SI4000 Summer 2010
Systems Engineering Colloquium
Naval Postgraduate School
15 July 2010

Simulation Supported Decision Making

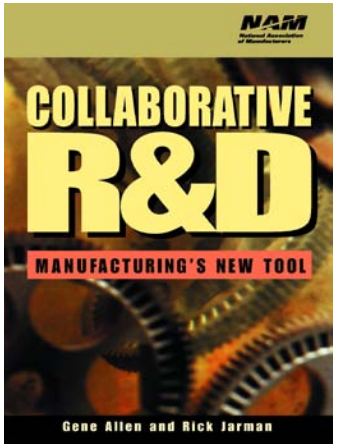
Gene Allen

Naval Surface Warfare Center Carderock Division

Gene Allen - Introduction

- **OBJECTIVE: TO PASS ON WHAT I KNOW on SIMULATION**
- **CAREER FOCUS:**

  HOW TO USE COMPUTERS TO DO HELP MAKE BETTER DECISIONS



- **PRACTICE USED:**

  User-driven COLLABORATION to commercialize technology
- **EXPERIENCE:**
  - **Private sector: 1993-2008 MSC Software, Decision Incite**
  - **Non-profit: 1989-1993 NCMS**
  - **Government:**
    - **Hill – Senate Majority leader staff, 1987-1989**
    - **U.S. Navy**
      - **Active (5 yrs) - Qualified Nuclear Engineer/SWO, CGN Command Duty Officer, Wartime OOD, Nuclear Training Officer, PreCom CRA**
      - **Reserve Officer on CNO staff (15 yrs - N4, OP611)**
      - **SSBN Acquisition (Booz, Allen & Hamilton - 3 yrs)**
      - **Presently with Navy Surface Warfare Center Carderock**
- **EDUCATION: Nuclear Engineering, MIT, 1978**

GEA Background in

Development of Computer Simulation Technology

1978-1986	1987-1992	1992-1997	1996-2002	2002-
Engineering Foundation	DICE DARPA Initiative in Concurrent Engineering	RRM Rapid Response Mfg	RDCS Robust Design Compute System	Stochastics
Newport News, EB, Westghouse GE	GEAE, GECRD CMU, RPI, WVU	Ford, GM, TI P&W, Boeing, Kodak Tech Suppliers	Boeing-Rocketdyne Ford MSC	MSC, IBM, Engineous, Ontonix, Decision Incite
Nuclear Navy	DARPA	NIST	DARPA	


41
Decisions Result in Actions


BEWARE OF OUTLIERS –
THEY ARE NOT
'JUST ANOMALIES'
Bad Decisions Result
when all possibilities
are not taken into account

Good Decisions

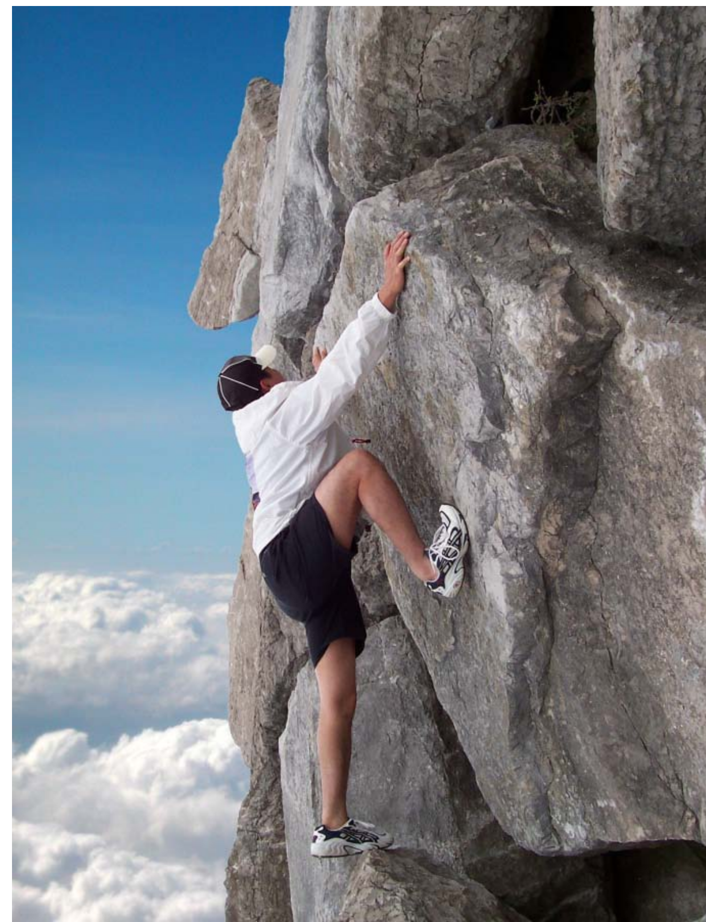
Based On Understanding:

- All possible results
- BEFORE taking action

Understanding from Knowledge

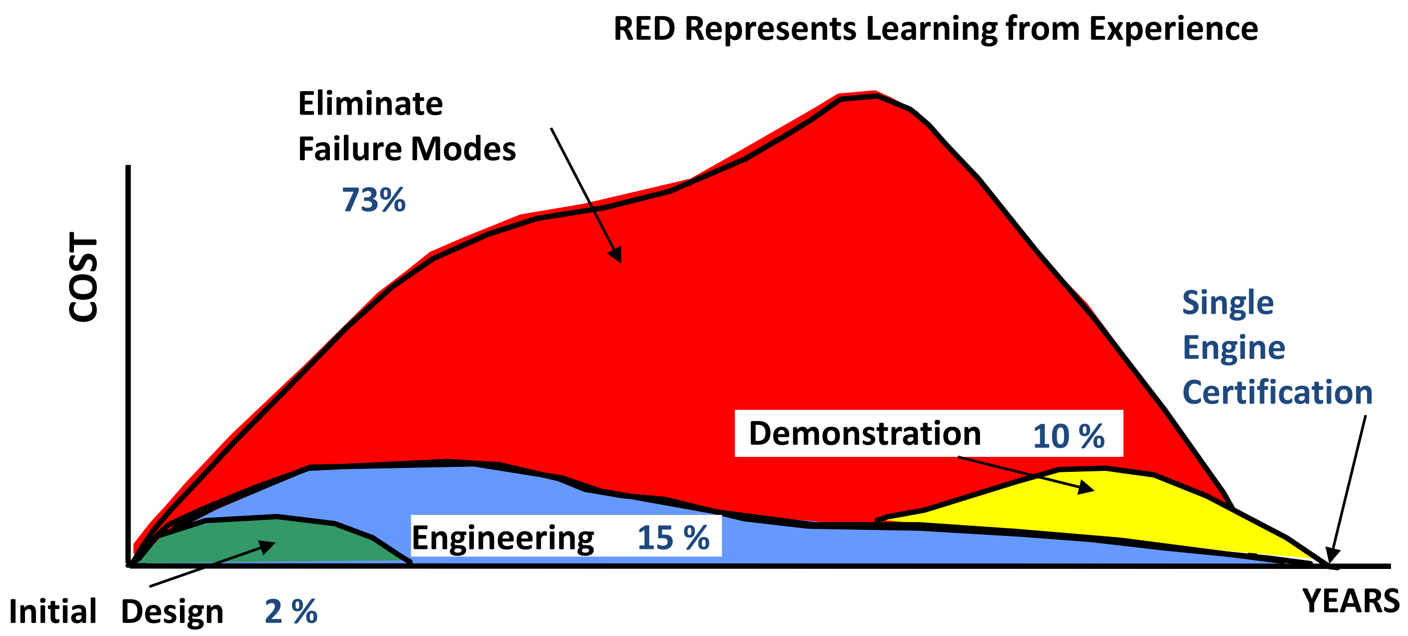
Knowledge is based on

Education & Experience



Experience

- **Takes Time and Money**
  - **Von Braun had hundreds of V2 failures before the Saturn V**



  - **Graph for Saturn V rocket engines: $2.2 Billion, 9 years**

Source: Rocketdyne

Engineering before Computers

Successful Engineering Cultures

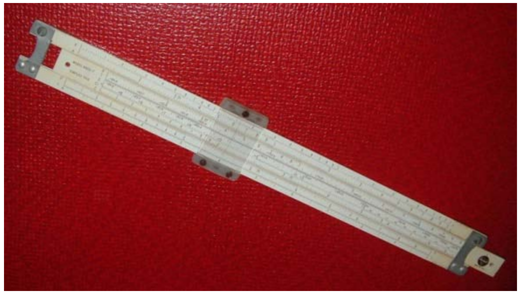
- U.S. Aerospace Industry
  - Produced: U-2, X-15, Saturn V, C-130, B-52

- U.S. Navy Nuclear Program
  - Decades of dynamic operations of hundreds of nuclear power plants without casualties

- Combined Theory and Practice

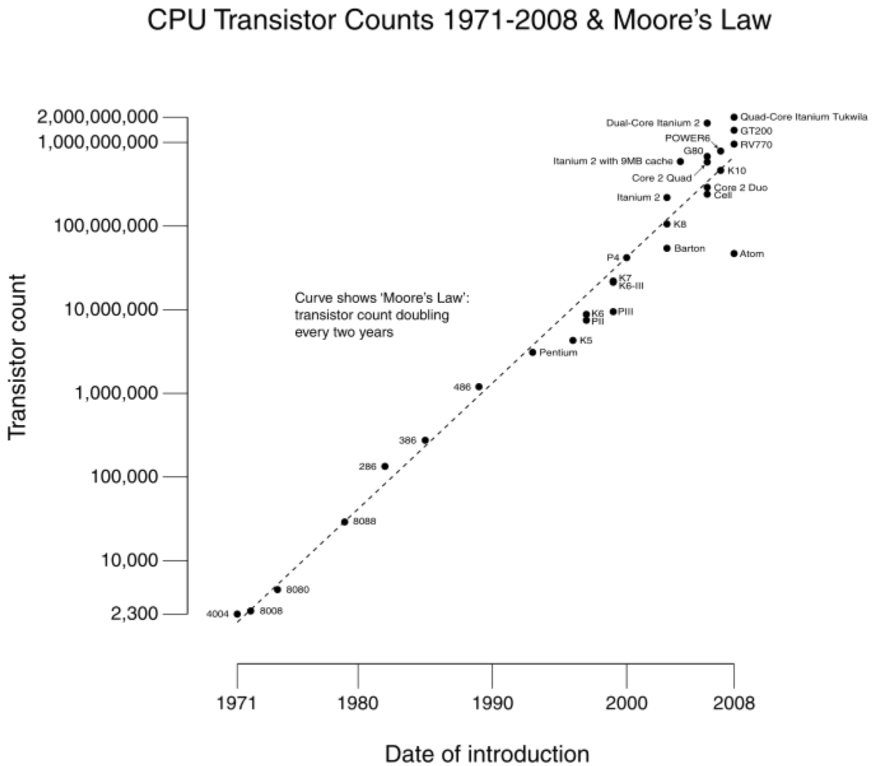
Tools Engineers use to help do better engineering:

- Slide Rule
- Calculator



Moore’s Law has provided an unprecedented capability

- “Commodity” computing
- CPU/hr costs in ¢
- <u>We have not yet taken advantage of this</u>



Simulation can *Accelerate* Experience

- Understand How Products/Processes Function
- New Processes to Forecast Risk
  - Complexity used as metric.
- Identify:
  - Major factors driving functionality or contributing to potential failure.
  - Combinations of factors that lead to unexpected situations (outliers) that lead to failures.
  - Trends towards unstable situations.
- These Abilities Exist Today!
  - Due to advances in compute capability

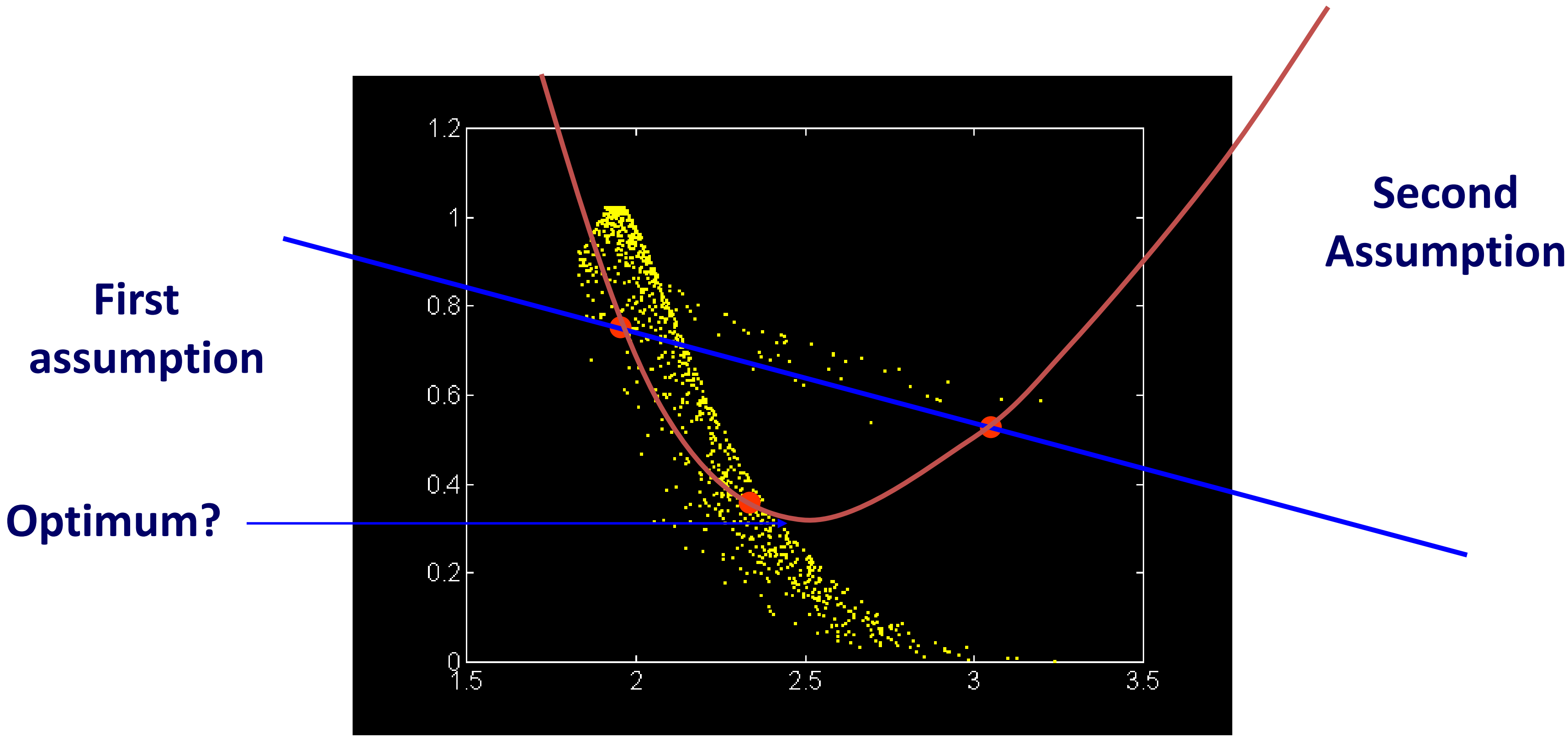
Commodity Computing can be used to provide tools to simulate reality to:

- Compliment experience
- Learn what we do not know

Tools need to be:

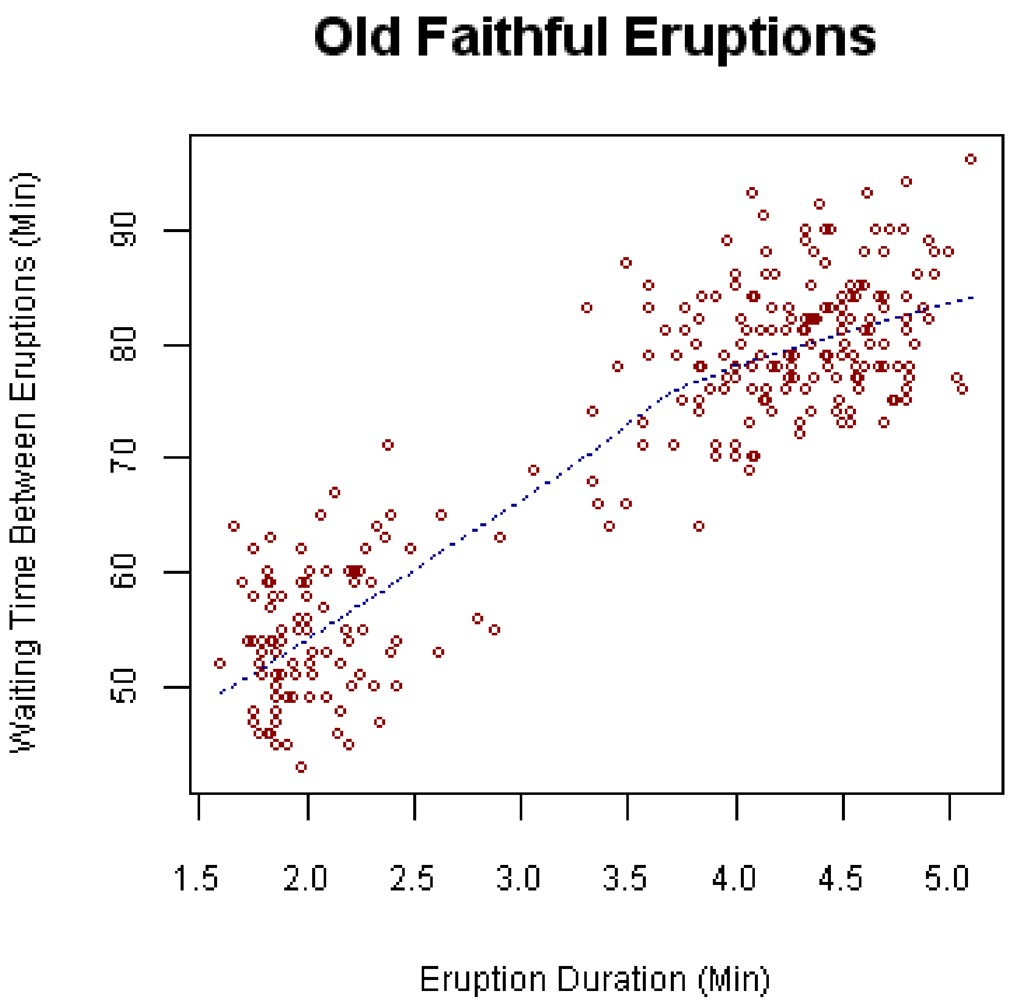
- Accurate
  - Results readily verified
  - Independent results replication
- Easy to Use
- Interoperable

Improved Understanding



Reality - Each Point:
- **May be the result of a test, or an analysis.**
- **Is the result of different combinations of variables.**

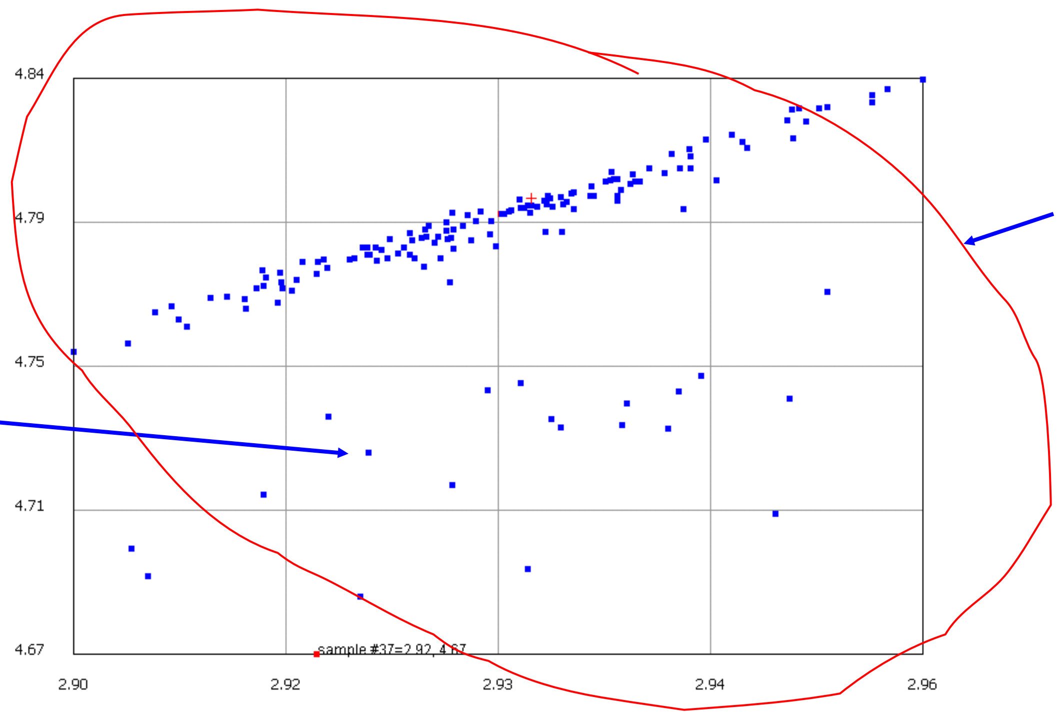
EXPERIENCE – Forms Engineering Knowledge Base



Analysis vs. Simulation

One run vs. Many runs

<u>Single</u>
computer
run =
<u>Analysis</u>



<u>Multiple</u>
computer
runs =
<u>Simulation</u>

Understanding = Knowing the topology and structure of the data cloud.

Simulation must address Variation

- **Material Variation Examples**

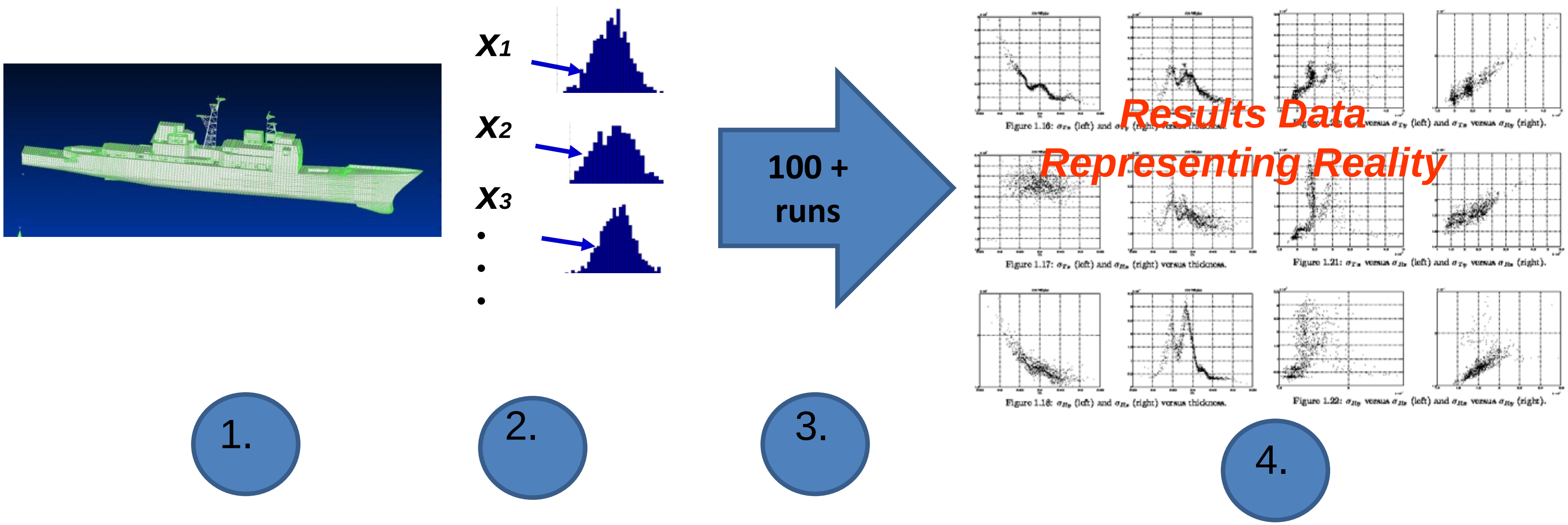
MATERIAL	CHARACTERISTIC	VARIATION
Metallic	Rupture	8-15%
	Buckling	14%
Carbon Fiber	Rupture	10-17%
Screw, Rivet, Welding	Rupture	8%
Bonding	Adhesive strength	12-16%
	Metal/metal	8-13%
Honeycomb	Tension	16%
	Shear, compression	10%
	Face wrinkling	8%
Inserts	Axial loading	12%
Thermal protection (AQ60)	In-plane tension	12-24%

- **Similar Variation with Geometry**
- **More Variation with Forces**

It’s the Way the World Is

Source: Klein, M., Schueller, G.I., et.al.,Probabilistic Approach to Structural Factors of Safety in Aerospace, Proceedings of the CNES Spacecraft Structures and Mechanical Testing Conference, Paris, June 1994, Cepadues Edition, Toulouse, 1994.

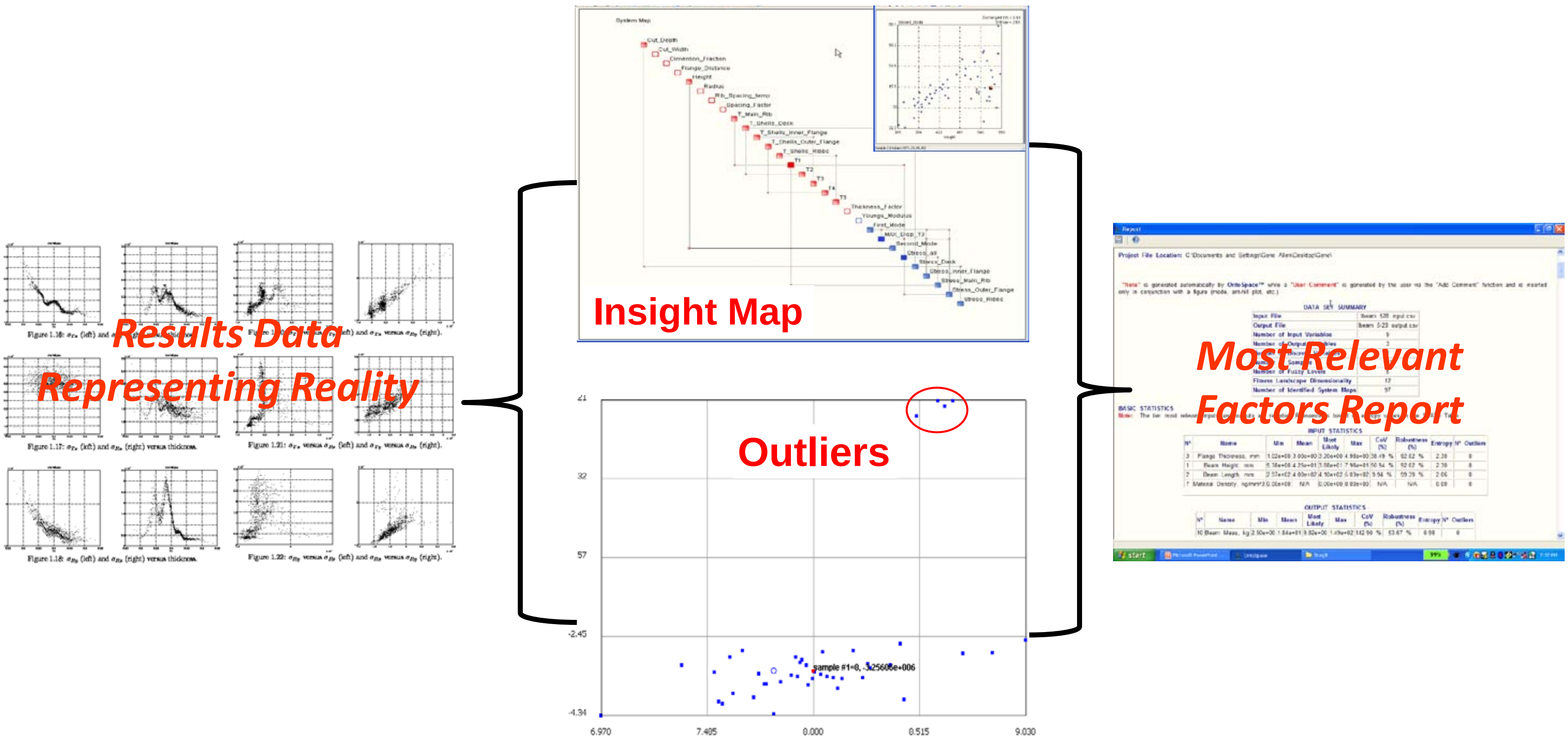
Simulation from an Analysis Model



PROCESS STEPS

1. **Start with a product or process computer analysis model**
2. **Replace all discreet inputs with ranges and distributions**
3. **Run model ≈ 100 + times randomly changing all variables**
4. **Result is a Multi-dimensional data cloud that represents reality**

Simulation Results to Information



The CHALLENGE

PREVENT FAILURES

Of Complex Systems & Programs

We Don't Know What We Don't Know

To LEARN WHAT WE DO NOT KNOW

We Need To

MINIMIZE ASSUMPTIONS

Common Foundation Assumptions:

- Continuity
- Mathematical Constructs
  - Gödel’s incompleteness theorems:
    - Any computable axiomatic system that is consistent, cannot be complete;
    - The consistency of the axioms cannot be proved within the system.
- General Over-Simplifications

