Sensitivity Analysis using Experimental Design in Ballistic Missile Defense

Jacqueline K. Telford
Johns Hopkins University Applied Physics Laboratory
Laurel, Maryland
jacqueline.telford@jhuapl.edu

Introduction

Experimental design is used so that:

- valid results from a study are obtained
- with the maximum amount of information
- at a minimum of experimental material and labor (in our case, number of runs).

Poorly designed experiment	Well designed experiment
150 runs (30 design points, each repeated 5 times)	128 runs (all at different design points)
⇒ 11 estimates of effects	⇒ 66 estimates of effects
⇒ 11/150 = 7% efficiency	⇒ 66/128 = 52% efficiency

Historical Perspective

- **The fundamental principles of experimental design are due primarily to R. A. Fisher, who developed them from 1919 to 1930 in the planning of agricultural field experiments at the Rothamsted Experimental Station in England.**
  - Replication
  - Randomization
  - Blocking
  - Analysis Methods
  - Factorial Designs

"To call in the statistician after the experiment is done may be no more than asking him to perform a postmortem examination; he may be able to say what the experiment died of."

Sensitivity Analysis

Overall Goal

Provide a quantitative basis for assessing technology needs for missile defense architectures.

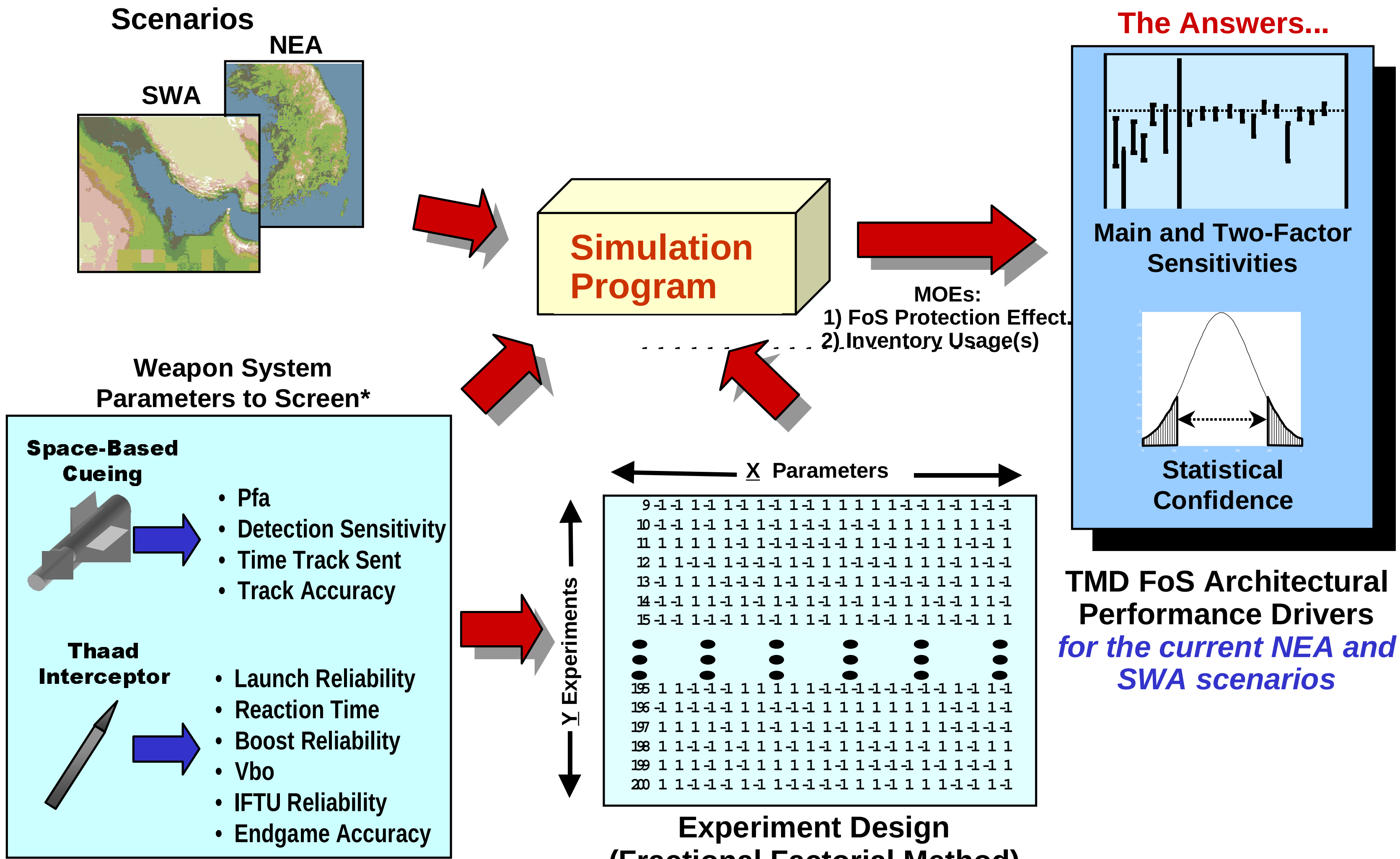
1. Screening Experiment

Use experimental design to identify the main performance drivers in the scenarios from among the many possible drivers.

