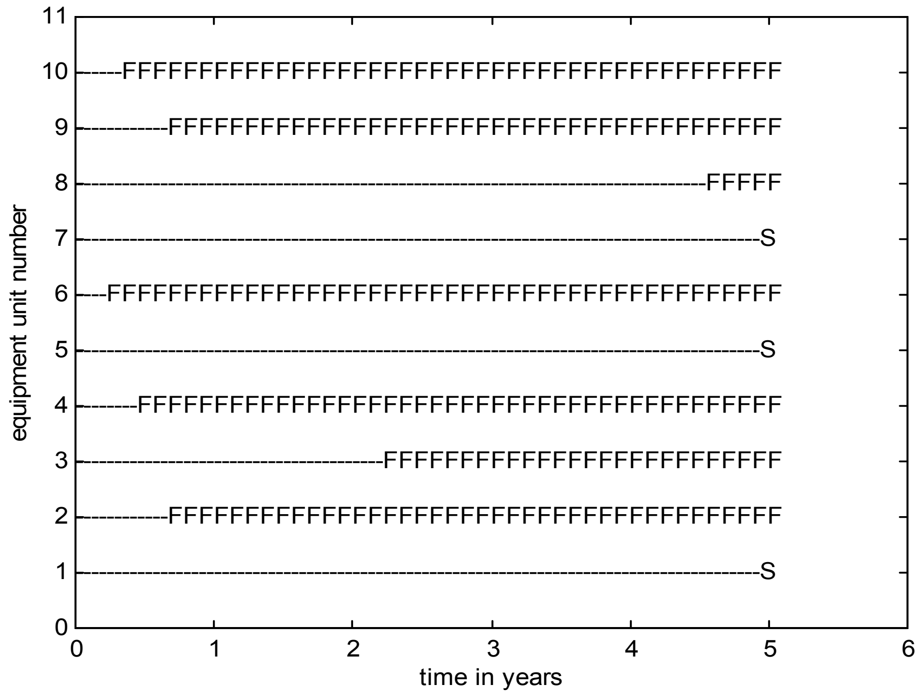

Figure 1. Example Time-to-Failures Clustered in First Year of Service

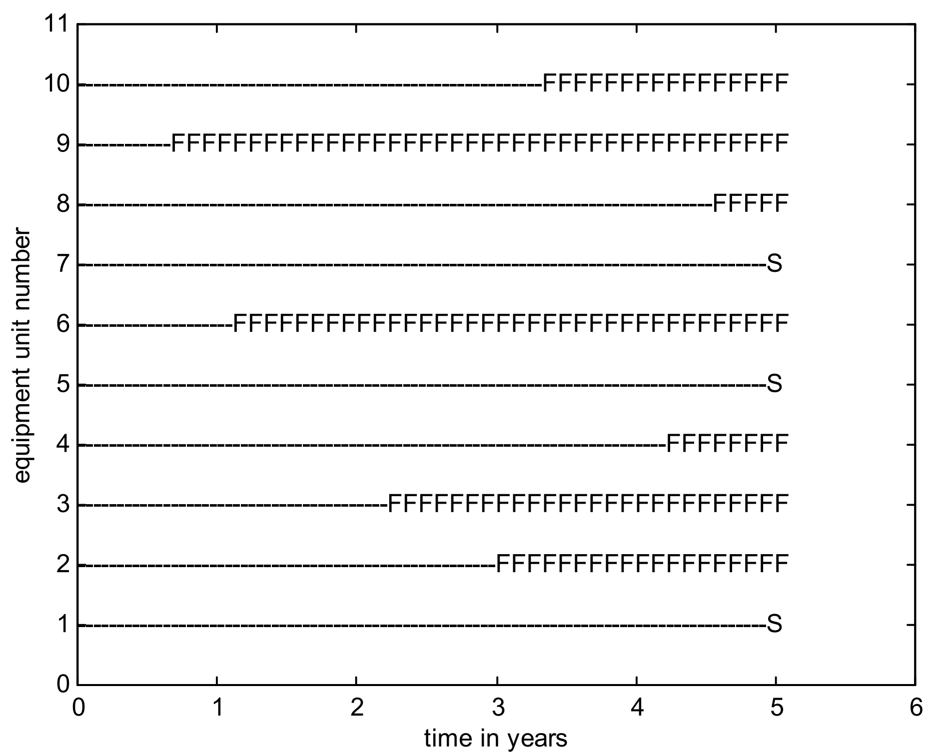

Figure 2. Example Time-to-Failures Spread Over 5 Years of Service

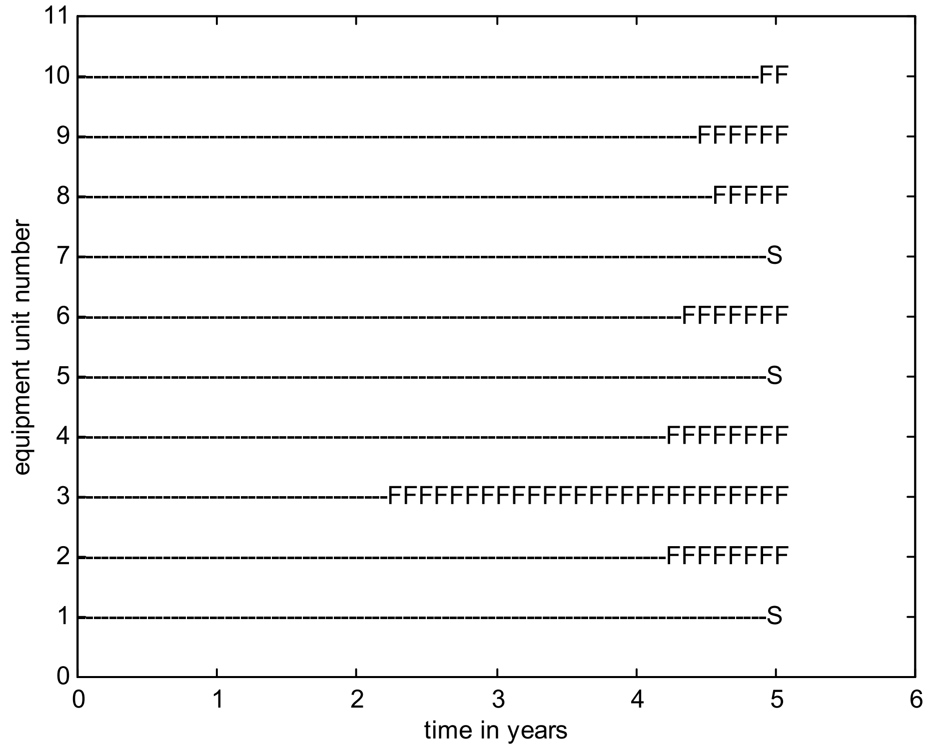

Figure 3. Example Time-to-Failures Clustered in Last Year of Service

Now consider the fact that the time-to-failure data captures all failures regardless of underlying cause. Therefore, $\lambda_D(t)$ estimated from the time-to-failure data along with PIF[2] is sufficient to model all types of failures. Given $\lambda_D(t)$ and PIF, and noting that $R(0) = 1 - PIF$, PFDavg can be computed from (1) above for any valid $\lambda_D(t)$[3].

4. *ESTIMATING $\lambda_D(t)$ FROM PROOF TEST DATA*

Because the PRV is normally closed, the "stuck-shut" failure mode cannot be detected while the PRV is installed within the process being protected. Identifying this failure mode requires a proof test. In a proof test, the PRV is removed from the process and is pressurized to determine the pressure at which the PRV opens. If the test pressure required to open the PRV the first time exceeds 150% of set pressure, the PRV is deemed to have "failed to open", i.e., the PRV is "stuck shut". This type of testing is costly and, while it identifies failed PRV, it does not provide an indication of *when* the failure occurred prior to the proof test.

[2] For mathematical completeness, we state for the record that PIF could easily be incorporated in $\lambda_D(t)$ as an impulse (Dirac delta function) at t=0. Here we treat PIF separately for simplicity of explanation and because, as a practical matter, it would be handled separately in numerical computation of PFDavg.

[3] Reliability theory imposes certain constraints on $\lambda_D(t)$ which must be met. Specifically, $\lambda_D(t)$ must be ≥ 0 for all t and ${}_0\int^{\infty} \lambda_D(s)\, ds \rightarrow \infty$.

Now consider the case where we do not have time-to-failure information but only proof test results. Re-examine Figures 1 - 3 and assume that we can only observe the state of each unit of equipment at t = 5 years. This is equivalent to conducting a proof test at T_P = 5 years and recording whether the unit successfully opened (S) or failed (F) to open. Note that the proof test results would be identical for all three distinctly different failure patterns. Thus, proof test data analysis, *based solely on the operating time accumulated prior to proof test and the outcome (S/F) of the test*, **cannot** identify a unique $\lambda_D(t)$ underlying the actual failure patterns since these patterns are not observable. However, subject to appropriate analysis, proof test data can produce a series of candidate $\lambda_D(t)$ and PIF combinations that are all statistically consistent with the observed proof test data. Each different candidate $\lambda_D(t)$ and PIF combination will potentially result in a different PFDavg. Determining which one of the candidate $\lambda_D(t)$ and PIF combinations best describes the true underlying failure pattern (and corresponding PFDavg) will require additional information and analysis techniques.

