MORS

Contractor Disclosure Form 712A – Deadline: 2 June 2014
MORS Symposium
16-19 June 2014, Hilton Mark Center, Alexandria, VA
Fax completed form to 703-933-9066 or email to liz@mors.org

Abstract 595

PART I **Author Request** - The following author(s) request authority to disclose the following presentation at the MORS Symposium with subsequent publication in the MORS Final Report, and posting on the MORS website, if applicable.

Principal Author: Thomas A Donnelly	Other Author(s): N/A
Principal Author's **Organization** and complete mailing address: SAS Institute Inc. 27 Farmingdale Ln Newark, DE 19711	Principal Author's Signature: X Thomas A. Donnelly Date: 28 May 2014
	Phone: 302-489-9291 FAX: 919-677-4444
	Email: tom.donnelly@jmp.com

Title of Presentation:
Efficient Modeling and Simulation (M&S) Using Design of Experiments (DOE) Methods

This presentation is: ☐ SECRET ☐ SECRET//REL TO FVEY ☐ CONFIDENTIAL ☐ CONFIDENTIAL//REL TO FVEY ■ UNCLASSIFIED ☐ Other ________ and will be presented in:

■ Tutorial ☐ List all WG(s) #:

This work was performed in connection with a government contract.	☐ YES (Complete Parts I, II, & III)
This presentation is based on material developed by the author as part of company-approved research e.g. IR&D and was NOT done under a government contract.	☐ YES (Complete Parts I, II & III)
This presentation was NOT done under a government contract, contains no government information, is my own work and is approved for public release.	■ YES (Complete Part I only)

Efficient Simulation Using Design of Experiments (DOE)

June 16, 2014
82nd MORSS Tutorial
Alexandria, VA

Tom Donnelly, PhD
SAS Institute Inc.
JMP Systems Engineer & Co-insurrectionist

tom.donnelly@jmp.com

Outline

- Overview of Design of Experiments (DOE)
  - Slides 1 to 27 – background – many may be skipped if most people have an understanding of DOE
- Efficient M&S Using DOE
  - Sequential traditional DOE
  - Sequential space-filling DOE

My Background

Design of Experiments (DOE) for 29 Years

- ‘83-’87 Honeywell, Inc., Engineer

  First saw the power of DOE in 1984 – career changing event

- ‘87-’99 ECHIP, Inc., Partner & Technical Director

  200+ DOE courses, on-site at 40+ companies - many chemical/food/pharma - requiring mixture/formulation DOE

- ‘99-’05 Peak Process, LLC, Consultant

- **‘05-’08 US Army, Edgewood Chemical Biological Center (ECBC), Physicist, Analyst, & Co-insurrectionist**

  DOE with Real data and Modeling & Simulation data

- Dec. ’08 Joined the SAS Institute Inc., Customer Advocate

  Work in DOE and Federal Government domains
  – Data Visualization, Data Mining and their synergy with DOE
  – Support DoD sites, National Labs, & Gov’t Contractors

Projects Using DOE at U.S. Army ECBC CY05-08

Detection, Decontamination & Protection

- JPM Nuclear Biological Chemical Contamination Avoidance (NBCCA) - Whole Systems Live Agent Test (WSLAT) Team support to the Joint Biological Point Detection System (JBPDS)
- Agent Fate wind tunnel experiments
- Decontamination Sciences Team
  - Contact Hazard Residual Hazard Efficacy Agent T&E Integrated Variable Environment (CREATIVE) - real and simulation data
  - Modified vaporous hydrogen peroxide (mVHP) decontamination – real data
- Smoke and Target Defeat Team
  - Pepper spray characterization – real data
  - Obscurant material evaluation (with OptiMetrics, Inc.) – simulation data
- U.S. Army Independent Laboratory In-house Research (ILIR) on novel experimental designs used with simulations
  - **Re-analysis of U.S. Air Force Kunsan Focused Effort BWA simulation data**
  - CB Sim Suite used for sensitivity analysis of atmospheric stability
- U.S. Marine Corps Expeditionary Biological Detection (EBD) Advanced Technology Demonstration (ATD)
  - Chamber testing of detectors – real data
  - CB Sim Suite sensor deployment studies – simulation data
- U.S. Navy lead on Joint Expeditionary Collective Protection (JECP)
  - Swatch and chamber testing – real data
  - **Computational Fluid Dynamics (CFD) – simulation data**

Take Away Materials

- PDFs available
  - Today’s Slides
  - White Paper - “Efficient Modeling & Simulation of Biological Warfare Using Innovative Design of Experiments Methods”
  - MORSS Tutorial Summary 11 X 17 Handout
    “DOE for Real-World Problems”

Why Use Design of Experiments Methods with Simulation Experiments?

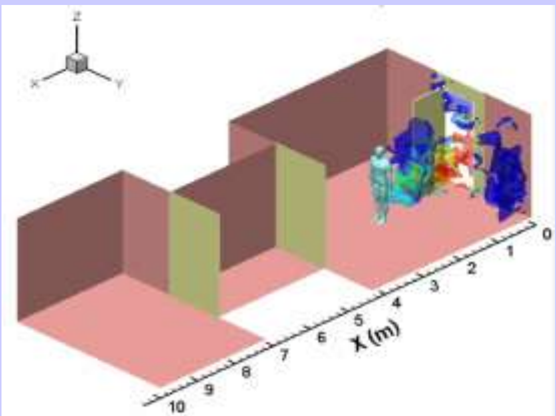
Quicker answers, lower costs, solve bigger problems

- Obtain a fast surrogate model of the simulation
  - Individual simulations can run for hours, days, weeks
    - Computational Fluid Dynamics (CFD)
    - Simulation runs in real-time
  - Numbers of factors can be very large (40+)
  - Numbers of simulations needed can be large (thousands in many cases)
  - Simulations can be stochastic requiring many replications
- Surrogate model yields a fast approximation of the simulation
  - more rapidly answer “what if?” questions
  - do sensitivity analysis of the control factors
  - optimize multiple responses and make trade-offs
- By running efficient subsets of all possible combinations, one can – for the same resources and constraints – *solve bigger problems*
- By running sequences of designs one can be as *cost effective as possible* & run *no more trials than are needed* to get a useful answer

Long Running Physics-Based Simulations

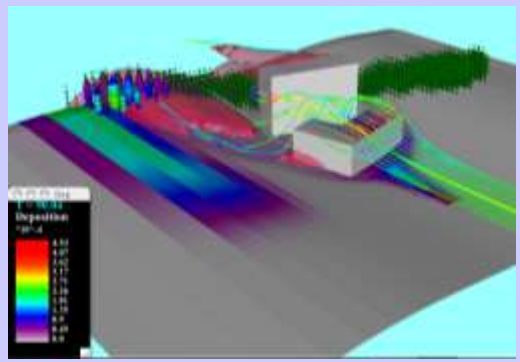
Detailed Physics Models can require a great deal of runtime to generate a short period of simulation time.

Computational Fluid Dynamics (CFD) Models



Developed for Interior
Moving Man in Simulation
8M cells
10 Seconds of Simulation
64 CPUs – 4K slower
12 Hours of Runtime

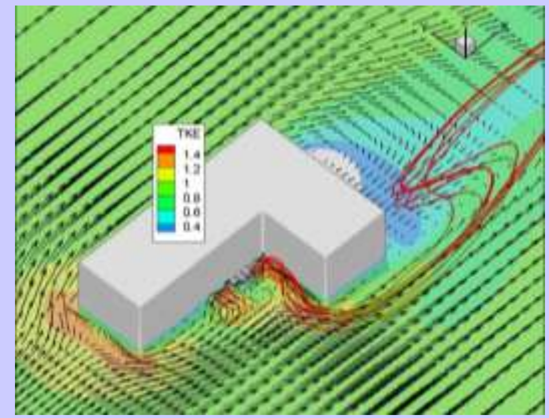
Detailed Ingress/Egress, Internal Airflow and Convection



Developed for Exterior
Stationary Grids
1.5M Cells
30 Seconds of Simulation
Single CPU – 20K slower
7 Days of Runtime

External CW Deposition/ Evaporation, Vegetation, Solar Heating

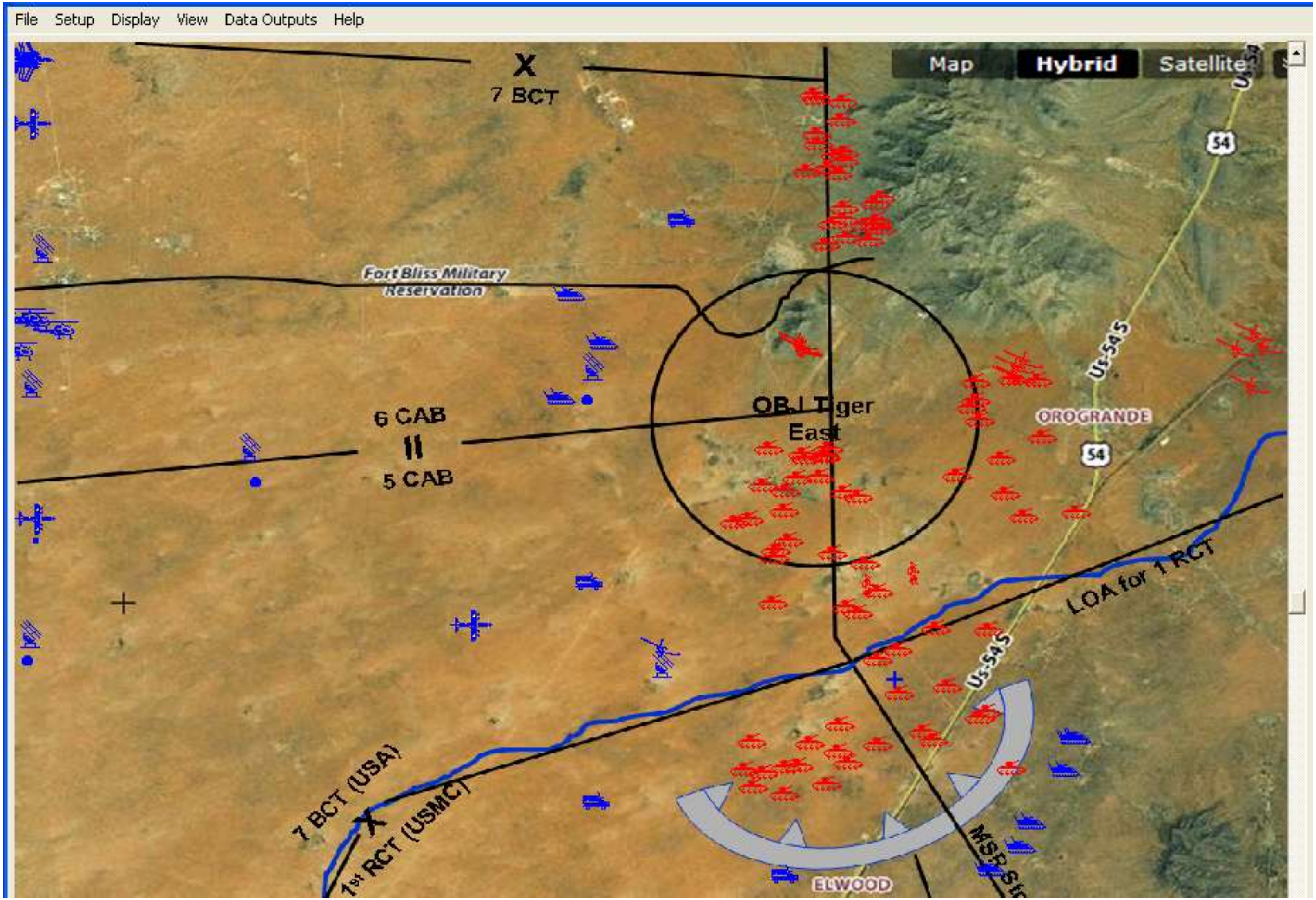
Lagrangian-Particle



Developed for Exterior
Stationary Grids
TBD Cells
Min-Hours of Simulation
Single CPU
Minutes-Days of Runtime

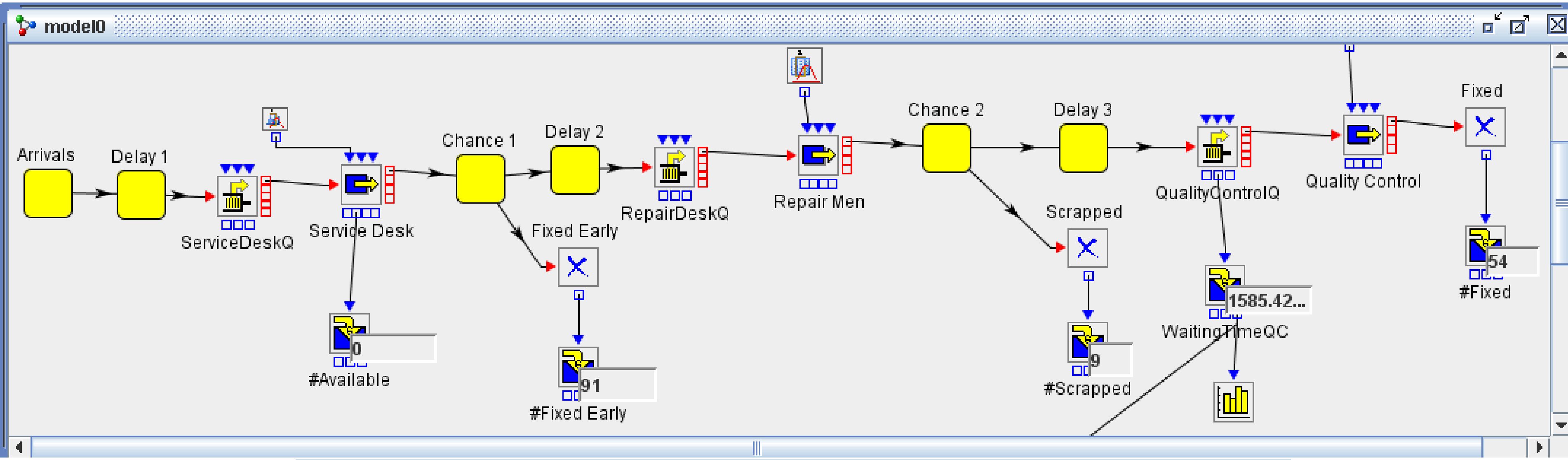
Speed, Flexibility, More User Friendly, V&V

Agent-Based Simulations



Stochastic Simulations with Many Replicates

Discrete Event Simulations

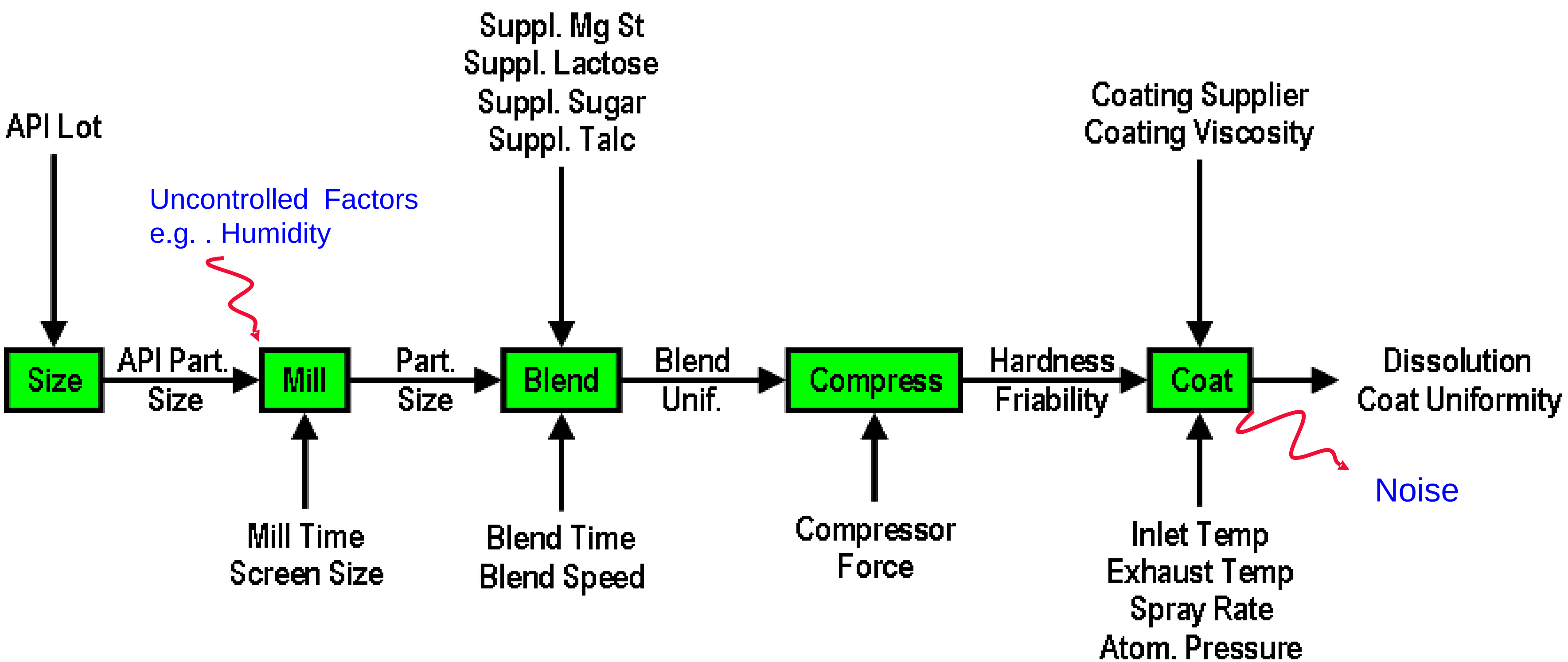


experiment0

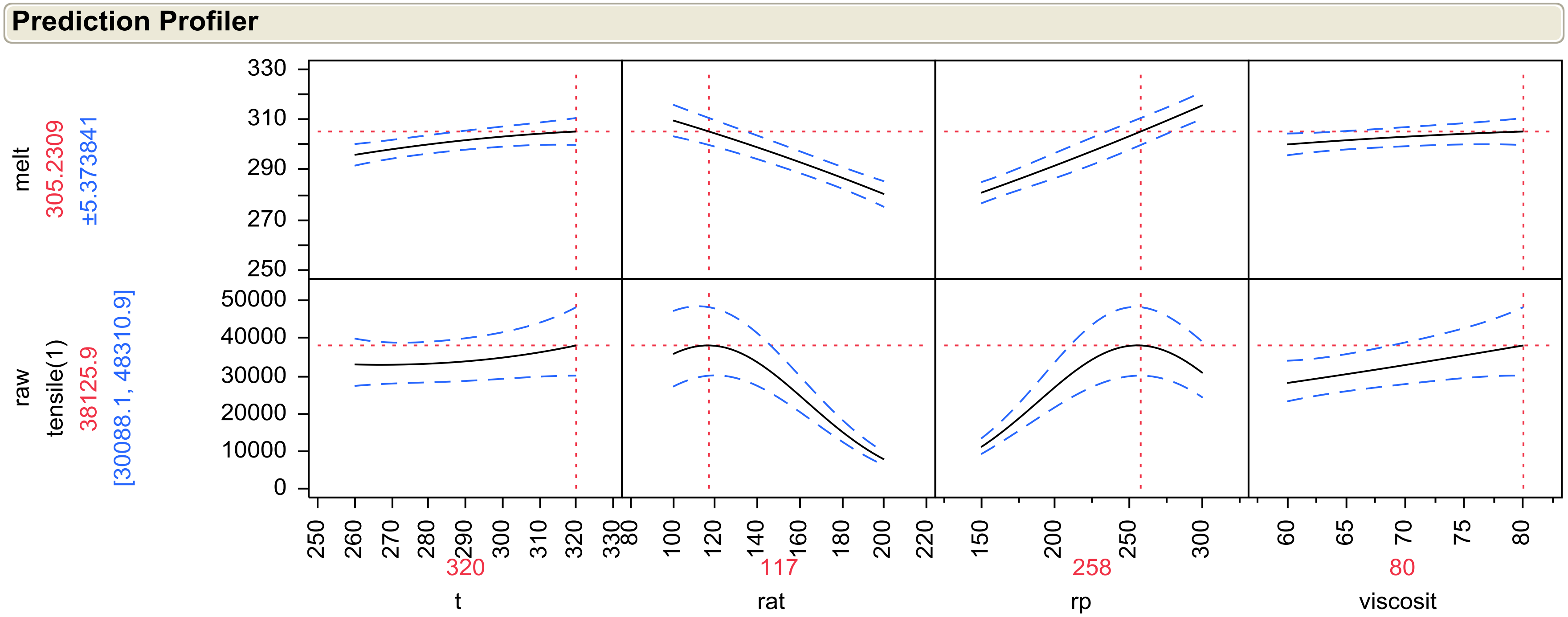
Point...	StartT...	EndTi...	Num...	Num...	Num...	Repli...	Num...	AvgUt...	AvgW...	AvgUti...	AvgUt...	AvgW...	AvgW...
point 1	0	2,700	1	1	3	5	55.8	98.1...	640....	32.90...	22.7...	1.681...	0.11...
point 2	0	2,700	3	2	1	5	105.6	61.8...	3.65...	16.39...	67.8...	0.0	16.7...
point 3	0	2,700	2	3	1	5	100.2	88.3...	84.8...	10.92...	67.8...	0.0	16.7...
point 4	0	2,700	2	1	3	5	100.6	88.5...	97.3...	32.90...	22.7...	1.681...	0.11...
point 5	0	2,700	2	1	1	5	100.2	88.3...	84.6...	32.78...	67.8...	0.233...	16.7...
point 6	0	2,700	3	1	2	5	105.8	61.9...	8.69...	32.90...	34.1...	1.382...	0.83...
point 7	0	2,700	2	2	2	5	100.4	88.4...	97.9...	16.47...	34.1...	0.020...	0.83...
point 8	0	2,700	2	2	3	5	100.6	88.5...	98.4...	16.46...	22.7...	0.094...	0.11...
point 9	0	2,700	1	1	1	5	55.8	98.1...	621....	32.78...	67.8...	0.233...	16.7...
point ...	0	2,700	3	3	3	5	105.8	61.9...	9.32...	10.97...	22.7...	0.001...	0.11...
point ...	0	2,700	1	3	2	5	55.8	98.1...	641....	10.98...	34.1...	4.305...	0.83...
point ...	0	2,700	1	2	1	5	55.8	98.1...	621....	16.39...	67.8...	0.0	16.7...

Classic Definition of DOE

- Purposeful control of the inputs (factors) in such a way as to deduce their relationships (if any) with the output (responses).