2. Response Surface Experiment

Use experimental design to determine the shape (linear or curved) of the effects and interactions between the effects on the response variable for the main drivers to performance.

Process



9	-1	-1	1	-1	1	-1	1	-1	1	-1	1	1	1	1	1	-1	-1	1	-1	1	-1	-1
10	-1	-1	1	-1	1	-1	1	-1	1	-1	-1	1	-1	-1	1	1	1	1	1	1	1	-1
11	1	1	1	1	1	-1	1	-1	-1	-1	-1	-1	1	1	-1	1	-1	1	1	-1	-1	1
12	1	1	-1	-1	1	-1	-1	-1	1	-1	1	1	-1	-1	-1	1	1	1	-1	1	-1	-1
13	-1	1	1	1	-1	-1	-1	1	-1	1	-1	-1	1	1	-1	-1	-1	1	-1	1	1	-1
14	-1	-1	1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1	-1	1	1	-1
15	-1	-1	1	-1	-1	1	1	-1	1	1	-1	1	1	-1	-1	1	-1	1	-1	-1	1	1
⋮																						
195	1	1	-1	-1	-1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	1	-1	1	-1
196	-1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	1	1	1	1	1	1	-1	-1	1	-1
197	1	1	1	1	-1	1	1	-1	1	-1	1	-1	1	-1	1	1	1	1	-1	-1	-1	-1
198	1	1	-1	-1	1	-1	1	1	-1	1	-1	1	1	-1	-1	1	-1	1	1	-1	1	1
199	1	1	1	-1	1	-1	1	1	1	1	1	-1	1	-1	-1	-1	1	-1	1	-1	-1	1
200	1	1	-1	-1	-1	1	-1	1	-1	-1	-1	-1	1	1	-1	1	1	1	-1	-1	1	-1

Experiment Design (Fractional Factorial Method)

TMD FoS Architectural Performance Drivers ***for the current NEA and SWA scenarios***

*Note: Examples-only, not a complete list.

Polynomial Models for Sensitivity Analysis

- **Simple Additivity:**
  - $P.E. = b_o + \Sigma\, b_i X_i$ (i = 1,…, p factors)
  - X_i = -1 or +1 (coded values for the factor with a span wide enough that should result in a lower P.E. and a higher P.E. if there is an effect)

- **Two-way Interactions:**
  - $P.E. = b_o + \Sigma\, b_i X_i + \Sigma\, b_{ij} X_i X_j$ ($i \neq j$) <u>many</u> b_{ij} terms

- **Quadratic with two-way interactions:**
  - $P.E. = b_o + \Sigma\, b_i X_i + \Sigma\, b_{ij} X_i X_j + \Sigma\, b_{ii} X_i^2$
  - requires more than two levels for each factor

Factorial Designs

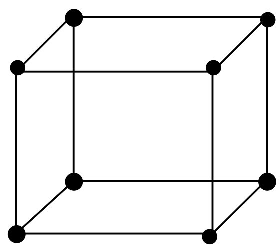
- **Varies many factors <u>simultaneously</u>, not the "change-one-variable-at-a-time" method.**
- **Checks for <u>interactions</u> (non-additivity) among factors.**
- **Shows the results over a <u>wider</u> variety of conditions.**
- **<u>Minimizes</u> the number of computer simulation runs for collecting information.**
- **<u>Built-in replication</u> for the factors to minimize variability due to random variables - usually no design point is replicated, all different points in the design matrix.**

Screening Designs

Full Factorial Design

(R. A. Fisher - 1926)

2^3 = 8 points

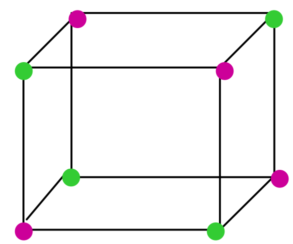


Fractional Factorial Design

(Yates/Cochran/Finney - 1930's)

2^{3-1} = 4 points in each 1/2 fraction

$\Downarrow$



Use either the Green or Purple points

Huge efficiencies for large numbers of dimensions, such as 2^{11-4}, 2^{47-35}, or $2^{121-113}$.