5. *PROOF TEST DATA ANALYSIS METHODS*

There are currently two analysis methods available for application to proof test data, Quantal Response Analysis (QRA) [2, 6] and Statistical Signature Analysis (SSA) [4, 5]. The discussion below compares and contrasts them.

5.1 *Quantal Response Analysis*

In QRA, the analyst assumes the form $\lambda_D(t)$ will take. For example, $\lambda_D(t) = \lambda_D$, a constant, $\lambda_D(t) = At^n + b$, $n > 0$, a growth function, etc. The proof test data is then sorted and grouped, the grouped data is transformed, curve fitting methods are applied to find a best fit for the integral of $\lambda_D(t)$, and, if a valid curve fit is found, the unknown parameters in the assumed form of $\lambda_D(t)$ are estimated.

QRA has a number of drawbacks. It requires experience to appropriately group the data. The results of the analysis can be very sensitive to the particulars of the data grouping and different groupings may lead to different fitted curves. Because $\lambda_D(t)$ must fulfill certain conditions, the fitted curves are also constrained. As a consequence, it is sometimes difficult or impossible to find a curve fit that also meets the required constraints. When no valid curve fit is discovered, it does not mean that one does not exist; merely that one has not been found. Conversely, if only one valid curve fit is discovered, it does not mean it is the only valid fit; merely the only one that has been found. Further, when no valid curve fit is discovered, QRA does not indicate what to do next to try to find a valid fit. QRA requires that the proof test times, T_P, have a reasonable amount of variation in time before curve fitting can occur. For example, if an operating company faithfully carried out proof testing every 2 years plus or minus a month, the proof test times would be too tightly clustered to allow for effective curve fitting. Lastly, if new proof test data becomes available and is added to the data set, it is likely the data must be completely re-analyzed to take into account the new data.

Although in theory, QRA could handle any assumed valid form for $\lambda_D(t)$, as a practical matter, complex $\lambda_D(t)$ would lead to the need for very complicated curve fitting that would likely not be successfully achieved. To date, when applying QRA to proof test data, only two forms of $\lambda_D(t)$ have been considered individually: constant failure rate and failure rate growth. In particular, no attempt has been made to assume an infant mortality form for $\lambda_D(t)$ or to include more than one type of failure rate function in $\lambda_D(t)$, e.g., to include both a declining failure rate on one time interval and a constant failure rate over another time interval.