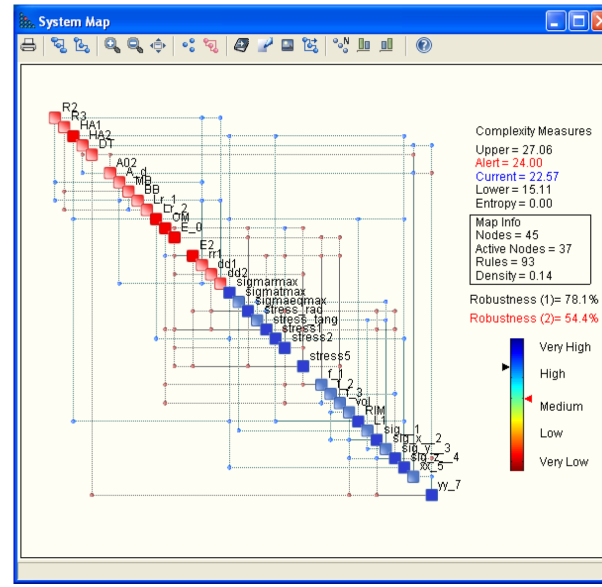
Assumption that Violates Physics

$$\left.\frac{\partial f}{\partial x}\right|_{\substack{y=a\\ z=b}}$$

This is NOT true

$$\boxed{\begin{aligned} y&=a\\ z&=b \end{aligned}}$$

From Analysis Models to Knowledge

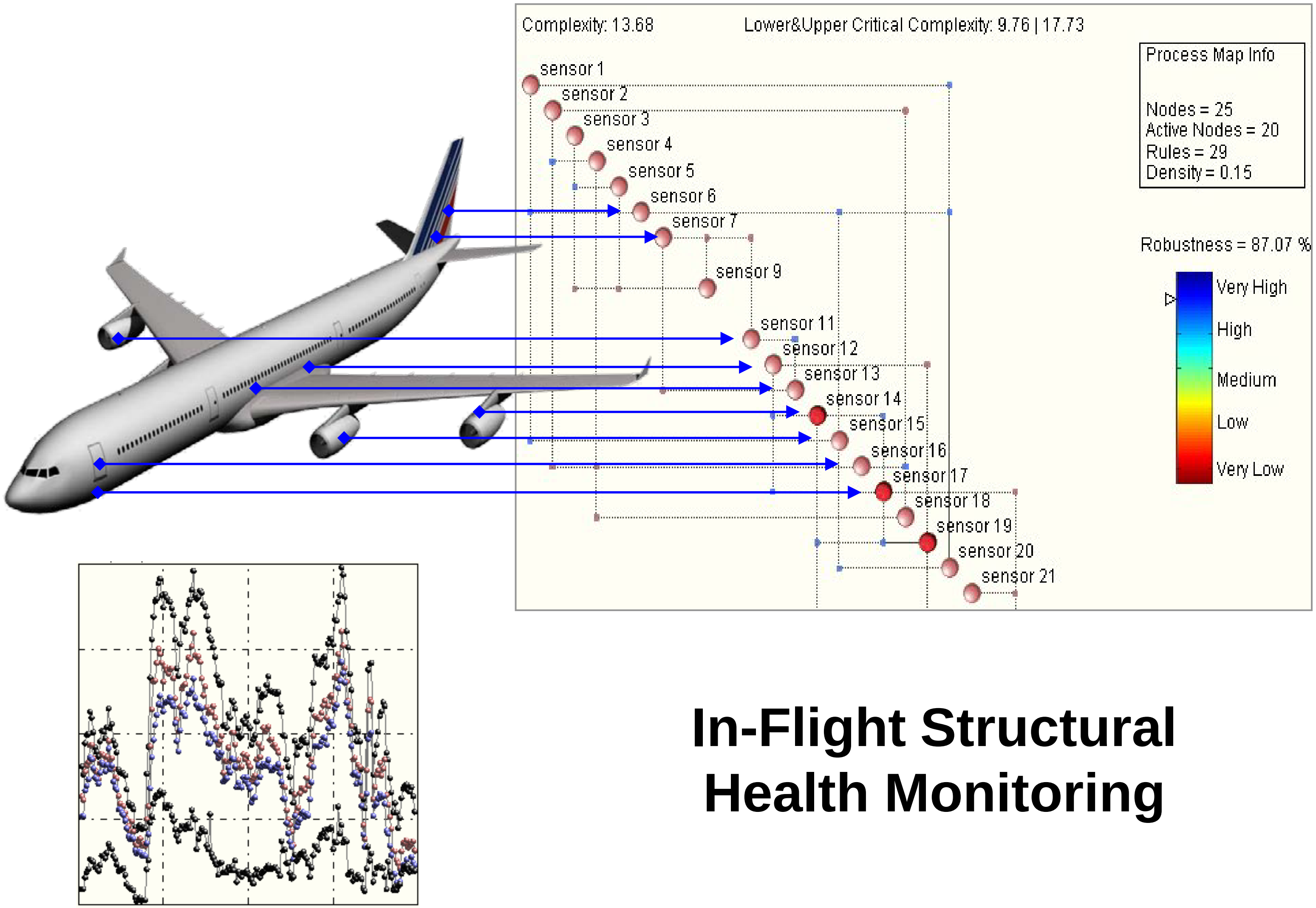


Model → **Simulation** → **Data** → **System Map**

The System Map shows:

- Relationships between parameters
- Complexity levels
- Outliers

From Raw Data to Knowledge

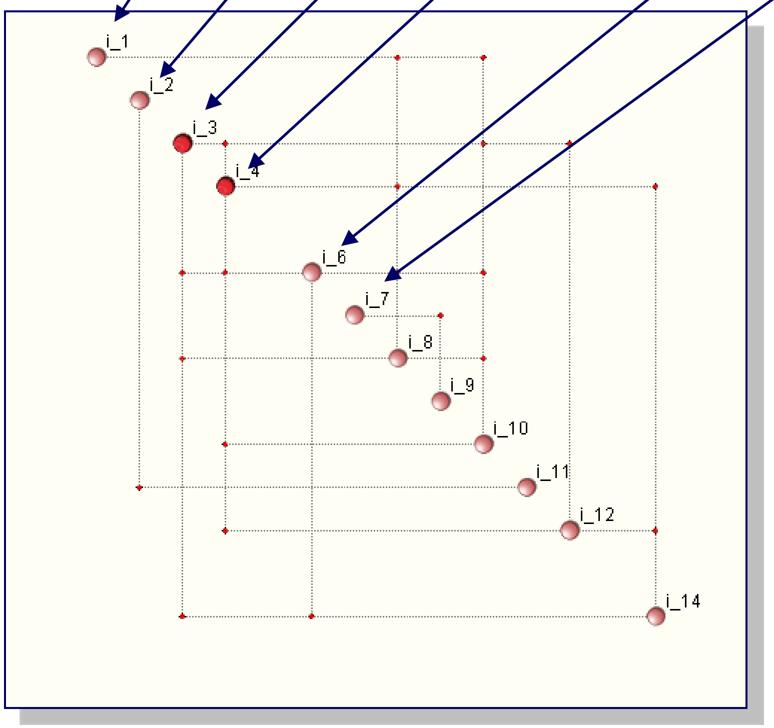


In-Flight Structural Health Monitoring

Complexity and System Maps

Complexity is a system attribute which reflects how information is organized and flows. Complexity can be measured and managed.

	Variable 1	Variable 2	Variable 3	Variable 4	Variable 5	Variable 6	Variable 7	Variable 8	Variable 9	Variable 10	Variable 11	Variable 12	Variable 13	Variable 14
measurement 1	2.7202888e+003	7.4722596e+001	3.3732330e+001	1.0184837e+002	9.0967402e+001	1.3025740e+000	5.8491109e+002	2.0883948e+003	5.7746246e+002	2.0745664e+003	7.7027814e+001	4.7343477e+001	9.3212680e+001	1.4980173e+000
measurement 2	2.7202888e+003	7.4722596e+001	3.3732330e+001	1.0184837e+002	9.0967402e+001	1.3025740e+000	5.8491109e+002	2.0883948e+003	5.7746246e+002	2.0745664e+003	7.7027814e+001	4.7343477e+001	9.3212680e+001	1.4980173e+000
measurement 3	2.7472810e+003	7.5254340e+001	3.2769872e+001	9.9563658e+001	9.0359902e+001	1.2958549e+000	5.7800692e+002	2.1200529e+003	5.7044970e+002	2.1071991e+003	7.7609260e+001	4.4859857e+001	9.2513456e+001	1.4691249e+000
.														
.														
measurement 498	2.6963742e+003	7.4267584e+001	3.2300819e+001	9.9709857e+001	9.0880438e+001	1.2846033e+000	6.0231885e+002	2.0989936e+003	5.9500231e+002	2.0844376e+003	7.6470372e+001	4.7246972e+001	9.2445500e+001	1.4934932e+000
measurement 499	2.7161385e+003	7.4648496e+001	3.1284797e+001	9.7180980e+001	8.9747118e+001	1.2855390e+000	5.7932586e+002	2.1128403e+003	5.7174746e+002	2.1006561e+003	7.7008998e+001	4.2349411e+001	9.3013352e+001	1.4359544e+000
measurement 500	2.6995332e+003	7.5451365e+001	3.4156452e+001	1.0258146e+002	9.0543986e+001	1.3110797e+000	5.7014063e+002	2.0590153e+003	5.6257964e+002	2.0467060e+003	7.7843365e+001	4.4955206e+001	9.2165592e+001	1.4766825e+000



A system can be represented as:

- **Nodes with each node representing a characteristic of the system, and**
- **Links between nodes representing the relationships between characteristics.**

System maps are built on analysis models or raw user data.

- **Links are established automatically using Ontospace software.**

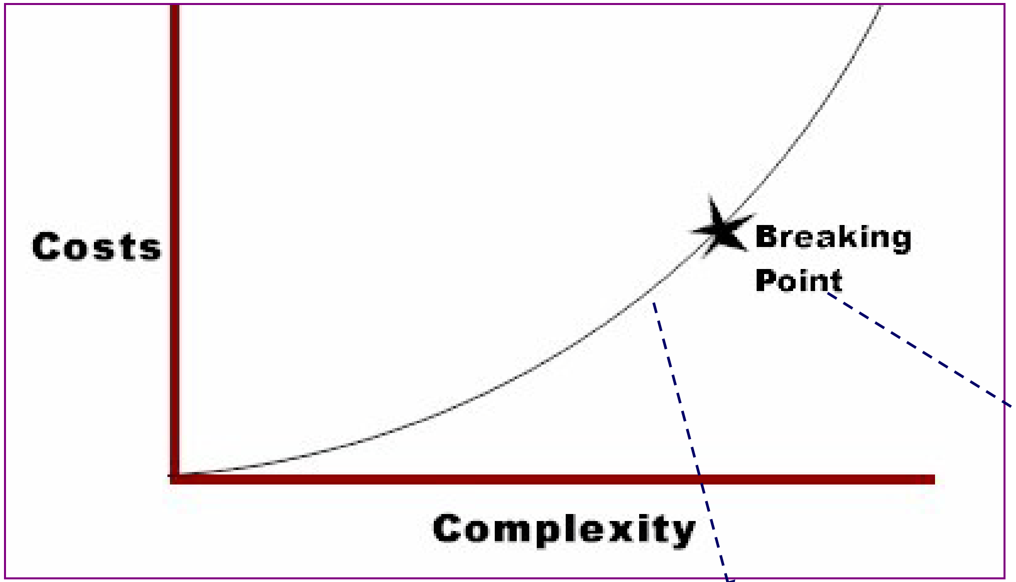
The Complexity of a system is based on the number and nature of the nodes and links.

Complexity x Uncertainty = Fragility
Applies to product and enterprise:

$$C_{\text{business model}} \times U_{\text{environment}} = F_{\text{enterprise}}$$

Managing Complexity

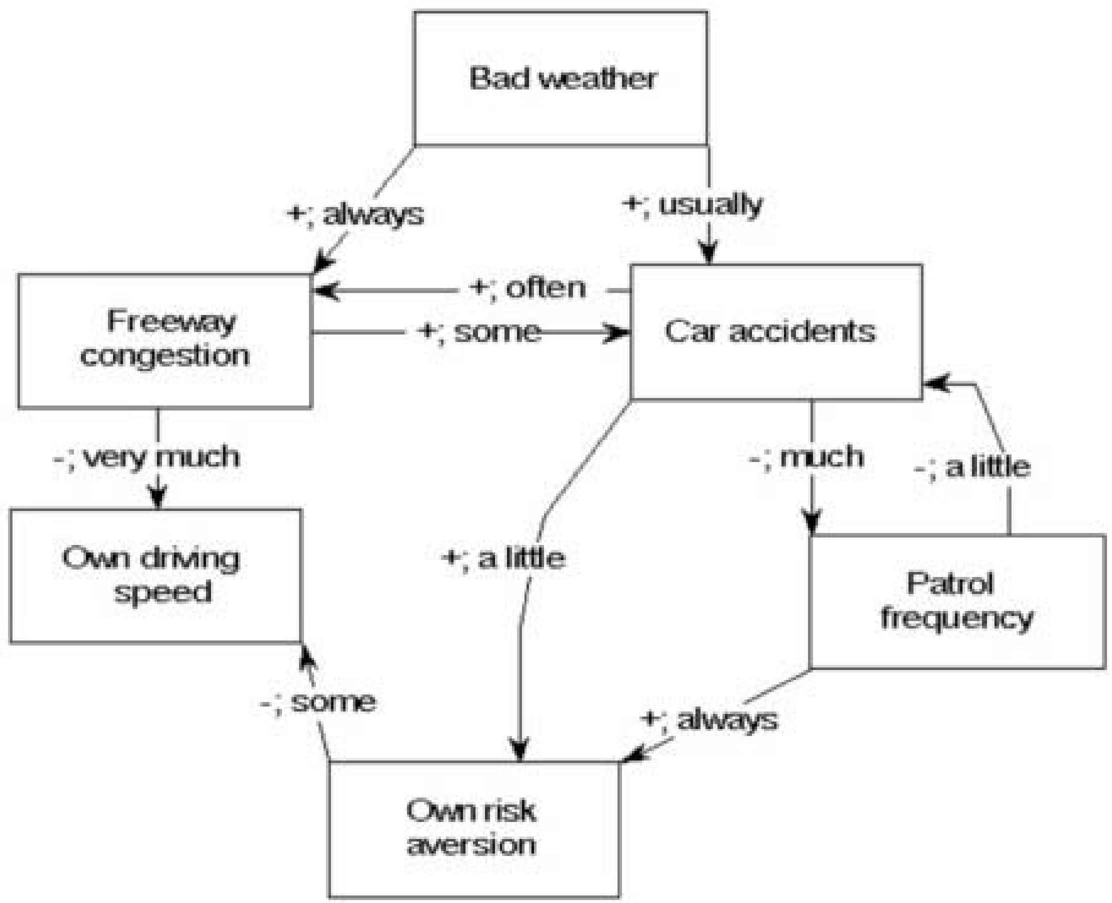
- More complex systems provide greater functionality and can be more stable in having redundancy and flexibility.
- However, complexity cannot grow indefinitely, and has a maximum, called critical complexity. A system becomes fragile and vulnerable near this point.



As complexity increases, the cost of managing the complexity increases at an exponential rate until the system finally collapses.

Thomas Frey, The da Vinci Institute

FUZZY COGNITIVE MAP

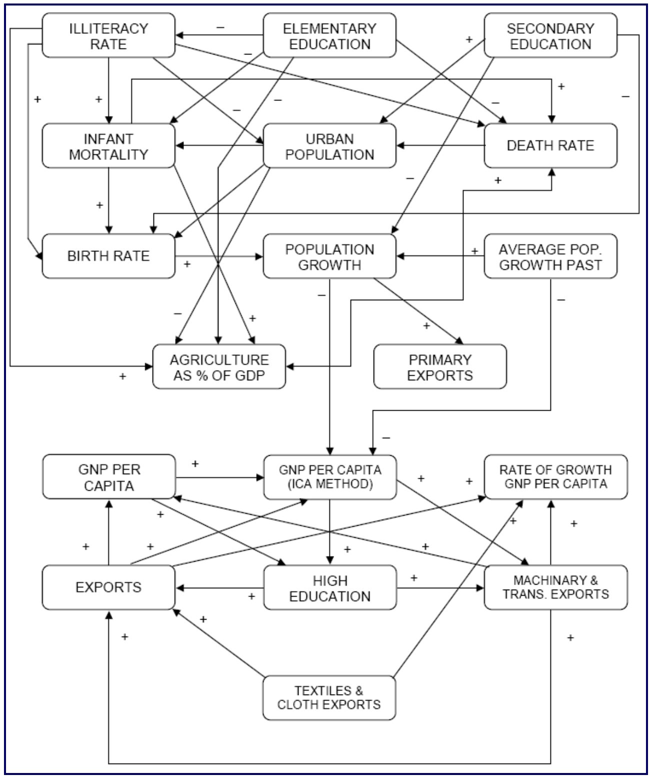


This Fuzzy Cognitive Map shows how bad weather can affect how fast you drive on a California freeway in the daytime.

From "*Fuzzy Thinking*" by Bart Kosko (Flamingo, 1994).

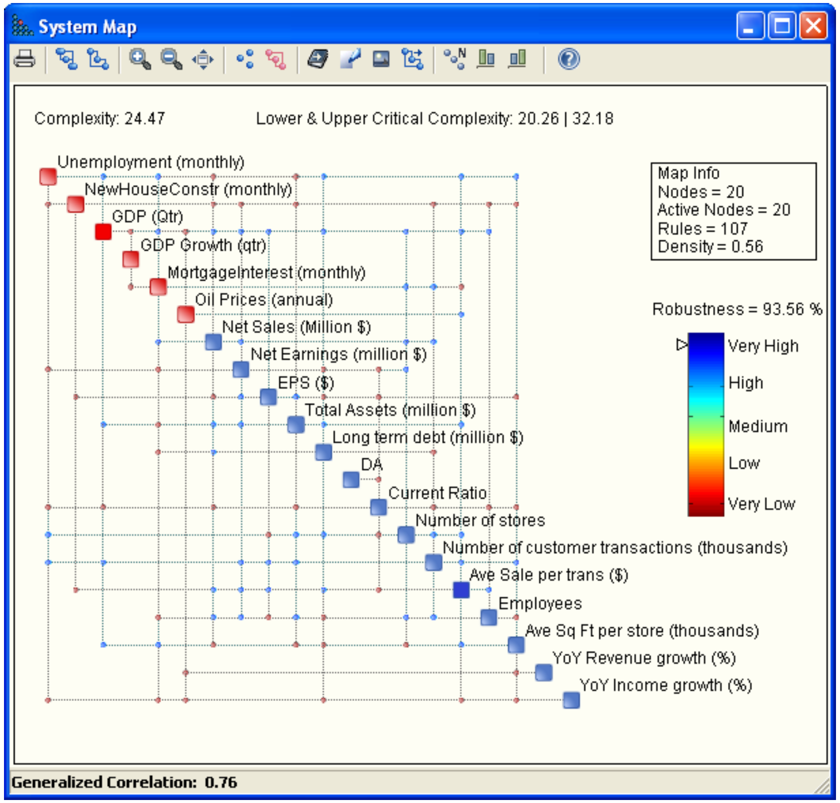
Decision Maps for Understanding Influence and Information Flow

Past



Hand drawn, subjective

Today



Computer generated, extracts information from data

Simulation - Input to Knowledge

- Can compliment education and experience
- Can be validated by all members of a team

A good team will beat a superstar

Simulation-generated Information Infrastructure

- The pieces are in place
- Computer HW, SW, IT, Web technologies
- Advance "Proven" and new methods to leverage these capabilities:
  - Incorporate variability
  - Manage complexity
  - Methods are scalable

  provides transparent credibility
- Expertise

USE INFORMATION FROM SIMULATION

TO LEARN What We Do Not Know

Inputs

- Less Assumptions
- More Variables

Outputs - Identify WHAT'S IMPORTANT

- Significant Variables
- Outliers

Process

- Scalable to people and disciplines
- Replicable

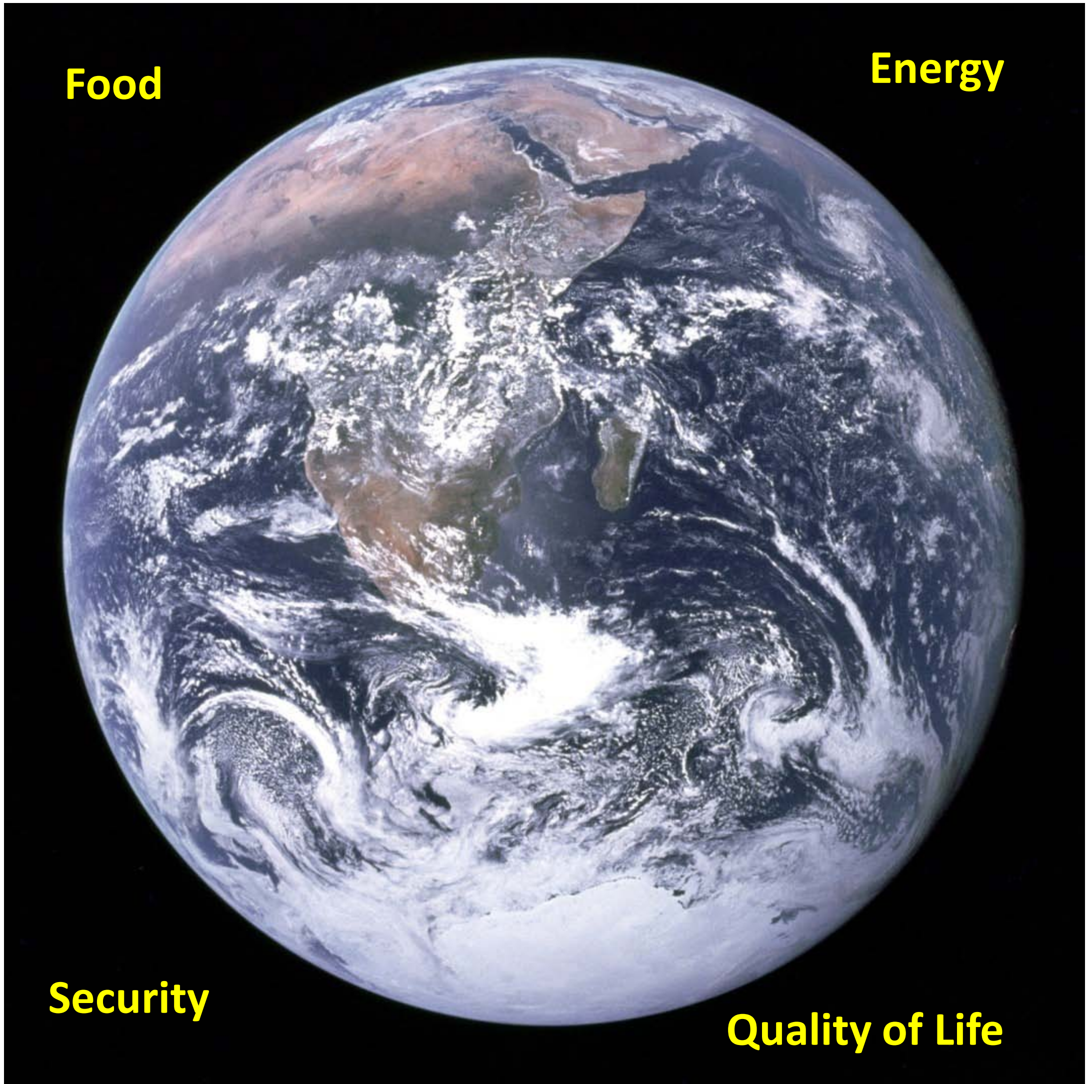
Simulation for Better Understanding

- Leverage Commodity Computing
- Augment and Compliment Experience

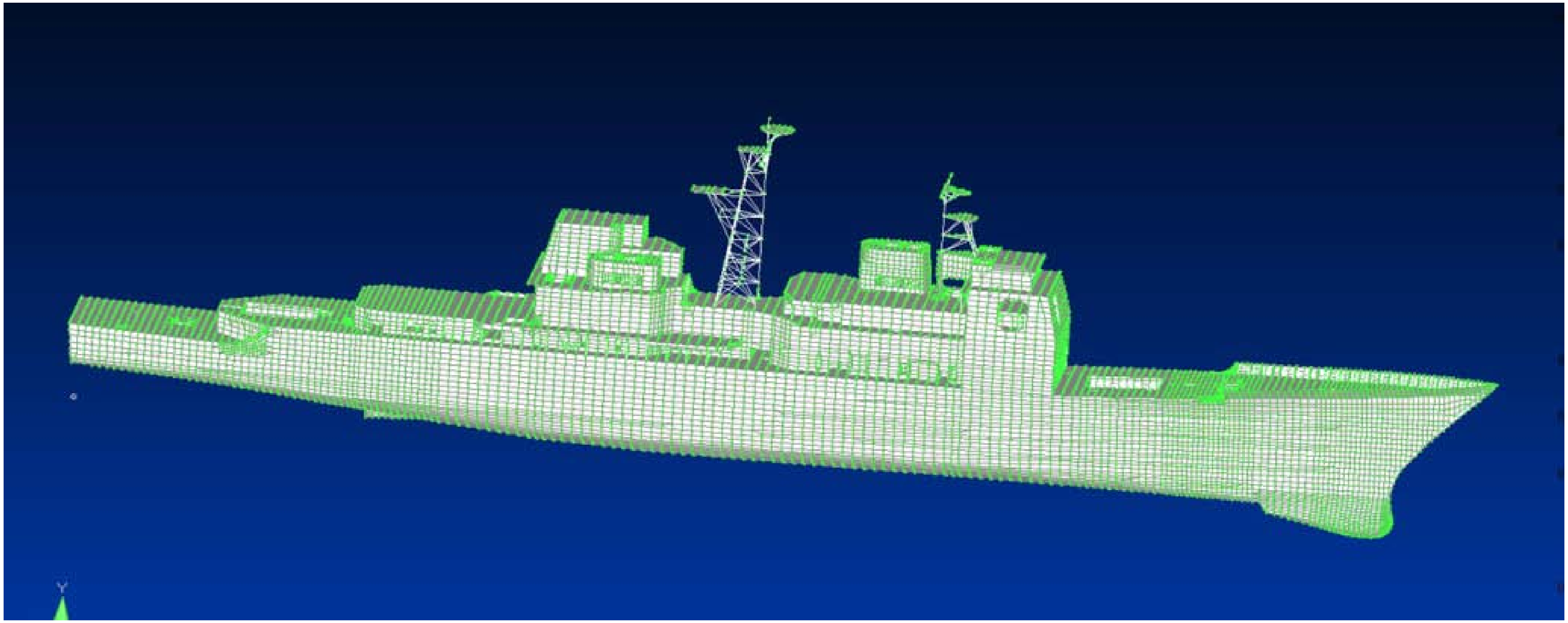
Can we learn from Virtual Experience?

Simulation for Better Understanding

Can We Afford not to LEARN?



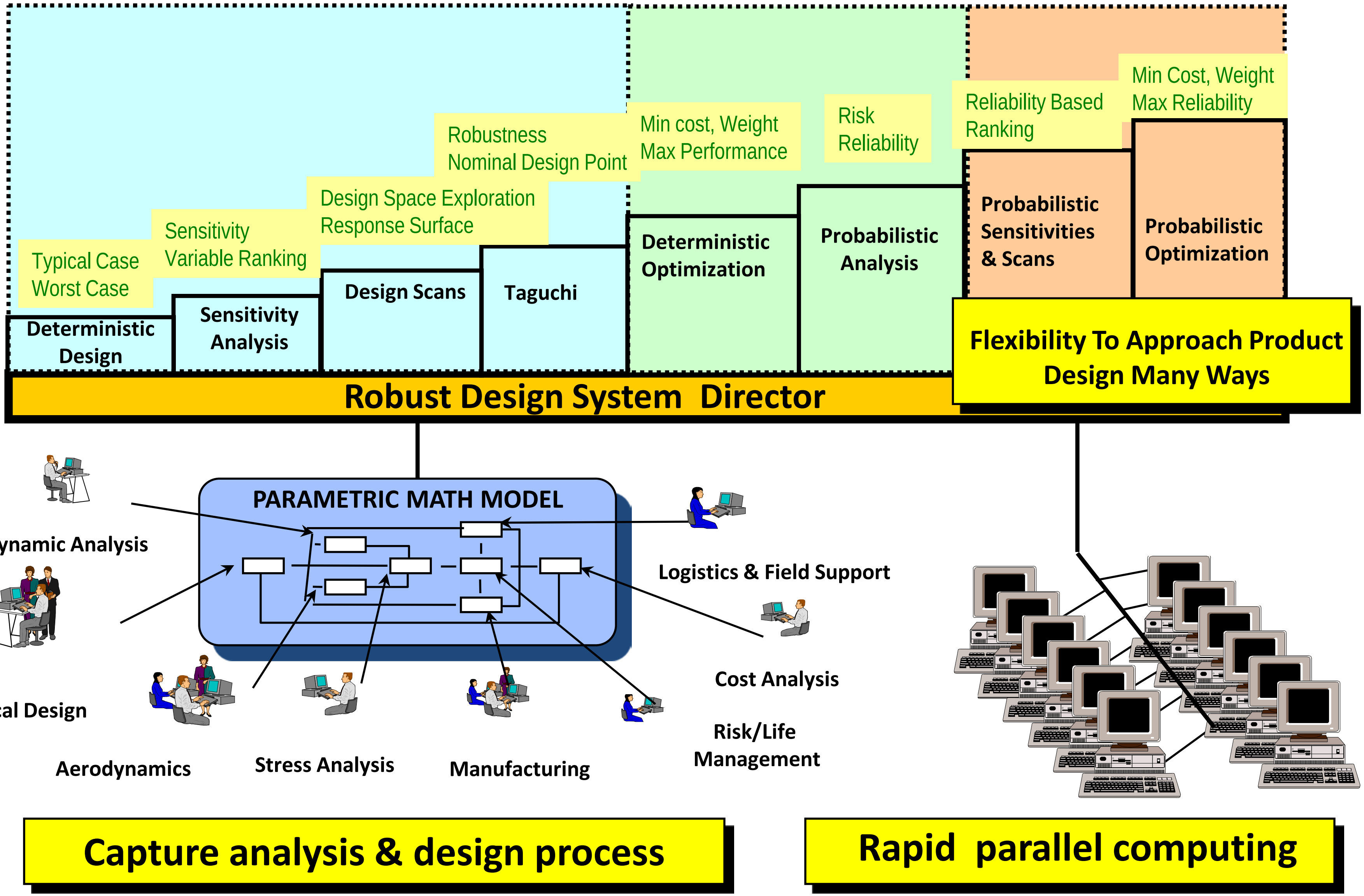
Back-Up Slides



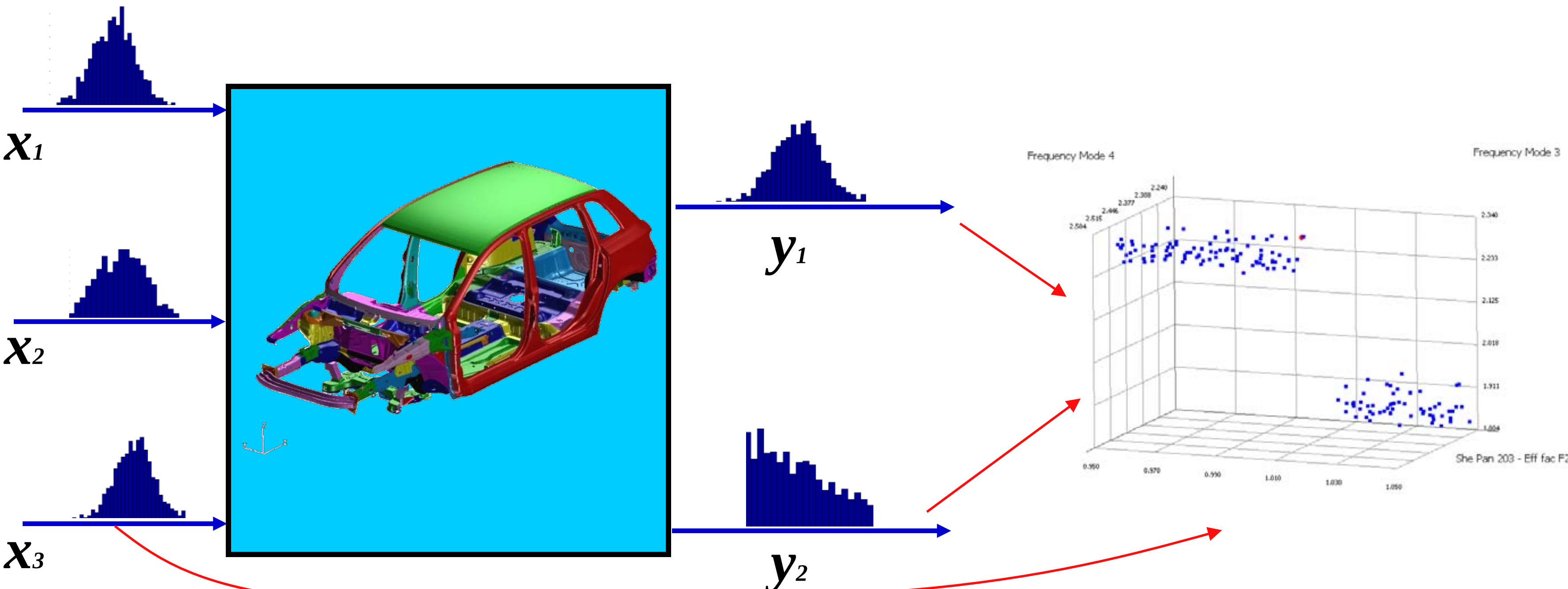
CG 53 FEM from ABS

- 118,855 elements
- 3 loading conditions:
  - Still Water, Sagging and Hogging in 25 ft waves
- 243 MB file takes 33 minutes to run on laptop

MSC Robust Design Overview



MONTE CARLO METHOD



Sources of Variability

- Material Properties
- Loads
- Boundary and initial conditions
- Geometry imperfections
- Assembly imperfections
- Solver
- Computer (round-off, truncation, etc.)
- Engineer (choice of element type, algorithm, mesh band-width, etc.)

