Evaluation of extreme events and reliability in the context of the Mars Sample Return mission

Richard Smith and Dawn Sanderson

Joint Statistical Meetings

Washington, DC, August 11, 2022

THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

The decision to implement Mars Sample Return will not be finalized until NASA's completion of the National Environmental Policy Act (NEPA) process. This document is being made available for information purposes only.

Acknowledgements

- This work was done in collaboration with JPL's "UQ Advisory Group" focusing specifically on the Mars Sample Return mission
- Thanks to other members of the group: Amy Braverman, Aaron Siddens, Kevin, Giuseppe, Ralph and others
- All opinions expressed here are those of the authors and not necessarily attributable to JPL, NASA or other members of the group

Objectives

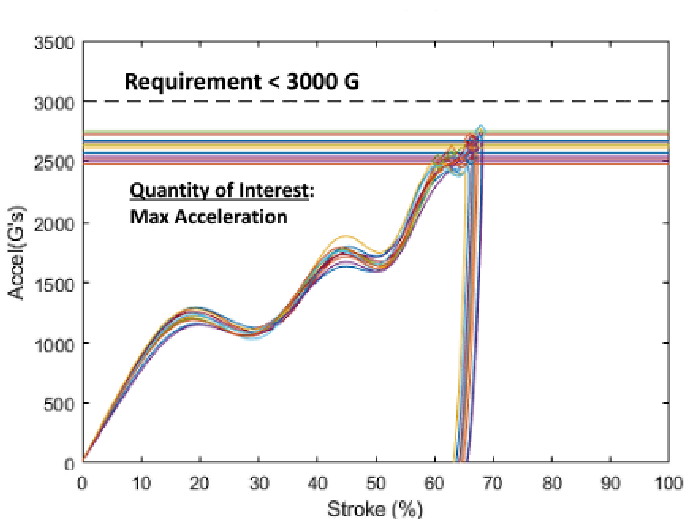
- Big Picture view: investigate how close the failure probability comes to the target of <10^{-6} and characterize uncertainty in the computation
- Little Picture: focus on one specific model (LS-DYNA model – next slide)
- Failure depends both on the strengths of the individual components (IM7 and Kevlar), and on how these components determine the quantity of interest ("Peak OS Acceleration")
- Our purpose: develop Bayesian methods that allow for model/parameter uncertainty as well as the inherent randomness of the system

Framework application to former off-nominal landing

FAILURE DEFINITION

CNA METRIC

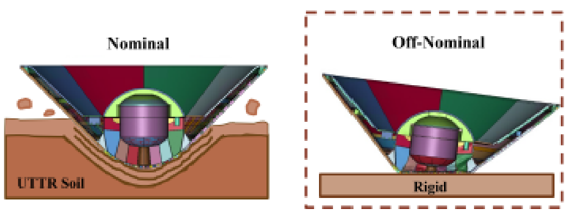
Quantity of interest (QoI)		Limit state	
COS Max Acceleration	–	3000 G	> 0



PROBABILITY COMPUTATION

Based on a LS-DYNA model (high-fidelity, 25 runs) + up to 4 low-fidelity models

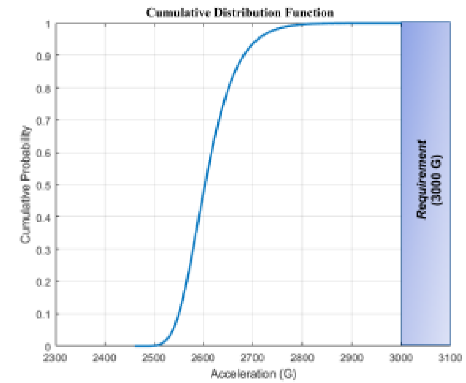
a. Low-fidelity models help speed up stochastic simulations.
b. High-fidelity model in the loop enables maintaining accuracy guarantees on the final result.



(Credit: Kevin Carpenter)

REQUIREMENT VERIFICATION

Simulations show the landing requirement is met.



Credit: Kevin Carpenter (Jet Propulsion Laboratory, California Institute of Technology)

Data I

- 15 variables measuring physical characteristics of the sample (7 for IM7, 8 for Kevlar)
- 5 of the variables measure elastic moduli (either Young's modulus or the shear modulus)
- The other 10 variables measure strengths in various modes (tensile, compressive, shear)
- Between 3 and 12 samples for each variable
- Previous work at JPL: fitting normal distributions for moduli, Weibull for the failure strengths