5.2 *Statistical Signature Analysis*

SSA takes a different philosophical approach than QRA to the task of analyzing proof test data. It is simply a "guess and check" method of exploring the proof test data. It allows the analyst to assume not only the form of $\lambda_D(t)$ but also to assume values for all of the parameters defining $\lambda_D(t)$ and to include either a zero or non-zero value for PIF. It then compares the number of failures actually observed in the proof test data to the number of failures that would likely be seen if the assumed $\lambda_D(t)$ and PIF were correct. Based on this comparison, a statistical signature is produced. The signature is valid if it fulfills certain properties in which case the assumed parameter values from $\lambda_D(t)$ and PIF are candidates to explain the observed data. If the signature is not valid, the assumed parameter values are not candidates to explain the observed data. However, the properties of the signature that make it invalid also indicate broadly what parameter values to try next in the search for a valid signature.

Unlike QRA, SSA does not require any data sorting. Even if the proof test times are highly clustered, SSA can process the data. The proof test data can be fed into the algorithm as it is collected and one point in the signature is computed and plotted for each proof test based on the results of the latest test and all previous tests in the data set. New data is simply added to the running analysis. There is no limitation on the complexity of $\lambda_D(t)$ that may be assumed and $\lambda_D(t)$ may be described mathematically piecewise over several intervals. Thus quite complex $\lambda_D(t)$ can be constructed from a series of simpler parts.

6. *APPLYING SSA TO PREVIOUSLY ANALYZED PRV PROOF TEST DATA*

6.1 *Proof Test Data Analyzed*

In [2, 3], three proof test data sets which met the data quality intent of the CCPS PERD Initiative[4] were analyzed using QRA to estimate the PRV constant failure rate,

[4] See Acknowledgement for contact information for and additional description of the CCPS PERD Initiative.

λ_D, and PIF. We now wish to revisit those data sets and apply SSA. Unfortunately, Data Set II from the original study is not retrievable in the form required for SSA. Thus, this paper is limited to re-analyzing Data Set I and used-valve[5] proof test data from a recently expanded Data Set III.

We have also gathered some additional general information about Data Sets I and III. Data Set I consists of PRV proof tests from a mix of clean and foul service, weighted toward clean service. These proof tests were primarily bench tests that included both in-house and out-sourced testing. The expanded Data Set III consists of PRV proof tests almost exclusively from clean service. All of the tests were in-house bench tests.

6.2 *Assumed Form of $\lambda_D(t)$*

SSA is capable of incorporating $\lambda_D(t)$ of any complexity and this report documents the application of SSA using a more complex form of $\lambda_D(t)$ than has previously been studied. Although we originally intended to include all characteristics of the traditional "bath tub curve" in our $\lambda_D(t)$, a closer examination of the two data sets to be utilized in the analysis showed that neither set contained any evidence of failure rate growth indicating end of life failures. Consequently, we limited the characteristics of $\lambda_D(t)$ in this study to a declining failure rate (modeling infant mortality over the time interval $[0, T_{IM}]$) in combination with a constant failure rate and either zero or non-zero PIF. Two different declining failure rates were considered, exponential and linear, to represent steep or shallow rates of decline. Clearly, other variations are possible. The two different forms of $\lambda_D(t)$ studied are illustrated in Figures 4 and 5 and their formulas with relevant parameters are given in Equations (4) and (5). For exponentially declining infant mortality

$$\begin{aligned} \lambda_D(t) &= c\,e^{-kt} && \text{for} \quad 0 \le t \le T_{IM} \\ &= \lambda_D && \text{for} \quad T_{IM} \le t \le \infty \end{aligned} \tag{4}$$

For linearly declining infant mortality

$$\begin{aligned} \lambda_D(t) &= d + m\,t && \text{for} \quad 0 \le t \le T_{IM} \\ &= \lambda_D && \text{for} \quad T_{IM} \le t \le \infty \end{aligned} \tag{5}$$

It appears that in either case, when an analyst applies SSA to the proof test data s/he must provide the values for 5 parameters: PIF, λ_D, T_{IM}, and c and k, or d and m. However, this is reduced to 4 values because of the requirement of continuity at the time value T_{IM}.

[5] Data Set III contained test data from both new and used valves. We considered only the used-valve data as the new valves had no operating time on them.