Solution:

Establish tolerances for the input and design variables.

Measure the system's response in statistical terms.

Monte Carlo Simulation Background

- **Allows engineers to introduce hundreds of thousands of stochastic variables into the problem, and still call the solver only 100 times to obtain correct results.**

- **The power lies in the fact that the cost, i.e. the number of solver calls, is independent of the number of variables in a problem.**

Why Stochastic Analysis

Address the Curse of Dimension

“Monte Carlo simulation was developed by the Los Alamos team (the people who developed the nuclear bomb for the US during the 1940's). They had high-dimensional integrals to solve, and traditional methods of numerical integration failed them because of the curse of dimensionality.

What is so fantastic about Monte Carlo simulation is the fact that its precision is proportional to the square root of the number of scenarios used, and THIS RESULT IS ENTIRELY INDEPENDENT OF THE NUMBER OF DIMENSIONS OF THE PROBLEM. Effectively, Monte Carlo simulation was developed to break the curse of dimensionality. The history of World War II might have been different if it were never invented.”

Glyn Holton

Monte Carlo Simulation Results

Number of 2D Views of Results = Sum of all integers from 1 to (Number of Variables -1)

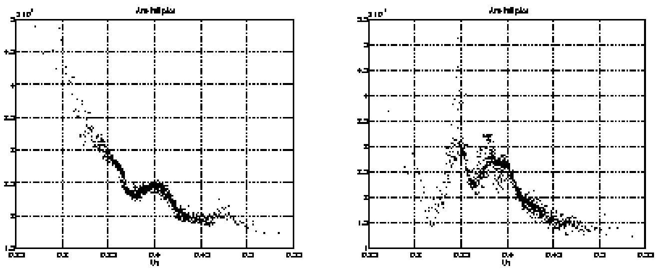


Figure 1.16: σ_{Tx} (left) and σ_{Ty} (right) versus thickness.

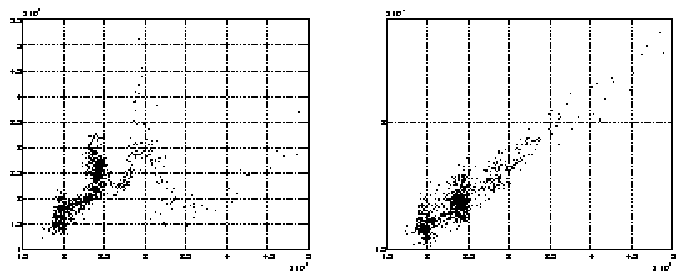
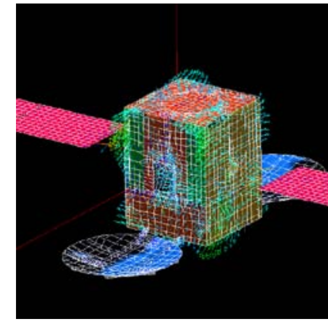


Figure 1.20: σ_{Tx} versus σ_{Ty} (left) and σ_{Tx} versus σ_{Ry} (right).



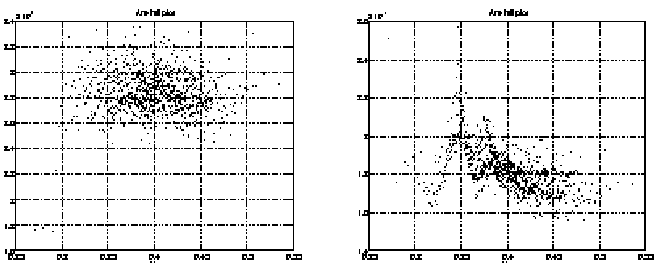


Figure 1.17: σ_{Tz} (left) and σ_{Rx} (right) versus thickness.

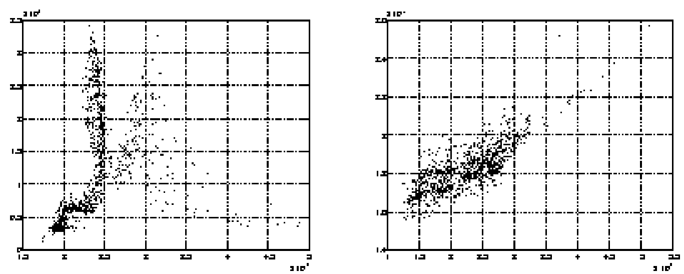


Figure 1.21: σ_{Tx} versus σ_{Rz} (left) and σ_{Ty} versus σ_{Rx} (right).

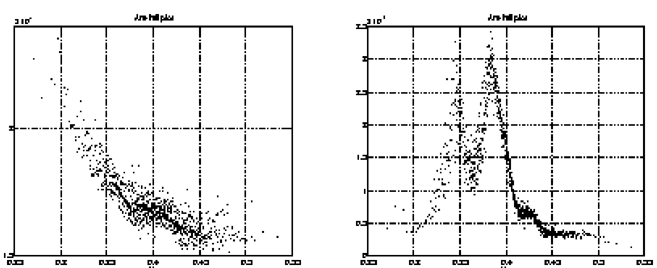


Figure 1.18: σ_{Ry} (left) and σ_{Rz} (right) versus thickness.

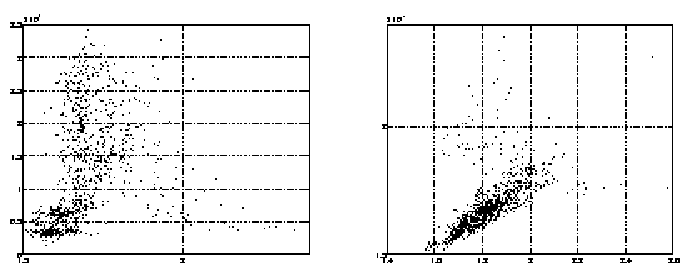


Figure 1.22: σ_{Ry} versus σ_{Rz} (left) and σ_{Rx} versus σ_{Ry} (right).

12 of the 78 2D views that resulted from a simulation with 6 outputs from a scan of 7 inputs with uniform distributions.

SIMULATION-SUPPORTED ENGINEERING

Key Validation Events

MSC Robust Design, 2002- 2005

- MSC CTO Dr. Ed Stanton focuses on probabilistic analysis following:
  - RDCS experience at MSC and Boeing
  - Attending 2001 Stochastic Simulation Conference in Germany
  - Crystal Ball use
- MSC hired Dr. Jacek Marczyk
- 2003 Innovation of the Year Award - German CAD-CAM Magazine
- Positive Feedback from Customers and Application Engineers
- Limited to Nastran applications
- Needed ability to change geometry, run with other solvers

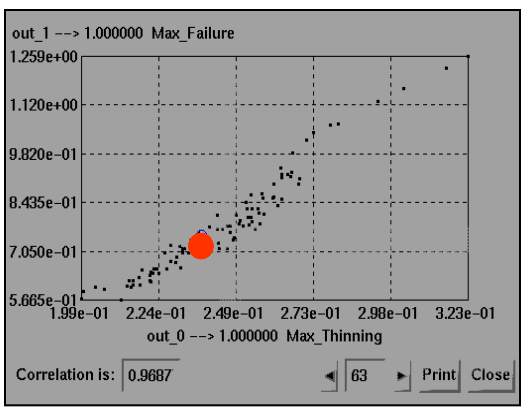
Process Validation

- Dr. Schueller - MIT 2005 keynote presentation
- Dr. Hazelrigg - NSF meeting
- MIT mathematician K. Keilmeyer –

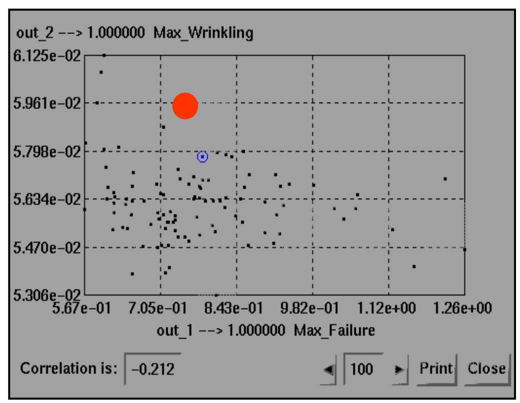
  "Monte Carlo Simulation is not elegant, it just gives the right answers."

Why Stochastic Analysis

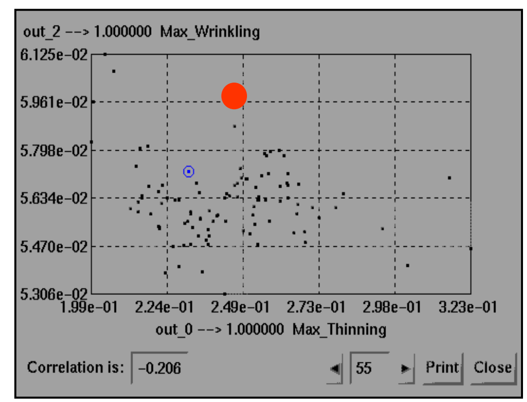
The most likely behavior (practically) never corresponds to the most likely values of input/design variables.



Max. Failure vs. Max. Thinning



Max. Failure vs. Max. Wrinkling



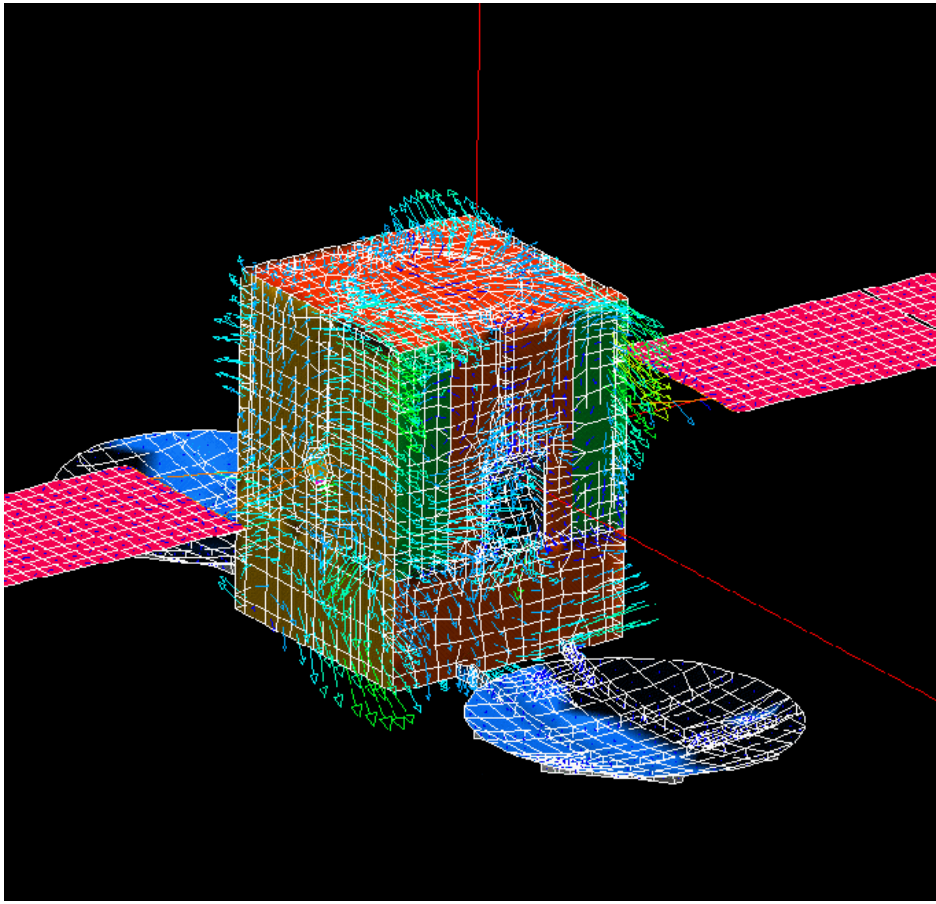
Max. Thinning vs. Max. Wrinkling

Result of Nominal Run

START WITH A GOOD MODEL

Accurately Capture:

- **Physics**
- **Loads and boundary conditions**
- **Material properties**
- **Geometry**



Remember **"All models are wrong. Some are useful."** J. Box

Modeling Processes:

- **In Analysis Software**
  - **NASTRAN, Ansys, STAR-CD, Dyna, etc.**
- **In CAD Software translated to Analysis Model**
  - **Catia/UG/Pro-E/… to Hypermesh/Patran/…**
- **Multi-Disciplinary Process Model Mapping**
  - **iSight/ModFrontier/ModelCenter/….**

Results Processing/Visualization

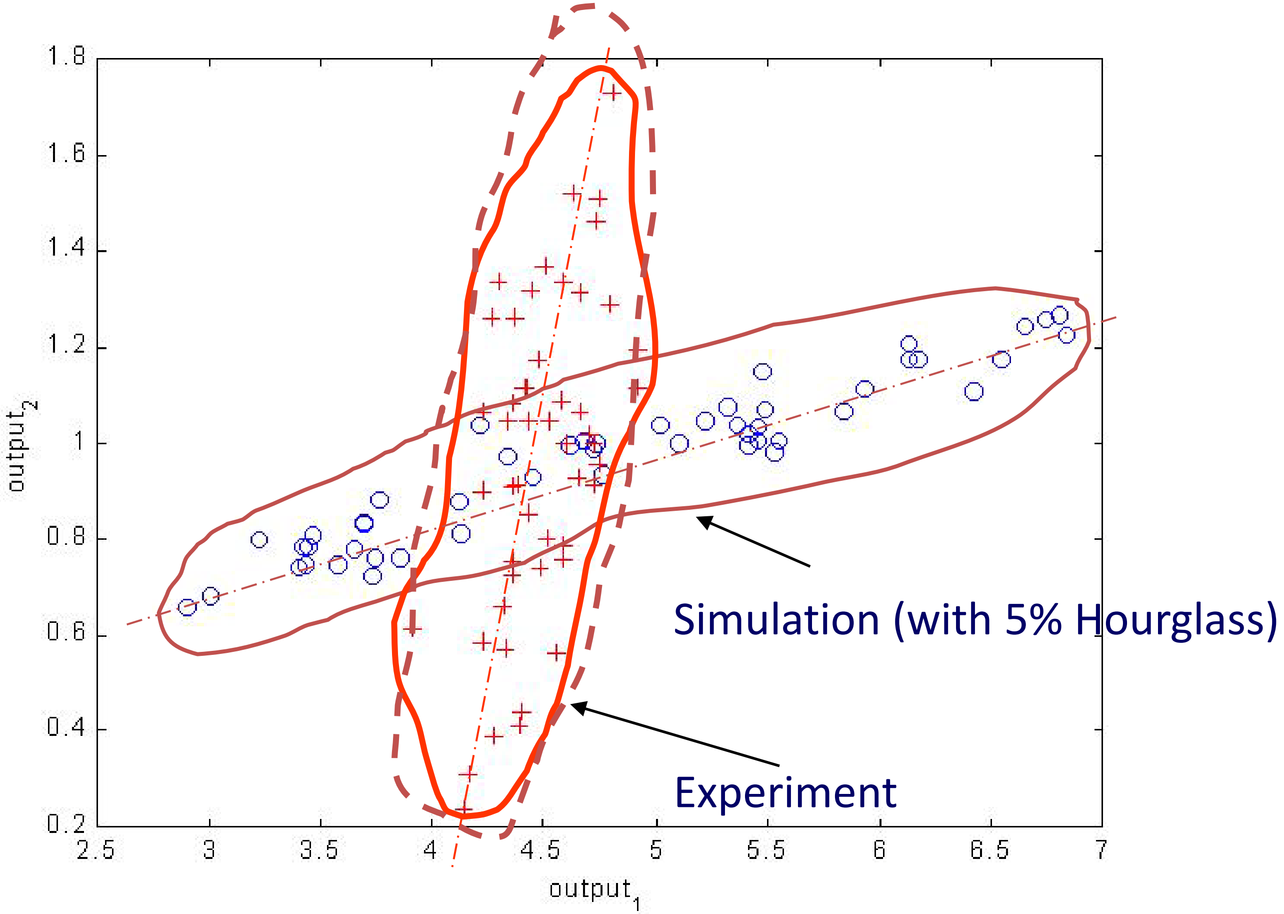
Remember **"All models are wrong. Some are useful."** J. Box

Model Verification

- Make sure the math is correct
- Model checking with probabilistic analyses
  - If solver runs 100 times the model is likely valid
- Good CAE models exist
  - Primarily done for forensic analysis
  - Represents one instance of reality
  - Stochastic analysis can provide information

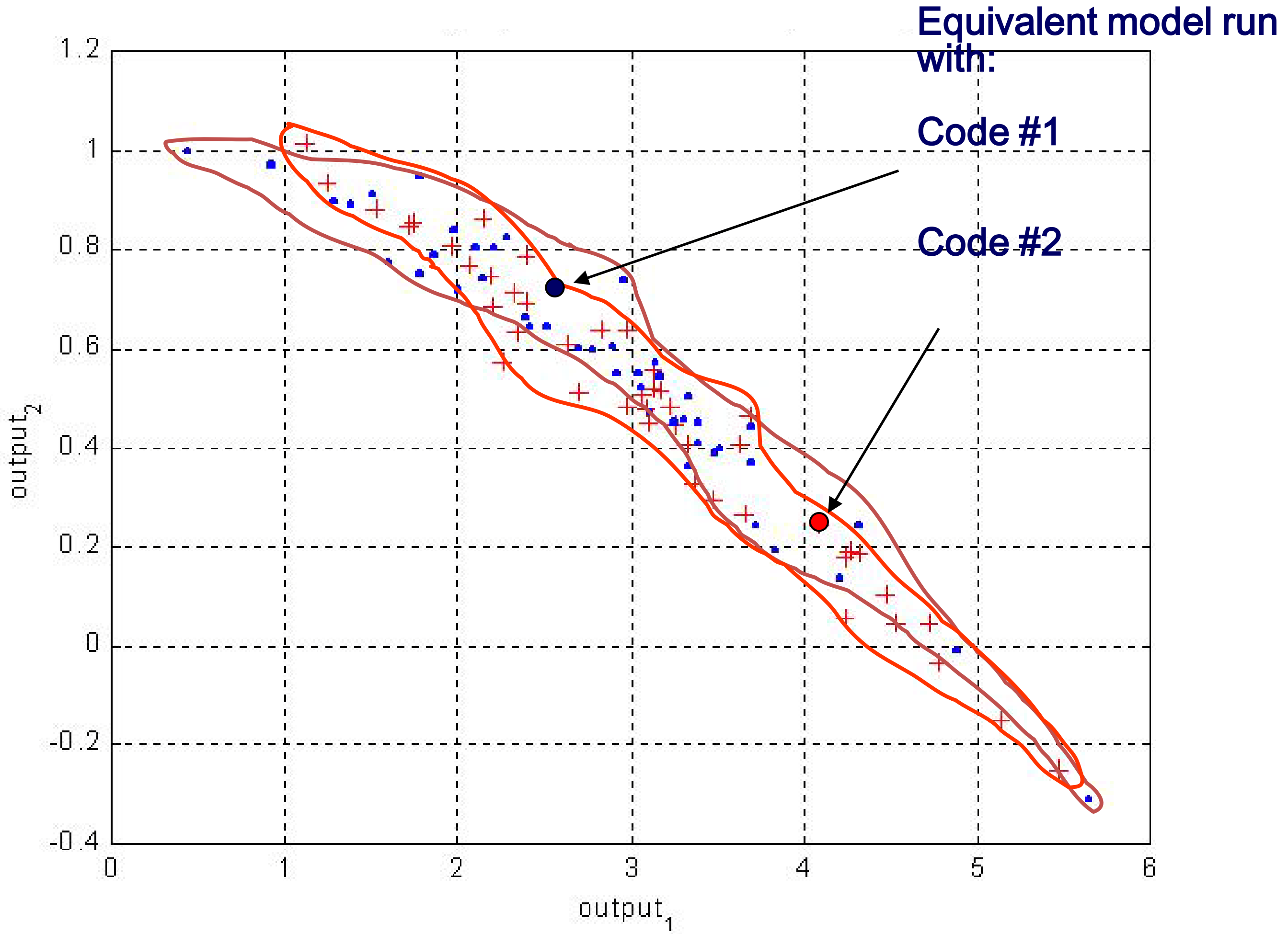
Remember **"All models are wrong. Some are useful."** J. Box

Quantifying model confidence



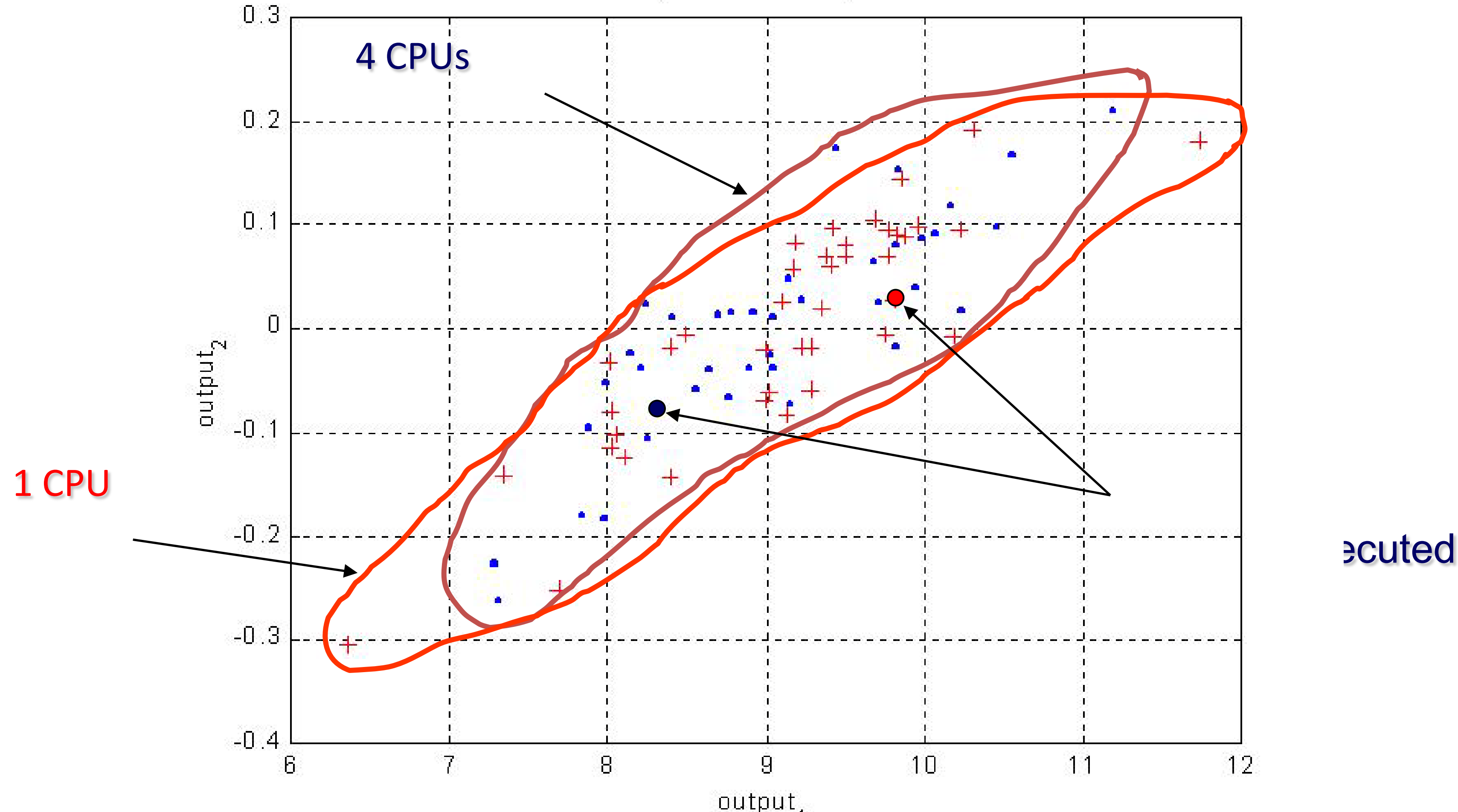
The two meta-models show excellent statistical equivalence (shape, aspect ratio, cog position. The difference in orientation has been found to be due to 5% HG energy in the FE model.

Which crash code?



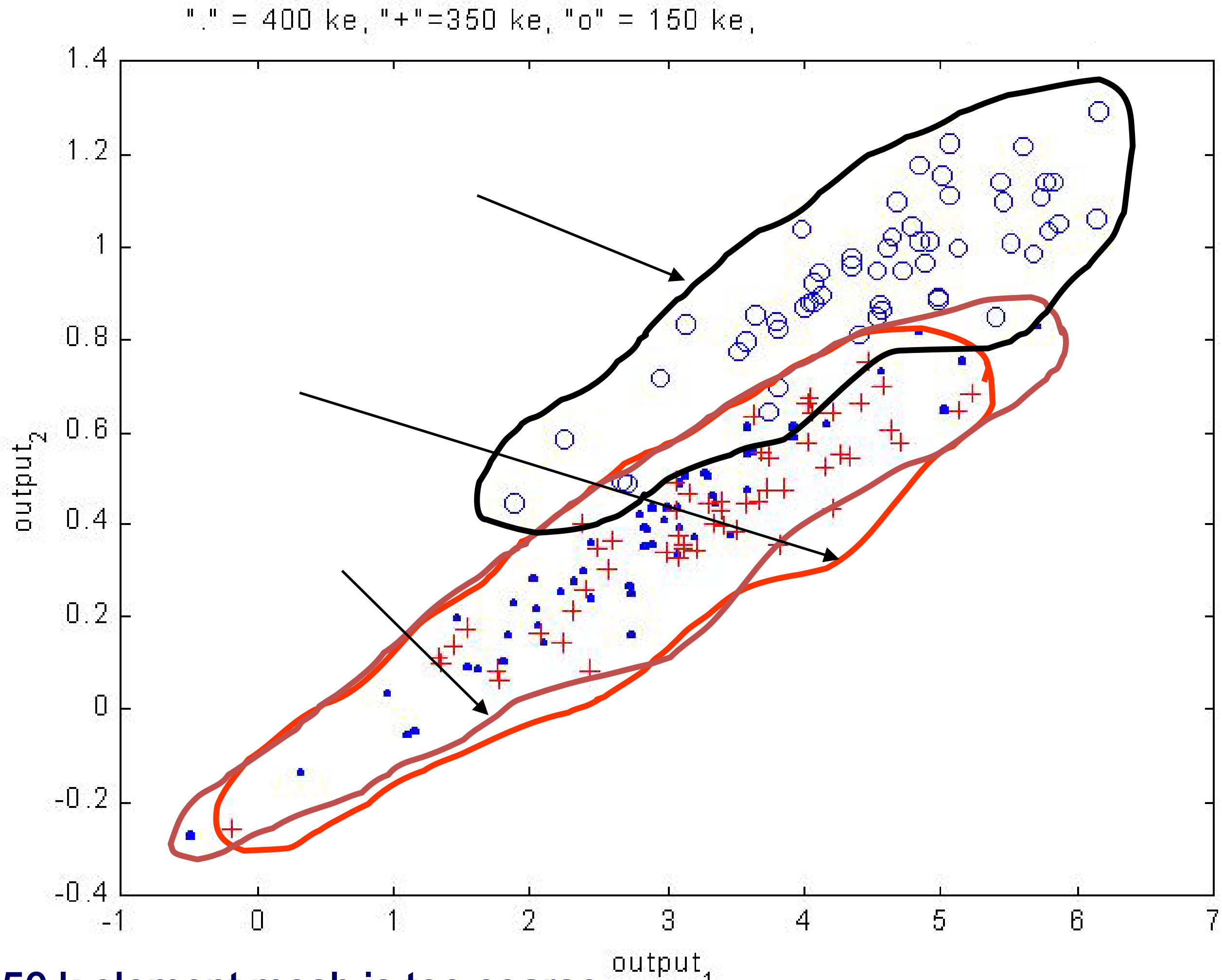
From a statistical point of view, the two codes are equivalent.