The "curse of dimensionality" is solved by fractional factorial designs.

Factorial Method: Full vs. Fractional

Full Factorial Design

- Varies P factors at two levels
- Requires 2^P computer runs
- If 47 factors ➡ 140 trillion runs !!
- Full information on:
  - main effects
  - two-way interactions
  - three-way, four-way, …, up to P- way interactions

Fractional Factorial Design

- Requires 2^{P-K} computer runs
- Only hundreds to thousands of computer runs for 47 factors
- Assumptions:
  - Monotonicity (not Linearity)
  - Few higher order interactions are significant
- The terms of the model may not be estimated separately, only the linear combinations of them

  ➡ Resolution Levels

Number of Runs Needed for Two-Level Fractional Factorials

Resolution Levels	
Resolution 2	Main effects (b_i) confounded with themselves.
Resolution 3	Main effects (b_i) not confounded with themselves, but with two-way effects (b_{ij}).
Resolution 4	Main effects (b_i) not confounded with two-factor effects, but two-way effects (b_{ij}) confounded with themselves.
Resolution 5	Main effects (b_i) and two-way effects (b_{ij}) not confounded with each other, but three-ways effects confounded with two-ways and four-ways with main effects.

Number of Runs Required for a Resolution 4 Fractional Factorial

Number of Factors	Minimum Number of Runs
1	2
2	4
3 – 4	$8 = 2^3$
5 – 8	$16 = 2^4$
9 – 16	$32 = 2^5$
17 – 32	$64 = 2^6$
33 – 64	$128 = 2^7$
65 – 128	$256 = 2^8$
129 – 256	$512 = 2^9$

Number of Runs Required for a Resolution 5 Fractional Factorial

Number of Factors	Minimum Number of Runs
1	2
2	4
3	$8 = 2^3$
4 – 5	$16 = 2^4$
6 – 7	$32 = 2^5$
8	$64 = 2^6$
9 – 11	$128 = 2^7$
12 – 17	$256 = 2^8$
18 – 22	$512 = 2^9$
23 – 31	$1{,}024 = 2^{10}$
32 – 40	$2{,}048 = 2^{11}$
41 – 54	$4{,}096 = 2^{12}$

Factors to be Screened

	Factors		Factors (continued)
1	Threat RCS	25	PAC III Reaction Time
2	SBIR Prob of Detection	26	PAC III Pk
3	SBIR Network Delay	27	PAC III Vbo
4	SBIR Accuracy	28	AEGIS Time to Acquire Track
5	SBIR Time to Form Track	29	AEGIS Time to Discriminate
6	THAAD Time to Acquire Track	30	AEGIS Time to Commit
7	THAAD Time to Discriminate	31	AEGIS Time to Kill Assessment
8	THAAD Time to Commit	32	AEGIS Prob of Correct Discrimination
9	THAAD Time to Kill Assessment	33	AEGIS Prob of Kill Assessment
10	THAAD Prob of Correct Discrimination	34	AEGIS Launch Reliability
11	THAAD Prob of Kill Assessment	35	AEGIS Reaction Time
12	THAAD Launch Reliability	36	AEGIS Pk
13	THAAD Reaction Time	37	AEGIS Vbo
14	THAAD Pk	38	Network Delay
15	THAAD Vbo	39	Lower Tier Minimum Intercept Altitude
16	PATRIOT Time to Acquire Track	40	Upper Tier Minimum Intercept Altitude
17	PATRIOT Time to Discriminate	41	ABL Reaction Time
18	PATRIOT Time to Commit	42	ABL Beam Spread
19	PATRIOT Prob of Correct Discrimination	43	ABL Atmospheric Attenuation
20	PAC II Launch Reliability	44	THAAD Downtime
21	PAC II Reaction Time	45	PATRIOT Downtime
22	PAC II Pk	46	AEGIS Downtime
23	PAC II Vbo	47	ABL Downtime
24	PAC III Launch Reliability		