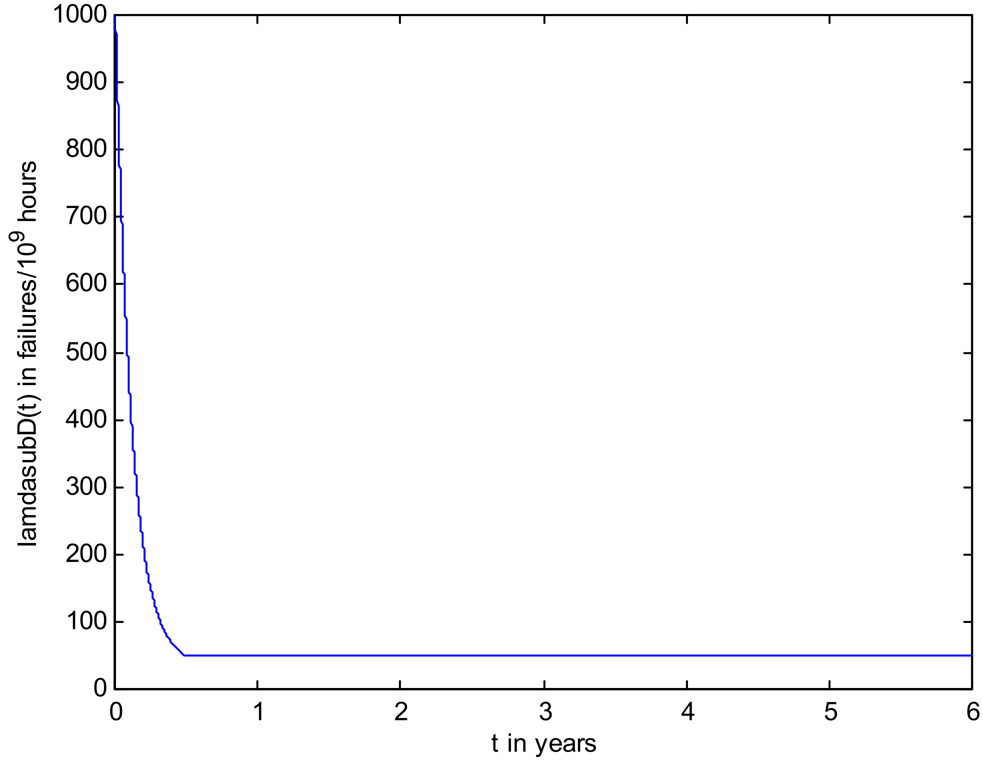

Figure 4. Exponentially Declining Infant Mortality and Constant Failure Rate

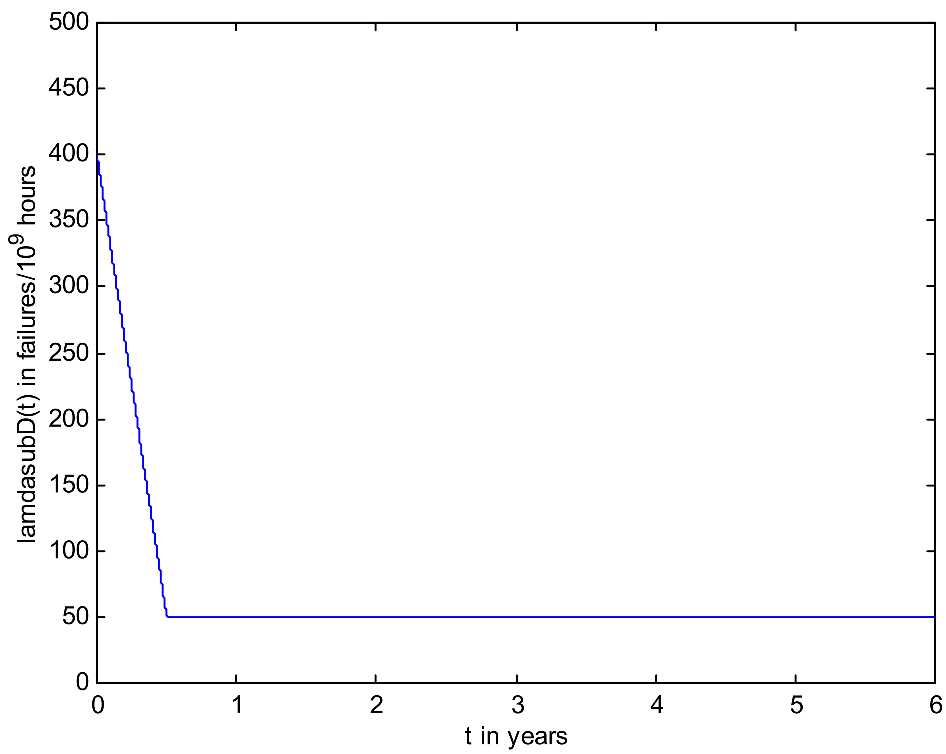

Figure 5. Linearly Declining Infant Mortality and Constant Failure Rate

From (4)

$$\lambda_D(T_{IM}) = c\, e^{-kTIM} = \lambda_D \tag{6}$$

which implies that

$$k = -(1/T_{IM})\ \ln(\lambda_D / c) \tag{7}$$

and from (5),

$$\lambda_D(T_{IM}) = d + m\, T_{IM} = \lambda_D \tag{8}$$

which implies that

$$m = (\lambda_D - d)/T_{IM}. \tag{9}$$

6.3 *Choosing Parameter Values*

Equipment failures due to the constant failure rate, λ_D, are usually considered to be caused by environmental stresses that exceed the average strength of the equipment during its normal lifetime. As such, we would expect that λ_D would be ever-present. Thus, in this study, λ_D always has a non-zero value. However, all other unconstrained parameters may assume either zero or non-zero values. In order to limit the number of parameters to be varied, we decided to fix T_{IM}, the end of the period of infant mortality failures at 0.5 years for this study. Note that only one infant mortality model can be considered at a time. Consequently, c and d cannot simultaneously assume non-zero values. Table 1 shows the combinations of parameter values considered.

Table 1. Combinations of Parameter Values Considered

λ_D	PIF	c	d
non-zero	0	0	0
non-zero	non-zero	0	0
non-zero	0	non-zero	0
non-zero	0	0	non-zero
non-zero	non-zero	non-zero	0
non-zero	non-zero	0	non-zero

6.4 *Methodology and Results*

To conduct this study we began with Data Set I and sought a value of λ_D that alone could account for all the failures observed in the proof test data, i.e., SSA was run with all parameter values other than λ_D set equal to zero. Figure 6 shows the statistical signature produced with λ_D = 514 FITS. We then selected other smaller values for λ_D and, by trial and error, found various combinations of other parameter values that resulted in statistical signatures similar to the one in Figure 6. Since these combinations of values produced similar signatures, they were deemed to be equally good explanations for the observed data.

