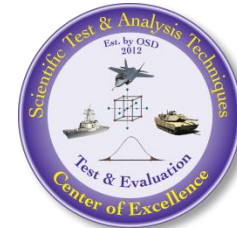# Assessment of radar detection performance using design of experiments; A case study

Luis A. Cortes, P.E.
Darryl Ahner, P.E., Ph.D.

OSD STAT COE
www.AFIT.edu/STAT
937-255-3636 x 4736

Scientia Prudentia et Valor

	Form Approved
Report Documentation Page	OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **2013**	2. REPORT TYPE	3. DATES COVERED **00-00-2013 to 00-00-2013**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Assessment of radar detection performance using design of experiments; A case study	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Air Force Institute of Technology,2950 Hobson Way,WPAFB,OH,45433**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**25**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

# Bottom Line Up Front

➤ This case study is an application of experimental design to the test and evaluation of surface radars.

➤ It builds upon work done by the Naval Surface Warfare Center, Corona Division.

➤ We look back into a test that was considered a landmark in M&S-based acquisition and contrast the way one objective was evaluated to the way it could have been evaluated with experimental design.

➤ In the process, we explore the attributes of a well designed test and demonstrate the utility of experimental design for planning, designing, executing, and analyzing a test.

➤ *What can we learn from the data? What could we have done differently? What can we do different next time?*

An experimental design approach contributes to
making the test more robust, efficient, and cost effective.
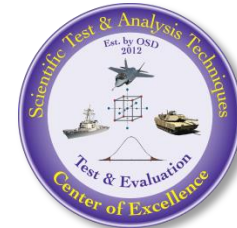
Tradeoff Space
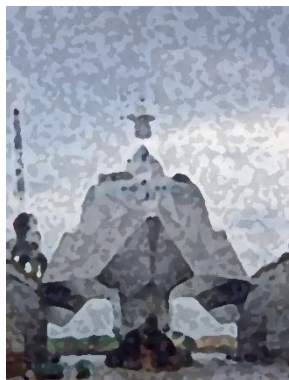
Risk of accepting a "bad" system

Adequate sample size

Effects on performance

Risk of rejecting a "good" system
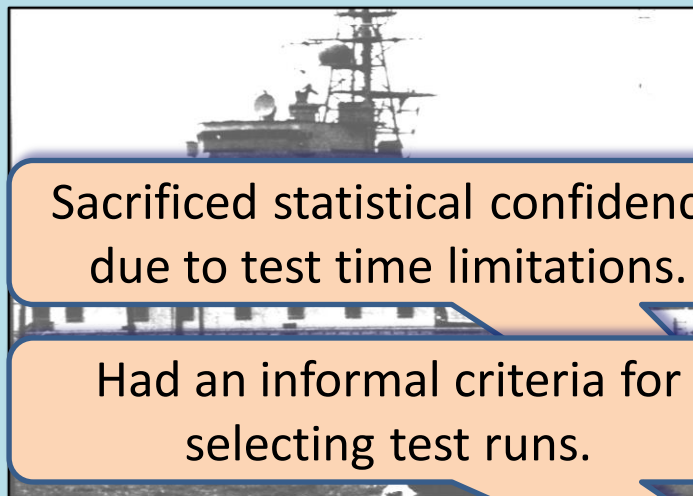
β

n

α

δ/σ

The *Central Problem of Test* is to determine the true nature of the system, in all possible scenarios, with a finite number of samples that yield valid conclusions while minimizing the risk of error.

3

Test Background

Test Operations

Sacrificed statistical confidence due to test time limitations.

Had an informal criteria for selecting test runs.

Analysis limited to pass/fail.

- Test events
  - 80 hrs total test time
  - 18 hrs of manned aircraft raids
  - 110 electronic attack (EA) techniques
  - 1900 simulated Anti Ship Cruise Missiles

One Objective

- ➤ Evaluate detection performance for a class of threat representative targets*

- ➤ Factors involved
  - Three target factors - A, B, F
  - Two environmental factors - C, D
  - One system factor - E

- ➤ Test strategy
  - 96 possible treatments
  - 30 samples per treatment required
  - 2880 total runs required
  - 96 hrs of test required-**not enough time!**
  - 670 runs conducted

- ➤ Assessment criteria - Pass/Fail

*Other objectives are beyond the scope of this brief; however, similar lessons apply.

Experimental Design Guidelines*

Plan

1. Formulate a clear statement of the problem.
2. Identify the proper responses to be analyzed.
3. Identify the factors and their levels.

Design

4. Choose an appropriate experimental design.

Execute

5. Perform the test as outlined in the test matrix.

Analyze

6. Perform the appropriate statistical data analysis.
7. Reach valid and practical recommendations.

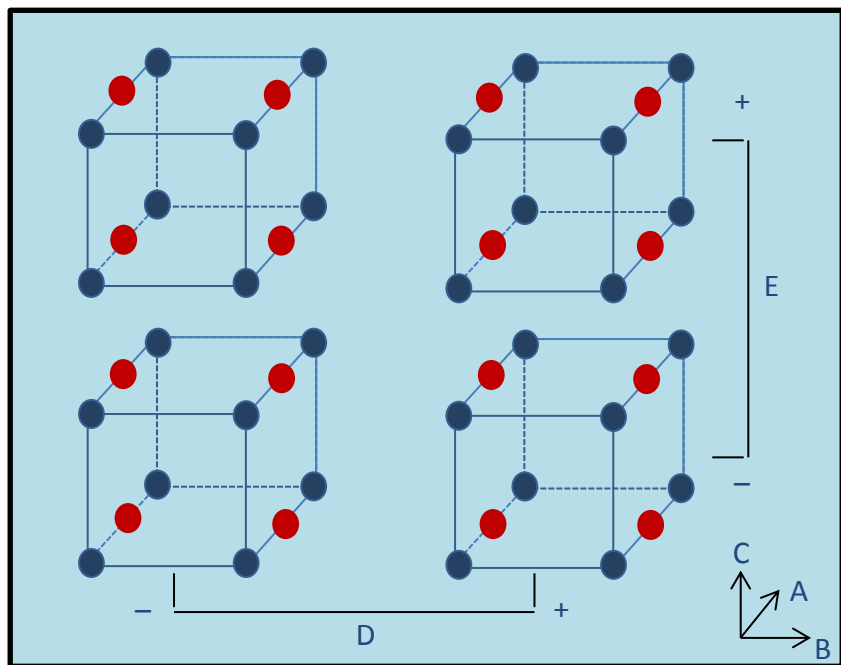* Montgomery, D. C. (2013), *Design and Analysis of Experiments,8th ed.,* John Wiley & Sons.