Screening Designs Used

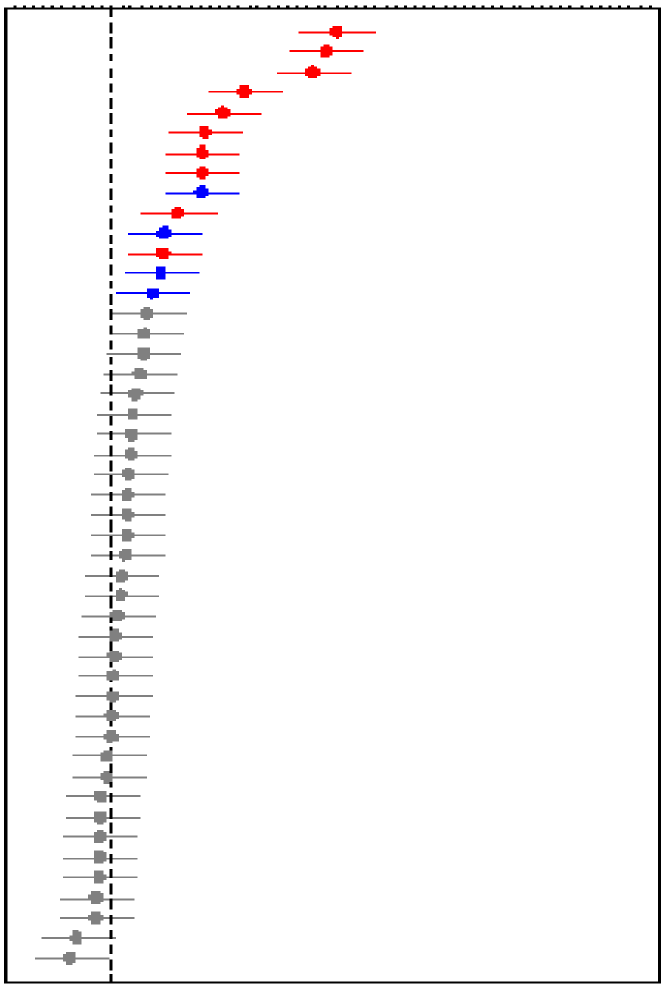
	Number Of Runs	Number of Two-Ways Estimated Separately	Resolution Level	Degrees of Freedom For Error
	128	0	4	17
	256	52	4.2	36
NEA ⇨	512	97	4.4	249
	1,024	146	4.6	712
	2,048	194	4.8	1,754
SWA ⇨	4,096	1,081 (all of them)	5	2,967

Selected a 2^{47-38} Resolution 4.4 Design (512 EADSIM runs) for NEA and a 2^{47-35} Resolution 5 Design (4,096 EADSIM runs) for SWA.

Approximately 350 additional runs were made for NEA to sort out combinations of two-way interactions: Recommend Resolution 5.