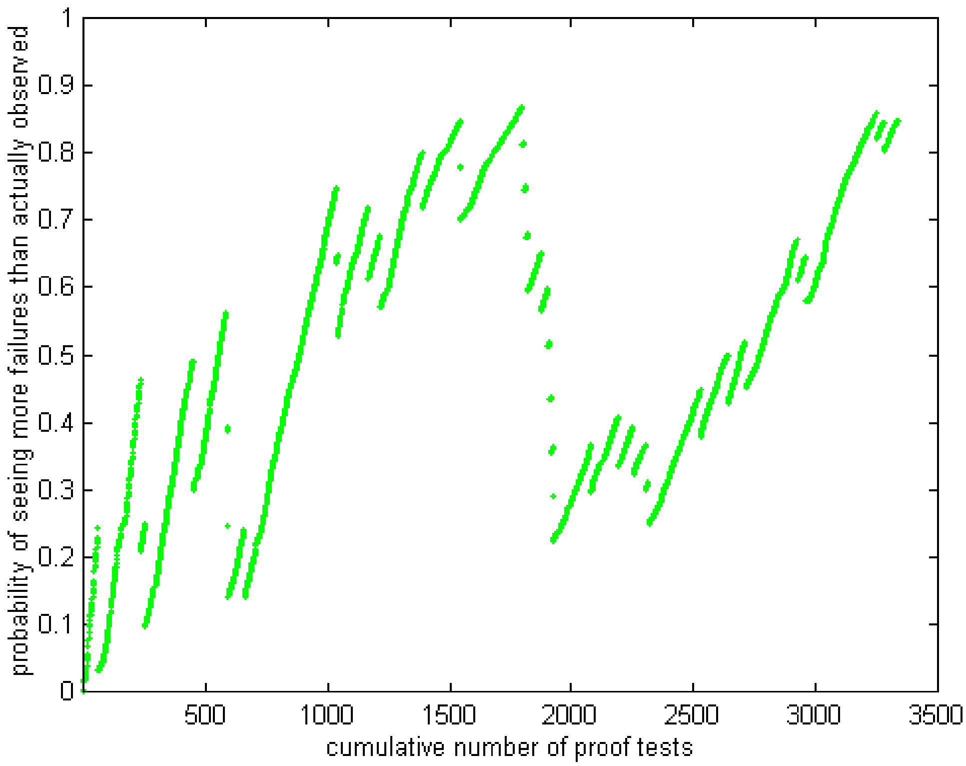

Figure 6. Statistical Signature: λ_D = 514 FITS and All Other Parameters Zero

We studied a wide variety of λ_D values including 400, 300, 200, 100, 75, 50, and 25 FITS producing far too much data to include in this report. We illustrate our findings for λ_D = 50 FITS, chosen because it is the midrange value for the generally accepted range of λ_D and this range is supported both by previous QRA and by FMEDA analyses [7]. Table 2 shows a sampling of combinations of parameter values. Note that these are only a sampling and many other combinations were discovered for any fixed value of λ_D. Figure 7 shows the signatures from all combinations in Table 2 overlaid, with the signature for Case A plotted in green and the signatures for Cases B - F plotted in blue. Note that the signatures for Cases B - F are virtually indistinguishable from one another and are very similar to the signature for Case A. (If Figure 7 is viewed in black and white, Case A is the narrower, lighter line that deviates from the thicker, darker line.)

Table 2. Sampling of Parameter Values Offering Equally Good Explanations of the Proof Test Data in Data Set I

Case	λ_D FITS	PIF	c FITS	d FITS
A	514	0	0	0
B	50	0.010065	0	0
C	50	0	13,500	0
D	50	0	0	4795
E	50	0.002	10,400	0
F	50	0.002	0	3860

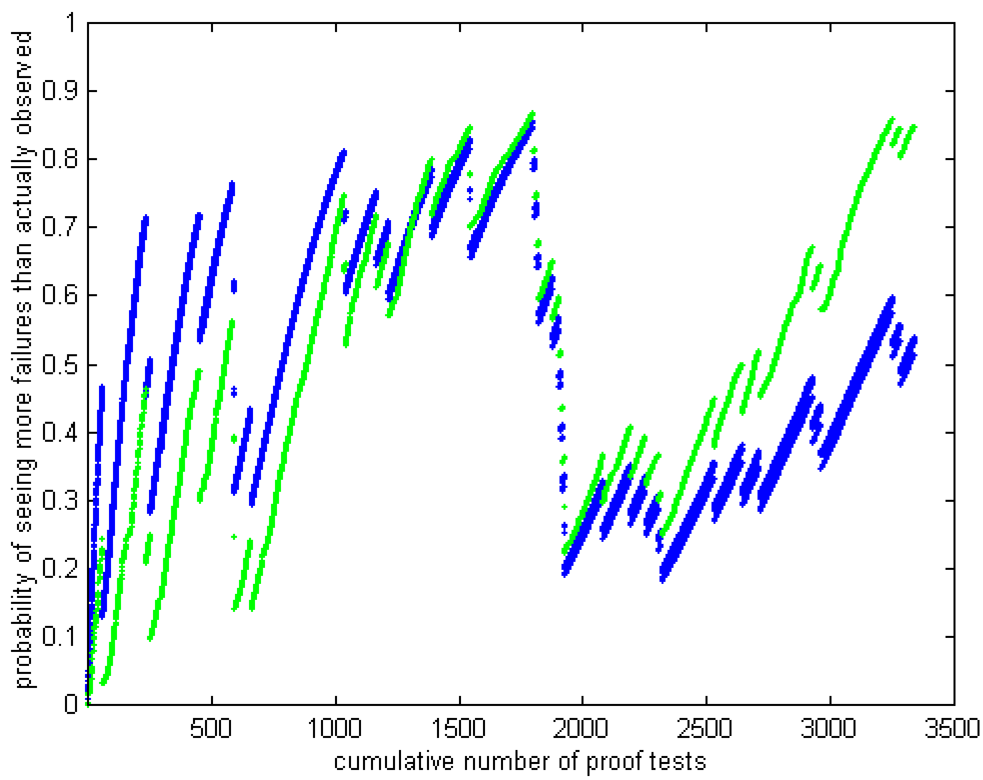

Figure 7. Overlay of All Statistical Signatures for Combinations in Table 2

We preceded similarly with the used-valve test data from the extended Data Set III. Table 3 shows a sampling of the combinations that produced essentially the same signature and Figure 8 shows all of the signatures overlaid. Note that in this case, the signature with λ_D = 500 and other parameters set to zero (Case G) is shown in green while the other cases from Table 3 are shown in blue. (If Figure 8 is viewed in black and white, Case G is the lighter line which deviates most from the other darker lines.)

