Sensitivity Strategy Supporting the Estimate of Extremely Low Probabilities

Cédric J. Sallaberry, Robert E. Kurth, Frederick W. Brust, Elizabeth A. Kurth-Twombly

Engineering Mechanics Corporation of Columbus (Emc^2)
3518 Riverside Drive - Suite 202
Columbus, OH 43221-1735
Phone: (614) 459-3200
Fax: (614) 459-6800

Context

- **The result of active degradation mechanisms, primarily Primary Water Stress Corrosion Cracking (PWSCC), presented challenges to the Leak Before Break (LBB) analysis**
- **The US Nuclear Regulatory Commission (USNRC) and the Electric Power Research Institute (EPRI) jointly developed a probabilistic fracture mechanics code to assess LBB: xLPR (extremely Low Probability of Rupture)**
- **The extremely low occurrence of events (10^{-6} range) requires optimization of the sampling based methods in order to generate adverse events when using a reasonable sample size**
- **In order to concentrate the sampling in the area of interest in the input space it is important to**
  - 1) Identify the uncertain input driving the output uncertainty
  - 2) Identify the region of interest for each of these inputs (or conjointly)
- **The following presents the proposed sensitivity analysis methodology developed with that purpose in mind**

Purpose of sensitivity Analysis

- **Sensitivity Analysis of a probabilistic model is used**
  - To understand the influence of input uncertainty to the response uncertainty
  - Mapping the uncertainty from the output of interest to the response uncertainty
  - These results **may** change if the response is changed
  - Ranking uncertain inputs by importance : how much they contribute to the output uncertainty

- **Local sensitivity analyses**
  - Estimate of gradients
    - Computationally intense when there are many uncertain variables
    - limited to the neighborhood of interest
- **Global sensitivity analysis considers the entire response variation. Analysis of Variance (ANOVA) considered to decompose the output variance into inputs contribution.**

Regressions analyses considered

- **Three regression analyses considered**
  - Stepwise rank regression analysis
    - Captures monotonic influence starting with one variable fitting and adding another input to the regression model until a stopping criterion is reached
    - Additive (no conjoint influence) and Linear/Monotonic
    - Works well in the majority of cases (most influences are monotonic)
  - Recursive partitioning (tree regression)
    - Looks at threshold values that may split the response uncertainty into high/low groups
    - Captures conjoint influence and non monotonic relations
    - May have a tendency to overfitting
  - MARS (Multi-Adaptive Regression Spline)
    - Spline regression in a stepwise way
    - Captures conjoint influence and non monotonic relations
    - Splines makes it inefficient when discrete variables are considered

Rationale in using multiple regressions

- All regressions have assumptions: the use of one regression type may bias the results and may miss some specific relations
- Complex systems considered with conditional results (with inspection, leak rate detection) may make the input space of interest disjoint
- More sophisticated regressions may over-fit when used on a large number of inputs (200 inputs ~ 40K possible combinations when considering two input interactions)
- Cost: more complex interpretation at the end since many results available
- Strategy developed to aggregate all results

Qualitative Analysis : Scatterplots

- Qualitative so the purpose of scatterplots is NOT to rank and compare uncertain inputs
- If large number of uncertain inputs, may not be appropriate to plot all of them
- However, no assumption is made. Any relation is visible. Helps increase confidence in regression analyses when displayed for the most influential inputs identified (up to 10)
- Graphical representation is more efficient and effective when communicating results to diverse audiences than tables of numbers

Example: Crack initiation and number of axial cracks

- **The initiation of a crack in a pressurized pipe at the weld level is modeled using the following equation**

$$t_{INI,nom} = \left(\frac{e^{\frac{Q}{RT}}}{A\sigma^n}\right)$$

$$A = A_{mult} \times A_i$$

- Where
  - Q is the activation energy
  - R the gas constant
  - T the temperature
  - σ the stress in the inner part of the weld
  - A the proportionality constant which is split into two component : weld to weld variation and within weld variation.
- **Two outputs considered : Probability of having a circumferential crack over 60 yrs. (rare event) and number of axial cracks occurring during the same period.**