Main Effects Sensitivities for Protection Effectiveness

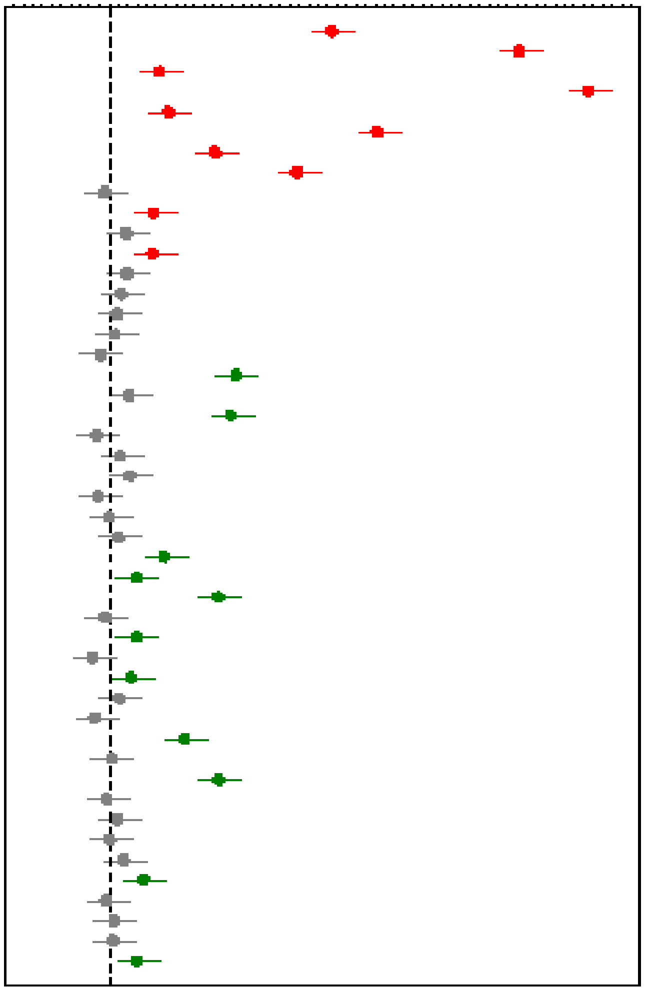
N.E. Asia



List of Factor Names Here

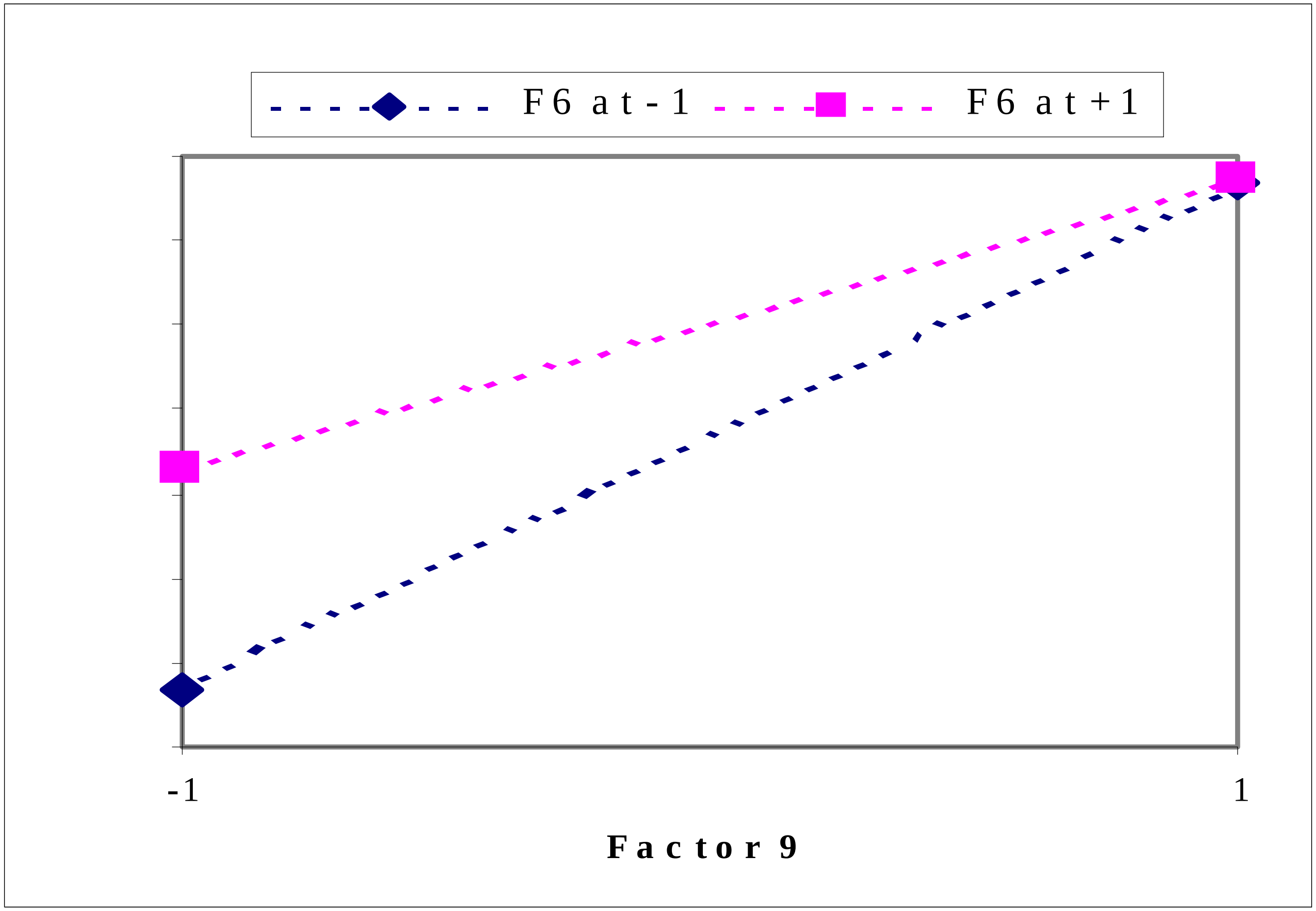
-
-
-
-
-
-
-
-
-

S.W. Asia



Significant in both theaters | Significant in NEA only | Significant in SWA only | Not significant

Two-way Interaction Result



Factor 6 and Factor 9 are not the same as in the table on Slide #11

Steps in Sensitivity Analysis

✓ 1. Screen a large number of factors at two levels each.

– Fractional Factorial Design (subset of all vertices of the p-dimensional hypercube)

– Resolution 5 if you can, otherwise Resolution 4

✓ 2. Determine important factors and combinations

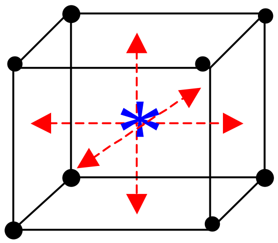
– Regression Analysis to estimate size of effects, test for statistical significance, and get confidence intervals

– 11 main effects were significant (and greater than a 1% P.E. effect), as well as several two-way interactions (which were combinations of significant main effects)

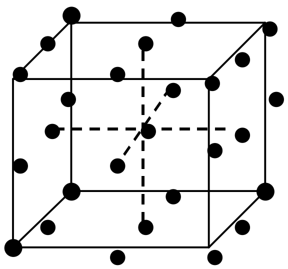
3. Establish a response surface using more than two levels of each important variable.