IM7 90 Deg Tensile Strength

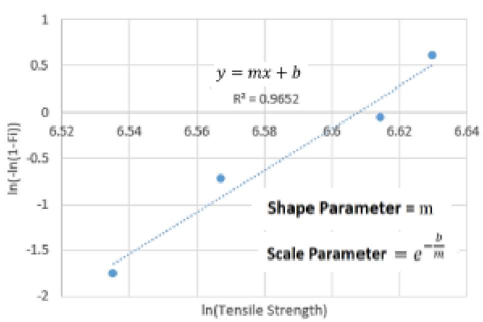
i	Tensile Strength (MPa)	Bernard's Estimator (Fi)	ln(-ln(1-Fi))	ln(Tensile Strength)
1	XXX	0.159091	-1.75289	XXX
2	XXX	0.386364	-0.71672	XXX
3	XXX	0.613636	-0.05027	XXX
4	XXX	0.840909	0.60883	XXX

$$U_{MR} = [1 + 0.72(n - 1.8)^{-1.21}]^{-1} + 0.008$$

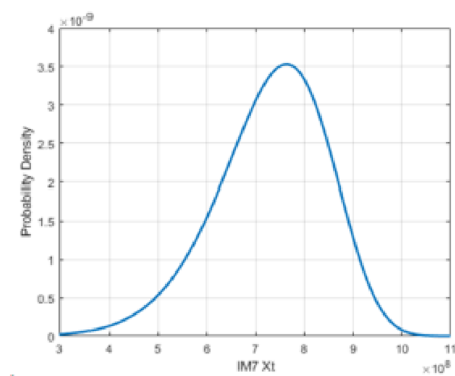
Ross's Modified Correcting Formula	
3 Samples	0.6419
4 Samples	0.79088

Scale Parameter (alpha) = xxx

Shape Parameter (beta) = xxx (UMR)



Zhang, L.F., Xie, M., and Tang, L.C. "Bias Correction for the Least Squares Estimator of Weibull Shape Parameter with Complete and Censored Data." *Reliability Engineering and System Safety*, 2006. doi:10.1016/j.ress.2005.09.010.

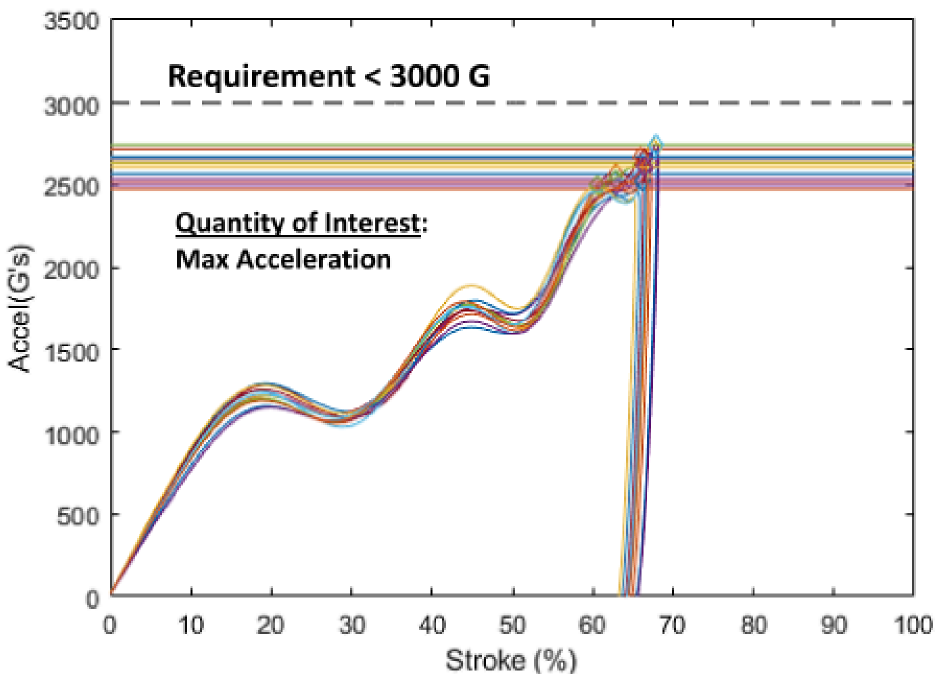


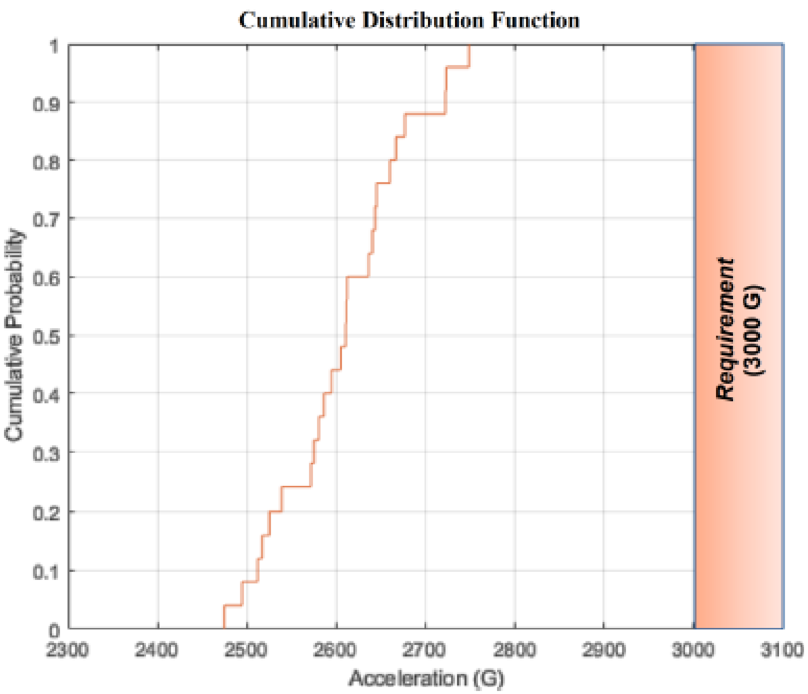
5

Credit: Kevin Carpenter (Jet Propulsion Laboratory, California Institute of Technology)