Experimental Design Approach
Step 1 - Problem Definition

Scientia Prudentia et Valor

Performance Requirements

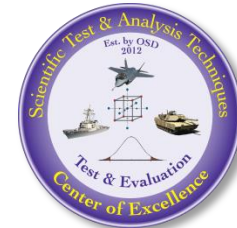## Statistical Parameters

Confidence level (α) - 0.05

Effect to detect (δ) – based on performance expectations

Variability (σ) - based on historical data

S/N (δ/σ) - 1.00 (for the case study)

We want to evaluate the effect of five factors* on detection performance.

*Six factors were of interest, but data for one factor was incomplete; therefore, the study was limited to five factors.
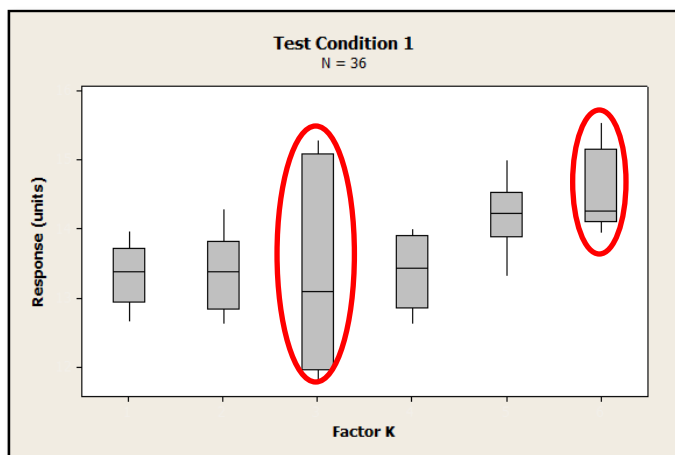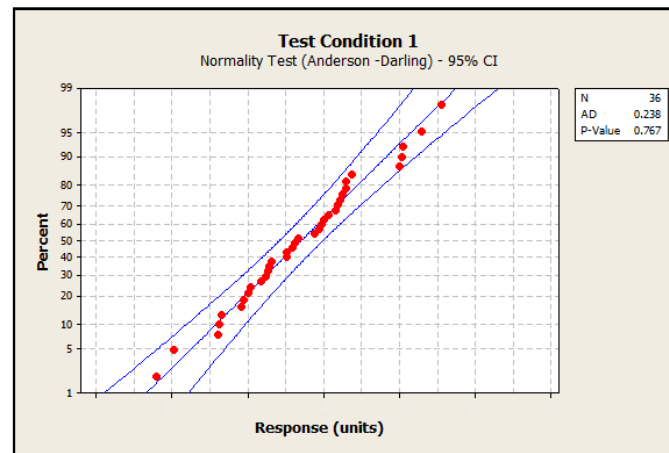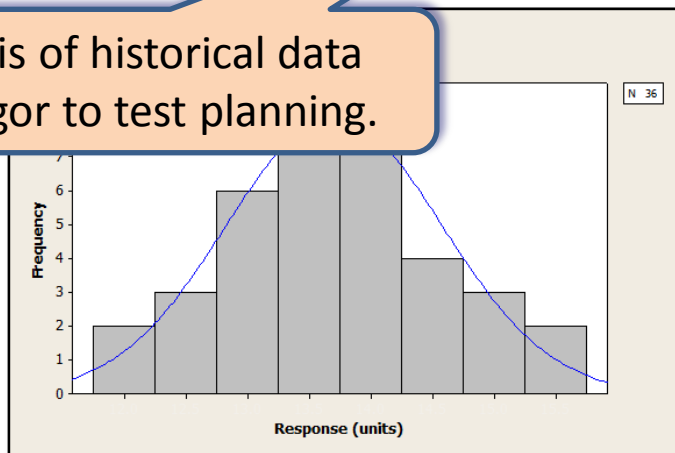
Scientia Prudentia et Valor

Analysis of historical data adds rigor to test planning.

Test Condition 1
Normality Test (Anderson -Darling) - 95% CI

N	36
AD	0.238
P-Value	0.767

Test Condition 1
N = 36

Test Condition 1
Paired T-test for Difference of Means $(K_i - K_j)$

	1	2	3	4	5	6
1	-	X	X	X	X	X
2	-	-	X	X	X	X
3	-	-	-	X	X	X
4	-	-	-	-	X	X
5	-	-	-	-	-	X
6	-	-	-	-	-	-

Alpha = 0.05; x = p-value < 0.05

7

Experimental Design Approach
Steps 2 & 3 - Responses and Factors

Continuous factors are preferred.

A single design can be used to evaluate multiple responses.

Factors
- A (Continuous, 3-levels)
- B (Categorical, 2-levels)
- C (Categorical, 2-levels)
- D* (Categorical, 2-levels)
- E* (Categorical, 2-levels)
- F^ (Fixed)

Covariates
Search Times, Power, Sensitivity, Material Readiness

Responses
- R1-Detection range
- R2-Transition-to-track range
- R3-Firm track range
- R4-Engagement Range
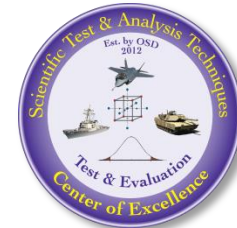
Select appropriate factors.

$$\text{Response} = f(\text{Factors}) + \varepsilon$$

Noise

Continuous responses are preferred.

* For large designs, these are hard(er)-to-change factors
^ Fixed during the original test due to test time limitations

Experimental Design Approach
Step 4 – Select a Design

Scientia Prudentia et Valor

Some Experimental Design Alternatives*
Completely Randomized Designs; Model – ME + 2FI; Power (1 std. dev.) at α = 0.05

Design	Runs	Center Points	Power (%) (ME)	VIF	DOF			Std. Error (FDS=0.8)
					Model	LOF	PE	
MR-Res IV	12	0	27-28	1.1 - 7.0	10	1	0	1.5
MR-Res IV	92	5	39-99	1.0 – 1.5	15	12	64	0.7
2_V^{5-1}	16	0	-	1.0	15	0	0	-
2_V^{5-1}	96	5	50–98	1.0	15	16	64	0.6
D-Optimal	21	0	54–57	1.1	15	5	0	1.0
2^5	32	0	76	1.0	15	16	0	0.7
2^5	112	5	> 80	1.0	15	32	64	0.5
2×2^5	64	0	-	1				
2×2^5	144	5	98-99	1				
4×2^5	192	4	99	1				

Evaluate several designs and select one that has good properties and that is appropriate for the problem.