Here are 4 Controls (inputs) & 2 Responses (outputs) and their empirical relationships (model)



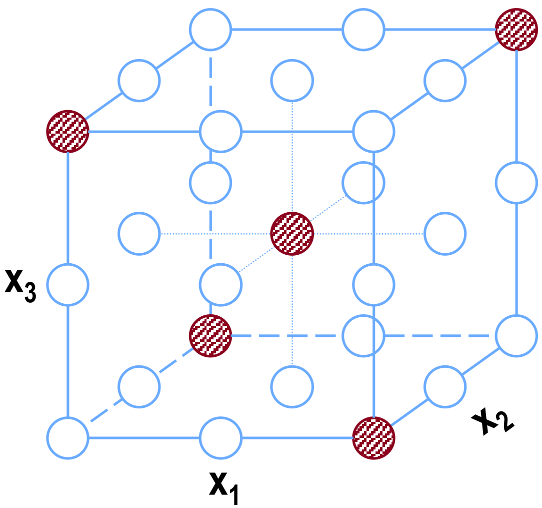
Get this Prediction Profiler as result of analyzing data collected for a DOE

Alternative Definition

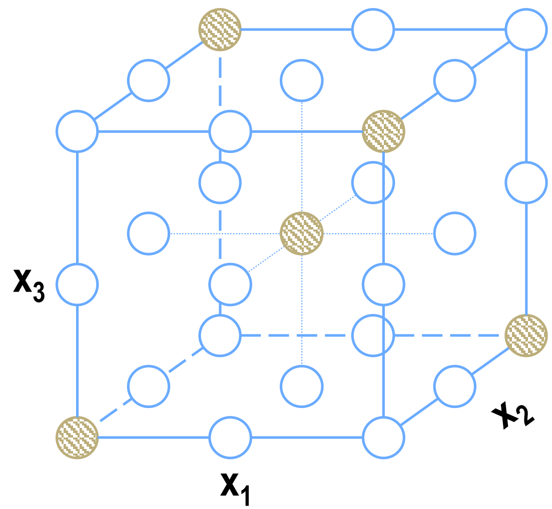
- A DOE is the specific collection of trials run to support a proposed model.
  - If proposed model is ***simple***, e.g. just main or 1st order effects ($\boldsymbol{x_1}$ **,** $\boldsymbol{x_2}$ **,** $\boldsymbol{x_3}$, etc.), the design is called a ***screening*** DOE
    - Goals include **rank factor importance** or find a “winner” quickly
    - Used with many (> 6?) factors at start of process characterization
  - If the proposed model is ***more complex***, e.g. the model is 2nd order so that it includes two-way interaction terms ($\boldsymbol{x_1x_2}$ **,** $\boldsymbol{x_1x_3}$**,** $\boldsymbol{x_2x_3}$, etc.) and in the case of continuous factors, squared terms ($\boldsymbol{x_1^2}$**,** $\boldsymbol{x_2^2}$**,** $\boldsymbol{x_3^2}$ **,** etc.), the design is called a ***response-surface*** DOE
    - Goal is generally to develop a **predictive model** of the process
    - Used with a few (< 6?) factors after a screening DOE

Response Surface DOE in a Nutshell

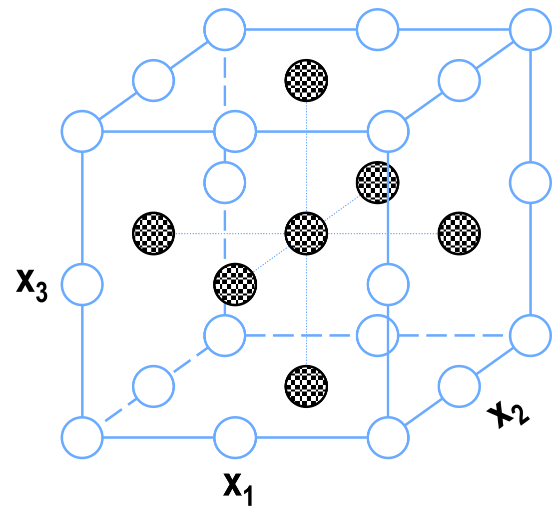
Block 1



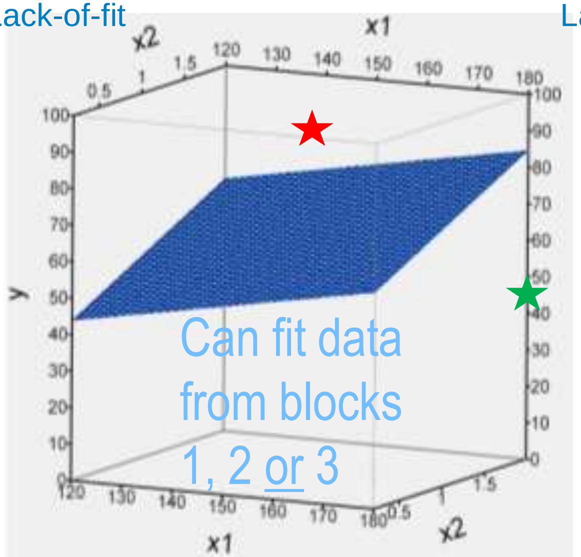
Block 2



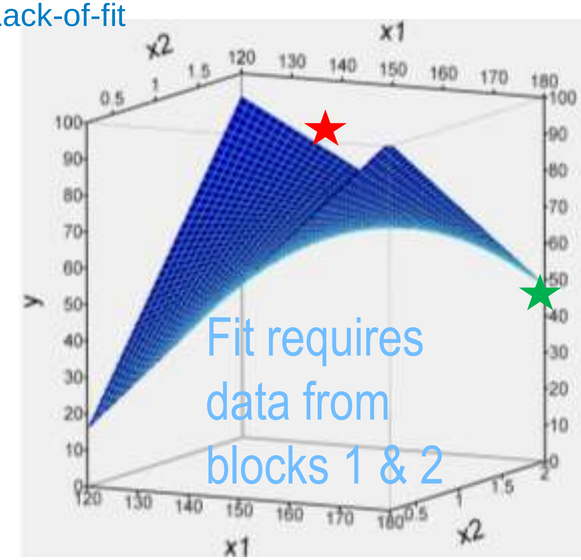
Block 3

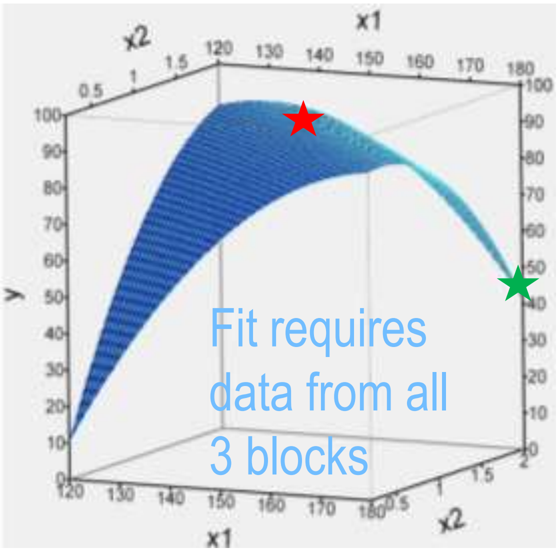


Lack-of-fit

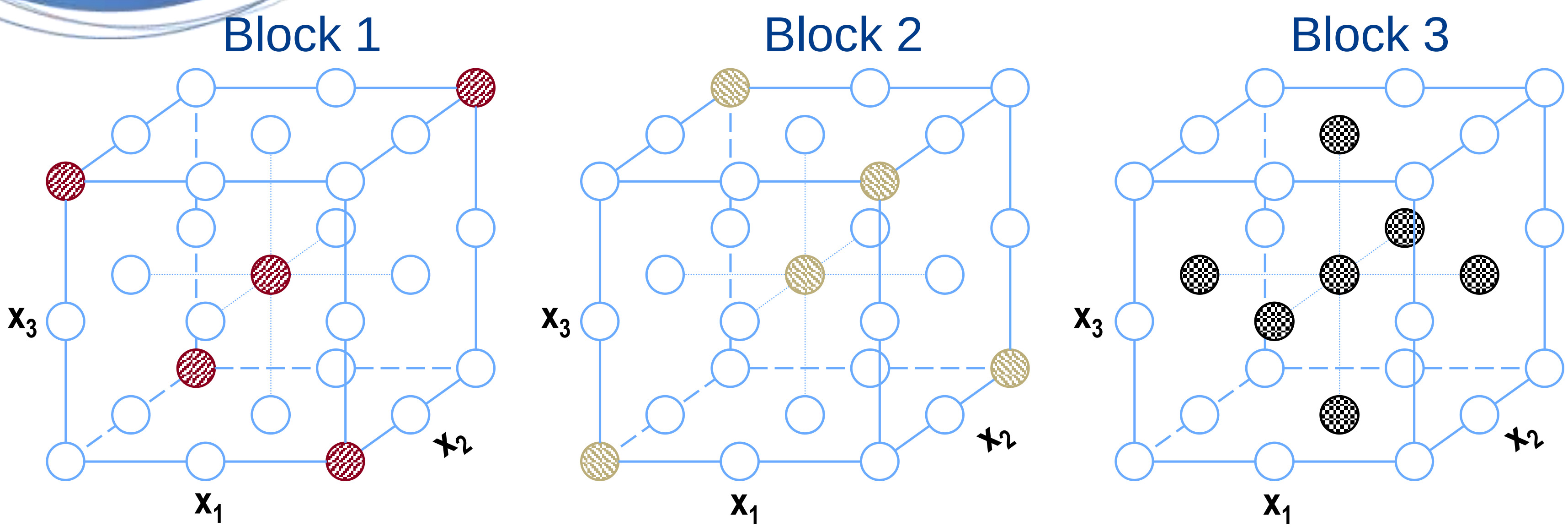


Lack-of-fit





Expensive Experimentation? Sequential DOE is Often Used



Block 1

$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$

Run this block 1st to:

(i) estimate the main effects*
(ii) use center point to check for curvature.

*May be all that are needed with appropriate physics-based scaling

Also called non-linear modeling

Block 2

$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$
$+ a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$

Run this block 2nd to:

(i) repeat main effects estimate,
(ii) check if process has shifted
(iii) add interaction effects to model if needed.

Block 3

$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$
$+ a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$
$+ a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2$

Run this block 3rd to:

(i) repeat main effects estimate,
(ii) check if process has shifted
(iii) add curvature effects to model if needed.

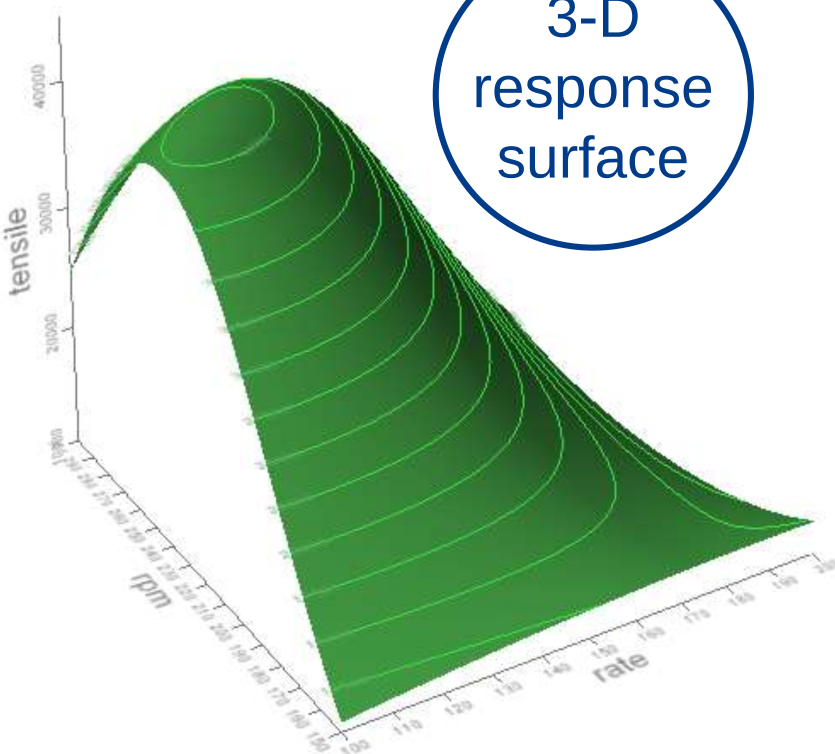
Why is Using DOE Important?

- *"One thing we have known for many months is that the spigot of defense funding opened by 9/11 is closing."*
- *"In the past, modernization programs have sought a 99 percent solution over a period of years, rather than a 75 percent solution over a period of weeks or months."*
  - Two quotes from the January 27, 2009 submitted statement of Secretary of Defense Robert M. Gates to the Senate Armed Services Committee.

- DOE is one of the more powerful tools we can use to efficiently accomplish our goals.
  - DOE yields the maximum information from the fewest experiments.
  - DOE often yields an 80% solution in less than 20% of the work.

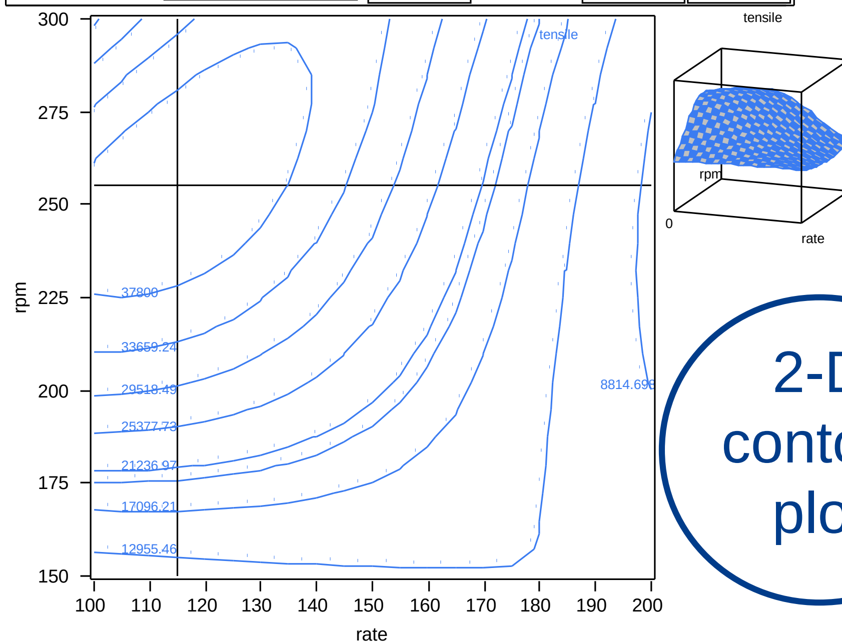
Response Surface & Contour Plot
(four control variables)

3-D response surface



Horiz	Vert	Factor	Current X
○	○	t4	320
⦿	○	rate	115
○	⦿	rpm	255
○	○	viscosity	80

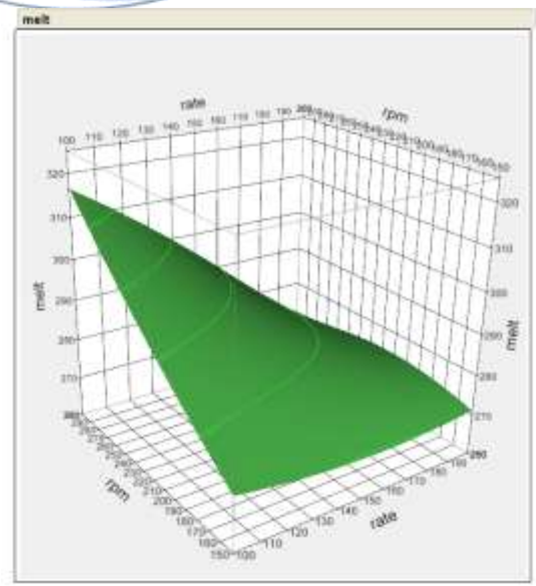
Response	Contour	Current Y	Lo Limit	Hi Limit
melt	250	305.35337	.	.
tensile	20000	41081.766	.	.

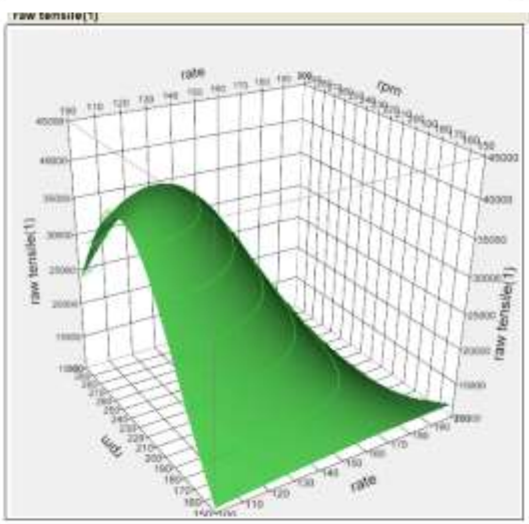


2-D contour plot

Response Surfaces & Contour Plots

(four control variables & two responses)

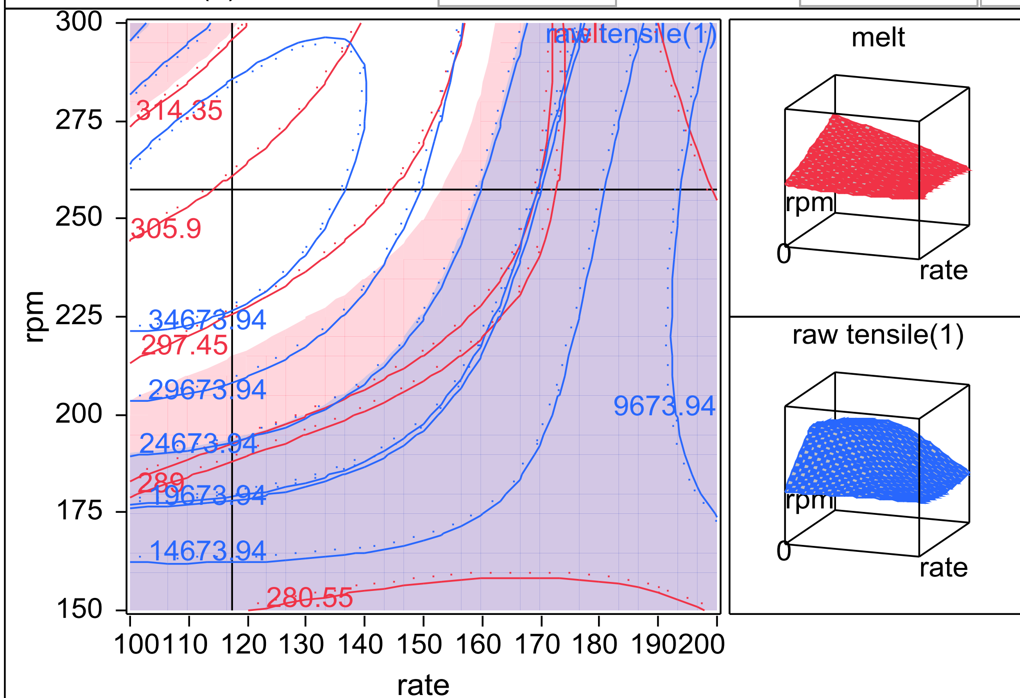




Contour Profiler

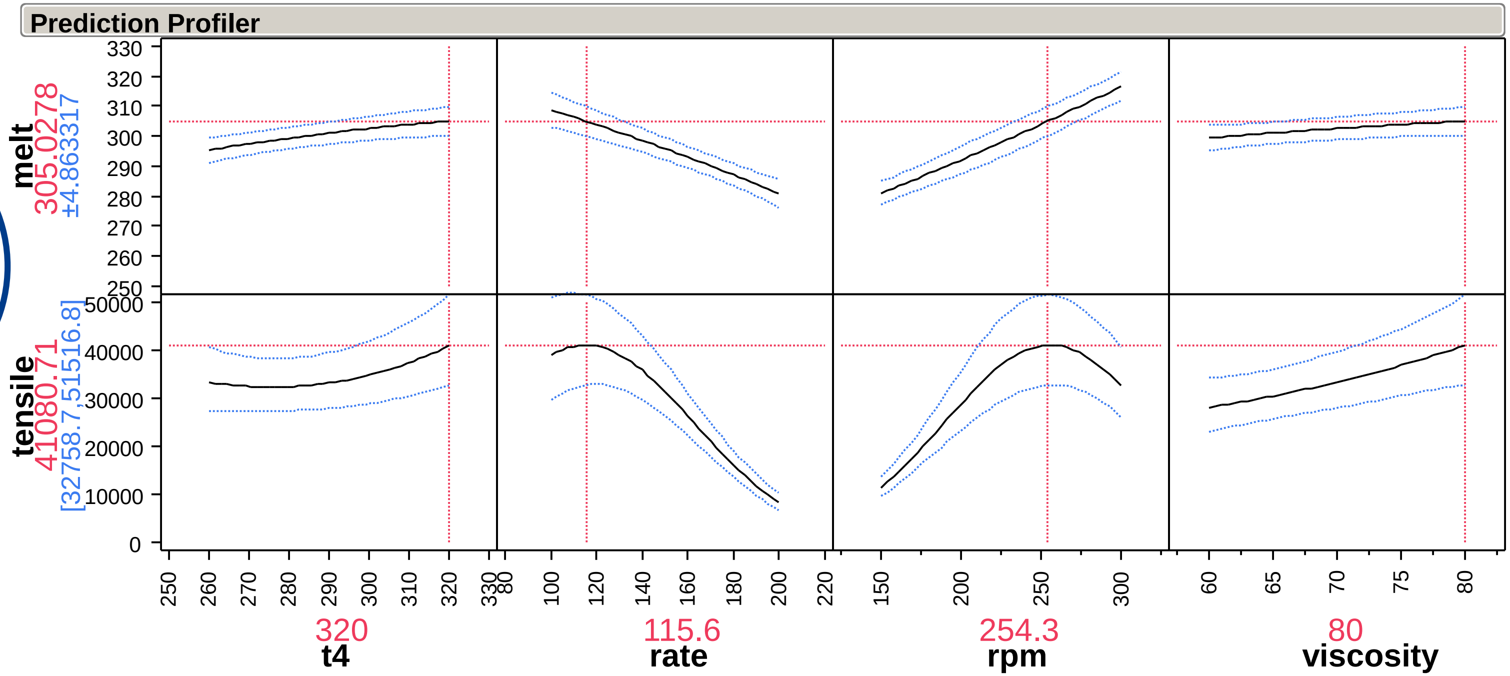
HorizVert	Factor	Current X
	t4	320
	rate	117.29498
	rpm	257.64505
	viscosity	80

Response	Contour	Current Y	Lo Limit	Hi Limit
melt	290	305.06654	295	315
raw tensile(1)	20000	38127.616	25000	40000

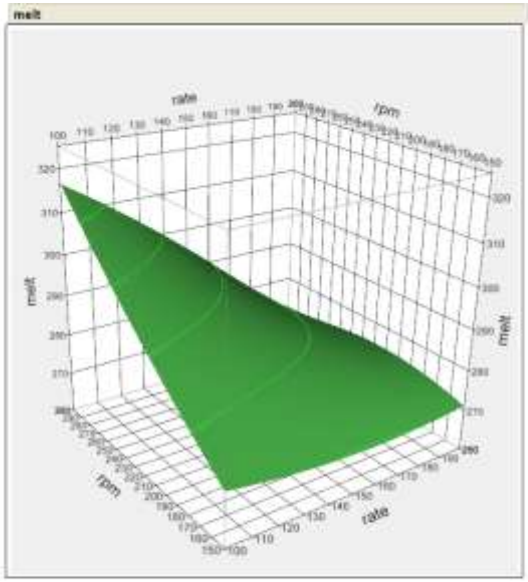


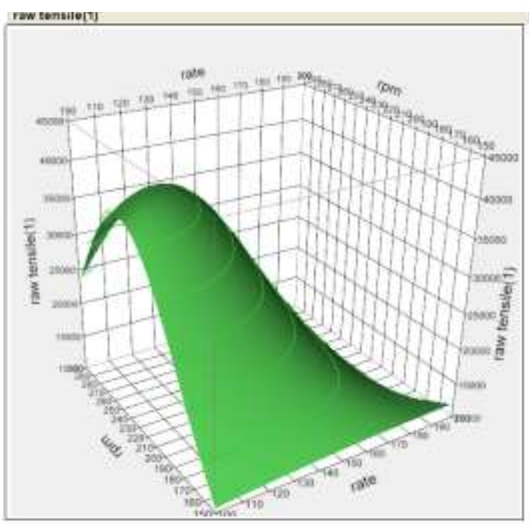
1-D Prediction Profiles are a Way to View Higher Dimensionality as "Interactive Small Multiples" - Here 4 Controls & 2 Responses