Data II

- UQ Input Data: 25 rows of X1-X15 data from latin hypercube sampling (LHS)
- UQ Output Data (next slide): Peak OS acceleration
- Requirement: a peak OS acceleration <3000 G





- Credit: Kevin Carpenter (Jet Propulsion Laboratory, California Institute of Technology)

Structure of this Presentation

- Part 1: Bayesian Weibull/Normal analysis of strength and modulus data
- Part 2: UQ analysis of the LS-DYNA experiment
- Part 3: End to end analysis of the probability of failure in this system (preliminary version!)

Weibull Distribution

- Cumulative distribution function (CDF):
  $F(y;\alpha,\beta) = 1 - \exp\left\{-\left(\frac{y}{\beta}\right)^{\alpha}\right\},\ y > 0,\ \alpha > 0,\ \beta > 0$

- Probability density function (pdf):
  $f(y;\alpha,\beta) = \frac{dF}{dy} = \alpha y^{\alpha-1}\beta^{-\alpha}\exp\left\{-\left(\frac{y}{\beta}\right)^{\alpha}\right\}$

- Negative log likelihood (NLLH): given sample $Y_1,\ldots,Y_n$,
  $\ell_n(\alpha,\beta) = -\sum_{i=1}^{n}\log f(Y_i;\alpha,\beta) = \sum_{i=1}^{n}\left\{-\log\alpha - (\alpha-1)\log Y_i + \alpha\log\beta + \left(\frac{Y_i}{\beta}\right)^{\alpha}\right\}$

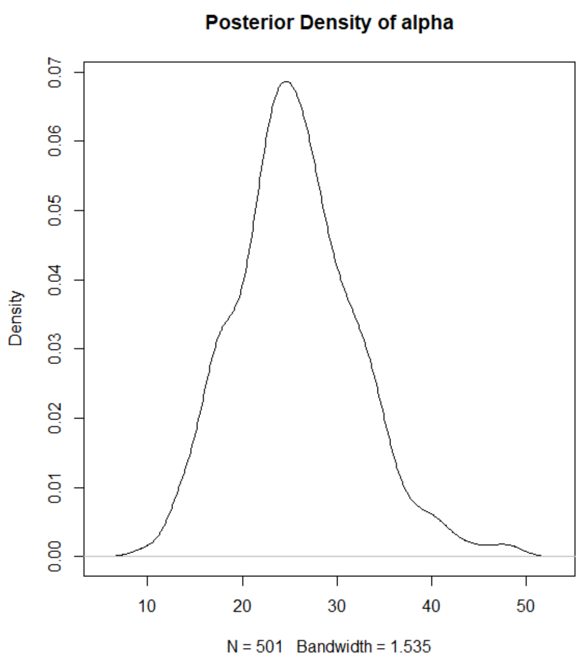
- Maximum likelihood estimation (MLE) usually computed by minimizing the NLLH: for given sample $Y_1,\ldots,Y_n$, choose α,β to minimize $\ell_n(\alpha,\beta)$

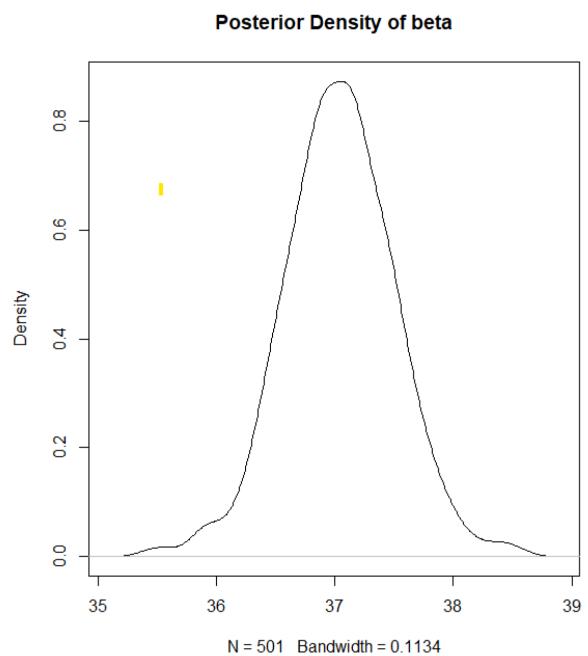
- Bayesian estimation: assume prior density $\pi_0(\alpha,\beta)$, then
  $\pi(\alpha,\beta|Y_1,\ldots,Y_n) = \frac{\pi_0(\alpha,\beta)\prod_{i=1}^{n} f(Y_i;\alpha,\beta)}{\int\int \pi_0(\alpha,\beta)\prod_{i=1}^{n} f(Y_i;\alpha,\beta)d\alpha d\beta}$