Legend:
ME – main effects
DOF – degrees-of-freedom
LOF – lack-of-fit
FDS – fraction of the design space

2FI – two factor interactions
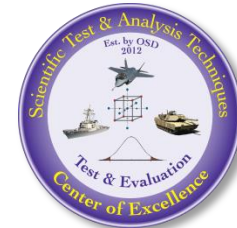VIF – variance inflation factor
PE – pure error

* Other designs were not explored due to data limitations.
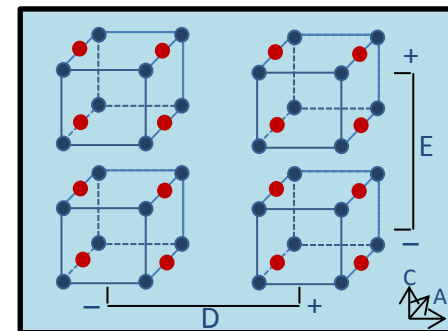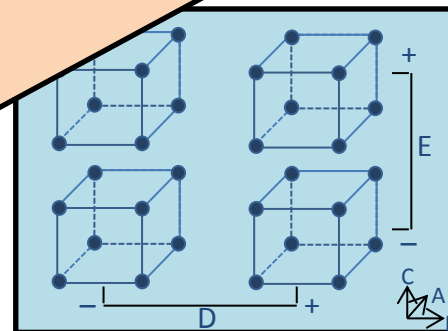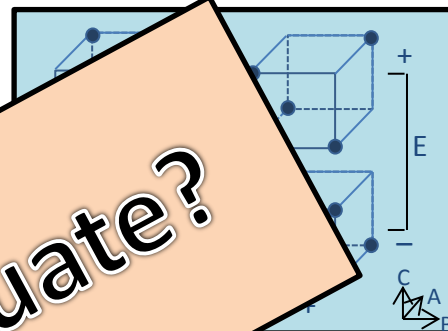
Scientia Prudentia et Valor

> ➤ Case I: 2_V^{5-1} fractional factorial
>   - 16 runs
>   - No degrees of freedom for estimating pure error, lack-of-fit, or test of significance

> ➤ Case II: 2^5 factorial
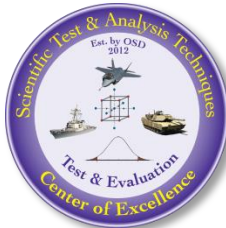>   - 32 runs
>   - No

> ... + center point – 4 reps.
>
> ... points allow testing for curvature and
> estimating pure error.

Are these designs adequate?

10

Scientia Prudentia et Valor

Run the test as specified in the test matrix.

Replicate, Randomize, and Block whenever possible.

...trix For Case III (2^5 factorial + ... point...

| Run Order	Std. Order	Block	A	B	C	D	E	y
96	**1**	**Blk1**	**1**	**B2**	**C2**	**D1**	**E1**	**xxx**
35	2	Blk1	-1	B1	C1	D2	E1	xxx
101	3	Blk1	1	B1	C2	D2	E1	xxx
46	4	Blk1	1	B2	C1	D2	E1	xxx
32	5	Blk1	1	B2	C2	D1	E1	xxx
107	6	Blk1	-1	B2	C1	D2	E2	xxx
89	7	Blk1	0	B1	C2	D2	E1	xxx
....	....	....	....	....	....	....	....	....
56	**n**	**Blk1**	**1**	**B2**	**C2**	**D1**	**E1**	**xxx**
....	....	....	....	....	....	....	....	....
192	192	Blk2	-1	B1	C1	D1	E1	xxx

11
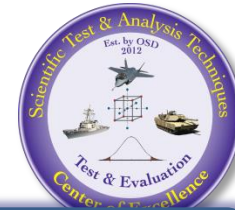
Scientia Prudenti...

Factor D is not significant—a model is suspect.

Experimental design affords studying interactions.

Analysis of Variance Table for Case I (2_V^{5-1} Fractional Factorial); p-value <0.1

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	167.26	5	33.45	25.70	< 0.0001	significant
A-A	66.59	1	66.59	51.15	< 0.0001	
B-B	65.29	1	65.29	50.15	< 0.0001	
C-C	26.37	1	26.37	20.26	0.0011	
D-D	1.70	1	1.70	1.31	0.2794	
BD	7.32	1	7.32	5.62	0.0392	
Residual	13.02	10	1.30			
Cor Total	180.28	15				

Reduced Empirical Model (Coded Factors)
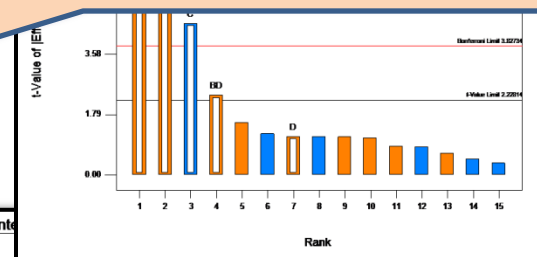
$$R1 = I + x_1A + x_2B - x_3C + x_4D + x_{24}BD$$