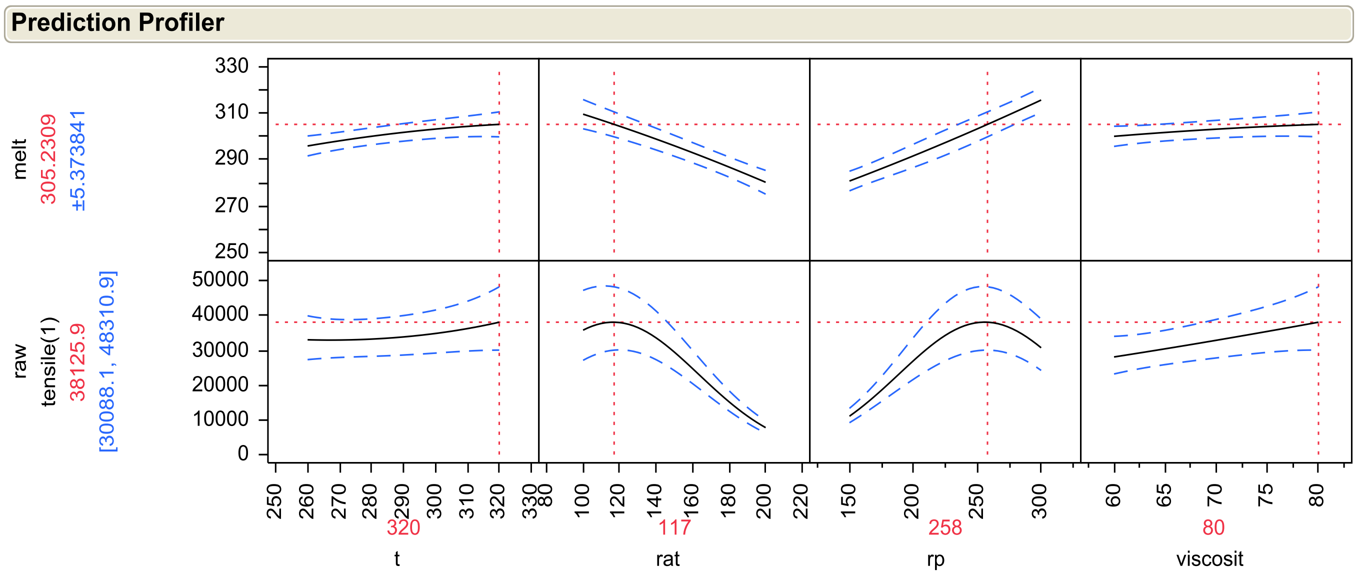
1-D profiler plots



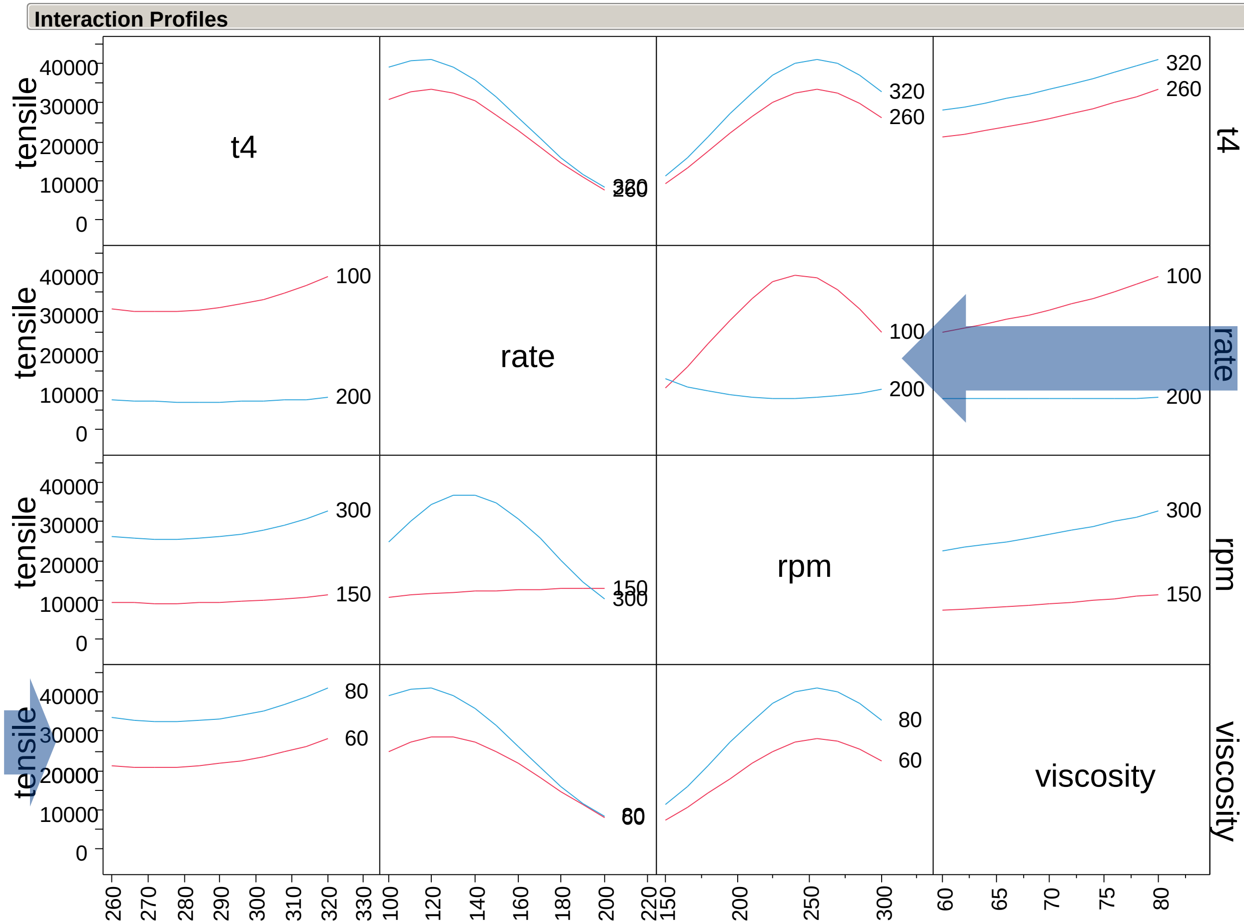
1-D Prediction Profiles are a Way to View Higher Dimensionality as "Interactive Small Multiples" - Here 4 Controls & 2 Responses







Interaction Profiles are Another Way to View Higher Dimensionality - Here 4 Controls and 1 Response



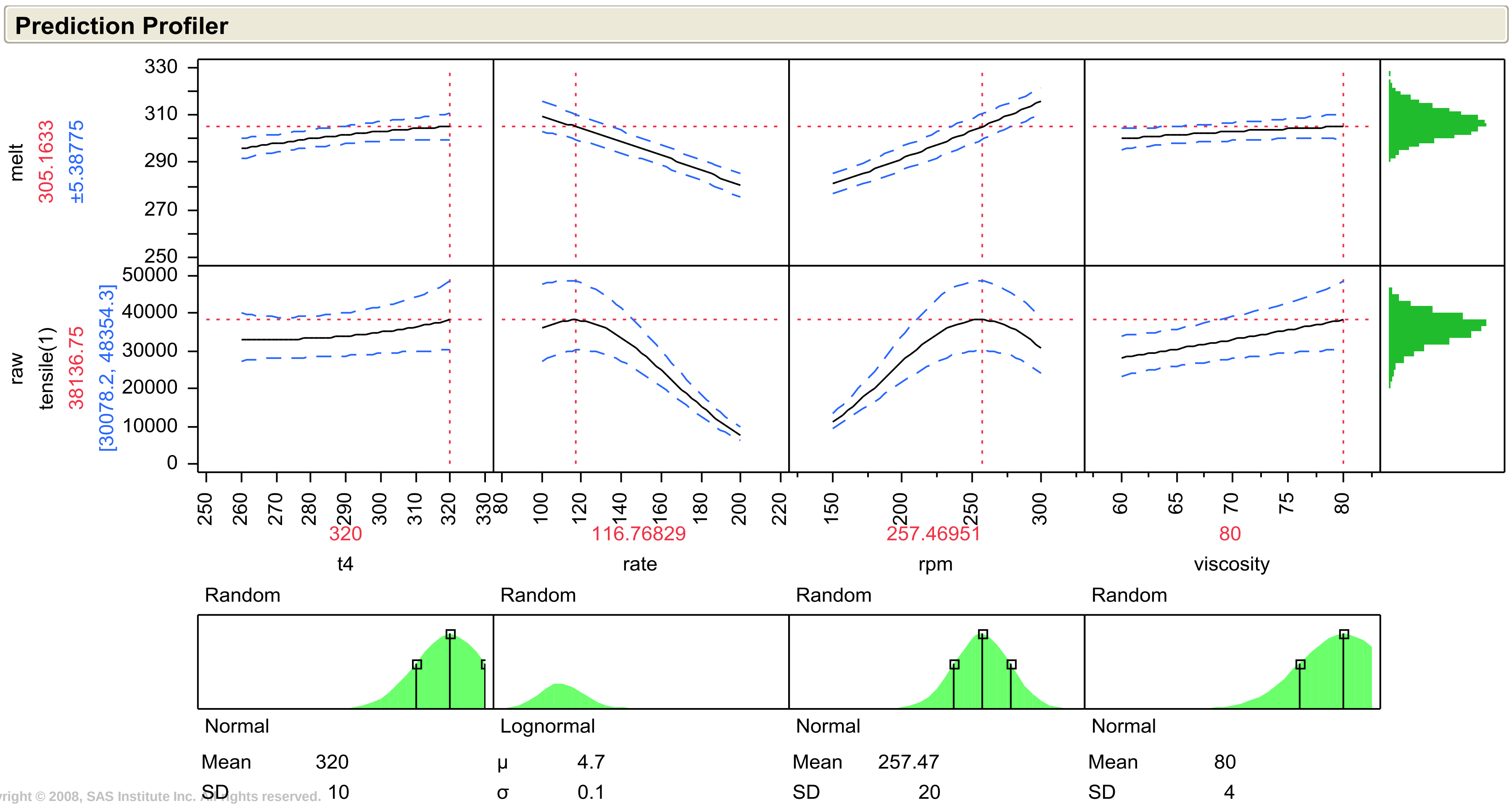
1-D plots at high & low of other factors

NOT Parallel indicates interaction

Parallel indicates NO interaction

Assess Uncertainty in Surrogate Model Predictions Even for a Non-Stochastic Simulation with No Replications

For non-stochastic simulations for which a surrogate model has been created, Monte Carlo simulations can be run using assumed distributions for inputs to better assess transmitted variation about the model point estimate.

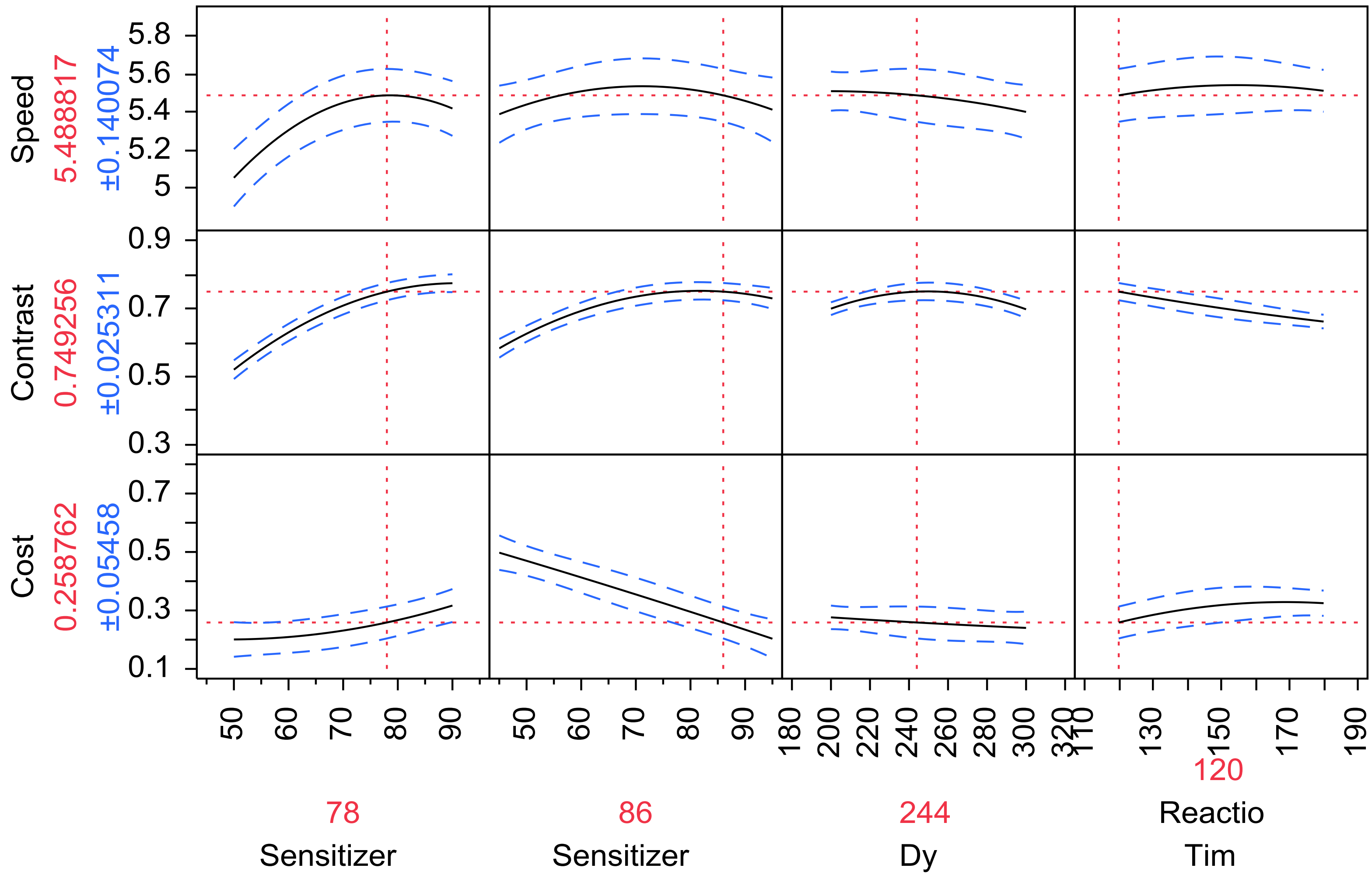


The Power of Predictive Modeling

- System with 2 primary MoPs and Cost
- Management wants to lower cost but maintain performance
- Multi-response and multi-factor process characterized using a JMP Custom DOE that supports interactive optimization – the trading off of performance and cost
- Management provided with visually interactive process knowledge that makes their decision making easier

Multiple Response Optimization

3 responses and 4 control factors



Leading Design of Experiments Texts

Box, G. E. P., Hunter, W. G., and Hunter, J. S. (2005), *Statistics for Experimenters*, 2nd ed., Wiley, New York
- The classic 1978 text recently revised

Wu, C. F. J. and Hamada, M. (2009), *Experiments, Planning, Analysis and Parameter Design Optimization*, , 2nd ed., Wiley, New York
- Both classic DOE approaches and orthogonal arrays & orthogonal main effects plans

Montgomery, D. C. (2009), *Design and Analysis of Experiments*, 7th ed., Wiley, New York
- **Popular text, solution book available, examples illustrated with DOE software.**
- **8th Edition includes newly developed screening approaches**
- **All problems worked in JMP software due out this year**

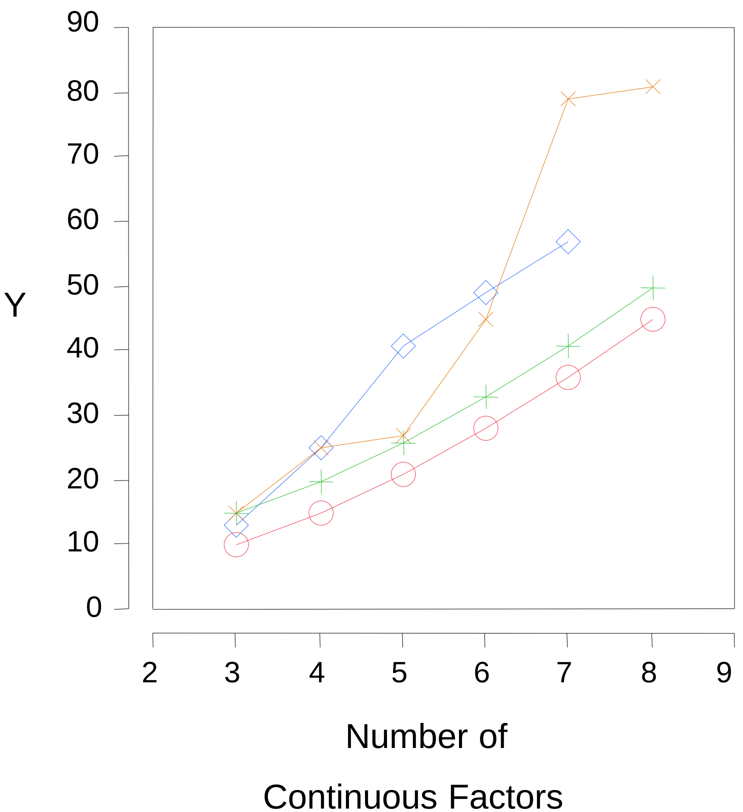
Texts Specifically on DOE for Computer Experiments:

Kleijnen, J. P. C. (2008), *DASE: design and analysis of simulation experiments*. Springer, New York.

Santner, T. J., Williams, B. J., and Notz, W. I. (2003), *The Design and Analysis of Computer Experiments*, Springer, New York

Fang, K. T., Li, R. Z., and Sudjianto, A. (2005), *Design and Modeling for Computer Experiments*, Chapman & Hall/CRC Press, New York

Unique Trials for 3 Response-Surface Designs and # Quadratic Model Terms vs. # Continuous Factors



If generally running 3, 4 or 5-factor fractional-factorial designs…

1. How many interactions are you NOTinvestigating?
2. How many more trials needed to fit curvature?
3. Consider two stages – Definitive Screening + Augmentation

Create a More Complex "Real-World" DOE

How many folks have any of these issues?

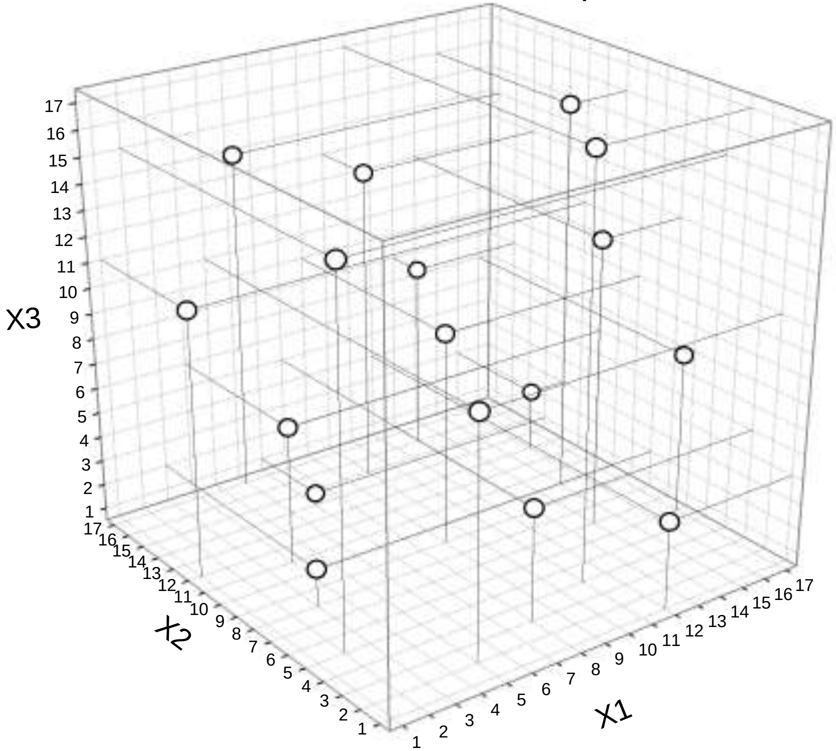
- Work with these different kinds of control variables/factors:
  - **Continuous/quantitative?** (Finely adjustable like *temperature*, *speed*, *force*)
  - **Categorical/qualitative?** (Comes in types, like material = *rubber*, *polycarbonate*, *steel* with mixed # of levels; 3 chemical agents, 4 decontaminants, 8 coupon materials…)
  - **Mixture/formulation?** (Blend different amounts of *ingredients* and the process performance is dependent on the *proportions* more than on the amounts)
  - **Blocking?** (e.g. "lots" of the same raw materials, multiple "same" machines, samples get processed in "groups" – like "eight in a tray," run tests over multiple days – i.e. variables for which there *shouldn't* be a causal effect
- Work with **combinations of these four kinds** of variables?
- Certain **combinations cannot be run**? (too costly, unsafe, breaks the process)
- Certain factors are **hard-to-change** (temperature takes a day to stabilize)
- Would like to **add onto existing trials**? (really expensive/time consuming to run)
- Characterize process or **run experiments using computer simulations**? (war gaming, agent-based, discrete event, computational fluid dynamics (CFD))
- Measure response **data in vicinity of physical limits**? (counts, hardness, resistivity can't fall below zero, or percentage yield or killed can't exceed 100%)

Two Classes of Designs for Two Types of Surrogate Modeling of Simulations

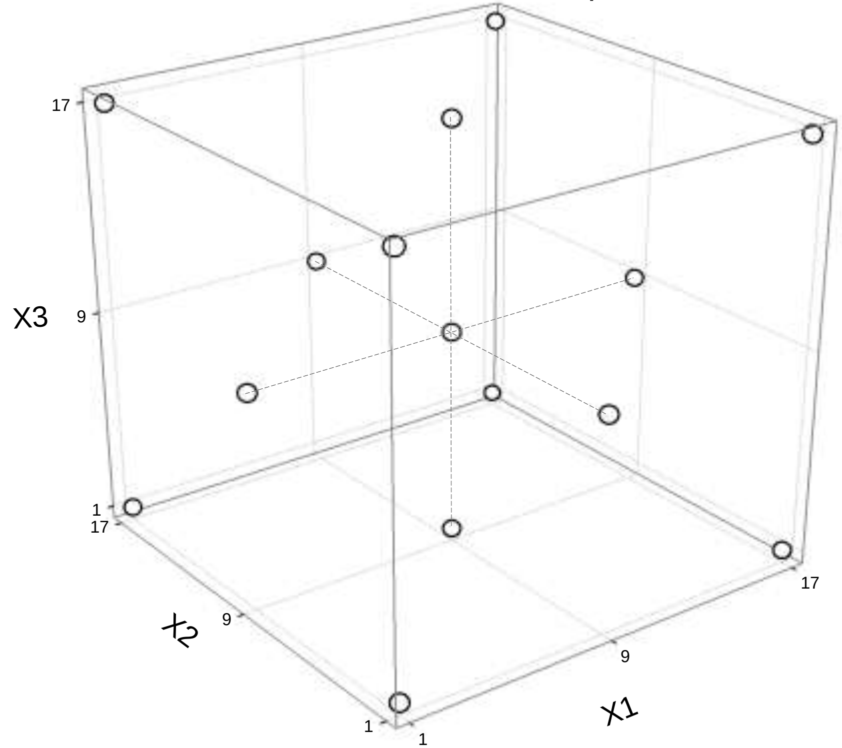
- “Traditional factorial/response surface” designs for polynomial modeling with categorical (qualitative) and continuous (quantitative) variables
  - Designs can be sequentially constructed to support increasingly complex models
  - Example featured here reanalyzes a simulation case matrix in which all combinations of 6 variable settings were originally run- a total of 648 = 6 X 3 X 3 X 3 X 2 X 2
  - References on Resolution V, Fractional-Factorial Designs for many (40+) factors
    - Mee, R. W. (2004), **Efficient Two-Level Designs for Estimating Main Effects and Two-Factor Interactions**, *Journal of Quality Technology*, 36, 400-412.
    - Sanchez, S.M. and Sanchez, P.J. (2005), **Very Large Fractional Factorial and Central Composite Designs**, *ACM Transactions on Modeling and Computer Simulation*, Vol. 15, No. 4, October 2005, Pages 362–377.
    - Xu, H. (2009), **Algorithmic Construction of Efficient Fractional Factorial Designs with Large Run Sizes**, *Technometrics*, (in press) http://www.stat.ucla.edu/~hqxu/pub/ffd2r3.pdf
- “Space-filling” designs primarily for use with continuous variables AND non-stochastic/deterministic responses
  - These designs can support “Gaussian Process” or “Kriging” spatial regression analysis – an interpolation technique, as well as linear regression – an approximation method

How are Space-Filling Designs Different from Traditional Designs?

Space-Filling Design
for 3 Variables with 17 Unique Trials



Response-Surface Design
for 3-Variables with 15 Unique Trials



Rather than emphasizing high leverage trials ("corners") for a simple polynomial model, space-filling designs "spread" their trials more uniformly through the space to better capture the local complexities of the simulation model.