Must Analyses be Repeatable?



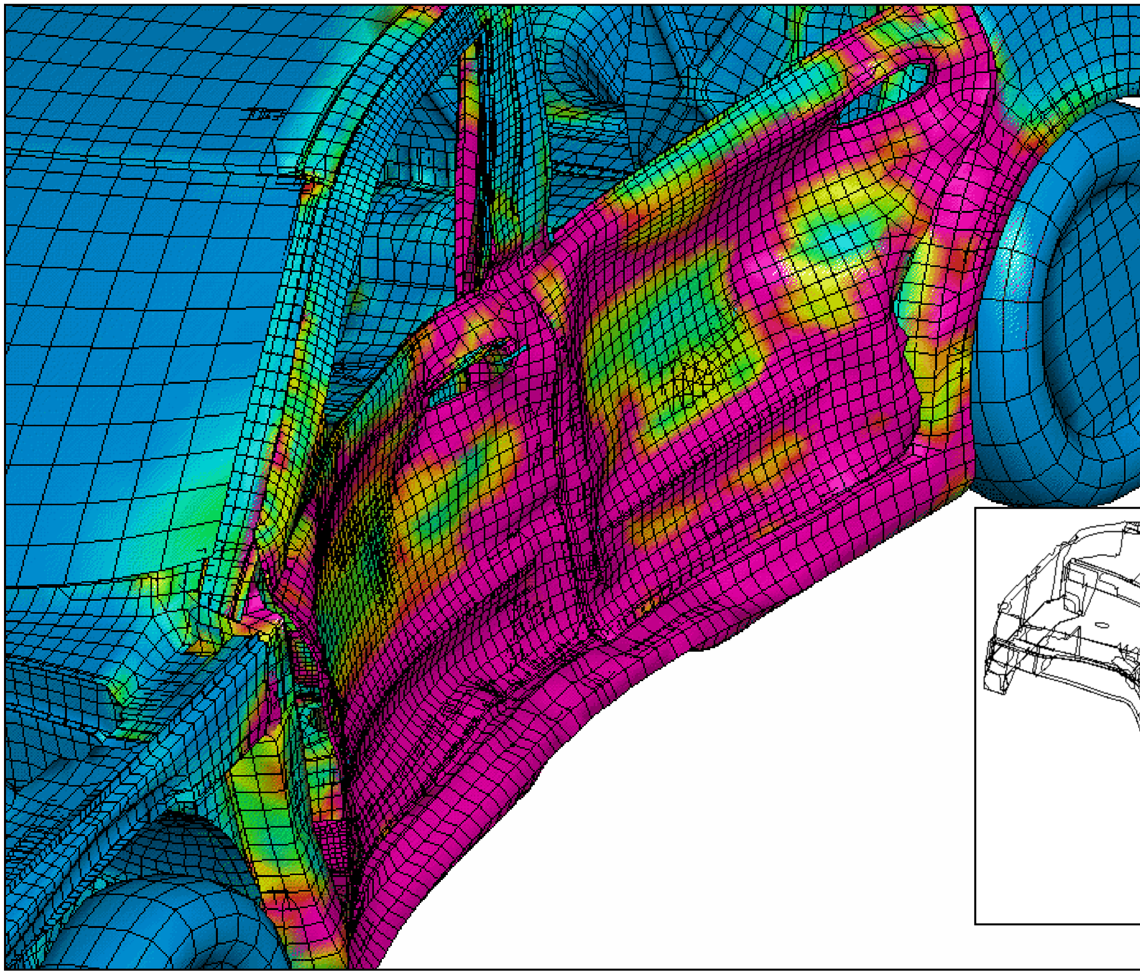
Given that from a statistical standpoint the meta-models are equivalent, the number of processors is irrelevant.

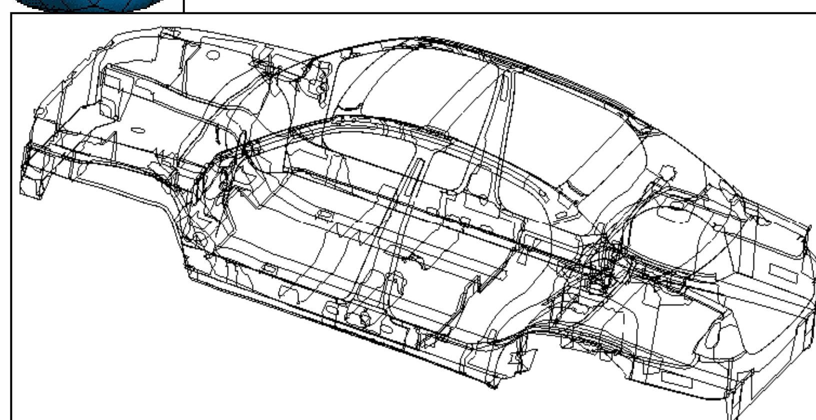
Models and number of elements



- 150 k element mesh is too coarse
- Other two meshes are essentially equivalent (i.e. no need for more than 350 k)
- Meshes of more than 350 k elements are too detailed

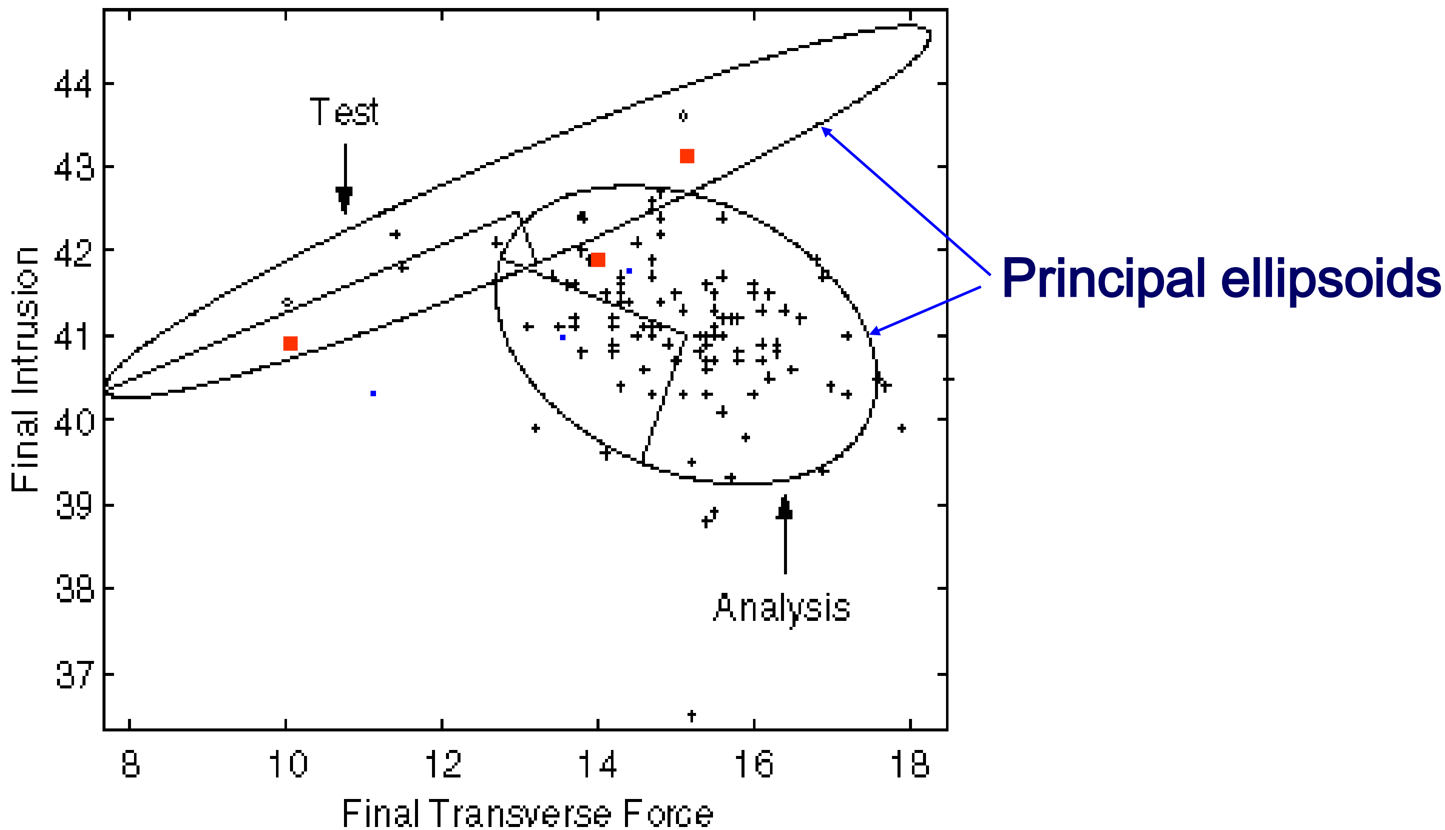
Model Validation - B-Pillar example





Courtesy of Italdesign

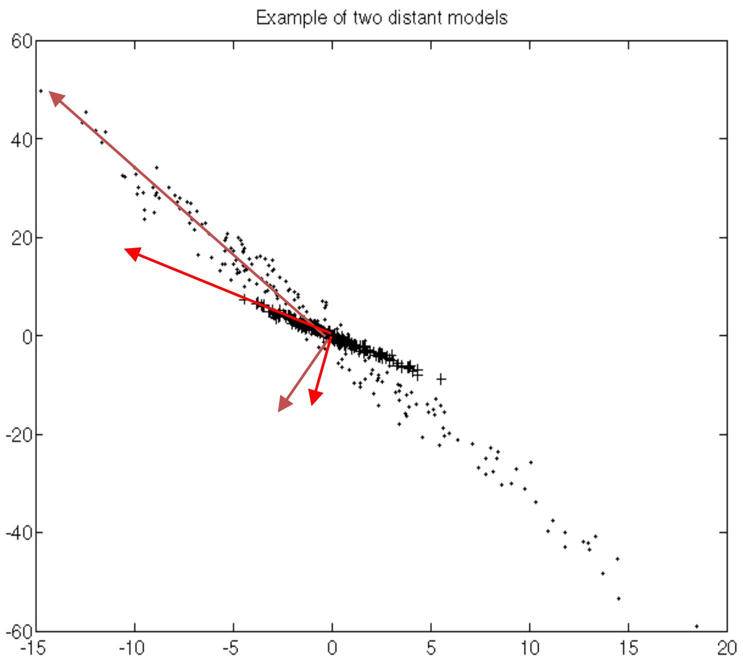
B-Pillar Model Validation



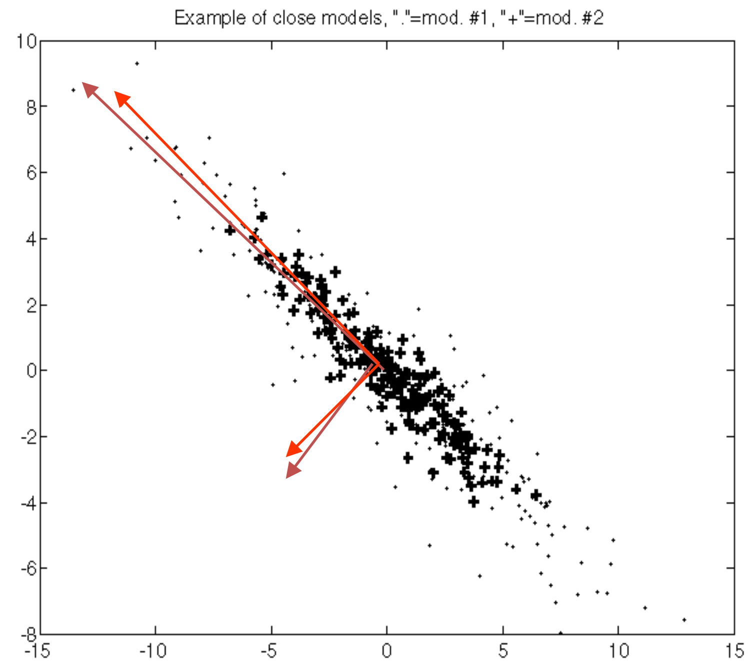
- The model, in its current configuration, is not valid.
- Three tests are insufficient to make any statements as to the model's validity.
- A single experiment cannot validate a model, it can merely verify it (or falsify it!).
- Validation requires multiple experiments and multiple calculations.

What Does Validation Really Mean

It is matching the covariance matrices of test and simulation meta-models



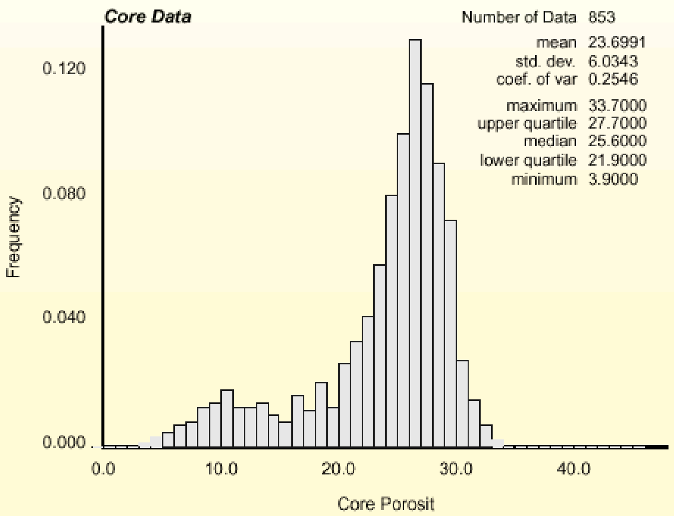
Means OK but differences in covariance and scatter

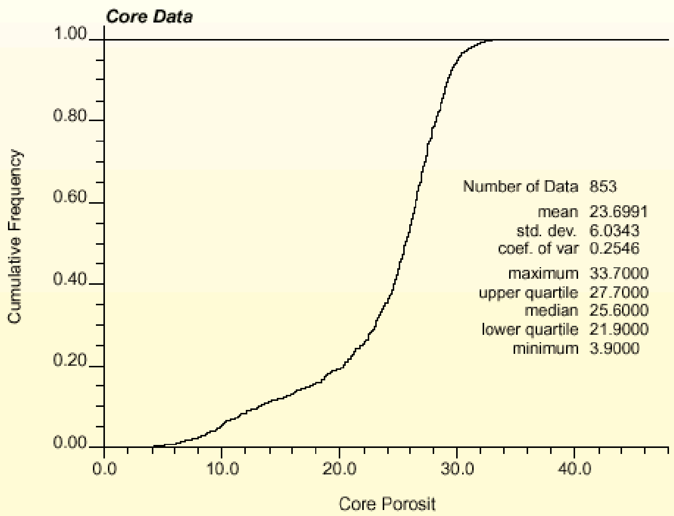


Means and covariance OK but not scatter (eigenvalues)

Monte Carlo Simulation

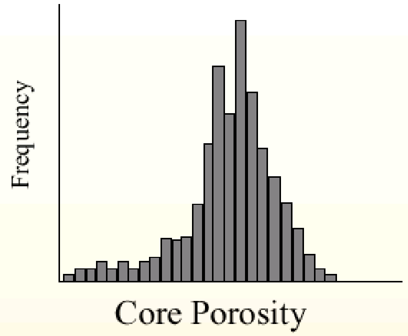
- Monte Carlo Simulation proceeds by reading quantiles from a cumulative distribution
- The procedure:
  - generate a random number between 0 and 1
  - read the quantile associated to that random number

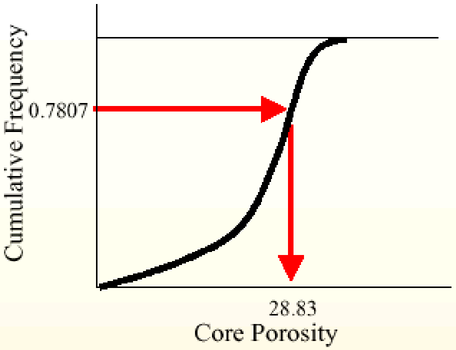




Monte Carlo Simulation

Example: simulate core porosity from its cdf.





- Monte Carlo Simulation / Stochastic Simulation / Random Drawing proceed by reading quantiles from a cumulative distribution

The procedure:

- generate a random number between 0 and 1 (calculator, table, program, ...
- read the quantile associated to that random number

For Example:

Random Number	Simulated Number
0.7807	28.83
0.1562	
0.6587	
0.8934	

...

Cumulative Distribution Function

- Cumulative distribution function cdf defined as:

$$F(z) = \text{Prob}\{Z \leq z\} \in [0,1]$$

This formula gives the area under the pdf of the RV Z, and is the probability that the RV Z is less than or equal to a threshold value of z.

- The probability of exceeding any of the threshold values of z can be written:

$$\text{Prob}\{Z > z\} = 1 - F(z)$$

- Properties of the cdf:
  - F(z) non decreasing
  - $F(z) \in [0,1]$
  - $F(-\infty) = 0$ and $F(\infty) = 1$

Cumulative Distribution Function (CDF) and Probability Distribution Function (PDF)

- The probability of Z occurring in an interval from a to b (where $b>a$) is the difference in the cdf values evaluated at points b and a:

$$\text{Prob}\{Z \in [a,b]\} = F(b) - F(a)$$

- Probability density function (pdf) is the derivative of the cdf, if it is differentiable:

$$f(z) = F'(z) = \lim_{dz \to 0} \frac{F(z+dz) - F(z)}{dz}$$

- The cdf can be obtained from integrating the pdf:

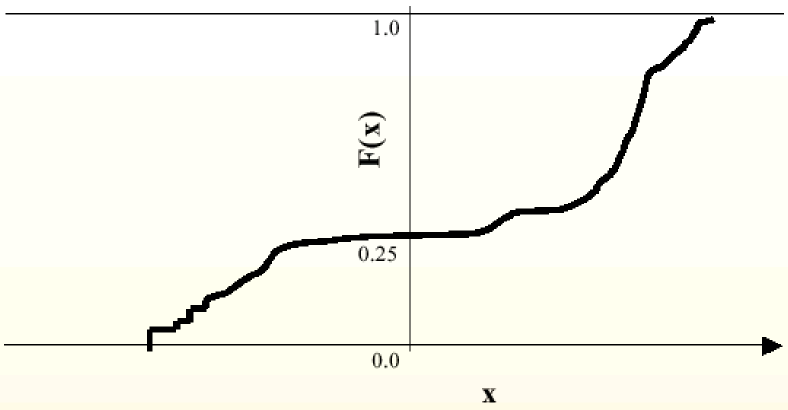
$$F(z) = \int_{-\infty}^{z} f(z)\,dz$$

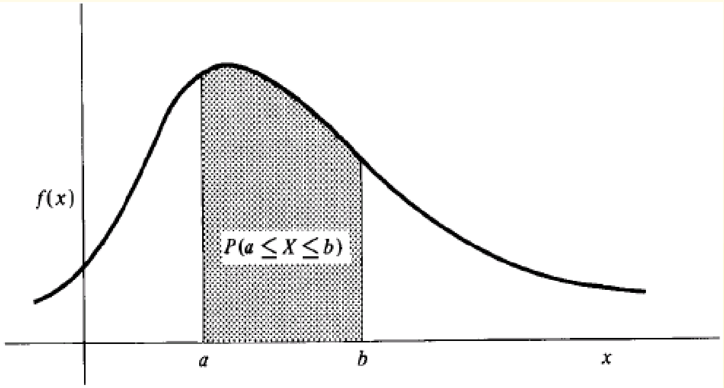
CDF and PDF

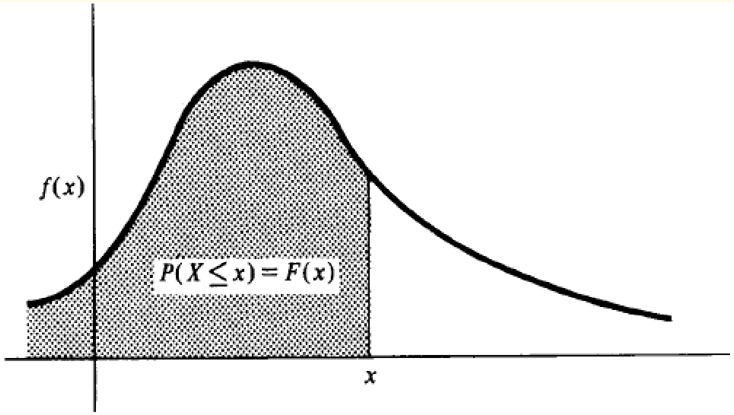
- Properties of pdf:

$$f(z) \geq 0$$

$$\int_{-\infty}^{\infty} f(z)dz = 1$$







Where Does the PDF come from?

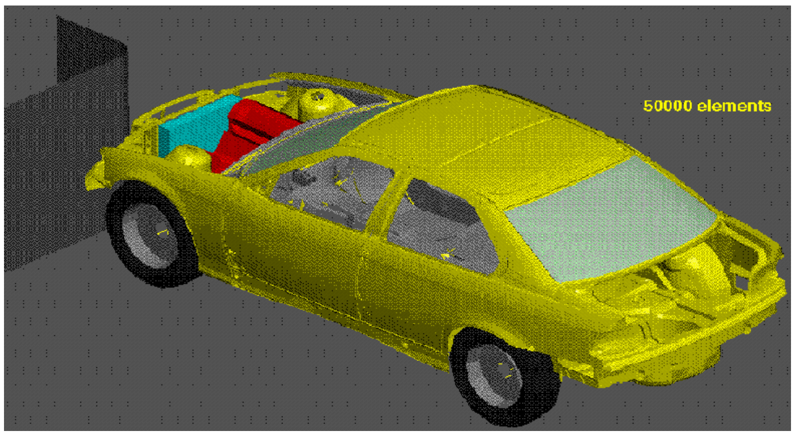
GOOD ENGINEERING

- We know from experience that the world is not discreet
- Replace all discreet values with ranges and a distribution

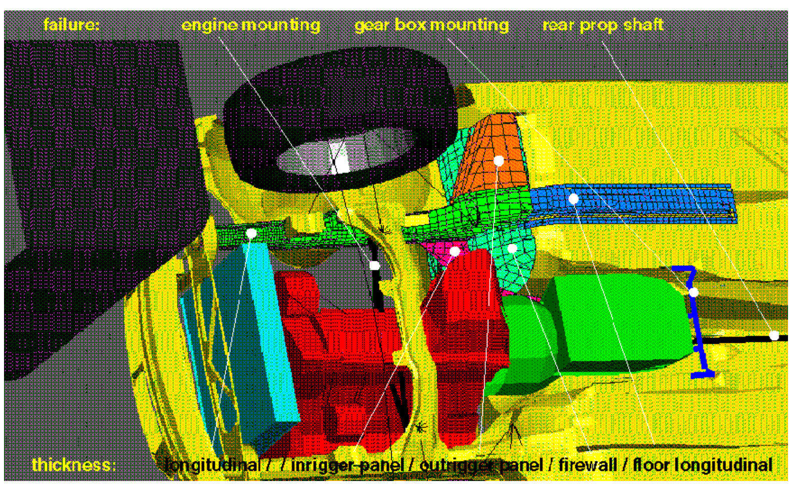
What is the real density of water?

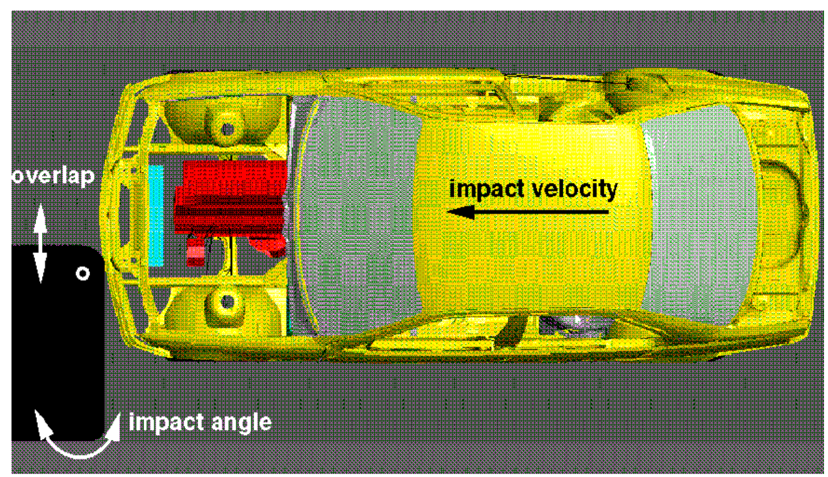
First World-wide Stochastic Crash

(BMW-CASA, August 1997)



- Stochastic material properties, thicknesses and stiffnesses (70 variables),initial and boundary conditions (angle, velocity and offset).
- 128 Monte Carlo samples on Cray T3E/512 (Stuttgart Univ.)
- 1 week-end of execution time.