$R^2 =$ ___ 0.8917   Pred. $R^2 = 0.8151$ Adeq. Precision = 17.23

The empirical model is useful for tactical decision aids, training, and performance assessment.
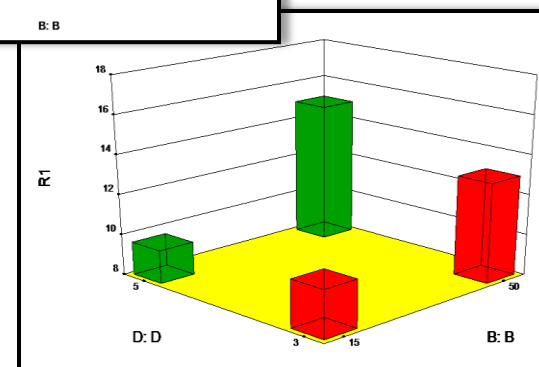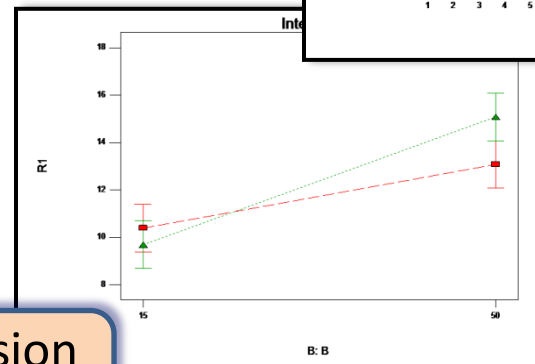
Reference Mechanistic Model

$$R = \left(\frac{1}{(S/N)_t} \times \frac{P_t \times G \times \lambda^2}{(4\pi)^3 \times k} \times \frac{\sigma \times F^2}{L_t \times L_r \times L_{bs}^2 \times L_a^2 \times L_s^2} \times \frac{1}{T_s} \times \frac{\tau \times N}{L_p} \right)^{1/4}$$
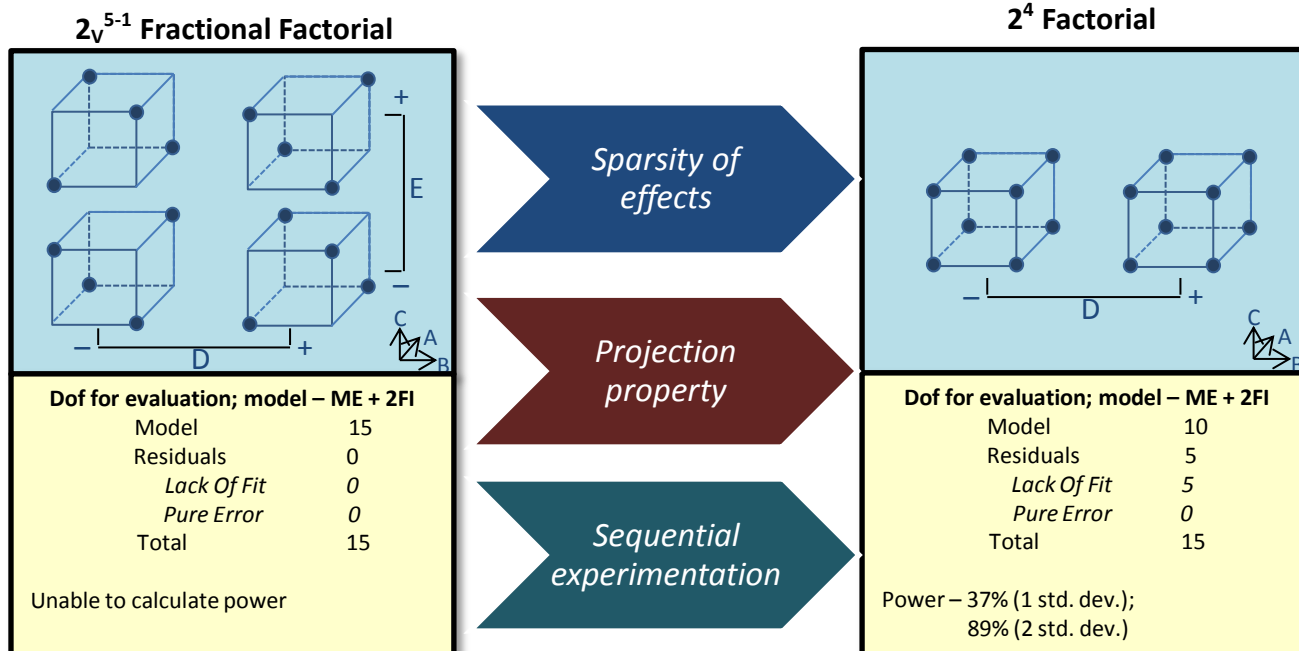
12

Scientia Prudentia et Valor
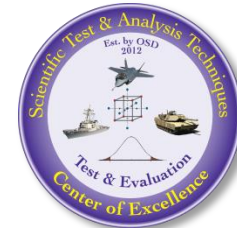
- ➢ Only factors A, B, and C and interaction BD are significant; factor E is dropped from consideration—*the sparsity of effects principle*.
- ➢ A Res V fractional factorial design contains a complete factorial in any subset of 4 factors—*the projection property*.
- ➢ We can combine the runs of fractional factorials to assemble a larger design (two blocks)—*sequential experimentation*
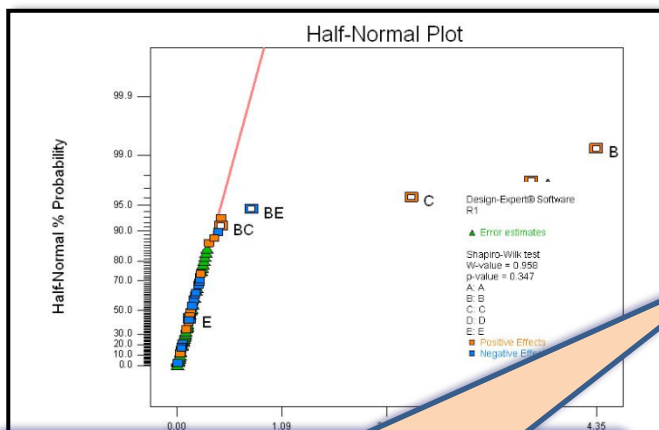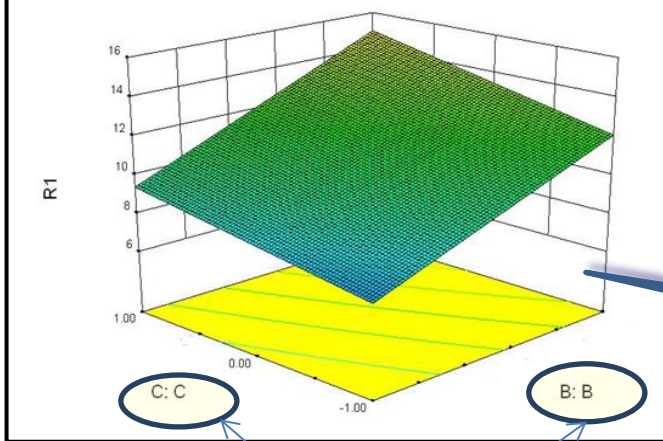

2_V^{5-1} Fractional Factorial

Dof for evaluation; model – ME + 2FI	
Model	15
Residuals	0
Lack Of Fit	*0*
Pure Error	*0*
Total	15

Unable to calculate power

Sparsity of effects

Projection property

Sequential experimentation

2^4 Factorial

Dof for evaluation; model – ME + 2FI	
Model	10
Residuals	5
Lack Of Fit	*5*
Pure Error	*0*
Total	15

Power – 37% (1 std. dev.);
89% (2 std. dev.)