Traditional Designs for Polynomial Modeling

- I used to say "If a "textbook" fractional-factorial, orthogonal array or response-surface design is available, then use it." Now I say, "If Definitive Screening or Minimal Alias design is available, then use it."
- Textbooks and web site catalogs do not always contain designs for categorical variables with:
  - all combinations of mixed numbers of levels (e.g. 3, 4, 5, and 21)
  - large numbers of levels for variables (e.g. 5+)
- Algebraic (Orthogonal Array) and algorithmic (D-optimal) computer generated designs can often be used
  - Orthogonal Arrays are good at yielding analysis with un-confounded estimates of the "main effects" when variables have many different levels
  - D-optimal designs are good for adding on the fewest additional trials to support higher order "interaction" terms in the model

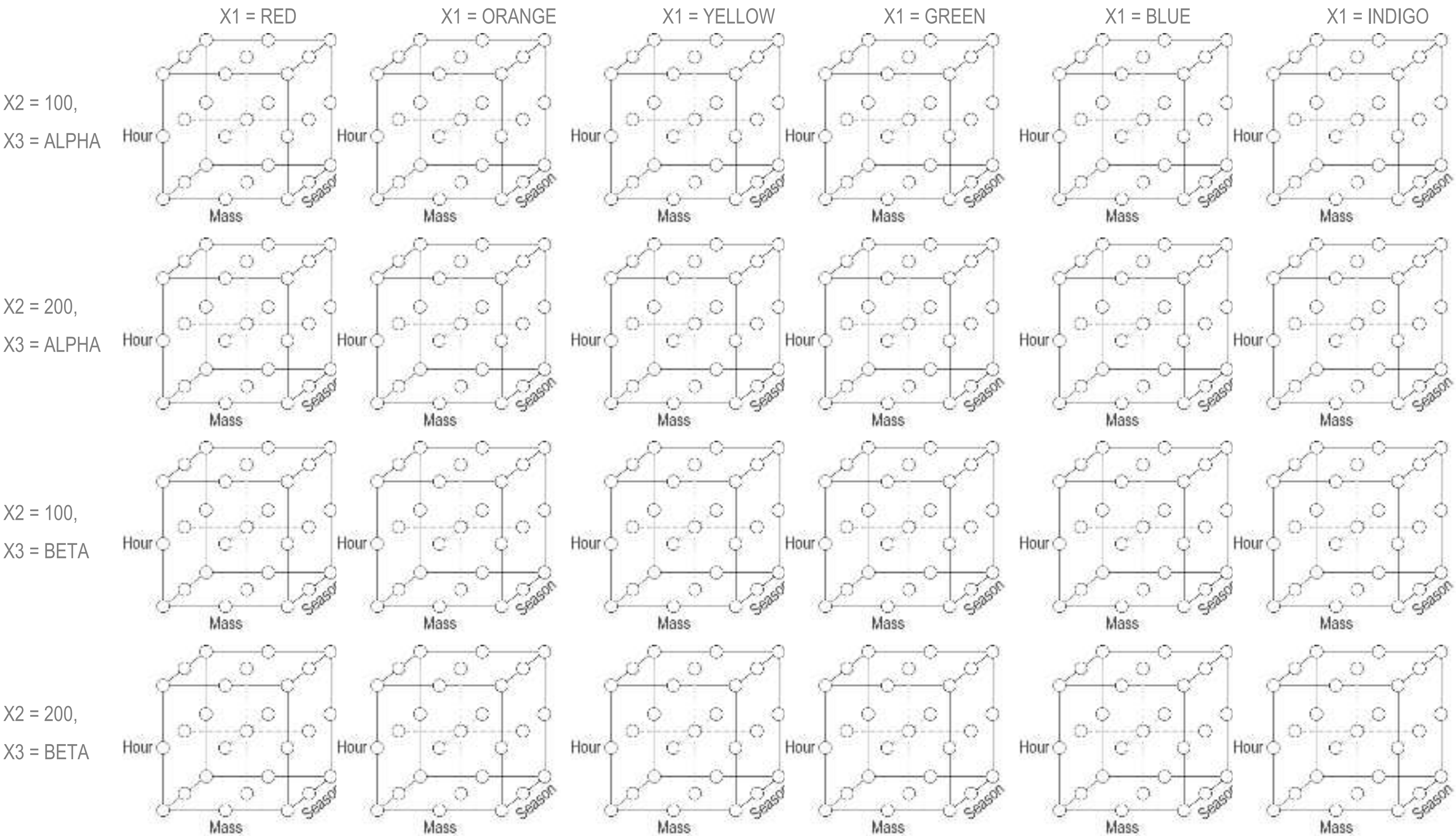
Sequential Designs

- Simulation experiments – Sequential designs are easily employed because “restricted randomization” is not an issue
  - Many simulations are deterministic
  - Even if stochastic (random), correlation with unknown factors is not possible
  - All factors are generally just as easy to change
  - Can still inexpensively add a blocking variable to test if “the code has been changed!”
- Real experiments – The issue of “restricted randomization” does arise making sequential experimentation a bit more complicated – but still possible to employ
  - Groups of trials run at different (even widely spaced) periods of time
    - Addressed using a *blocking* factor
  - Sometimes there are factors that are harder to change than others, e.g. *Oven Temperature*
    - Addressed using *split-plot* designs

Case Matrix as Used in Study of the Observed Response "Probability of Casualty" (PCAS)

Variable	# Levels	Levels
Agent Codes (X1)	6	A, N, T, H, R, Y (categorical)
Season	3	Winter, Summer, Spring/Fall (categorical)
Time of Attack (Hour)	3	0500, 1200, 2200 Local Time (continuous)
No. of TBMs & Spread Radius (X2)	2	1 TBM & 1 m, 2 TBMs & 1000 m (categorical)
Mass (relative)	3	1.00, 1.57, 2.00 (continuous)
Height of Burst (X3)	2	0, 10 m (continuous)
Total Cases	648	

All 648 Possible Combinations of Settings for 6 Variables (6 X 2 X 2 X 3 X 3 X 3)

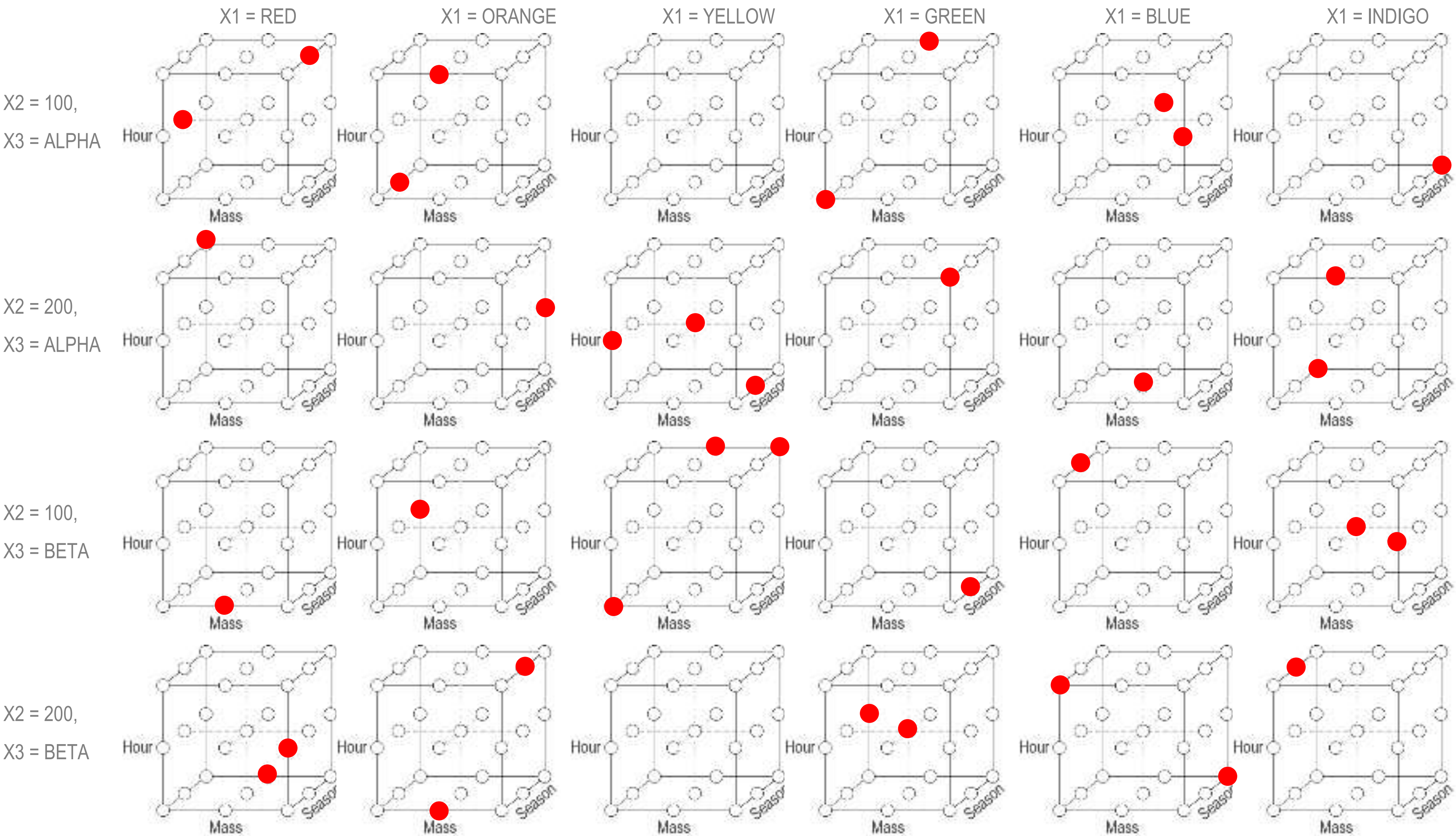


Four Stage Design Sequence

Stage 1	Stage 2	Stage 3	Stage 4
36 Total Simulations	**108** Total Simulations	**324** Total Simulations	ALL **648** Simulations
Design 1, 36 trials	Design 1, 36 trials	Design 1, 36 trials	Design 1, 36 trials
	Design 2, 72 trials	Design 2, 72 trials	Design 2, 72 trials
		Design 3, 216 trials	Design 3, 216 trials
Main effects only for ALL variables + some 2-way interactions	Stage 1 effects plus all 2-way interactions + some 3-way interactions	Stage 2 effects plus all 3-way interactions	Stage 3 effects plus ALL remaining 4-way, 5-way and 6-way interactions
5.6% of 648	16.7% of 648	50% of 648	Design 4, 324 trials NOTE: Length of this green box should be longer than shown

324 trials in Design 4 used as checkpoints for Designs 1, 2 & 3 →

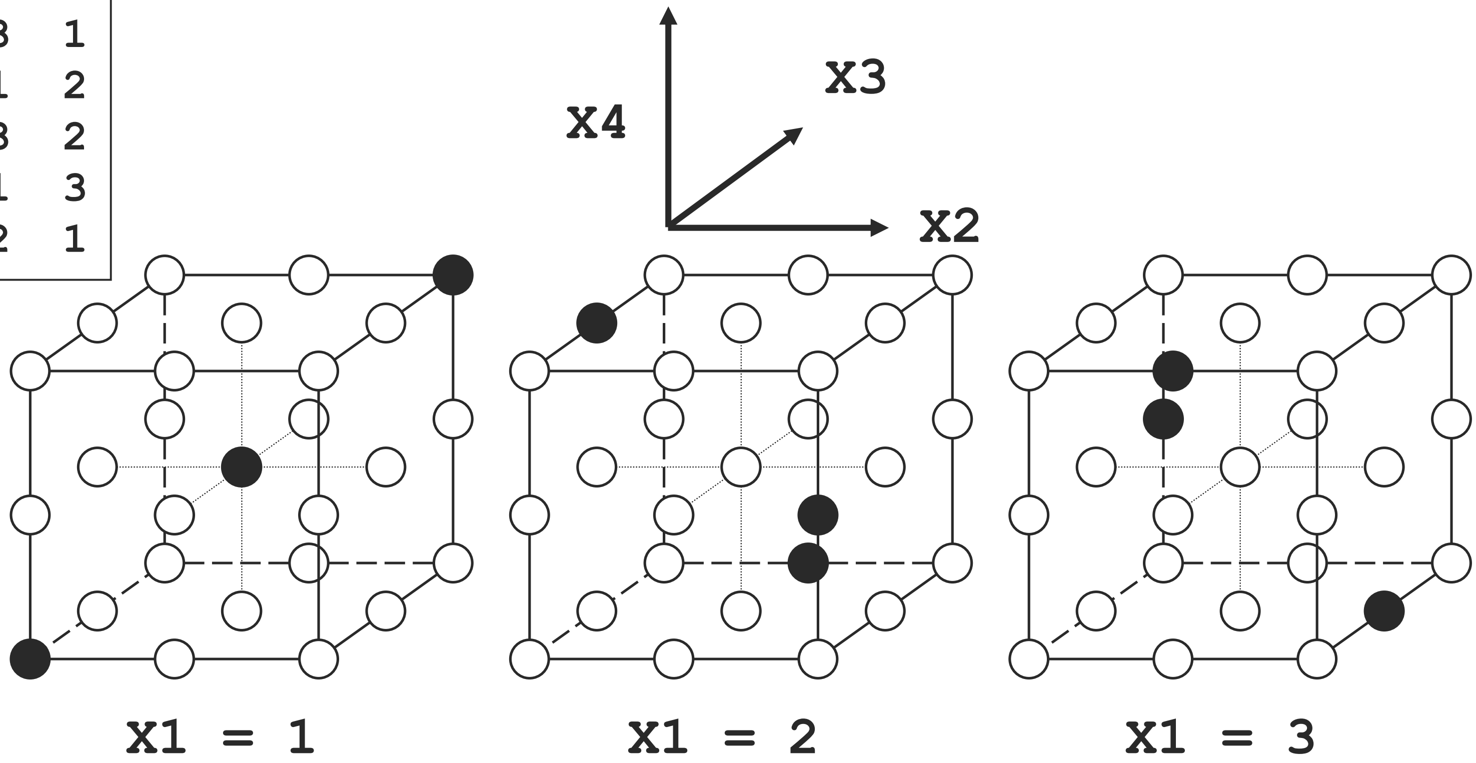
36 of All 648 Possible Combinations of Settings for 6 Variables (6 X 2 X 2 X 3 X 3 X 3)



Red Dots Mark the 36 Trials (an Orthogonal Array) Analyzed for Stage 1

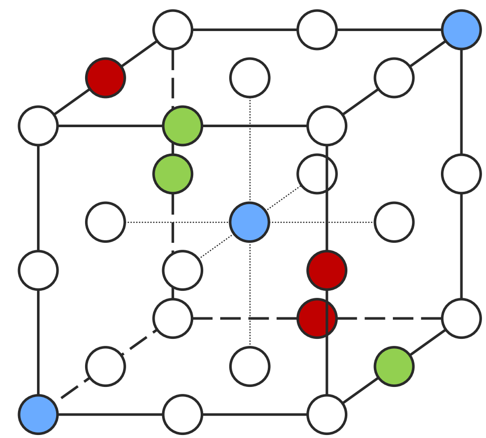
Locations of Trials for a 4-variable, 9-trial Orthogonal Array Design

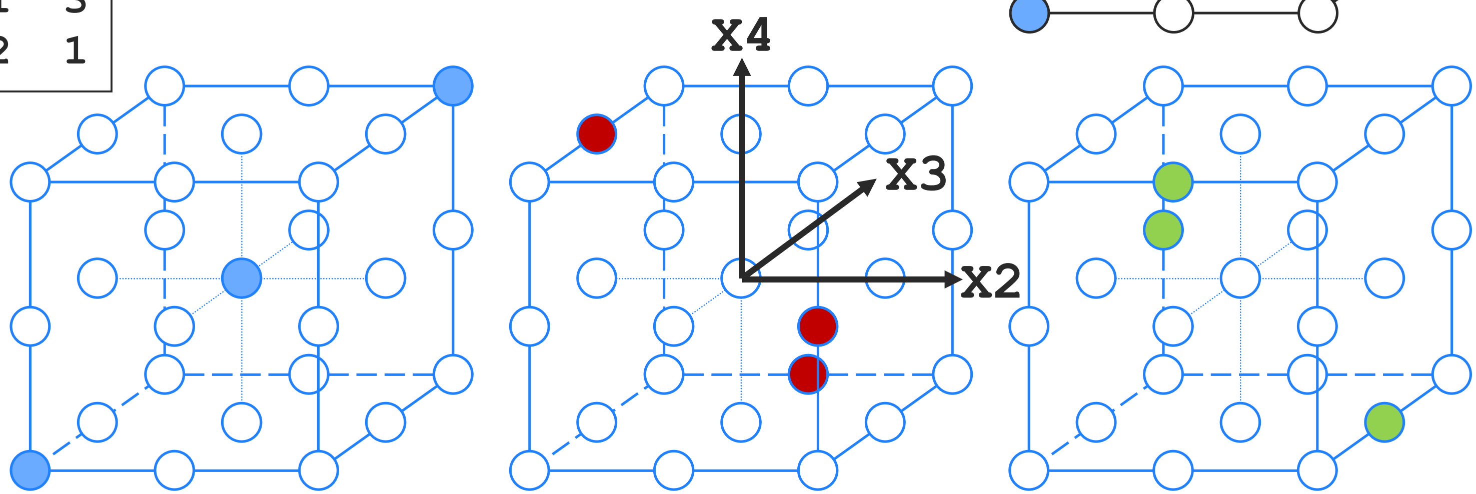
x1	x2	x3	x4
1	1	1	1
1	2	2	2
1	3	3	3
2	1	2	3
2	2	3	1
2	3	1	2
3	1	3	2
3	2	1	3
3	3	2	1



x1	x2	x3	x4
1	1	1	1
1	2	2	2
1	3	3	3
2	1	2	3
2	2	3	1
2	3	1	2
3	1	3	2
3	2	1	3
3	3	2	1

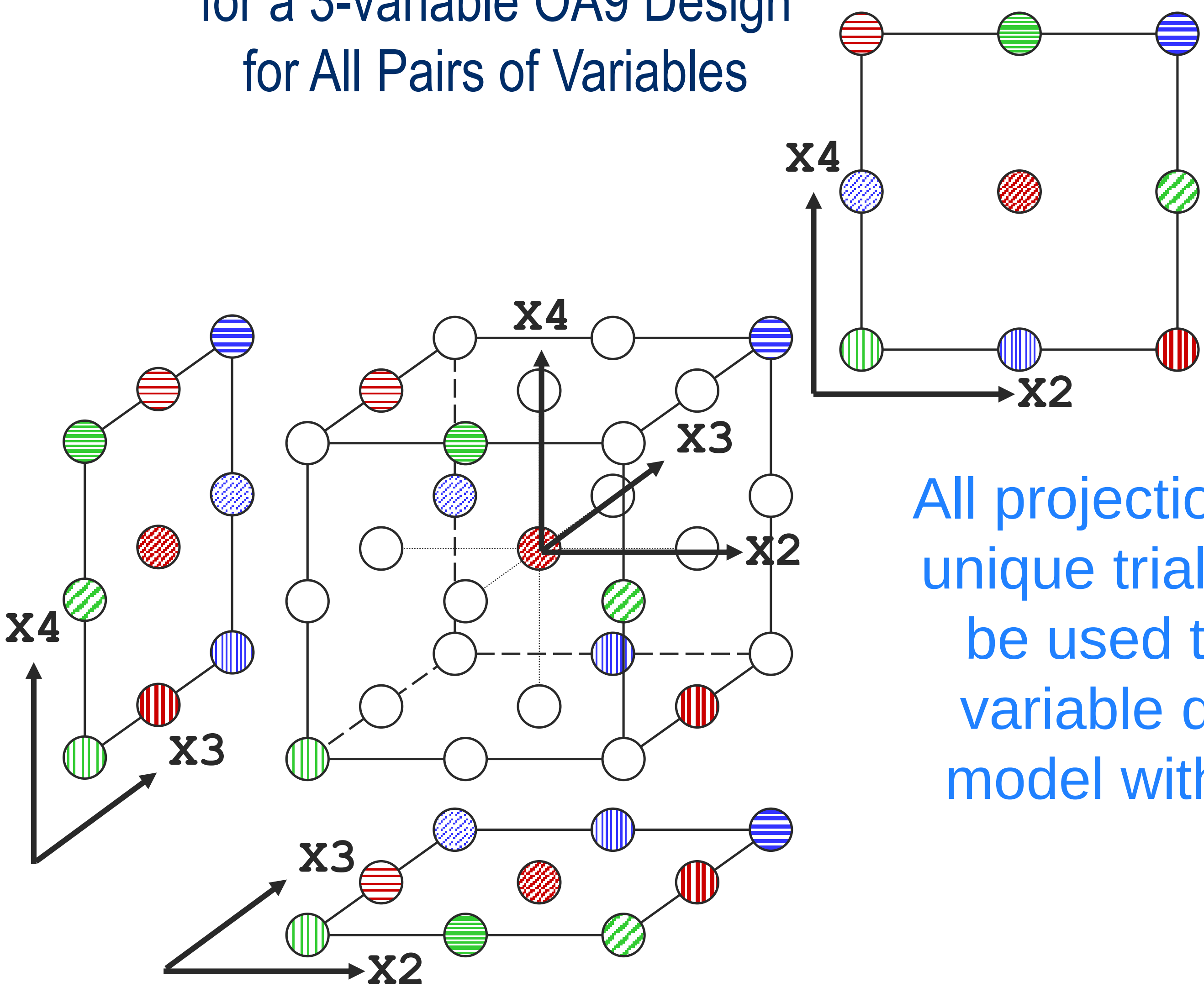
Delete **x1** and View Locations of Trials for a 3-Variable OA9 Design





x1 = 1   x1 = 2   x1 = 3

Projection of Trial Locations for a 3-variable OA9 Design for All Pairs of Variables



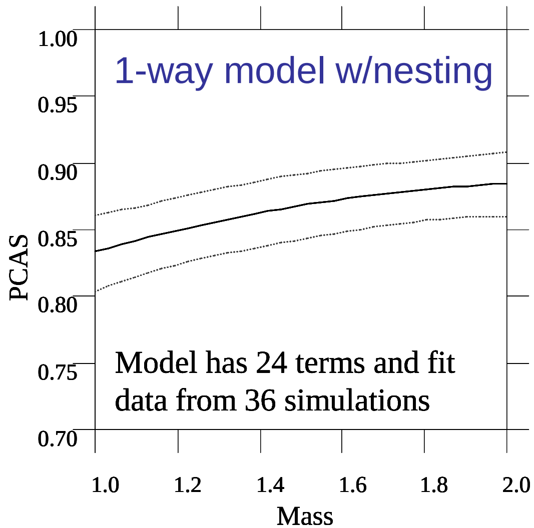
All projections have 9 unique trials that can be used to fit a 2-variable quadratic model with 6 terms

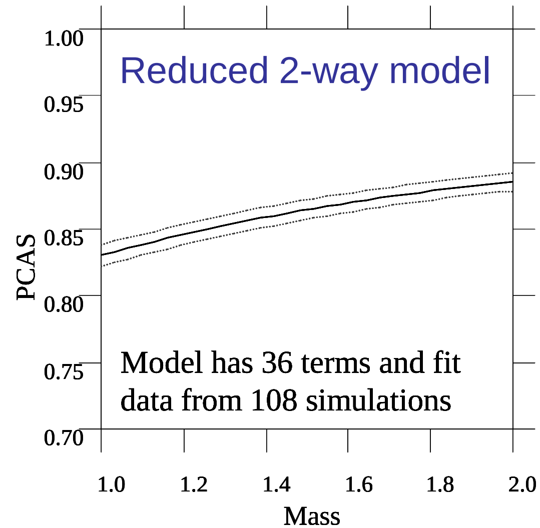
Can Get Designs from Different Sources

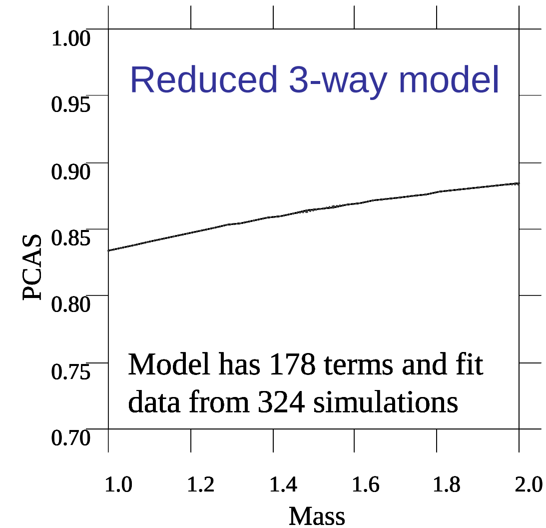
- Textbook
  - Limited number of catalogued solutions – experimenters frequently change their problem to match available designs
  - Variable settings are in coded units
- Web sites of designs
  - Greater number of catalogued solutions – but never all
  - Variable settings are in coded units
- Custom computer code
  - Can find solutions for previously un-catalogued cases
  - Variable settings are in coded units (-1, 0, 1)
- COTS Solution
  - Textbook and algorithmic code for generating custom designs
  - Variable settings in natural or laboratory units (120, 150, 180)

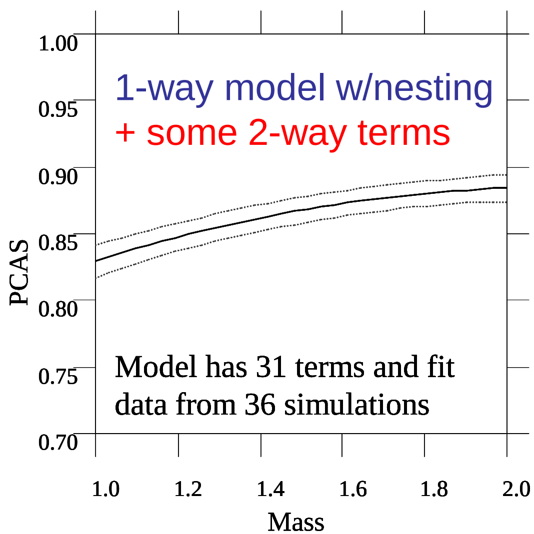
Predictions (w/95% Pred. Limits) of PCAS vs. Nested Mass and MunCnt_Spread for 1-way, reduced 2-way and reduced 3-way models

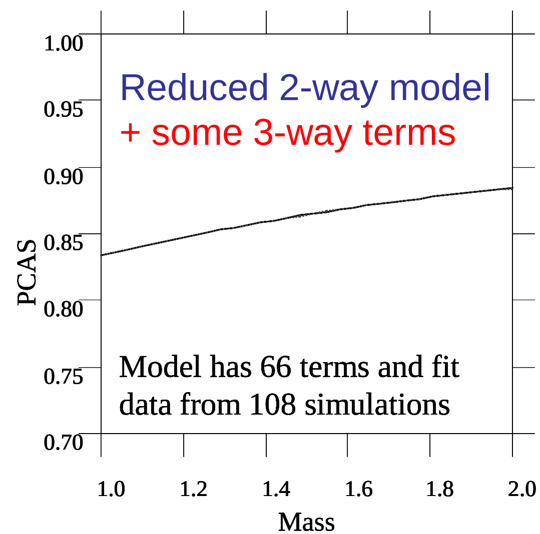
Predicted Probability of Casualty (PCAS) vs. Mass – with Mass Treated as a Continuous Variable – for 5 Different Models Fit to 3 Sets of Simulation Data











