SAVE PRINT

LA-UR- 03-7640

Approved for public release; distribution is unlimited.

Title:

OVERVIEW OF UNCERTAINTY QUANTIFICATION AND MODEL VERIFICATION AND VALIDATION (U)

Author(s):

François M. Hemez, Los Alamos National Laboratory, ESA-WR

Submitted to:

Invited Lectures at the Pan American Advanced Study Institute on Damage Prognosis, October 19-31, 2003, Florianopolis, Brazil

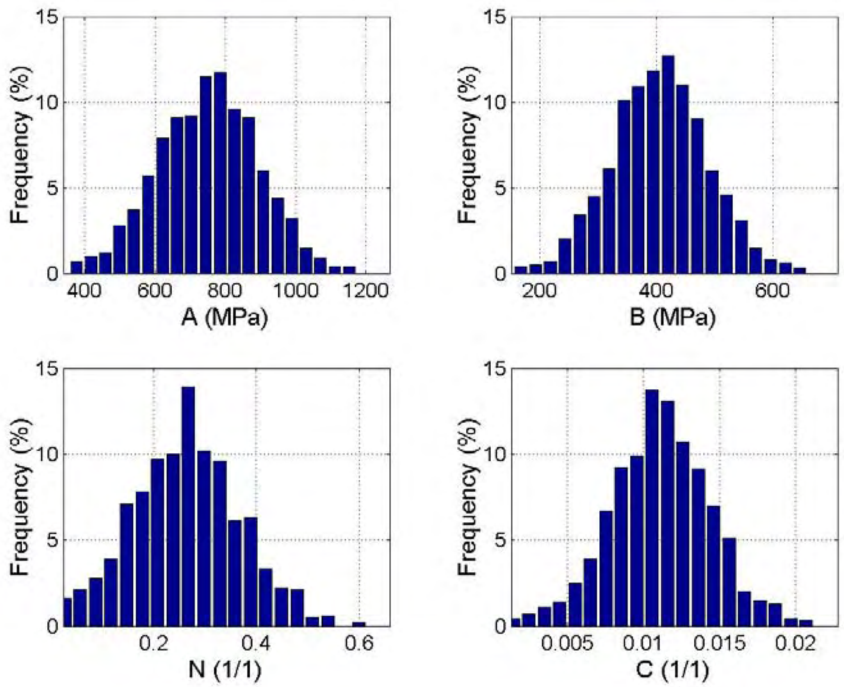
SAVE PRINT CLEAR FORM

Los Alamos
NATIONAL LABORATORY

QUANTIFICATION OF UNCERTAINTY

François M. Hemez

Los Alamos National Laboratory
Engineering Sciences and Applications
ESA-WR, Mail Stop T006
Los Alamos, New Mexico 87545
hemez@lanl.gov



APPROVED FOR UNLIMITED, PUBLIC RELEASE ON OCTOBER 16, 2003. LA-UR-03-7640. UNCLASSIFIED.

Outline

- What is Uncertainty Quantification?
- Statistical Sampling
- Design of Experiments
- Variance Analysis and Effect Screening
- Meta-modeling and Response Surfaces
- Non-probabilistic Methods

Uncertainty

- **Uncertainty is here defined in a broad sense …**
- **Uncertainty can be *aleatoric*: Randomness, variance. Can be better characterized, but cannot be reduced by taking more measurements or performing more simulations.**
- **Uncertainty can be *epistemic*: Lack-of-knowledge, vagueness. Can be reduced by collecting more information and evidence.**
- **Uncertainty can be *systematic*: Error, bias, residue. May come from an inappropriate functional form of the model, a lack-of-knowledge, etc.**

Activities of Uncertainty Quantification

- Propagate uncertainty from inputs to outputs

  → Use statistical sampling, design of experiments

- Explore a range of variations on the inputs

  → Use the design of experiments

- Study the effect of uncertainty

  → Use the analysis of variance, effect screening

- Meta-modeling is not associated to uncertainty, but it relies heavily of techniques such as the design of experiments and effect screening.

Uncertainty in Testing

- **Measurements collected during physical tests *always* exhibit an uncertainty that must be reported.**

- **Instead of saying *"The resonant frequency is 10.3 Hz"*, say *"The resonant frequency is 10.3 Hz +/- 0.1 Hz"*. These quantifies are random variables, they come with a mean and standard deviation (and possibly higher order statistical moments).**

- **Replicate experiments, either tests or simulations, are *the only way* to estimate the variability of a quantity, variability due to environmental changes, unknown interactions, or uncontrollable parameters.**

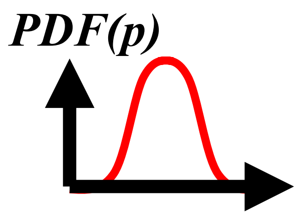
Outline

- What is Uncertainty Quantification?
- Statistical Sampling
- Design of Experiments
- Variance Analysis and Effect Screening
- Meta-modeling and Response Surfaces
- Non-probabilistic Methods

The Concept of Probability

- What is a probability?
- The *frequentist* concept is that a probability is the frequency of occurrence of an event occurring within a collective of "similar" events.

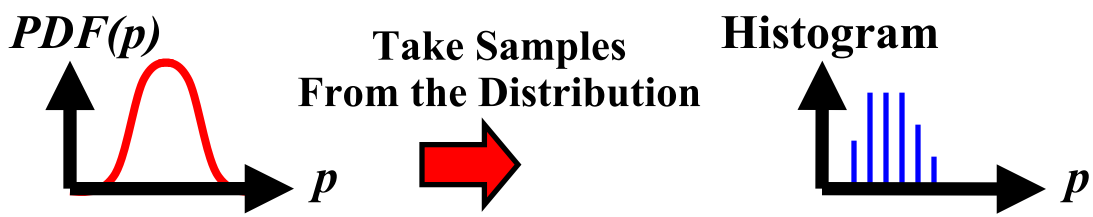
$$\text{Probability} = \frac{\text{Number of Occurrences}}{\text{Total Number of Events}}$$



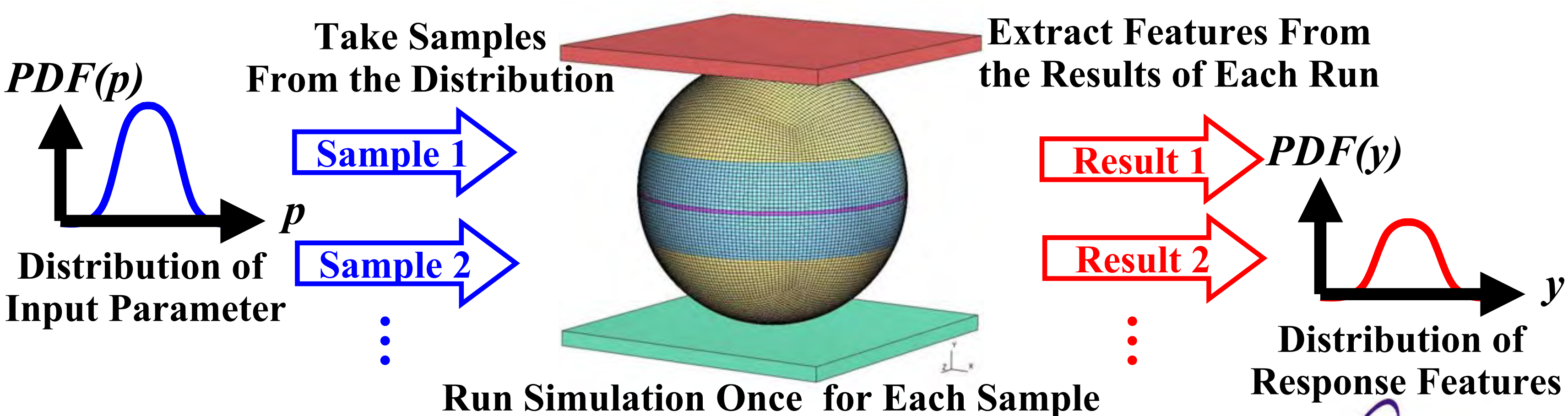
- The *Bayesian* concept is that a probability is an opinion, a degree of belief. *"What is the probability of life on Mars?" "What are the chances of rain?"*
  - Both satisfy the same basic axioms: P(Null)=0; P(All)=1; P(AUB) = P(A) + P(B) if A, B are disjoint events, etc.
  - UQ is, here, discussed assuming the frequentist concept.

What is Sampling?

- **The basic principle of sampling is to randomly draw values from a given probability distribution.**



- **General procedure for propagating uncertainty:**



Sampling Techniques

- **Sampling from probability distributions:**
  - Monte Carlo sampling
  - Latin Hypercube sampling
  - Markov Chain Monte Carlo
  - Many others …

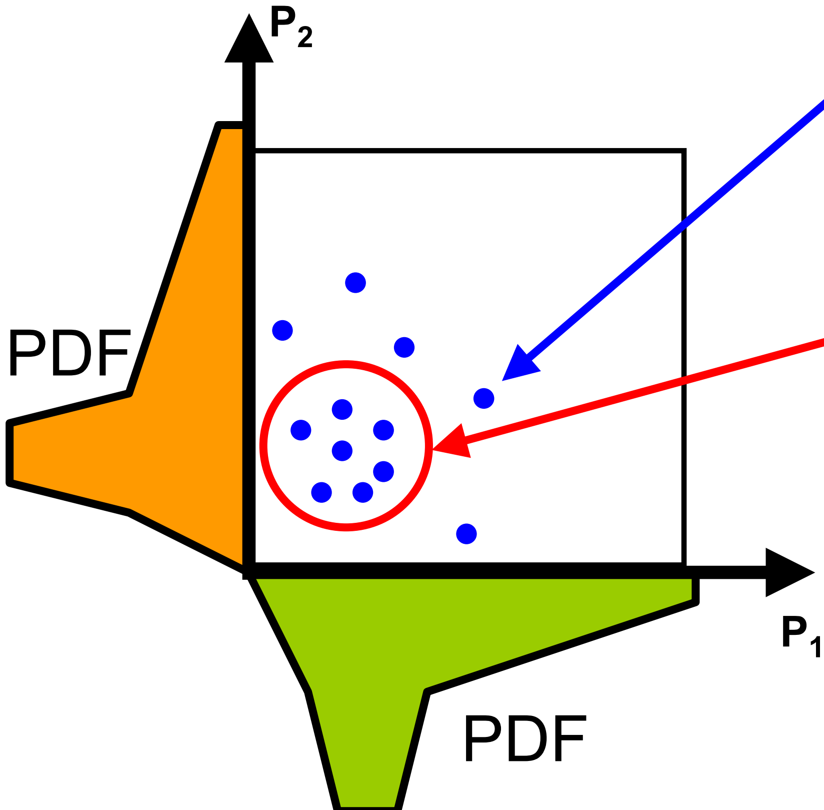
  Generally more efficient for uncertainty propagation.

- **Design of Experiments (DoE) techniques:**
  - Full-factorial design
  - Fractional factorial designs
  - Orthogonal arrays
  - Central composite designs
  - Many others …

  Generally more efficient for effect screening and meta-modeling.

Monte Carlo Sampling

- **Each sample is randomly sampled according to the joint probability density function.**

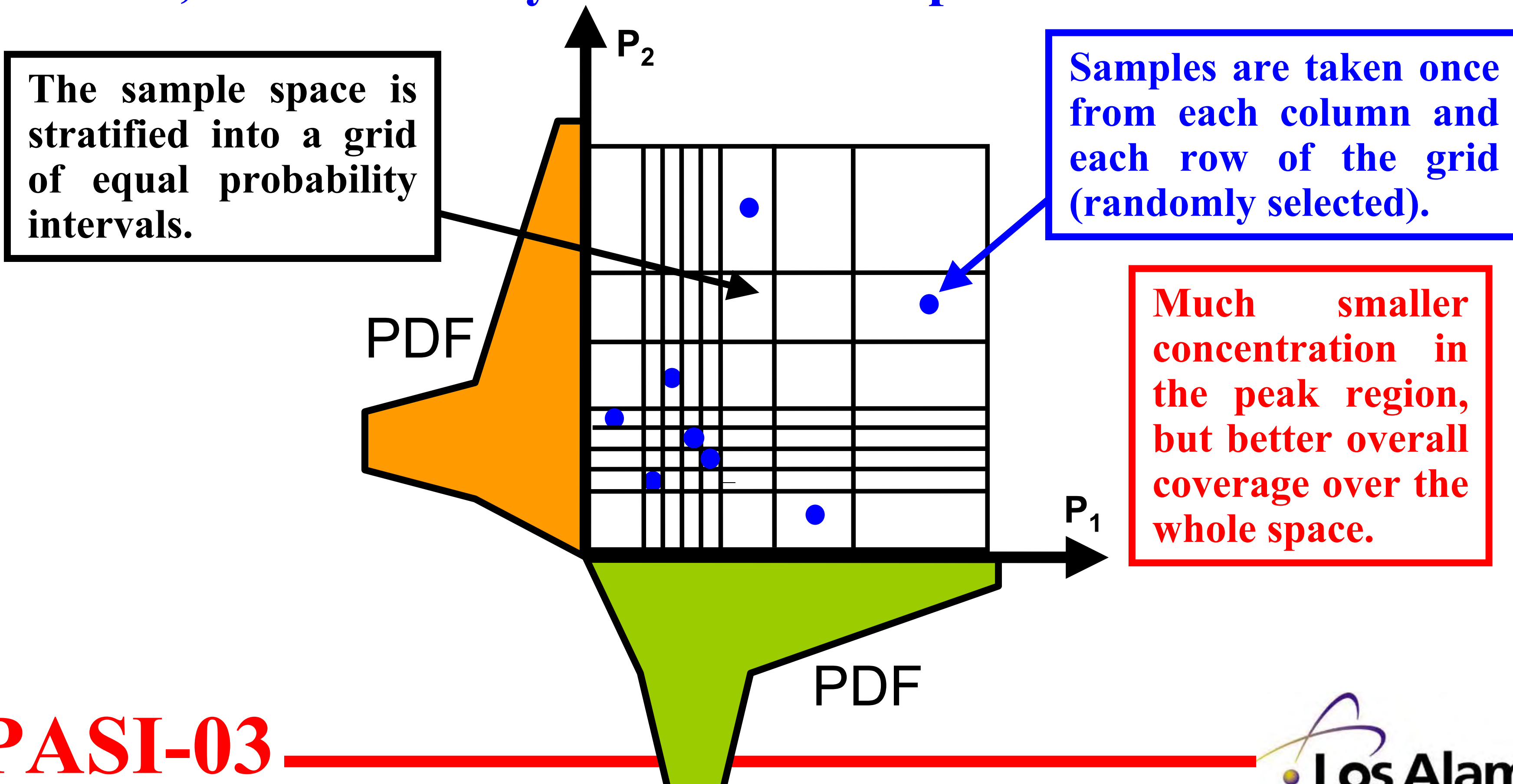


The points represent samples taken from the joint PDF of p_1 and p_2.

Note the concentration of samples around the area of highest joint probability. However, not very widespread coverage of the parameter space.

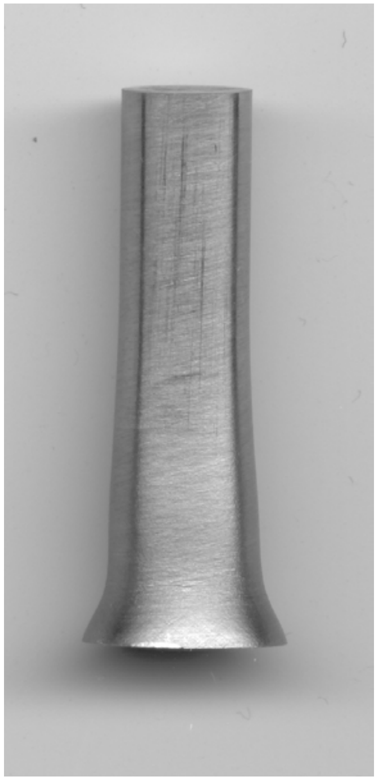
Latin Hypercube Sampling

- The Latin Hypercube Sampling (LHS) enhances Monte Carlo sampling by offering a faster convergence, that is, same accuracy with fewer samples.



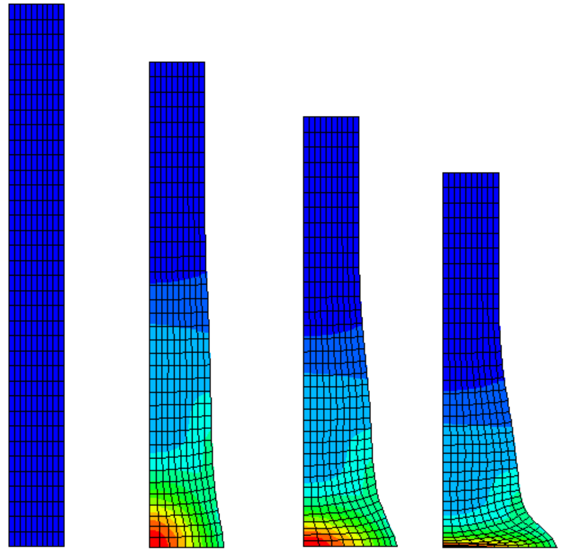
Taylor Anvil Impact

- The high-velocity impact of a sample of material against a rigid wall is simulated numerically.



- The response features (isotropic material) are the initial and final radii *R* and lengths *L* of the impacted sample.
- The coefficients α_1, α_2, α_3, and *N* define the Johnson-Cook model of plasticity:

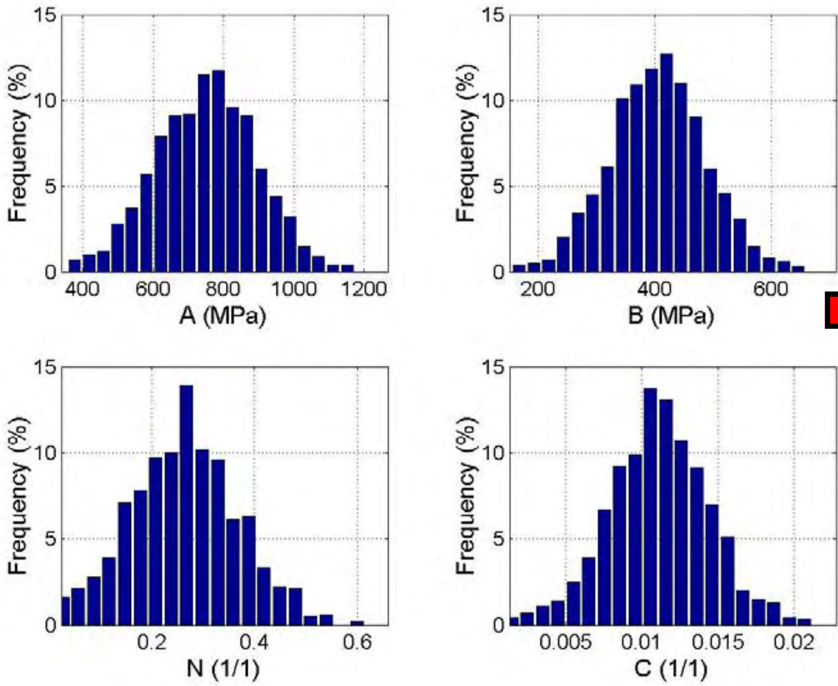
$$\sigma = (\alpha_1 + \alpha_2 \varepsilon_p^N)\left[1 + \alpha_3 \log\left(\frac{\partial \varepsilon_p}{\partial t}\right)\right]$$



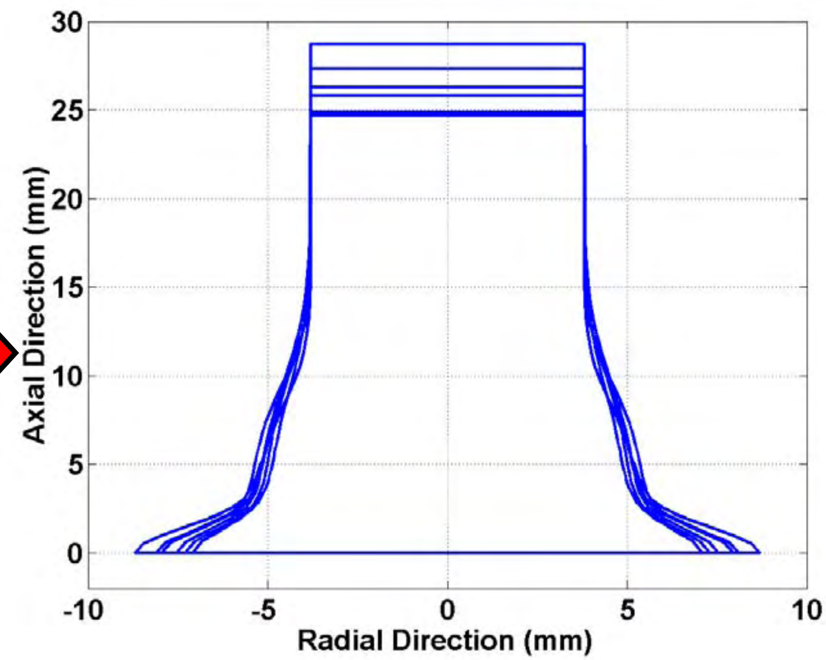
Simulation of the impact and resulting deformation profile.

Propagation of Uncertainty

- **A Monte Carlo simulation is performed by randomly drawing a 1,000 samples of parameters $(\alpha_1;\alpha_2;\alpha_3;N)$.**



Histograms of the sampled values of parameters.

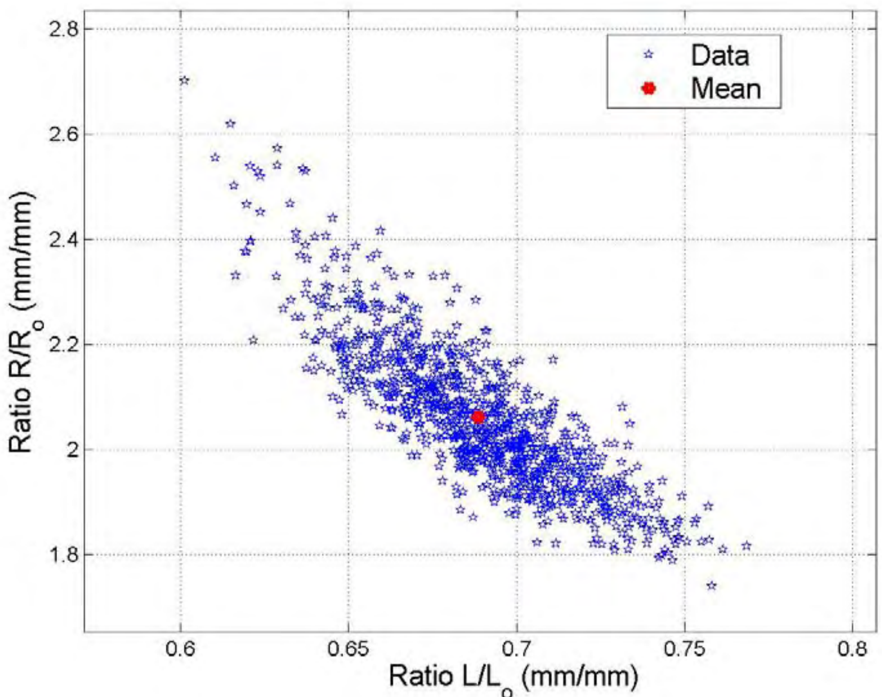


Example of six profiles obtained (from the distribution of 1,000 results).

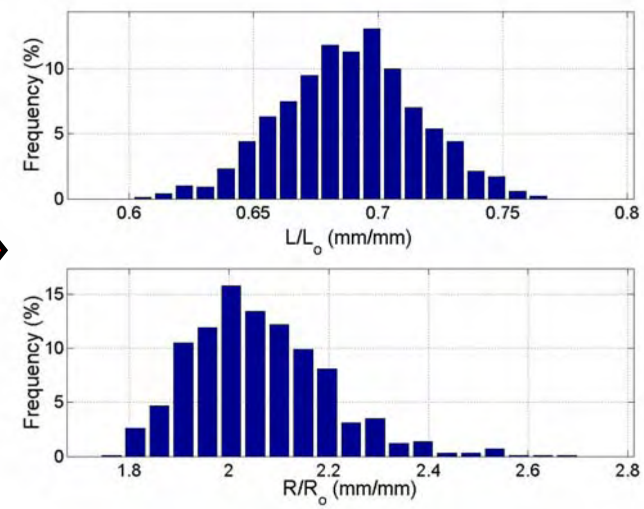
Propagation of Uncertainty

- **The output feature variability is estimated from the Monte Carlo simulation (1,000 runs). Statistics and correlation coefficients can be estimated next.**

Ratio of deformed-to-initial radius.



Ratio of deformed-to-initial length.



Histograms of features.