13

Scientia Prudentia et Valor

Half-Normal Plot

Analysis of Variance Table for Case I (2_V^{5-1} Fractional Factorial); p-value <0.1

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	627.68	6	104.61	139.00	< 0.0001	significant
A	216.38	1	216.38	287.51	< 0.0001	
B	303.20	1	303.20	402.85	< 0.0001	
C	94.92	1	94.92	126.11	< 0.0001	
E	0.27	1	0.27	0.36	0.5513	
BC	3.35	1	3.35	4.45	0.0393	
BE	9.56	1	9.56	12.71	0.0007	
Residual	42.90	57	0.75			
Lack of Fit	19.01	25	0.76	1.0	0.4742	not significant
Pure Error	23.89	32	0.75			
Cor Total	670.58	63				

Factor D has no significant effect on the response.

F-values consistent with complete randomization.

Reduced Empirical Model (Coded Factors)

$$R = I + x_1A + x_2B + x_3C - x_4E + x_{23}BC - x_{25}BE$$

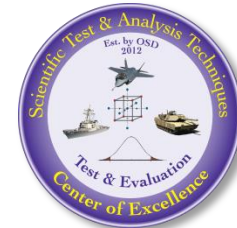$R^2 = 0.9360$ Adj. $R^2 = 0.9293$ Pred. $R^2 = 0.9193$ Adeq. Precision = 39.2
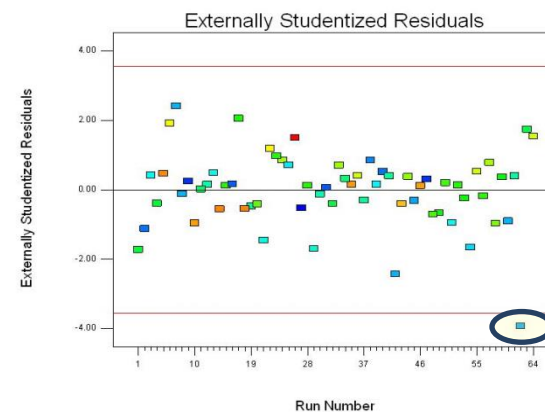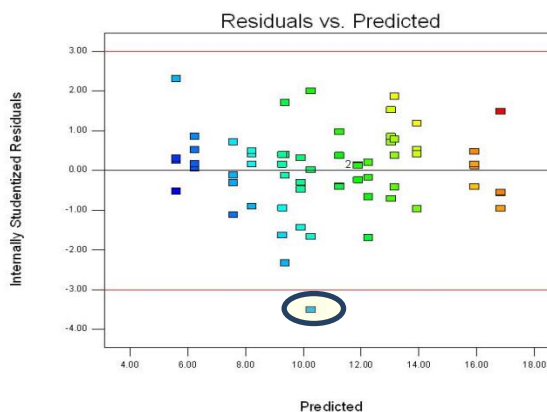
Continuous factors yield response surfaces.

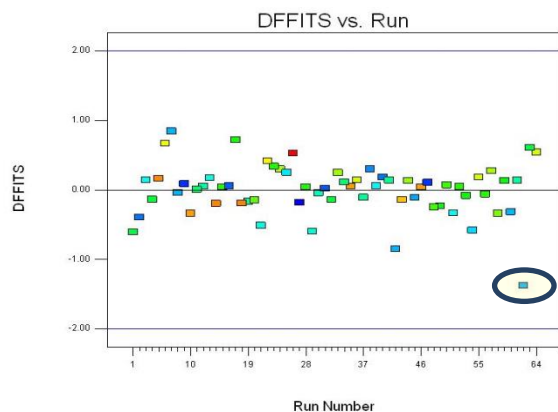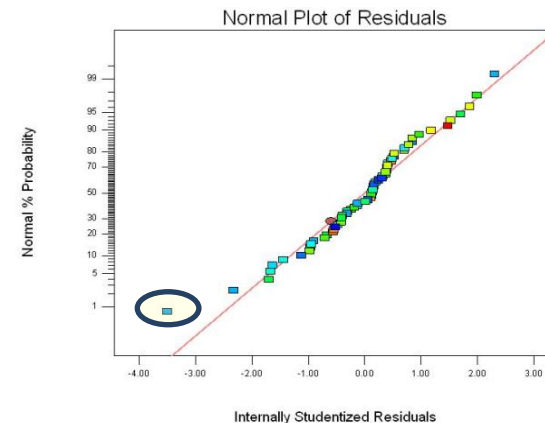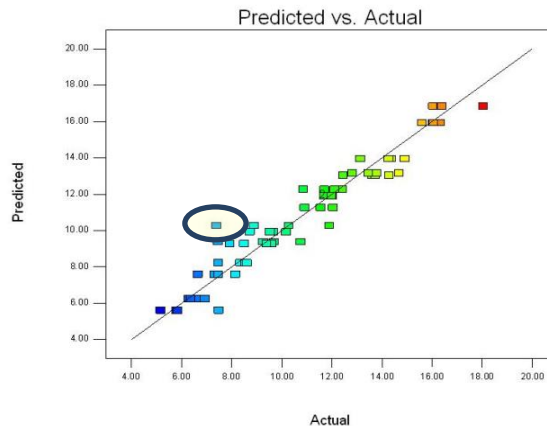Continuous

Scientia Prudentia et Valor

Design

2^5 Factorial
2 Replicates
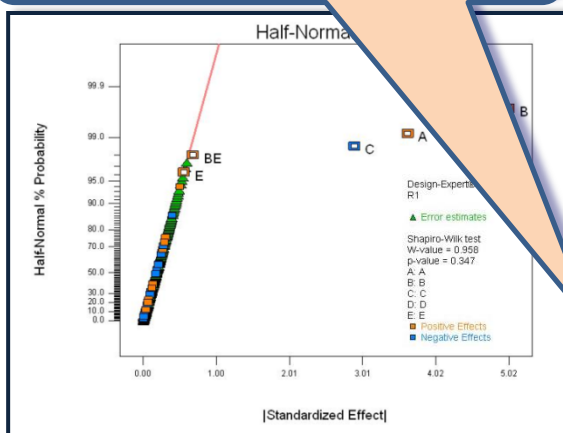
Validating the data and the statistical assumptions.