Response Surface Designs

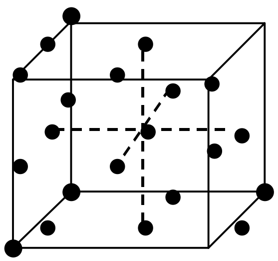
1. **Central Composite Design**
   – add points on the surfaces and at center to the two-level (fractional) factional.



2. **Three-Level Fractional Factorial Design**
   – three levels for each factor (-1, 0, +1) and uses a subset of the 3^P possible points.



3. **"Optimal" Design**
   – useful if:
   a. too many points needed for fractional factorial
   b. have an irregular design space



Number of Runs Needed for Three-Level Designs

Three-Level Fractional Factorial Resolution 5 Designs:

Number of Factors	Minimum Number of Runs
1	3
2	9
3	27
4 - 5	$81 = 3^4$
6 - 11	$243 = 3^5$
12 - 14	$729 = 3^6$
15 - 21	$2,187 = 3^7$
22 - ?	$6,561 = 3^8$

243 runs for 11 Factors

Central Composite Design:

10 replicates for each of the 22 faces of the hypercube plus 23 replicates of center of the cube (243 new design points).

D-Optimal Design:

Iterative search over 3^{11-1} (1/3 of total space) for 243 points to try to minimize $\det[(X'X)^{-1}]$, where X is the design matrix.

Comparisons among Three-Level Designs of 243 total design points

	3^{11-6} Fractional Factorial	Only Star and Center Points	D-optimal
$Det[(X'X)^{-1}]$	50	8 (no cross-products)	59
Standard Error:			
Main Effects	.0016	.0030	.0015
Two-way Interactions	.0019	--	.0016
Quadratic Effects	.0027	.0041	.0051

3^{P-K} design has the best balance for estimating effects.

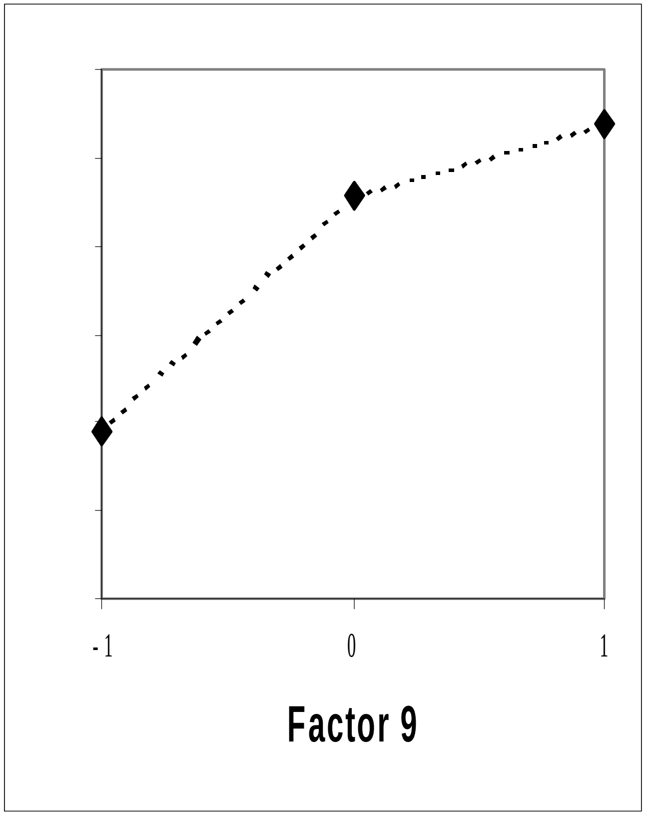
Results* for Response Surface Designs: Mature Theater/Force Level 4