- Practical computation: use Adaptive Metropolis Sampler (AMS) of Haario *et al.* (2001) [there are numerous variants, such as the Delayed Rejection Adaptive Metropolis (DRAM) algorithm]

Results for X11 (12 values)

- Y values (scaled): 34.29, 34.84, ... , 38.74
- MLE: $\hat{\alpha} = 27.52$ (SE=6.16), $\hat{\beta} = 37.03$ (SE=0.41); standard errors from observed Hessian matrix; same results using R functions `optim` and `nlm`
- AMS: 100,000 samples (2.9 seconds, acceptance rate 0.25)
  For α: posterior mean 25.69, posterior SD 6.40
  For β: posterior mean 37.04, posterior SD 0.47





Normal Distribution

- Cumulative distribution function (CDF):
  $G(y;\mu,\sigma) = \Phi\left(\frac{y-\mu}{\sigma}\right),\ \Phi =$ standard normal cdf, $\sigma > 0$

- Probability density function (pdf): $g(y;\mu,\sigma) = \frac{dG}{dy} \propto \frac{1}{\sigma}\exp\left\{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right\}$

- Negative log likelihood (NLLH): given sample $Y_1,\ldots,Y_n$,
  $\ell_n(\mu,\sigma) = -\sum_{i=1}^{n}\left\{\log\sigma + \frac{1}{2}\left(\frac{Y_i-\mu}{\sigma}\right)^2\right\}$

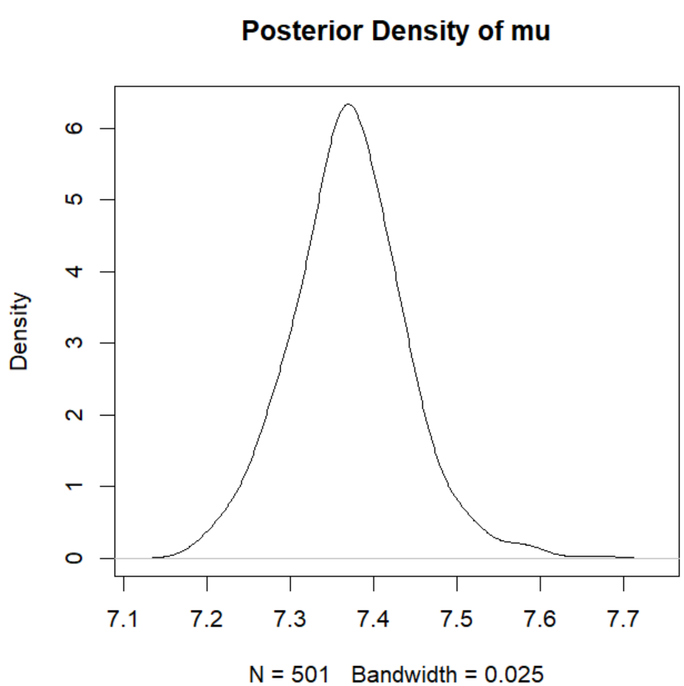
- Maximum likelihood estimation (MLE) by minimizing the NLLH: note that the MLE for σ is different from the usual sample SD (n rather than $n-1$ in numerator)

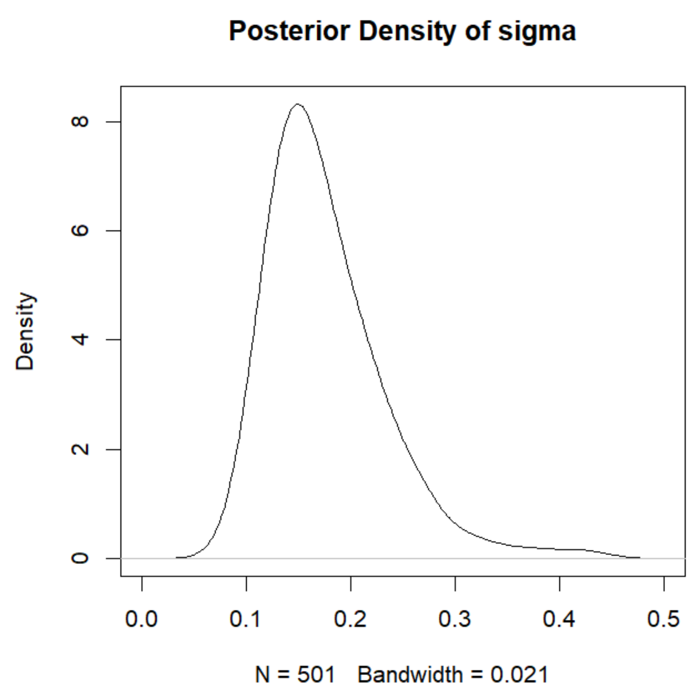
- Bayesian estimation: assume prior density $\pi_0(\mu,\sigma)$, then
  $\pi(\mu,\sigma|Y_1,\ldots,Y_n) = \frac{\pi_0(\mu,\sigma)\prod_{i=1}^{n} g(Y_i;\mu,\sigma)}{\int\int \pi_0(\mu,\sigma)\prod_{i=1}^{n} g(Y_i;\mu,\sigma)d\mu d\sigma}$

- Of course for this case we could directly use the sample mean and SD but we have computed the AMS for consistency with the Weibull analysis