Example : Regression results

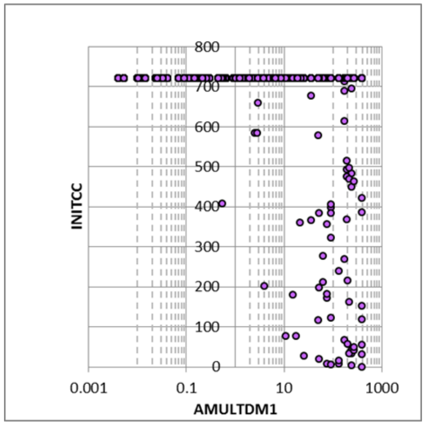
Coefficient of determination not as good when small number of events

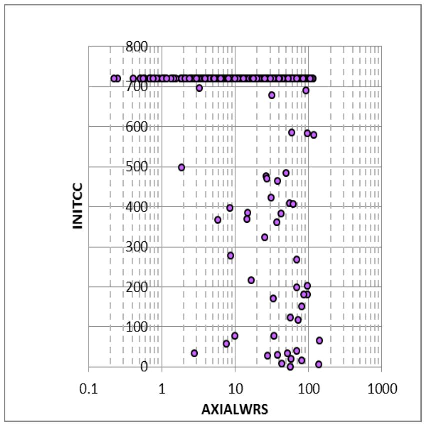
But the main drivers are found

Better regression when events are more common

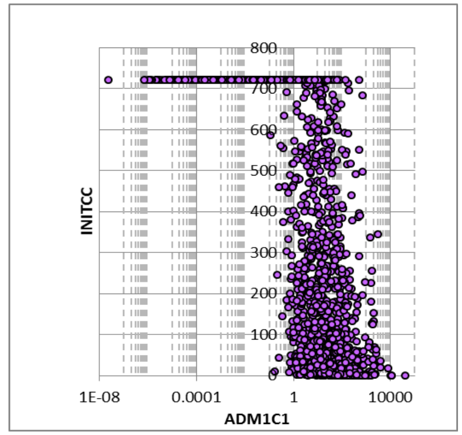
Example : Scatterplots for crack initiation