Scientia Prudentia et Valor

Factors D and E are not significant.

Analysis of Variance Table for Case III (4 x 2^5 Factorial + Center Points); p-value < 0.1

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	1524.59	5	304.92	141.81	< 0.0001	significant
A	423.11	1	423.11	196.78	< 0.0001	
B	636.09	1	636.09	295.83	< 0.0001	
C	191.10	1	191.10	88.88	< 0.0001	
E	2.90	1	2.90	1.35	0.2467	
BE	15.04	1	15.04	7.00	0.0089	
Curvature	22.76	8	2.84	1.32	0.2346	not significant
Residual	382.73	178	2.15			
Lack of Fit	52.47	34	1.54	0.67	0.9112	not significant
Pure Error	330.26	144	2.29			
Cor Total	1930.08	191				
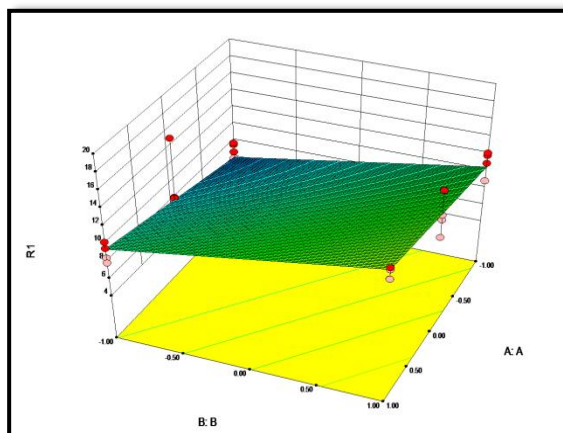
Reduced Empirical Model (Adjusted, Coded Factors)

$$R = I + x_1A + x_2B - x_3C + x5E + x_{25}BE$$

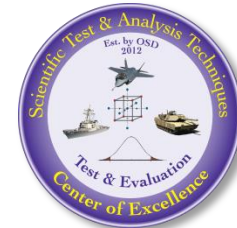$R^2 = 0.7899$ Adj. $R^2 = 0.7843$ Pred. $R^2 = 0.7775$ Adeq. Precision = 40.9
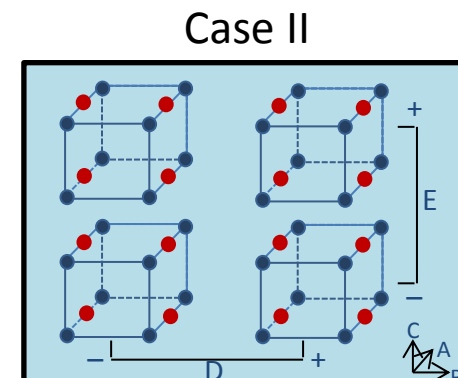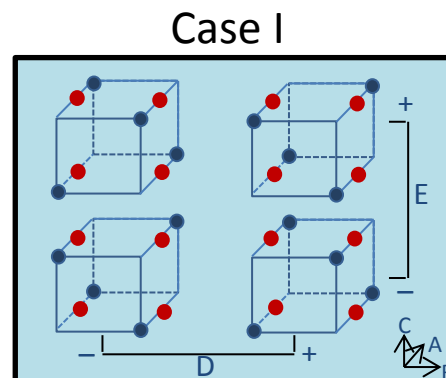
* Ref: Design Expert 8.0.7.1

Experimental Design Approach
Step 6 - Confirmation

Save a few runs for confirmation.

Empirical Error

A	B	C	E	Case I	Case II
0	-1	-1	-1	13.8	5.1
0	-1	-1	1	10.9	0.7
0	-1	1	-1	9.0	3.0
0	-1	1	1	0.2	1.6
0	1	-1	-1	7.1	1.6
0	1	-1	1	4.3	1.4
0	1	1	-1	13.6	0.6
0	1	1	1	3.2	2.0
Average				7.8	2.8

Case I   Case II

- ➤ Factor A was a 3-level factor.
- ➤ The designs for Case I and Case II used only the high and low settings (in blue), and not the center points (in red).
- ➤ The center points were used for confirmation.
- ➤ The Empirical Error is the difference between the average (5 runs) at the center points and the respective model predictions for those factor settings.
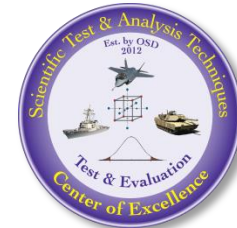
" All models are wrong, but some are useful. "