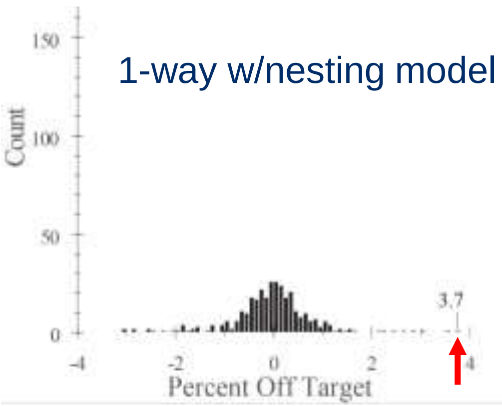
Five other variables were held constant at these settings:

Agent = R
Season = F
Time = 12
HOR = 0
#TBM &
Spread Radius = 1

Predicted Mass
95% Prediction Limits

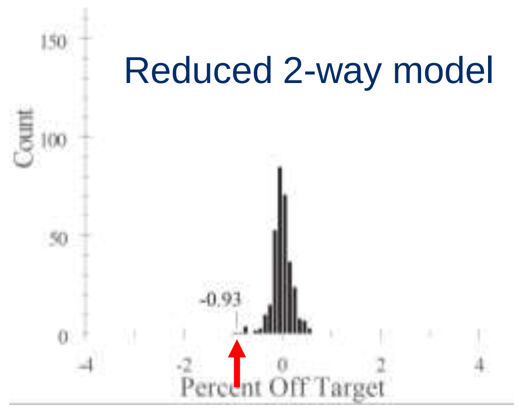
"Factor Sparsity" and "Effect Heredity" Used to Enhance Model Complexity

36 trials



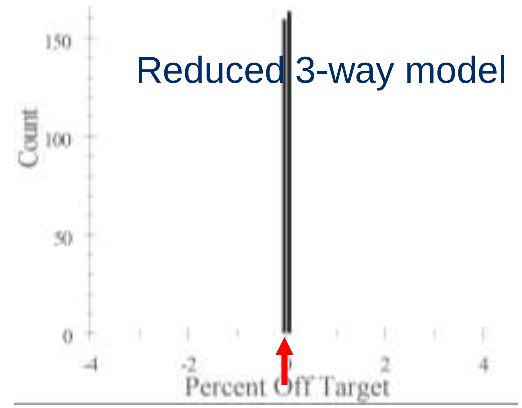
Worst Case = 3.7%
Half of Cases < 0.37%

108 trials

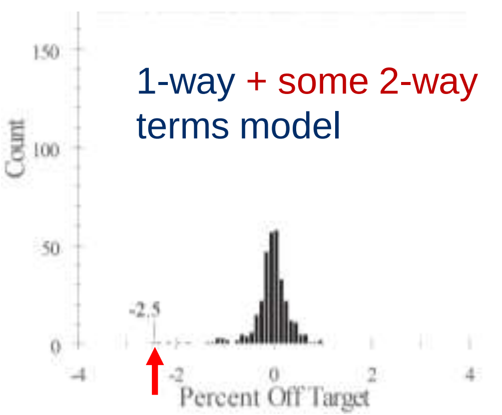


Worst Case = -0.93%
Half of Cases < 0.11%

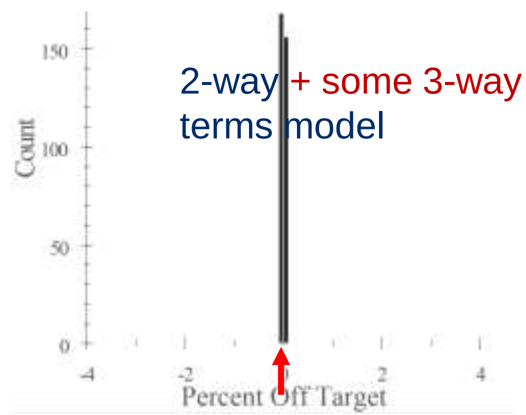
324 trials



Worst Case = -0.0081%
Half of Cases < 0.0007%



Worst Case = -2.5%
Half of Cases < 0.16%



Worst Case = -0.0251%
Half of Cases < 0.0010%

Factor Sparsity states only a few variables will be active in a factorial DOE

Effect Heredity states significant interactions will only occur if at least one parent is active

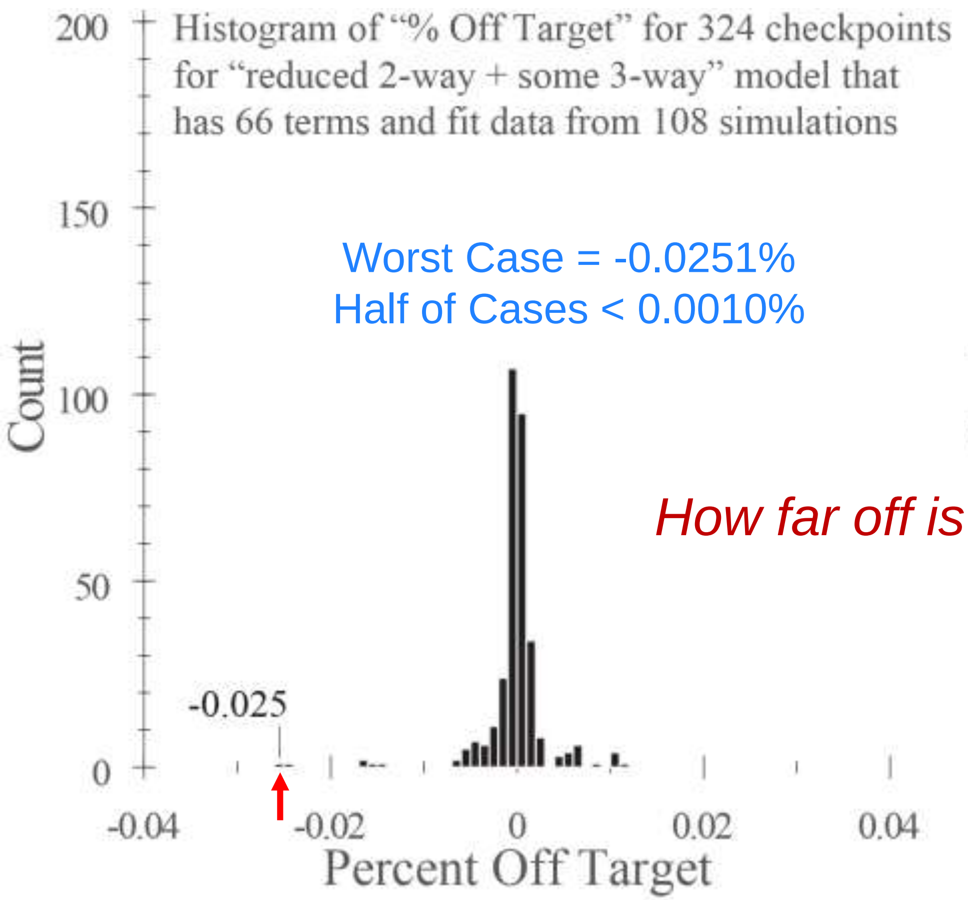
See Wu & Hamada, p. 112

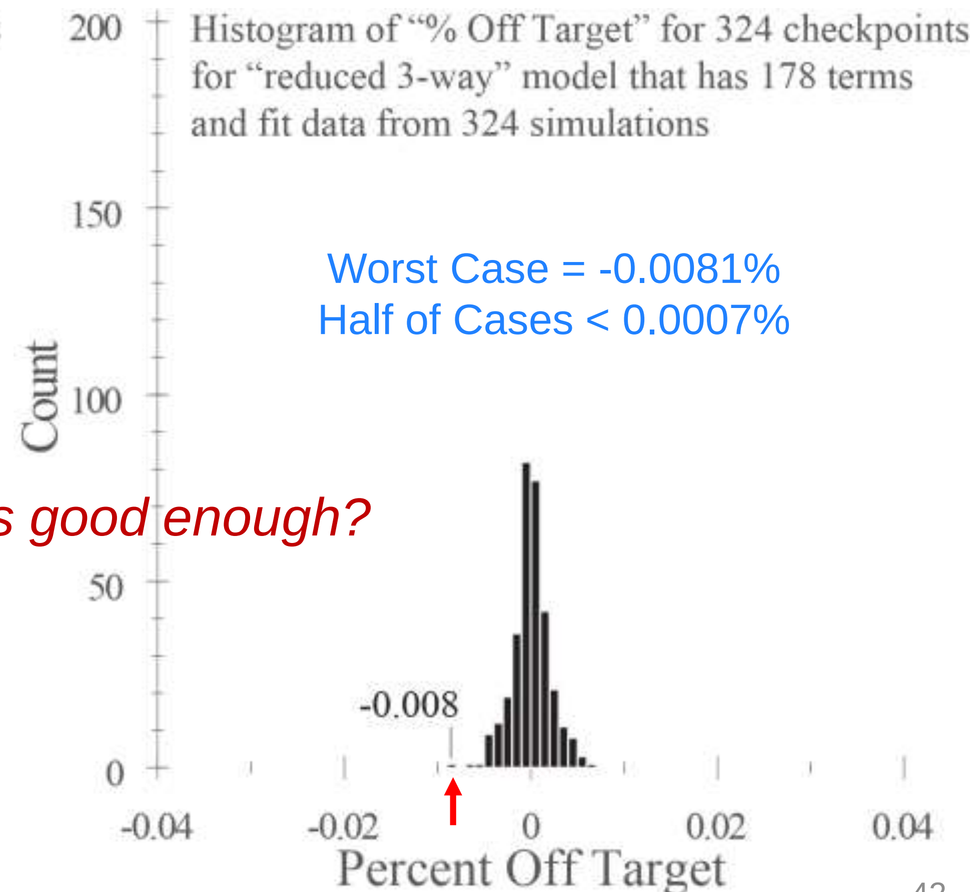
Oct. 1, 2007 visit by Profs. Wu & Joseph of GA Tech ISyE

Only a Fraction of All Possible Trials May be Required to Provide an Answer

108 trials **324 trials**

Higher Resolution (100X) Histograms of the "Percent Off Target" that Response Predictions Fell Relative to 324 Checkpoint Observations





How far off is good enough?

More Complexity - Better Fit
Fewer "Extraneous" Terms - Better Fit

Error Estimates for 8 Models Fit to Data Sets of 36, 108, & 324 Observations

Number of Trials	Model Used to Fit Data	Number of Model Terms	Residual SD (model error from data used in fit)	Cross-Validation RMS ("one-left out" error from data used in fit)	Checkpoint RMS (model error from 324 data values NOT used in fit)	Adjusted R-squared
36	1-way	14	0.043623	0.055802	0.037217	0.977
36	1-way w/nesting	24	0.026557	0.047269	0.035424	0.992
36	1-way w/nesting + some 2-way	31	0.008212	0.025188	0.016153	0.999
108	2-way	79	0.011197	0.022207	0.010772	0.998
108	reduced 2-way	36	0.008469	0.010933	0.008612	0.999
108	reduced 2-way + some 3-way	66	0.000045	0.000132	0.000179	1.000
324	3-way	242	0.000039	0.000078	0.000083	1.000
324	reduced 3-way	178	0.000037	0.000058	0.000064	1.000

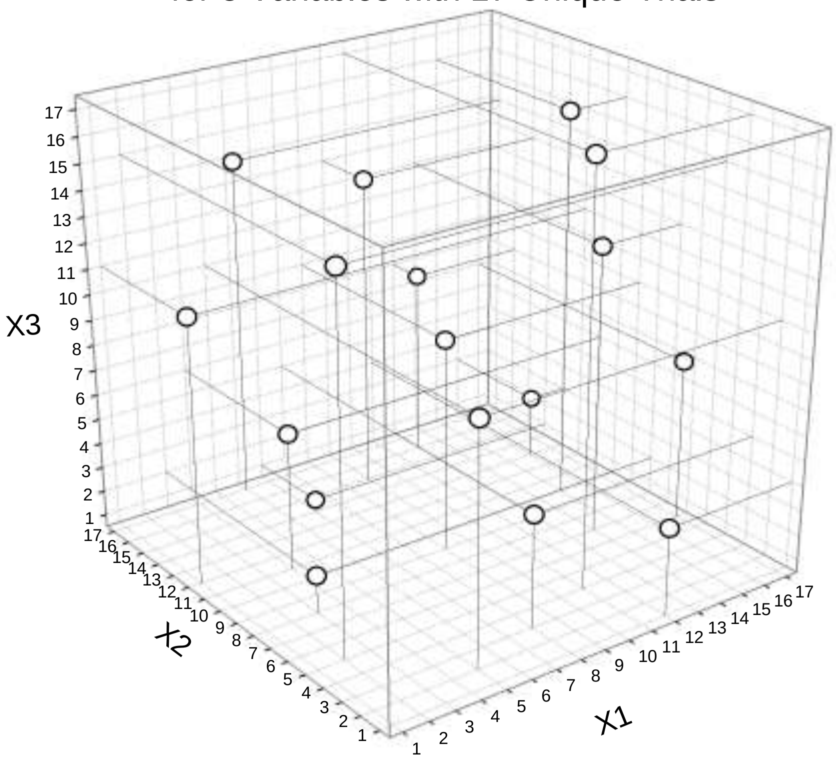
Higher Cross-Validation RMS May Be an Indicator of "Over Fitting"

Conclusions
Sequential Traditional Designs

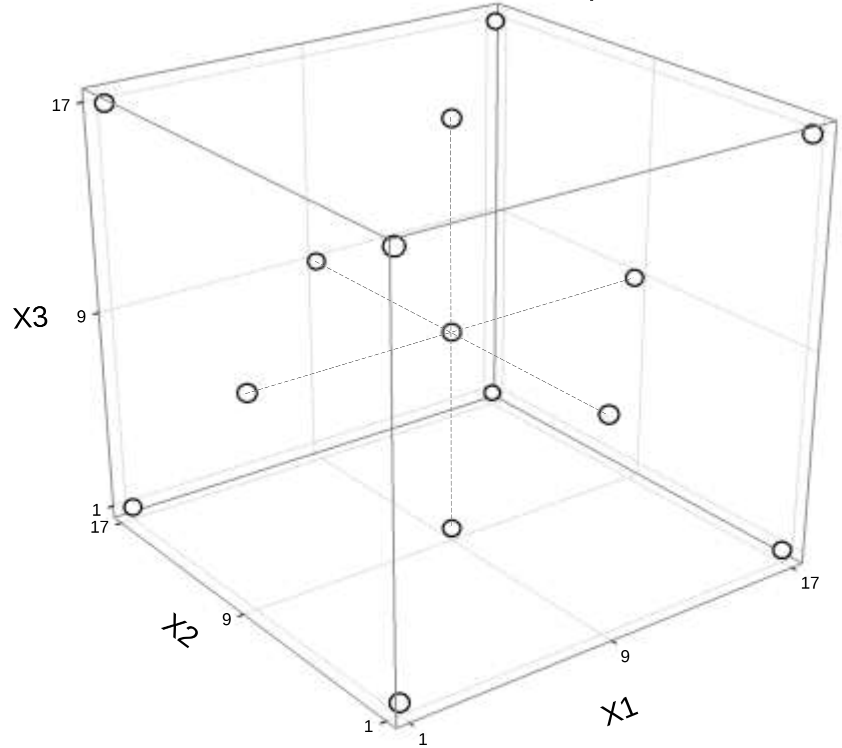
- Possible to get the 80% to 95% solution with less than 20% of the brute force running of all factor combinations
- Use of "factor sparsity" and "effect heredity" principles can help to get more information than the design was originally built to support
- Next stage trials can first be used as checkpoints for previous stages
- With improved efficiency over running all combinations, more factors can be studied with the same resources

How are Space-Filling Designs Different from Traditional Designs?

Space-Filling Design
for 3 Variables with 17 Unique Trials



Response-Surface Design
for 3-Variables with 15 Unique Trials



Rather than emphasizing high leverage trials ("corners") for a simple polynomial model, space-filling designs "spread" their trials more uniformly through the space to better capture the local complexities of the simulation model.

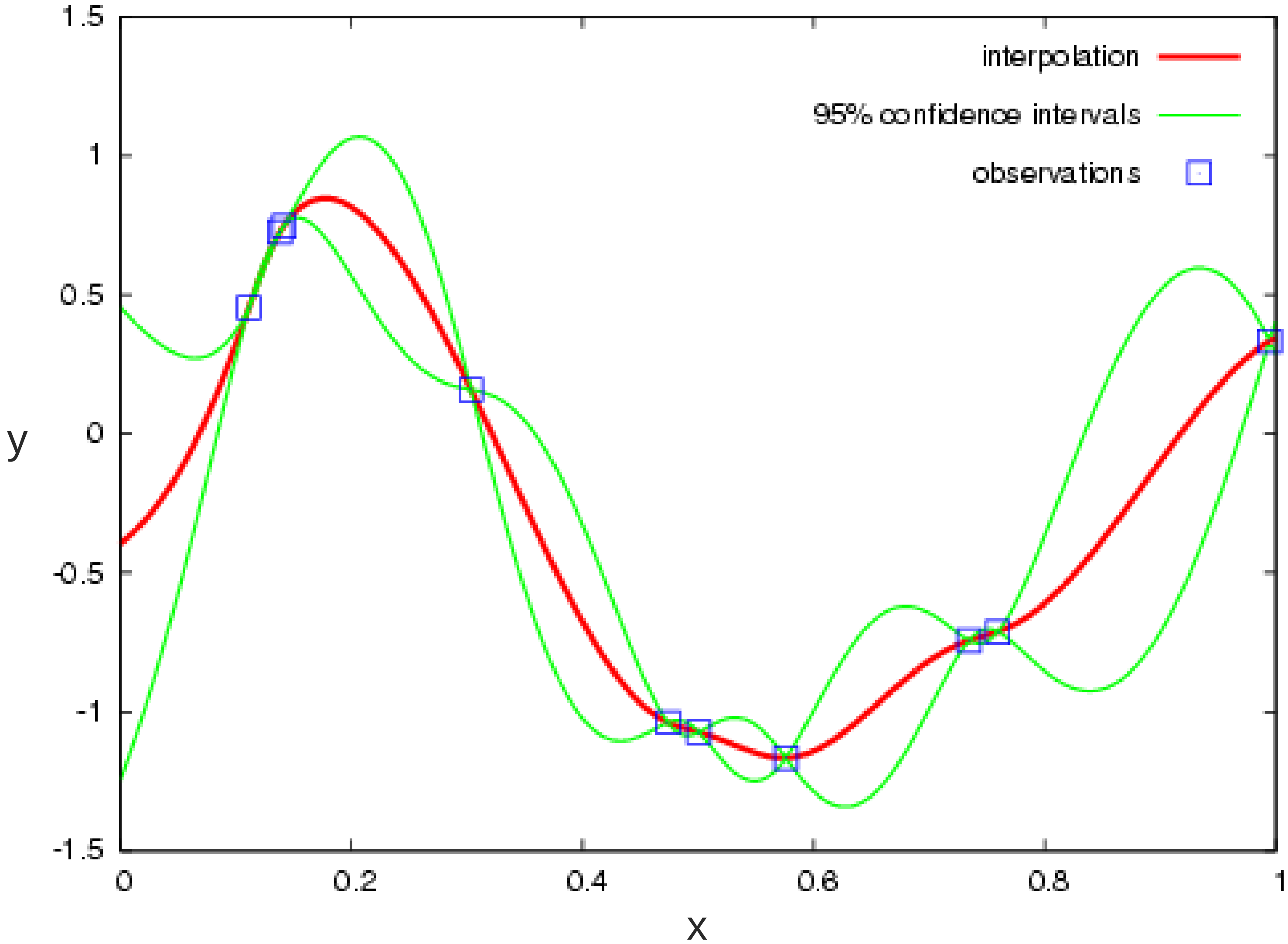
29 CFD Simulations Run – 17 Used to Metamodel & 12 Used as Checkpoints

17-trial Orthogonal Latin Hypercube (OLH) space-filling design settings used for creating the metamodel

12-trial Plackett-Burman screening design settings used as checkpoints – half just inside and half just outside design boundary (convex hull)

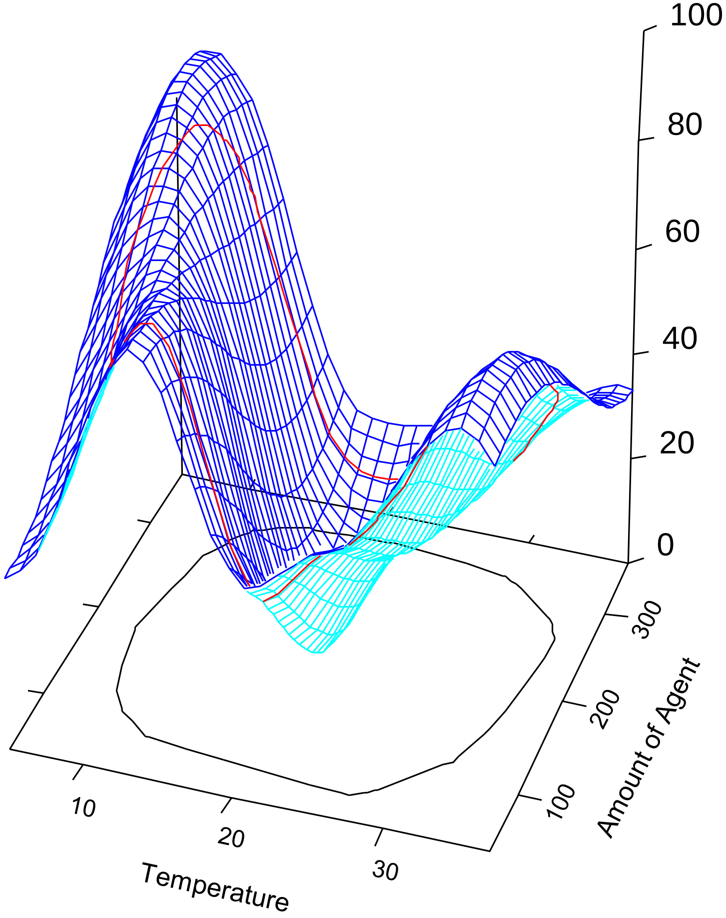
Trial	Time of Day	Temperature	Wind Speed	Wind Direction	Relative Humidity	Cloud Cover	
1	505	37	5.3	247.5	30	0.92	
2	165	13	5.6	281.25	10	0.32	
3	250	19	1.7	225	60	0.8	
4	335	25	2.9	360	55	0.14	
5	1100	35	3.5	202.5	35	0.02	- Min
6	1440	15	3.2	326.25	15	0.74	
7	930	11	6.2	236.25	80	0.44	
8	845	33	5	348.75	75	0.62	
9	760	21	3.8	270	50	0.5	- Mid
10	1015	5	2.3	292.5	70	0.08	
11	1355	29	2	258.75	90	0.68	
12	1270	23	5.9	315	40	0.2	
13	1185	17	4.7	180	45	0.86	
14	420	7	4.1	337.5	65	0.98	- Max
15	80	27	4.4	213.75	85	0.26	
16	590	31	1.4	303.75	20	0.56	
17	675	9	2.6	191.25	25	0.38	
18	972.5	26	3.05	298.125	62.5	0.65	Inside
19	547.5	16	4.55	241.875	62.5	0.65	Outside
20	972.5	26	3.05	241.875	37.5	0.65	Outside
21	547.5	26	4.55	298.125	37.5	0.35	Outside
22	972.5	16	4.55	298.125	62.5	0.35	Inside
23	547.5	16	3.05	241.875	37.5	0.35	Inside
24	547.5	26	4.55	241.875	62.5	0.65	Outside
25	972.5	16	4.55	298.125	37.5	0.65	Inside
26	547.5	26	3.05	298.125	62.5	0.35	Inside
27	547.5	16	3.05	298.125	37.5	0.65	Outside
28	972.5	16	3.05	241.875	62.5	0.35	Outside
29	972.5	26	4.55	241.875	37.5	0.35	Inside

Kriging Fit in 1-D Showing Interpolation and Confidence Intervals on Prediction



Kriging Analysis of Random Data!

10-Variable Metamodel Prediction



Off-Axis Variable Settings

Time wrt Sunset = 360
Wind Speed = 3.8
Wind Direction = 270
Humidity = 50
Cloud Cover = 0.50
Log_{10}(Duration) = 1.0
Latitude (coded) = 17
Longitude (coded) = 17

NOTE: This is a plot of Kriging analysis of the 100 integers between 0 and 99 randomly assigned to 100 space-filling design trials.

The “noise” has been fit perfectly! This is why one should only use this technique with non-stochastic or nearly non-stochastic data!

Seminal Paper on “Space-Filling” DOE for Computer Experiments

- **Design and Analysis of Computer Experiments**
  Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn, H.P.
  Statistical Science 4. 409-423, 1989

- Textbooks on this topic include:
  - Santner, T. J., Williams, B. J., and Notz, W. I. (2003), *The Design and Analysis of Computer Experiments*, Springer, New York
  - Fang, K. T., Li, R. Z., and Sudjianto, A. (2005), *Design and Modeling for Computer Experiments*, Chapman & Hall/CRC Press, New York
  - Kleijnen, J. P. C. (2008), *DASE: design and analysis of simulation experiments*. Springer, New York.