Results for X1 (8 values)

- Y values (scaled): 7.23, 7.64, ... , 7.35
- Usual sample estimates $\bar{y} = 7.37$, $s_y = 0.16$ (standard errors 0.06, 0.04)
  MLE: $\hat{\mu} = 7.37$ (SE=0.05), $\hat{\sigma} = 0.15$ (SE=0.04).
- AMS: 100,000 samples (2.2 seconds, acceptance rate 0.34)
  For μ: posterior mean 7.37, posterior SD 0.07
  For β: posterior mean 0.17, posterior SD 0.05





UQ Analysis: Gaussian Process Modeling for Computer Experiments

- Raw data (rescaled): 25 runs of a LHS experiment on variables X0001,...,X0015, output is "Peak OS Acceleration" (original units divided by 1000)

- The problem: determine probability that Peak OS Acceleration$>$ 3 when the input variables are random

Gaussian Process Approach (Sacks et al. 1989, Kennedy & O'Hagan 2001)

- Objective: Represent an output variable z_i in terms of input variables that I denote by $s_{ik},\ k = 1, \ldots, K$

- In this experiment, each $s_{ik},\ i = 1, \ldots, 25$ is one of X0001–X0015, rescaled to mean 0 and variance 1

- The model also includes linear covariates X though for this analysis we don't use them (assume constant-mean response)

- Model $\mathbf{Z} \sim \mathcal{N}_n[X\beta, \Sigma]$ where $\mathcal{N}$ denotes n-dimensional normal, $X\beta$ is mean response (for us: $\mu\mathbf{1}$) and Σ is covariance matrix

- $\Sigma = \alpha V(\boldsymbol{\theta})$ where $\alpha > 0$ and a very popular choice for V is

$$v_{ij} = \exp\left\{-\sum_{k=1}^{K}\left(\frac{s_{ik}-s_{jk}}{\rho_k}\right)^2\right\}$$

where $\rho_1, \ldots, \rho_K$ are "spatial range" parameters. In practice I write $\theta_k = \log \rho_k$ and $\boldsymbol{\theta} = \begin{pmatrix} \theta_1 & \theta_2 & \ldots & \theta_K \end{pmatrix}$.

Connections with Spatial Statistics

- Refs: RLS notes (http://rls.sites.oasis.unc.edu/postscript/rs/envnotes.pdf); books by Cressie, Stein and several others

- Assuming $\mathbf{Z} \sim \mathcal{N}_n[X\beta, \alpha V(\boldsymbol{\theta})]$ where X is $n \times q$, define Define

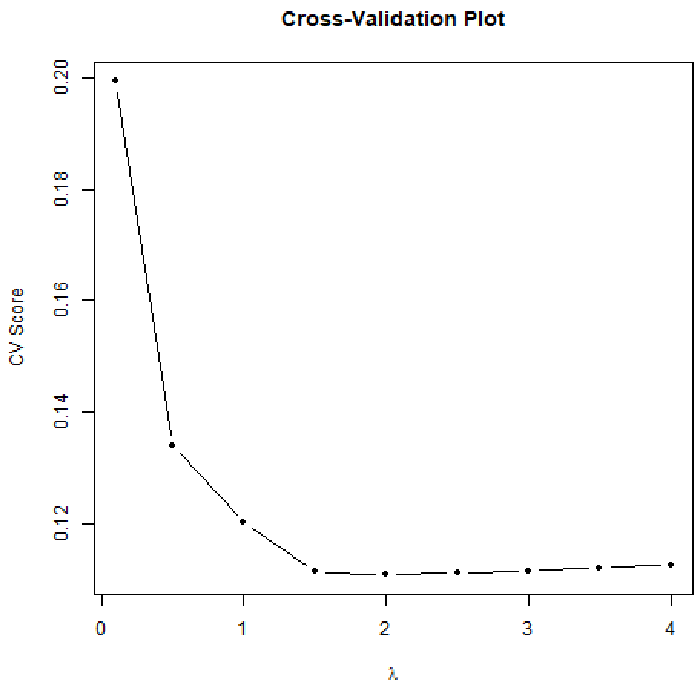
$$\begin{aligned} H &= V(\boldsymbol{\theta})^{-1} - V(\boldsymbol{\theta})^{-1}X(X^TV(\boldsymbol{\theta})^{-1}X)^{-1}X^TV(\boldsymbol{\theta})^{-1}, \\ G^2(\boldsymbol{\theta}) &= \mathbf{Z}^T H \mathbf{Z}, \\ \widehat{\alpha} &= G^2(\boldsymbol{\theta})/(n-q), \\ \ell^*(\boldsymbol{\theta}) &= \frac{n-q}{2}\log G^2(\boldsymbol{\theta}) + \frac{1}{2}\log\det(X^TV(\boldsymbol{\theta})^{-1}X) + \frac{1}{2}\log\det(V(\boldsymbol{\theta})) \end{aligned}$$