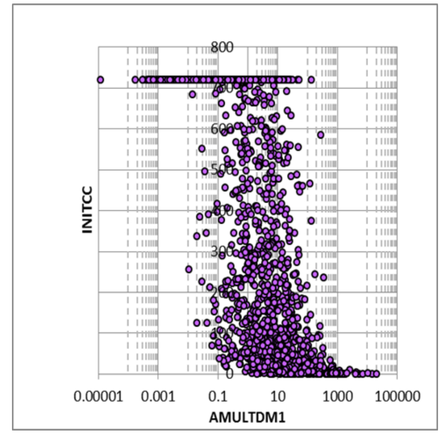
Plant 1

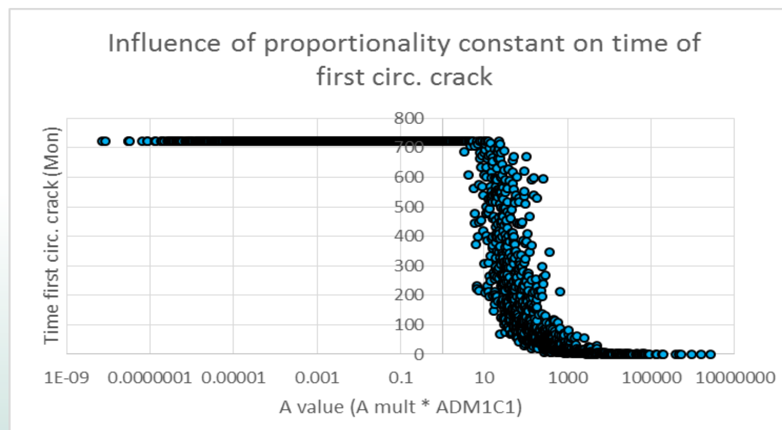




Plant 2



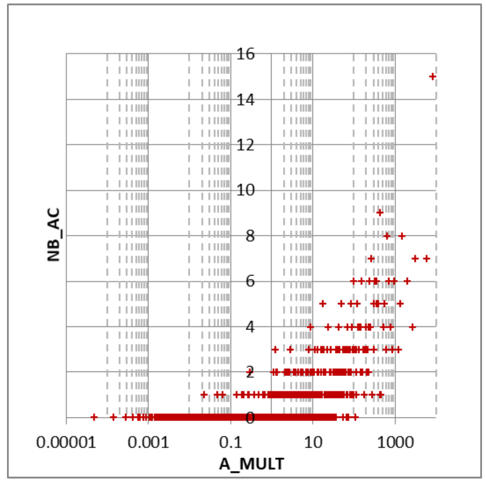


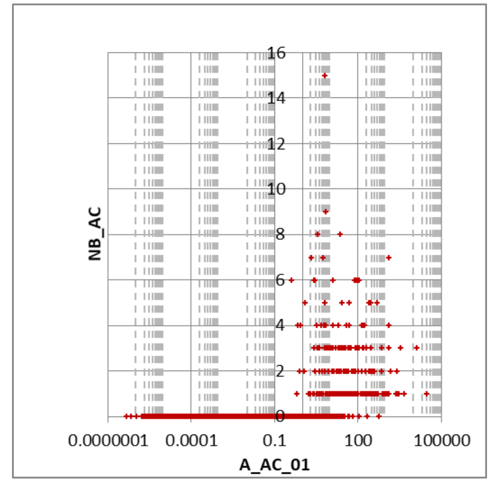


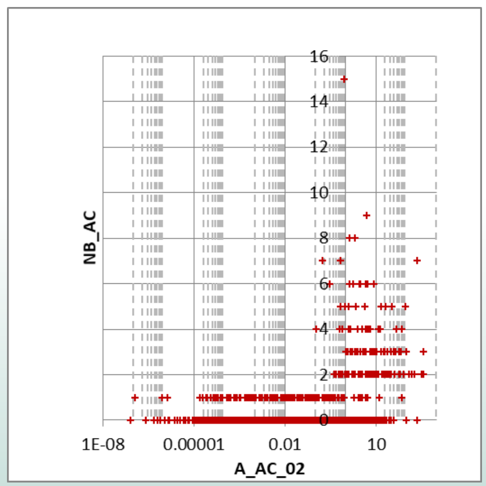
Example : Scatterplots for number of axial cracks

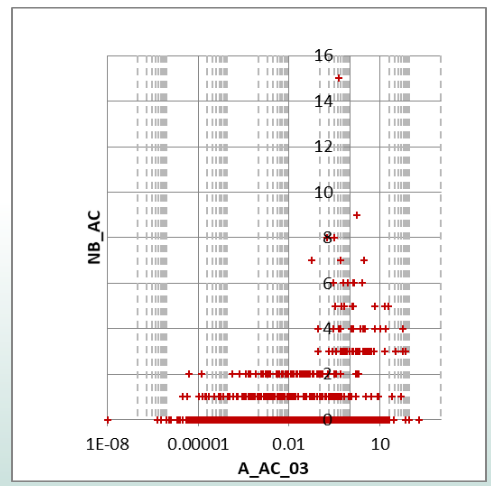
- **Multiplier affects all runs**

- **A_{01} separates 0 to 1 axial crack**
- **A_{02} separates 1 to 2 axial cracks**
- **A_{03} separates 2 to 3 axial cracks**









Conclusion

- **Sensitivity Analysis provides**
  - A better understanding of the model
  - Checking for potential errors in the framework construction (if the results make no sense)
  - Prioritizing research to improve the analysis
- **In this project we also use it to identify the main drivers so we can apply importance sampling on them.**
- **Any single technique may miss an important part of the analysis. The use of a suite of techniques and regression increase the confidence in the analysis.**
- **The use of graphical representation both increases confidence and simplifies communication of results. It also helps to identify what the region of importance is.**