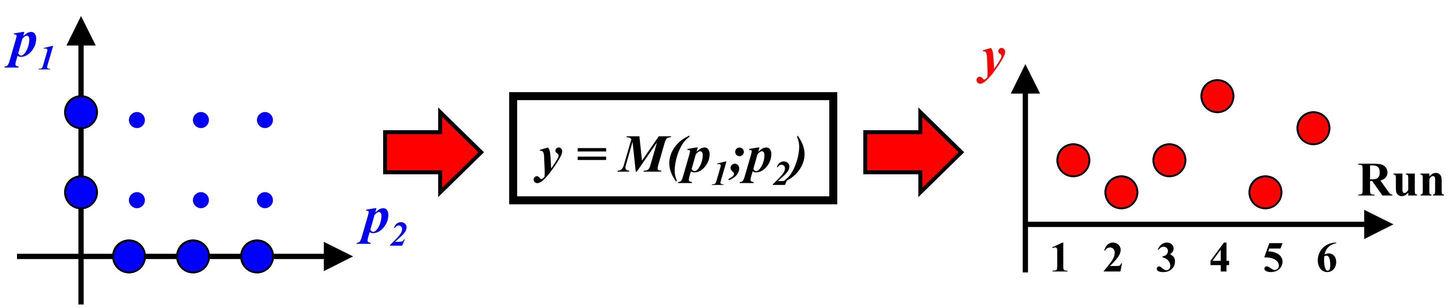
Outline

- What is Uncertainty Quantification?
- Statistical Sampling
- Design of Experiments
- Variance Analysis and Effect Screening
- Meta-modeling and Response Surfaces
- Non-probabilistic Methods

Design of Experiments (DoE)

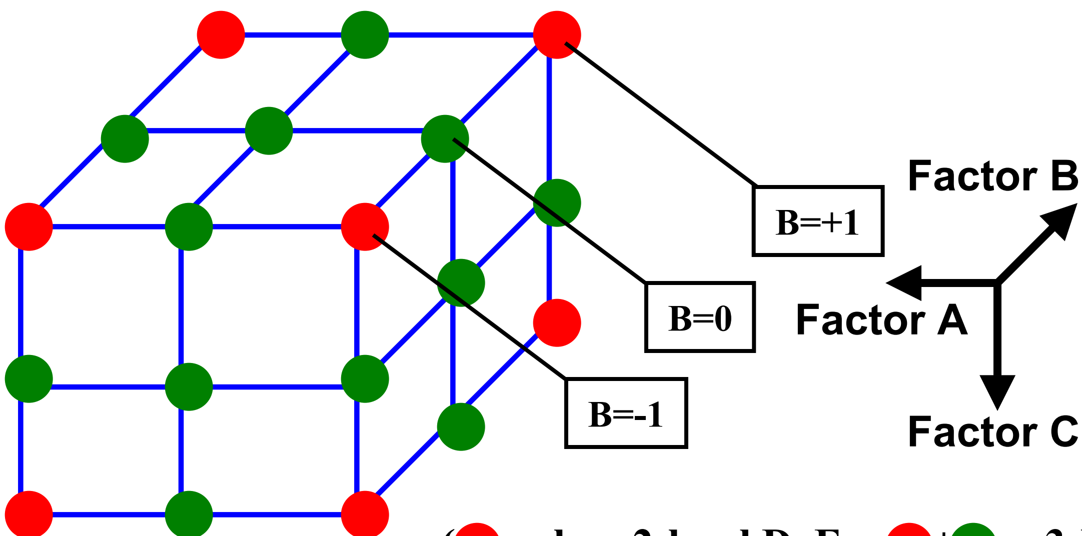
- **A Design of Experiments (DoE) provides another way to propagate uncertainty from the inputs $p_1 \ldots p_N$ of a physical test or numerical simulation to an output *y*. The (joint) probability distribution of $p_1 \ldots p_N$ must be known.**

- **DoE also provides a way to explore the *total variability* of inputs $[p_1^{(min)};p_1^{(max)}] \ldots [p_N^{(min)};p_N^{(max)}]$ and quantify what the effect of this variability is on the output *y*.**



Full-factorial DoE

- A full-factorial Design of Experiments involving N input parameters (or *factors*) p_1 … p_N that can each assume k values (or *levels*), requires k^N runs.
- Illustration of a 3-factor, 3-level DoE:



Computer Run	Factor A	Factor B	Factor C
1	+1	+1	+1
2	+1	0	+1
3	+1	-1	0
4	+1	+1	0
5	+1	0	-1
…	…	…	…
27	-1	-1	-1

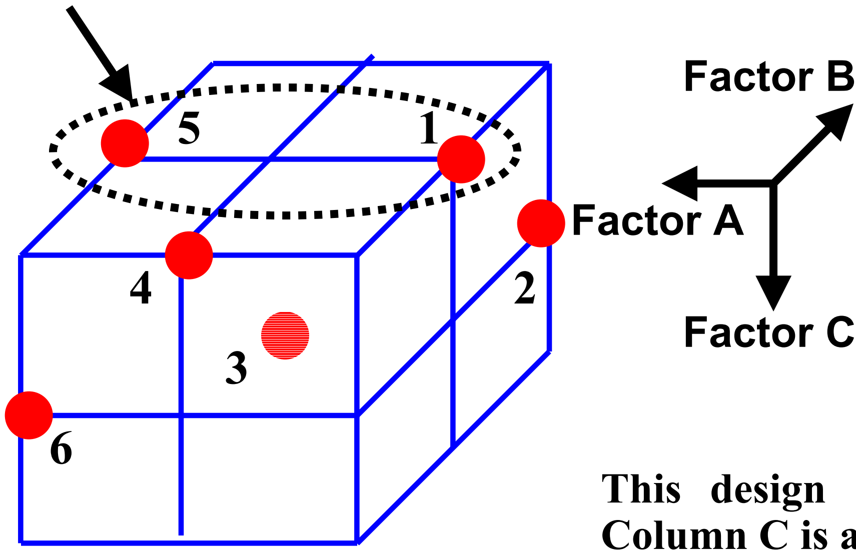
(Note: Values are scaled: Maximum value is "+1"; Minimum value is "-1".)

(● only = 2-level DoE; ●+● = 3-level DoE.)

Fractional Factorial DoE

- **A fractional factorial design is simply a subset of the combinations that populate the full-factorial DoE.**

With runs #1 and #5, the linear effect of the factor A on the output feature *y* can be studied.

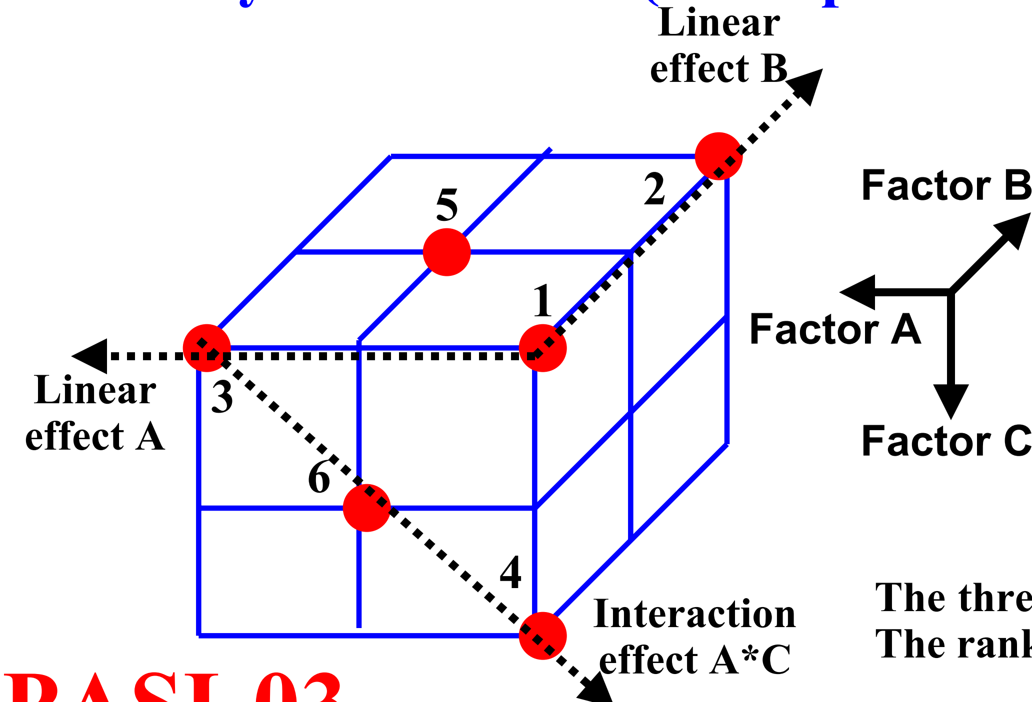


Computer Run	Factor A	Factor B	Factor C
1	-1	0	-1
2	-1	+1	0
3	0	0	0
4	0	-1	-1
5	+1	0	+1
6	+1	-1	0

This design features a very undesirable property. Column C is a linear combination of columns A and B.

Orthogonal Array DoE

- **An orthogonal array is a fractional factorial DoE that guarantees a full-rank design matrix (and guarantees an alias-free DoE). Aliasing refers to the fact that some of the interactions between factors can be compounded by other effects (not represented in the DoE).**



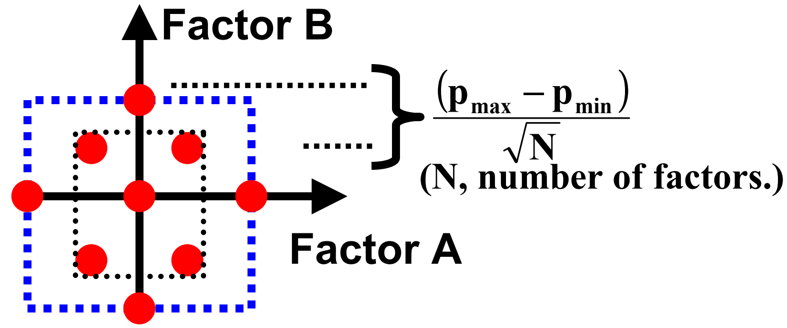
Computer Run	Factor A	Factor B	Factor C
1	-1	-1	-1
2	-1	+1	-1
3	+1	-1	-1
4	-1	-1	+1
5	0	0	-1
6	0	-1	0

The three columns are orthogonal to each other.
The rank of the design matrix is equal to 3.

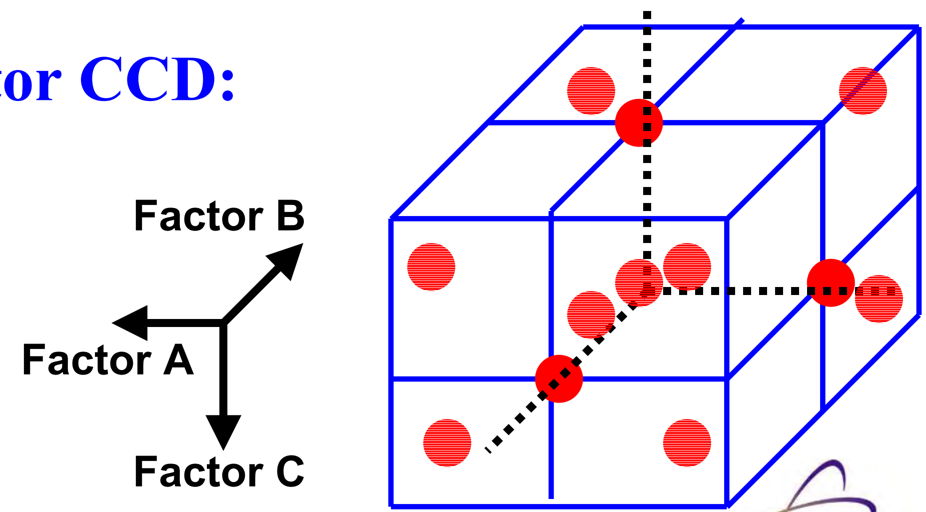
Central Composite DoE

- Central Composite Designs (CCD) distribute points at the design space's boundary and analyze the linear and quadratic effects.

- Illustration of 2-factor CCD:

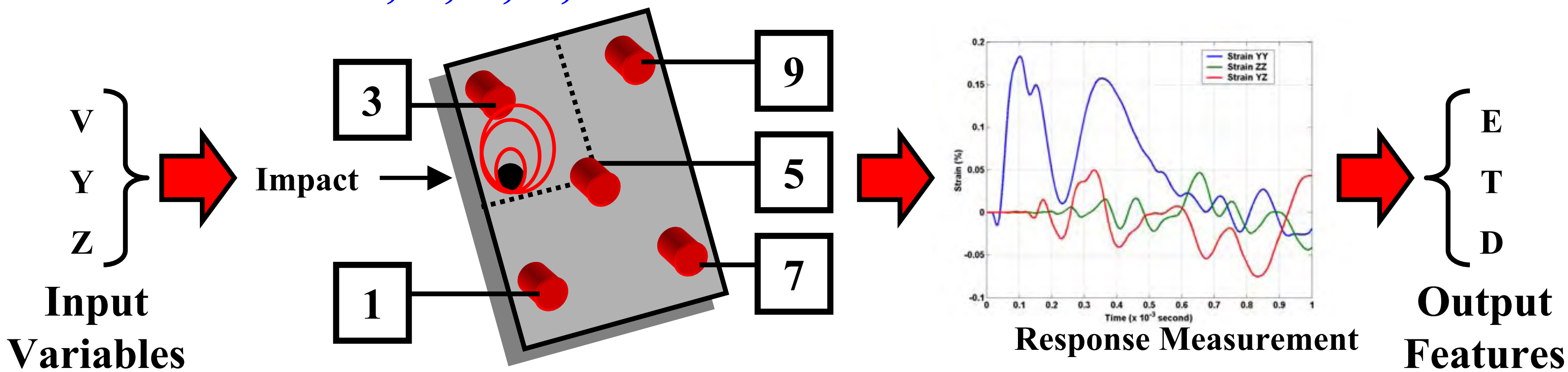


- Illustration of 3-factor CCD:



Composite Plate Impact

- **The response of a composite plate to the impact of a projectile is simulated using a finite element model. The three input variables are *V*, *Y*, and *Z*. The output features *y* are the temporal moments *E*, *T*, *D* at locations 1, 3, 5, 7, and 9.**



- **Two designs are considered: Central composite design (CCD-20 with 20 computer runs); Taguchi orthogonal array (TOA-25 with 25 computer runs).**

Normalization

- **The three inputs or "factors" *V*, *Y*, *Z* are normalized between –1.0 and +1.0 for dimensionless analysis.**

- **CCD design:**

Factor	Description	-1.0 Level	+1.0 Level
1	Velocity (V)	30.0 m/s	80.0 m/s
2	Y-location (Y)	2.4 inch	8.4 inch
3	Z-location (Z)	2.4 inch	8.4 inch

- **Taguchi design:**

Factor	Description	-1.0 Level	+1.0 Level
1	Velocity (V)	10.0 m/s	100.0 m/s
2	Y-location (Y)	0.0 inch	10.8 inch
3	Z-location (Z)	0.0 inch	10.8 inch

$$x^{(n)} = \frac{\left(2x - x_{\max} - x_{\min}\right)}{\left(x_{\max} - x_{\min}\right)}$$

Equation for Normalization

Three-factor, Five-level CCD-20

Computer Run	Factor 1 (V)	Factor 2 (Y)	Factor 3 (Z)
1	0.00	+1.80	0.00
2	0.00	0.00	0.00
3	0.00	0.00	0.00
4	-1.80	0.00	0.00
5	+1.00	-1.00	+1.00
6	0.00	-1.80	0.00
7	+1.80	0.00	0.00
8	+1.00	+1.00	+1.00
9	0.00	0.00	0.00
10	0.00	0.00	+1.80

Computer Run	Factor 1 (V)	Factor 2 (Y)	Factor 3 (Z)
11	-1.00	-1.00	-1.00
12	0.00	0.00	-1.80
13	0.00	0.00	0.00
14	0.00	0.00	0.00
15	0.00	0.00	0.00
16	-1.00	+1.00	-1.00
17	+1.00	+1.00	-1.00
18	-1.00	+1.00	+1.00
19	+1.00	-1.00	-1.00
20	-1.00	-1.00	+1.00

(Coded levels are shown.)

Three-factor, Five-level TOA-25

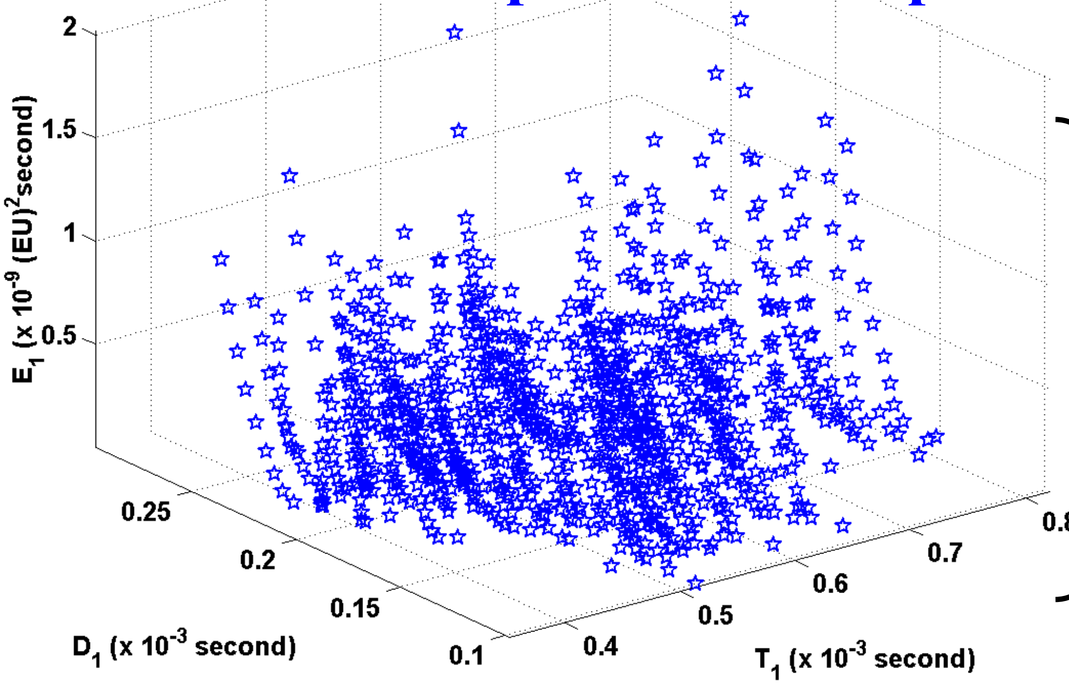
Computer Run	Factor 1 (V)	Factor 2 (Y)	Factor 3 (Z)
1	-1.00	-0.11	-0.11
2	-0.11	+0.33	-1.00
3	+0.78	+0.33	-0.11
4	+0.78	+0.78	+0.33
5	-1.00	-1.00	-1.00
6	-1.00	+0.33	+0.33
7	-0.56	-1.00	-0.56
8	+0.78	-1.00	+0.78
9	+0.33	-1.00	+0.33
10	-0.56	+0.78	-1.00
11	+0.33	+0.33	-0.56
12	-0.11	-0.56	+0.33
13	+0.33	+0.78	-0.11

Computer Run	Factor 1 (V)	Factor 2 (Y)	Factor 3 (Z)
14	+0.78	-0.11	-0.56
15	-1.00	+0.78	+0.78
16	+0.78	-0.56	-1.00
17	-0.11	+0.78	-0.56
18	+0.33	-0.11	-1.00
19	-0.56	-0.56	-0.11
20	-0.56	+0.33	+0.78
21	+0.33	-0.56	+0.78
22	-0.11	-1.00	-0.11
23	-0.11	-0.11	+0.78
24	-1.00	-0.56	-0.56
25	-0.56	-0.11	+0.33

(Coded levels are shown.)

Total Variability

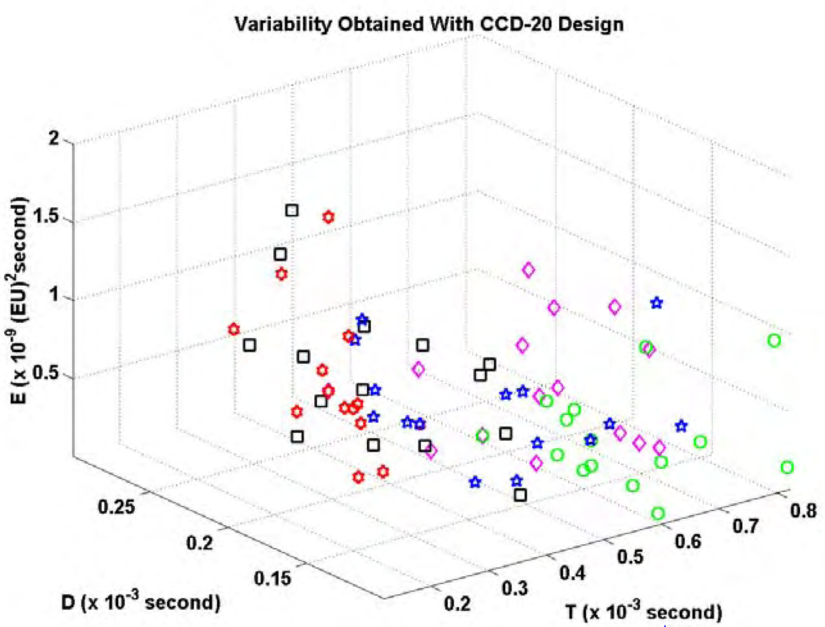
- **Varying the three factors *V*, *Y*, *Z* "spreads" the output features *E*, *T*, *D*. A pre-requisite to study the input-output relationship is to assess which combinations of factors are responsible for explaining this variability.**

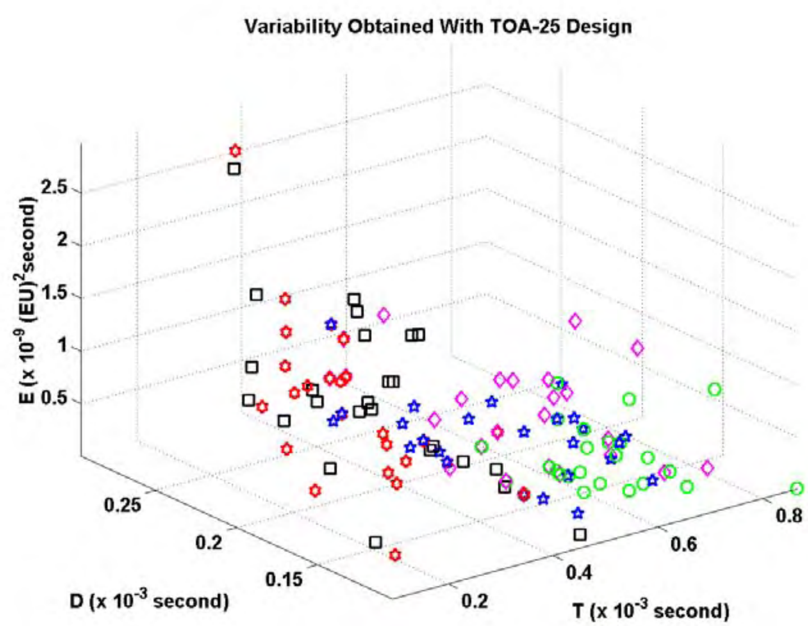


"Spread" of output features obtained from the 1,000 computer simulations (3-factor, 10-level full factorial DoE) at sensor 1.

CCD-20 & TOA-25 Variability

- **Variability of output features *E*, *T*, *D* obtained at sensors 1, 3, 5, 7, and 9 with the CCD-20 design (left) and TOA-25 design (right).**





Sensors: 1 3 5 7 9

Outline

- What is Uncertainty Quantification?
- Statistical Sampling
- Design of Experiments
- **Variance Analysis and Effect Screening**
- Meta-modeling and Response Surfaces
- Non-probabilistic Methods

Local Sensitivity Analysis

- **Sensitivity analysis studies the effect that a change of input parameter *p* has on the output feature *y*.**

- **The conventional approach in engineering sciences is to calculate the gradient of *y* with respect to *p*, *dy/dp*. Finite differences can be used, for example:**

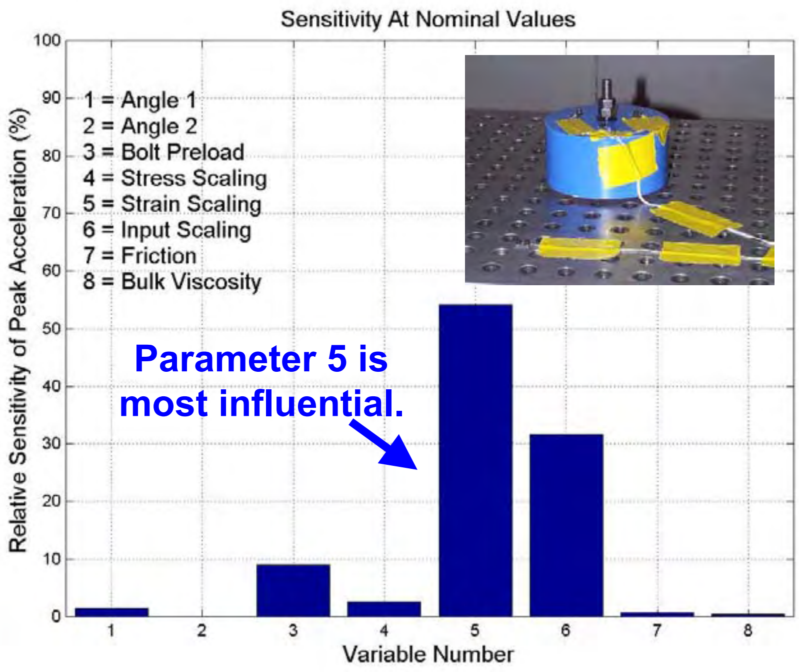
Central Differences: $$\left(\frac{\partial y}{\partial p}\right)_{p=p_o} = \frac{y(p_o+\Delta p)-y(p_o-\Delta p)}{2\Delta p} + O\left(\Delta p^2\right)$$

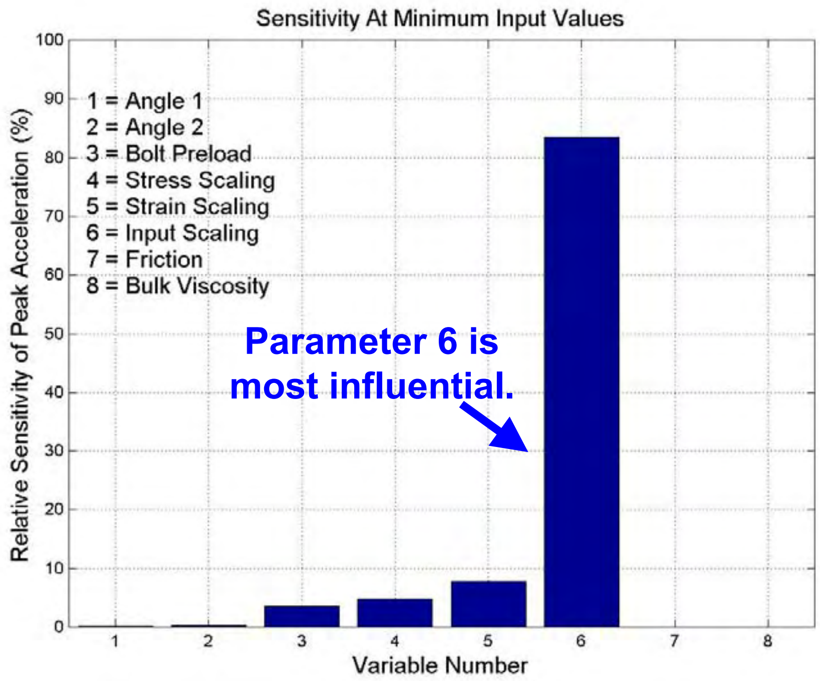
- **Be careful about convergence! How to select Δp?**

- **Generally, Δp should be "small" compared p_o. It is common practice to chose $\Delta p = p_o/10$ to $\Delta p = p_o/100$.**

Local Sensitivity Analysis

- **Finite difference-based derivatives only provide local information, about one specific point p_o in the design space, and in one specific direction.**





The Cost of Estimating Derivatives

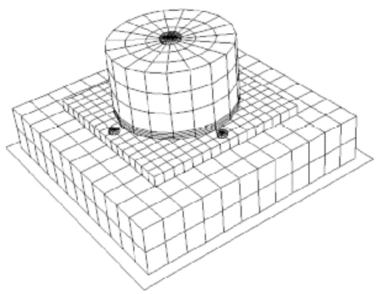
- **Estimating finite differences can be expensive if one wants to do it right (meaning, satisfactory accuracy).**
- **For example, to solve an optimization problem of *N=6* variables with the conjugate gradient algorithm, a total of *N* iterations *x (1+2N)* function evaluations-per-iteration is required. That's 6 x 13 = 78 FE runs.**
- **Often, such cost is comparable, even superior, to the cost of performing a DoE and fitting a meta-model. For example, a 6-variable, 3-level Taguchi array can be analyzed with as little as 27 runs. Computations such as sensitivity analysis, parameter optimization, etc., are cheap once the meta-model is obtained.**

Global Variability

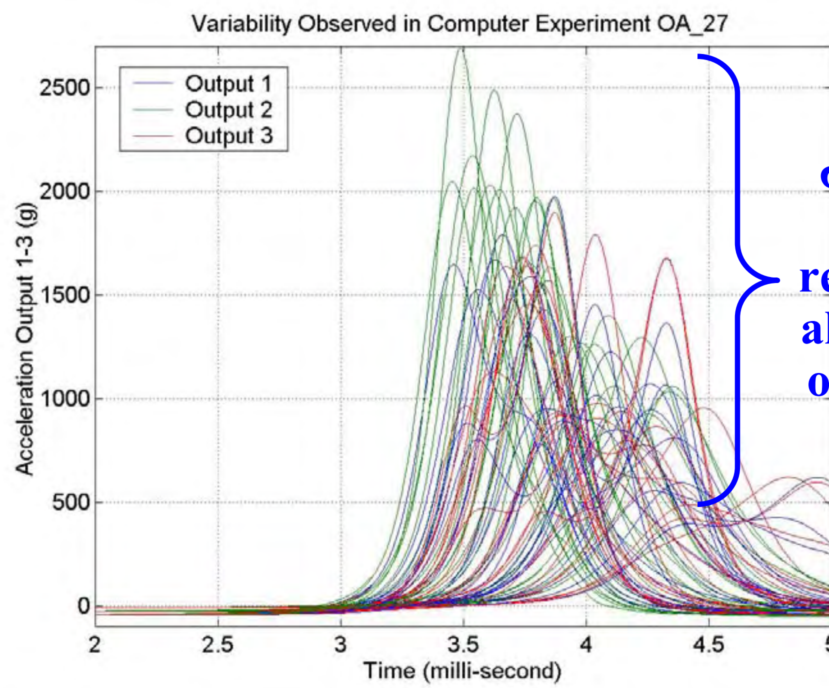
- **Global sensitivity analysis addresses a different problematic than gradient calculation. The problem is two-fold: Understand the input-output relationship and assess the effects of variability of inputs *p* on outputs *y*.**

Inputs

p_1
p_2
p_3
p_4
p_5
p_6
p_7
p_8



Finite Element Modeling

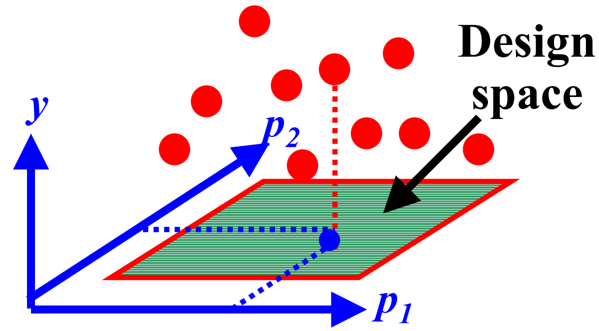


σ(*y*), Output variability resulting from all the sources of variability.