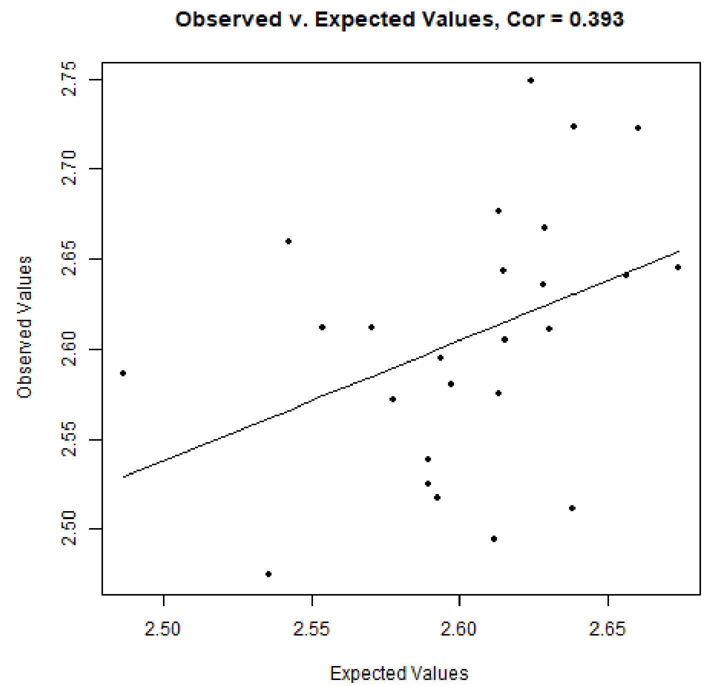
- $e^{-\ell^*(\boldsymbol{\theta})}$ is called the *restricted likelihood function* and the quantity $\widehat{\boldsymbol{\theta}}$ that minimizes $\ell^*(\boldsymbol{\theta})$ is called the *REML estimator*. Bayesian inference also starts with $\ell^*(\boldsymbol{\theta})$

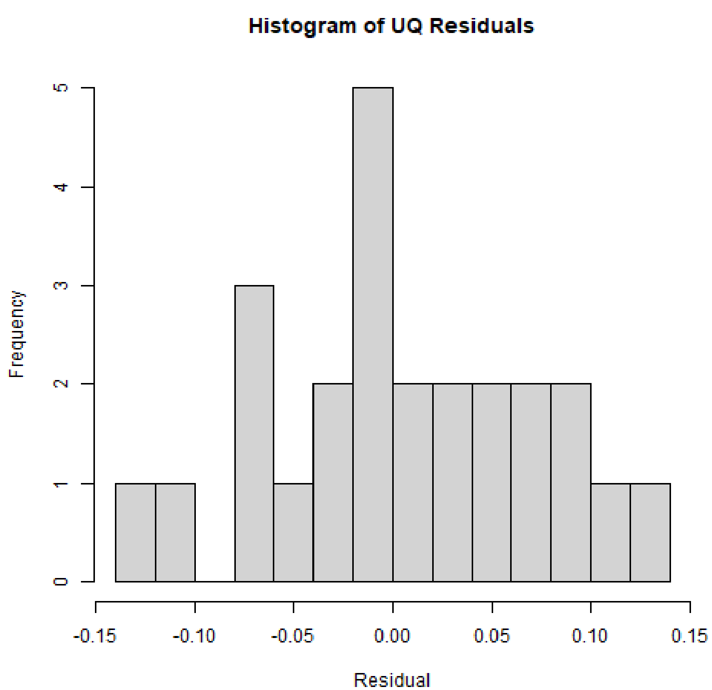
- Kriging formulas: To predict a new value z_0 with mean $x_0^T\beta$, variance σ_0^2 and $\mathrm{Cov}(z_0, \mathbf{Z}) = \boldsymbol{\tau}$, calculate $\widehat{z}_0 = (\mathrm{x}_0 - X^T\Sigma^{-1}\boldsymbol{\tau})^T\widehat{\beta} + \boldsymbol{\tau}^T\Sigma^{-1}\mathbf{Z}$ with mean squared prediction error $\sigma_0^2 - \boldsymbol{\tau}^T\Sigma^{-1}\boldsymbol{\tau} + (\mathrm{x}_0 - X^T\Sigma^{-1}\boldsymbol{\tau})^T(X^T\Sigma^{-1}X)^{-1}(\mathrm{x}_0 - X^T\Sigma^{-1}\boldsymbol{\tau})$.

