SAND2007-6555P

Joint Modeling of Degradation and Failure Rate Occurrence Using Bayesian Networks

Open PSA Workshop

October 3rd, 2007

Bobby Middleton

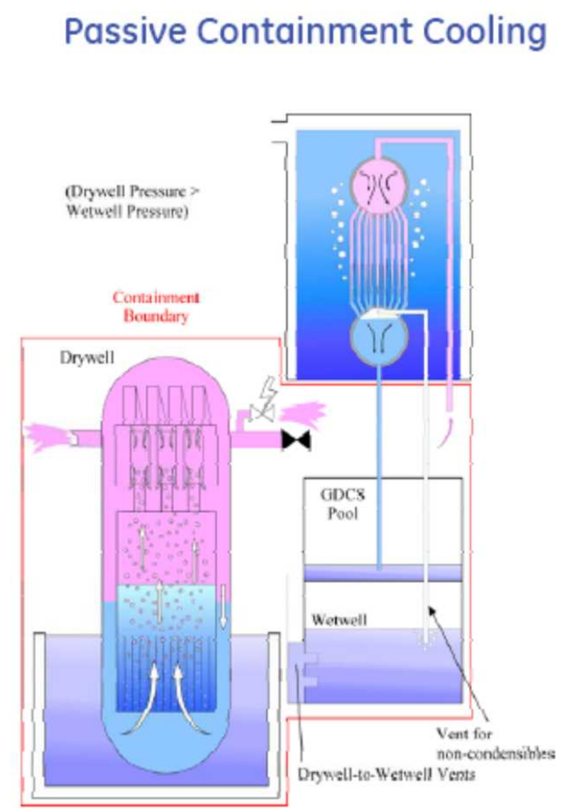
Sandia National Laboratories

Outline

- Motivation
- Technologies involved
  - Bayesian network background
  - MCMC sampling
- The process
  - Using available data
  - The model
  - The results
- Path forward

Motivation for Research

- Gen III/III+ Nuclear Reactors
  - Incorporate passive system designs
  - As many as 30+ COLs planned
  - At least one COL submitted (South Texas Project)
- Next Generation Nuclear Reactors
  - Will rely heavily on passive systems
  - No consensus methodology pertaining to PRA for passive systems



Motivation for Using Bayesian Networks

- Technique inherently uses all available information
  - Physical models
  - Expert judgment
  - Data
- Technique inherently produces results that quantify uncertainties
  - Accounts for measurement uncertainties
  - Accounts for model uncertainties
  - Accounts for variability among "individuals" in a population
- Allows hierarchical structure to account for different levels of model "importance"

Bayesian Networks

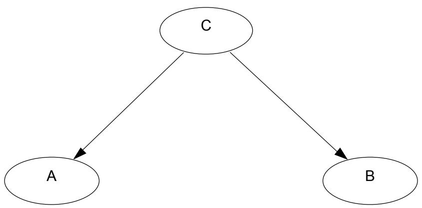
- **Based on Bayes' Rule**

$$P(B|C) = \frac{P(C|B) \times P(B)}{P(C)}$$

$$P''(\theta) = \frac{P'(X|\theta) \times P'(\theta)}{\int_{\text{all}\,\theta} P'(X|\theta) \times P'(\theta)\, d\theta}$$

- **Utilizes concept of conditional independence**

$$P(A \cap B|C) = P(A|C) \times P(B|C)$$



MC Sampling

- **Suppose we want to evaluate the integral of h(x)dx.**
- **Choose a probability distribution, w(x). Then:**

$$I \equiv \int h(x)dx = \int \frac{h(x)}{\omega(x)} \omega(x)dx$$

$$I \approx I_N \equiv \frac{1}{N} \sum_{t=0}^{N-1} \frac{h(x_t)}{\omega(x_t)}$$

Sandia National Laboratories

Pseudo-random Sampling

- **Pseudo-Monte Carlo**
  - **developed in nuclear weapons programs in the 1940's**
  - **let $I^s = [0,1]^s$ be a s -dimensional cube and let $f(t)$ be defined on I^s**
  - **let $(x_1,\ldots,x_N)$ be a *pseudo*-random sample of N points from I^s where**