Method 1 – ESA-LFU Load-Coupled Dynamic Analysis

Element/Material Type	***Property***	***C.o.V. (σ/μ)***
Isotropic Material	Young's modulus Poisson's Ratio Shear modulus Mass density	8% 10% 12% 4%
Orthotropic Shell Element Material	Young's Modulus Shear Modulus	8% 12%
Solid Element An-Isotropic Material	Material Property Matrix Mass Density	12% 4%
Simple Beam	Section dimension	5%
Layered Composite Material	Non-Structural Mass Thickness of Plies Orientation Fibre Angle	8% 12% 1.5°
Spring Element	Stiffness	8%
Shell Element	Membrane thickness Non-Structural Mass	4%
Concentrated Mass	Mass	3%
Damping	Modal Damping Structural Damping	40% 25%
	# of Random Variables	*≈ 1300*

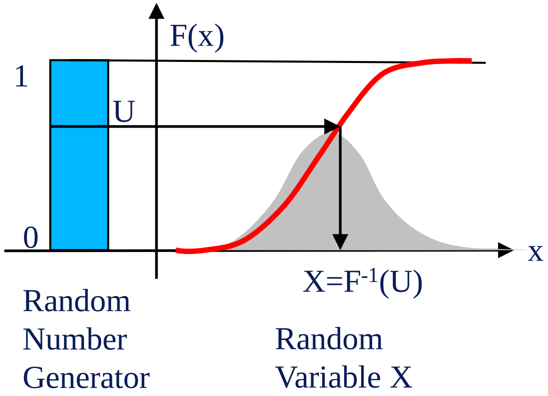
Gaussian

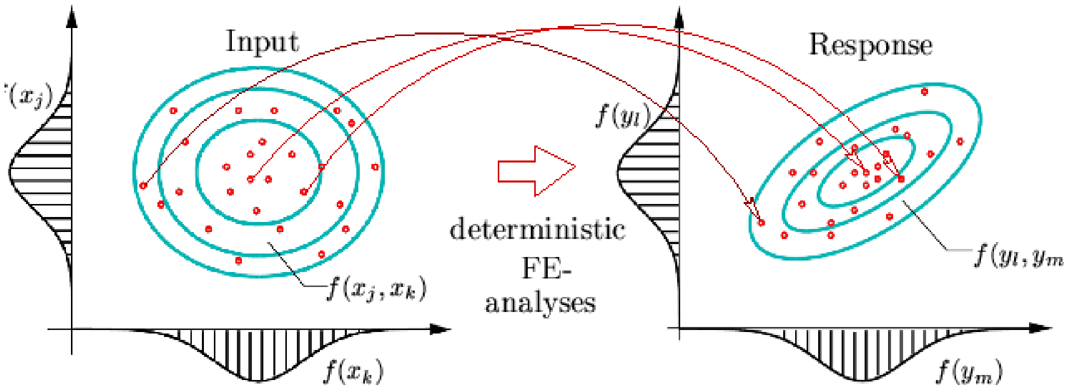
Log-Normal

Types of uncertainties: **Material Properties, Geometrical Properties**

Efficient Monte Carlo Simulation

MCS is the most general and versatile method to process uncertainties

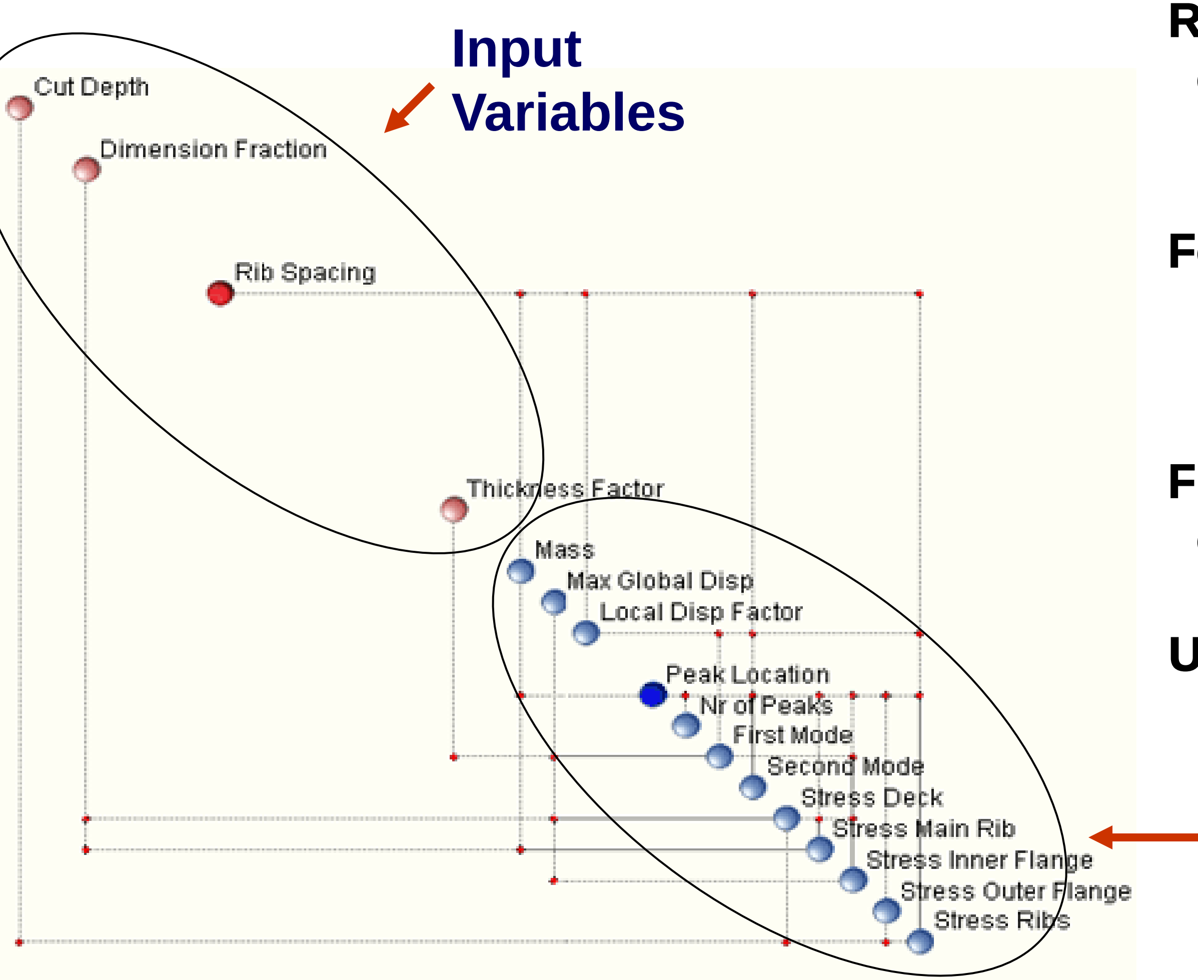




As few as 25-50 simulations provide information on the robustness and variability of the response due to uncertanties of the input parameters

Prof. G.I. Schueller, Third MIT Conference on Computational Fluid and Solid Mechanics, June 2005

Correlation Maps –
Filter Complexity while Modeling Reality



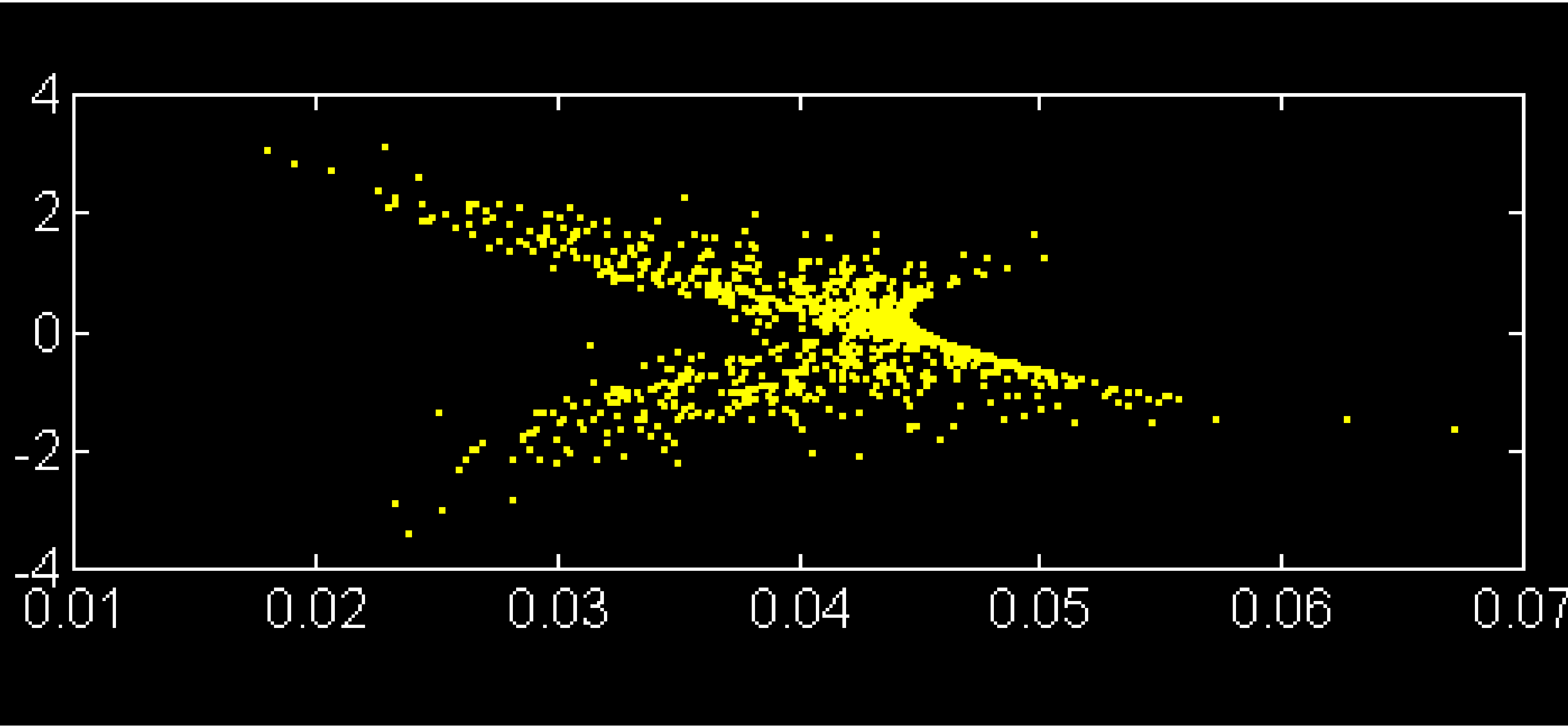
Ranks input variables and output responses by correlation level

Follows MIT-developed Design Structure Matrix model format

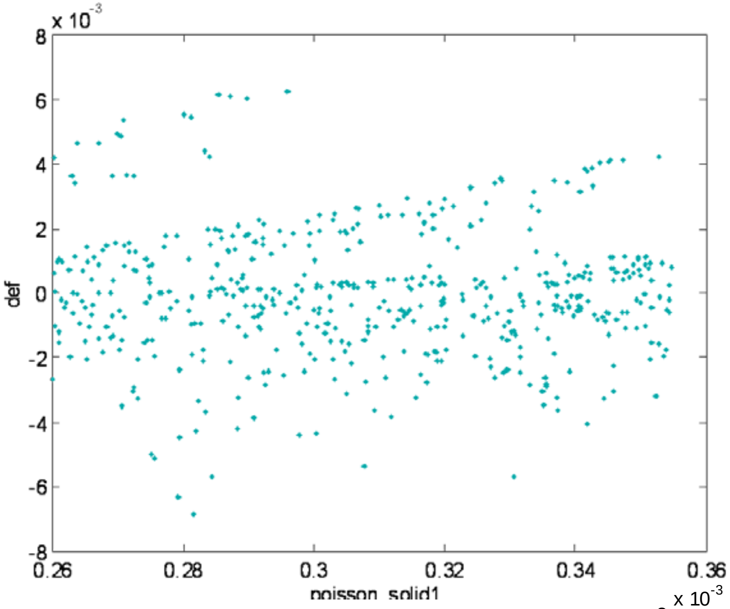
Filters Variables Based on Correlation Level

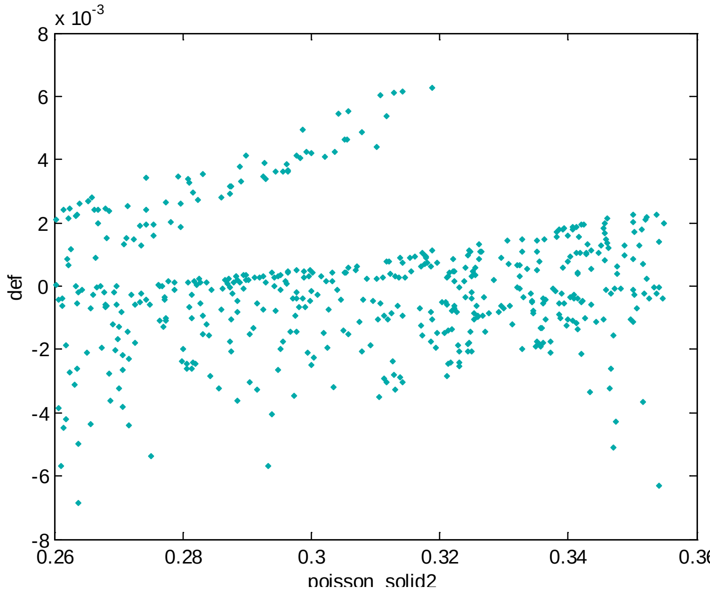
Understand How Things Work

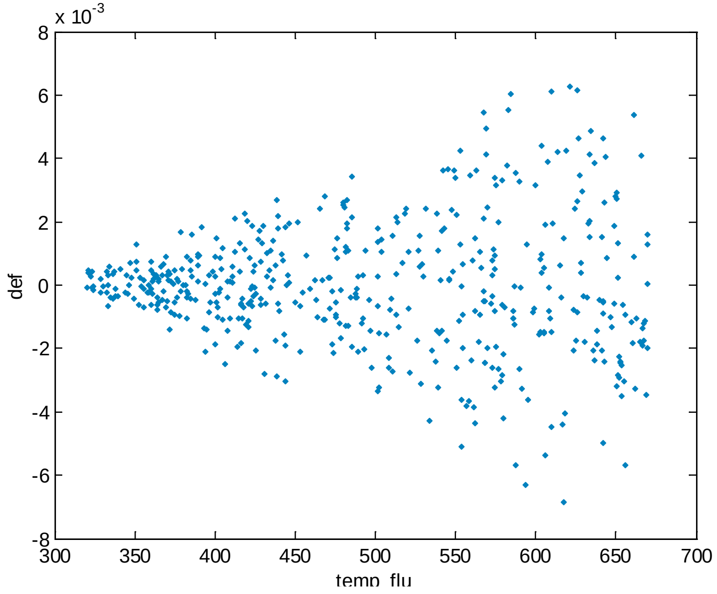
Design Exploration with Fitness Landscapes



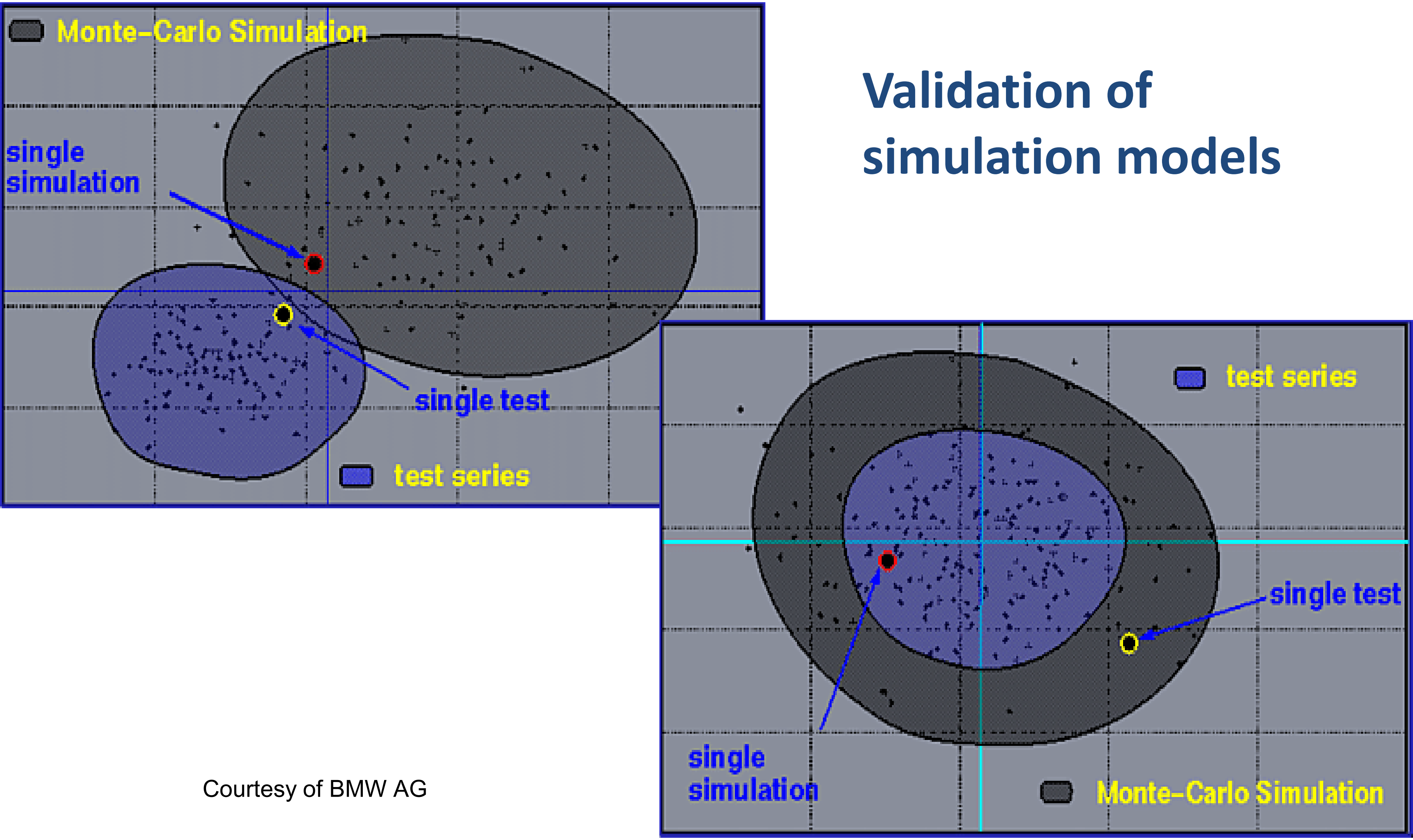
Design Exploration with Fitness Landscapes







Simulation Design Improvement



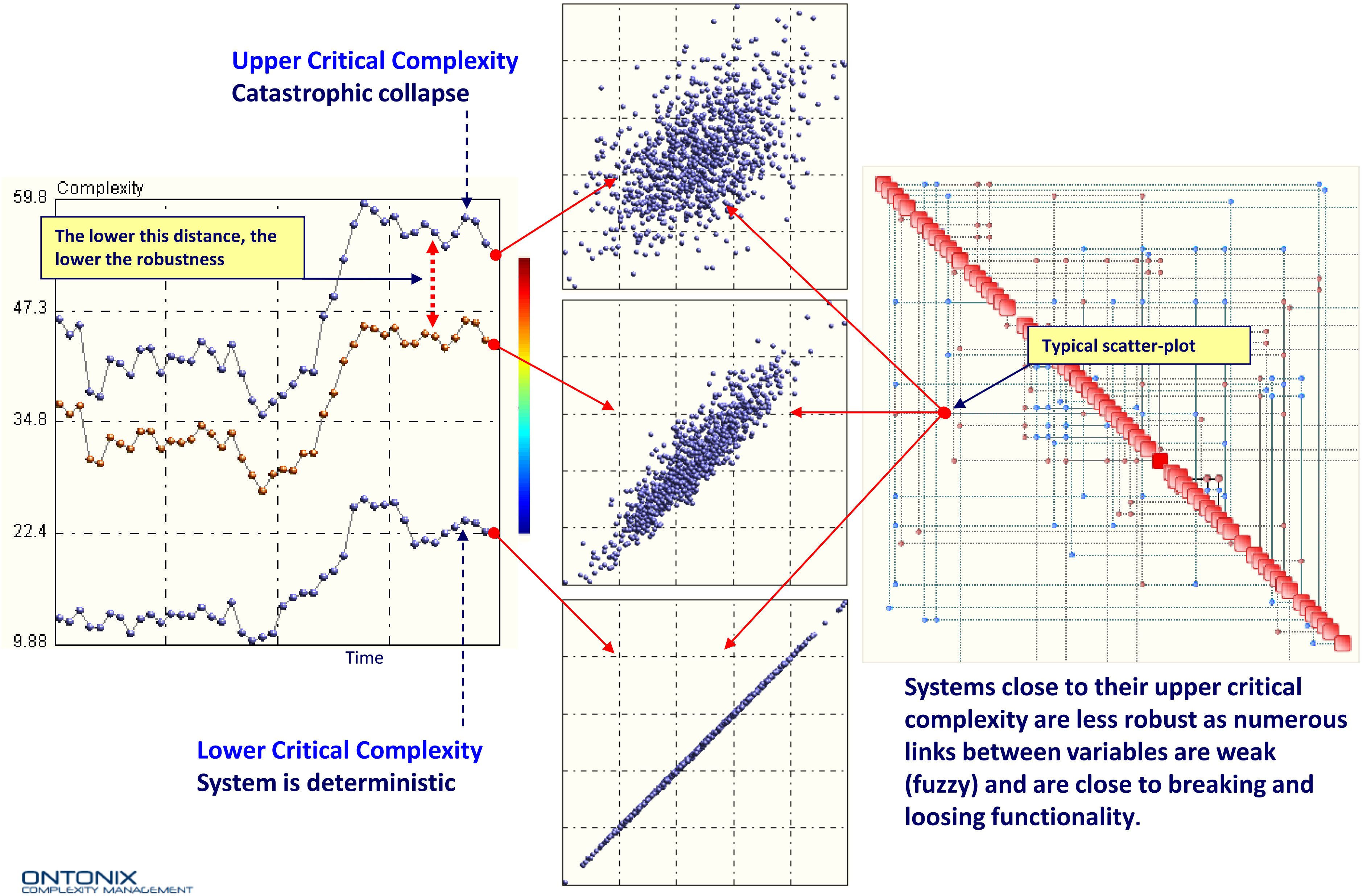
Validation of simulation models

Courtesy of BMW AG

What is Complexity?

- Complexity is an attribute which characterizes every system, just like energy or momentum. It can be measured, and therefore managed.
- However, complexity cannot grow indefinitely and has a maximum. Close to this maximum, called critical complexity, the system becomes fragile and vulnerable.
- Critically complex systems are very difficult to manage and can easily run out of hand.
- The risk exposure of any dynamical system can be measured and understood in an innovative way via complexity.

Complexity and Robustness

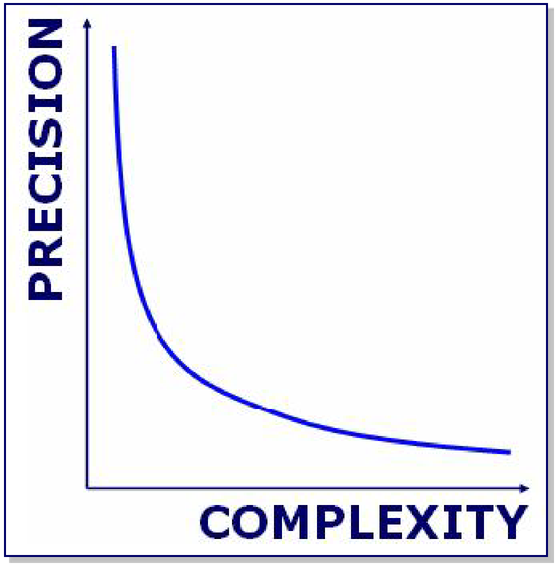


Systems close to their upper critical complexity are less robust as numerous links between variables are weak (fuzzy) and are close to breaking and loosing functionality.

ONTONIX COMPLEXITY MANAGEMENT

Complexity Principles

- **Principle of Complexity**:
  When the complexity and uncertainty of an engineering system increase, our ability to predict its behavior diminishes until a threshold is reached beyond which accuracy and significance become almost mutually exclusive.



- **Principle of Incompatibility**:
  High precision is incompatible with high complexity.

L. Zadeh, UCLA

Complexity x Uncertainty = Fragility

- When uncertainty meets high complexity, the result is fragility. Simple systems cope better with uncertainty than highly complex systems.
- Highly complex systems are more exposed to the effects of uncertainty because of the countless ways they can process information. They can fail in many ways, often due to apparently innocent causes.
- Uncertainty in the environment cannot be avoided. We must learn to live with it. Hence the need to manage complexity.
- Fragility is a prelude to risk, holistic risk management can be accomplished via complexity management.

Complexity x Uncertainty = Fragility

$$C_{design} \times \left(U_{manufacturing} + U_{environment}\right) = F_{product}$$

- A highly sophisticated design will result in a fragile product if:
  - The manufacturing process is of poor quality
  - The environment is very “turbulent”
- Hence, a more robust product requires:
  - A high-quality manufacturing process, or
  - A less severe environment in which to function, or
  - A less “ambitious” initial design

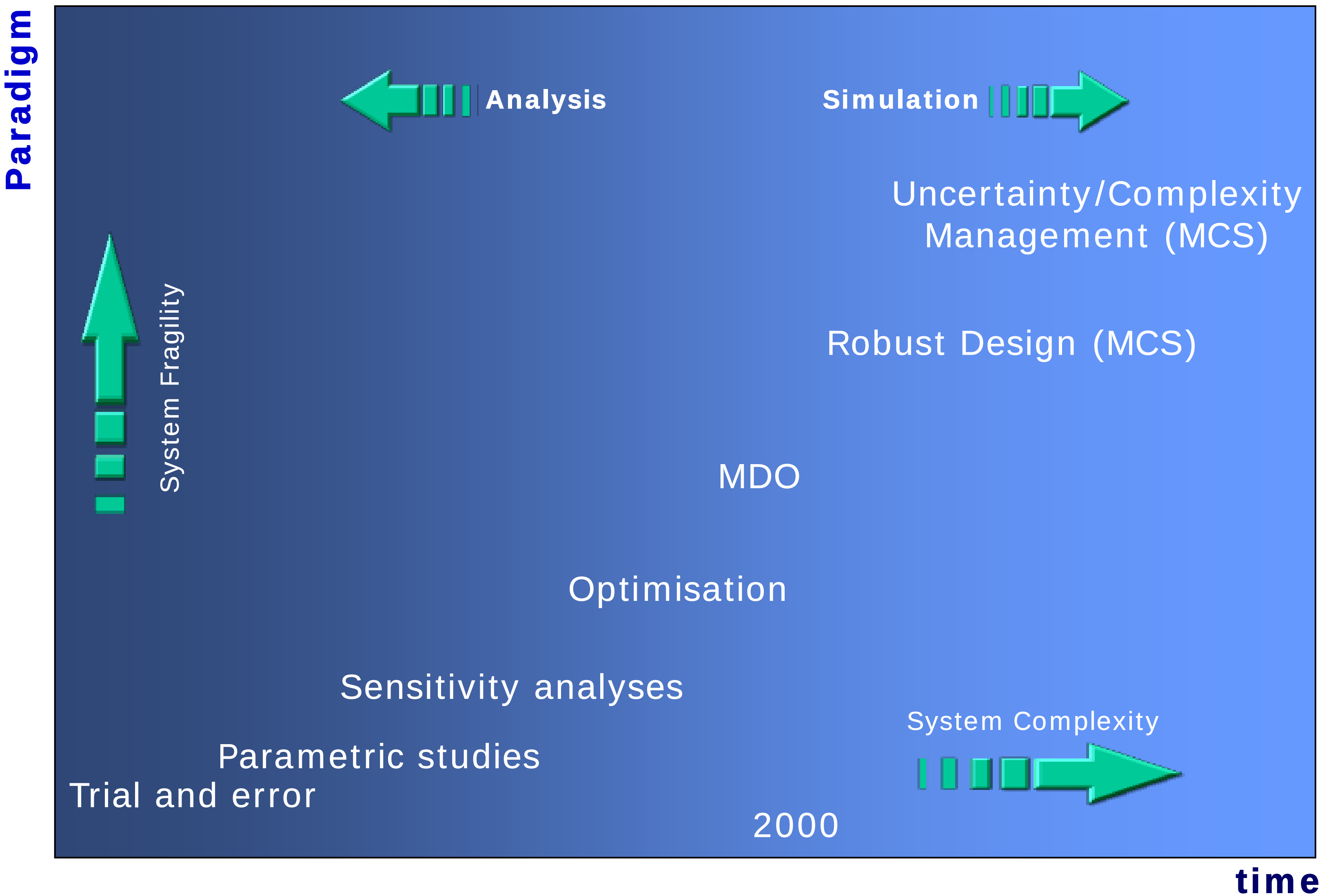
ONTONIX
COMPLEXITY MANAGEMENT

Complexity-Based Design

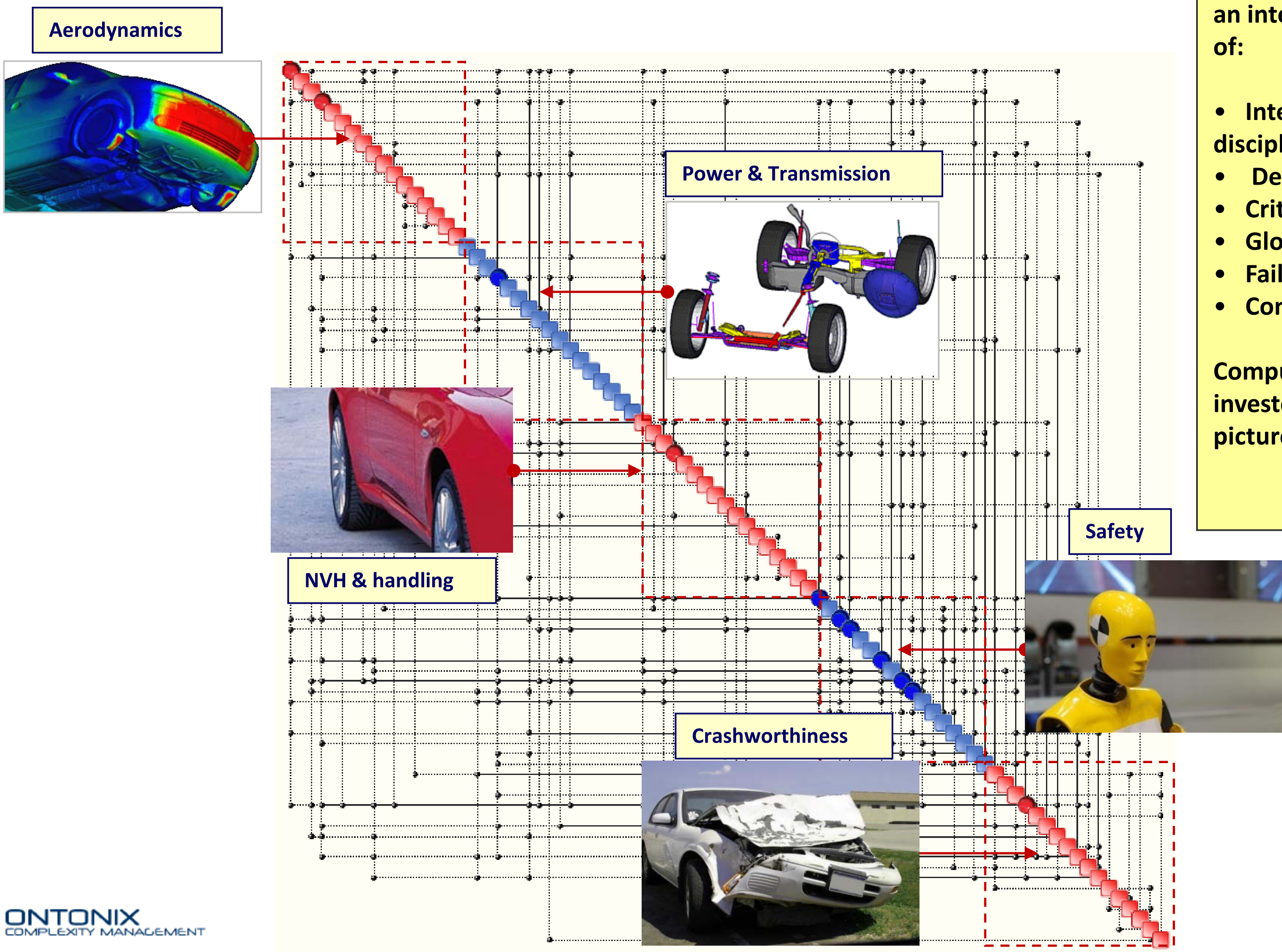
A less complex solution is generally:

- Less expensive to design and engineer
- Less expensive to manufacture
- Less expensive to service (replace broken components, etc.)
- Cheaper
- Easier to operate
- Less fragile. This means:
  - Less warranty costs
  - Less recalls
  - Less law-suits

Evolution of CAE Paradigms



Holistic CAE



Process Maps which gives users an integrated and holistic view of:

- Interaction between disciplines
- Degrees of coupling
- Critical variables
- Global robustness measures
- Failure modes
- Complexity

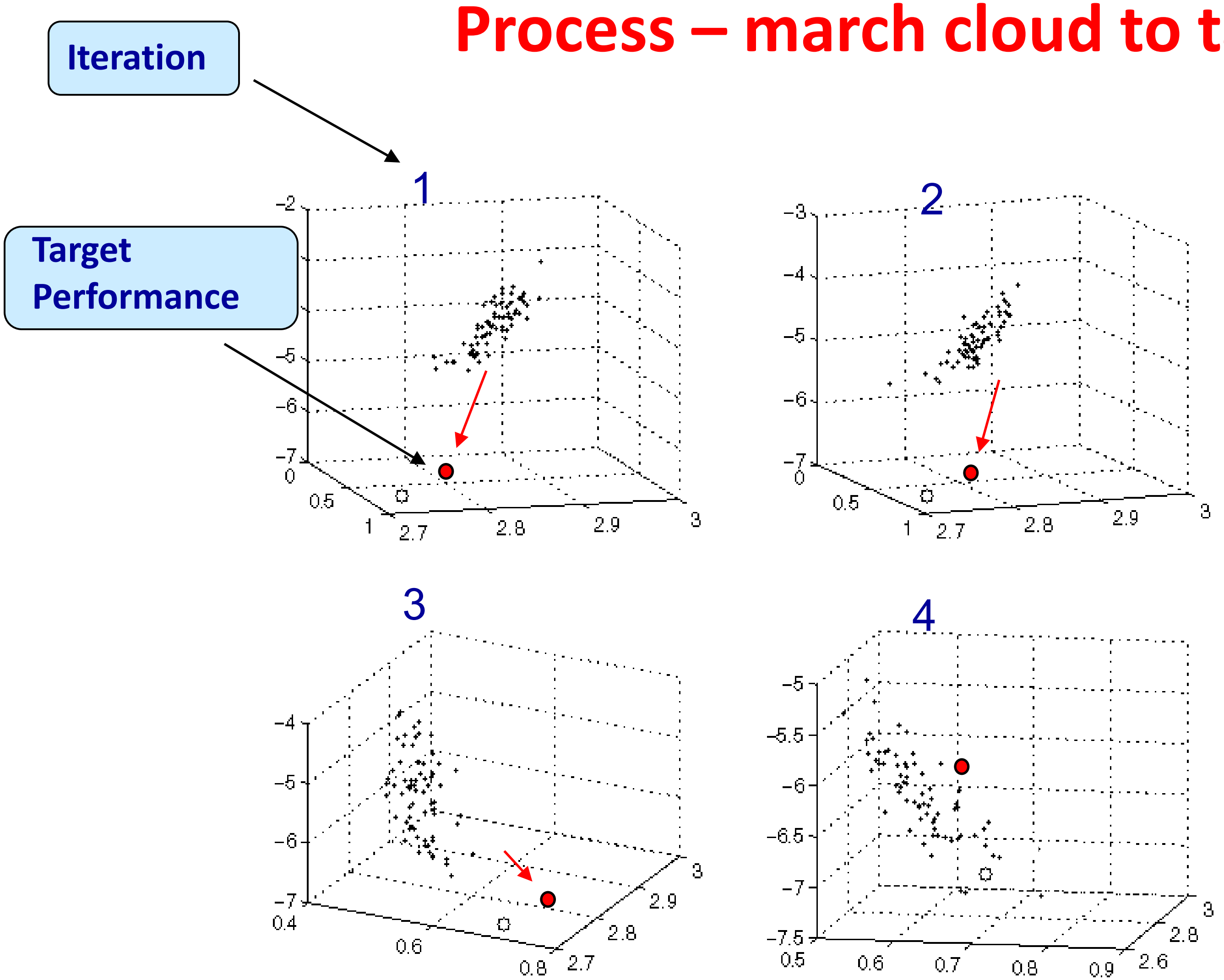
Compute power should be invested in getting the big picture and not in hair-splitting.