Software Tools for Kriging Analysis

(that I know about…)

- JMP® (called Gaussian Process modeling)
- ECHIP® (called Smoothing analysis)
- SYSTAT® (called Kriging analysis)
- Matlab® Toolbox Modules
  - Design and Analysis of Computer Experiments (DACE)
  - SUrrogate MOdeling (SUMO)
    - Contains DACE as well as another Kriging tool and many other surrogate modeling methods
- PErK (code available from authors of 2003 text by Santner, et. al.)
- "Blind" Kriging – R code potentially available from GA Tech
- The Gaussian Processes Website: http://www.gaussianprocess.org
- Code to do Bayesian Hierarchical Gaussian Process (BHGP) modeling by combining simulation and real experimental data is available from Prof. Peter Qian of the University of Wisconsin
- Code for Nested and Sliced Latin Hypercube Designs also available from Prof. Qian..

Websites for Designs, Software & Publications

- http://harvest.nps.edu/
  The Simulation Experiments & Efficient Design (SEED) Center for Data Farming at Naval Postgraduate School
  - Designs
    - Nearly Orthogonal Latin Hypercubes (NOLH) and
    - Resolution V, Fractional Factorials for many factors
  - Agent-Based Simulation Software
    - Pythagoras
    - MANA (Map Aware Non-uniform Automata)
  - Many Papers for Download and Links to INFORMS and WSC
- http://www.research.att.com/~njas/oadir/index.html
  Library of Orthogonal Arrays maintained by Neil J.A. Sloane
- http://support.sas.com/techsup/technote/ts723.html
  Library of Orthogonal Arrays maintained by Warren F. Kuhfield

Publications Discussing Recent Advances in Metamodeling Simulations

- **Blind Kriging: A New Method for Developing Metamodels**,
  Joseph, V.R., Hung, Y., and Sudjianto, A.,
  ASME Journal of Mechanical Design, 130, 031102-1-8, 2008
- **Gaussian Process Models for Computer Experiments With Qualitative and Quantitative Factors**,
  Qian, P.Z.G., Wu, H., and Wu, C.F.J.,
  Technometrics, 50 (4), 383-396, 2008
- **Bayesian Hierarchical Modeling for Integrating Low-Accuracy and High-Accuracy Experiments,**
  Qian, P. Z. G. and Wu, C. F. J.,
  Technometrics, 50 (2), 192-204, 2008
- **Regression-Based Inverse Distance Weighting for Multivariate Interpolation,**
  Joseph, V.R., and Kang, L.,
  (submitted) Preprint May 2009
- **Nested Latin Hypecube Designs,**
  Qian, P. Z. G.
  Biometrika, 96, 957-970, 2008

Publications listed on web page of Prof. Peter Qian

2011

19. Zhou, Q., Qian, P. Z. G. and Zhou, S (2010), A Simple Approach to Emulation for Computer Models With Qualitative and Quantitative Factors, *Technometrics*, Accepted.
18. He, X. and Qian, P. Z. G. (2010), Nested orthogonal array based Latin hypercube designs, *Biometrika*, Accepted.
17. Xu, X., Haaland, B. and Qian, P. Z. G. (2010), Sudoku based space-filling designs, *Biometrika*, Accepted.

2010

16. Tang, Q. and Qian, P. Z. G. (2010), ``Enhancing the Sample Average Approximation Method with U Designs,''*Biometrika*, 97, 947-960.
15. Haaland, B., Min, W., Qian, P. Z. G. and Amemiya, Y. (2010), ``A Statistical Approach to Thermal Management of Data Centers under Steady State and System Perturbations,''*Journal of the American Statistical Association*, 105, 1030-1041.
14. Qian, P. Z. G. and Ai, M. (2010),``Nested Lattice Sampling: A New Sampling Scheme Derived by Randomizing Nested Orthogonal Arrays,''*Journal of the American Statistical Association*, 105, 1147-1155.
13. Haaland, B. and Qian, P. Z. G. (2010), ''An Approach to Constructing Nested Space-Filling Designs for Multi-Fidelity Computer Experiments,''*Statistica Sinica*, 20, 1063-1075.
12. Wang, F., Hwang, Y., Qian, P. Z. G. and Wang, X. (2010), ``A Statistics-Guided Approach to Precise Characterization of Nanowire Morphology,'' *American Chemical Society Nano* , 4, 855-862.

2009

11. Qian, P. Z. G. and Wu, C. F. J. (2009), “Sliced Space-Filling Designs,” *Biometrika*, 96, 945-956.
10. Qian, P. Z. G.(2009), “Nested Latin Hypercube Designs,”*Biometrika*, 96, 957-970.

Publications listed on web page of Prof. Peter Qian

2009 (continued)

9. Qian, P. Z. G., Ai, M. and Wu, C. F. J. (2009), "Construction of Nested Space-Filling Designs,"*Annals of Statistics*, 37, 3616-43.

8. Qian, P. Z. G., Tang, B. and Wu, C. F. J. (2009), "Nested Space-Filling Designs for Experiments with Two Levels of Accuracy," *Statistica Sinica*, 19, 287-300.

2008

7. Qian, P. Z. G., Wu, H. and Wu, C. F. J. (2008), "Gaussian Process Models for Computer Experiments with Qualitative and Quantitative Factors," *Technometrics*, 50, 383-396.

6. Qian, P. Z. G. and Wu, C. F. J. (2008), "Bayesian Hierarchical Modeling for Integrating Low-accuracy and High-accuracy Experiments," *Technometrics*, 50, 192-204.

5. Mukerjee, R., Qian, P. Z. G. and Wu, C. F. J. (2008), "On the Existence of Nested Orthogonal Arrays," *Discrete Mathematics*, 308, 4635-4642.

2007 and Earlier

4. Negrut, D., Qian, P. Z. G., and Khude, N. (2007),``Building Gaussian Process Based Metamodels Using Variable-fidelity Experiments for Dynamic Analysis of Mechanical Systems,''*Proceedings of the 2007 ASME International Mechanical Engineering Congress and Exposition*, Seattle, WA.

3. Qian, Z., Seepersad, C., Joseph, R., Allen, J. and Wu, C. F. J. (2006), "Building Surrogate Models with Detailed and Approximate Simulations," *ASME Journal of Mechanical Design*, 128, 668-677.

2. Qian, Z., Jiang, W. and Tsui, K. (2006), "Churn Detection via Customer Profile Modeling," *International Journal of Production Research*, 44, 2913-2933.

1. Qian, Z. and Shapiro, A. (2006), "Simulation-based Approach for Estimation in Latent Variable Models," *Computational Statistics and Data Analysis*, 51, 1243-1259.

Publications listed on web page of Prof. Peter Qian

Submitted

Xu, X., Hwang, Y., Kim, T., Wang, F., Wang, X. and Qian, P. Z. G. (2011), Sequential Synthesis of Nanomaterials via Level-expansion.

Hwang, Y., Khude, N., Qian, P. Z. G. and Negrut, D. (2010), Statistical Emulation of Multi-fidelity Simulations of Mechanical Dynamics Systems.

Xiong, S., Qian, P. Z. G. and Wu, C. F. J. (2010), Sequential Design and Analysis of High-accuracy and Low-accuracy Computer Codes.

Deng, X., Lin, C. and Qian, P. Z. G. (2010), Designs for the Lasso.

Zhou, Q., Qian, P. Z. G. and Zhou, S. (2010), Surrogate Modeling of Multistage Assembly Processes Using Integrated Emulation.

Haaland, B. and Qian, P. Z. G. (2010), Accurate Emulators for Large-Scale Computer Experiments.

Deng, X. and Qian, P. Z. G. (2010), Sliced Cross-validation for Efficient Estimation of the Error Rate of a Classification Rule.

Qian, P. Z. G. and Amemiya, Y. (2010), A Structural Equation Method for Modeling Multivariate Temperature Data from Data Center Computer Experiments.

Qian, P. Z. G. (2010), Sliced Latin Hypercube Designs.

Li, J. and Qian, P. Z. G. (2010), Construction of Nested (Nearly) Orthogonal Designs for Computer Experiments.

Qian, P. Z. G. and Ai, M. (2010), Nested Asymmetric Lattice Samples.

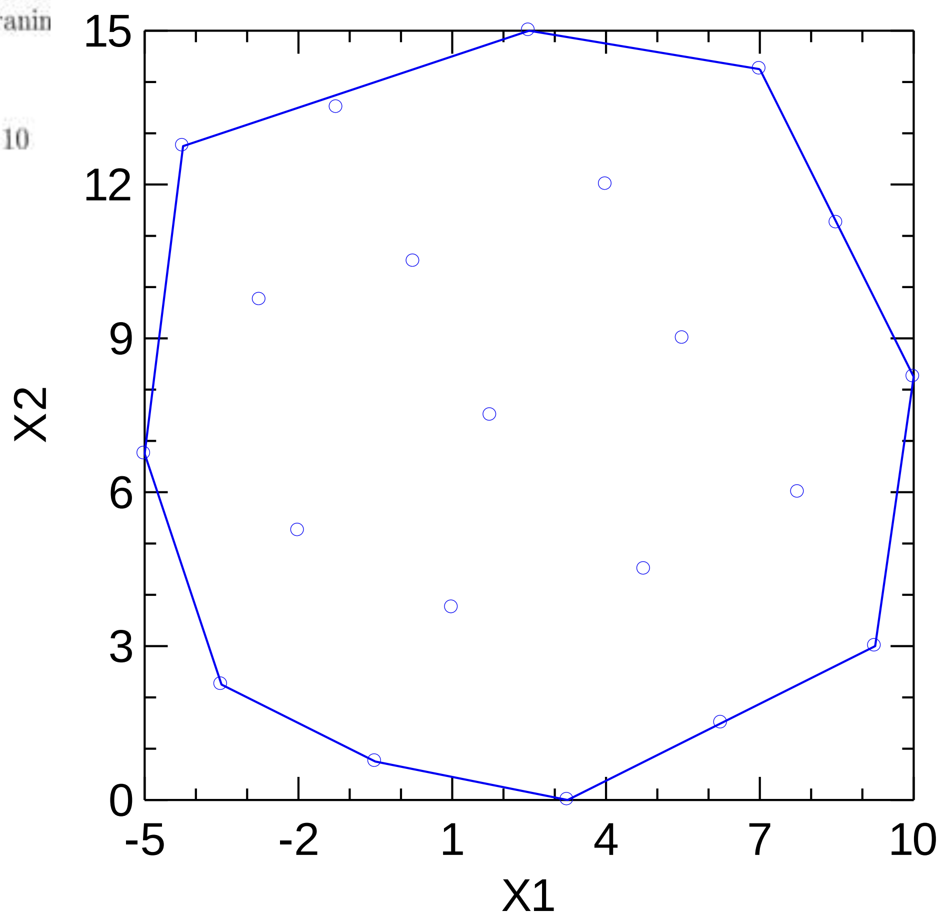
Example Latin Hypercube Design and Data Calculated with Branin Function from Santner, Williams and Notz (2003)

C.3 Examples

The following examples demonstrate many possible uses of PErK. The responses for these examples are based on the *Branin function*. The Branin function is the real-valued function of two variables

$$y_B(x_1, x_2) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos(x_1) + 10$$

Trial	x1	x2	Y_B
1	7.75	6	35.80951
2	1	3.75	14.86287
3	10	8.25	31.41880
4	4.75	4.5	19.87899
5	2.5	15	141.88566
6	-3.5	2.25	99.43335
7	3.25	0	3.88973
8	-5	6.75	97.47380
9	-4.25	12.75	6.27060
10	6.25	1.5	19.85914
11	8.5	11.25	95.50587
12	7	14.25	181.74214
13	-0.5	0.75	49.39445
14	-2	5.25	23.13762
15	0.25	10.5	43.09524
16	9.25	3	2.82392
17	-2.75	9.75	3.61474
18	5.5	9	75.79100
19	4	12	104.11175
20	-1.25	13.5	43.33586
21	1.75	7.5	23.39797



Example Latin Hypercube Design, Data Calculated with Branin Function and Plots from Kriging Analysis

C.3 Examples

The following examples demonstrate many possible uses of PErK. The responses for these examples are based on the *Branin function*. The Branin function is the real-valued function of two variables

$$y_B(x_1, x_2) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos(x_1) + 10$$

Trial	x1	x2	Y_B
1	7.75	6	35.80951
2	1	3.75	14.86287
3	10	8.25	31.41880
4	4.75	4.5	19.87899
5	2.5	15	141.88566
6	-3.5	2.25	99.43335
7	3.25	0	3.88973
8	-5	6.75	97.47380
9	-4.25	12.75	6.27060
10	6.25	1.5	19.85914
11	8.5	11.25	95.50587
12	7	14.25	181.74214
13	-0.5	0.75	49.39445
14	-2	5.25	23.13762
15	0.25	10.5	43.09524
16	9.25	3	2.82392
17	-2.75	9.75	3.61474
18	5.5	9	75.79100
19	4	12	104.11175
20	-1.25	13.5	43.33586
21	1.75	7.5	23.39797

Plot from textbook of Branin Function

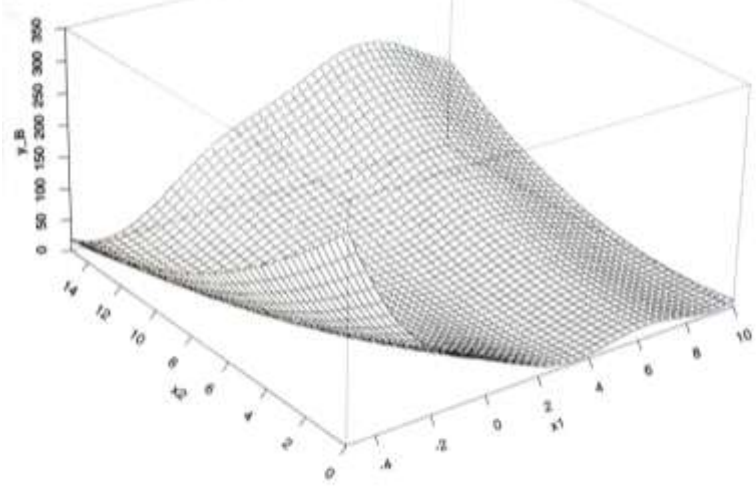
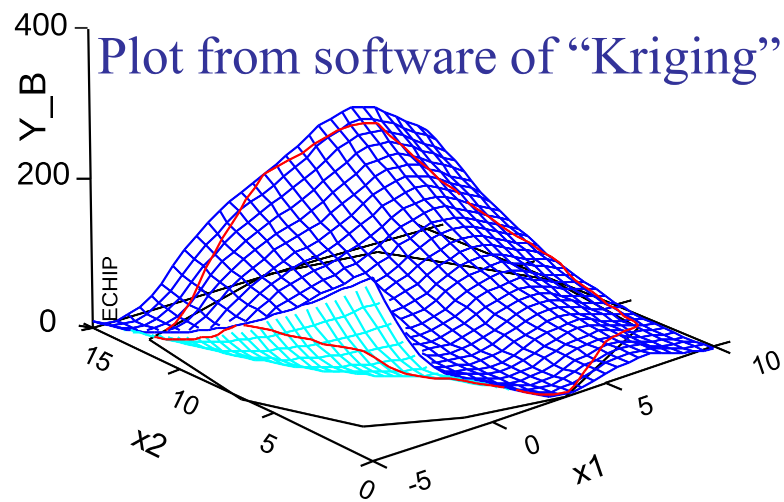


FIGURE C.1. The Branin function on $[-5, 10] \times [0, 15]$

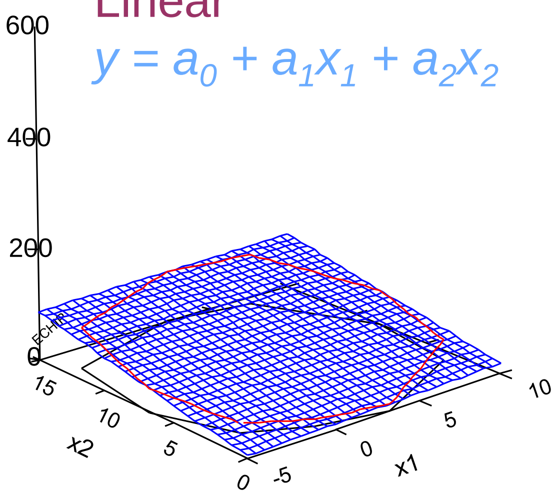
Plot from software of "Kriging" fit



Comparing Surfaces for Increasingly Complex Polynomial Models Fit to Data from the Branin Function

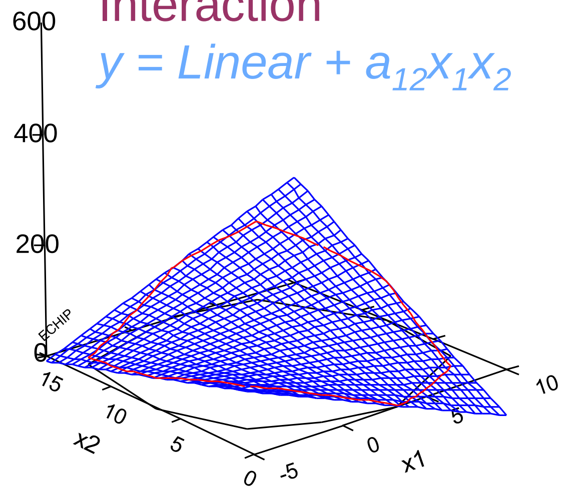
Linear

$y = a_0 + a_1x_1 + a_2x_2$



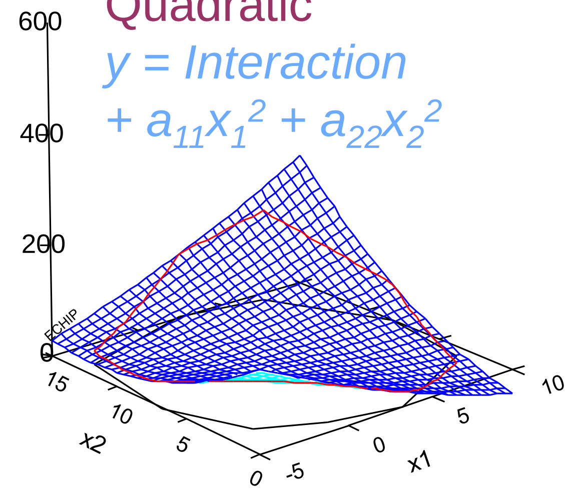
Interaction

$y = Linear + a_{12}x_1x_2$



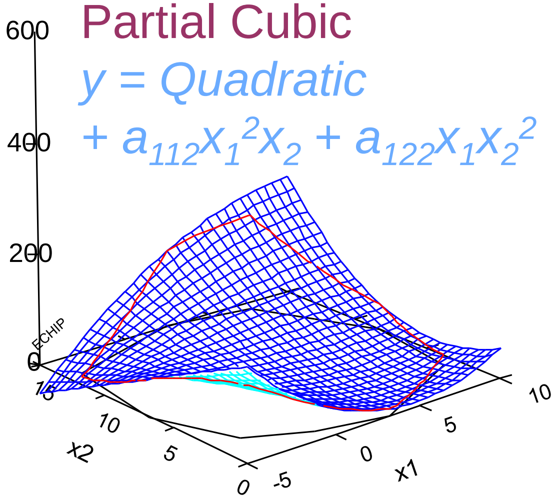
Quadratic

$y = Interaction + a_{11}x_1^2 + a_{22}x_2^2$



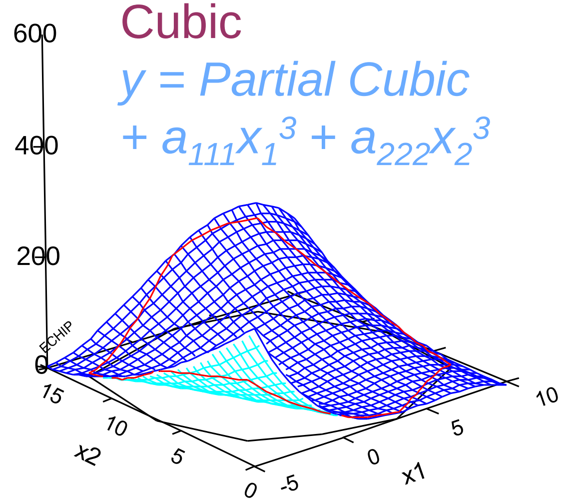
Partial Cubic

$y = Quadratic + a_{112}x_1^2x_2 + a_{122}x_1x_2^2$

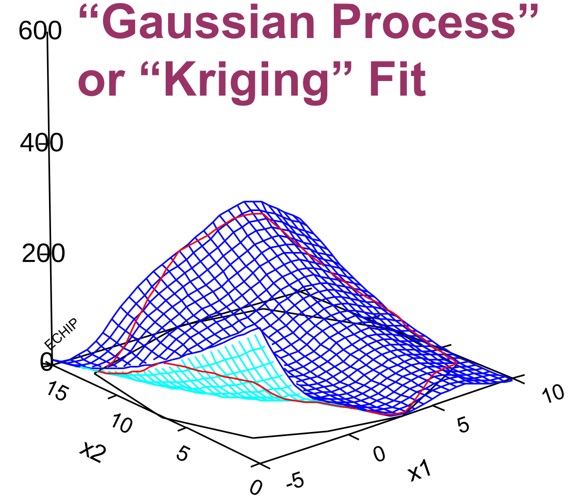


Cubic

$y = Partial\ Cubic + a_{111}x_1^3 + a_{222}x_2^3$



"Gaussian Process" or "Kriging" Fit



The full *cubic* model closely approximates the Branin function, but still cannot capture the ripples seen in the fit using the Kriging method.

Simulations Being Used had at Least 10 Variables

- Branin function example is trivial. With 2 control variables the full cubic model has 10 terms.
- What if your simulation has 10 control variables?
  - Full cubic model has 166 terms!
  - And still may not be complex enough to accurately approximate the simulation

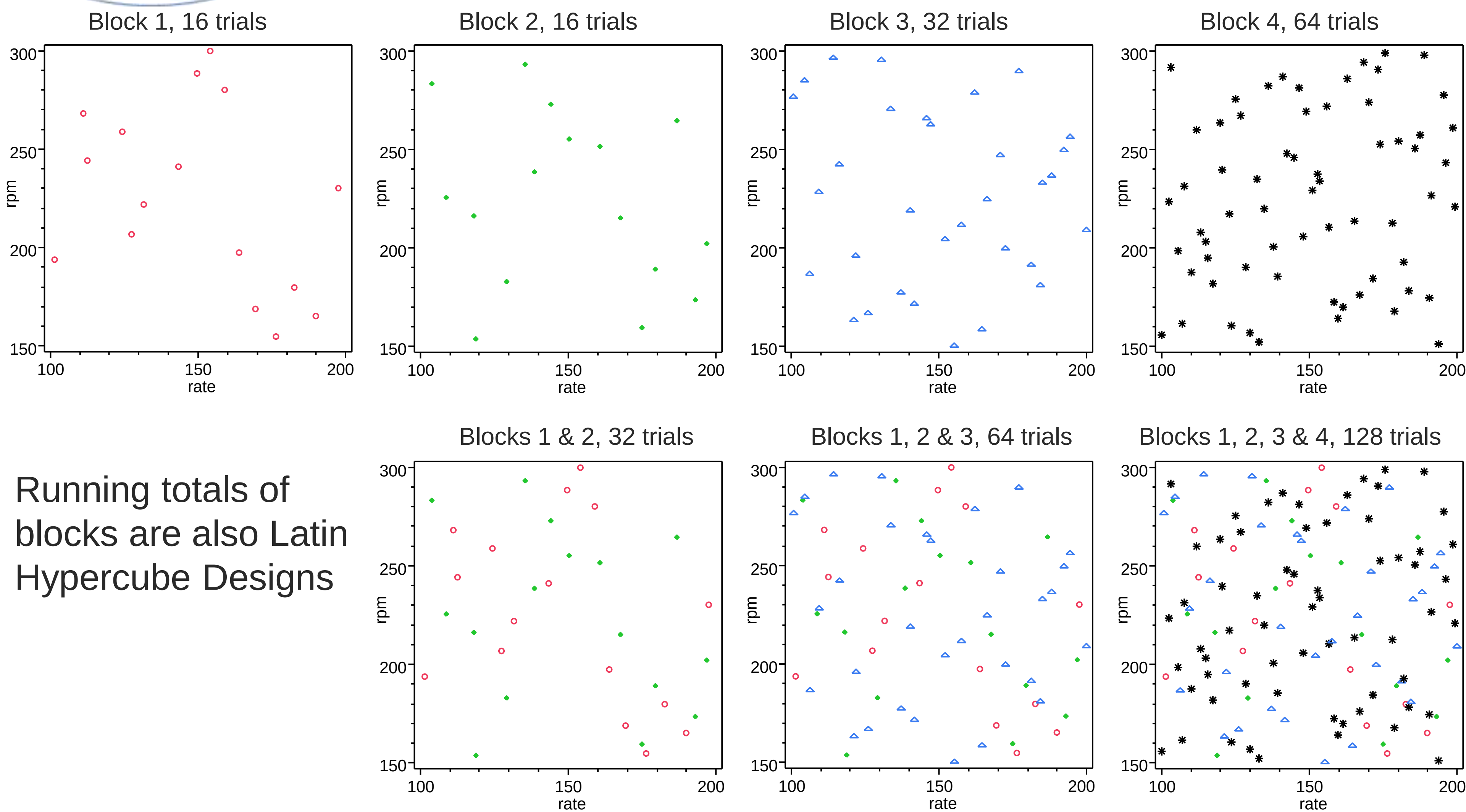
Why Is a Sequential Approach So Useful?