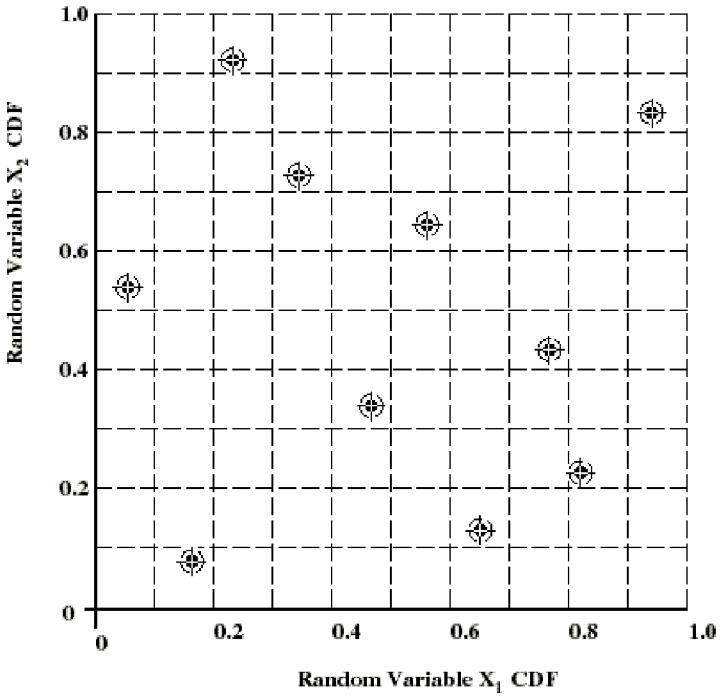
$$x_n = a x_{n-1} \bmod(m)$$

$$e.g.,\ x_n = 16807 x_{n-1} \bmod(2^{31} - 1)$$

  - **x_i/m is a pseudo-random number on the interval [0,1]**
  - PROS**:**
    - **sampling can be conducted sequentially (easy to add new samples)**
    - **error bounds not dependent on dimension *s***
  - CONS**:**
    - **<u>Probabilistic error bounds</u> depend on equidistribution of sample points in sample space $O(n^{-1/2})$**
    - **no methodical means of constructing sample to achieve error bound, therefore rate of convergence is very slow**

Latin Hypercube Sampling

- **Latin Hypercube Sampling**
- **also based on pseudo-random sampling**
- **form of stratified sampling in which the samples are 'forced' to be dispersed across the support space**
- **number of samples dictates the number of regions**
- PROS:
  - **significant reduction in number of samples compared to traditional MC**
- CONS:
  - **samples do not provide good uniformity across**
  - **samples can not be generated sequentially**



Quasi-random Monte Carlo (MCMC)

- A Quasi-random sample is commonly referred to as a low-discrepancy sequence.
- Low discrepancy sequence is one that places sample points nearly uniformly in the sample space of interest.
- Low-discrepancy $\rightarrow$ low integration error
- Deterministic error bounds –$O(N^{-1}(\log N))$
- Variety of sequences
  - Halton (simple, leaped, RR2)
  - Hammersley
  - Fauer
  - Sobol

Sandia National Laboratories

MCMC Applied to Bayes'

- **Suppose we have a model defined by:**

$$Y \sim N(\mu, \tau)$$

$$\mu = ax + bt + \varepsilon$$

$$\tau \sim G(\alpha, \beta)$$

$$a \sim N(\mu_a, \tau_a)$$

$$b \sim N(\mu_b, \tau_b)$$

$$\varepsilon \sim N(0, \tau_\varepsilon)$$

Gibbs Sampler (MCMC Sampler)

$$P\left(a^{t1} \middle| b^{t0}, \varepsilon^{t0}, \tau^{t0}, y\right) = \frac{P\left(b^{t0}, \varepsilon^{t0}, \tau^{t0}, y \mid a^{t0}\right) P\left(a^{t0}\right)}{\int P\left(b^{t0}, \varepsilon^{t0}, \tau^{t0}, y \mid a^{t0}\right) P\left(a^{t0}\right) da^{t0}}$$

$$P\left(b^{t1} \middle| a^{t1}, \varepsilon^{t0}, \tau^{t0}, y\right) = \frac{P\left(a^{t1}, \varepsilon^{t0}, \tau^{t0}, y \mid b^{t0}\right) P\left(b^{t0}\right)}{\int P\left(a^{t1}, \varepsilon^{t0}, \tau^{t0}, y \mid b^{t0}\right) P\left(b^{t0}\right) db^{t0}}$$

.