ONTONIX
COMPLEXITY MANAGEMENT

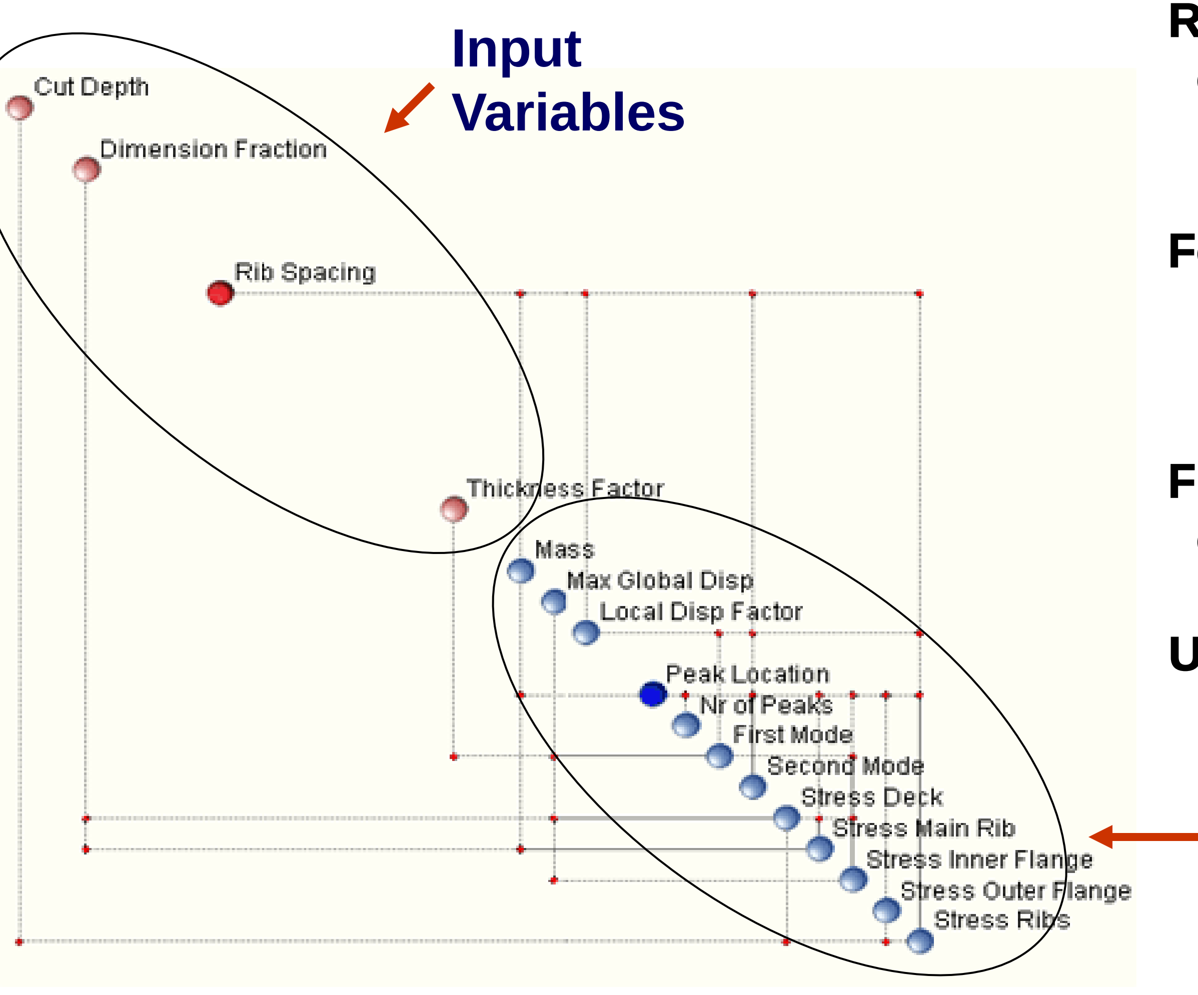
Future Trends in CAE

- **No surrogates – FE Models are already surrogates. Why work with models of models?**
- **Move from (slow & expensive) pursuit of optimal solutions to (fast & cheap) identification of acceptable compromises.**
- **Go holistic: Less local details more global patterns – Nature works like that!**
- **Measure robustness.**
- **Measure model credibility – the success of CAE hinges on it!**

Simulation Design Improvement Process – march cloud to target



Correlation Maps – Filter Complexity while Modeling Reality



Ranks input variables and output responses by correlation level

Follows MIT-developed Design Structure Matrix model format

Filters Variables Based on Correlation Level

Understand How Things Work

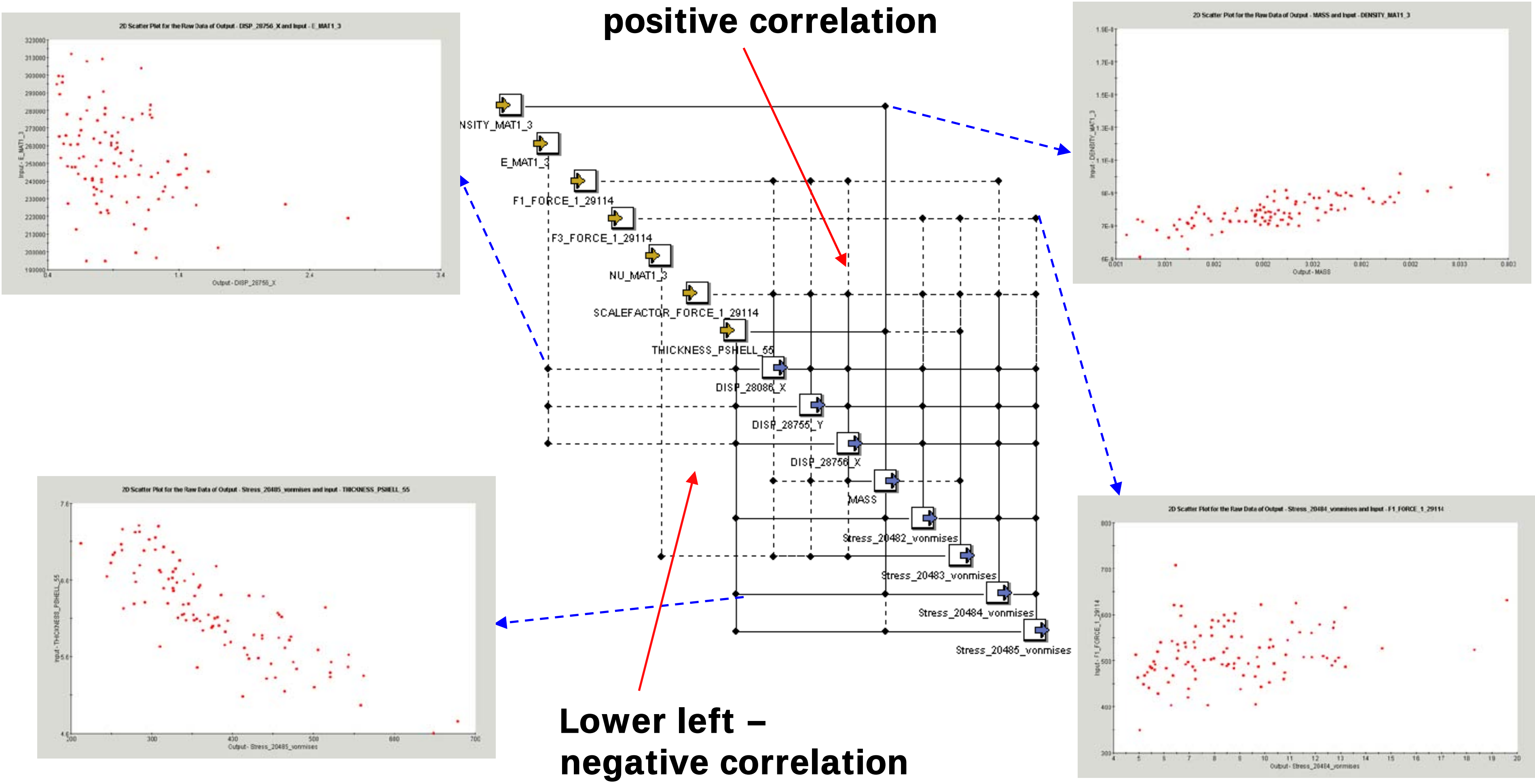
Correlation Maps: Tools for Understanding

- Displays condensed information from hundreds of analysis runs.
- Correlation Map = Structured Information = Knowledge
- A Correlation Map helps an engineer:
  - Understand how a system works.
    - How information flows within the system.
    - how variables and components correlate.
  - Make decisions on how a design may be improved.
    - Identify dominant design variables.
    - Use as input for stochastic design improvement.
  - Find the weak points in a system.
  - Find redundancies in a design.
  - Identify rules that govern the performance (“if A and B then C”).

There are NO algorithms to learn. The engineer concentrates on engineering, not on numerical analysis.

A Correlation Map

Upper right –
positive correlation



Lower left –
negative correlation

Correlation

- **Correlation supersedes sensitivity.**
- **Correlation** between two variables **expresses the strength of the relationship** between these variables **while taking account of the scatter in ALL the other variables** in the system.
- It is possible to compute correlations between any pair of variables (input-output, output-output, etc., where input is a design or noise variable, and output is a performance, like stress or frequency).
- **Knowledge of the correlations in a system is equivalent to understanding how that system works.**

Pearson and Spearman Correlatio

The Pearson, or linear correlation is given by:

$$r_{XY} = \frac{\sum_{k=1}^{N}(X_k - \bar{X})(Y_k - \bar{Y})}{\sqrt{\sum_{k=1}^{N}(X_k - \bar{X})^2}\sqrt{\sum_{k=1}^{N}(Y_k - \bar{Y})^2}},$$

The Spearman, or rank correlation is given by:

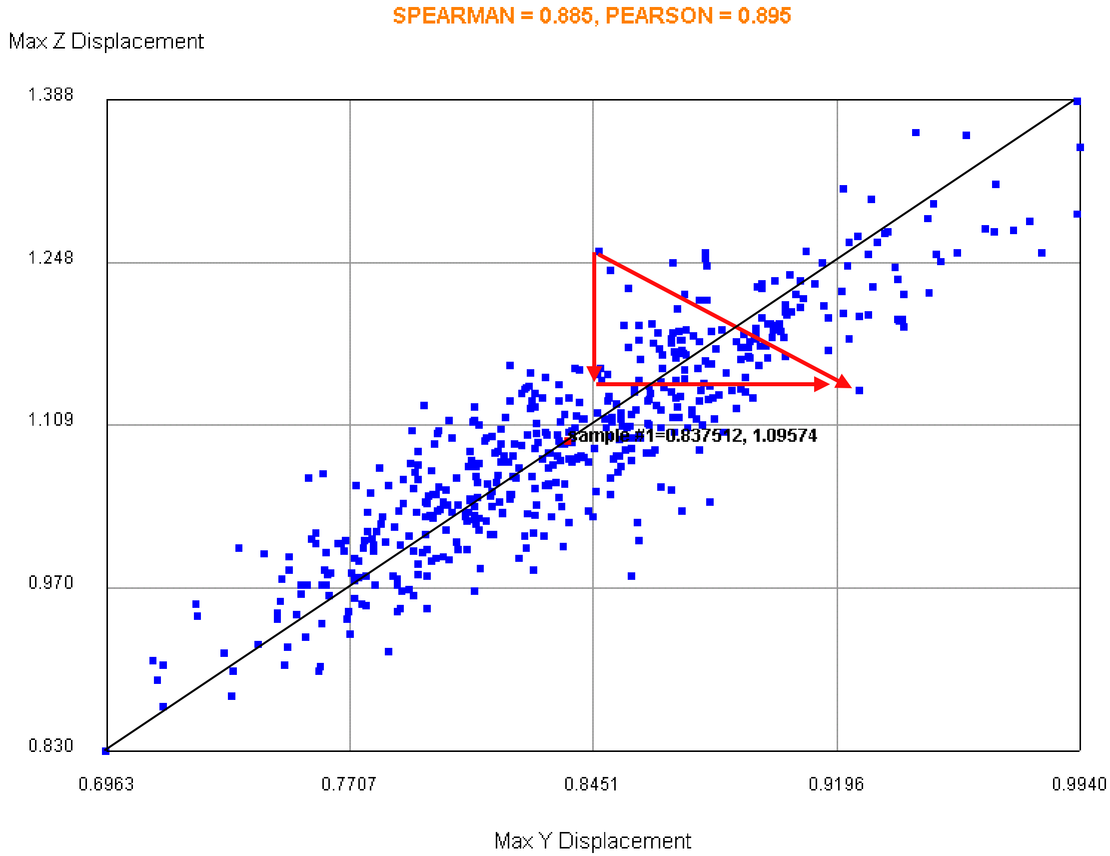
$$s_{XY} = \frac{\sum_{k=1}^{N}(R_k - \bar{R})(S_k - \bar{S})}{\sqrt{\sum_{k=1}^{N}(R_k - \bar{R})^2}\sqrt{\sum_{k=1}^{N}(S_k - \bar{S})^2}},$$

In the Spearman correlation, variable values are replaced with the corresponding ranks.

Correlation

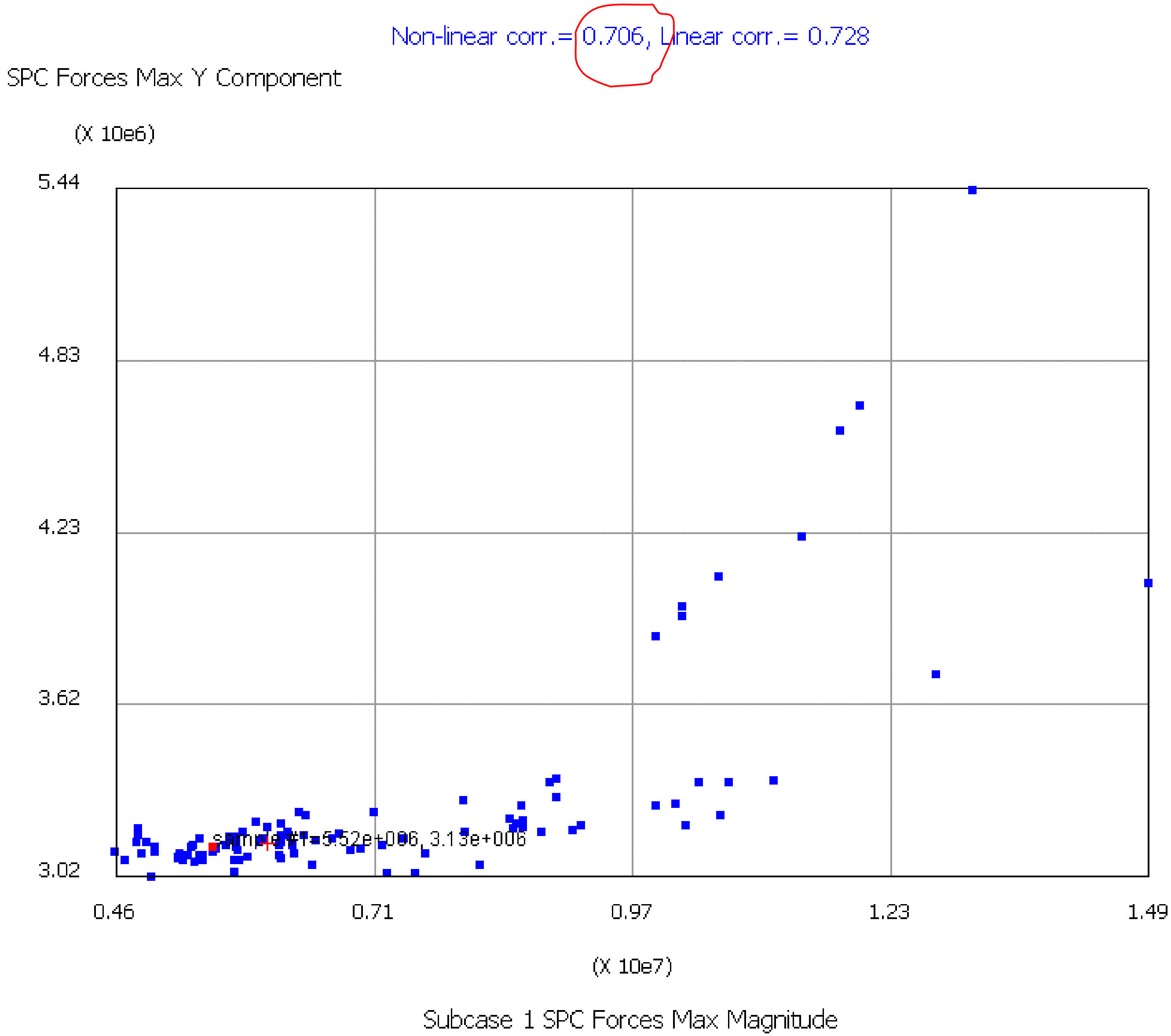
- **Positive**. When one variable increases the other increases too.
- **Negative**. When one variable increases, the other decreases.
- **Linear** Correlation (or Pearson correlation) – means that scatter plot (also known as ant-hill plot) is "cigar shaped" (elliptical).
- **Non-linear** Correlation (or Spearman correlation) – means that scatter plot is irregular (e.g. has clusters).
- If the Spearman and Pearson correlations are not similar in terms of magnitude, then you can be sure that the input-output transfer properties of a system are non-linear.

Correlation

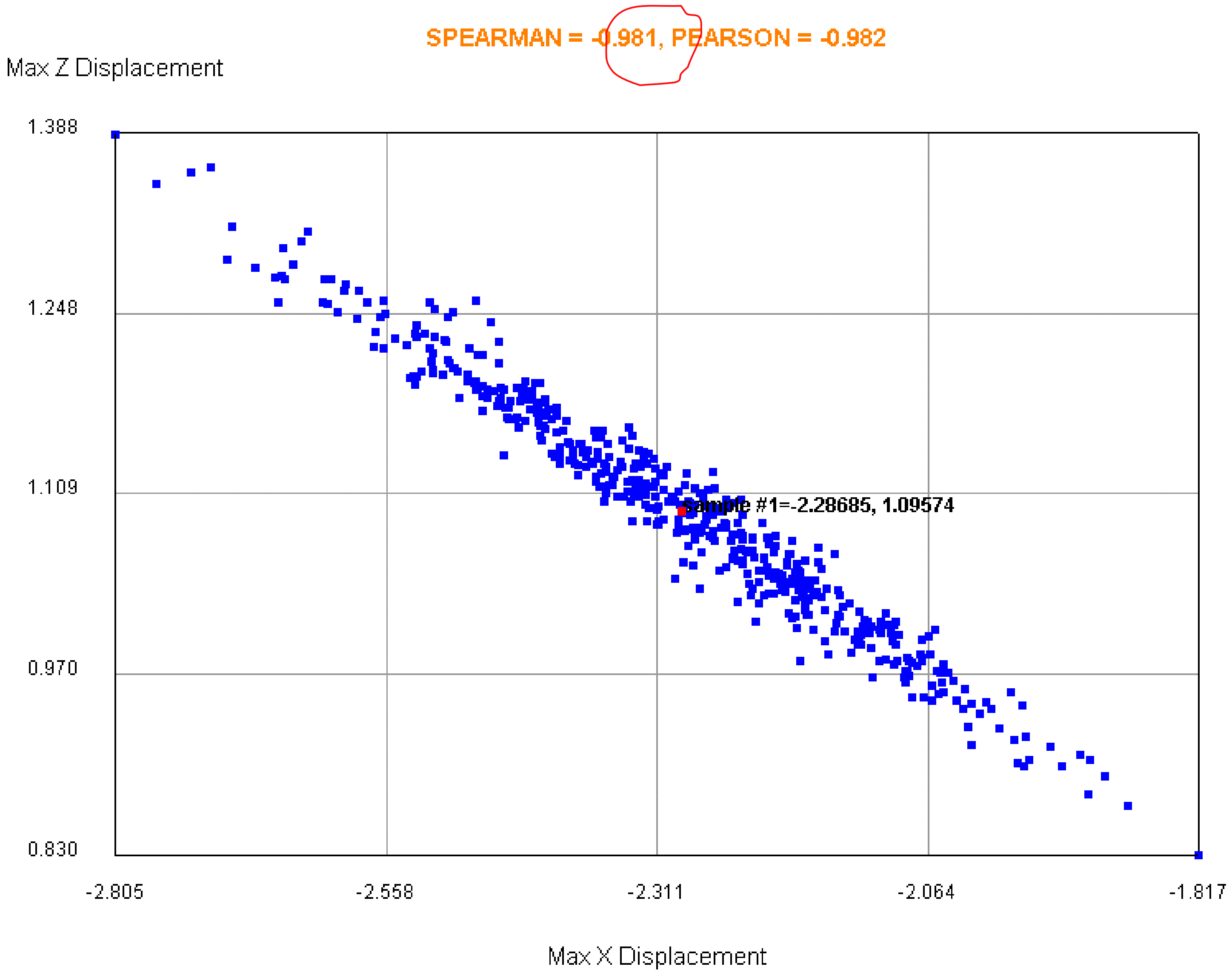


- A large value of correlation doesn't guarantee that a variable will increase if the other increases by a small amount - (as in the above case).
- A correlation is only an indication of a general trend, that can be verified only over large intervals.
- Handle with care!