Table 3. Sampling of Parameter Values Offering Equally Good Explanations of the Proof Test Data in Data Set III

Case	λ_D FITS	PIF	c FITS	d FITS
G	500	0	0	0
H	50	0.01289	0	0
I	50	0	19,075	0
J	50	0	0	6,450
K	50	0.005	10,720	0
L	50	0.005	0	3,975

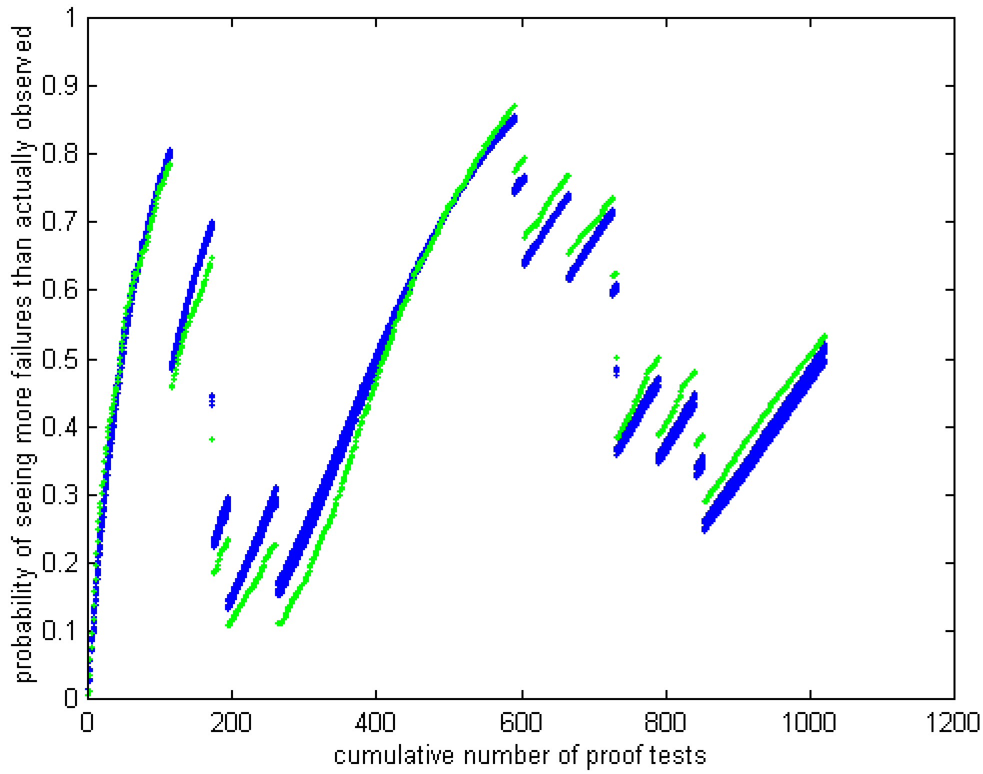

Figure 8. Overlay of All Statistical Signatures for Combinations in Table 3

7. *CALCULATING PFDavg WITH MORE COMPLEX $\lambda_D(t)$*

PFDavg is correctly defined in (1) for any valid $\lambda_D(t)$. However, with more complex $\lambda_D(t)$, the actual calculation of PFDavg for any value of T_P becomes more complicated and may require numerical integration of all or part of the integral. The exact formulas for PFDavg are derived below for the cases of $\lambda_D(t)$ given in (4) and (5) and are the means by which we computed PFDavg for Figures 9 and 10 below.

7.1 *PFDavg with Exponentially Declining Infant Mortality (Equation (4))*

Recall that PFDavg is given by

$$\text{PFDavg} = (1/T_P)\int_0^{T_P} \text{PFD}(t)\,dt = (1/T_P)\int_0^{T_P} 1 - R(t)\,dt$$

$$= 1 - (1/T_P)\int_0^{T_P} R(t)\,dt \qquad 0 \le T_P < \infty. \qquad (10)$$

Now, for the case of exponentially declining infant mortality, R(t) is given by

$$R(t) = (1 - \text{PIF}) \exp\left[-\int_0^{t} \lambda_D(s)\,ds\right] \qquad 0 \le t < \infty \qquad (11)$$

$$= (1 - \text{PIF}) \exp\left[-\int_0^{t} c\, e^{-ks}\,ds\right] \qquad 0 \le t < T_{IM} \qquad (12)$$

$$= (1 - \text{PIF}) \exp\left[-\int_0^{T_{IM}} c\, e^{-ks}\,ds - \int_{T_{IM}}^{t} \lambda_D\,ds\right] \qquad T_{IM} \le t < \infty \qquad (13)$$

and

$$(1/T_P)\int_0^{T_P} R(t)\,dt = (1/T_P)\int_0^{T_P} (1 - \text{PIF}) \exp\left[-(c/k)(1 - e^{-kt})\right] dt \qquad 0 < T_P \le T_{IM} \qquad (14)$$

$$= (1/T_P)\int_0^{T_P} \{(1 - \text{PIF}) \exp[-(c/k)(1 - e^{-kTIM})]\}\ \exp\left[-\int_{T_{IM}}^{t} \lambda_D\,ds\right] dt \qquad T_{IM} \le T_P < \infty. \qquad (15)$$

Note that (14) must be integrated numerically and that, in (15), the term enclosed by { } is simply a constant. Thus, (15) reduces to

$$(1/T_P)\int_0^{T_P} R(t)\,dt = (1/T_P)\{\ \}\int_0^{T_P} \exp[-\lambda_D (t - T_{IM})]\,dt \qquad T_{IM} < T_P < \infty \qquad (16)$$

$$= (1/T_P)\{\ \}(1/\lambda_D) \exp[+\lambda_D T_{IM}]\,[e^{-\lambda DTIM} - e^{-\lambda DTP}] \qquad T_{IM} < T_P < \infty. \qquad (17)$$

Depending on the value of T_P, we substitute either (14) or (17) into (10) to evaluate PFDavg for the case of exponentially declining infant mortality.