.

.

$$P\left(\tau^{t1} \middle| a^{t1}, b^{t1}, \varepsilon^{t1}, y\right) = \frac{P\left(a^{t1}, b^{t1}, \varepsilon^{t1}, y \mid \tau^{t0}\right) P\left(\tau^{t0}\right)}{\int P\left(a^{t1}, b^{t1}, \varepsilon^{t1}, y \mid \tau^{t0}\right) P\left(\tau^{t0}\right) d\tau^{t0}}$$

Sandia National Laboratories

Our Problem

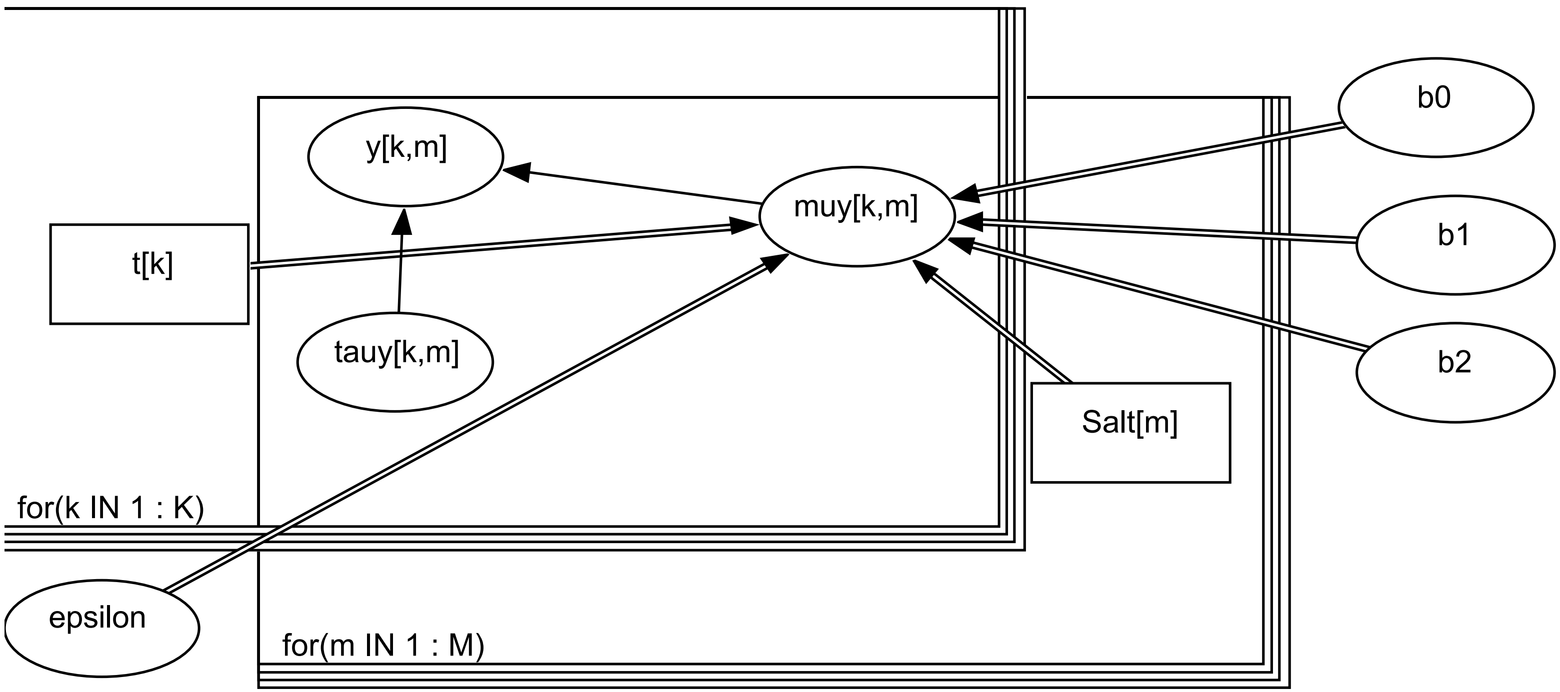
- Used available data…
- Multiple resistors placed in various environments:
  - Temperature
  - Salt content
  - Humidity
- Measurements of resistance recorded over time
- Failure time recorded
- Want model to:
  - Predict degradation state
  - Predict probability of failure at time t1 given no failure at time t0