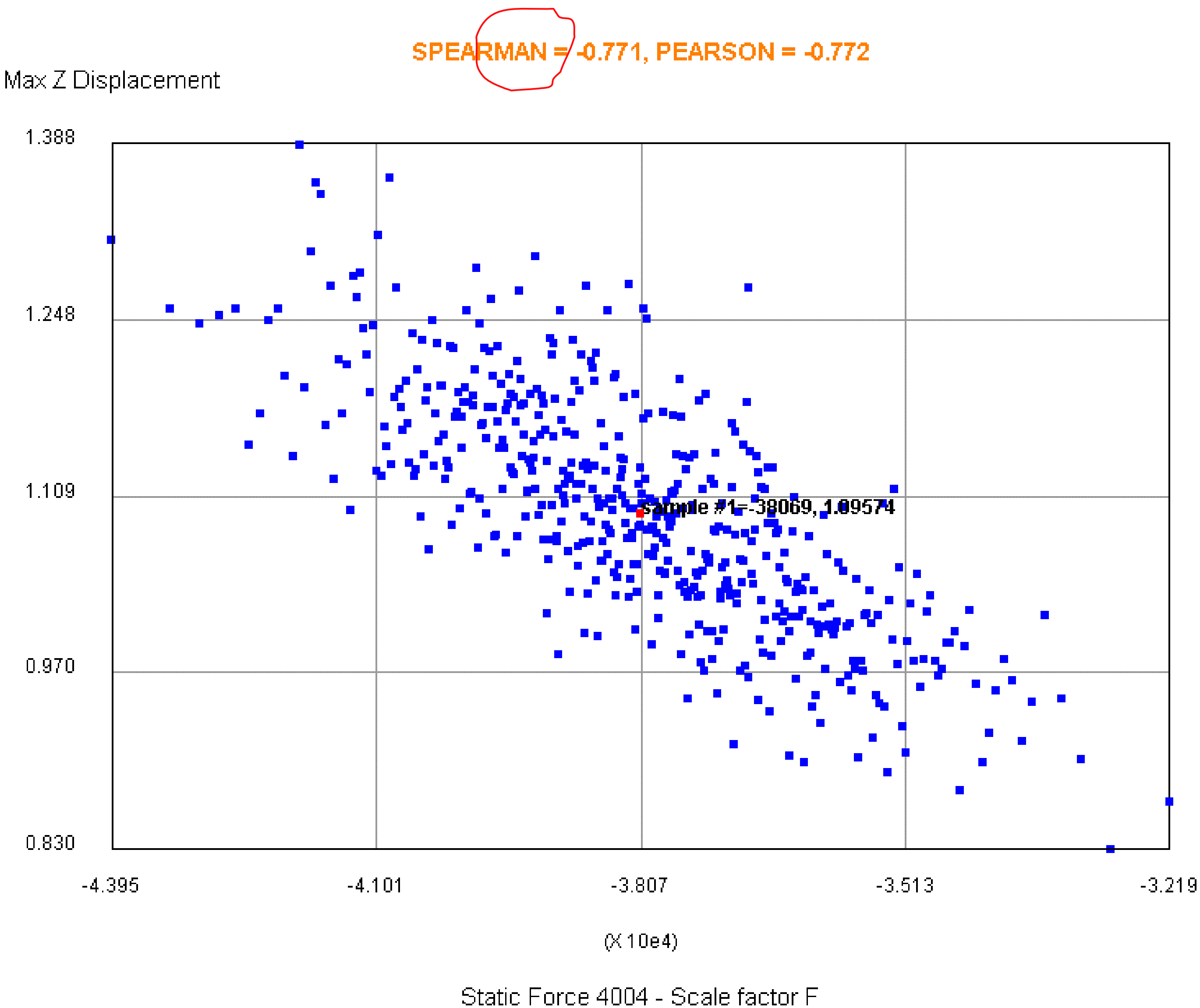
Non-linear Correlation



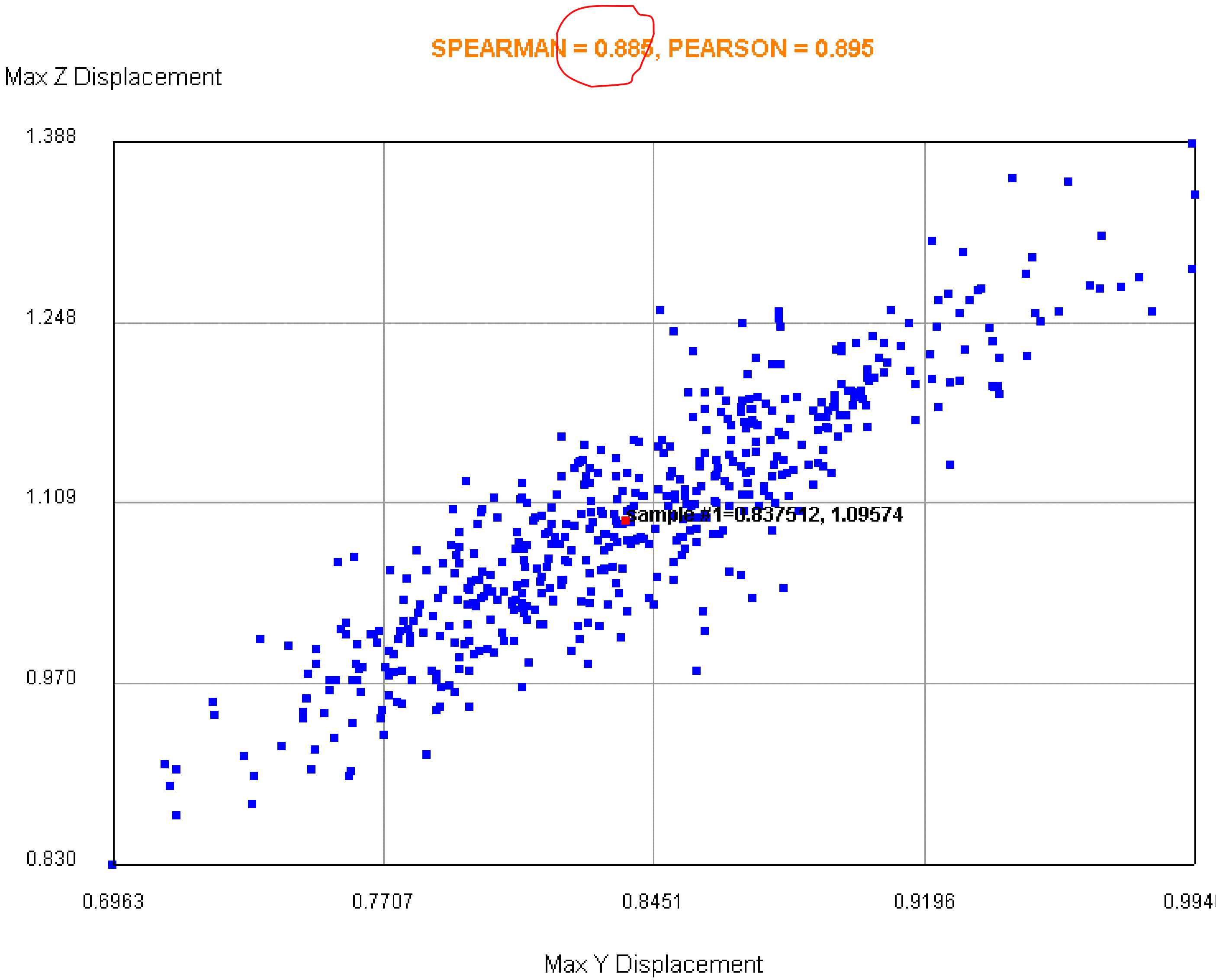
Large Negative (Linear) Correlation



Medium Negative Correlation

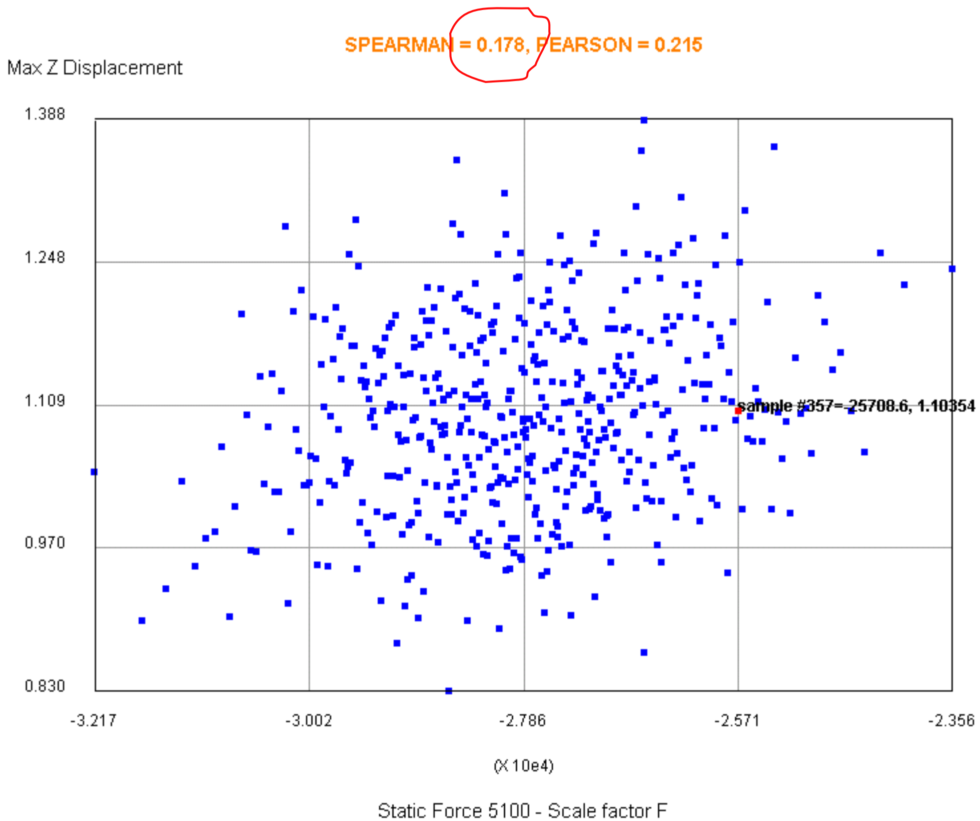


Large Positive Correlation

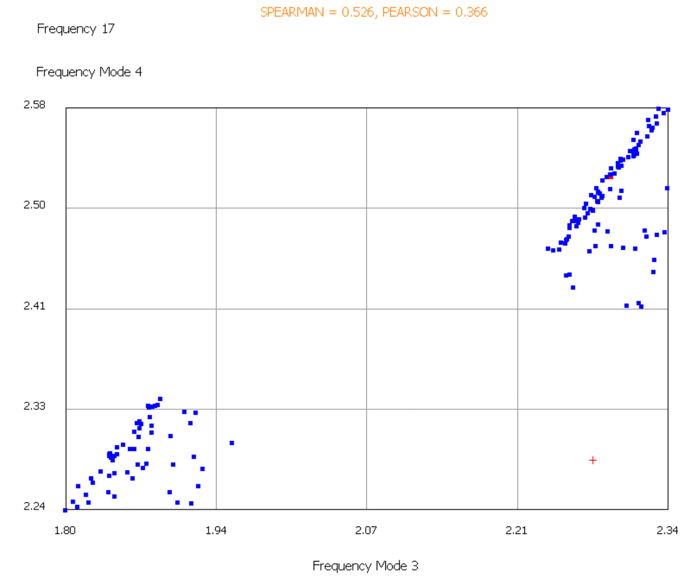


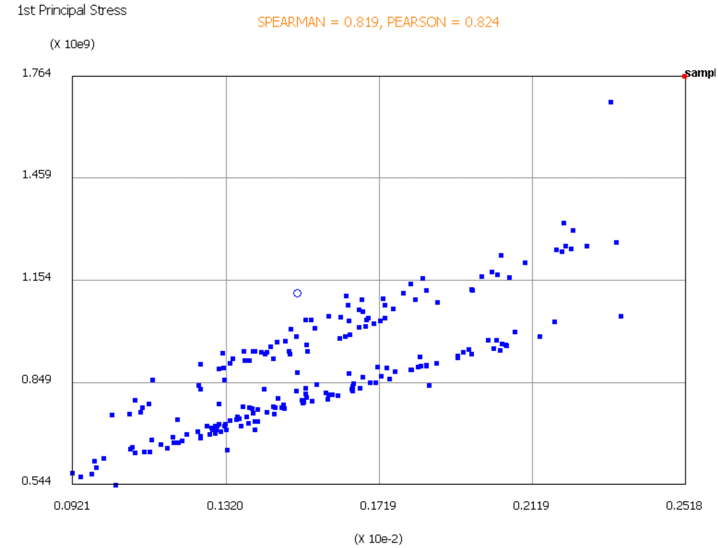
Correlation

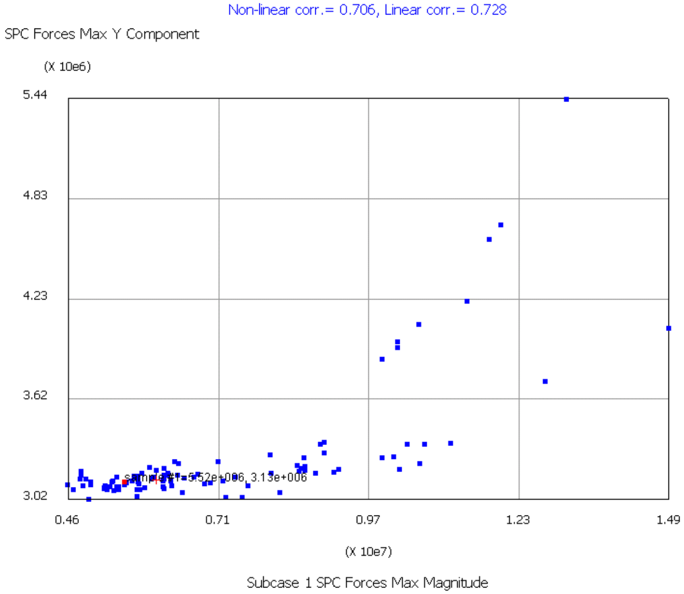
- **Correlation values below 0.5 may be neglected, as they represent meaningless (chaotic) relationships between variables. In these cases, the scatter in the other variables masks the relationship between the two in question.**

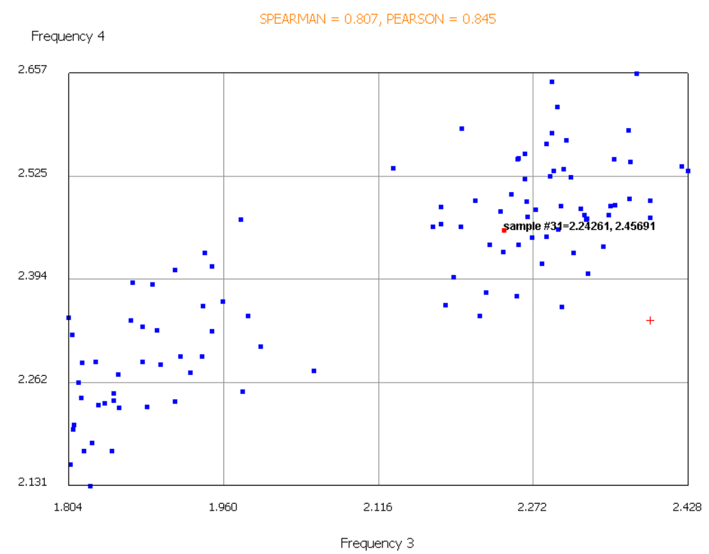


Clustering (Bifurcations)

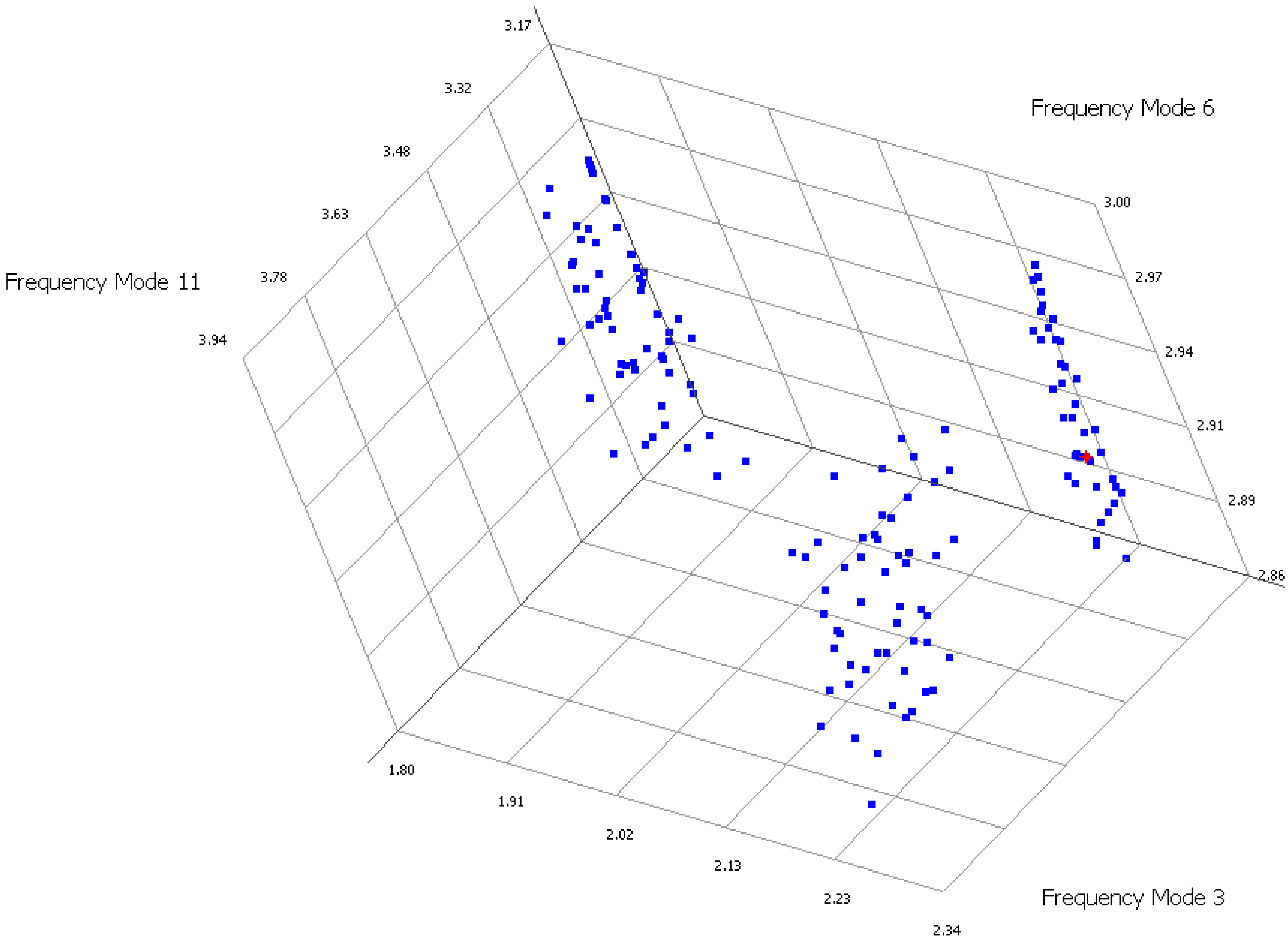




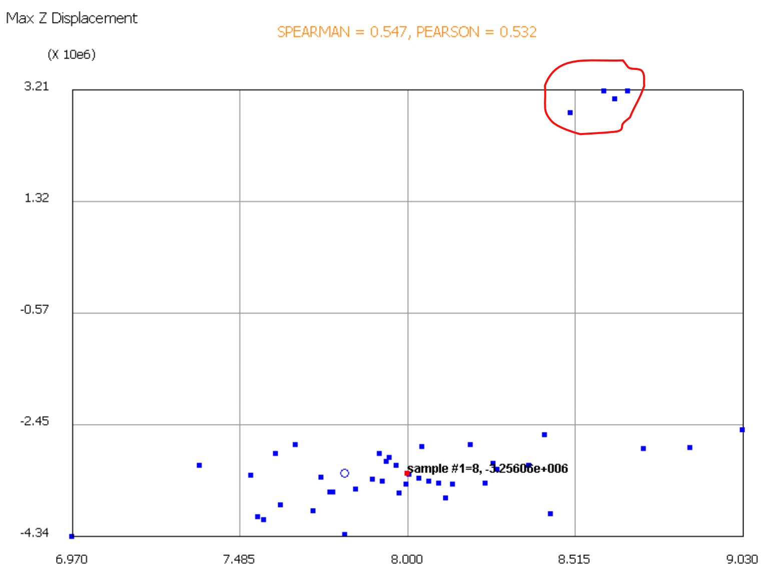




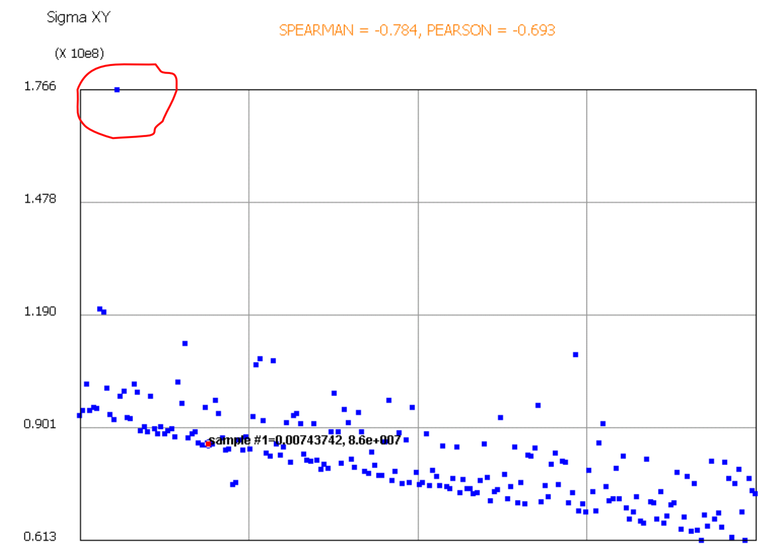
Clustering (Bifurcations)



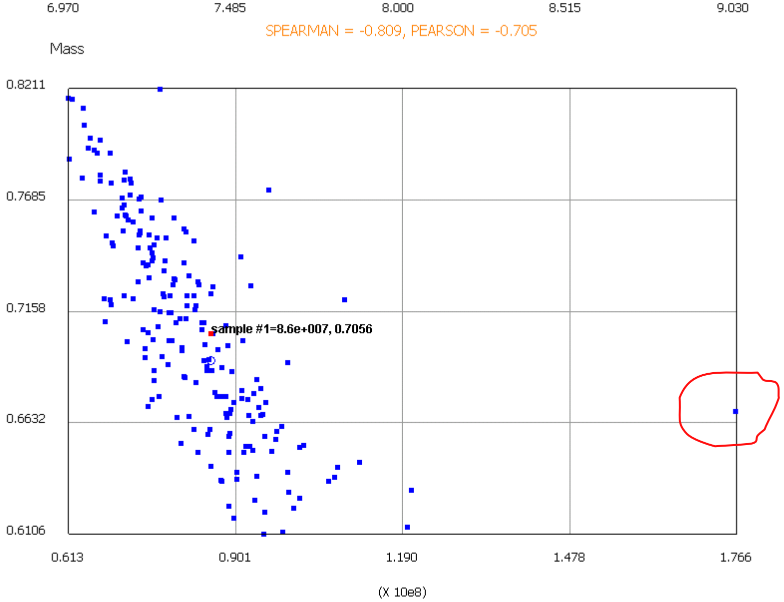
Outliers



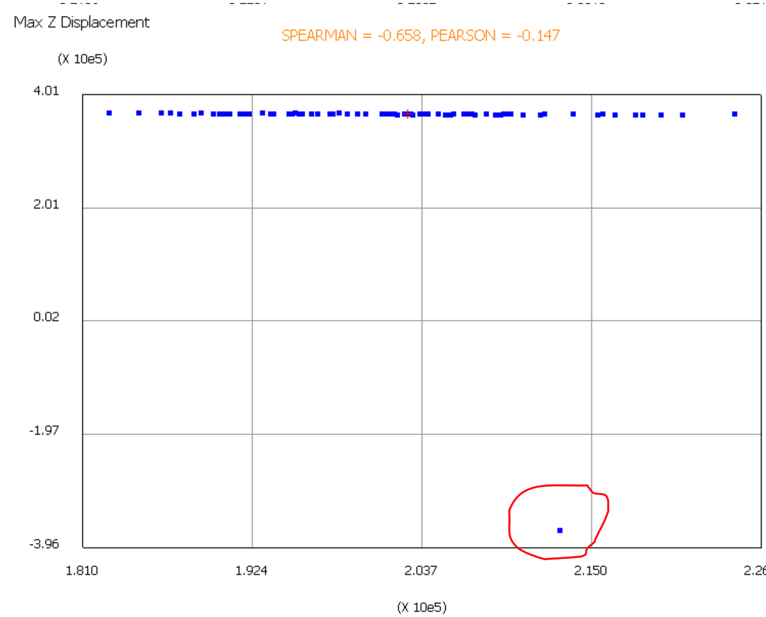
Max Z Displacement
SPEARMAN = 0.547, PEARSON = 0.532
(X 10e6)
3.21
1.32
-0.57
-2.45
-4.34
6.970
7.485
8.000
8.515
9.030



Sigma XY
SPEARMAN = -0.784, PEARSON = -0.693
(X 10e8)
1.766
1.478
1.190
0.901
0.613

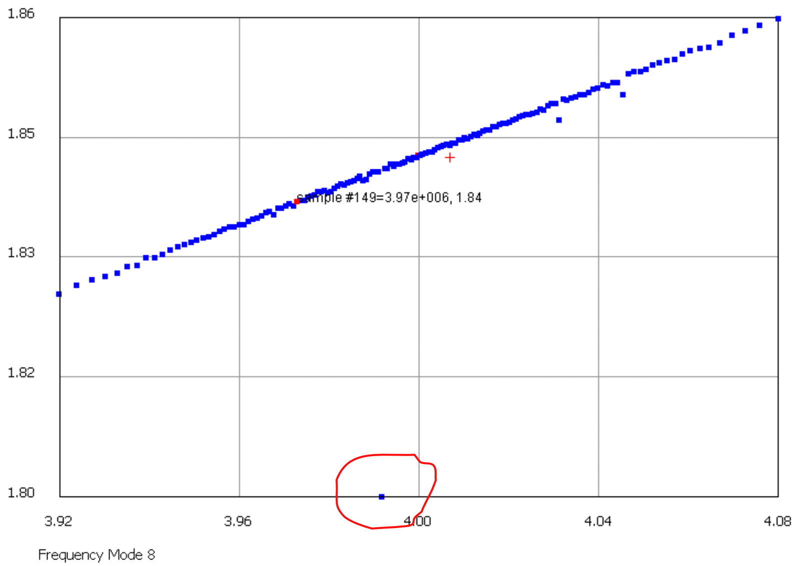


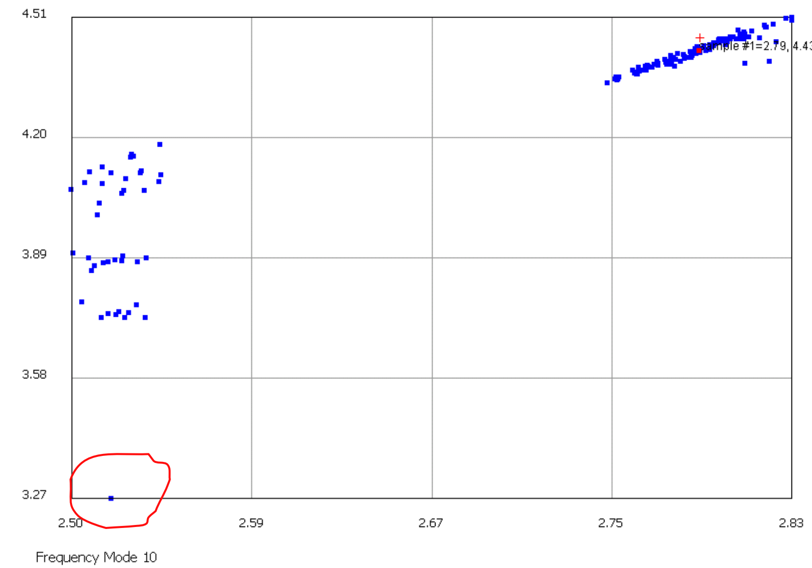
SPEARMAN = -0.809, PEARSON = -0.705
Mass
0.8211
0.7685
0.7158
0.6632
0.6106
sample #1=8.6e+007, 0.7056
0.613
0.901
1.190
1.478
1.766
(X 10e8)

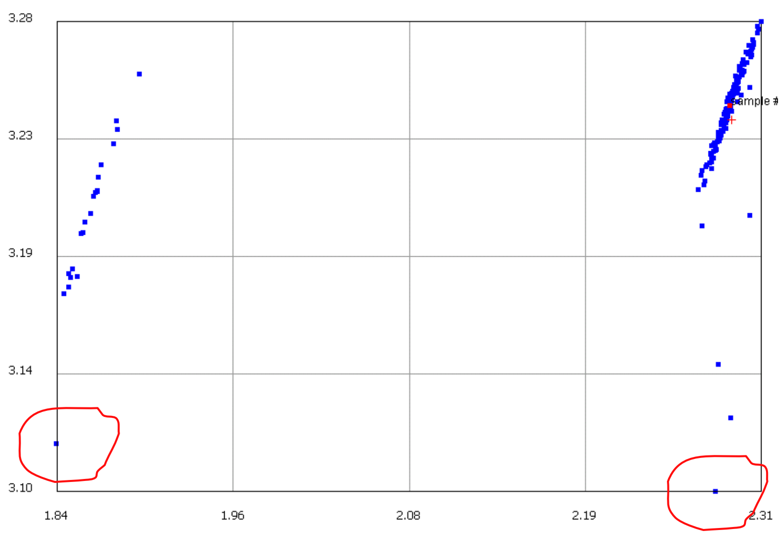


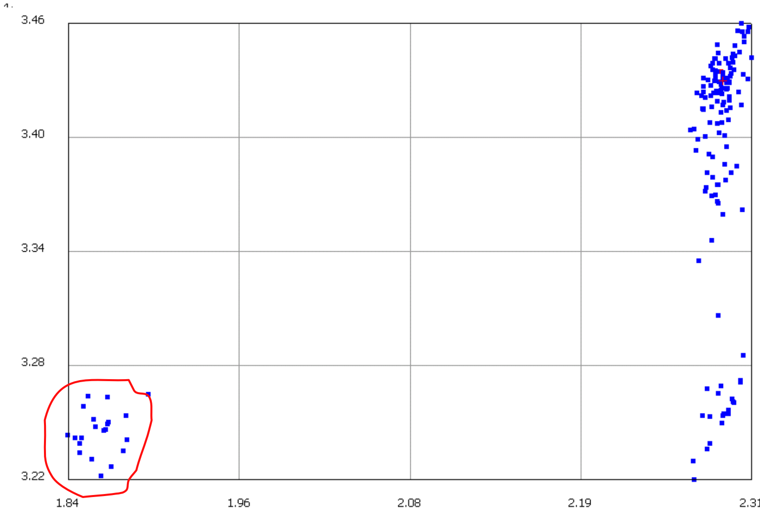
Max Z Displacement
SPEARMAN = -0.658, PEARSON = -0.147
(X 10e5)
4.01
2.01
0.02
-1.97
-3.96
1.810
1.924
2.037
2.150
2.264
(X 10e5)

Outliers

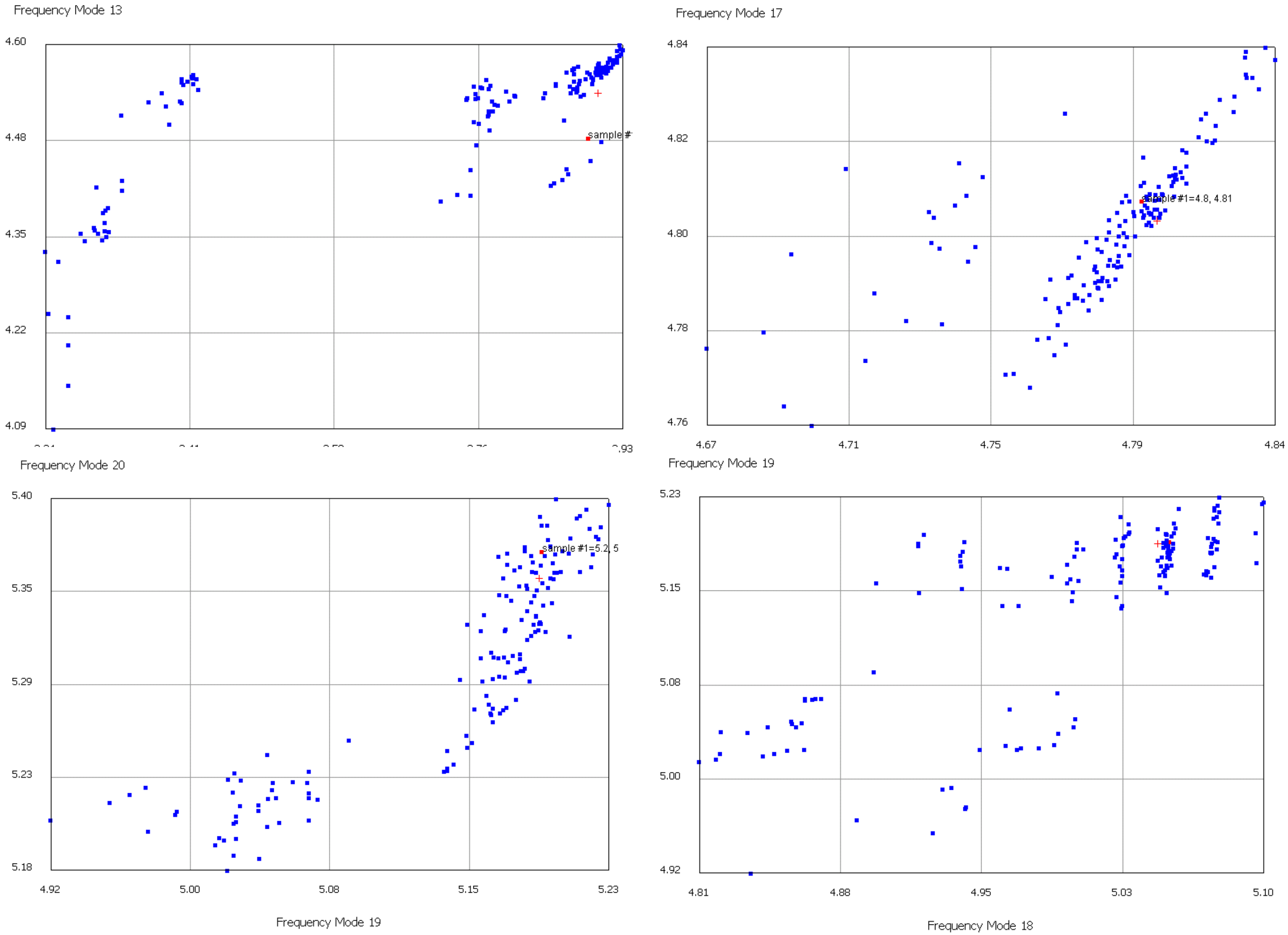

1.86
1.85
1.83
1.82
1.80
3.92
3.96
4.00
4.04
4.08
sample #149=3.97e+006, 1.84
Frequency Mode 8


4.51
4.20
3.89
3.58
3.27
2.50
2.59
2.67
2.75
2.83
sample #1=2.79, 4.43
Frequency Mode 10