7.2 *PFDavg with Linearly Declining Infant Mortality (Equation (5))*

The derivation for this case follows the derivation above except that in (12) and (13) we substitute the model for linearly declining infant mortality. Thus, in this case, R(t) is given by

$$R(t) = (1 - PIF) \exp\left[- \int_0^t c + ms \, ds \right] \qquad 0 \leq t < T_{IM} \qquad (18)$$

$$= (1 - PIF) \exp\left[- \int_0^{T_{IM}} c + ms \, ds - \int_{T_{IM}}^t \lambda_D \, ds \right] \qquad T_{IM} \leq t < \infty \qquad (19)$$

Again, as above,

$$(1/T_P) \int_0^{T_P} R(t)\, dt = (1/T_P) \int_0^{T_P} (1 - PIF) \exp[-c\,t - 0.5\, m\, t^2]\, dt \qquad 0 < T_P \leq T_{IM} \qquad (20)$$

$$= (1/T_P) \int_0^{T_P} \{(1 - PIF) \exp[-cT_{IM} - 0.5m\, T_{IM}^2]\} \exp[- \int_{T_{IM}}^{t} \lambda_D \, ds]\, dt \qquad T_{IM} \leq T_P < \infty. \qquad (21)$$

Again, note that (20) must be integrated numerically and that, in (21), the term enclosed by { } is simply a constant. Thus, (21) reduces to

$$(1/T_P) \int_0^{T_P} R(t)\, dt = (1/T_P) \{\ \} \int_0^{T_P} \exp[-\lambda_D (t - T_{IM})]\, dt \qquad T_{IM} < T_P < \infty \qquad (22)$$

$$= (1/T_P) \{\ \} (1/\lambda_D) \exp[+ \lambda_D T_{IM}] [e^{-\lambda DTIM} - e^{-\lambda DTP}] \qquad T_{IM} < T_P < \infty. \qquad (23)$$

Depending on the value of T_P, we substitute either (20) or (23) into (10) to evaluate PFDavg for the case of linearly declining infant mortality.

8. *RESULTS FOR PFDavg*

Given that each of the combinations of parameter values in Table 2 or in Table 3 produces essentially the same statistical signatures, all are equally good explanations for their respective failure data sets. But each combination of parameter values potentially produces a different assessment of PFDavg. Figures 9 and 10 show plots of PFDavg vs. T_P for the parameter combinations given in Tables 2 and 3, respectively.

In Figure 9, note that for Cases B – F, which all have λ_D = 50 FITS, the PFDavg plots are discernibly different for values of T_P < 1 year, but show little difference for T_P

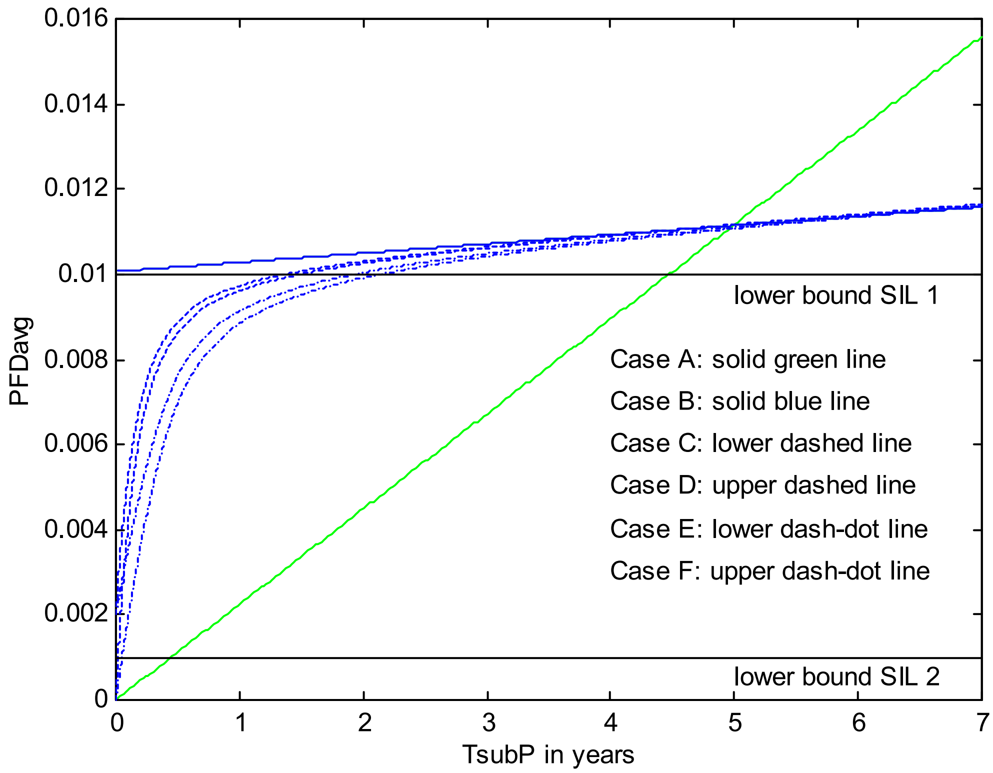

Figure 9. Plots of PFDavg for All Parameter Combinations in Table 2

between 2 and 4 years. Once T_P > 4 years, the PFDavg computed for these cases are approximately equal. For T_P < 1 year, Cases C – F support SIL 2 behavior. For T_P > 4.5 years, the data support SIL 1 behavior only regardless of the model used to explain the observed failures in Data Set I.

Figure 10 shows that Cases I – L support SIL 2 behavior only for values of T_P < 6 months. For T_P > 6 months, the data support SIL 1 behavior only regardless of the model used to explain the observed failures in used-valve data from extended Data Set III.