Sandia National Laboratories

General Approach for Degradation

- Due to time constraints, limit model to time (t_k) and single time-independent covariate (Salt content)
- Assume measurements are Gaussian distributed with a mean equivalent to the “true” value and measurement error determined by a precision that is Gamma distributed
- Assumed “true” value is linear in time and salt content with model noise that is normally distributed
- Assumed coefficients are normally distributed

DAC for Degradation Model



General Approach for Failure Rate

- Assume failure is Bernoulli distributed
- Define:
  - d_{km}~Bern(PI(0,t))
  - d_{km}=0 if mth resistor is working at time t_k
  - d_{km}=1 if mth resistor is not working at time t_k, but was working at time t_{k-1}
  - d_{km}=NA otherwise
- Assume proportional hazards model of failure

$$PI(t) \equiv 1 - \int_0^t e^{-\lambda u} du$$

Failure Rate Model, cont.

- **Assume failure rate of mth component is equal to a "population" failure rate multiplied by a factor that is specific to the mth component**

$$\lambda_{km} = \lambda_{k0} e^{a_0 + a_1 * muy_{km} + a2 * Salt_m}$$

- **Define**

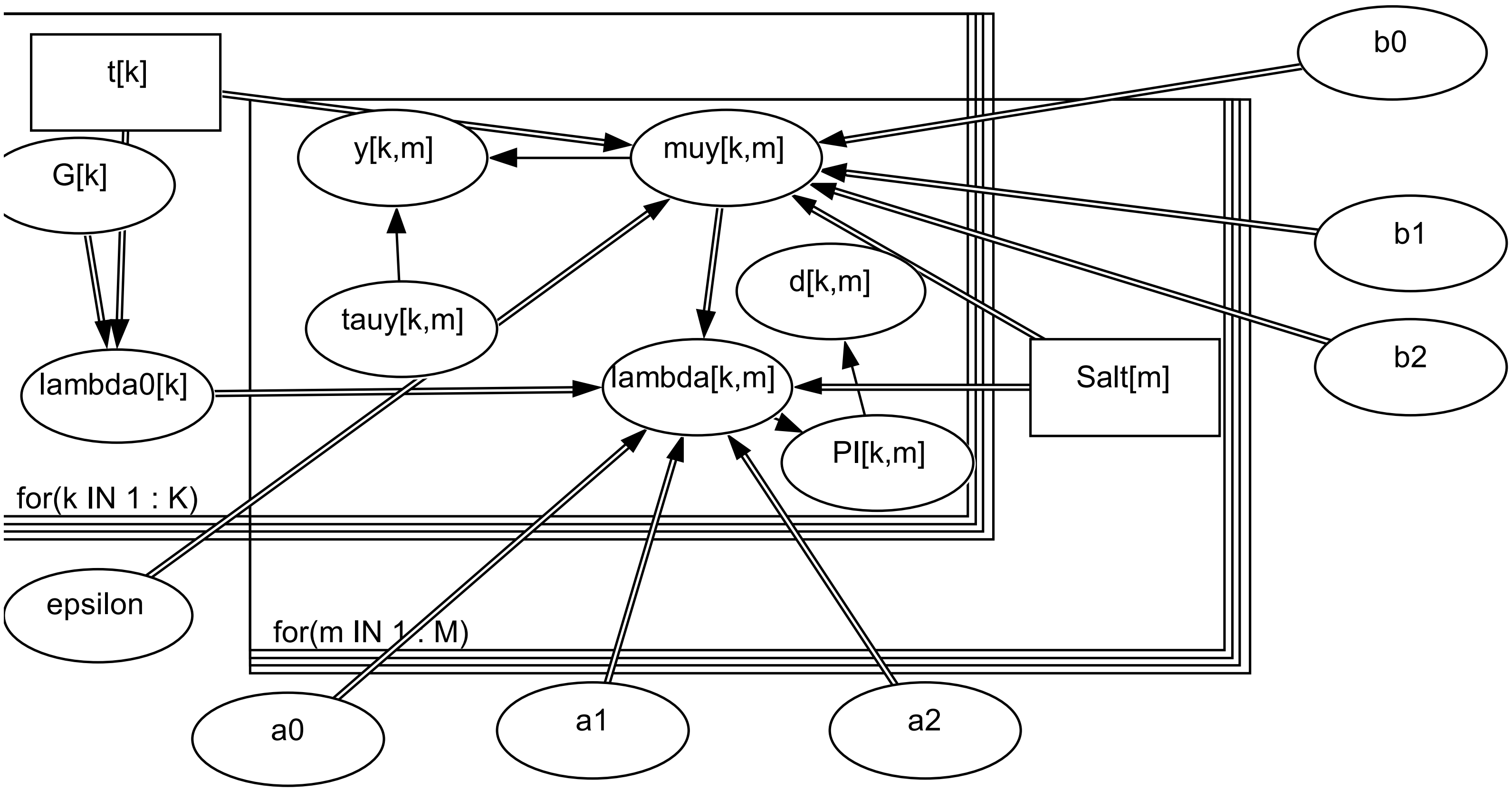
$$G = \int_0^t \lambda_0(u) du$$

$$\frac{dG}{dt} \sim Gamma(\alpha, \beta)$$

Technique Wrap-up

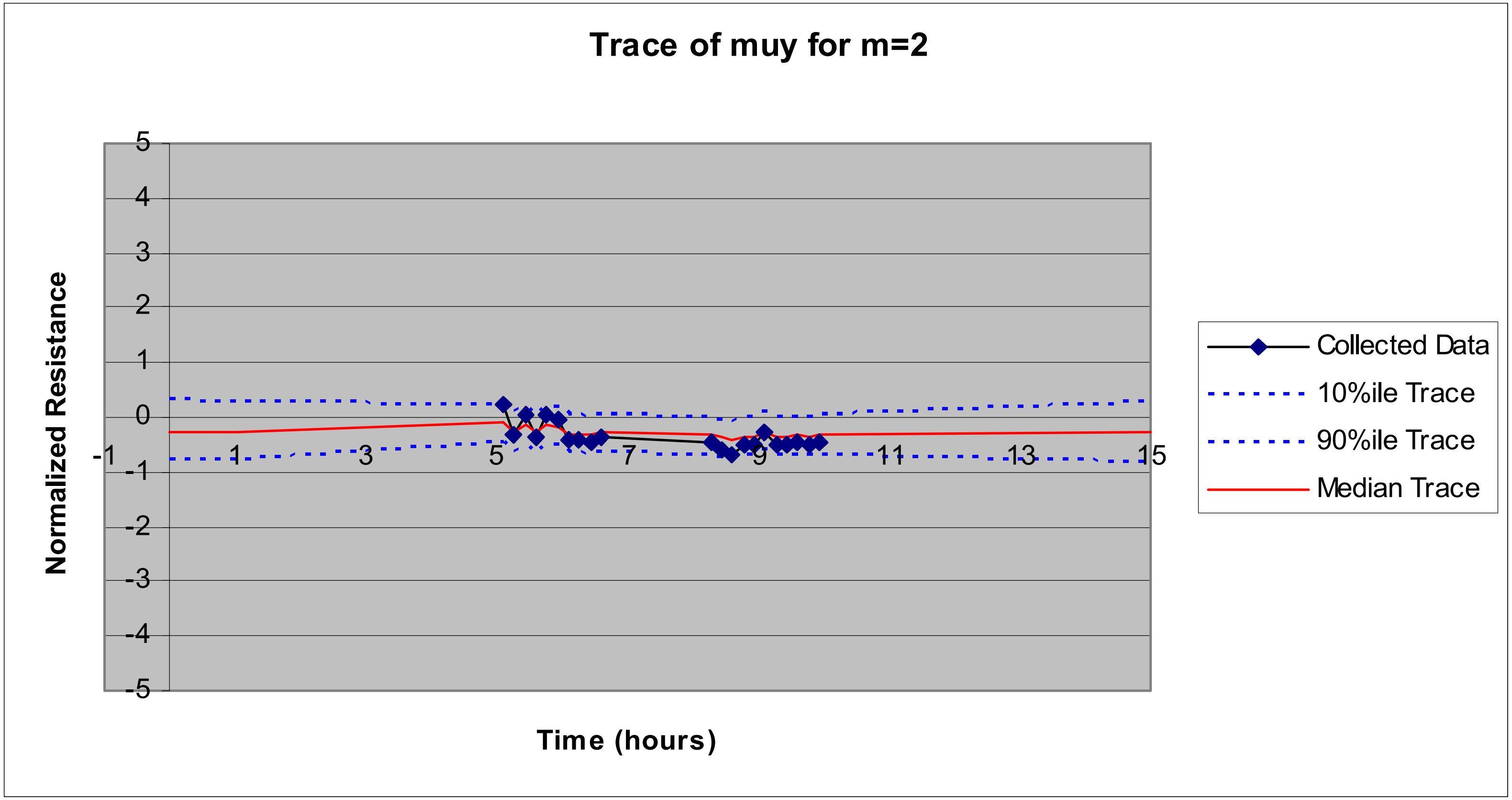
- Note that the assumed “true” value of the resistance is used in the failure rate model
- So, we have a joint model of degradation and failure rate
- Prior distributions can be input for all unknown parameters
- Data can be used to update the parameters using Bayes’ rule
- Hierarchies can be built to account for different levels of parameter interaction
  - “Population” failure rate
  - Component specific factor