Analysis of Variance (ANOVA)

- **Numerical model:**

$$y = M(p_1; p_2)$$



- **Does knowing p_1 reduce the variability?**

$$\sigma^2(\hat{y}) + E[\sigma^2(y|\, p_1)] = \sigma^2(y)$$

Variability of $\hat{y} = E[y|\, p_1]$

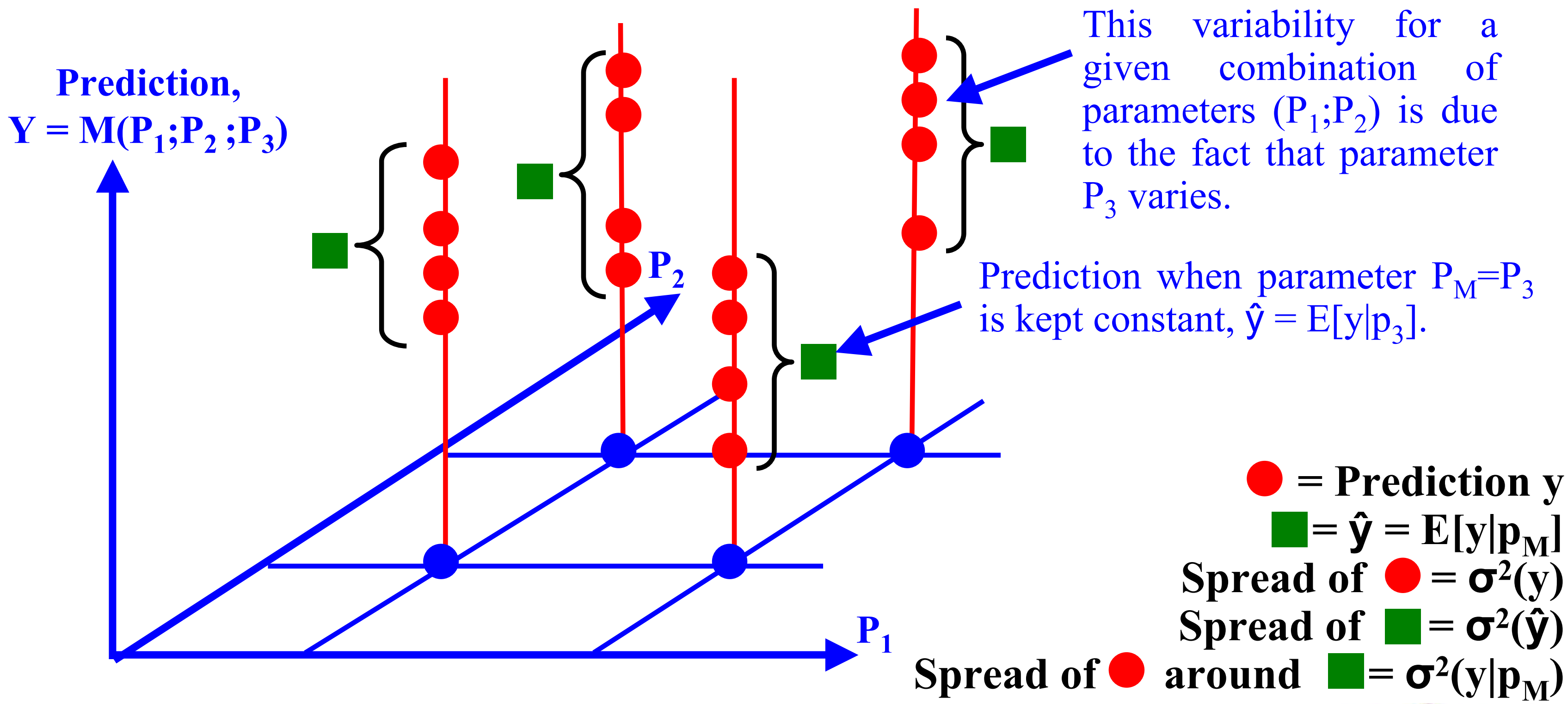
Importance of input parameter p_1 to explain the total variability of $y^{(\#)}$

Total variability

(#)The importance of the input parameter p_1 is indicated by the difference $\sigma^2(y)$-$\sigma^2(\hat{y})$ or the correlation ratio $\eta^2 = \sigma^2(\hat{y})/\sigma^2(y)$.

Analysis of Variance (ANOVA)

- **The figure represents the equation $\sigma^2(\hat{y}) + E[\sigma^2(y|p_M)] = \sigma^2(y)$. Simply speaking, an ANOVA compares the spread of green squares to the spread of red circles:**

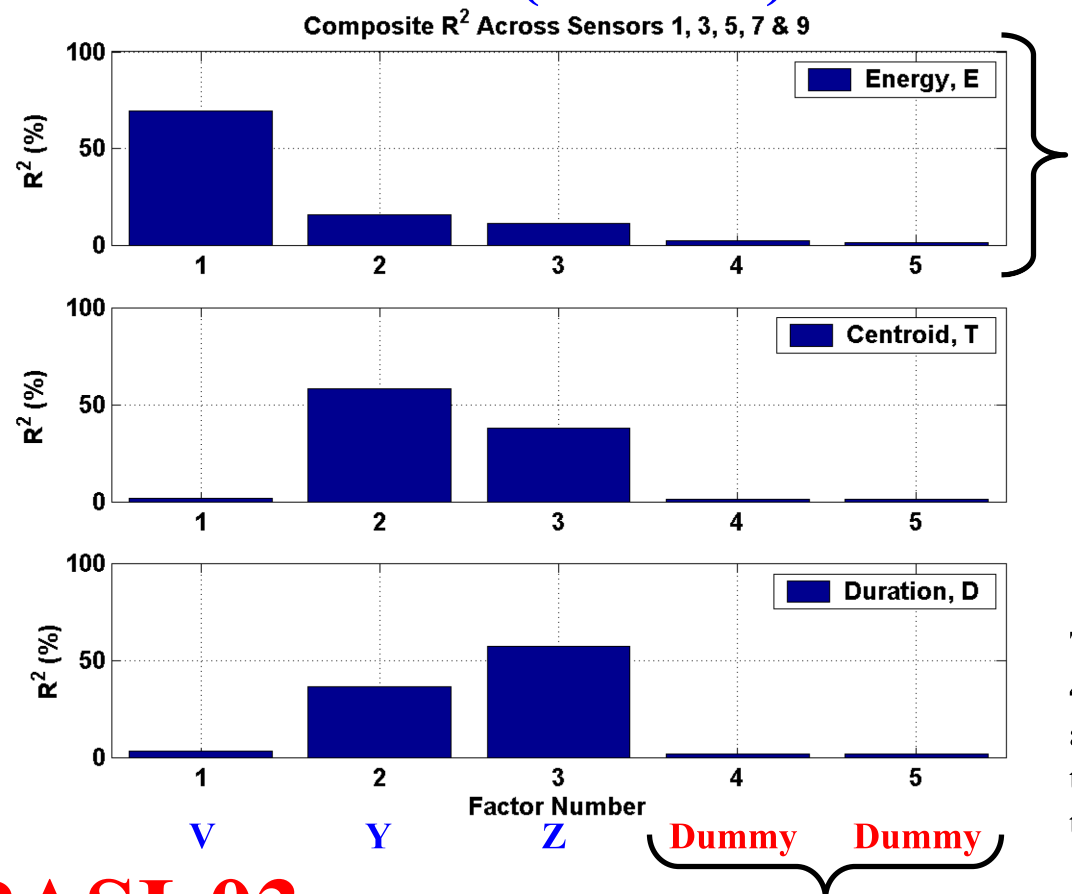


Analysis of Variance (ANOVA)

- Assume that the factors {p} of the model y=M(p) are partitioned in two subsets $\{p_M\}$ and $\{p_O\}$. Assume that only the factors $\{p_O\}$ vary. The factors $\{p_M\}$ are known, they do not vary.
  - The prediction becomes $\hat{y}=E[y|p_M]$. The quantity $E[y|p_M]$ is called a "conditional expectation" because the model's output feature y is conditioned on the knowledge of input factors p_M.
- How does knowing p_M reduce the variability?
  - Obviously, if $\hat{y}$ and y both exhibit the same variability, then knowing the factors p_M does not help to reduce the variability. If that's the case, the factors p_M might as well be kept constant and equal to their nominal values.
- How to measure the importance of factors p_M?
  - The importance of p_M is indicated by the difference $\sigma^2(y)-\sigma^2(\hat{y})$ which is equal to $E[\sigma^2(y|p_M)]$. The correlation ratio $\eta^2 = \sigma^2(\hat{y})/\sigma^2(y)$ can equivalently be estimated with an ANOVA technique.

ANOVA Based on FF-1000

- **The composite plate impact problem is analyzed using a full-factorial (FF-1000) with 10^3 = 1,000 runs.**

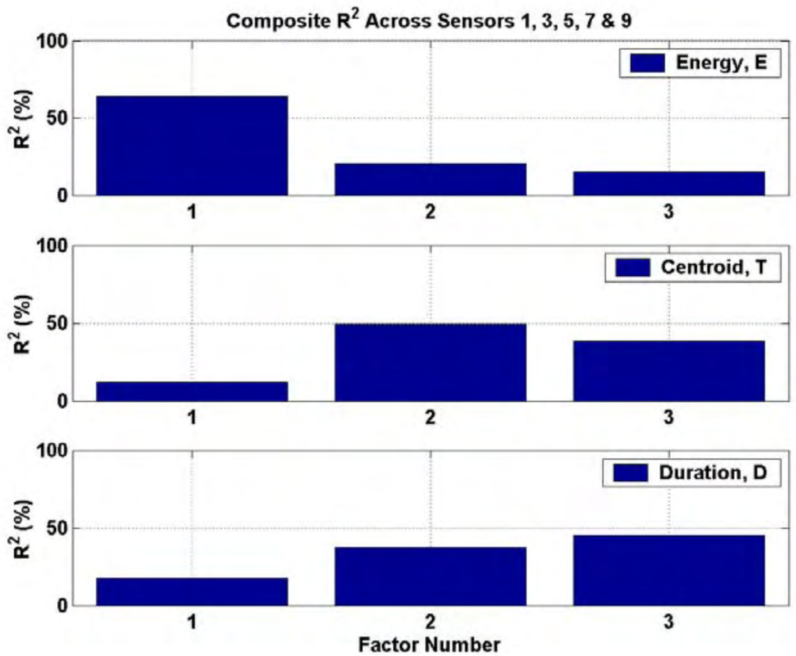


This result indicates that the impact velocity *V* is the most significant factor for explaining the variability of the energy feature *E* across sensors 1, 3, 5, 7 and 9.

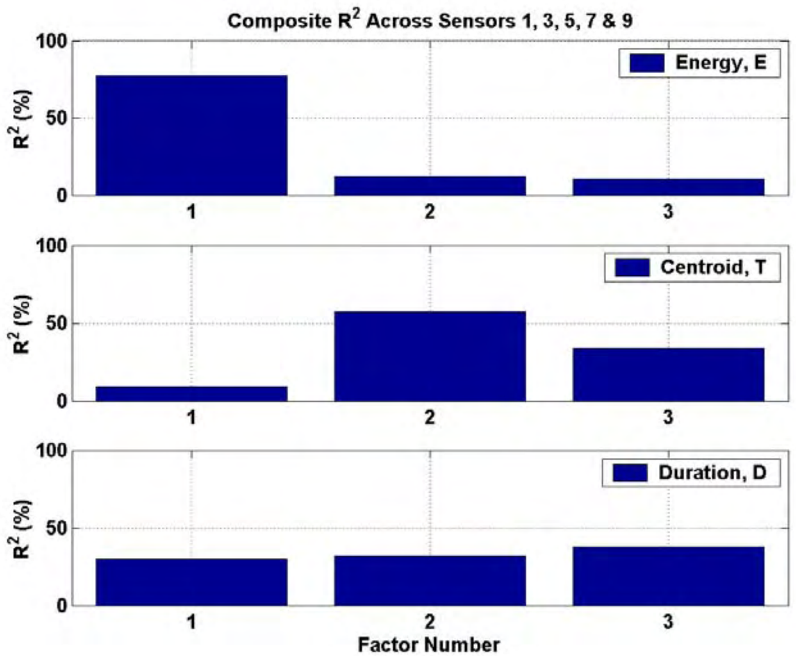
Two "dummy" variables (labeled factors 4 & 5) are added as a sanity check. They are initialized at random and, therefore, they should not be significant to explain the variability of the features.

ANOVA Based on CCD-20 & TOA-25

- The CCD-20 design (left) and TOA-25 design (right) provide the same significance factors as the full-factorial analysis, using 40 to 67 times less computer runs than the 1,000-run full-factorial analysis.



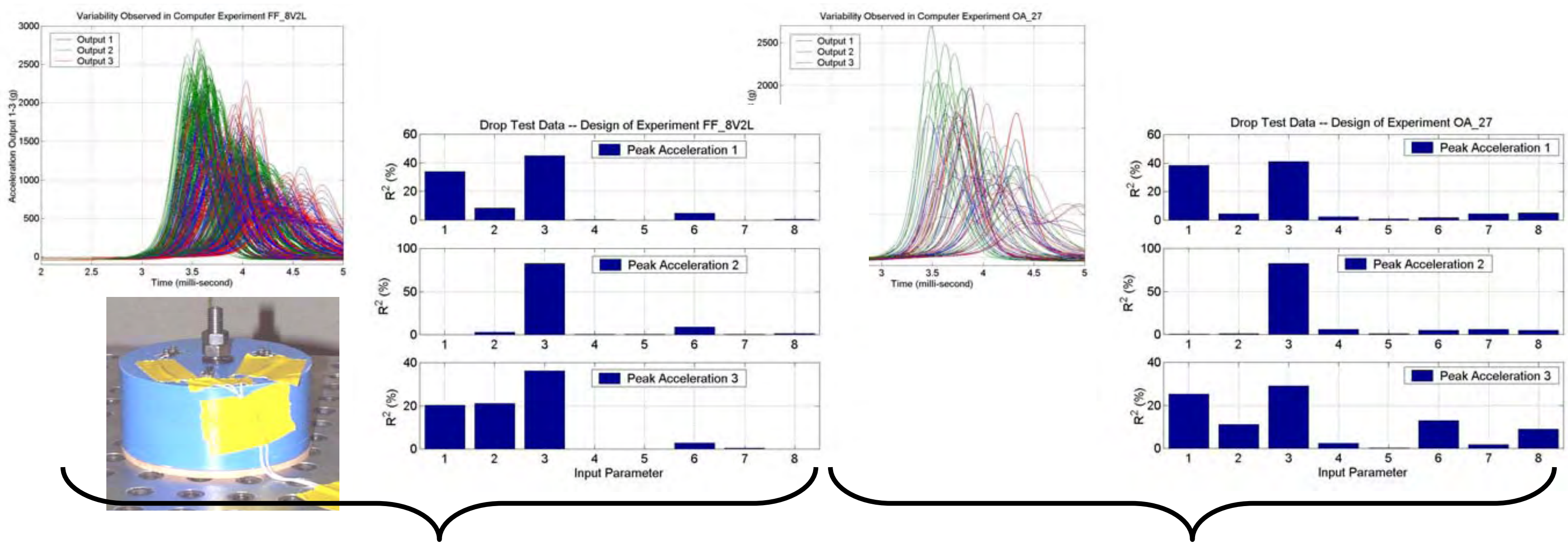
Central Composite Design, CCD-20



Taguchi Orthogonal Design, TOA-25

Effect Screening

- **Effect screening refers to the identification of those effects (*A*, *B*, *A*B*, A^2, A^3*B^2, etc.) which are most responsible for explaining the total output variability.**

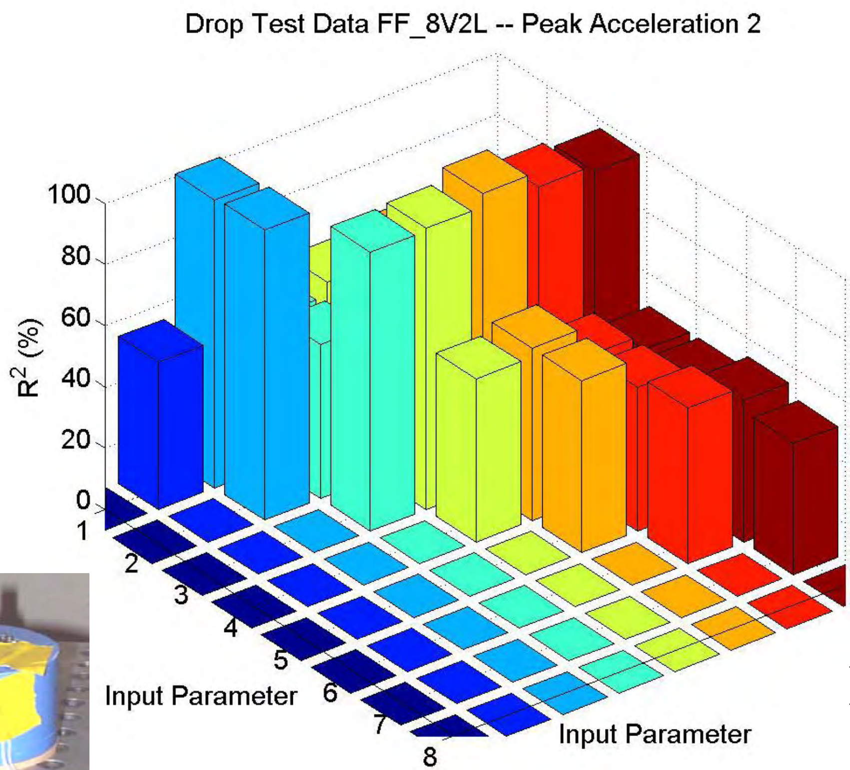


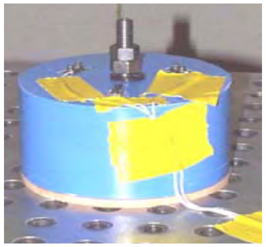
8-factor, 2-level full-factorial analysis (256 finite element evaluations)

Taguchi orthogonal array (27 finite element evaluations)

2^{nd} Order Effect Screening

- 2^{nd} order effect screening refers to the identification of interaction and quadratic effects: *A*B*, *A*B*C*, *A^2*.





- The R^2 statistic estimates the correlation ratio of an analysis of variance.
- It can be estimated for linear or higher-order effect screening, if enough data are available.

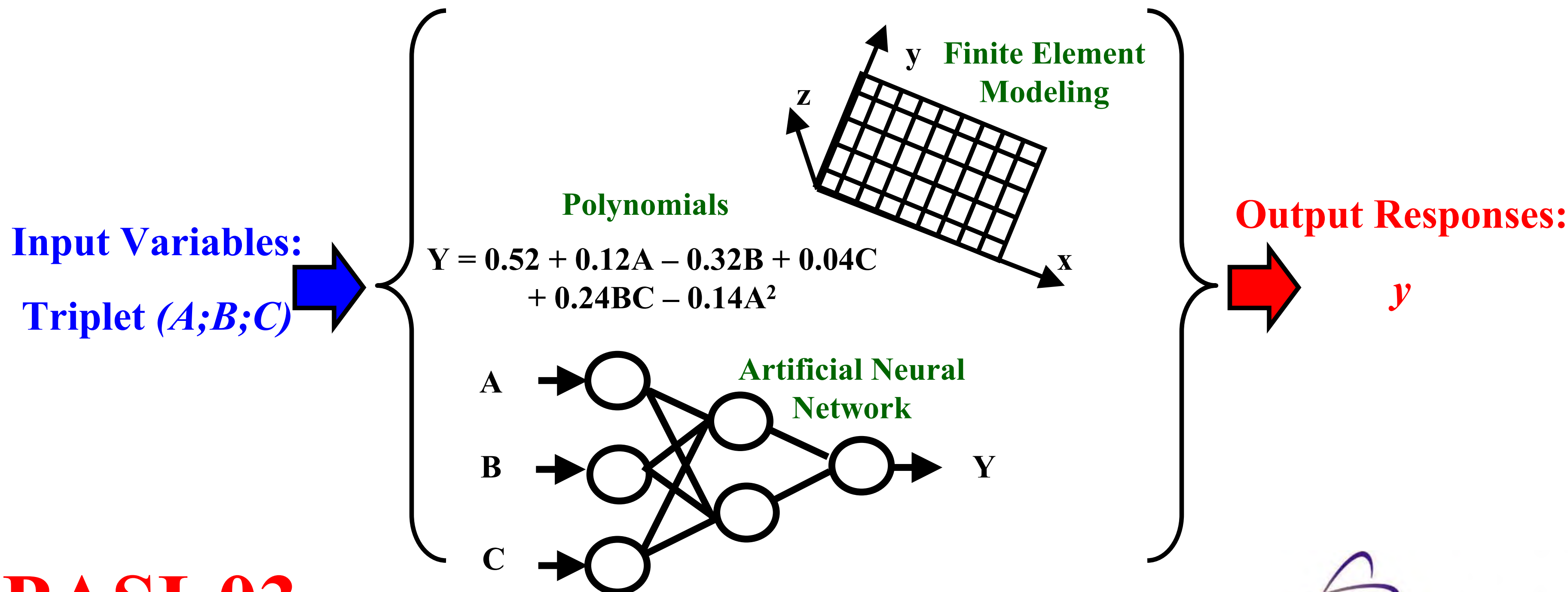
$$R^2 = 1 - \frac{\sum_{l=1 \cdots N_{level}} \sum_{j=1 \cdots N_{data}^{(l)}} \left(y_j^{(l)} - \bar{y}^{(l)} \right)^2}{\sum_{j=1 \cdots N_{data}} \left(y_j - \bar{y} \right)^2}$$

Outline

- What is Uncertainty Quantification?
- Statistical Sampling
- Design of Experiments
- Variance Analysis and Effect Screening
- **Meta-modeling and Response Surfaces**
- Non-probabilistic Methods

Meta-modeling

- A ***meta-model*** (a.k.a. surrogate, response surface, fast-running) captures the ***input-output relationship*** over a domain of interest and can be evaluated at a fraction of the cost of the "physics-based" numerical model.



Why Develop a Meta-model?

- **A meta-model is, by definition, cheap to evaluate.**
- **It focuses on the input-output relationship, without including the details about the geometry, topology, materials, loading, etc.**
- **Meta-models can be used for estimating gradients (local derivatives), uncertainty propagation, parameter calibration, etc.**
- **Developing a meta-model through effect screening and with the help of a design of experiments can be much cheaper than calculating the derivatives.**

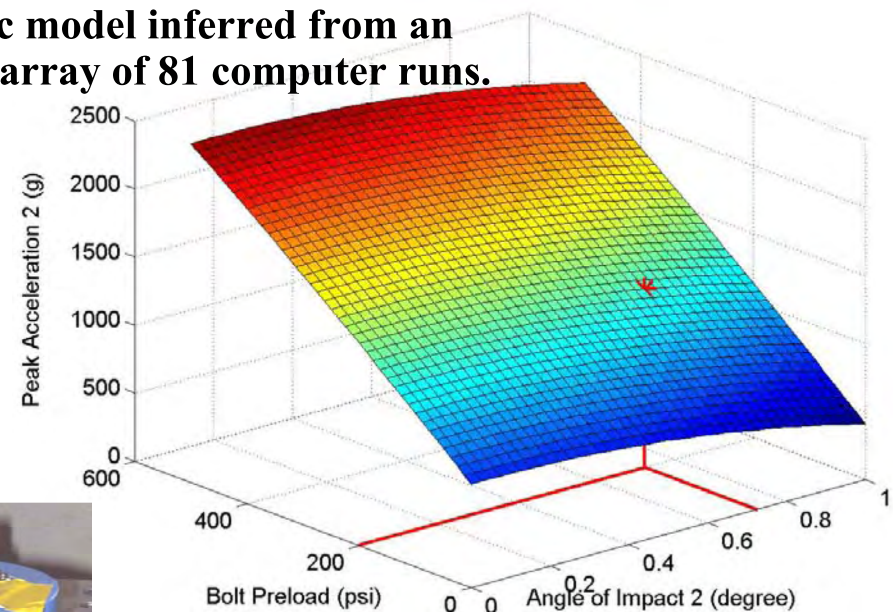
$$\ddot{x}_2^{peak} = \begin{Bmatrix} -1{,}538.2 \\ 43.6 \\ 288.4 \\ 2.4 \\ 2{,}552.8 \\ -391.3 \\ -307.1 \\ -0.0006 \\ 665.7 \\ -0.5 \\ -452.4 \\ 1.5 \end{Bmatrix}^T \begin{Bmatrix} 1 \\ a_1 \\ a_2 \\ P_{bolt} \\ s_I \\ a_1^2 \\ a_2^2 \\ P_{bolt}^2 \\ a_1 * a_2 \\ a_2 * P_{bolt} \\ a_2 * s_I \\ P_{bolt} * s_I \end{Bmatrix}$$

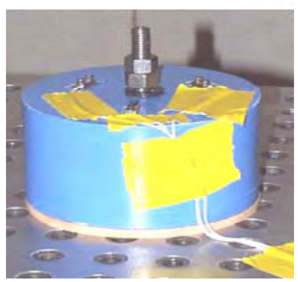
Polynomial Surrogates

- **Once the effects that best explain the variability have been identified, the coefficients of a polynomial that best-fits the data can be easily computed.**

$$\ddot{x}_2^{peak} = \begin{Bmatrix} -1{,}538.2 \\ 43.6 \\ 288.4 \\ 2.4 \\ 2{,}552.8 \\ -391.3 \\ -307.1 \\ -0.0006 \\ 665.7 \\ -0.5 \\ -452.4 \\ 1.5 \end{Bmatrix}^T \begin{Bmatrix} 1 \\ a_1 \\ a_2 \\ P_{bolt} \\ s_I \\ a_1^2 \\ a_2^2 \\ P_{bolt}^2 \\ a_1 * a_2 \\ a_2 * P_{bolt} \\ a_2 * s_I \\ P_{bolt} * s_I \end{Bmatrix}$$

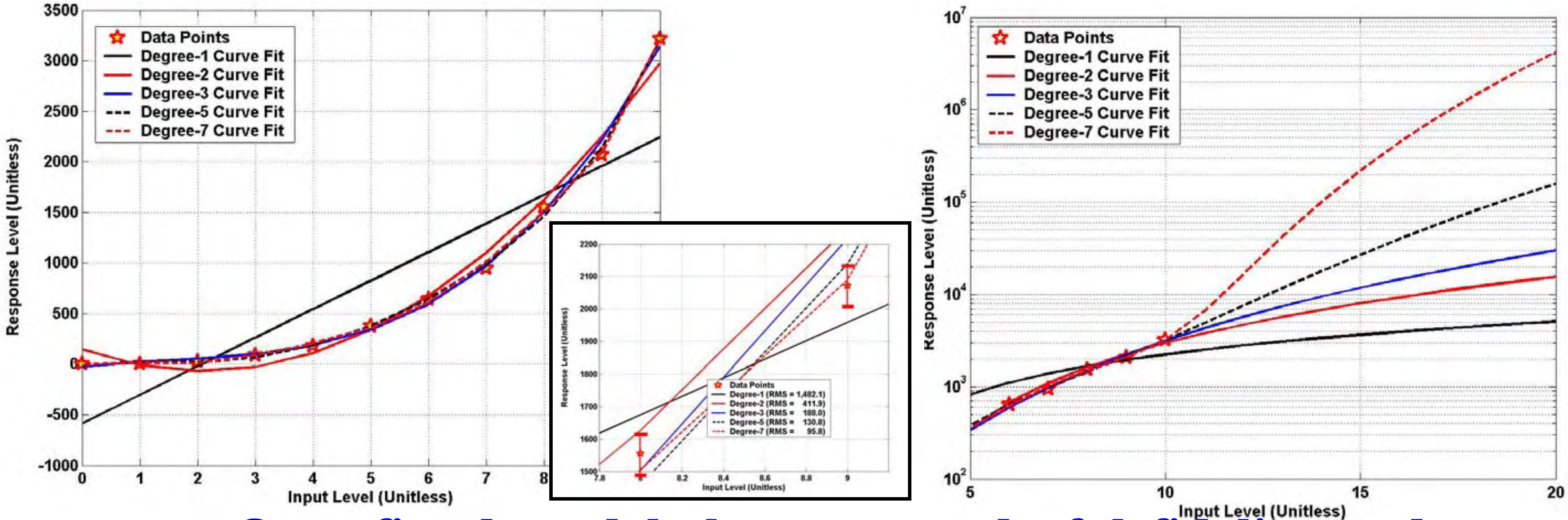
Quadratic model inferred from an orthogonal array of 81 computer runs.