George Box
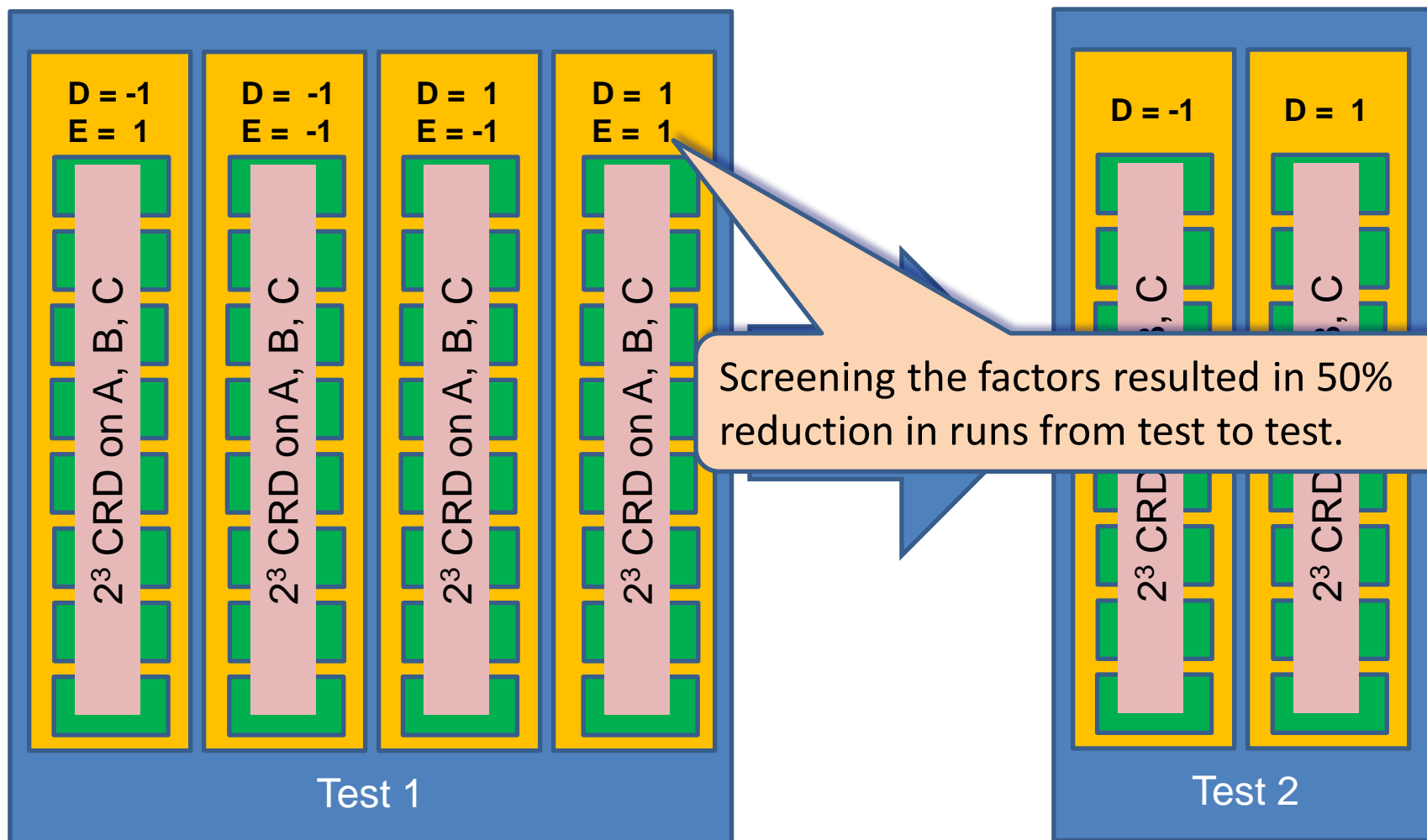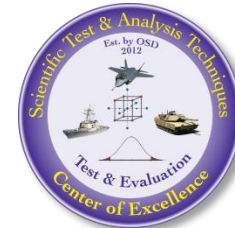
17

Scientia Prudentia et Valor

Test 1

D = -1, E = 1 — 2^3 CRD on A, B, C
D = -1, E = -1 — 2^3 CRD on A, B, C
D = 1, E = -1 — 2^3 CRD on A, B, C
D = 1, E = 1 — 2^3 CRD on A, B, C

Test 2

D = -1 — 2^3 CRD on A, B, C
D = 1 — 2^3 CRD on A, B, C

Screening the factors resulted in 50% reduction in runs from test to test.

18

Scientia Prudentia et Valor

Half-Normal Plot
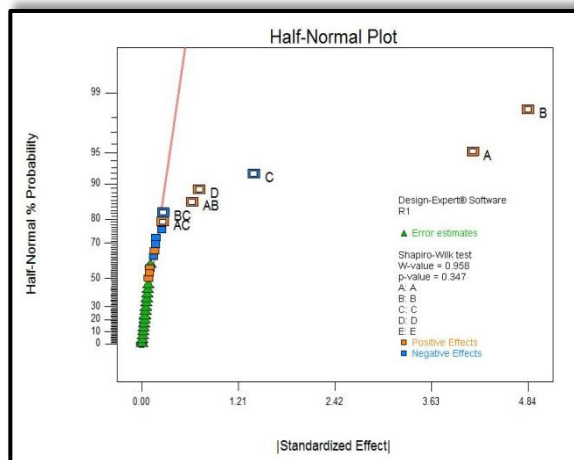
Analysis of Variance Table for Test 2 (2 x 2⁴ Factorial); p-value < 0.1

| Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	350.07	7	50.01	310.41	< 0.0001	significant
A	137.99	1	137.99	856.50	< 0.0001	
B	187.65	1	187.65	1164.74	< 0.0001	
C	15.83	1				
D						
AB	3.23	1	3.23	20.06	0.0002	
AC	0.58	1	0.58	3.57	0.0710	
BC	0.59	1	0.59	3.67	0.0674	
Residual	3.87	24	0.16			
Lack of fit	1.68	8	0.21	1.53	0.2231	not significant
Pure Error	2.19	16	0.14			
Cor Total	353.93	31				

Factors D now is significant.

Detect Range
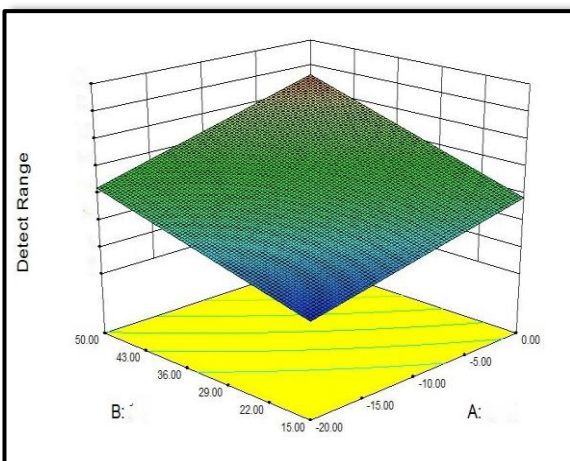
Reduced Empirical Model (Coded Factors)

$$R = I + x_1A + x_2B - x_3C + x_4D + x_{12}AB + x_{13}AC - x_{23}BC$$

$R^2 = 0.9891$ Adj. $R^2 = 0.9859$ Pred. $R^2 = 0.9806$ Adeq. Precision = 55.5

* Ref: Design Expert 8.0.7.1

Scientia Prudentia et Valor

ANOVA Table for the split-plot experiment (Test 1 vs. Test 2)