Joint Model

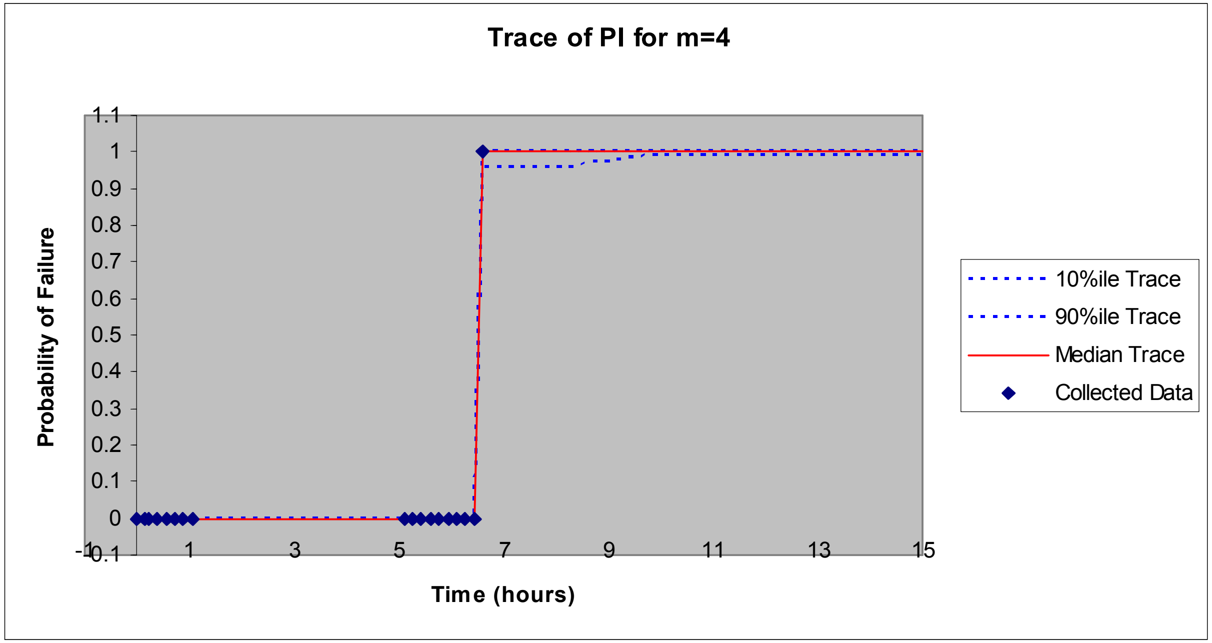


t[k]
b0
muy[k,m]
y[k,m]
G[k]
b1
d[k,m]
tauy[k,m]
b2
lambda0[k]
lambda[k,m]
Salt[m]
PI[k,m]
for(k IN 1 : K)
epsilon
for(m IN 1 : M)
a0
a1
a2

Degradation Results



Failure Rate Results



Path Forward

- STILL PRODUCING AND EVALUATING RESULTS
- Refine model to include all covariates
- Calculate mutual information of input parameters and output in order to assess coverage of model
- Develop real-time capability
- Currently working to adapt to Digital I&C applications