The Dangers of Over-fitting

- **The higher the degree of the polynomial, the better it will tend to fit the data. This is called *over-fitting*.**



- **Over-fitted models have a wonderful fidelity-to-data but they make horrible predictions (extrapolation). Goodness-of-fit indicators are available to detect over-fitting, like the MSE divided by the # of effects.**

Verifying Effect Screening and Meta-models

- **How to verify that the "right" variables have been screened, that the most significant effects have been retained, that the meta-model has the right form?**

- **After the most significant effects have been identified, a model that includes them can be fitted to the data. The Mean Square Error (MSE) indicator assesses the *goodness-of-fit* to data.**

$$MSE(y) = \frac{100}{N\sigma_y^2} \sum_i (y_i - \hat{y}_i)^2$$

- **The contribution of each effect can be assessed with an importance factor.**

$$s_\theta = 100 \frac{\sigma_\theta^2}{\sigma_Y^2}$$

(σ_θ^2 is the variance of the effect considered, and σ_Y^2 is the total variance).

Outline

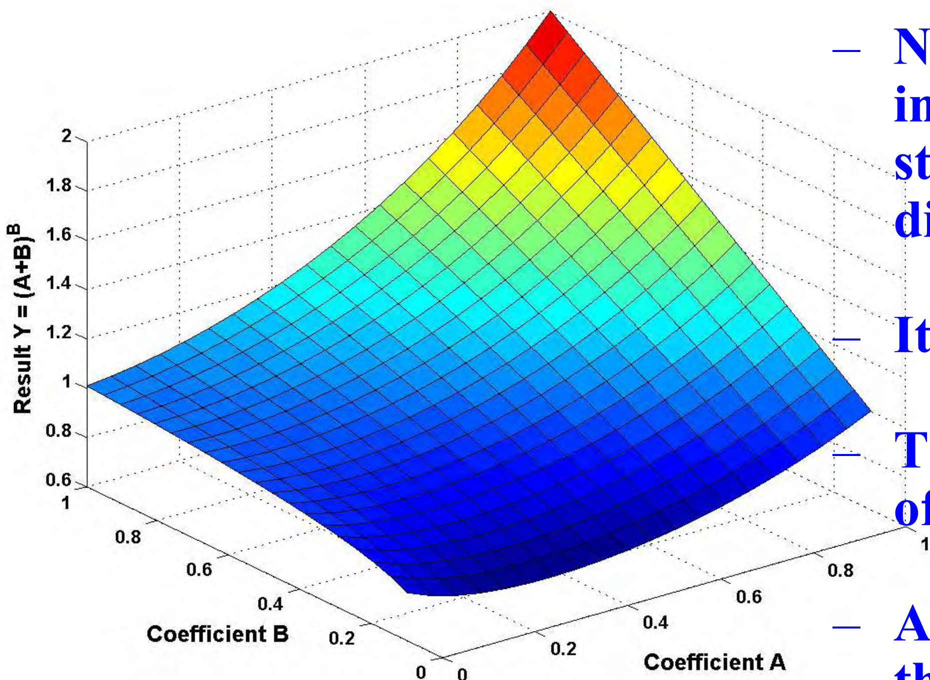
- What is Uncertainty Quantification?
- Statistical Sampling
- Design of Experiments
- Variance Analysis and Effect Screening
- Meta-modeling and Response Surfaces
- **Non-probabilistic Methods**

Why Non-probabilistic Uncertainty?

- **It may be difficult to represent the uncertainty associated to certain type of information using the theory of probability.**
- **Examples are the epistemic uncertainty (lack-of-knowledge), expert judgment, rare events (large, catastrophic Earthquakes), vagueness (*"It is warm."*).**
- **Assuming more than is *really* known can lead to wrong conclusion.**
- **Assuming a uniform probability density function (or another PDF) might be more than is really known. Beware of the so-called "principle of indifference".**

A Simple Example

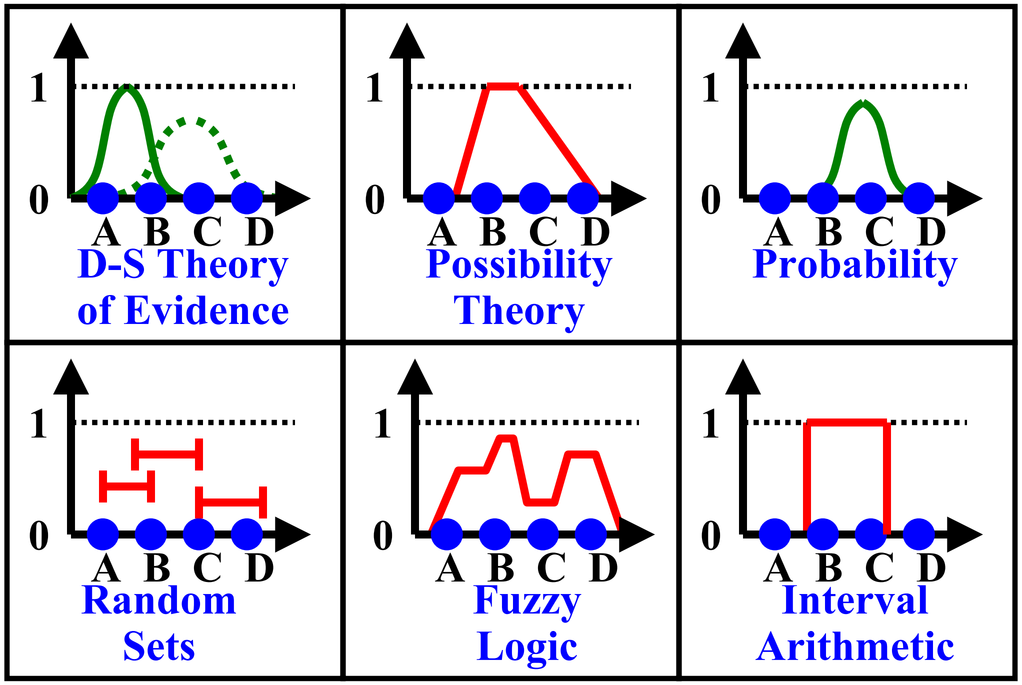
- *"The quantity y is given by $y = (a+b)^b$, where a is between [0;1] and b is between [0.1;1]. What can be said about y exceeding the critical value of y_C=1.8?"*



- **No statistical distribution can be inferred from the above problem statement, not even the uniform distribution.**
- **It is not even stated that *a*, *b* vary!**
- **The area where $y>y_C$ is only 1.04% of the total area *[0;1]x[0.1;1]*.**
- **A 10^{+3}-run Monte Carlo simulation that assumes uniform distributions for *a*, *b* returns *Prob[$y>y_C$] = 2.10%*.**

General Information Theory

- Each source of uncertainty should be represented with the most appropriate mathematical theory.
- The difficulty is to then combine them and propagate the non-probabilistic uncertainty through the code.



Interval Arithmetic

- **Uncertainty can be represented as an interval when very few data points are available.**
  - ***"Three measurements of X provide the values 1.5, 1.2, and 1.6"*****. All that can be said,** ***based on the current evidence*****, be is that** ***X*** **belongs to the interval** ***[1.2;1.6]*****, or that X belongs to an interval of unknown size, centered somewhere between 1.2 and 1.6.**

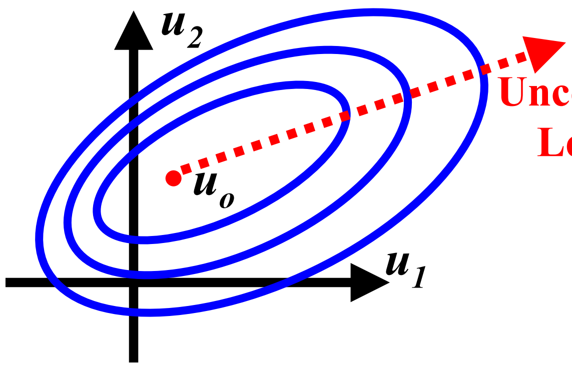
- **Propagating intervals is easy for simple operations.**
  - $A*B$ **in** $[A_{MIN}*B_{MIN};A_{MAX}*B_{MAX}]$, A/B **in** $[A_{MIN}/B_{MAX};A_{MAX}/B_{MIN}]$ **for positive numbers, etc.**

- **Propagating intervals through "black-box" codes requires the resolution of—computationally expensive—minimum and maximum optimization problems.**

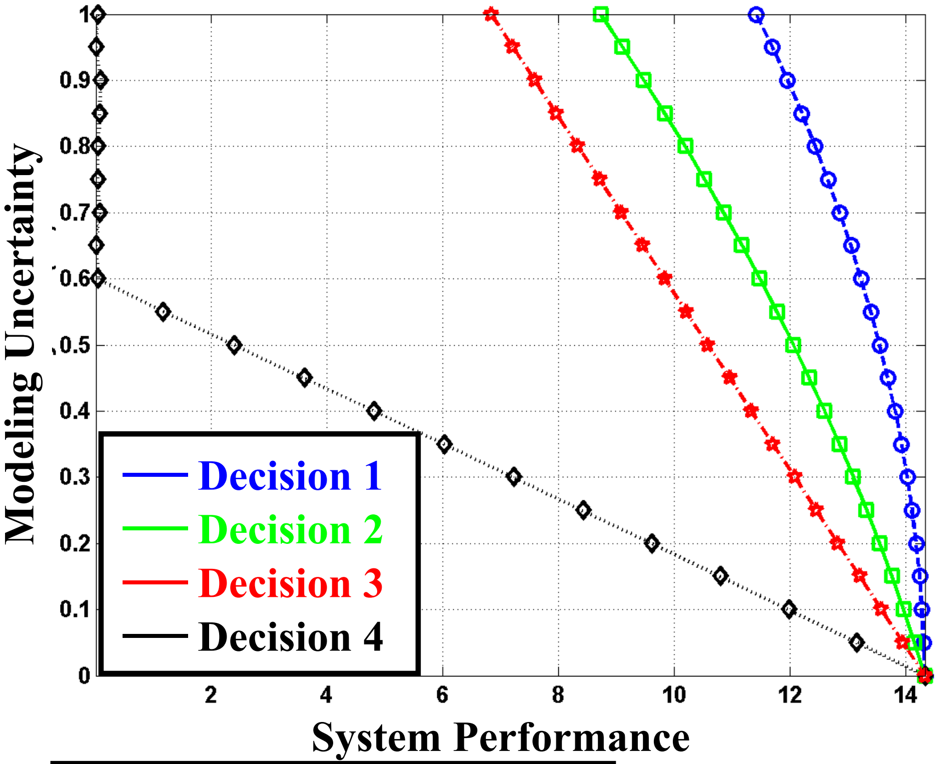
Convex Models of Information-gap

- **Information-gap theory models the *gap of knowledge* between what is currently known and what must be known to make a decision.**

Convex models are convenient for computational purposes.



In the information-gap theory, the best decision maximizes the *robustness-to-uncertainty*, not the performance.



(Note: Conceptual illustration, but the numbers come from a real analysis.)

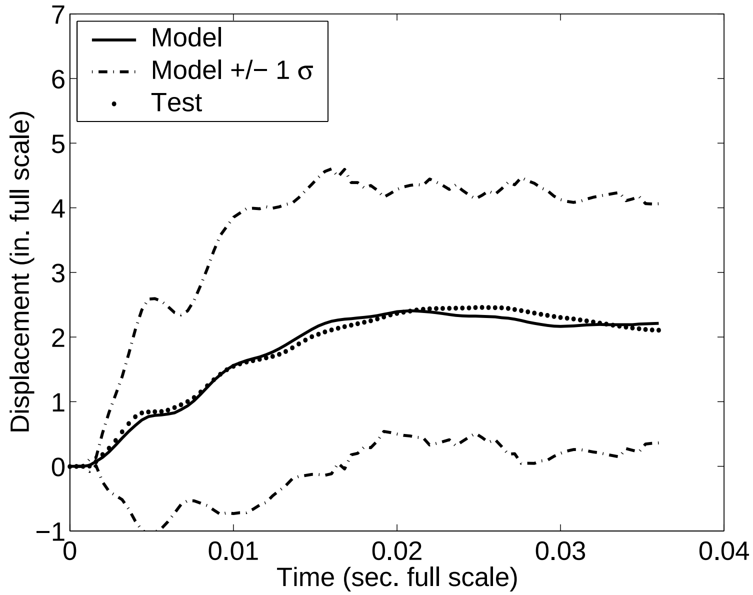
Summary

- Global effect analysis is critical to understand how uncertainty influences the simulation's output.
- Design of experiment techniques efficiently propagate uncertainty through numerical simulations.
- Large-size computational models can be replaced by fast-running surrogate models.
- It might be more efficient to devote the computational resource to the understanding of the variability and input-output relationship upfront, instead of trying to estimate gradients and perform parameter calibration as soon as the model runs.

VERIFICATION AND VALIDATION OF COMPUTATIONAL MODELS

François M. Hemez

Los Alamos National Laboratory
Engineering Sciences and Applications
ESA-WR, Mail Stop T006
Los Alamos, New Mexico 87545
hemez@lanl.gov



APPROVED FOR UNLIMITED, PUBLIC RELEASE ON OCTOBER 16, 2003. LA-UR-03-7640. UNCLASSIFIED.

Los Alamos
NATIONAL LABORATORY

Outline

1. Verification Activities
    (a) Code verification
    (b) Solution verification

2. Test-analysis Correlation
    (a) Feature extraction
    (b) Correlation metrics

3. Finite Element Model Updating
    (a) Formulation of updating
    (b) Spatial incompleteness

4. Model Validation
    (a) Definitions
    (b) Predictive accuracy assessment

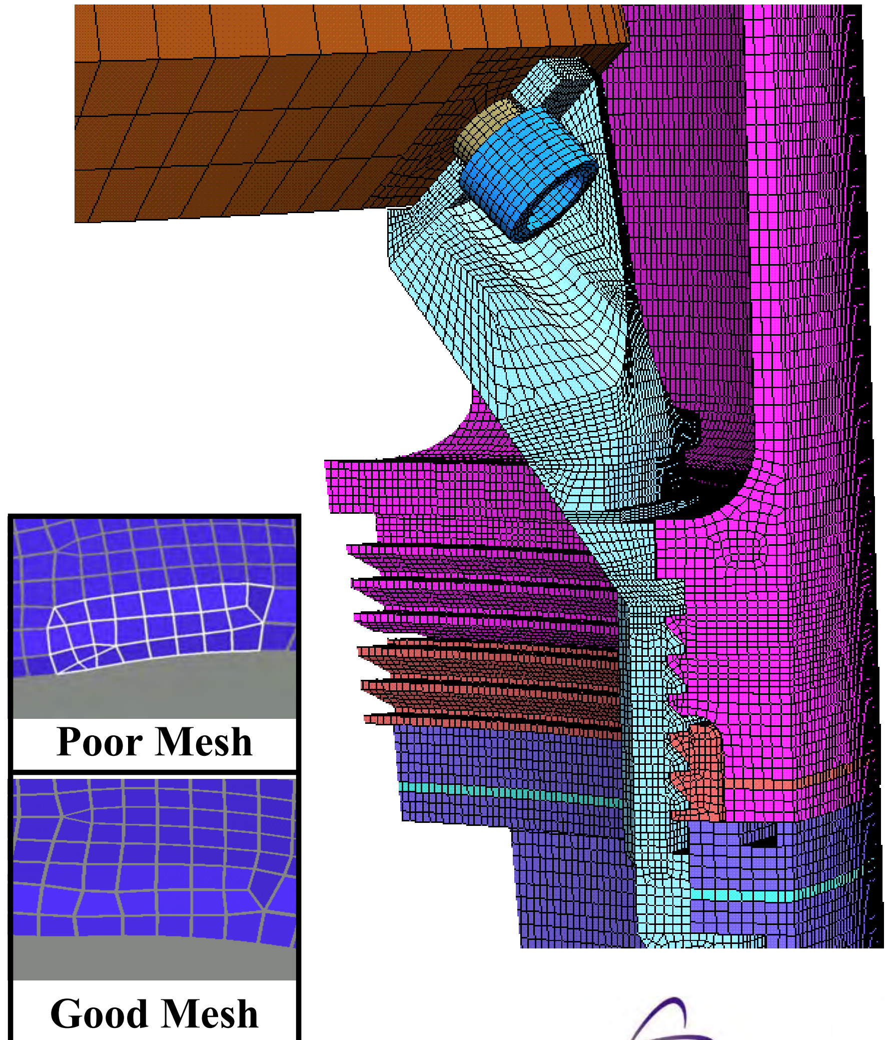
Convergence

- **Numerical methods are *always* approximations.**
  - **A finite element solution is a solution to the *weak*, or "averaged", equations of motion. It is different from the solution of the *strong*, or "enforced point-by-point", formulation.**
  - **Okay, the two are equivalent for a wide range of elliptic problems, that's what the Lax-Milgram Theorem states.**
  - **It does not mean that, at least *locally*, the solution might not be erroneous. It may even be wrong everywhere!**
  - **The basic principle behind the wide-spread practice of discretization is that the approximation *converges* to the "true-but-unknown" solution as the size $h \rightarrow 0$. Caution: Convergence may not be uniform throughout the mesh!**

Typical Problems

- **Stress concentrations are obtained at the “corners” of a mesh or the tip of a propagating crack.**
- **Check poor aspect ratios.**
- **A choice of time discretization Δt and element size h is inadequate to capture a specific wave speed, C_{Max}.**
- **Information propagates in the mesh at the maximum speed of $(h/\Delta t)$. Therefore, better make sure that $(h/\Delta t)>C_{Max}$.**

⇨ *Constraint, $\Delta t < (h/C_{Max})$.*



Verification

- **Verification is the assessment that the computer code does what it is supposed to be doing.**
  - **No programming bugs, no overwriting of values stored in memory, no round-off errors, no truncation errors, etc.**

⇨ ***Code verification.***

- **Calculations should always be verified as well.**
  - **Is the computational mesh adequate for the purpose intended? Has the solution converged? What is the maximum error associated to the solution?**

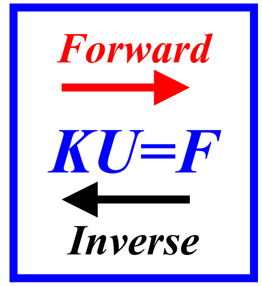
⇨ ***Solution verification.***

Code Verification

- **Good programming practices:**
  - **Do a good job the first time.**
  - **Don't use common blocks.**
  - **Be modular. Break-down the code into functions.**
  - **Include a lot of comments.**
  - **Verify the code every 1,000 lines written or less.**
  - **Track the versions of your code.**

The Method of Manufactured Solutions

- The Method of Manufactured Solutions (MMS) is the most well-known in computational fluid and solid mechanics to automatically verify a computer code.
  - Assume that the problem being solved is *KU=F*.
  - Select a solution *U**. Impose *U** as a prescribed displacement field and calculate the corresponding forces *F*.
  - Then, impose the forces *F* and calculate the corresponding displacement field *U*.
  - At the end, *U* better be equal to *U**!



- The MMS is very sensitive: it will "find" small errors. But it will not tell you where the errors come from. Another significant drawback is that implementation in a "black-box" code can be challenging.

Calculation Verification

- **Mesh convergence can theoretically be verified using *error estimators*.**

$$y^{True} = y(h) + \underbrace{\alpha h^p}_{} + O(h^{p+1})$$

Order of the convergence is *p*.

(Note: *y* denotes the quantity to predict.)

- **In practice, the calculation is performed twice using two meshes (one is called the *coarse mesh*, one is called the *fine mesh*) and solutions are compared.**

- **Mesh convergence depends on what is computed.**
  - **Modal frequencies converge much faster than mechanical or thermal stresses. In computational fluid dynamics, the coefficient of lift converges faster than the coefficient of drag.**

Richardson Extrapolation

- **Always obtain three solutions with a coarse mesh size (or time step) h_C, a medium mesh size (or time step) h_M, and a fine mesh size (or time step) h_F:**

$$y^{True} \approx y(h_C) + \alpha h_C^p, \quad y^{True} \approx y(h_M) + \alpha h_M^p, \quad y^{True} \approx y(h_F) + \alpha h_F^p$$

- **The order of convergence p can be estimated as:**

$$p \approx \frac{\log\left(\frac{y(h_M) - y(h_C)}{y(h_F) - y(h_M)}\right)}{\log(r)} \quad \textbf{where} \quad r = \frac{h_C}{h_M} = \frac{h_M}{h_F}$$

- **The "true-but-unknown" y^{True} can be estimated as:**

$$y^{True} \approx \frac{r^p y(h_F) - y(h_M)}{r^p - 1}$$

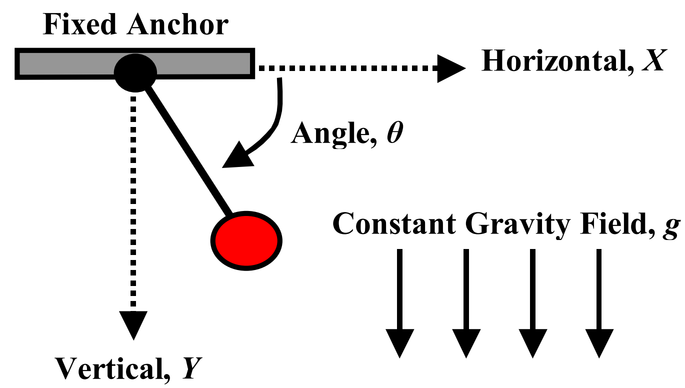
Grid Convergence Index:

$$GCI_{(C \to M)} = 100 \left| \frac{y(h_C) - y(h_M)}{y(h_C)} \right| \left(\frac{\beta}{r^p - 1} \right)$$

(where β is a constant, typically, $1 \leq \beta \leq 3$.)

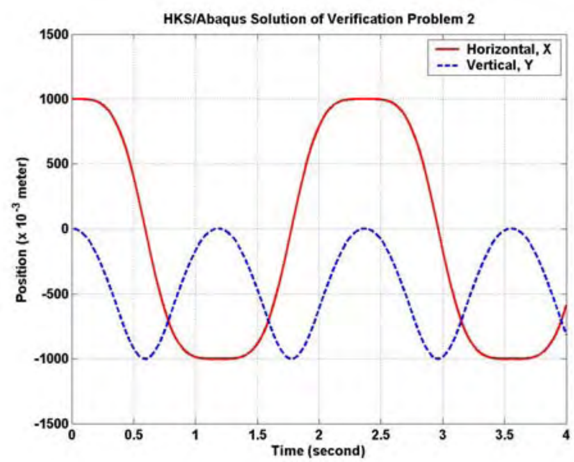
Example of Verification Problem

- **Single degree-of-freedom, large-angle pendulum subjected to gravity.**



	Coarse Solution	Medium Solution	Fine Solution
Time Increment	10^{-2} sec.	10^{-3} sec.	10^{-4} sec.
Number of Samples	400	4,000	40,000
Simulation Time	4.0 sec.	4.0 sec.	4.0 sec.

Criterion	Value Estimated	
Order of convergence	1.0093	
Grid Convergence Index	Coarse-to-medium	Medium-to-fine
	$5.25\ 10^{-2}$ %	$5.13\ 10^{-3}$ %



HKS/Abaqus Explicit (version 6.2), with $\Delta t=10^{-4}$ sec.