Source	dof	SS	MS	F_0	F	Significant
R (Test)	1	16.02	16.02	330.90	10.13	Significant
WP Error	2	0.10	0.05	0.35	*	
A	1	119.74	119.74	872.03	4.38	Significant
B	1	147.15	147.15	1071.65	4.38	Significant
C	1	27.08	27.08	197.25	4.38	Significant
AB	1	3.45	3.45	25.09	4.38	Significant
AC	1	0.00	0.00	0.03	4.38	Not significant
BC	1	1.16	1.16	8.41	4.38	Significant
RA	1	0.26	0.26	1.89	4.38	Not significant
RB	1	1.73	1.73	12.60	4.38	Significant
RC	1	2.99	2.99	21.77	4.38	Significant
SP Error	19	2.61	0.14	*	*	*
	31	322.28				

A split-plot design was used to compare detection performance between the tests. Factor R is significant—there is a difference between the radar systems.

20

Summary

➤ Experimental design is the integration of well defined and structured scientific strategies for gathering empirical knowledge using statistical methods for planning, designing, executing, and analyzing a test.

➤ Experimental design provides a comprehensive understanding of the tradeoffs in the techno-programmatic domains: risks, cost, and utility of information.

➤ Experimental design can help reducing test assets, shortening the test schedule, and providing more information to the warfighter and decision makers.
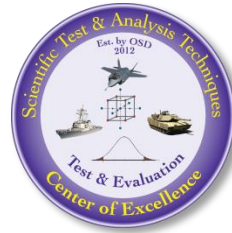
➤ Experimental design adds rigor and discipline to T&E.

Conclusions

What could we have done differently?

Only 16 runs

+ center points

+ axial points

(may be)

# References

➢ Bisgaard, S., Fuller, H. T., & Barrios, E. (1996). Quality quandaries: Two-level factorials run as split-plot experiments. *Quality Engineering*, 8, 705-708

➢ Coleman, D. D., & Montgomery, D. C. (1993). A systematic approach to planning for a designed industrial experiment. *Technometrics*, 35, 1-12.

➢ Cortes, L. A., & Bergstrom, D. (2012). Using design of experiments for the assessment of surface radar detection performance and compliance with critical technical parameters. *80th MORS Symposium*, June 2012.

➢ Fisher, R. A. (1962). The place of the design of experiments in the logic of scientific inference. C*olloques Internationaux du Centre National de la Recherche Scientifique,* 110, 13-19.

➢ Freeman, L. J., Ryan, A. G., Kensler, J. L. K., Dickinson, R. M., & Vining, G. G. (2013). A tutorial on the planning of experiments. *Quality Engineering*, 25, 315-332.

➢ Johnson, R. T., Hutto, G. T., Simpson, J. R., & Montgomery, D. C. (2012). Designed experiments for the defense community, *Quality Engineering*, 24, 60-79.

➢ Montgomery, D. C. (2013), *Design and Analysis of Experiments,8th ed.,* John Wiley & Sons.

➢ Simpson, J. R., Listak, C. M., & Hutto, G. T. (2013). Guidelines for planning and evidence for assessing a well-designed experiment. *Quality Engineering*, 25, 316-355.

STAT T&E Center of Excellence
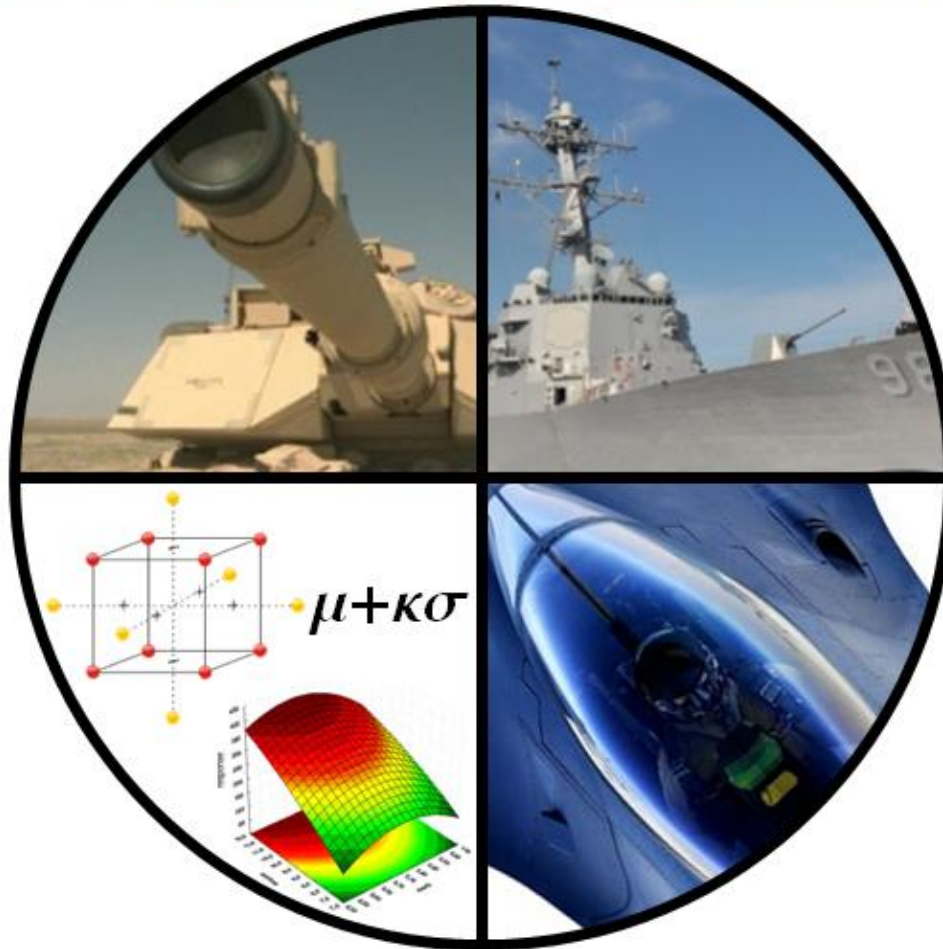
Scientia Prudentia et Valor

Knowledge, insight, and value