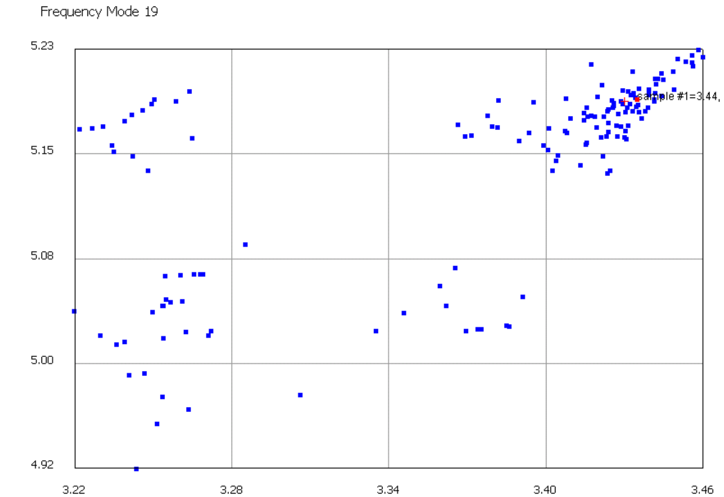

3.28
3.23
3.19
3.14
3.10
1.84
1.96
2.08
2.19
2.31
sample #

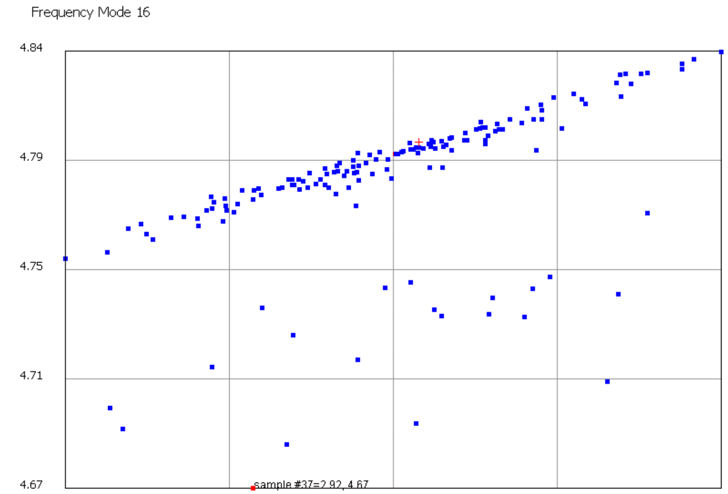

3.46
3.40
3.34
3.28
3.22
1.84
1.96
2.08
2.19
2.31

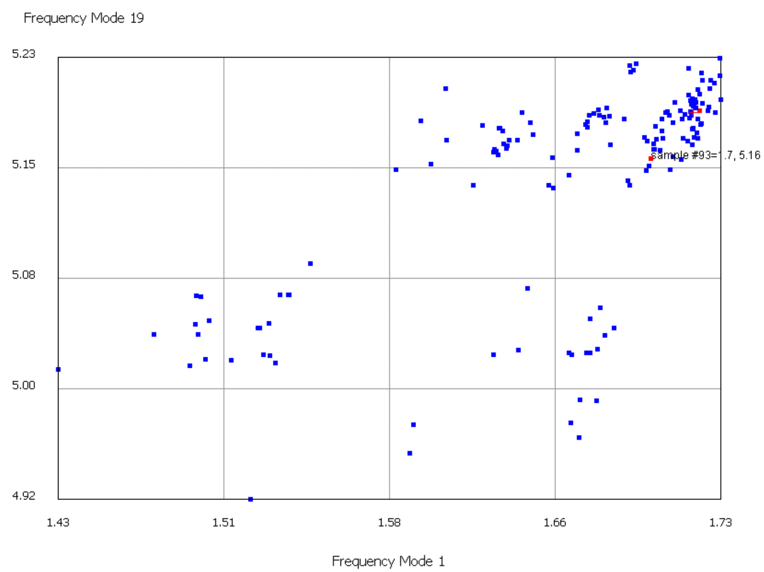
Other Pathologies

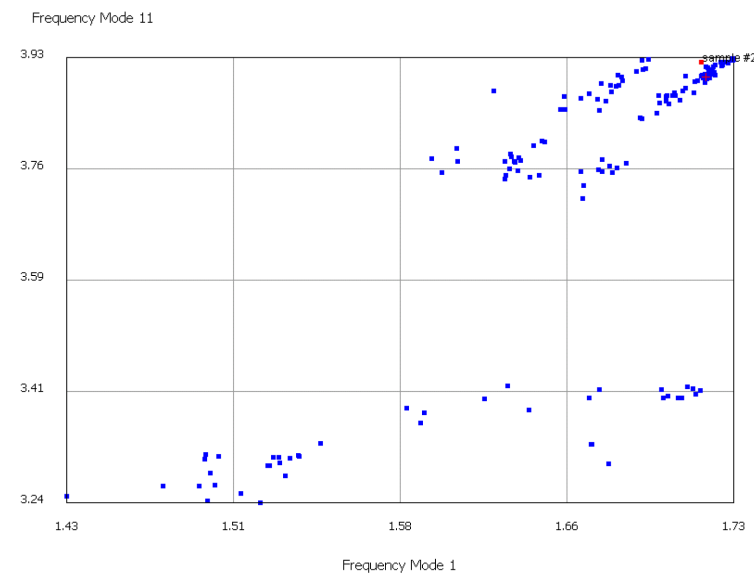


More pathologies





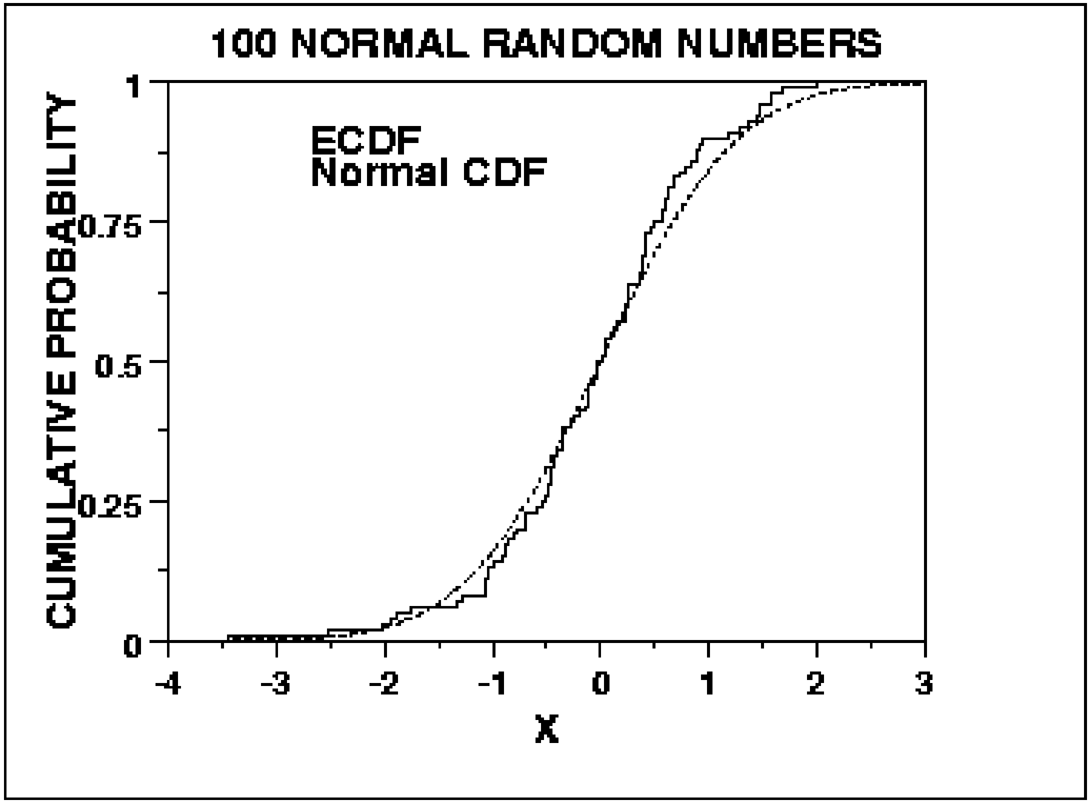




Relationship Robustness

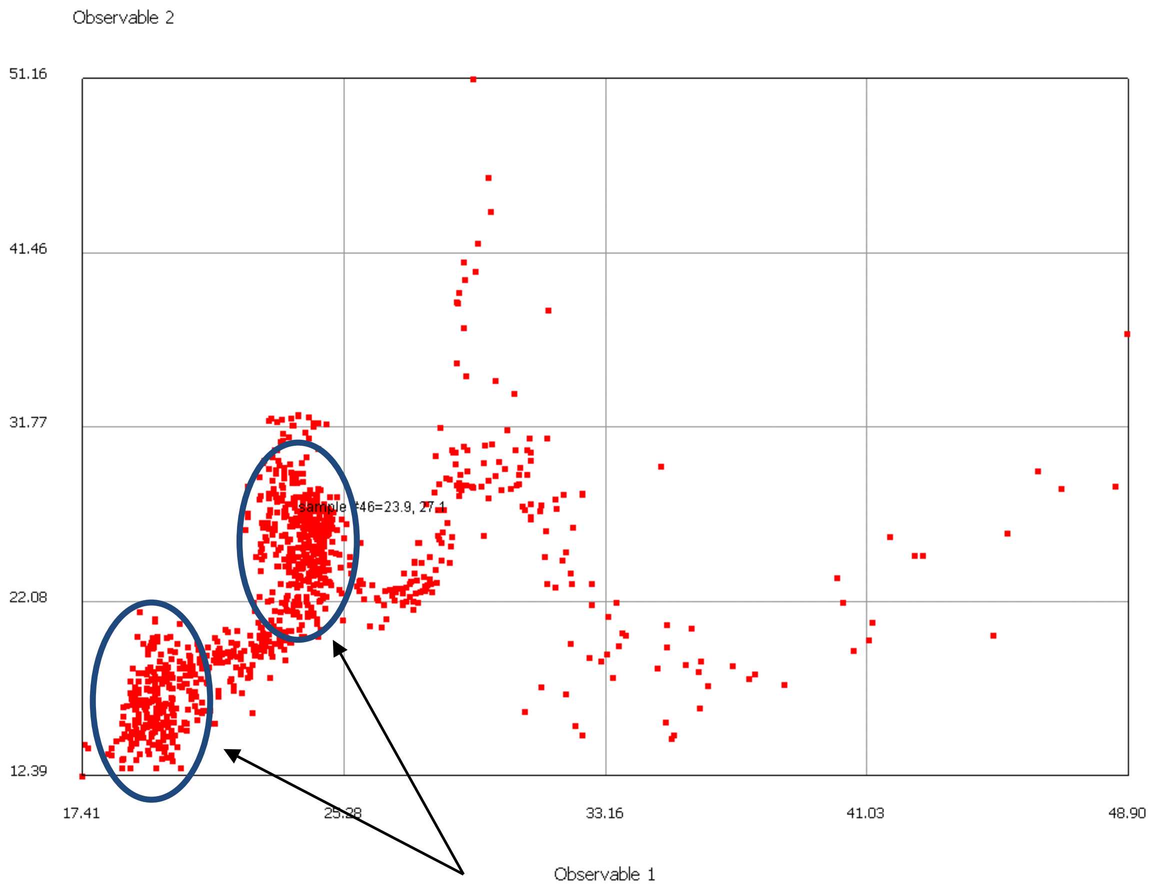
- The following forms of non-robustness exist:
  - Outliers
  - Clustering (multi-modal systems)
  - Discontinuities (jumps) in CDFs – high skewness
  - Any other pathology in ant-hills
- A good measure of robustness is the K-S distance (Kolmogorv-Smirnov) between the actual Cumulative Distribution Function (CDF) and the one corresponding to an equivalent Gaussian CDF.
- Another good measure of robustness is the CONVEXITY of the corresponding meta-model. Non robust systems are non-convex.
- An easy way to detect potential non-robustness is to check for large differences in the linear and non-linear correlation coefficients.
- Large scatter does not imply low robustness.

Kolmogorov-Smirnov (K-S) test



The Kolmogorov-Smirnov (K-S) test is based on the empirical distribution function (ECDF) compared to a normal cumulative distribution function. The K-S test is based on the maximum distance between these two curves.

Example of Non-Robust system

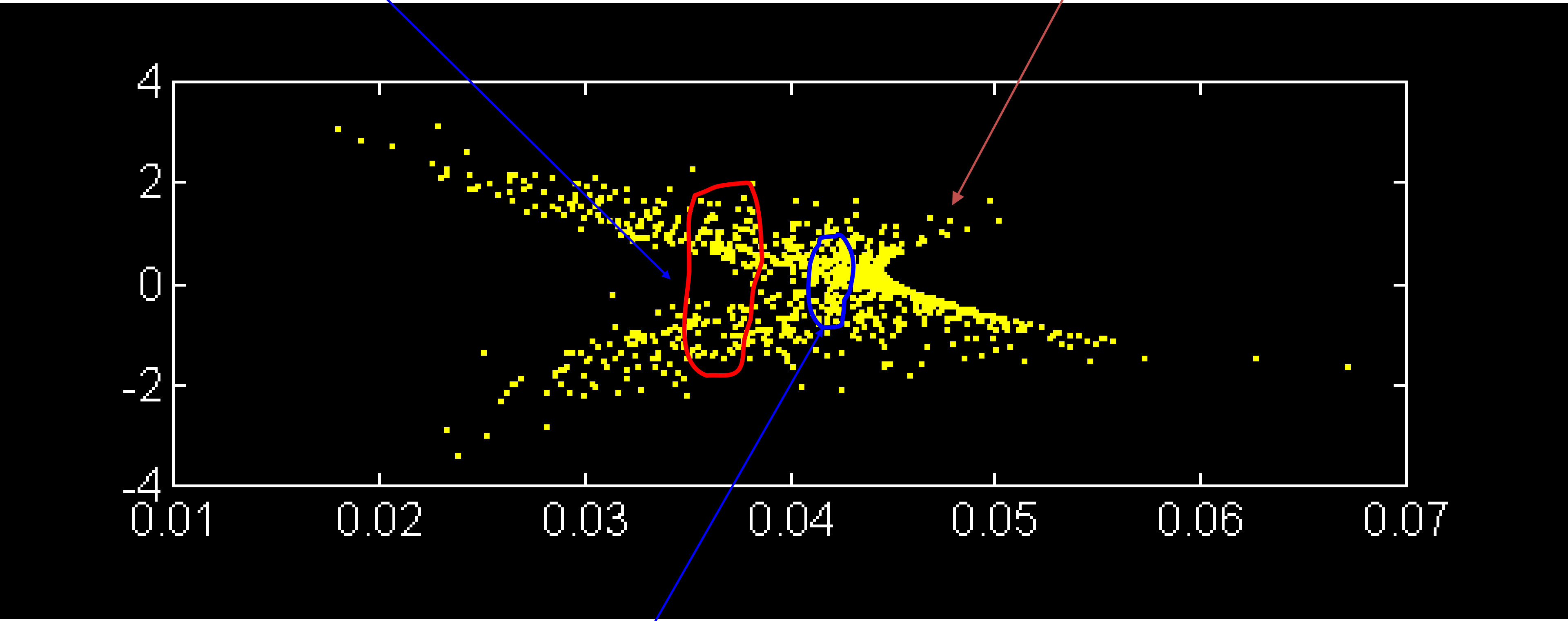


Most likely behavior (two possible modes)

Design Improvement and Robustness: Navigation of Fitness Landscapes

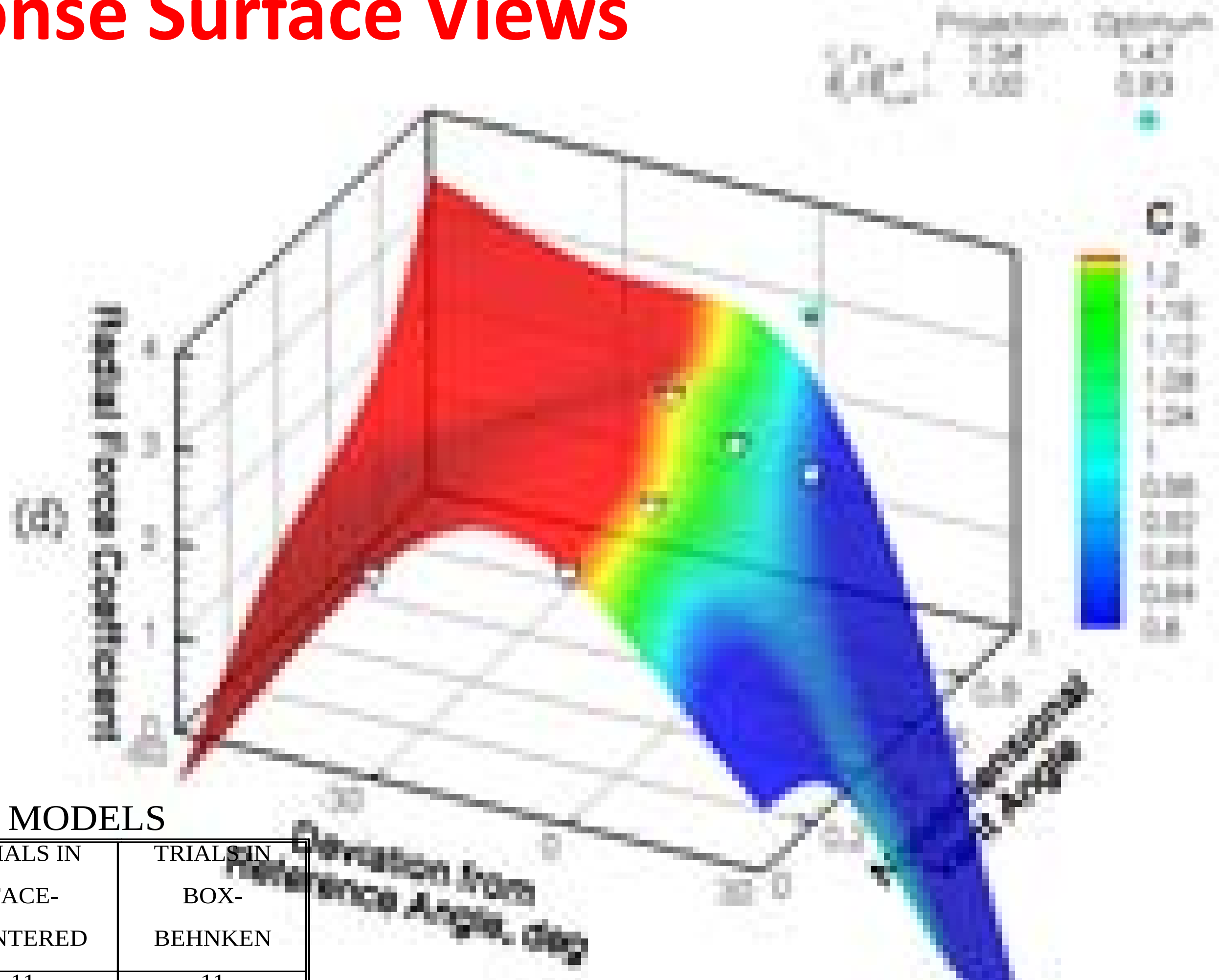
Initial design is non-robust (has two clusters, PDF is bi-modal). Must move cloud away from bifurcation.

Fitness landscape



Robust design: one cluster (no bifurcations possible)

Response Surface Views



Assume Continuity

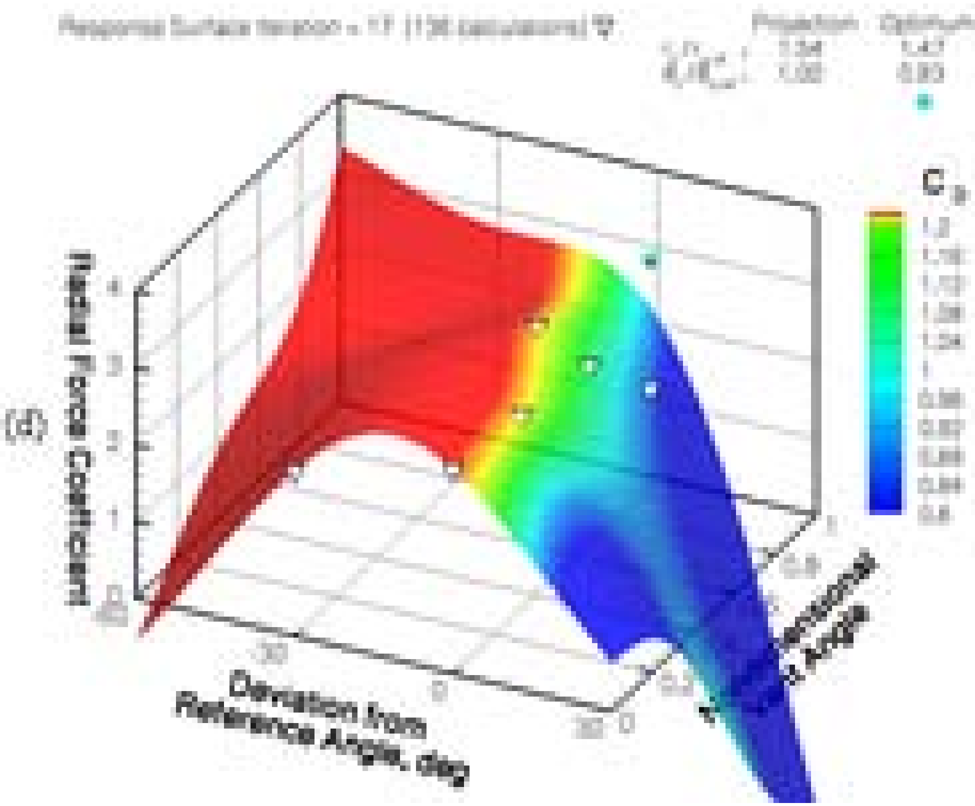
SIZE OF RESPONSE SURFACE MODELS

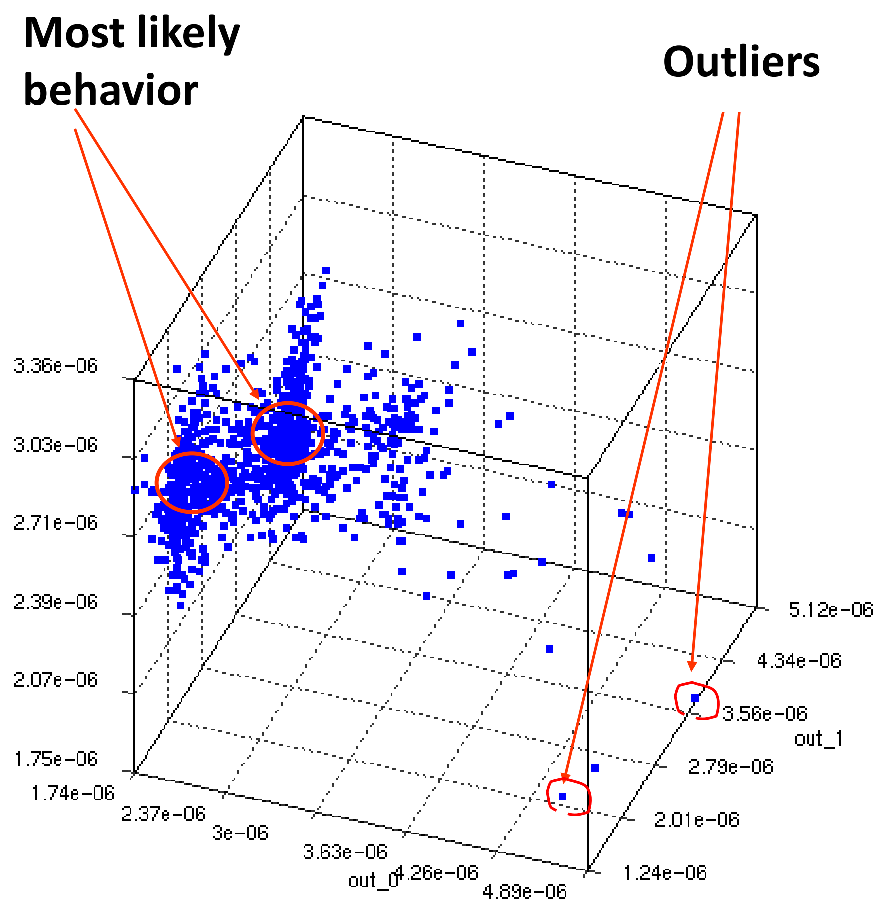
NUMBER OF FACTORS	RUNS IN 3 LEVEL FACT.	COEF. IN FULL QUAND	TRIALS IN FACE-CENTERED	TRIALS IN BOX-BEHNKEN
2	9	6	11	11
3	27	10	17	15
4	81	16	27	27
5	243	21	45	46
6	729	28	81	54
7	2187	36		62

Ohio State University FST 650 Lecture 4B

Sample Size a function of # of variables

Where is the Information?





Information from Correlation Maps

- **Minimum use of surrogates – process uses the full analysis model and is unlimited in terms of number of variables.**
- **Helps extract knowledge that is embedded in their analysis models - leverages years of investment made in the FEM and CFD grids.**
- **Actually "computes knowledge" and presents it to the engineer in the form of the Correlation Map.**
  - **Knowledge is an organized set of related design rules**
  - **Precisely what a Correlation Map is.**
- **No need to be a specialist in numerical analysis to take full advantage of this capability.**
- **Identifies what is IMPORTANT.**

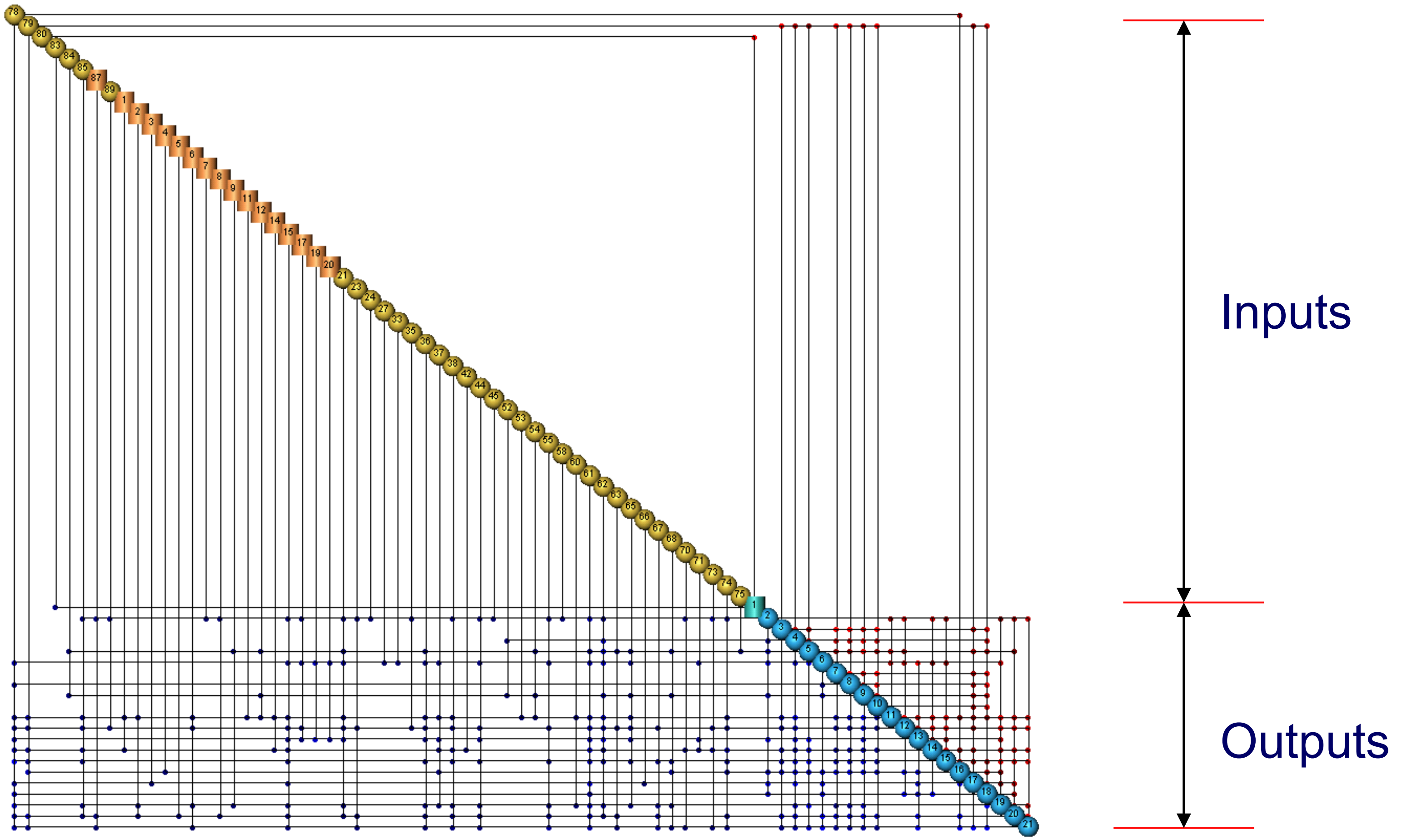
Correlation Map

- **A Correlation Map = Structured Information**
- **A Correlation Map helps an engineer:**
  - **Understand how a system works (how information flows within the system, how variables and components correlate).**
  - **Make decisions as to how the design may be improved.**
  - **Find the weak points in a system.**
  - **Find redundancies in a design.**
  - **Identify rules that govern the performance (“if A and B then C”).**
- **Correlation Maps help structure information better, to give it “topology” and a sense of dimension.**
- **Structured information (i.e. links between rules) is equivalent to knowledge.**
- **The following slides show an examples of Correlation Maps and their salient features.**

Design Structure Matrix

- **Format developed at MIT to maximize the data displayed in one view**
- **The Design Structure Matrix (DSM) has evolved into a tool to perform both the analysis and the management of complex systems.**
- **It enables the user to model, visualize, and analyze the dependencies among the entities of any system and derive suggestions for the improvement or synthesis of a system.**
- **Focus is on the visualization format**

Correlation Maps – Structured Information



Correlation Maps reflect how all system attributes (outputs) react to simultaneous changes in all of the input variables. It also shows relationships between outputs. Correlation maps can be very complex.