Summary

- **Computer codes are generally verified by checking their predictions against each other or against the analytical solutions of simple benchmark problems.**
- **Techniques such as the method of manufactured solutions and Richardson's extrapolation are popular in computational fluid dynamics, not so much in solid mechanics / structural dynamics.**
- **Lots of theoretical work by mathematicians in the area of error estimator, little practical applications.**
- **Practice sanity checks, talk to an experienced analyst.**

Outline

1. Verification Activities
   (a) Code verification
   (b) Solution verification

2. Test-analysis Correlation
   (a) Feature extraction
   (b) Correlation metrics

3. Finite Element Model Updating
   (a) Formulation of updating
   (b) Spatial incompleteness

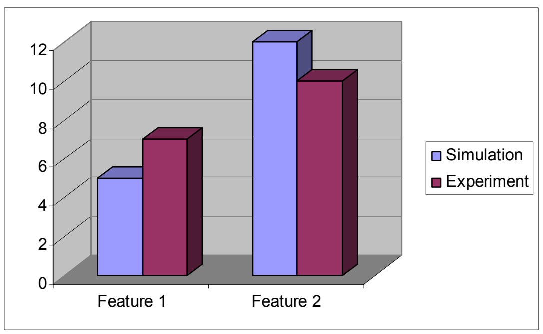
4. Model Validation
   (a) Definitions
   (b) Predictive accuracy assessment

Test-analysis Correlation

- **Test-analysis Correlation (TAC) is the process of assessing the consistency—or lack-of-consistency—between the measured and predicted responses.**

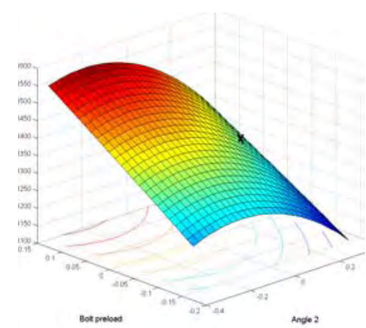
Discrete Point to Discrete Point

$$E = \left\| \left(f_p - f_e \right) \right\|_2$$



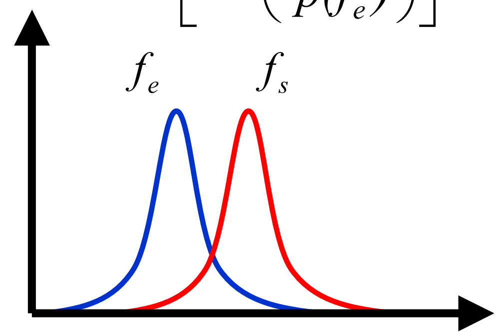
Discrete Point to Response Surface

$$E = \min\left(\left\| f_e - S_p \right\|_2 \right)$$



Distribution to Distribution

$$E = E\left[\log\left(\frac{p(f_s)}{p(f_e)} \right) \right]$$

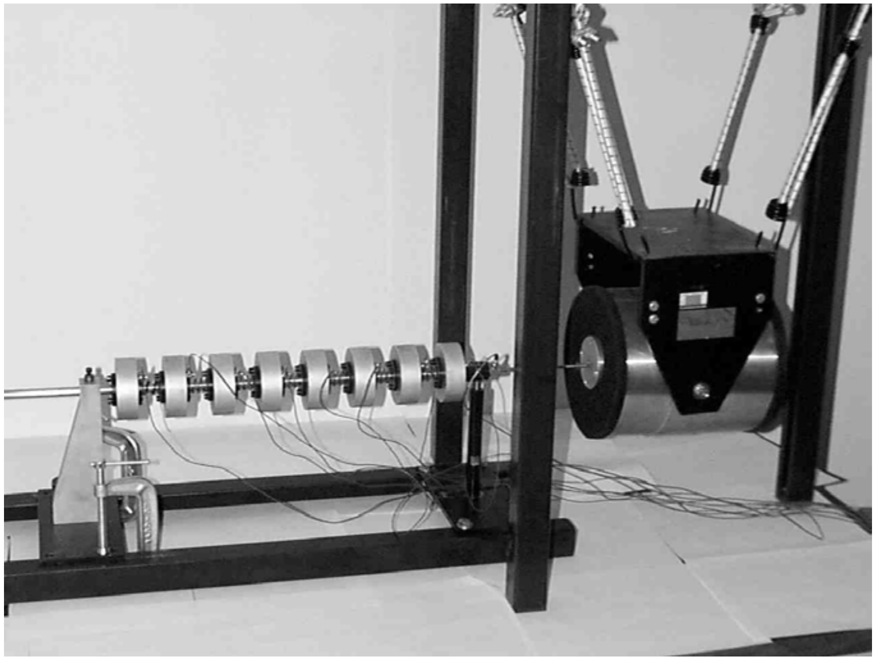


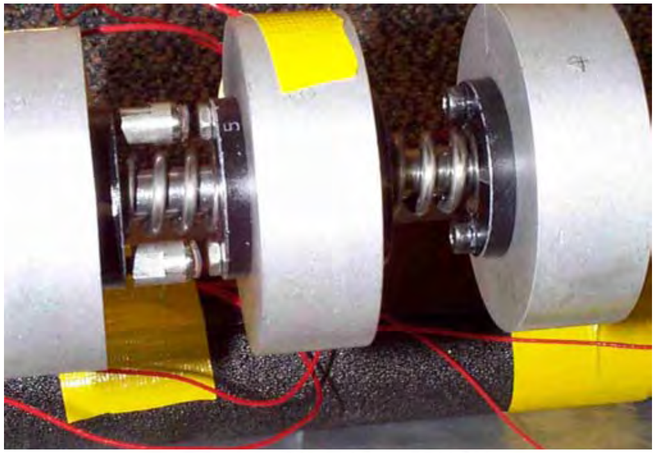
- **Fidelity *metrics* assess the "error" between measured and simulated *features*.**

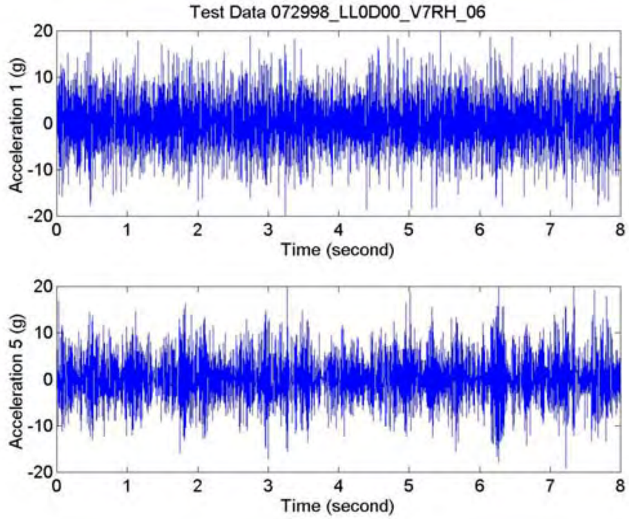
Features

- **Models may include a high level of refinement, such as the geometrical details, loading, material characteristics, etc.**
- **The output can be spatially/temporally distributed, with predictions such as acceleration, stress history at every node of the mesh, energy distribution, etc.**
- **There are only a few *features* of interest to the analyst. In (linear, modal) structural dynamics, the features are generally the mode shapes, resonant frequencies and damping ratios. In shock response, usual features are peak acceleration values and characteristics of the shock response spectrum.**

LANL 8-DOF Mass and Spring System

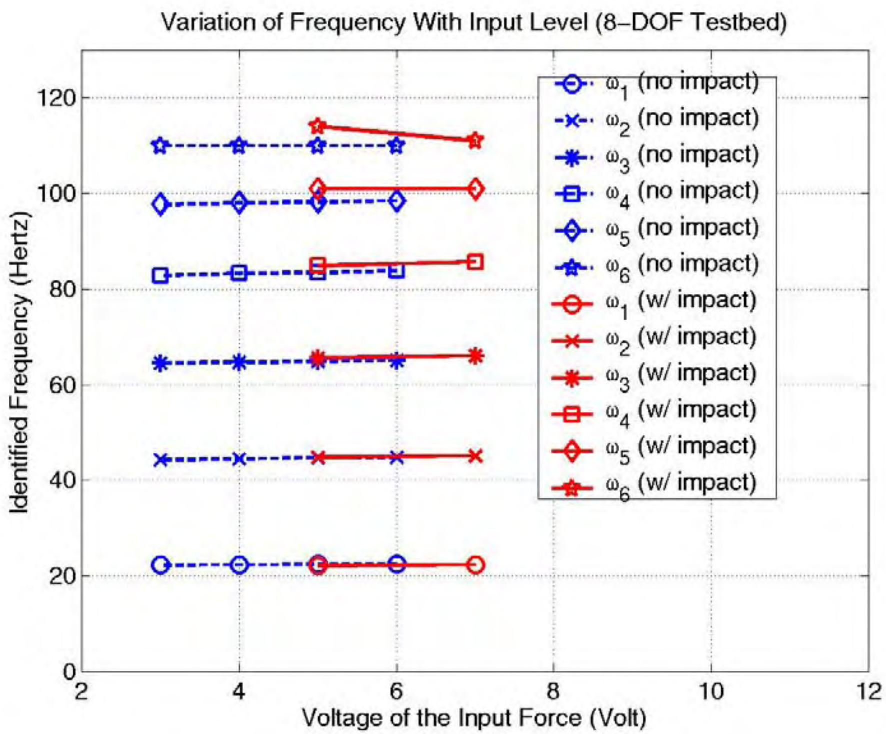






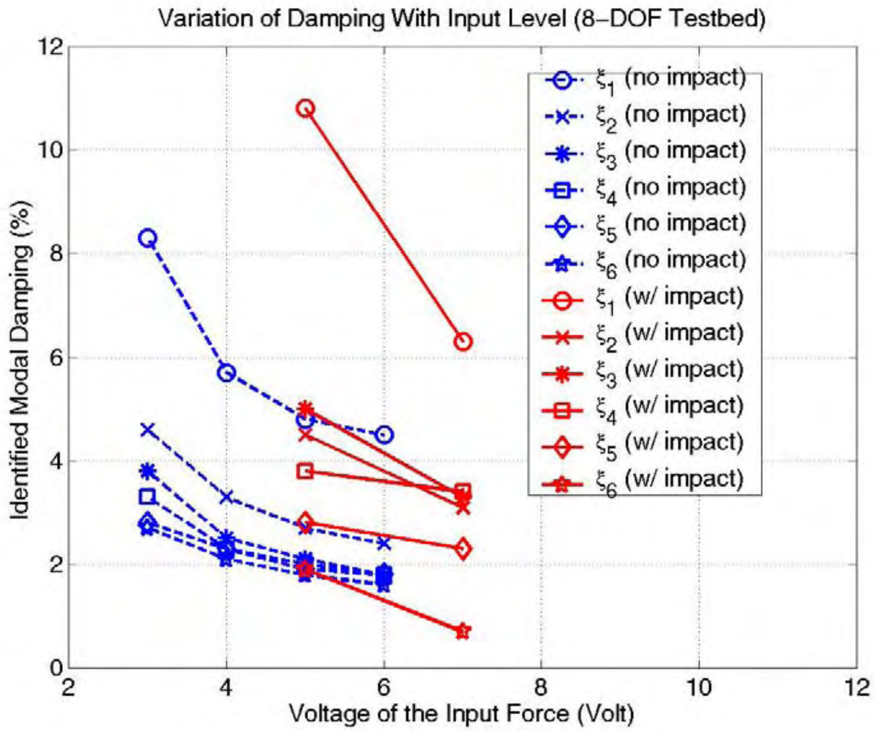
Comparison of Frequencies

- **The identified frequencies are constant as the input force level is increased or the impact mechanism is added.**
  - **The fact that little variation of the modal frequencies is observed seems to make sense because the system remains essentially the same at all input force levels and even when masses come into contact with each other.**



Comparison of Damping Coefficients

- **Modal damping ratios decrease significantly as more force is inputted to the system. Damping increases when the impact mechanism is present.**
  - **For the purely linear system (no impact), damping is reduced at higher force levels, which indicates less energy dissipation.**
  - **For some of the modes, modal damping is increased when the masses come into contact with each other.**



Types of Dynamics

- **The feature analyzed must reflect the type of response.**
  - ***Linear, stationary, Gaussian vibrations:***
    Direct and inverse Fourier transforms; Power spectral density; Input-output transfer functions; Frequency responses; Modal parameters.
  - ***Transient and shock response:***
    Peak values; Energy content; Decrement and exponential damping; Shock response spectrum; Temporal moments.
  - ***General-purpose time series analysis:***
    Auto-regressive and moving average models; Time-frequency transforms; Wavelet transform; Statistics; Principal component decomposition.
  - ***Unstable, chaotic response:***
    State-space maps, Poincaré maps; Time-frequency, higher-order transforms; Symmetric dot pattern; Fractal analysis.

What's a Good Feature?

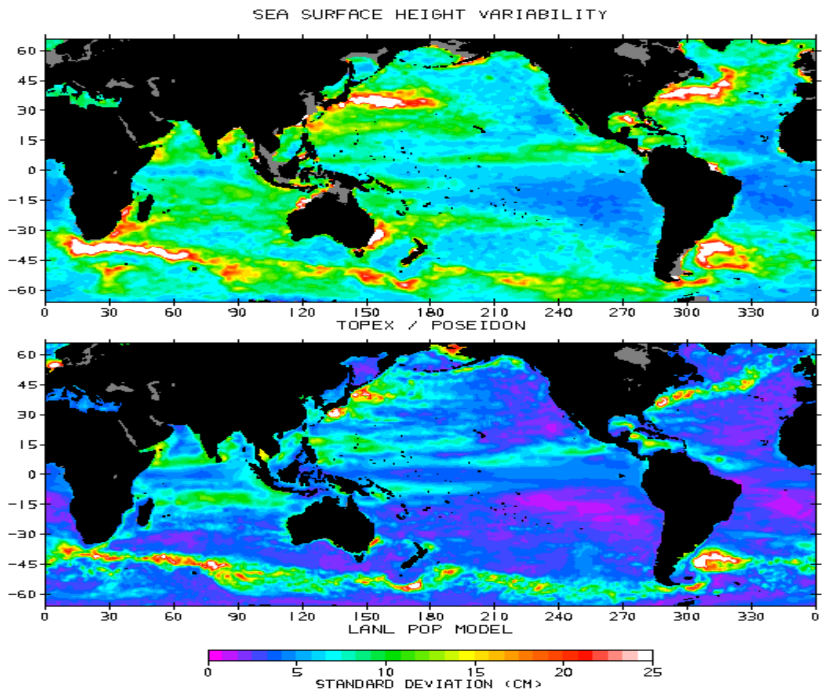
- **Attention must be paid to *dimensionality* in order to facilitate data processing and statistical testing.**
- **Feature selection should reflect the type of dynamical response: Stationary? Linear? Nonlinear? Keep in mind those features that assess the degree to which your *a priori* opinion about the system is verified.**
- **Feature selection depends strongly on the application.**
- **Comparing features is a different problem, it involves the choice of a *metric* for the comparison.**

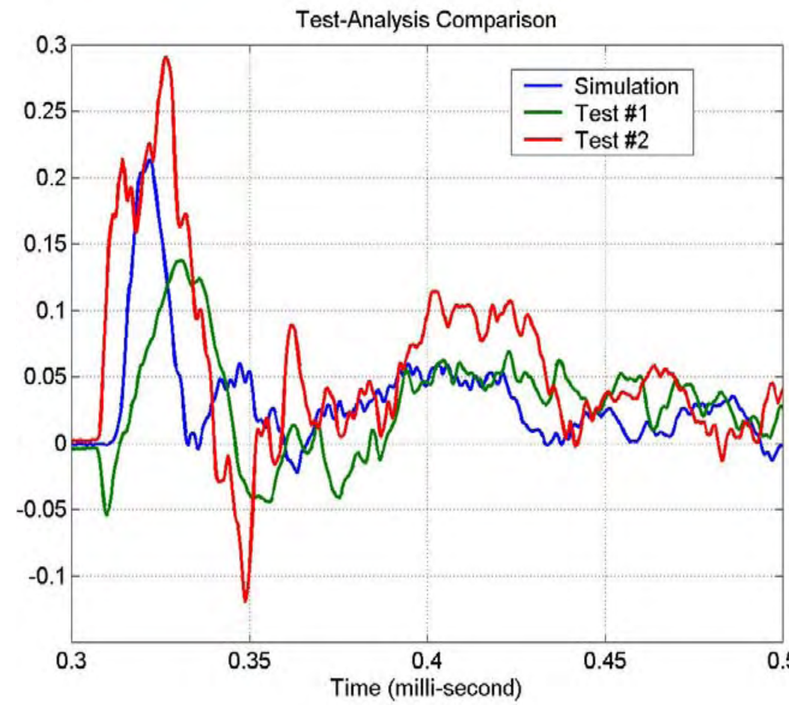
Comparison of Features

- Comparison is generally performed using the "view-graph norm." It is intuitive but non-informative.
- In (linear, modal) structural dynamics, comparison relies almost exclusively on the correlation of mode shapes and frequencies.
- Comparison in other computational sciences (such as physics, chemistry, biology, climatology) is generally restricted to the "view-graph norm."
- To research the most advanced methods, read the literature in statistical sciences, neural networks, image processing and pattern recognition.

The "View-graph" Norm

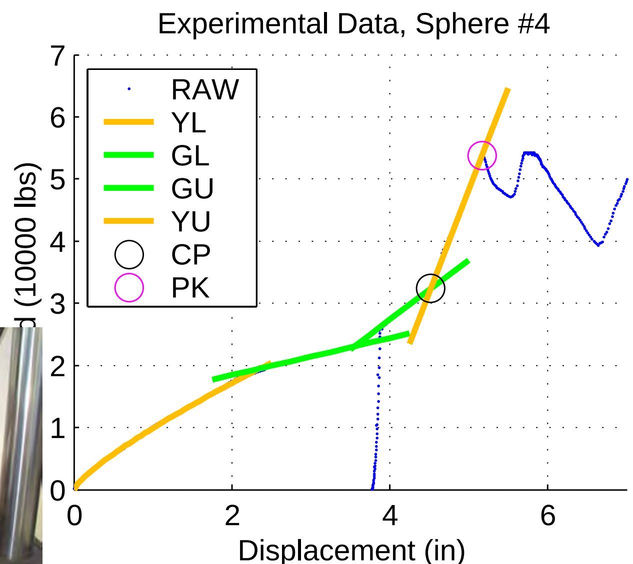
- The "view-graph" norm is unavoidable but it may not be very informative. Always back-up TAC with a quantitative assessment. Use quantitative metrics whenever possible.

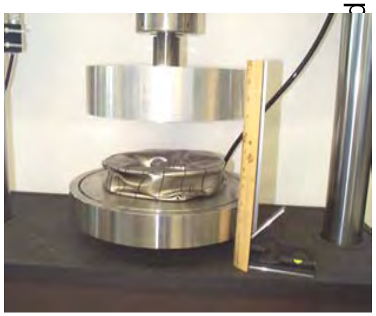


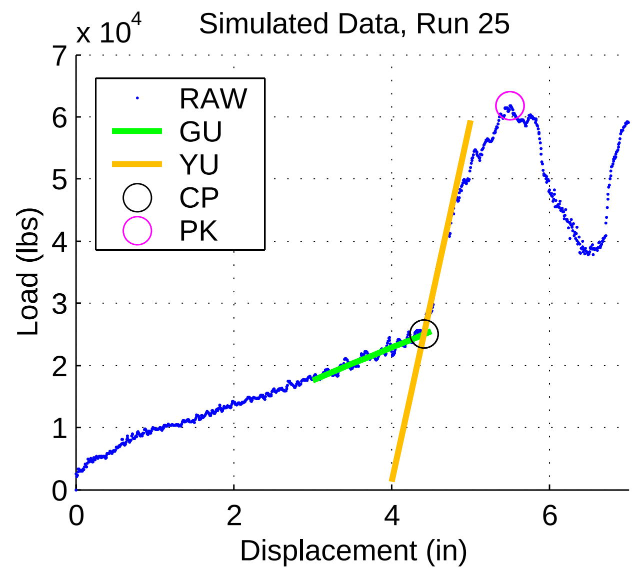


Single to Single

- Feature to feature comparison can be more informative than time series (or response curve) comparison if the right feature is selected. Error and correlation metrics are easier to define and compute with single features.





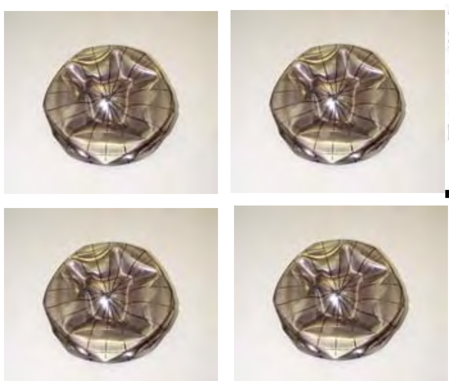


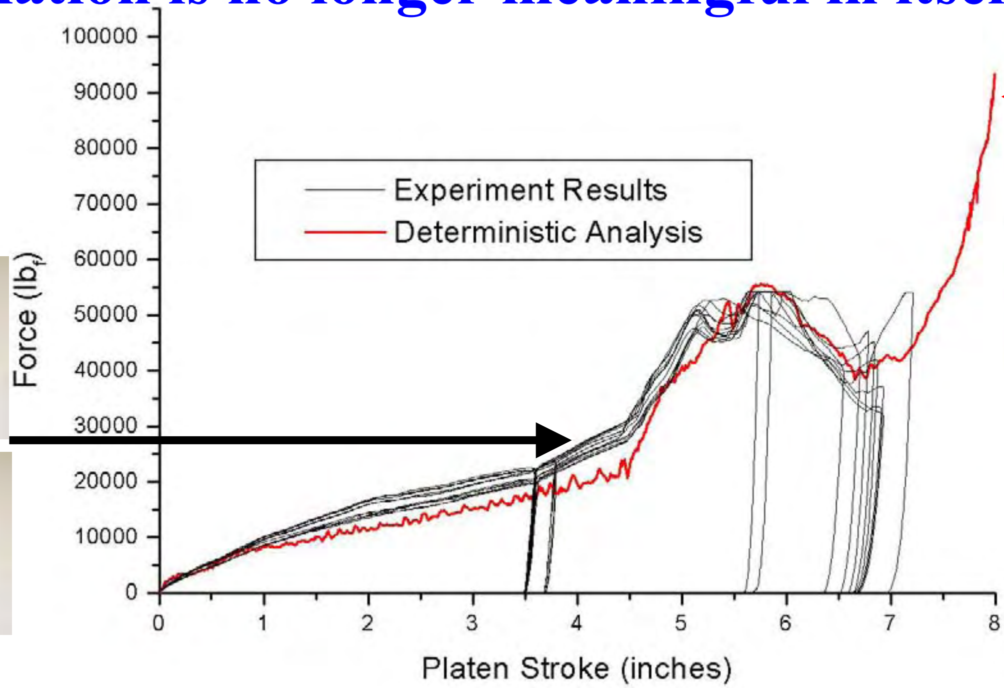
PAS
DAMAGE I

Single to Distribution

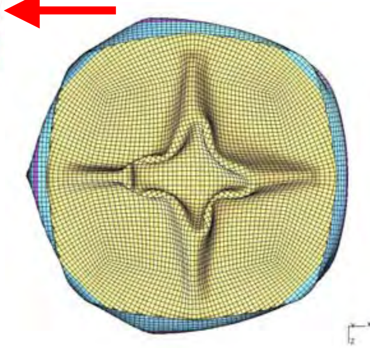
- **A single, deterministic prediction is compared to the family of curves obtained by performing *replicate* or multiple experiments. Agreement between the prediction and a single measurement extracted from the population is no longer meaningful in itself.**

Tests are repeated several times.



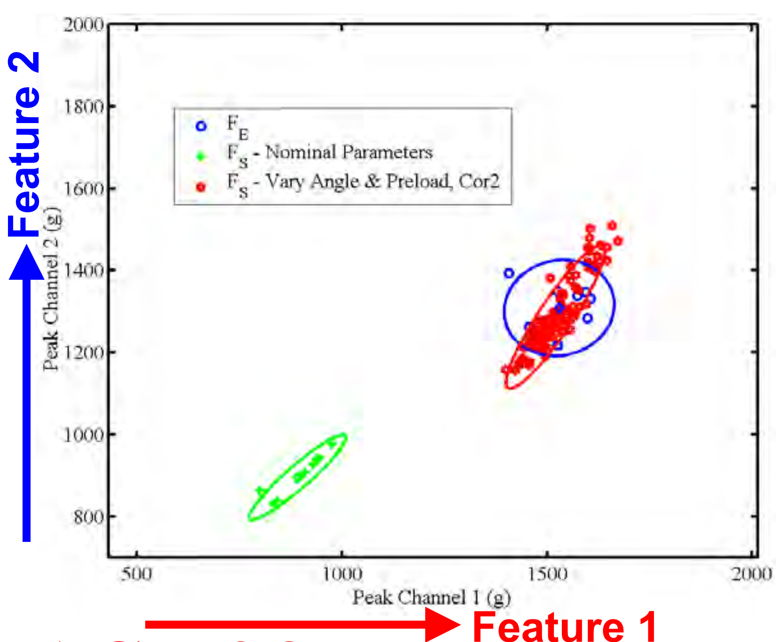


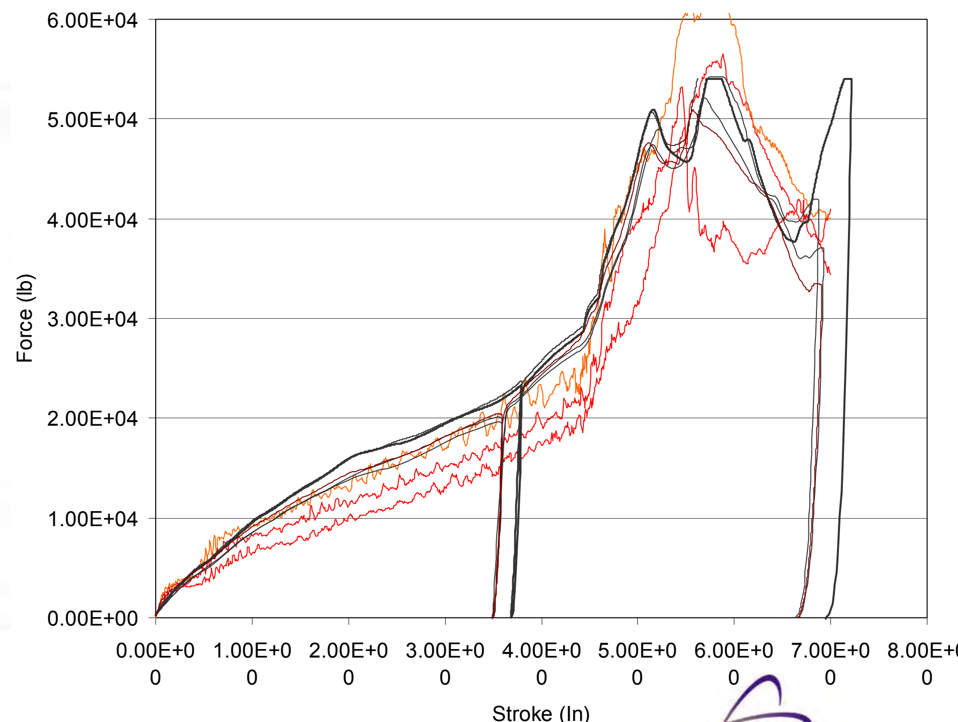
Result of a single FE analysis.