We wanted to not just do sensitivity analysis of the factors, but **provide an interactive surrogate model of the long-running simulation so that analysts could evaluate "what if?" scenarios.**

The problem was that the Computational Fluid Dynamics models we were looking to run could take a week on a single CPU or **12 hours on 50 CPU cluster**. With on the order of 10 factors we expected to need to run on the order of **100 simulations**. **This meant it could be weeks or months before we could start our analysis.**

Nested Latin Hypercube Designs gave us a way to start analyzing data after about the first 20% of the simulations were run. We also wanted to be able to run just enough simulations to achieve a surrogate model accuracy of 90%.

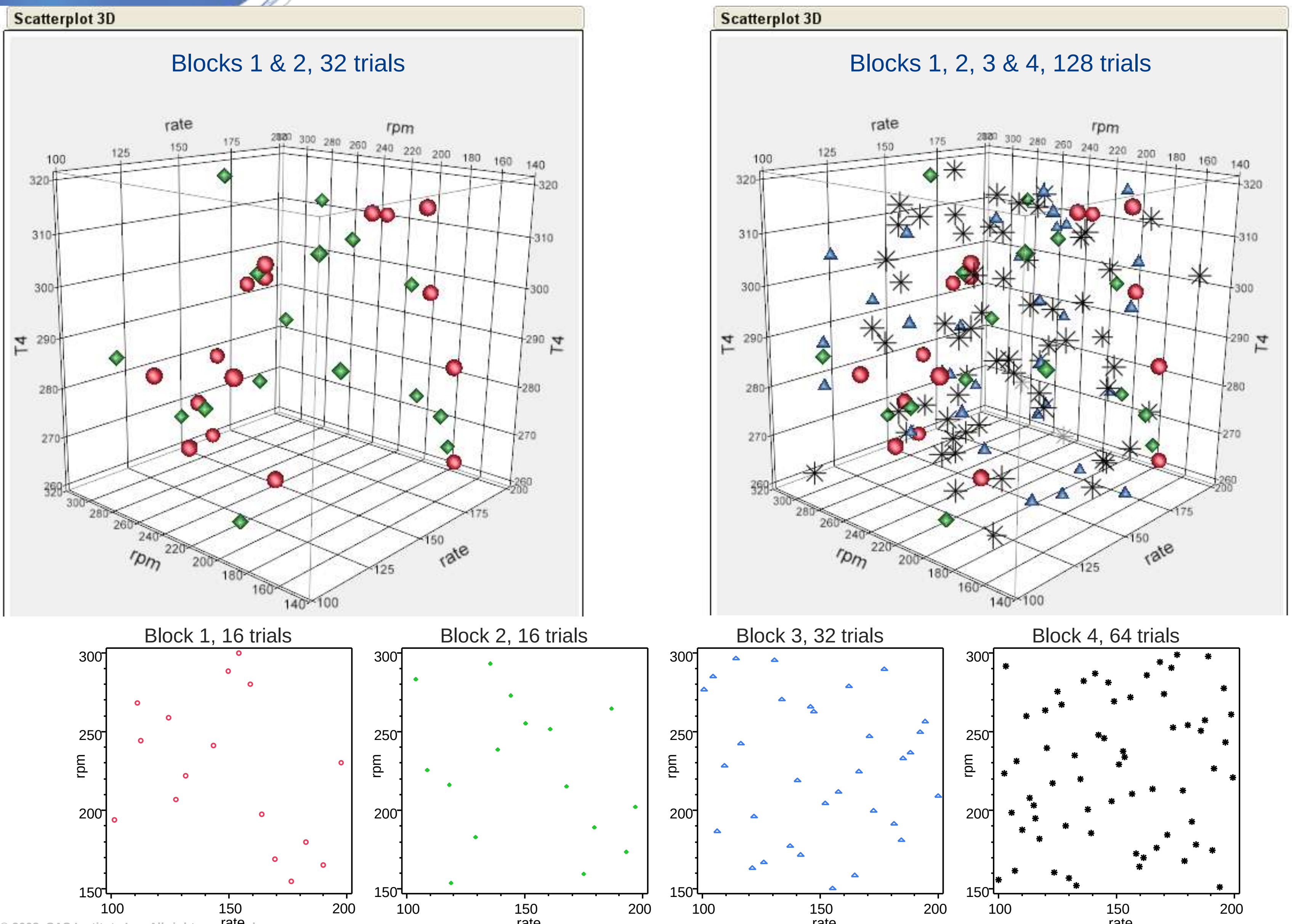
Projections of Trial Locations in 2 factors for a 10-factor, 128-trial, Nested Latin Hypercube Design* (NLHD) with 4 Blocks



Running totals of blocks are also Latin Hypercube Designs

*Generated with Matlab Code Received from Prof. Peter Qian of U of Wi.

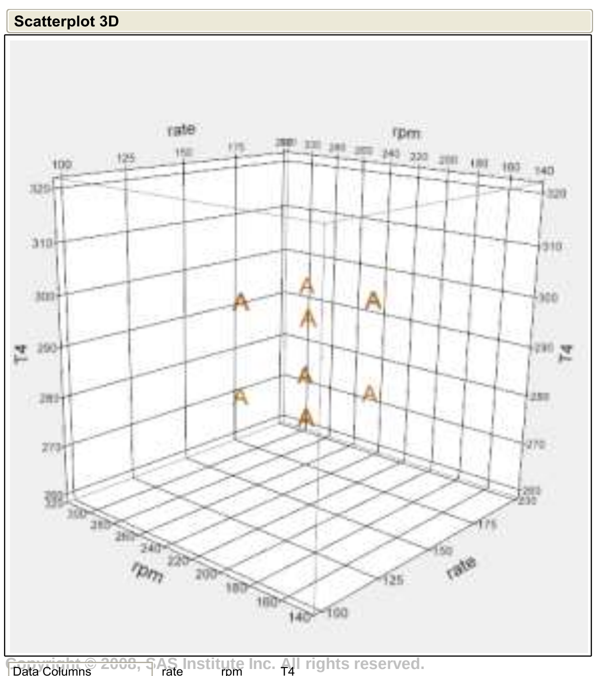
Projections of Trial Locations in 3 factors for a 10-factor. 128-trial, Nested Latin Hypercube Design (NLHD) in 4 Blocks

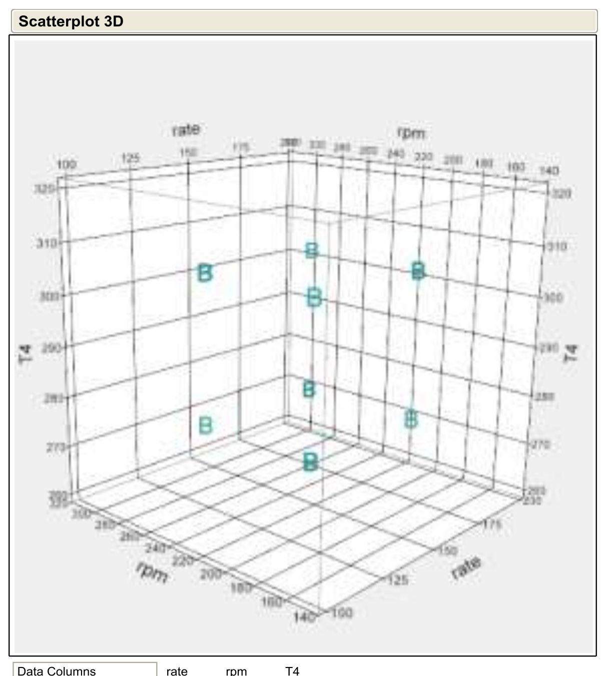


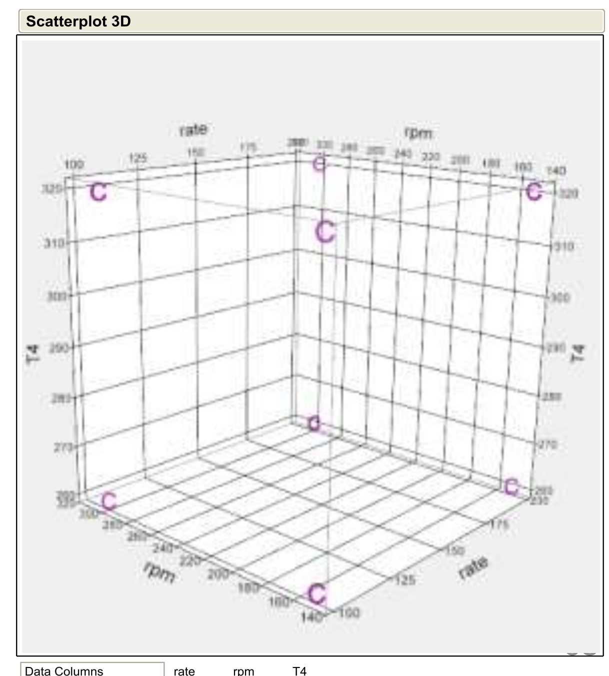
Why Run Simulations in Sequential Blocks?

The point of running this sequence of blocks is to be able to evaluate the surrogate model after each stage to see how accurately it is predicting observed values of 3 sets of checkpoint trials. If it proves to be sufficiently accurate, then subsequent blocks of simulation trials need not be run.

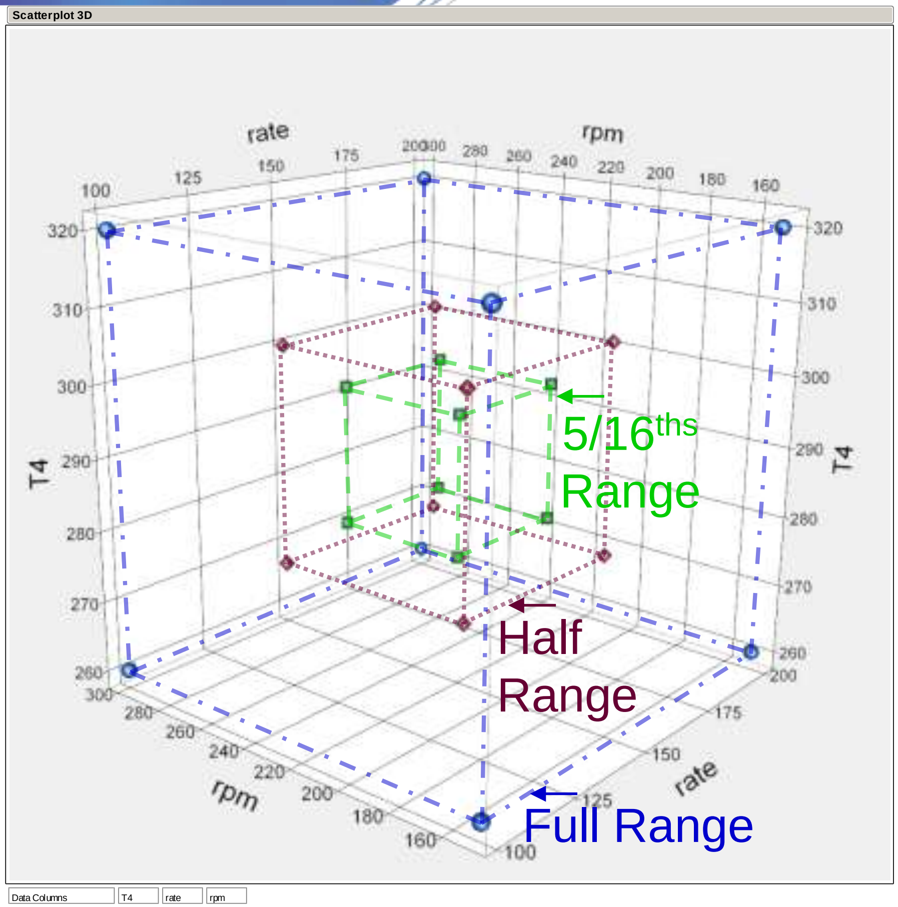
Without the NLHD approach one has to choose the “right” size space-filling design in order to get useful results. If you choose too small a design, one has to start over with a larger design.

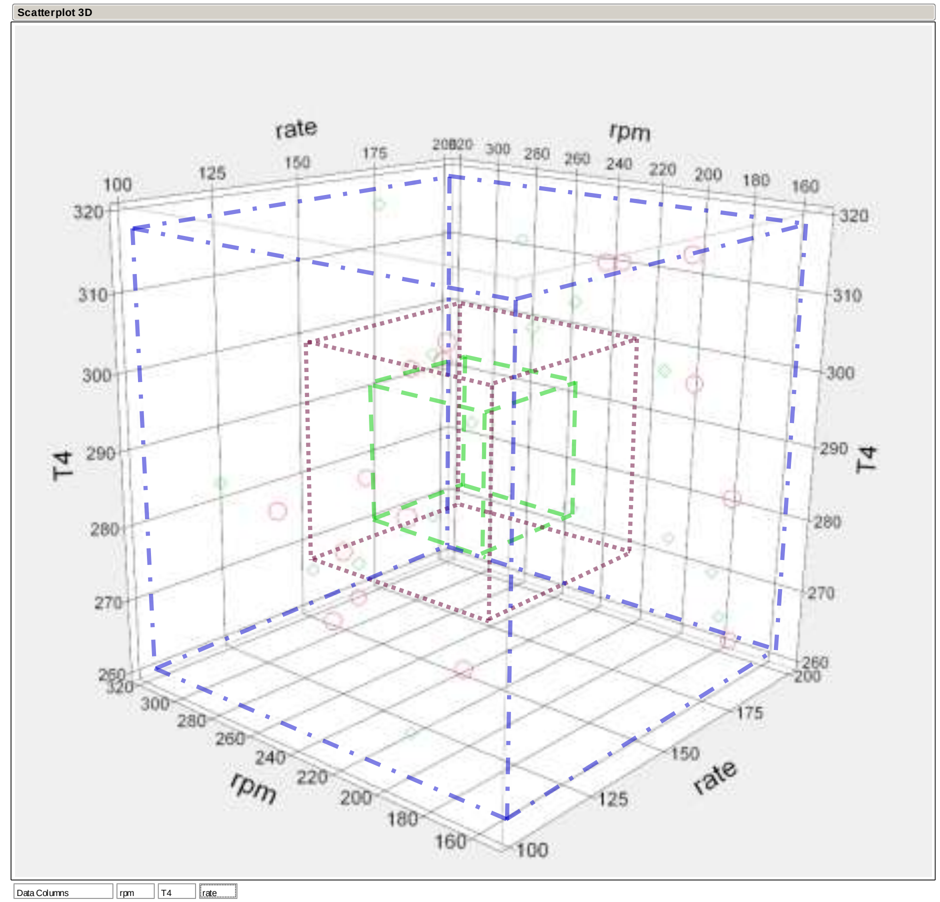






Relative Ranges for 3 Sets of 12 Checkpoints Shown in a 3-factor Space and Superimposed on Block 1 & 2 trials

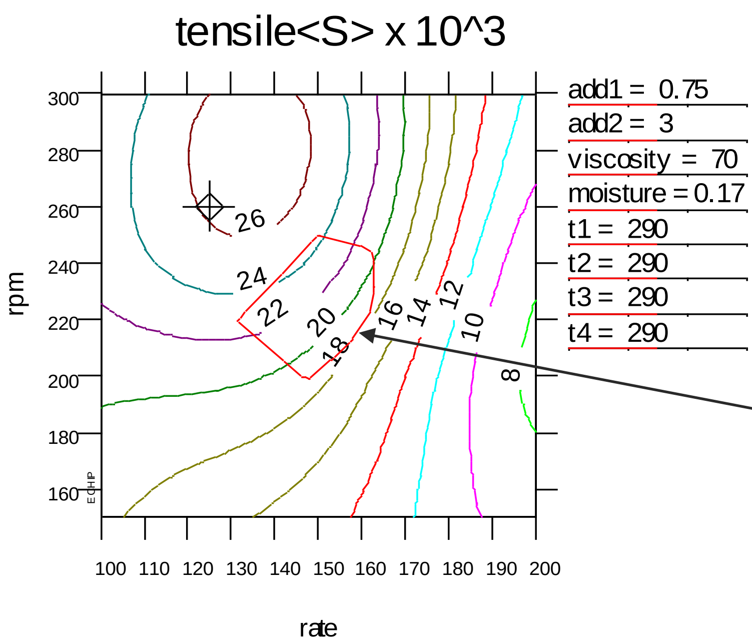




In the full design space over 10 factors there are 10 dimensions and 1024 corners. The 12 trials in a Plackett-Burman design populate only about 0.1% of these combinations of settings.

Today, I would use Definitive Screening Design with 21 trials for 10 factors but also get information at midpoints of each factor.

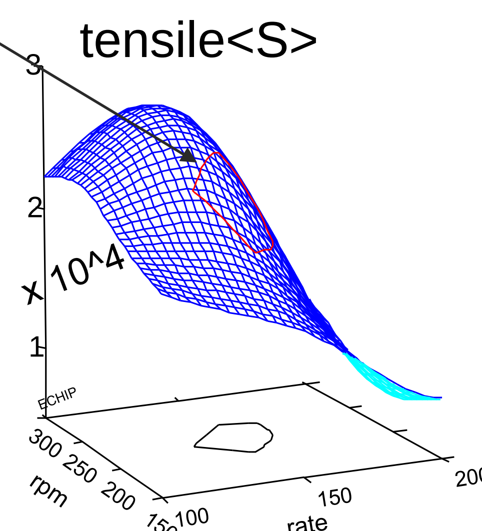
Used 10-factor Process Based on a Transcendental Function as the “Simulation” to Evaluate Improvement in Accuracy



rate=125.00		rpm=260.00
Value	LowLimit	HighLimit
26259.40	-1.#J	1.#J

The 10-dimensional design space is only sparsely covered by the initial 16-trial NLHD Block. As a result only a small fraction of the full design region is valid for interpolation with the Kriging analysis.

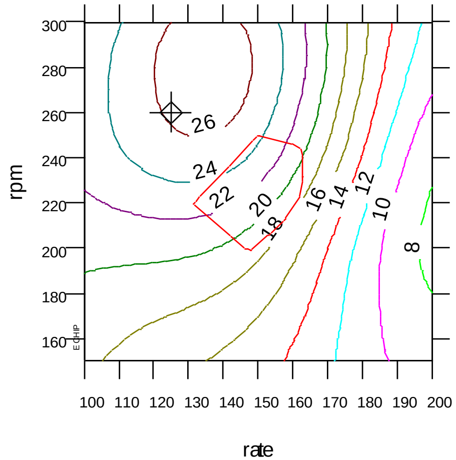
Red polygon marks boundary between regions of interpolation (inside) and extrapolation (outside). Statistical name for the design boundary is the “Convex Hull.”



NLHD, 16 trials Block 1 only

Outside Design ← Note that this entire plot is extrapolation.

tensile<S> x 10^3

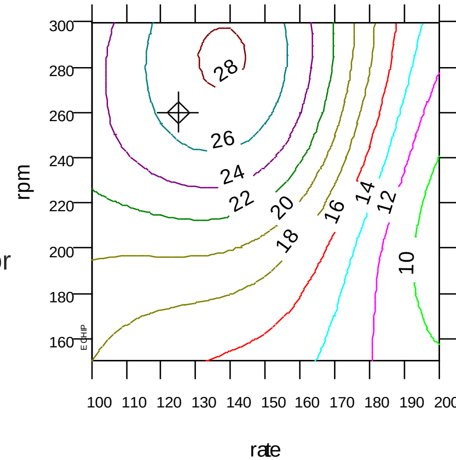


add1 = 0.75
add2 = 3
viscosity = 70
moisture = 0.17
t1 = 290
t2 = 290
t3 = 290
t4 = 290

All off-axis factor settings @ mid range value

rate=125.00		rpm=260.00
Value	Low Limit	High Limit
26259.40	-1.#J	1.#J

tensile<S> x 10^3



add1 = 0.75
add2 = 3
viscosity = 75
moisture = 0.17
t1 = 290
t2 = 290
t3 = 290
t4 = 305

viscosity and *t4* factor settings @ 75% of range

Six of 8 off-axis factor settings @ mid range value

rate=125.00		rpm=260.00
Value	Low Limit	High Limit
26826.46	-1.#J	1.#J

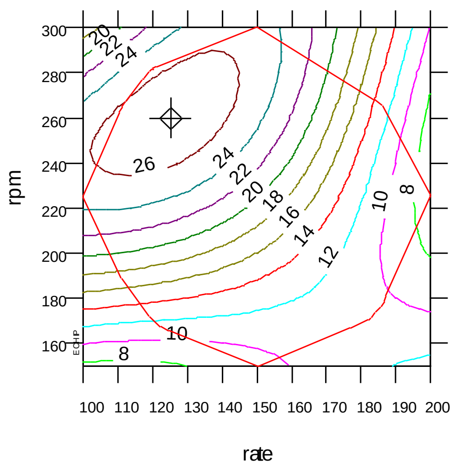
COEFFICIENTS	SD	P	CONDITION	TERM
18903.8				0 CONSTANT
-6319.34	8425.11	0.4870-	0.559	1 add1
-882.924	1658.5	0.6173-	0.703	2 add2
139.372	143.796	0.3769-	0.820	3 viscosity
-25974.4	19170.6	0.2334	0.813	4 moisture
32.8283	74.8572	0.6793-	0.515	5 t1
21.6944	79.5117	0.7959-	0.488	6 t2
61.0479	88.6668	0.5218-	0.432	7 t3
44.7973	81.6762	0.6070-	0.477	8 t4
-128.299	42.7544	0.0301	0.550	9 rate
96.4185	27.5245	0.0172	0.546	10 rpm

N trials = 16

Closer to 1.000 the CONDITION is, the closer the term is to being orthogonal

NLHD, 16 + 12PB = 28

tensile<S> x 10^3

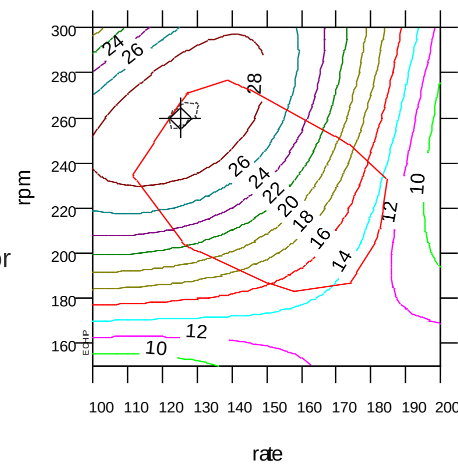


add1 = 0.75
add2 = 3
viscosity = 70
moisture = 0.17
t1 = 290
t2 = 290
t3 = 290
t4 = 290

All off-axis factor settings @ mid range value

rate=125.00		rpm=260.00
Value	Low Limit	High Limit
27257.18	-1.#J	1.#J

tensile<S> x 10^3



add1 = 0.75
add2 = 3
viscosity = 75
moisture = 0.17
t1 = 290
t2 = 290
t3 = 290
t4 = 305

viscosity and *t4* factor settings @ 75% of range

Six of 8 off-axis factor settings @ mid range value

rate=125.00		rpm=260.00
Value	Low Limit	High Limit
30078.62	-1.#J	1.#J

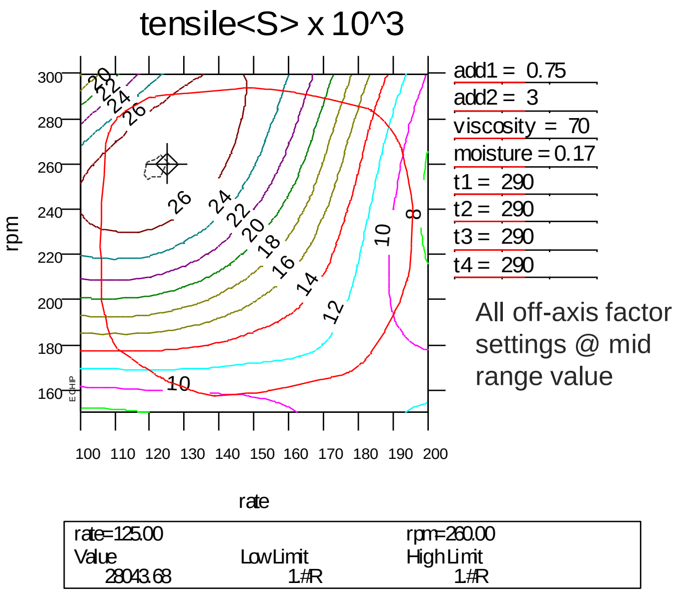
COEFFICIENTS	SD	P	CONDITION	TERM
15760.2				0 CONSTANT
-721.328	6720.15	0.9158-	0.958	1 add1
966.423	1647.02	0.5651-	0.974	2 add2
65.0763	163.369	0.6953-	0.986	3 viscosity
4164.81	21651	0.8497-	0.989	4 moisture
-7.23148	54.8311	0.8966-	0.973	5 t1
1.39981	54.9105	0.9800-	0.973	6 t2
-38.7242	56.0048	0.4986-	0.951	7 t3
47.8879	54.5957	0.3926-	0.981	8 t4
-49.521	32.9741	0.1515	0.976	9 rate
43.0792	21.8001	0.0646	0.971	10 rpm

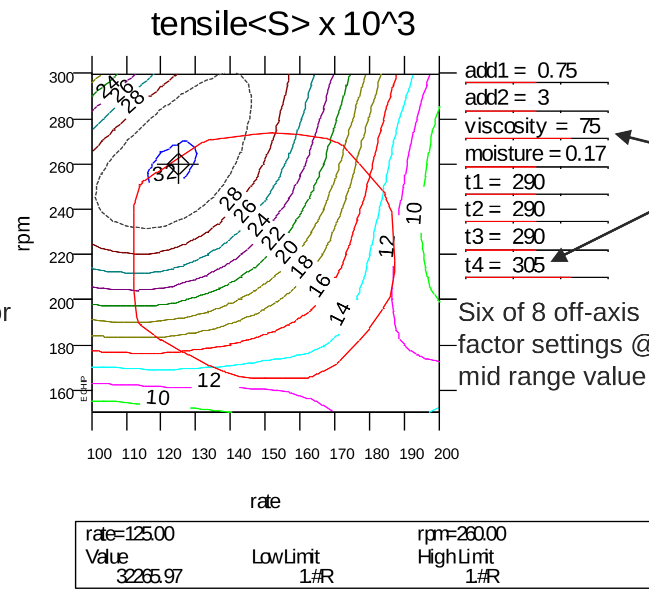
N trials = 28

Closer to 1.000 the CONDITION is, the closer the term is to being orthogonal

Inclusion of checkpoints – here the 12 over the full range of the factors – increases the size of the design boundary and the volume of interpolation region.