9. *DISCUSSION*

Clearly, the SSA method opens up many new possibilities for modeling equipment failure rates from proof test data. Its current shortcoming is that, alone, it cannot tell us which one of the many models, identified as being statistically consistent with the data, actually describes the true underlying failure mechanisms generating the observed proof test data. It is likely that other independent means to estimate or predict some of the parameters in the $\lambda_D(t)$ model will narrow the possibilities that need to be considered. For example, by using FMEDA techniques to predict the constant portion of the failure rate, the remaining parameter ranges may be significantly reduced. Investing in the development of additional analysis techniques is worthwhile as these techniques will likely produce useful models that allow designers to better optimize safety system designs. Finally, finding ways to identify the $\lambda_D(t)$ which most accurately explains the underlying failure mechanisms is also important when we seek to reduce the occurrences of dangerous failures.

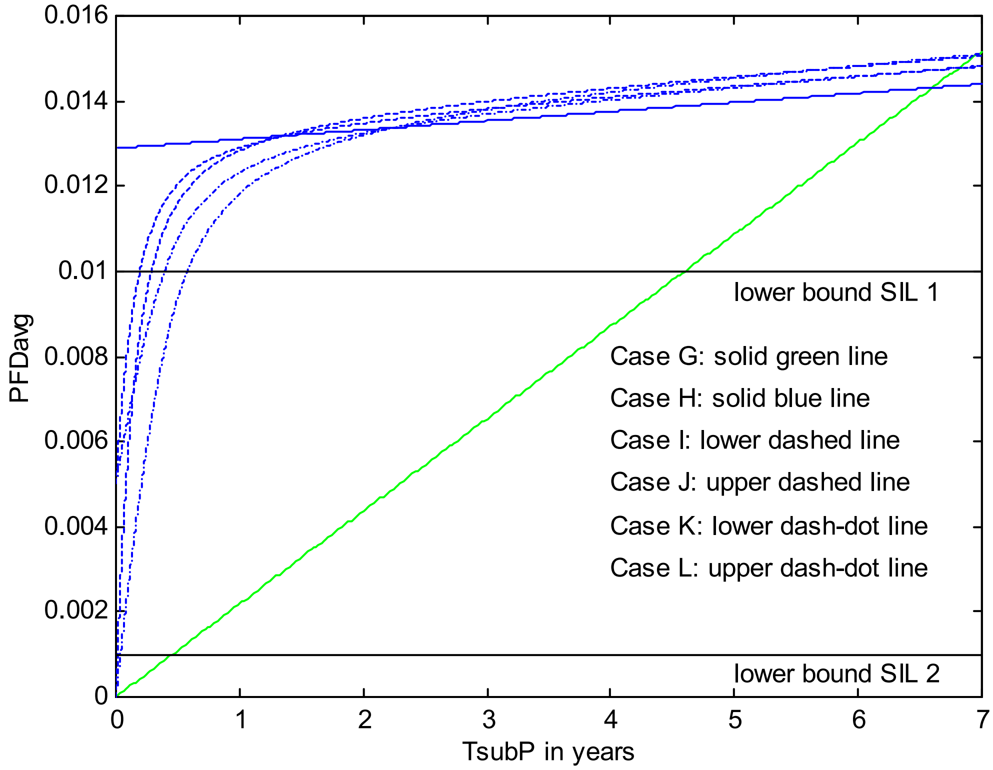

Figure 10. Plots of PFDavg for All Parameter Combinations in Table 3

The results reported here apply to the particular data sets analyzed. We do not know what results or conclusions might be observed if additional data sets become available for analysis. But it is critical that *appropriate* data be collected *and* analyzed and the results shared. There is significant potential to reduce costs while maintaining safety levels or even reduce costs while improving safety if we can accurately understand and identify the underlying sources of dangerous failure.

REFERENCES

1. "TUV Certificate" (2006), Report No. 432/21205342, 30.01.2006, TUV Rheinland Group.

2. Bukowski, J. V. (2007), "Results of Statistical Analysis of Pressure Relief Valve Proof Test Data Designed to Validate a Mechanical Parts Failure Database," Technical Report, September, exida, Sellersville, PA.

3. Bukowski, J. V., and Goble, W. M. (2009), "Analysis of Pressure Relief Valve Proof Test Data," *AIChE Journal Process Safety Progress*, March 2009.

4. Bukowski, J. V. (2008), "A New Method for Analyzing Proof Test Data," Technical Report, September, exida, Sellersville, PA.

5. Bukowski, J. V. and Goble, W. M. (2009), "Unexpected Results from the Analysis of PERD Proof Test data & the Implications for Pressure Relief Valve Safety," *Proceedings of the 64th Annual Instrumentation Symposium for the Process Industries*, January, College Station, TX.

6. Sheesley, J. H., Thomas, H. W. and Valenzuela, C. A. (1995), "Quantal Response Analysis of Relief Valve Test Data," *ASQC 49th Annual Quality Congress Proceedings*, Cincinnati, OH, pp. 741-748.

7. Bukowski, J. V. and Goble, W. M. (2009), "Validation of a Mechanical Component Constant Failure Rate Database," *Proceedings Annual Reliability and Maintainability Symposium*, January, Fort Worth, TX.

ACKNOWLEDGEMENT

The author wishes to thank the CCPS PERD Initiative for providing the PRV proof test data used to test the statistical signature analysis method described this report. This initiative is comprised of operating companies, consultants, equipment manufacturers, and engineering companies committed to developing the technical guidance and tools necessary to capture on-going performance data, enabling operation of an equipment reliability database. The objective of PERD is to make high quality, valid, and useful data available to participating HPI and CPI companies to support availability, reliability, and equipment design improvements, maintenance strategies, and life cycle cost determination. For more information, contact Dave Belonger, CCPS PERD staff consultant at dbelonger@verizon.net or Hal Thomas, PERD Chairman, at thomashw@airproducts.com.