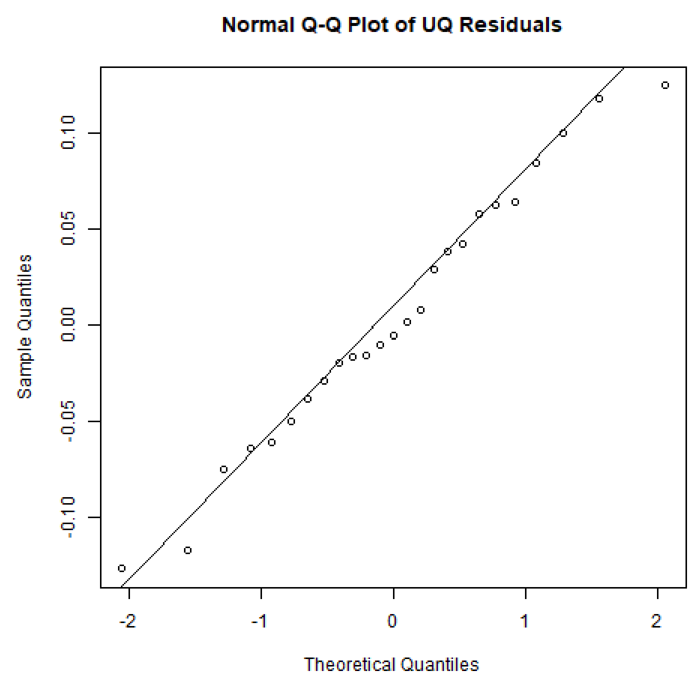
Application to this dataset

- 25 datapoints with 15 unknowns — already problematic!
- Tried three R optimization routines: optim with method="Nelder-Mead", optim with method="BFGS", nlm. Wildly varying solutions, none has positive definite hessian
- Alternative: try a "regularized REML" minimize $\ell^*(\boldsymbol{\theta}) + \lambda \sum (\theta_k - \bar{\theta})^2$.
- Seems to work when $\lambda > 0$, still some pathological behavior with optim routine, nlm not sensitive to starting values
- Use cross-validation to select best $\lambda = 2$.
- Diagnostics: Cross-validation plot, plot of z_0 against $\hat{z}_0$, histograms and QQ plots of $z_0 - \hat{z}_0$


Cross-Validation Plot
CV Score
0.20
0.18
0.16
0.14
0.12
0
1
2
3
4
λ

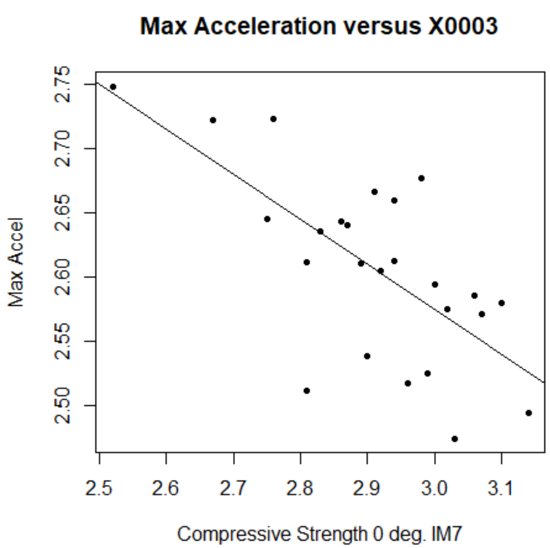

Observed v. Expected Values, Cor = 0.393
Observed Values
2.75
2.70
2.65
2.60
2.55
2.50
Expected Values


Histogram of UQ Residuals
Frequency
0
1
2
3
4
5
-0.15
-0.10
-0.05
0.00
0.05
0.10
0.15
Residual


Normal Q-Q Plot of UQ Residuals
Sample Quantiles
0.10
0.05
0.00
-0.05
-0.10
-2
-1
0
1
2
Theoretical Quantiles