Distribution to Distribution

- **Distributions are compared when populations of testing and FE analysis results are available. Beyond the visual assessment and trend comparison, statistical testing is necessary to assess the *lack of consistency* between the two populations.**





Metrics

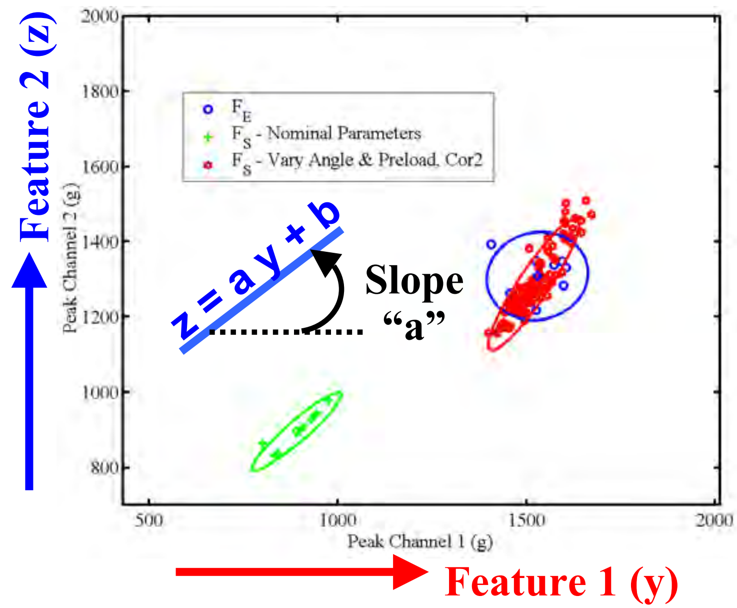
- **The metric is a measure of "distance" whose purpose is to estimate the *fidelity error*, that is, the extent to which the numerical model reproduces the observed data.**

- **RMS error "ε":**

$$\varepsilon = \sqrt{\frac{1}{N}\sum_{k=1\cdots N}\left(y_k^{Test} - y_k\right)^2}$$

- **Regression metric "*a*":**

  Regression metrics assess the degree of correlation between two (or more) features.



Comparing Modal Parameters

- **The error between modal frequencies is generally formulated in terms of relative error.**

$$\text{Error} = 100\,\frac{\omega_{\text{Test}} - \omega_{\text{Model}}}{\omega_{\text{Test}}} \quad (\%)$$

- **To make sure that the same vectors are compared, the Modal Assurance Criterion (MAC) can be estimated. The MAC measures the degree of correlation between two vectors.**

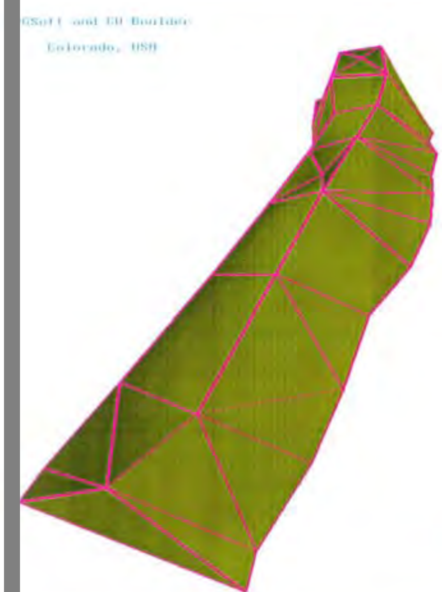
$$\text{MAC} = 100\,\frac{\left(u_{\text{Test}}^{T} u_{\text{Model}}\right)^2}{\left(u_{\text{Test}}^{T} u_{\text{Test}}\right)\left(u_{\text{Model}}^{T} u_{\text{Model}}\right)} \quad (\%)$$

- **In practice, very high MAC values are needed to provide confidence that the two vectors are the same mode (>90%).**

Mode Pairing

First measured torsion mode.

Second measured bending mode.



First measured bending mode.

First predicted bending mode.

Low MAC value (20%-to-60%)

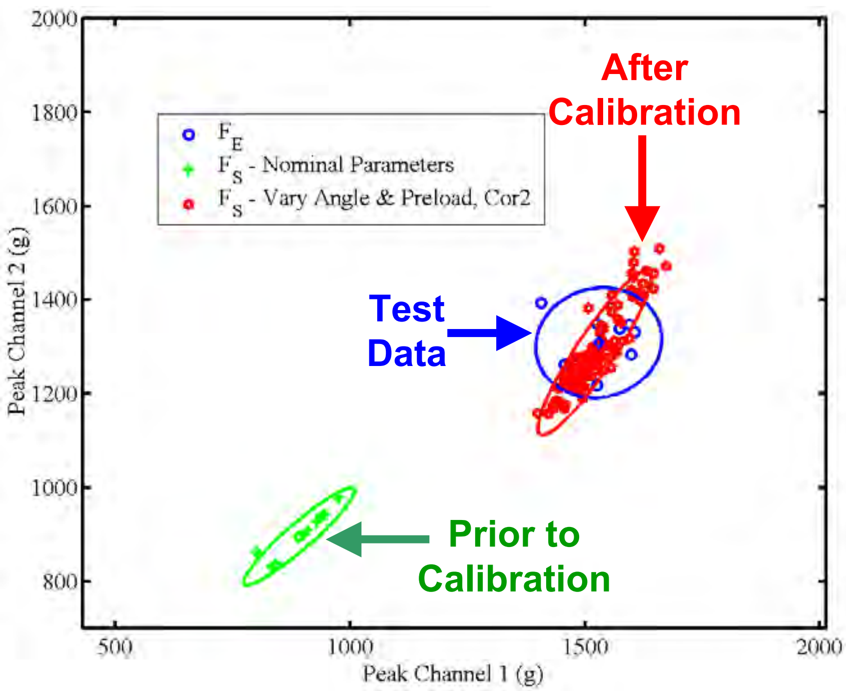
High MAC value (80%-to-95%)

Statistical Testing

- **Statistical testing assesses the consistency between populations of features. It assesses whether or not the "parent" populations have similar characteristics.**
  - **The *T-test for significantly different means* rejects the hypothesis that the means are the same.**
  - ***Variance-based analysis* (Pearson correlation ratio, R^2 statistic) estimate the correlation between two populations of features.**
  - **The *F-test for significantly different variances* rejects the hypothesis that the variances are the same.**
  - **The *Chi-square test* rejects the hypothesis that the predicted and measured features are drawn from the same parent population.**

Mahalanobis Distance

- **The Mahalanobis metric calculates the "distance" between a single point and a population. The reference is the population. It can be generalized to multiple dimensions, and can be associated to a Chi-square test.**



$$\varepsilon^2 = \left(\frac{\bar{y}^{Test} - y}{\sigma^{Test}}\right)^2$$

$$\varepsilon^2 = \left(\bar{y}^{Test} - y\right)^T \left(\Sigma^{Test}\right)^{-1} \left(\bar{y}^{Test} - y\right)$$

$$\Sigma^{Test} \approx \left(Y^{Test} - \bar{y}^{Test}\right)\left(Y^{Test} - \bar{y}^{Test}\right)^T$$

$$Y^{Test} = \begin{bmatrix} a^{(1)} & a^{(2)} & \cdots & a^{(N)} \\ b^{(1)} & b^{(2)} & \cdots & b^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ x^{(1)} & x^{(2)} & \cdots & x^{(N)} \end{bmatrix}, \bar{y}^{Test} = \begin{Bmatrix} \bar{a} \\ \bar{b} \\ \vdots \\ \bar{x} \end{Bmatrix}$$

Summary

- **Modal-based features, MAC correlations and RMS errors are the core tools of conventional test-analysis correlation for linear, modal dynamics.**

- **Maintaining a clear distinction between *features*, test-analysis *comparison* (visual) and test-analysis *correlation* metrics (quantitative) helps to understand the process of test-analysis correlation.**

- **Formulating the metrics as statistical tests (instead of conventional "minimum distance", RMS errors) helps when dealing with variability issues.**

Outline

1. Verification Activities
   (a) Code verification
   (b) Solution verification

2. Test-analysis Correlation
   (a) Feature extraction
   (b) Correlation metrics

3. Finite Element Model Updating
   (a) Formulation of updating
   (b) Spatial incompleteness

4. Model Validation
   (a) Definitions
   (b) Predictive accuracy assessment

FE Model Updating

- **At this point, we have built a finite element model, performed numerical simulations, and extracted features from the output. We have also conducted physical experiments to measure the same features.**
- **Finally, we have compared the two sets of features (measured vs. predicted) and have assessed that the predictions are not *accurate* enough.**
- **Finite element model updating techniques are semi-automatic optimization techniques that can *identify and correct the modeling error*.**

Formulation

- **Select the appropriate features *y* and formulate a fidelity metric *ε* for measuring the correlation between test data and FE predictions.**

- **Optimize the parameters *p* of the finite element model ("calibration", "tuning", etc.) to minimize the metric.**

$$\min_{\{p\}} \sum_{k=1\cdots N} \left(y_k^{Test} - y_k(p)\right)^T W_k^{-1} \left(y_k^{Test} - y_k(p)\right)$$

- **Constraints can be added to guide the search of a solution, avoid local minima and physically unrealistic adjustments.**

Examples

- **If the purpose of the model is to predict the modal frequencies and replicate modal tests have been performed, the updating criterion can be defined as:**

$$\min_{\{p\}} \left(\frac{\overline{\omega}^{Test} - \omega(p)}{\sigma^{Test}} \right)^2$$

- **If the modeling error is unknown, criteria based on error *residuals* can be defined to obtain a *mapping* of the modeling error over the computational mesh:**

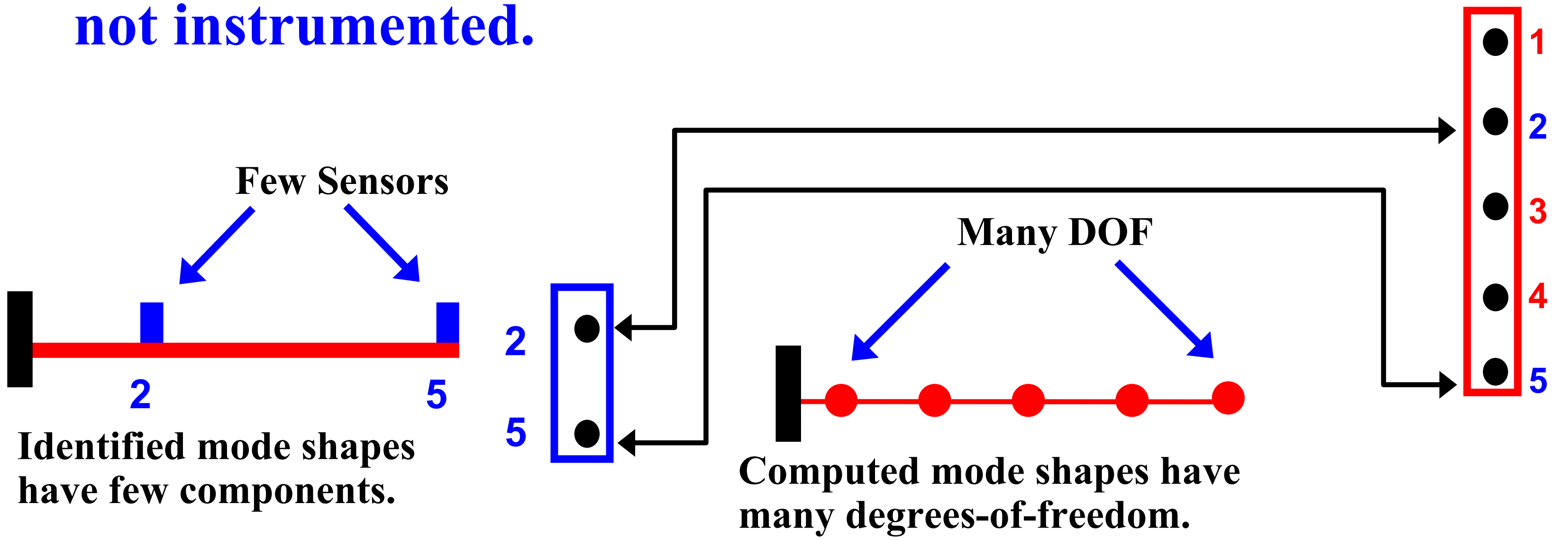
$$\min_{\{p\}} \sum_{k=1\cdots N} R_k^T(p)\, W_k^{-1}\, R_k(p)$$

$$\text{where} \quad R_k(p) = \left(K(p) - \omega_k^{Test^2} M(p) \right) u_k^{Test}$$

Caution: Quantities must have the same size!

Spatial Incompleteness

- Spatial incompleteness stems from the fact that some of the degrees-of-freedom of the computational model are not instrumented.



- Measurements generally come in lesser quantity than grid points of a mesh. Measurements are not located at the same coordinates as the grid points of the mesh.

Dealing With Spatial Incompleteness

- **How to account for the fact that only a subset of the model's degrees-of-freedom are instrumented?**

$$\{u\} = \begin{Bmatrix} u_M \\ u_O \end{Bmatrix}$$

u_M: **Instrumented (Measured)**

u_O: **Non-measured (Unknown)**

- **The FE matrices can be condensed (likewise, the measured quantities can be expanded) by projecting them in a user-defined subspace (denoted by *R* below):**

$$[T] = \begin{bmatrix} Id \\ R \end{bmatrix}$$

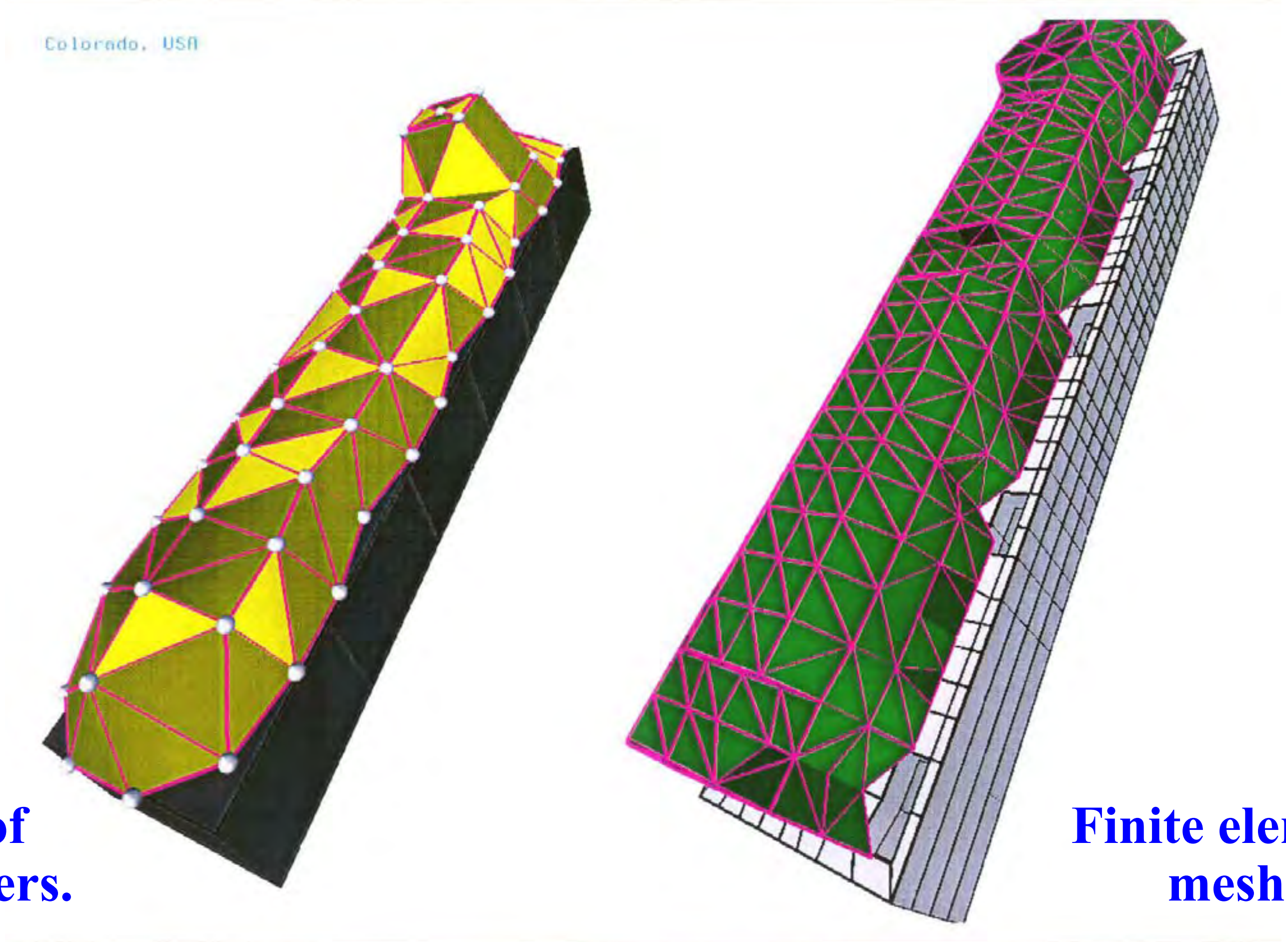
Id: **Instrumented DOF**

R: **Non-measured DOF**

Matrix Reduction $[K_{Reduced}] = [T]^T [K][T]$

Vector Expansion $\{u_{Expanded}\} = [T]\{u_M\}$

An Industrial Example



Location of accelerometers.

Finite element mesh.

Original TAC

- **Original test-analysis correlation:**

Identified Mode Number	Identified Frequency (Hertz)	Computed Mode Number	Computed Frequency (Hertz)	MAC (%)	Frequency Error (%)
2	1,392.7	1	1,467.9	72.8	5.4
5	2,244.2	2	2,406.5	71.5	7.2
6	2,359.2	3	2,902.6	46.7	23.0
8	2,482.5	4	3,212.3	35.3	29.4

- **The original model is too stiff. Mass error in the model accounts for about 2% of the frequency error. Also noticeable are the poor MAC values. Several modes identified are not predicted by the FE model.**

Formulation of the Updating Problem

- **In this example, FE model updating is formulated as the minimization of the out-of-balance force residuals:**

$$R_j = \left(\begin{bmatrix} K_{MM} & K_{MS} \\ K_{SM} & K_{SS} \end{bmatrix} - \omega_j^2 \begin{bmatrix} M_{MM} & M_{MS} \\ M_{SM} & M_{SS} \end{bmatrix} \right) \begin{Bmatrix} u_{Mj} \\ u_{Oj} \end{Bmatrix}$$

Measured (u_{Mj})

Non-measured (Unknown) (u_{Oj})

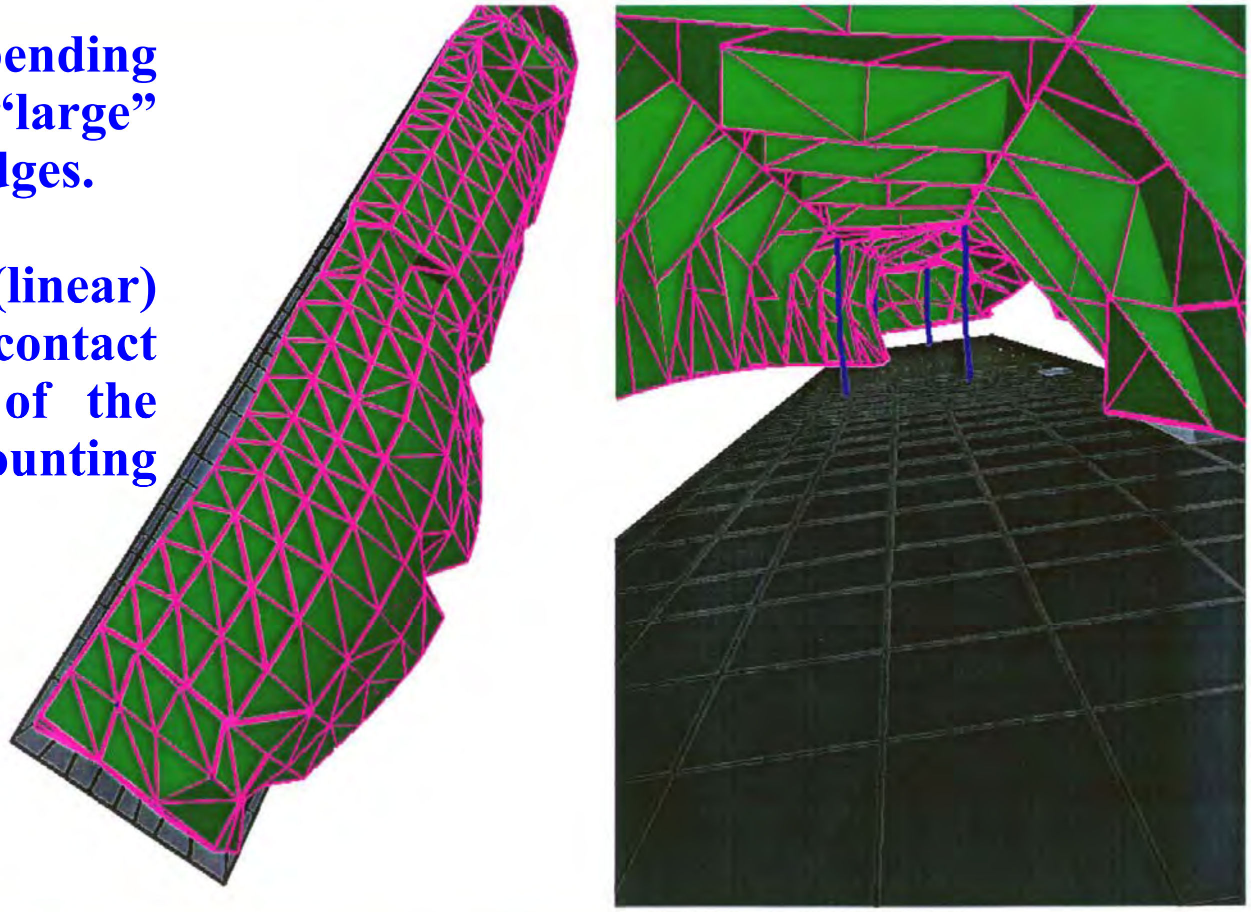
The FE matrices depend on the (unknown) adjustment *dp*.

- **The difficulty is to solve the optimization problem. The mass and stiffness matrices depend on the unknown adjustments *dp*. Part of the mode shape is measured, but part is also unknown.**

- **Which parameters *p* of the model should be updated?**

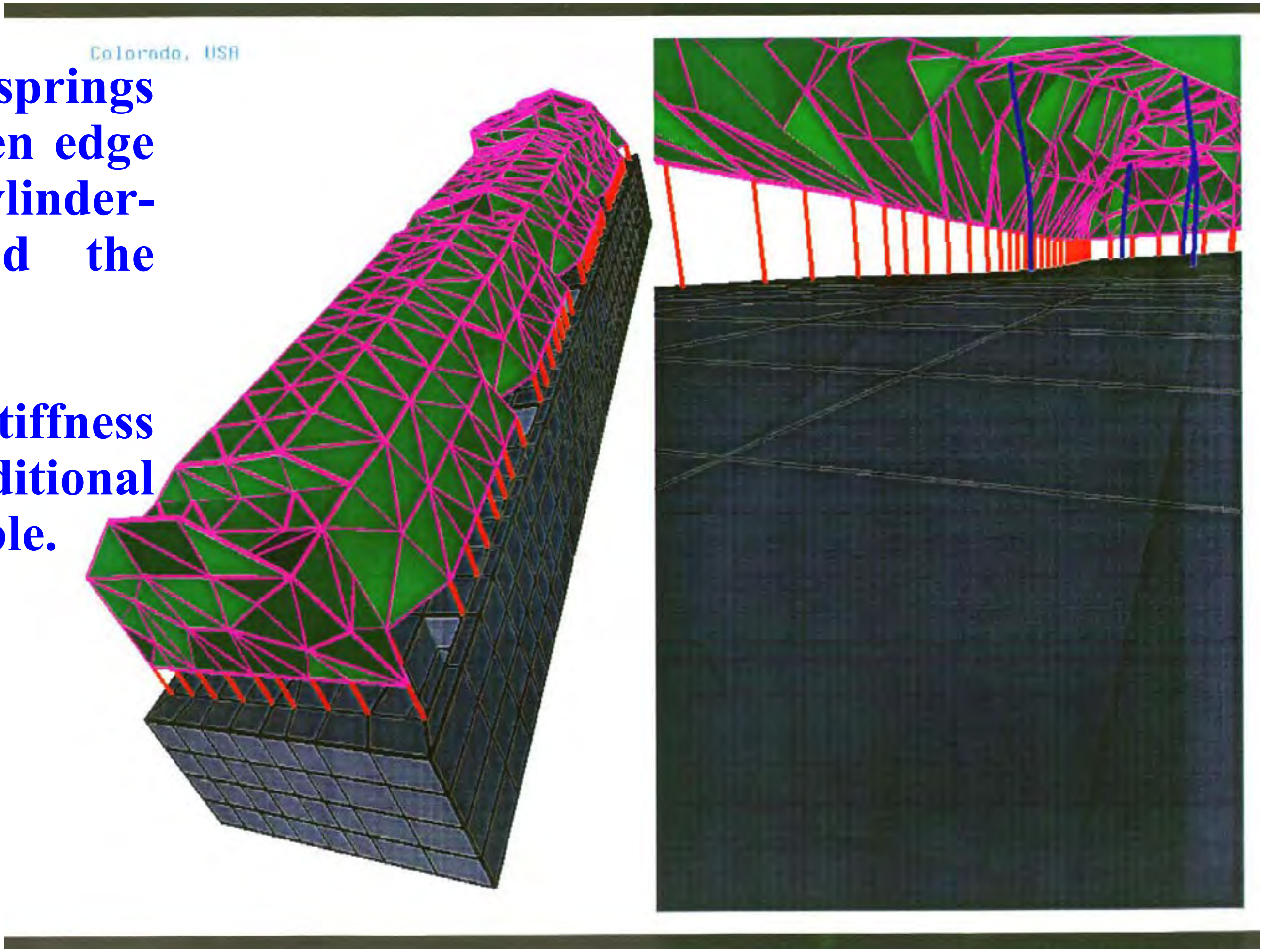
Deflection of the Cylinder-head Cover's Edge

- Some of the bending modes show a "large" deflection of the edges.
- Nothing in the (linear) model prevents contact and penetration of the edges with the mounting plate.
- Is this a problem?