NLHD, 16+16+32+64=128
Blocks 1, 2, 3 & 4



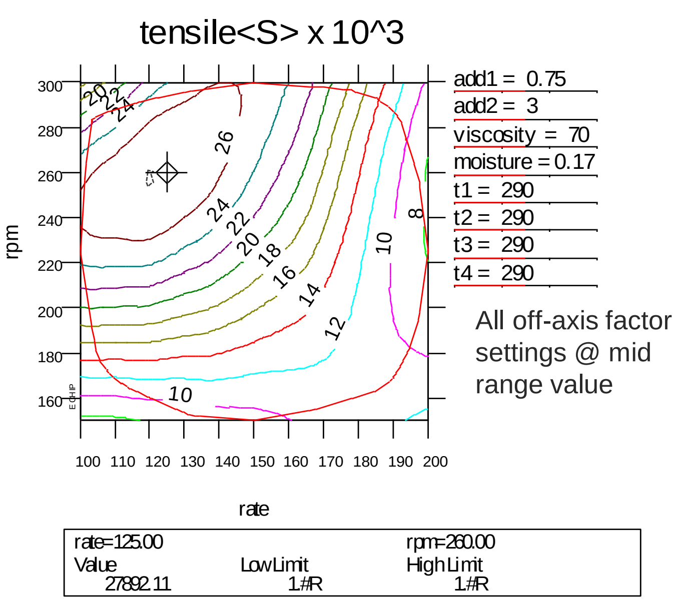


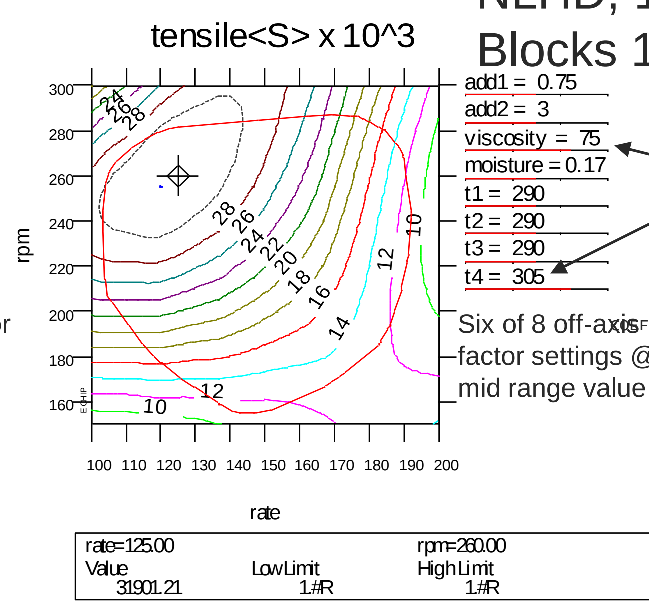
viscosity and *t4* factor settings @ 75% of range

COEFFICIENTS	SD	P	CONDITION	TERM
17631.9				0 CONSTANT
-682.306	2233.14	0.7605-	0.945	1 add1
884.808	540.923	0.1046	0.975	2 add2
300.046	54.081	0.0000	0.976	3 viscosity
-3680.05	7202.41	0.6104-	0.977	4 moisture
5.82534	18.0142	0.7470-	0.976	5 t1
-11.1487	18.0473	0.5379-	0.975	6 t2
31.8328	18.7587	0.0924	0.938	7 t3
86.4875	17.8697	0.0000	0.984	8 t4
-130.079	10.864	0.0000	0.971	9 rate
90.3335	7.22455	0.0000	0.974	10 rpm

N trials = 128

NLHD, 16+16+32+64+36ckps=164
Blocks 1- 4, & 3 sets of checkpoints





viscosity and *t4* factor settings @ 75% of range

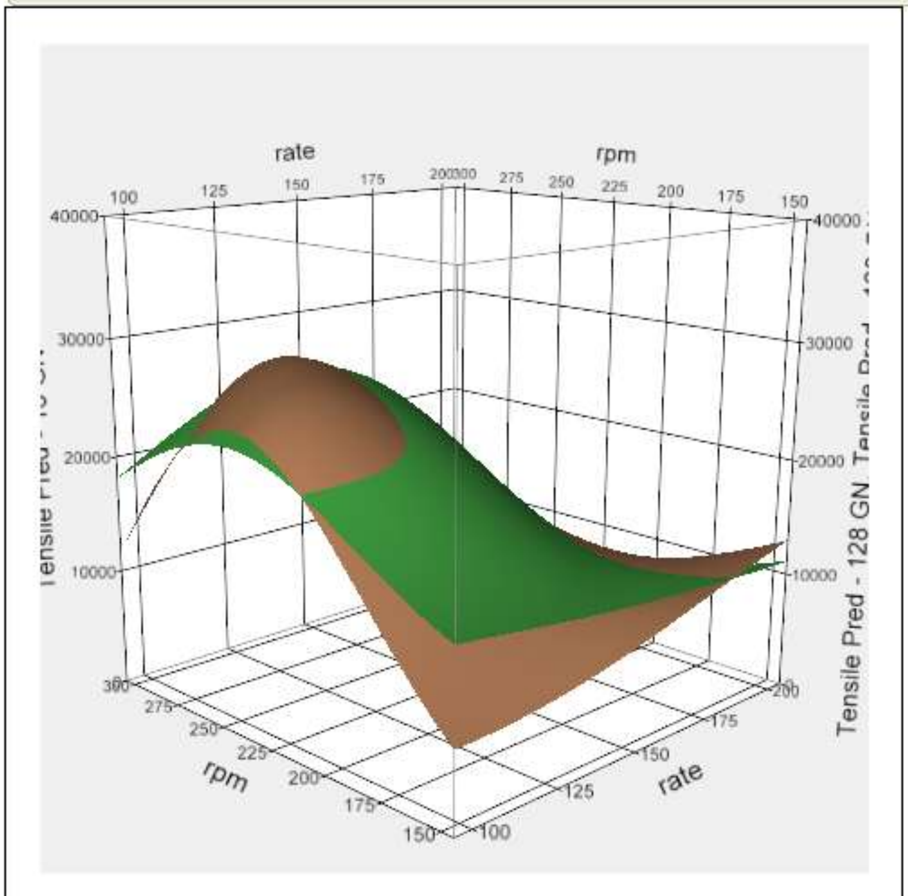
COEFFICIENTS	SD	P	CONDITION	TERM
17357				0 CONSTANT
-843.041	2246.05	0.7079-	0.973	1 add1
1016.04	553.662	0.0684	0.987	2 add2
226.532	55.3433	0.0001	0.987	3 viscosity
2098.09	7370.9	0.7763-	0.988	4 moisture
8.98137	18.4363	0.6268-	0.988	5 t1
-16.0393	18.4583	0.3862-	0.987	6 t2
5.38914	18.7981	0.7747-	0.969	7 t3
88.5134	18.3523	0.0000	0.992	8 t4
-106.215	11.086	0.0000	0.986	9 rate
76.3942	7.37976	0.0000	0.987	10 rpm

N trials = 164

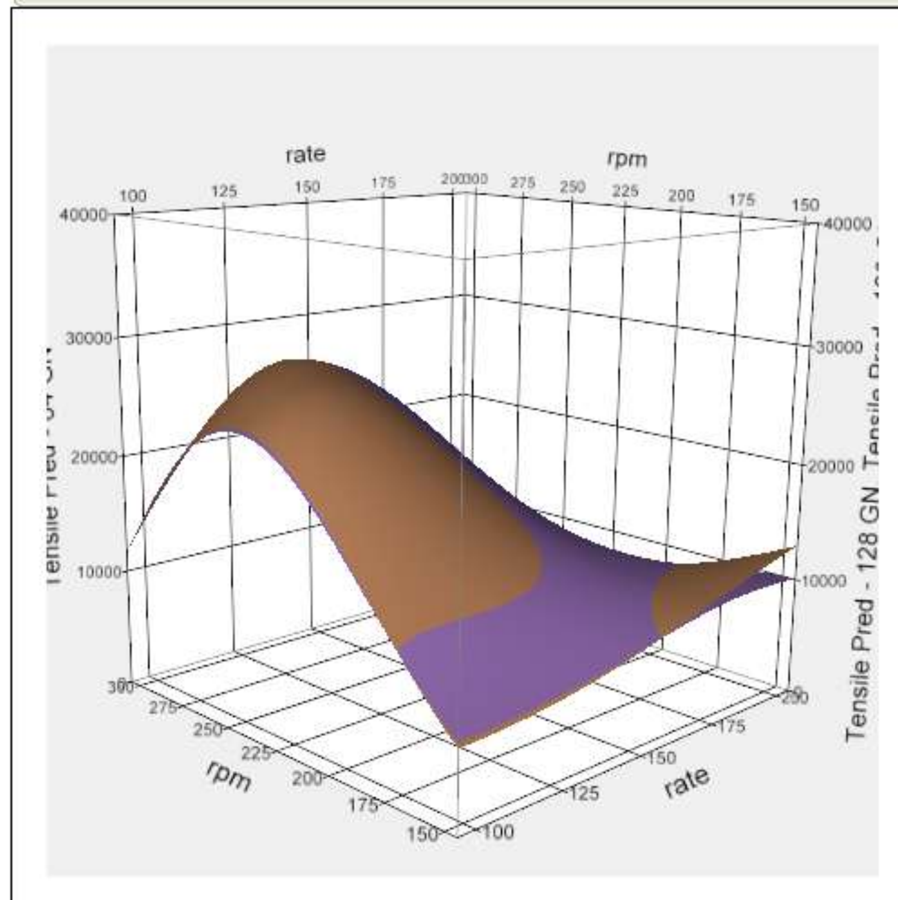
Compare Response Surfaces for fit of 16 vs. fit of 128 trials (left) and for fit of 64 vs. fit of 128 trials (right)

Stage 1 fit of16 trials colored green
Stage 4 fit 128 trials colored brown
Stage 3 fit 64 trials colored purple

Surface Plot

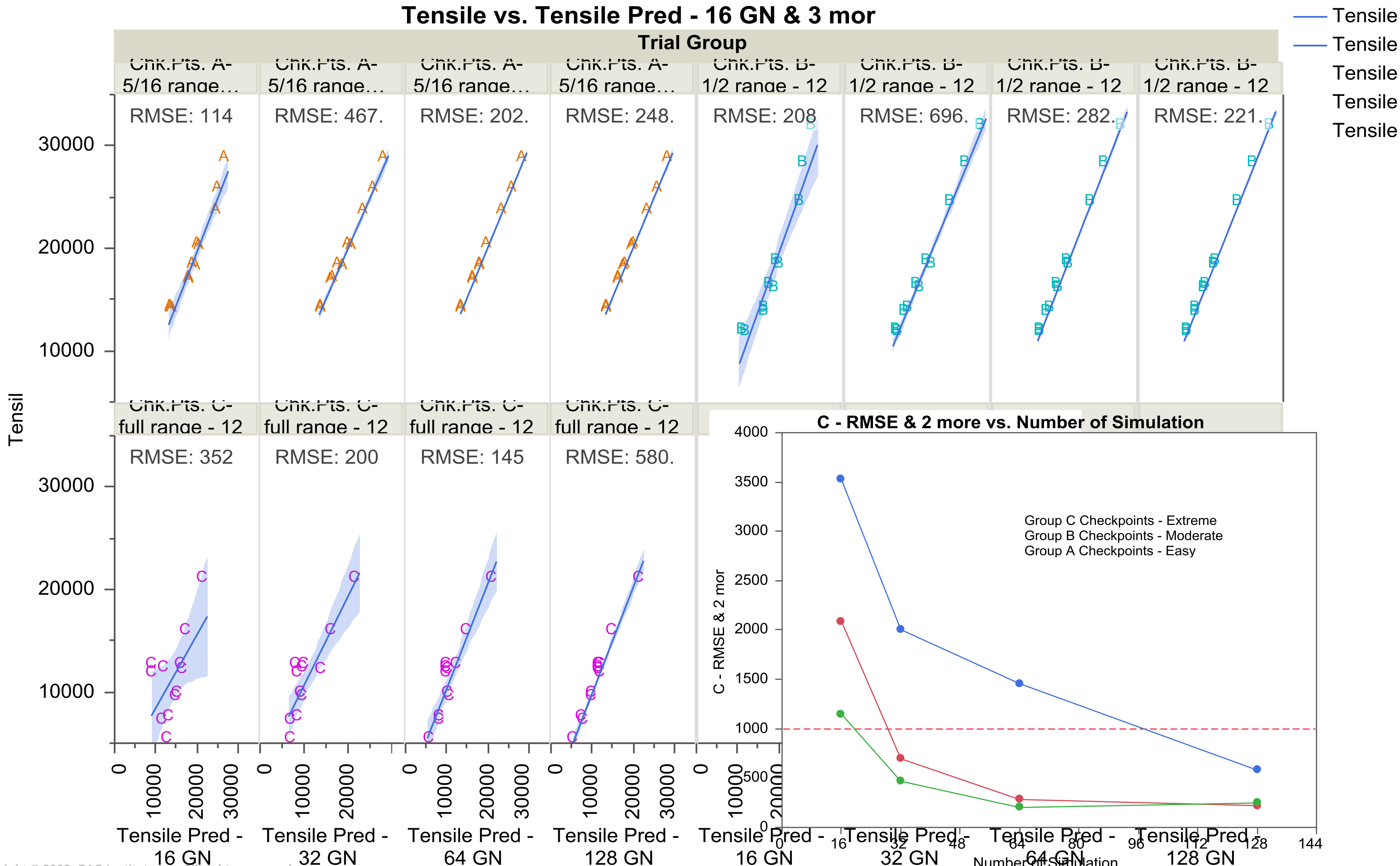


Surface Plot

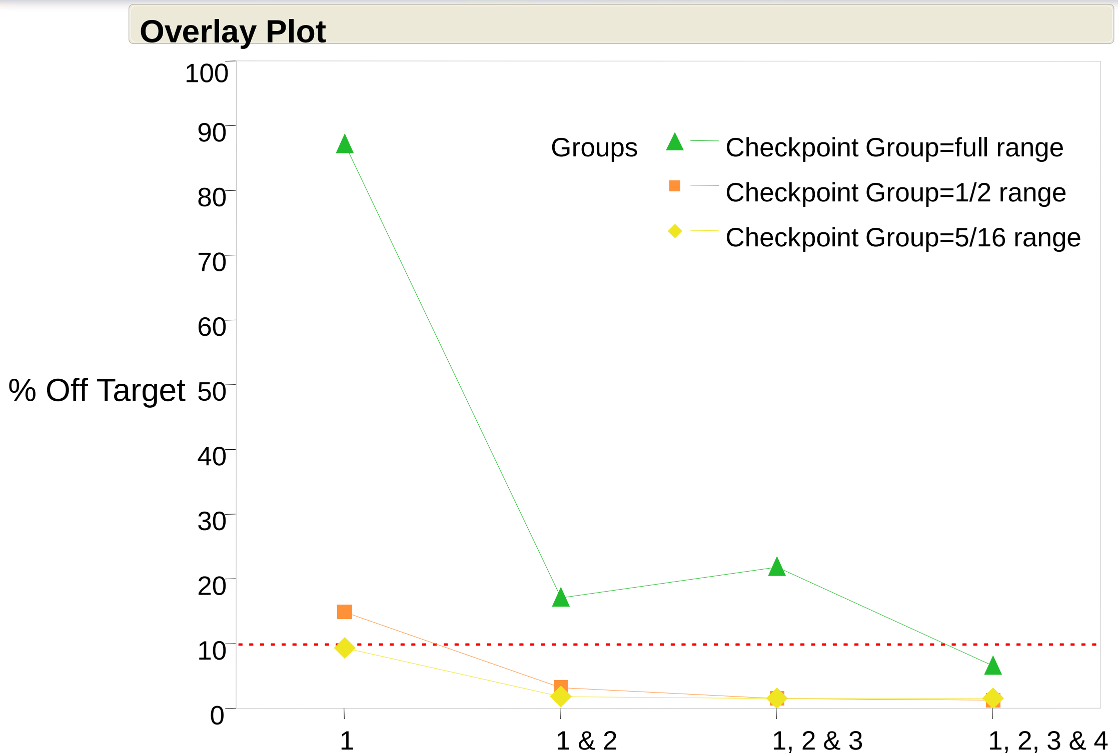


Plots of Actual vs. Predicted (Simulation vs. Surrogate) by Checkpoint Group for 4 Stages of Analysis of NLHD

Checkpoint Groups A & B show diminishing return in prediction improvement for running past stage 3



Accuracy of Surrogate Predictions for 3 Groups of Check-points Yielding Marginal, Moderate and Extreme Extrapolation

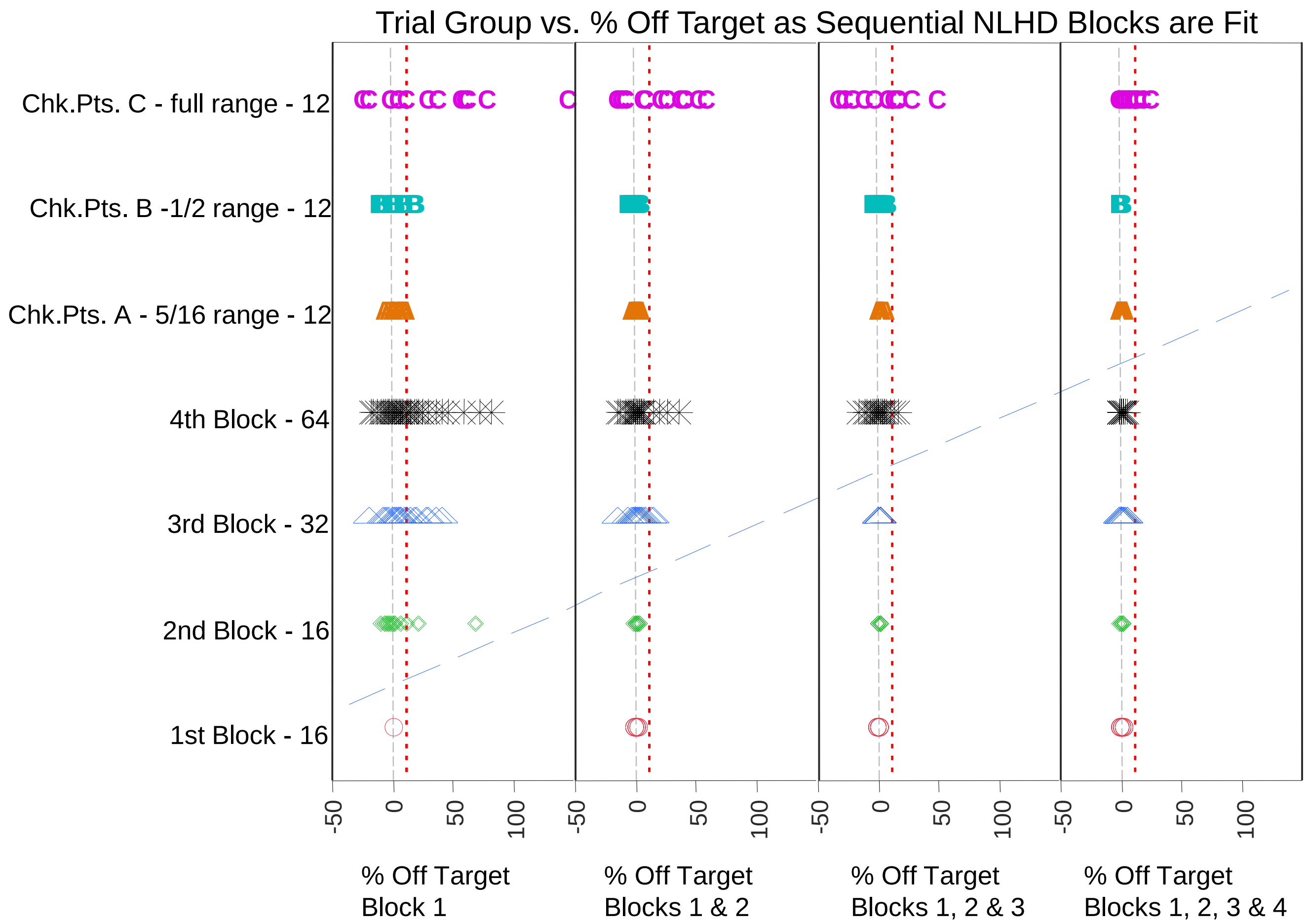


Percent Off Target - Root Mean Square of 12 Checkpoints				
Blocks	1	1 & 2	1, 2 & 3	1, 2, 3 & 4
5/16 range	9.39	2.08	1.72	1.53
1/2 range	14.94	3.33	1.79	1.27
full range	87.16	17.17	21.96	6.72

Percent Off Target - Worst Case of 12 Checkpoints				
Blocks	1	1 & 2	1, 2 & 3	1, 2, 3 & 4
5/16 range	17.13	4.52	3.48	2.74
1/2 range	33.74	7.11	-3.38	2.31
full range	225.70	34.69	46.98	16.66

Each checkpoint group consisted of a 12-trial Plackett-Burman DOE. The ranges of the factors relative to the ranges used for the NLHD were 5/16ths (marginal extrapolation), half (moderate extrapolation) and full (extreme extrapolation).

Accuracy of Surrogate Predictions for 3 Groups of Check-points Yielding Marginal, Moderate and Extreme Extrapolation



% Off Target for Checkpoints NOT Included in Model Fit

% Off Target for Points Fit with Gaussian Process Modeling

Conclusions
Sequential Space-Filling Designs

- NLHD designs can be run sequentially so that surrogate model accuracy can be evaluated after each block and decision made as to whether or not to move forward with the next block
- Generally as more NLHD blocks are run, the surrogate model accuracy increases
- Inclusion of extreme (full range) extrapolation checkpoints will expand interpolation volume of Kriging analysis – assuming Kriging analysis remains stable
- Caveat: These conclusions were reached using a moderately complex transcendental function in lieu of a CFD simulation model that is believed to do a good job of stressing extrapolation with the surrogate model..

Why Use Design of Experiments Methods with Simulation Experiments?

Quicker answers, lower costs, solve bigger problems

- Obtain a fast surrogate model of the simulation
  - Individual simulations can run for hours, days, weeks
    - Computational Fluid Dynamics (CFD)
    - Simulation runs in real-time
  - Numbers of factors can be very large (40+)
  - Numbers of simulations needed can be large (thousands in many cases)
  - Simulations can be stochastic requiring many replications
- Surrogate model yields a fast approximation of the simulation
  - more rapidly answer "what if?" questions
  - do sensitivity analysis of the control factors
  - optimize multiple responses and make trade-offs
- By running efficient subsets of all possible combinations, one can – for the same resources and constraints – *solve bigger problems*
- By running sequences of designs one can be as *cost effective as possible* & run *no more trials than are needed* to get a useful answer