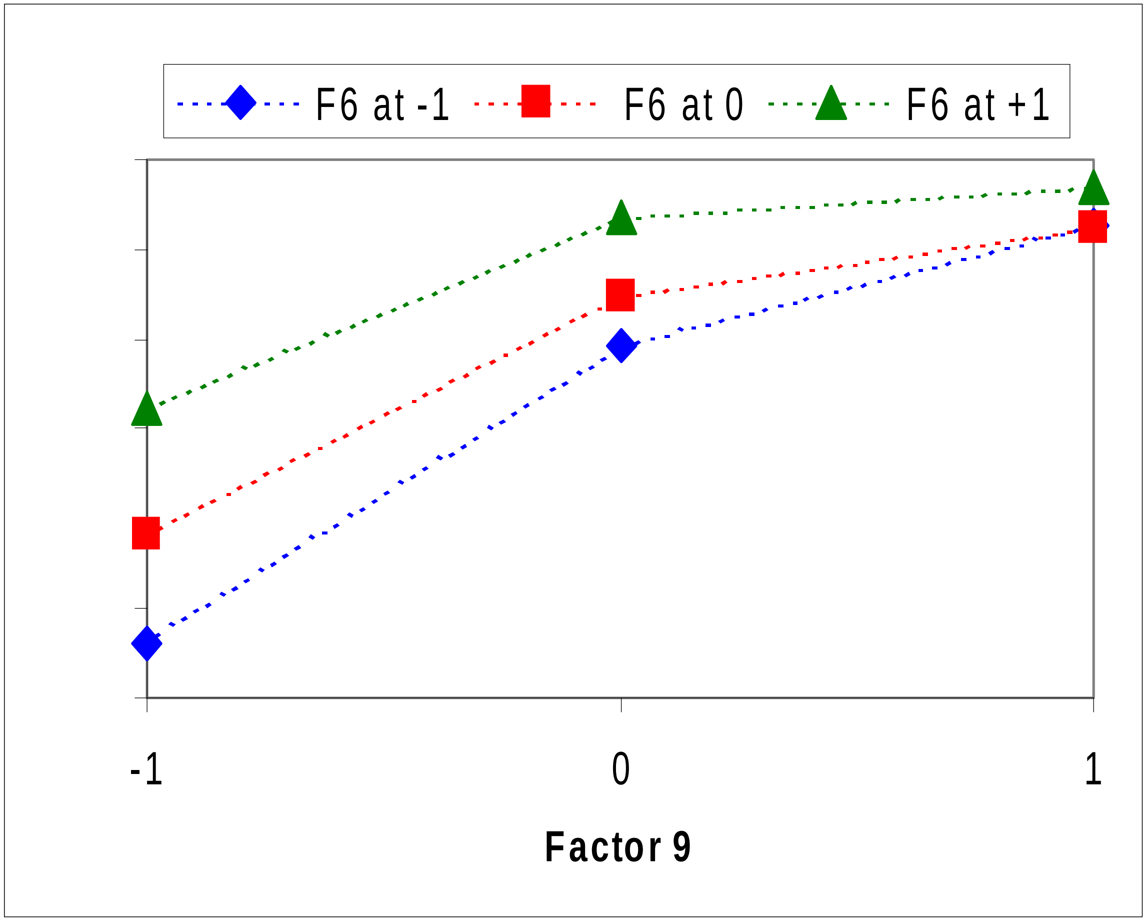
	3^{11-6} Fractional Factorial	Central Composite**	D-optimal
Main Effects	11	11	8
Two-way Interactions	7	5	8
Quadratic Effects	6	4	2

* Statistically significant at 5% level and effect > 1%.