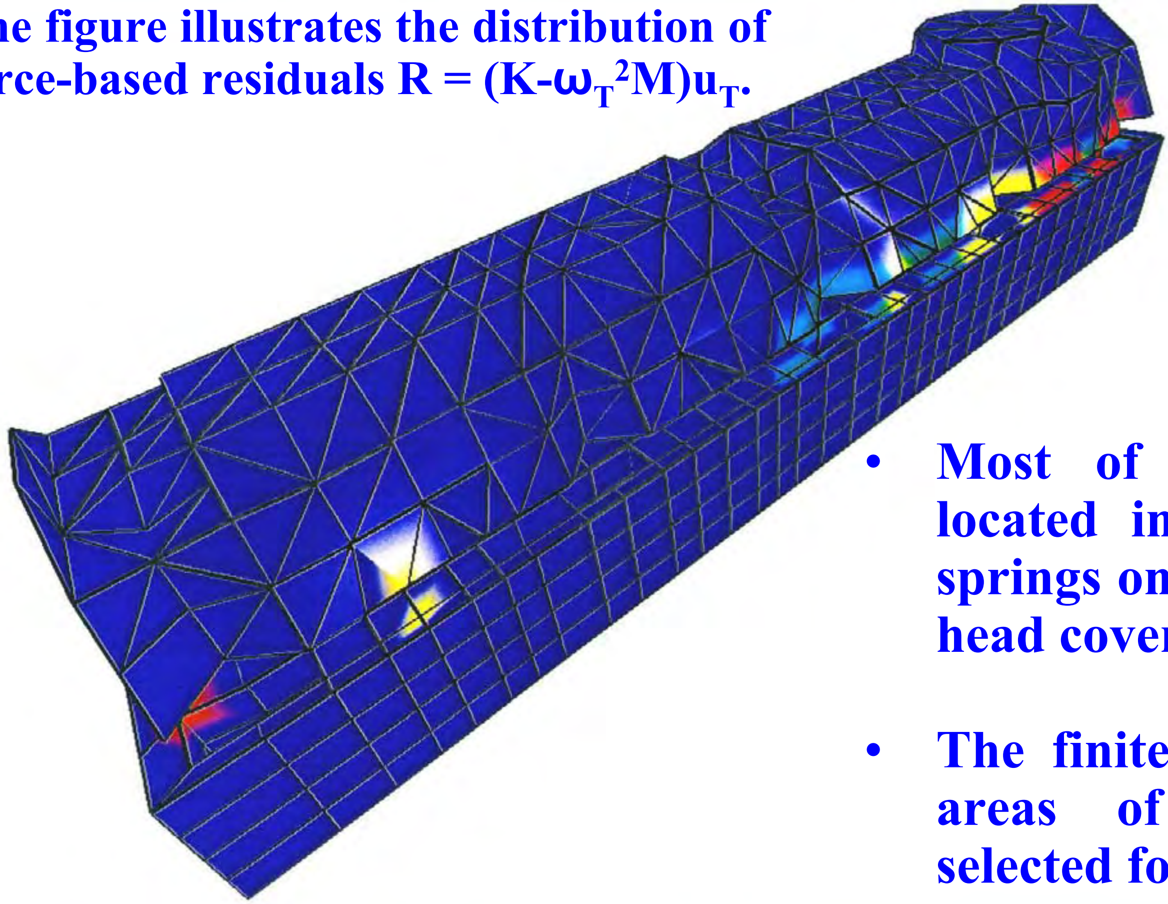
Parameterization

- Simple linear springs are added between edge nodes of the cylinder-head cover and the mounting plate.
- Each spring stiffness becomes an additional calibration variable.



Mapping of the Modeling Error

The figure illustrates the distribution of force-based residuals $R = (K-\omega_T^2 M)u_T$.



- **Most of the modeling error is located in the vicinity of contact springs on one edge of the cylinder-head cover.**
- **The finite elements located in the areas of dominant error are selected for parametric correction.**

Final TAC

- Original test-analysis correlation:

Identified Mode	Identified Freq. (Hz)	Computed Mode	Computed Freq. (Hz)	MAC (%)	Frequency Error (%)
2	1,392.7	1	1,467.9	72.8	5.4
5	2,244.2	2	2,406.5	71.5	7.2
6	2,359.2	3	2,902.6	46.7	23.0
8	2,482.5	4	3,212.3	35.3	29.4

- Final test-analysis correlation:

Identified Mode	Identified Freq. (Hz)	Computed Mode	Computed Freq. (Hz)	MAC (%)	Frequency Error (%)
2	1,392.7	1	1,412.5	65.7	-1.4
5	2,244.2	2	2,250.6	72.1	0.3
6	2,359.2	3	2,378.2	63.0	0.8
8	2,482.5	4	2,563.2	69.6	-3.2

Summary

- **Features for finite element model updating should not be restricted to modal parameters, especially if the dynamics are nonlinear, non-stationary, etc.**

- **Likewise, metrics should not be restricted to the deterministic RMS error. Use correlation coefficients, statistical testing.**

- **Finite element model updating techniques can help understanding why predictions are in disagreement with measurements. However, usefulness for "real", "industrial" applications remains to be demonstrated with more than a few isolated studies.**

- **A more efficient alternative to FE model updating is to perform a design of experiments, effect screening, replace the FE simulation by a meta-model, *then* optimize the parameters and propagate the uncertainty from inputs to outputs.**

Outline

1. Verification Activities
   (a) Code verification
   (b) Solution verification

2. Test-analysis Correlation
   (a) Feature extraction
   (b) Correlation metrics

3. Finite Element Model Updating
   (a) Formulation of updating
   (b) Spatial incompleteness

4. Model Validation
   (a) Definitions
   (b) Predictive accuracy assessment

A Few Comments on V&V …

- **The weakness of science-based prediction is *credibility*, or accumulating evidence of the predictive accuracy of numerical simulations.**
- **Verification and Validation (V&V) is the scientific method to provides *evidence* of predictive accuracy.**
- **There is never a formal proof that a model is valid. There is only lack-of-evidence that a model is invalid.**
- **A non-valid model may still be useful. (Example: Greek astronomy models developed by Aristotle, ~300 BC.)**
- **V&V is not about *truth*, it is about *control*.**

What is Model Validation?

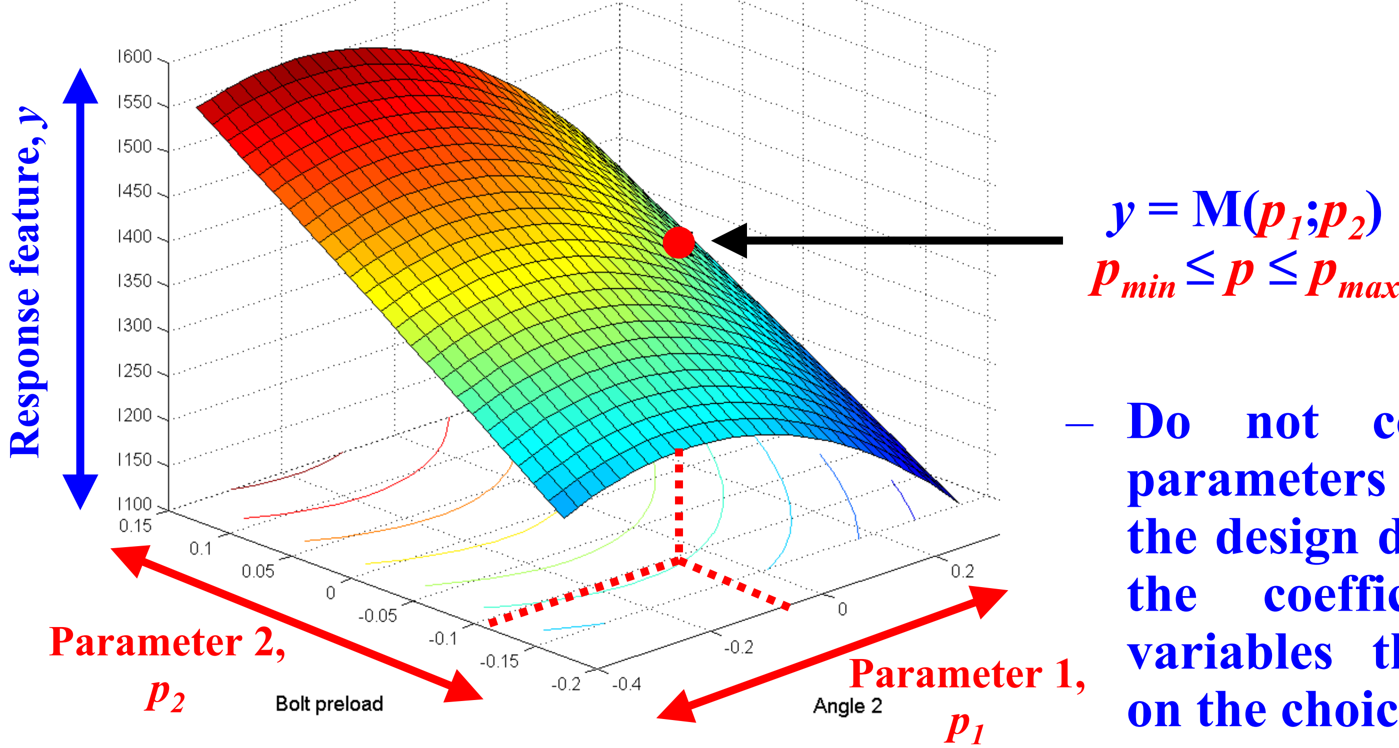
- "The process of determining the degree to which a computer simulation is an accurate representation of the real world, from the perspective of the intended uses of the model." —DoD Modeling and Simulation —DoE ASCI Program

- "Solving the right equations." —Roache (1998)

- "The substantiation that a model within its *domain* of applicability possesses a *satisfactory* range of *accuracy* consistent with the intended *applications* of the model." —Schlesinger (1979)

Design Space

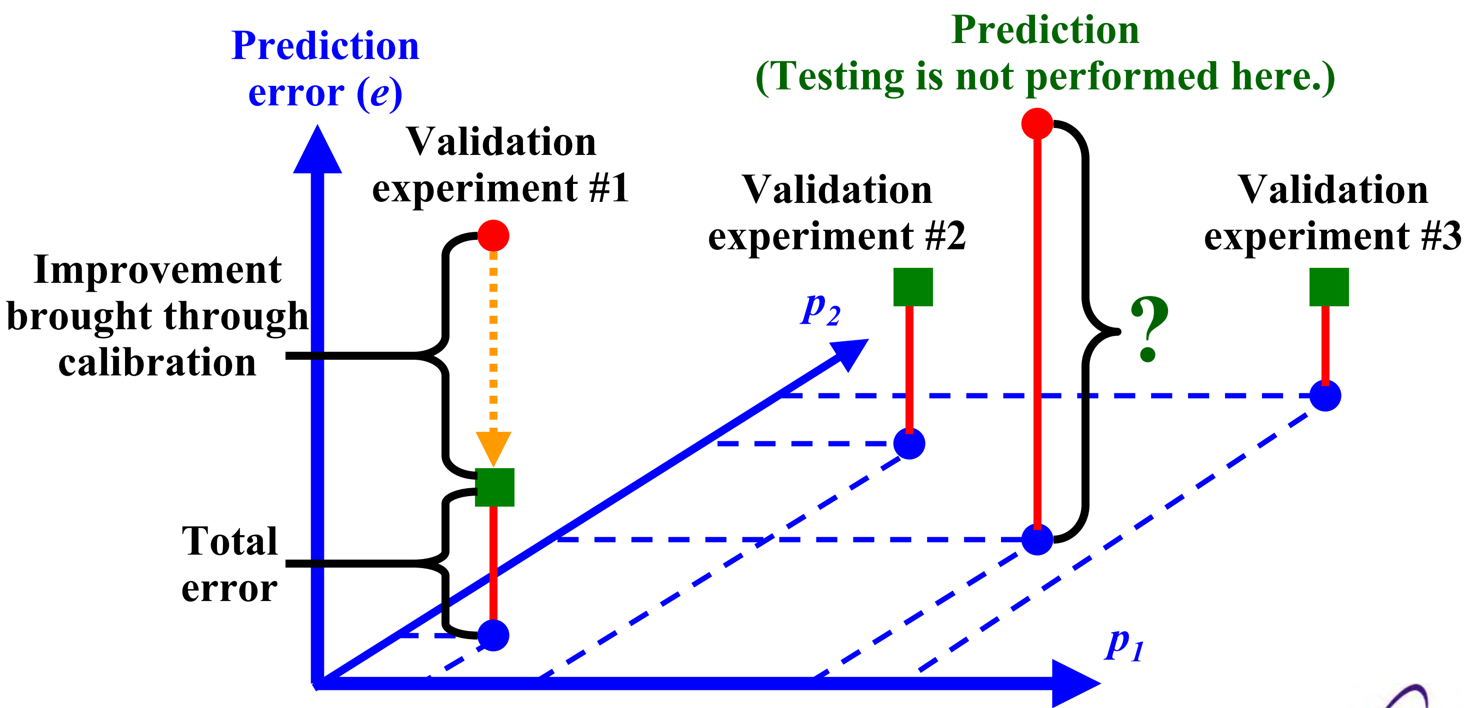
- **Control parameters of a numerical simulation or physical experiment vary within specified ranges. Their variation defines the design space.**



– **Do not confuse the parameters that define the design domain with the coefficients and variables that depend on the choice of model.**

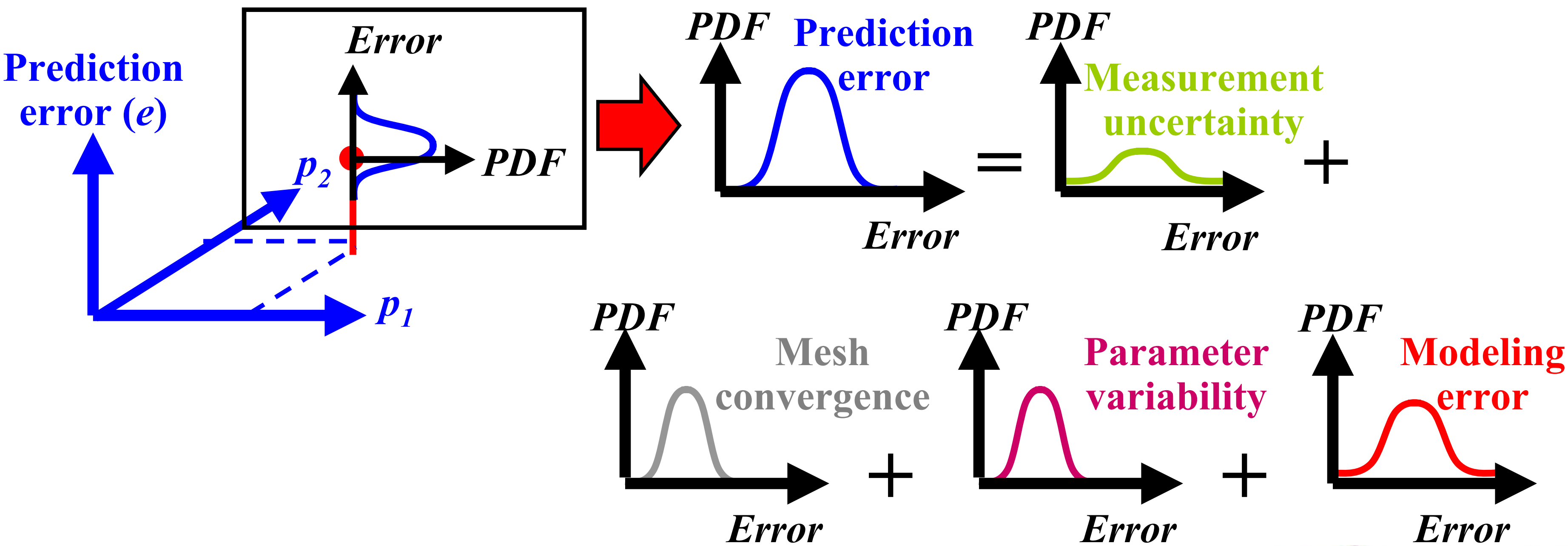
Assessment of Predictive Accuracy (1 of 2)

- **The components of *predictive accuracy assessment* are the model $y = M(p)$, design space, measurements y^{Test} and prediction error $e = ||y^{Test} - y||$.**



Assessment of Predictive Accuracy (2 of 2)

- **The uncertainty introduced by the modeling assumptions (*modeling error*) must be assessed and quantified, as well as uncertainty from the tests, calculations, and parameter variability. (There might be other sources of uncertainty.)**



Steps of Model Validation

- Code verification activities.
- Convergence of the numerical solution.
- Feature extraction.
- Local sensitivity study (finite difference-based).
- Design of computer experiments.
- Global sensitivity (variance-based), effect screening.
- Fast-running, meta-models.
- Test-analysis comparison and correlation.
- Model revision and parameter calibration.
- Uncertainty propagation and assessment.
- Prediction accuracy assessment.

Possibly, go back.

Three Examples

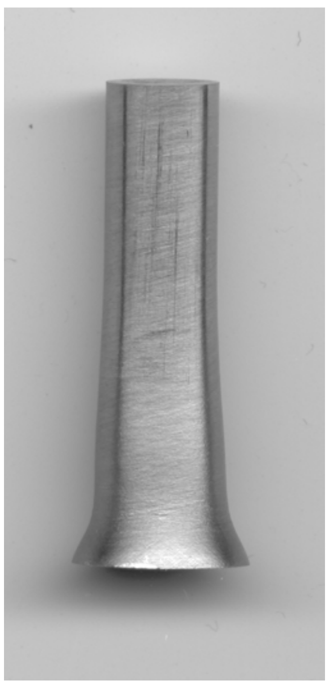
- **Example I: Simulation of a Taylor anvil impact**
  (Rate and temperature-dependent model of plasticity)
  → Illustration of an assessment of prediction accuracy

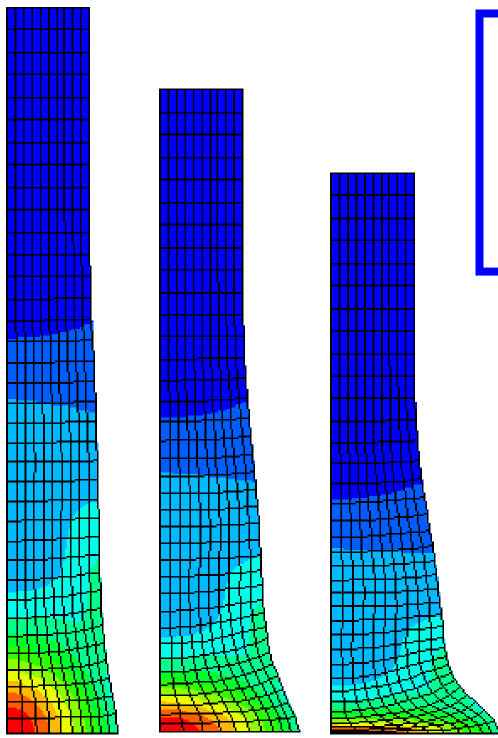
- **Example II: Energy dissipation of a threaded joint**
  (Energy dissipation of a shock through a threaded interface)
  → Illustration of effect screening and meta-modeling

- **Example III: Impact through a foam material**
  (Stress-strain behavior of a crushable foam)
  → Illustration of a quantification of epistemic uncertainty

Example I: The Taylor Anvil Impact

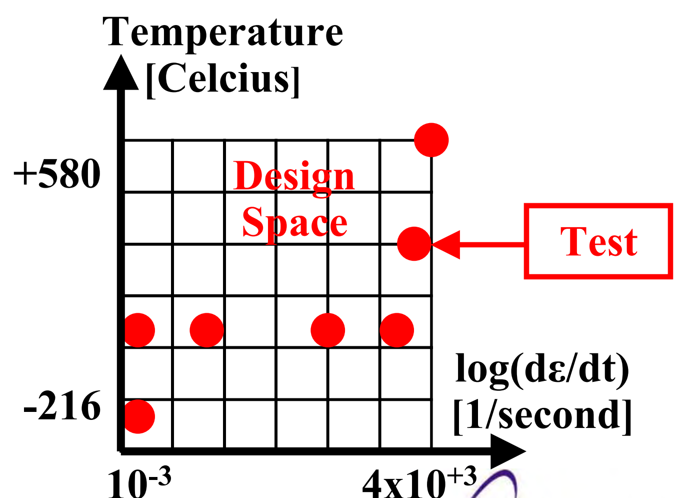
- **A strain-rate, temperature-dependent Zerilli-Amstrong model for the plasticity of high-strength HSLA-100 steel is implemented in a numerical simulation.**





Zerilli-Amstrong Plasticity Model

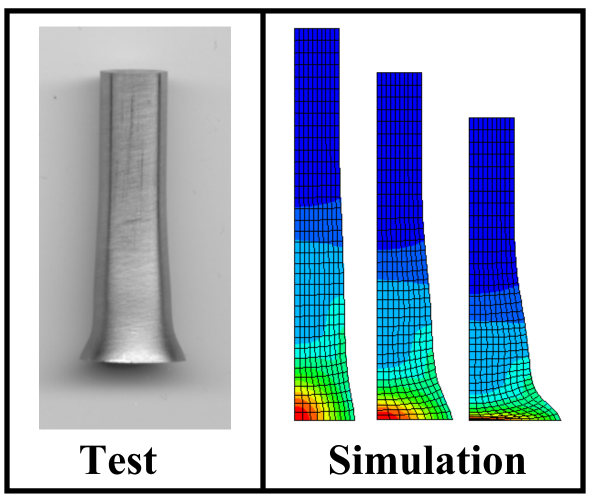
$$\varepsilon_P = C_o + C_1 e^{-C_3 T + C_4 T \log\left(\frac{d\varepsilon}{dt}\right)} + C_5 \varepsilon^N$$

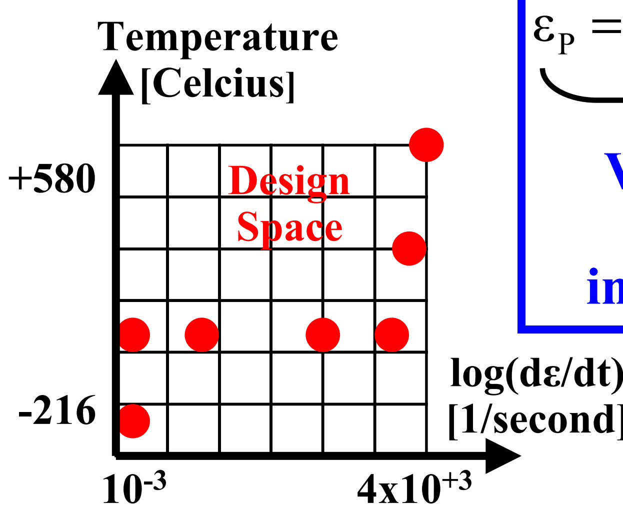


Taylor Anvil Impact (2 of 5)

- **The constitutive model must be validated within the design space defined by various combinations of strain-rates and temperatures.**

- **Be careful not to confuse the *control parameters* that define the design space and the *calibration variables*.**





$$\varepsilon_P = C_o + C_1 e^{-C_3 T + C_4 T \log(\frac{d\varepsilon}{dt})} + C_5 \varepsilon^N$$

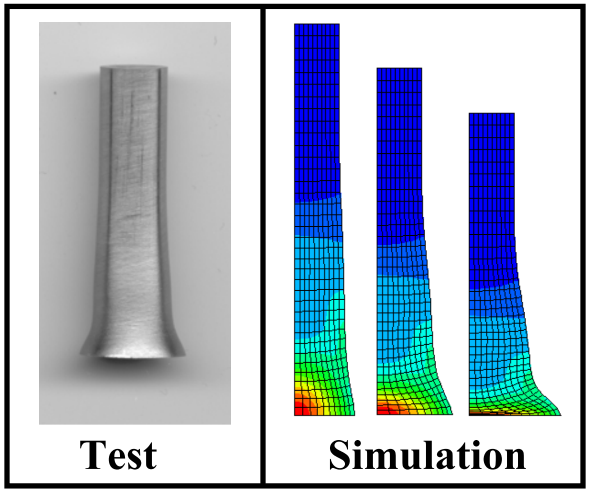
Validation requires the calibration of six independent coefficients.

Taylor Anvil Impact (3 of 5)

- **Coefficients of the Z-A model are calibrated using Bayesian inference (assuming normal distributions).**

$$\chi^2 = -2\log(\text{Prob}(p|y^{\text{Test}}))$$

$$\chi^2 = \sum_{k=1\ldots N} \left(\frac{y_k^{\text{Test}} - y_k(p)}{\sigma_y^{\text{Test}}} \right)^2 + \left(\frac{p - p_o}{\sigma_p} \right)^2$$



Posterior correlations.

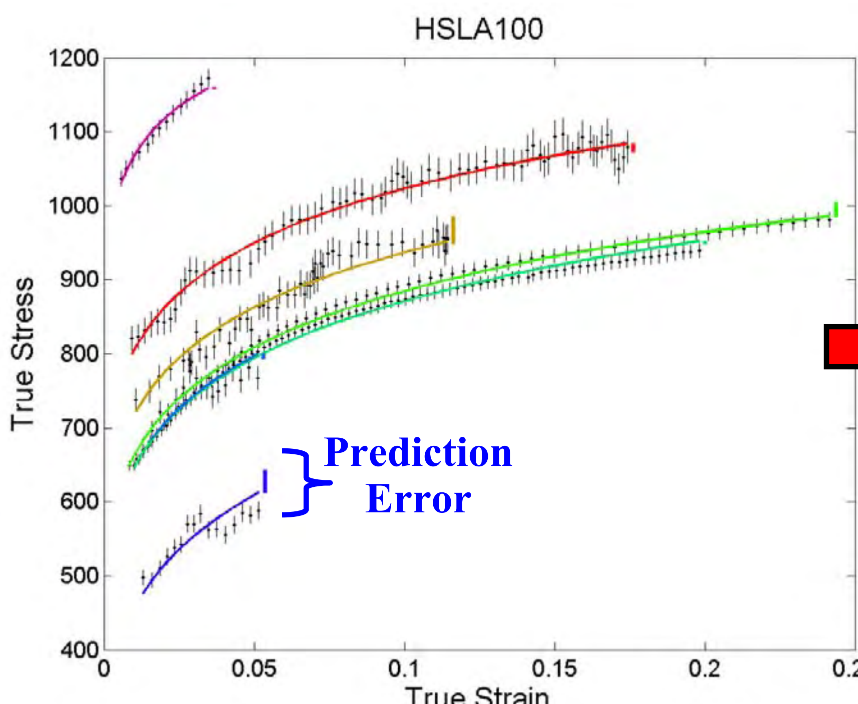
$C_{i,j}$	C_1	C_3	C_4	C_5	N
C_o	-8.3	+37.2	+20.7	-48.8	+26.7
C_1		+34.4	+31.1	+8.2	+13.0
C_3			+80.2	+45.3	-62.1
C_4				+27.1	-46.6
C_5					-86.0

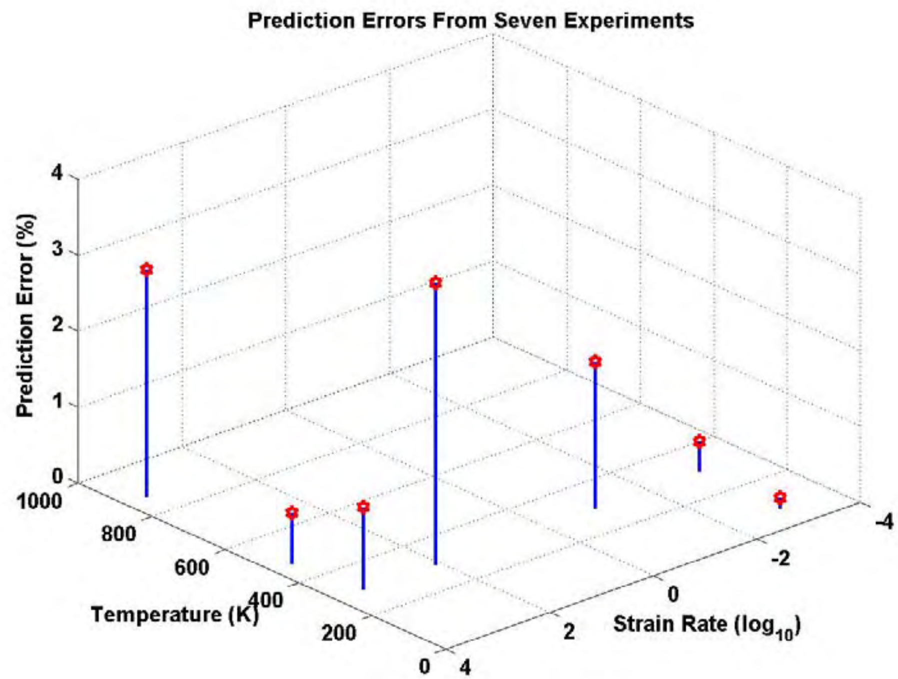
Calibrated variables.

p_i	C_o	C_1	C_3	C_4	C_5	N
Nominal	175.00	950.00	3.00e-3	8.50e-5	675.00	0.275
Calibrated	102.55	954.29	4.08e-3	11.73e-5	996.16	0.247
Deviation	32.90	62.72	0.59e-3	2.87e-5	22.38	0.021

Taylor Anvil Impact (4 of 5)

- **After the coefficients of the Z-A model have been calibrated using the seven physical experiments, test-analysis correlation is performed.**

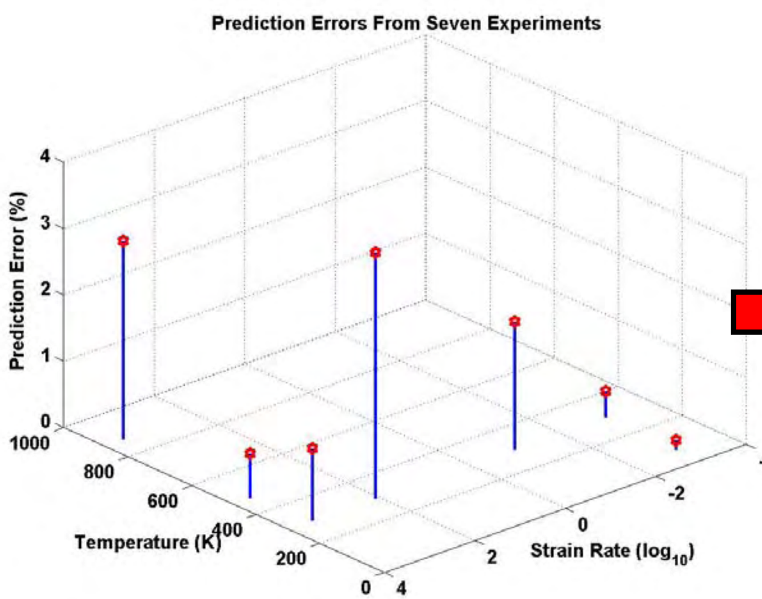


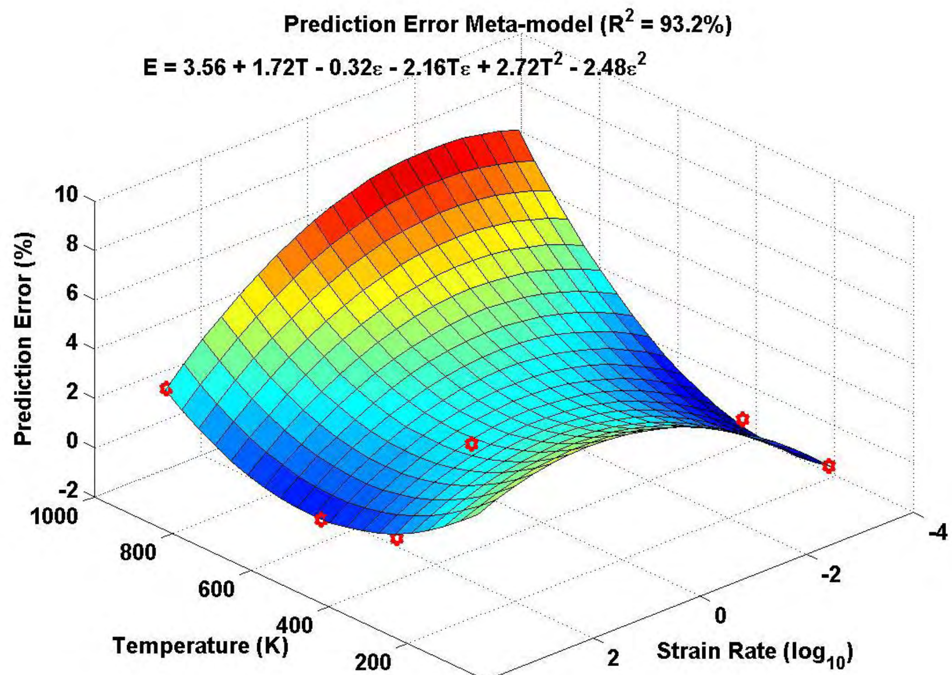


Note: The fidelity metric is the RMS error between measurements and predictions.