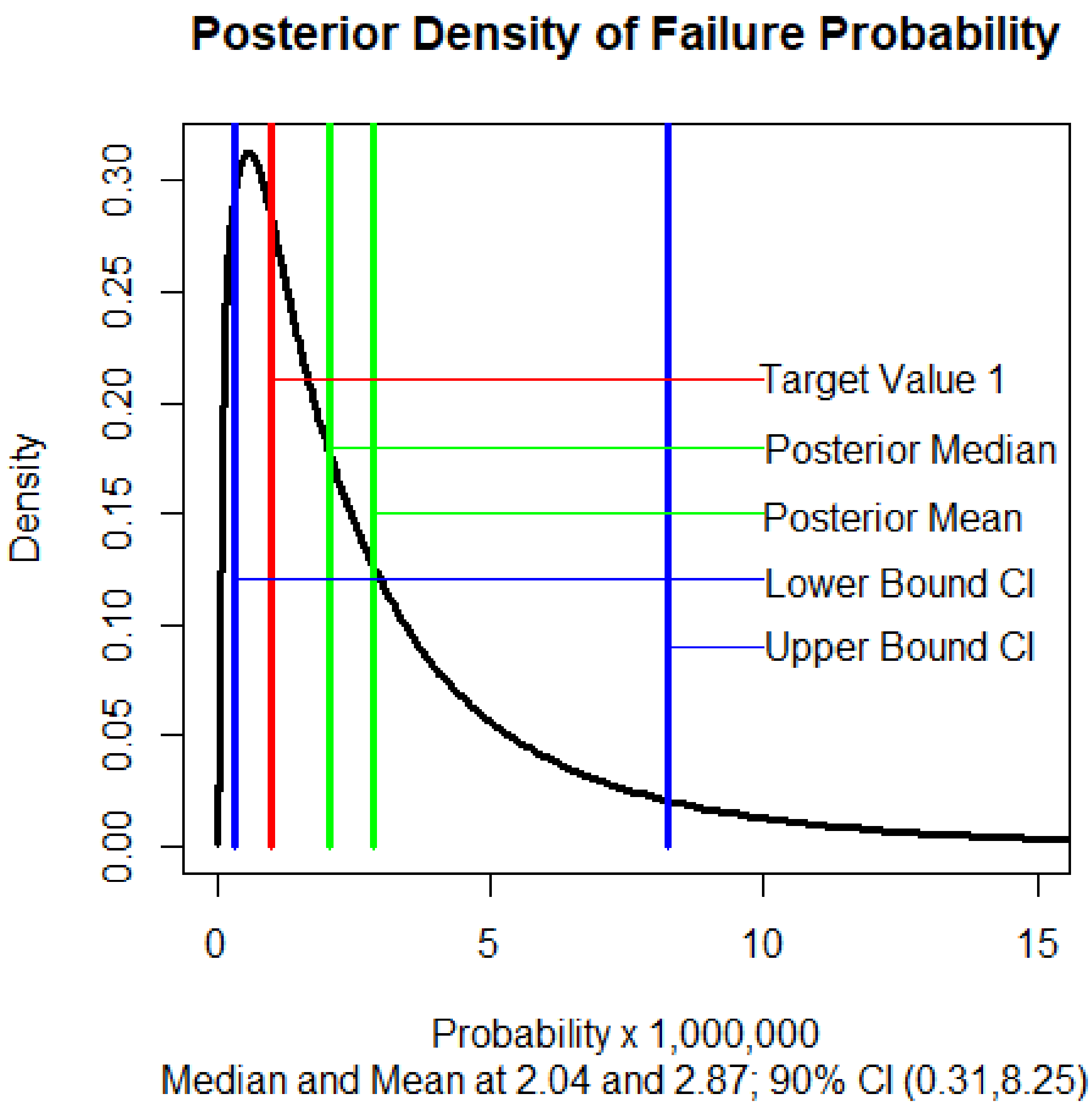
Critique

- The variance of the noise is greater than that of the signal.
- It does appear that the residuals closely follow a normal distribution.
- One possible issue is too narrow a range of input values:
  - We would get greater signal to noise ratio with a wider range of inputs
  - There are no simulated values in the critical region (in this scaling, $z > 3$ — observed range is 2.474 to 2.749)
  - Kriging with constant mean is an *interpolator*, which means it will never produce a predicted value outside the range of data (and input values far from the test dataset will simply be predicted back to the overall mean)
  - Devil's advocate argument: if we can't produce a critical value in a simulation, how can understand the conditions under which this might occur in reality?
- In fact, a linear regression fits this dataset better!
  - Regress z on X0001–X0015 with standard variable selection: X0003 is by far the strongest predictor (however, the minimum AIC model has 6 predictors)



Simulation Strategy

- Fix $N = 2000$, $M = 1000$ (or larger). Want to calculate $\Pr(z > z_{\text{crit}})$ for some critical value z_{crit}
- For $i = 1$ to N:
  - For $k = 1, \ldots, K = 15$, select either μ_k, σ_k (normal cases) or α_k, β_k (Weibull cases) from the respective posterior distributions of the kth input variable. *This is one sample from the posterior distribution.*
  - For $j = 1$ to M:
    - For $k = 1$ to 15, simulate one value of the normal or Weibull distribution corresponding to the kth input variable. This is our trial set of inputs $\mathbf{s}_0$.
    - Use the UQ model to obtain the mean predictor $\hat{z}_0$ and its RMSE S_0 at predictor $\mathbf{s}_0$.
    - Exceedance probability $1 - \Phi\left(\frac{z_{\text{crit}} - \hat{z}_0}{S_0}\right)$ (extension to be considered later: use a longer-tailed distribution, e.g. t).
  - Average over $j = 1, \ldots, M$ to obtain one estimate $\hat{p}_{\text{crit}}^{(i)}$ for the probability of exceeding z_{crit}.
- The set of values $\hat{p}_{\text{crit}}^{(n)}$, $i = 1, \ldots, N$ is treated as a sample from the posterior distribution of p_{crit}.



Posterior Density of Failure Probability
Density
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
5
10
15
Target Value 1
Posterior Median
Posterior Mean
Lower Bound CI
Upper Bound CI
Probability x 1,000,000
Median and Mean at 2.04 and 2.87; 90% CI (0.31,8.25)

Overview and Summary

- This is a *parametric* statistical analysis — normal, Weibull, GP distributions with parameters $\mu_k,\ \sigma_k,\ \alpha_k,\ \beta_k$ for each component, $\boldsymbol{\theta}$ for the GP model
- Known parameters: calculate exact p_{crit} by simulation
- Unknown parameters: p_{crit} depends on parameters, but we integrate different parts of the analysis to calculate a *posterior distribution* $\pi(p_{crit} \mid \mathbf{D})$ where $\mathbf{D}$ is data (not yet fully implemented for $\boldsymbol{\theta}$)
- How to interpret $\pi(p_{crit} \mid \mathbf{D})$? Various metrics, e.g. posterior mean or median, calculate a 95% credible interval, etc.
- The analysis is still dependent on the validity of various parametric models
- The analysis is also dependent on high-quality **data**. A sophisticated statistical model cannot make up for the lack of high-quality data