** Includes 4,096 additional runs from two-level screening design.

Quadratic Effects and Two-Way Interactions at Three Levels





Factor 6 and Factor 9 are not the same as in the table on Slide #11

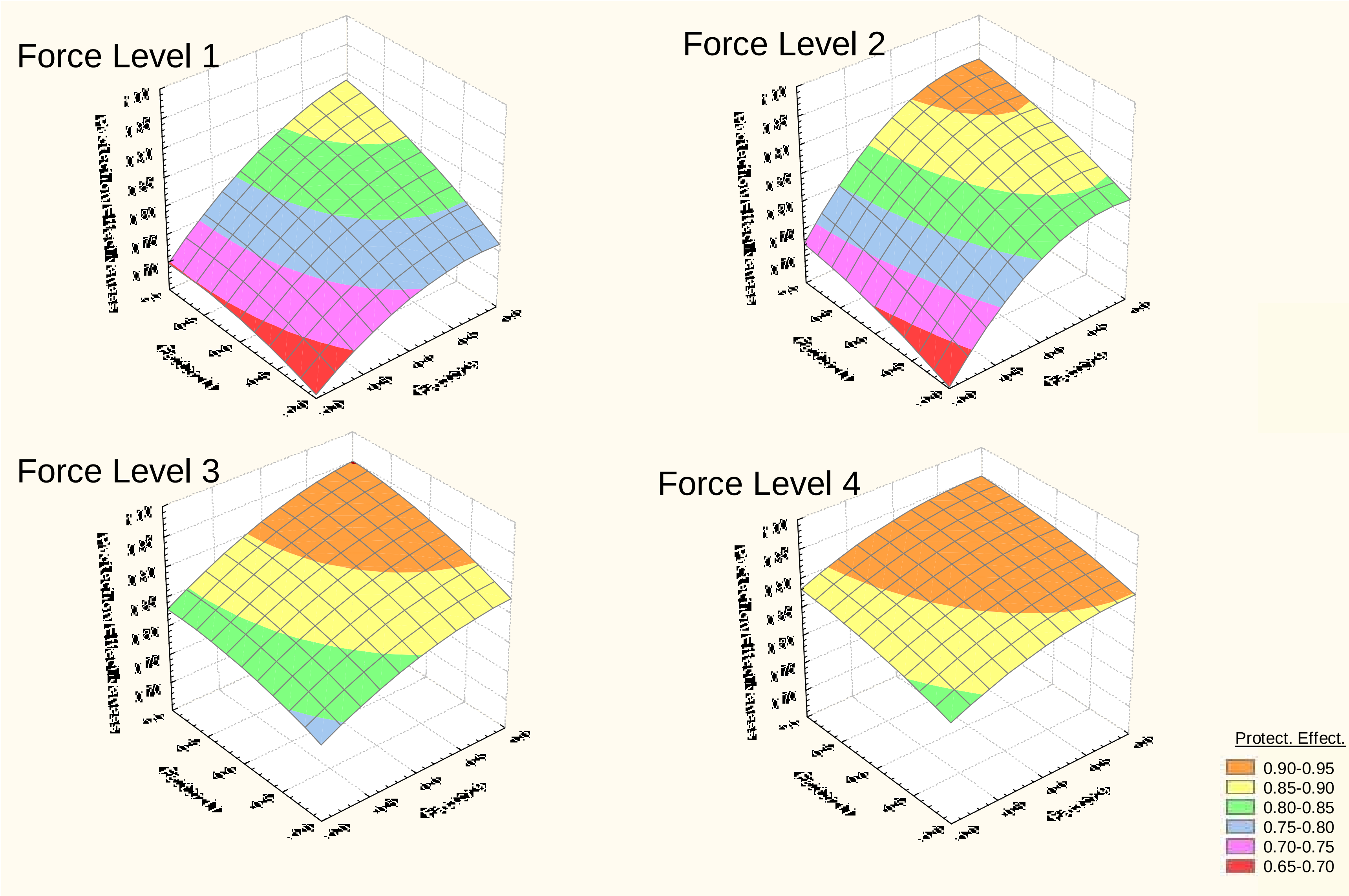
Fitted Model using 3^{11-6} Fractional Factorial Results

11 Factors were selected in the Screening Experiment (those color coded as red, blue, or green in the Main Sensitivities graph).

$$
\begin{aligned}
P.E. = .938 &+ .035X_9 + .026X_{11} + .017X_5 + .016X_2 \\
&+ .015X_6 + .014X_1 + .012X_7 + .011X_4 \\
&+ .007X_3 + .006X_8 \\
&- .011X_6X_9 - .007X_8X_9 - .007X_2X_5 - .006X_5X_7 \\
&- .005X_3X_9 + .005X_5X_6 - .005X_1X_5 \\
&- .019X_9^2 - .011X_5^2 - .009X_{11}^2 - .008X_4^2 \\
&- .006X_3^2 - .006X_2^2
\end{aligned}
$$

Effects are actually twice as large as coefficients since X_i = -1 and +1 (range of 2)

Response Surfaces by Force Levels: Factor 9 and Factor 11



Recommendations for a Sensitivity Analysis

1. **Screening Experiment:**

   Use Two-level Fractional Factorial design

   - **Resolution 5 if number of factors < 32 for 1,024 runs (if you can do more runs, you can have more factors)**
   - **Resolution 4.x otherwise**
   - **Replicates only at the center of the design [(0,0,0,…,0)] especially if no Response Surface as follow-on work**

2. **Response Surface:**

   Use Three-level Fractional Factorial design

   - **Resolution 5**

Resources

Textbooks:

Box, G.E.P., W.G. Hunter, and J.S. Hunter, *Statistics for Experimenters,* Wiley, 1978.

Montgomery, D. C., *Design and Analysis of Experiments*, Wiley, multiple editions.

Box, G.E.P. and N. R. Draper, *Empirical Model Building and Response Surfaces*, Wiley, 1987.

DO NOT USE Law and Kelton's fractional factorial design or analysis methods in *Simulation Modeling and Analysis* !

Software Used:

***SAS*, Version 8, for experimental design and analysis and for confidence interval graphs.**

***Statistica* for response surface graphs.**

Always Use A Designed Experiment!

"It is easy to conduct an experiment in such a way that no useful inferences can be made."

William G. Cochran and Gertrude M. Cox,
Experimental Designs, 1950