Taylor Anvil Impact (5 of 5)

- **A statistical, polynomial model is developed to estimate the constitutive model's accuracy away from settings where physical tests have been performed.**

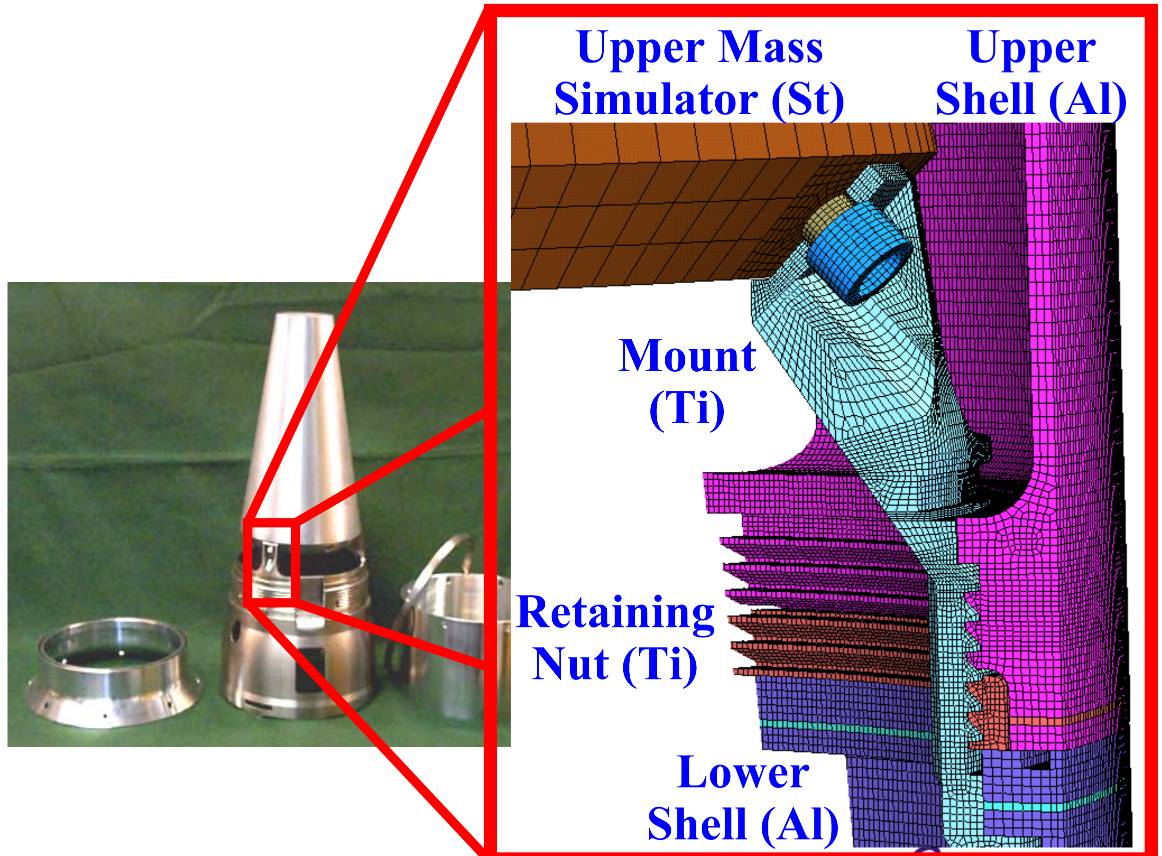




Note 1: Error meta-model: $e(T;S_R) = 3.56 + 1.72T - 0.32S_R - 2.16TS_R + 2.72T^2 - 2.48S_R^2$.
Note 2: Significance of the curve-fit to the seven data points, $R^2 = 93.2\%$.

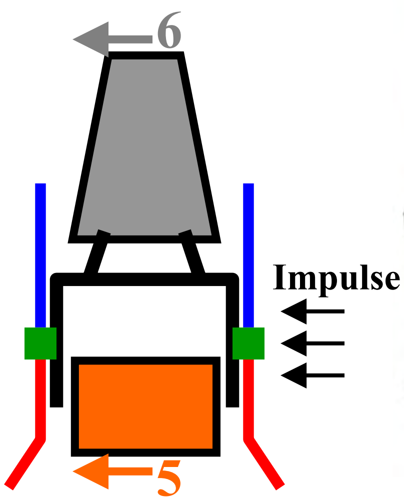
Example II: Threaded Joint Modeling

- This system is an assembly of components subjected to preloads and held together by a large thread.

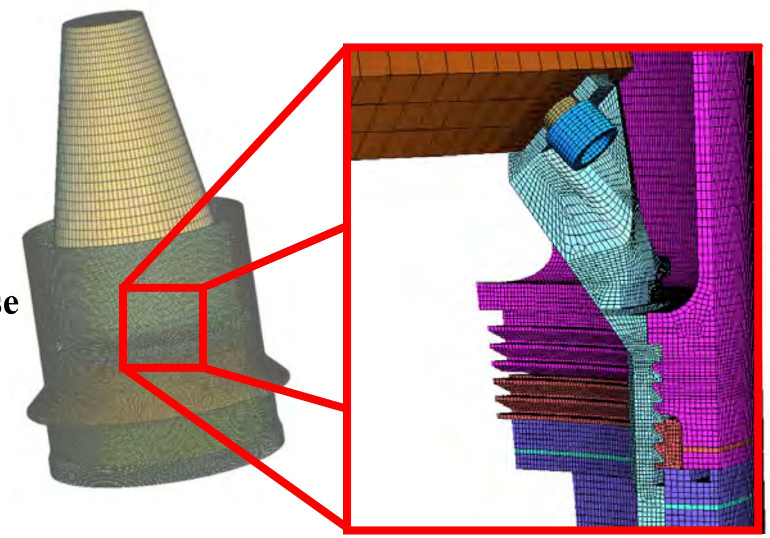


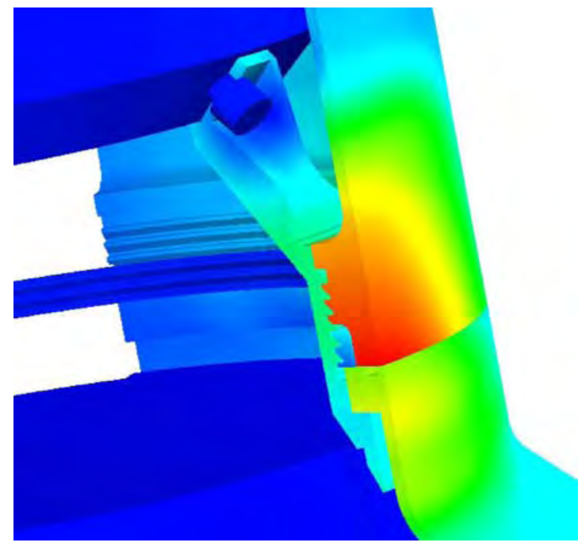
Threaded Joint Modeling (2 of 3)

- The LLNL/ParaDyn explicit simulation currently counts over 1.4 Million elements, 480 contact pairs, and 6 Million degrees-of-freedom.



Sensors 5 & 6


Detail of the Computational Mesh

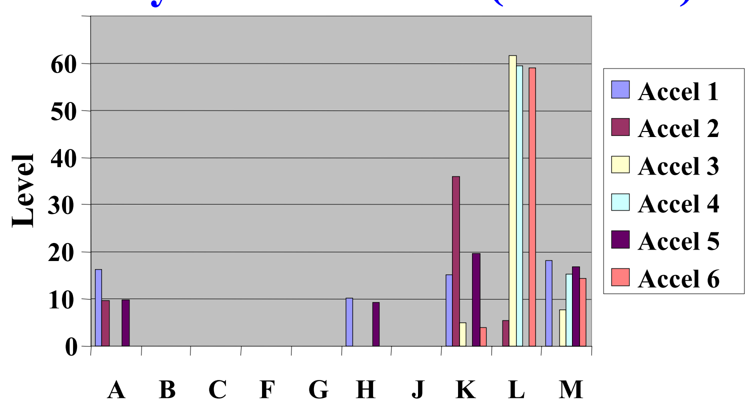

Displacement Contour

- Each run requires 4 hours to simulate 3 x 10^{-3} sec. of response on 504 processors of ASCI Blue Mountain.

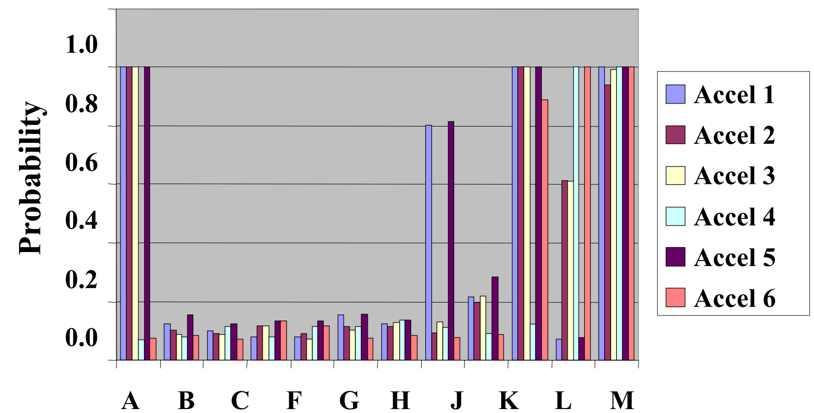
Threaded Joint Modeling (3 of 3)

- The random variables of the simulation are screened using designs of experiments and analysis of variance.

Analysis of Variance (ANOVA)



Bayesian Effect Analysis

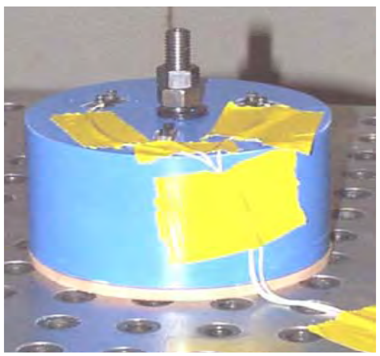


12 Random Variables of the Threaded Assembly Simulation

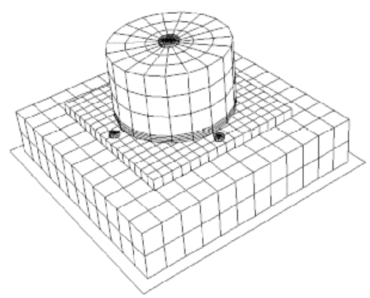
Preloads:	Static friction:	Kinetic friction:	Input Loading:
• A, Tape joint	~~• D, Al/Al static~~	~~• H, Al/Al kinetic~~	• M, Impulse level
~~• B, Retaining nut~~	~~• E, Ti/Ti static~~	• J, Ti/Ti kinetic	
~~• C, Upper shell~~	~~• F, Al/Ti static~~	• K, Al/Ti kinetic	
	~~• G, St/Ti static~~	• L, St/Ti kinetic	

Example III: Impact Through Foam

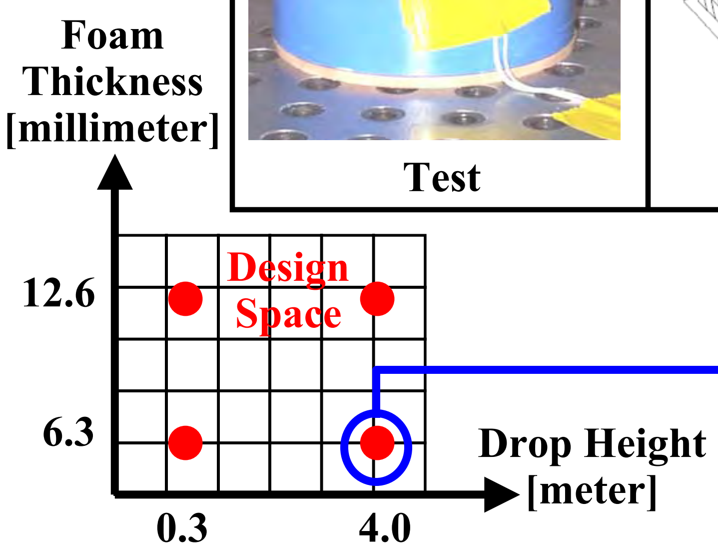
- **The numerical simulation predicts the acceleration signal transmitted through a layer of crushable foam material during an impact.**

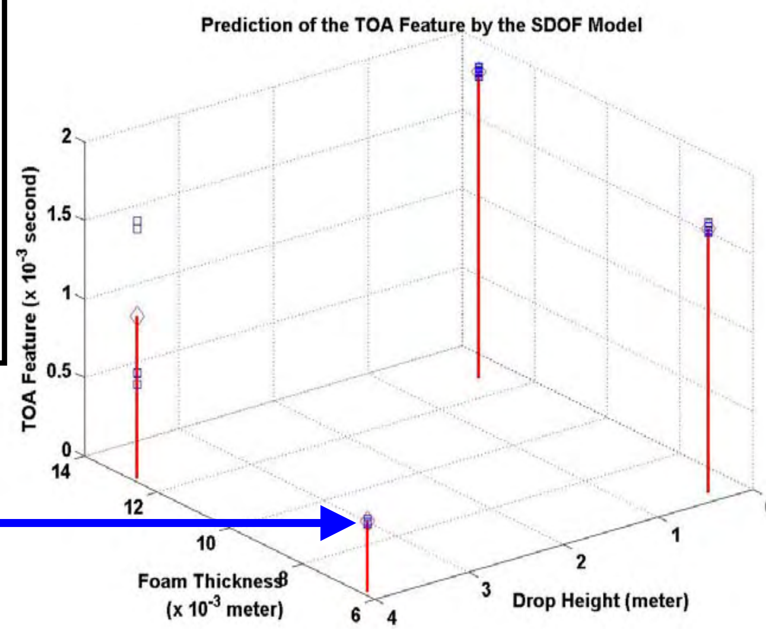


Test



Simulation





Impact Through Foam (2 of 3)

- **Statistical meta-models are developed to estimate the prediction accuracy away from testing conditions.**

Measurement of Features, y^{Test}

↓

Error metric

$$e(h;d) = \left(y^{Test} - y(h;d)\right)^T S_{yy}^{-1} \left(y^{Test} - y(h;d)\right)$$

↑

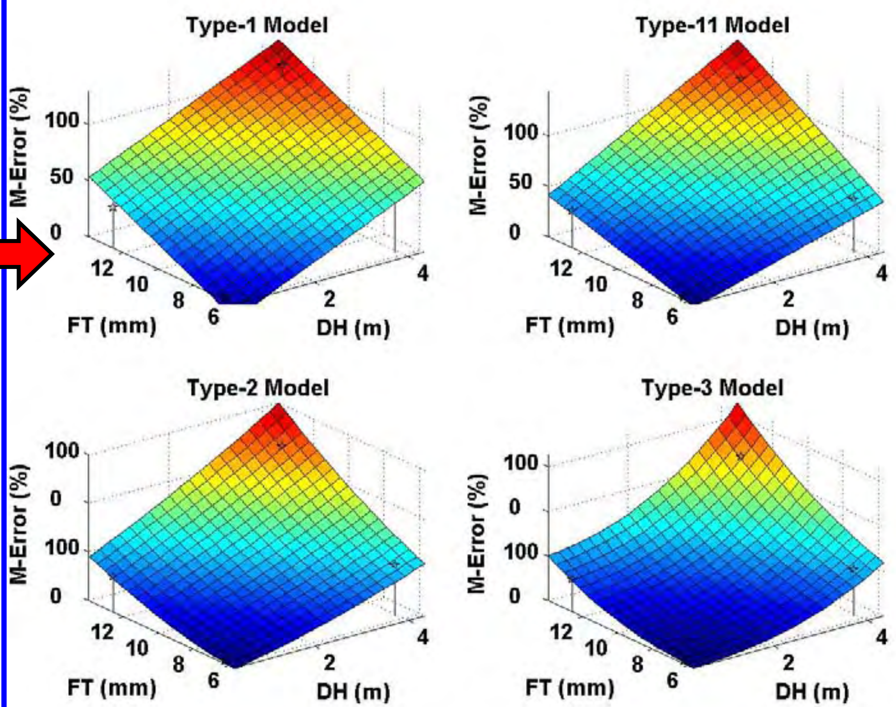
Prediction of Features, y

↑

Finite Element Simulation

↑

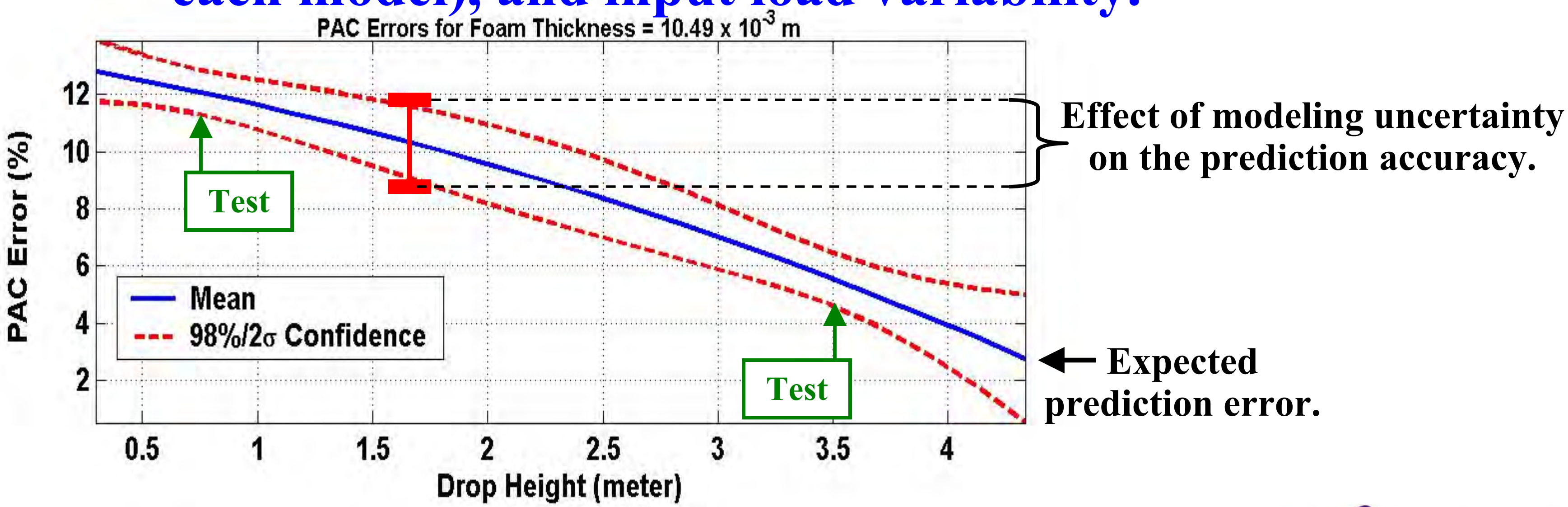
Material Model Type-1 OR **Material Model Type-11** OR **Material Model Type-2** OR **Material Model Type-3**



Meta-models of Prediction Accuracy

Impact Through Foam (3 of 3)

- Lack-of-knowledge is usually ignored (by making an assumption). Here, the effect of lack-of-knowledge on predictions is assessed.
- Uncertainty includes lack-of-knowledge about the type of material model, parameter variability (for each model), and input load variability.



Summary

- A model is validated when its predictive accuracy has been assessed throughout the design space of interest.
- Test-analysis correlation, finite element model updating and calibration are useful tools but a calibrated model is not necessarily a validated model.
- Model validation is application dependent.
- Model validation is still an area of active research to a great extent, especially when it comes to assessing the lack-of-knowledge and its effects. Uncertainty quantification is a critical component of predictive accuracy assessment.

Resources

- **Conferences:**
  - SAMO - Sensitivity Analysis of Model Output
    (Every 3 years; Next March 2004 in Santa Fe, NM; www.samo2004.org)
  - IMAC - Conference on Structural Dynamics
    (Every year; Next January 2004 in Dearborn, MI; www.sem.org)
  - SDM - AIAA/ASME Structural Dynamics Conference
    (Every year; Next April 2004 in Palm Springs, CA; www.aiaa.org)
- **Books:**
  - Patrick Roache (1998, code verification in fluid dynamics)
  - Hemez, Doebling, Anderson (2003?, V&V model in engineering)
- **Web resources:**
  - ASME PTC-60 Committee on Validation and Verification for Computational Solid Mechanics (www.usacm.org/vnvcsm/)
- **Short-courses:**
  - Bill Oberkampf (wloberk@sandia.gov)
  - Los Alamos Dynamics (www.la-dynamics.com)

References for UQ, V&V (1 of 2)

1. AIAA, "Guide for the Verification and Validation of Computational Fluid Dynamics Simulations," American Institute of Aeronautics and Astronautics, AIAA-G-077-1998.
2. Anderson, M.C., Hasselman, T.K., "Minimizing Parameter Correlation in Nonlinear Model Updating," *Proceedingof IMAC XVIII*, San Antonio, Texas, 2000.
3. Bishop, C.M., Neural Networks for Pattern Recognition, Oxford University Press, New York, 1998.
4. Chen, J.C., Wada, B.K., "Criteria for Analysis-Test Correlation of Structural Dynamic Systems," *Journal of Applied Mechanics*, June 1975, pp. 471-477.
5. Collins, J.D., Hart, G.C., Hasselman, T.K., Kennedy, B., "Statistical Identification of Structures," *AIAA Journal*, Vol. 12, No. 2, 1974, pp. 185-190.
6. Hasselman, T.K., Coppolino, R.N., Zimmerman, D.C., "Criteria for Modeling Accuracy: A State-of-the-Practice Survey," *Proceedings of IMAC XVIII*, San Antonio, Texas, 2000.
7. Hasselman, T.K., Anderson, M.C., Wenshui, G., "Principal Components Analysis for Nonlinear Model Correlation, Updating and Uncertainty Evaluation," *Proceedings of IMAC XVI*, Santa Barbara, California, 1998, pp. 664-651.

References for UQ, V&V (2 of 2)

8. Hemez, F.M., Doebling, S.W., “A Validation of Bayesian Finite Element Model Updating for Linear Dynamics," *Proceedings of IMAC XVII,* Kissimmee, Florida, 1999, pp. 1545-1555.

9. Hemez, F.M., Wilson, A.C., Doebling, S.W., “Design of Computer Experiments For Improving an Impact Test Simulation,” *Proceedings of IMAC XIX,* Kissimmee, Florida, 2001, pp. 977-985.

10. Hemez, F.M., Doebling, S.W., “Inversion of Structural Dynamics Simulations: State-of-the-art and Orientations of the Research,” *25th ISMA International Conference on Noise and Vibration Engineering*, Leuven, Belgium, Sept. 13-15, 2000.

11. McKay, M.D., Beckman, R.J., Conover, W.J., “A Comparison of Three Methods For Selecting Values of Input Variables in the Analysis of Output From a Computer Code,” *Technometrics*, Vol. 21, No. 2, 1979, pp. 239-245.

12. Mottershead, J.E., Friswell, M.I., “Model Updating in Structural Dynamics: A Survey,” *Journal of Sound and Vibration*, Vol. 162, No. 2, 1993, pp. 347-375.

13. Myers, R.H., Montgomery, D.C., Response Surface Methodology: Process and Product Optimization Using Designed Experiments, Wiley Inter-science, New York, 1995.

14. Saltelli, A., Chan, K., Scott, M., Editors, Sensitivity Analysis, John Wiley & Sons, 2000.