LA-UR- 11-06836

Approved for public release; distribution is unlimited.

Title: Visualizing Trade-offs Between Multiple Objectives: Tools to Help Decision-Makers

Author(s): Christine Anderson-Cook

Intended for: Taiwan International Statistical Symposium
Taipei, Taiwan
December 16-19, 2011

Los Alamos
NATIONAL LABORATORY
EST. 1943

Title: Visualizing Trade-offs Between Multiple Objectives: Tools to Help Decision-Makers

Invited Speaker: Christine M. Anderson-Cook, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

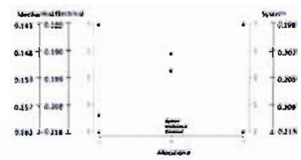
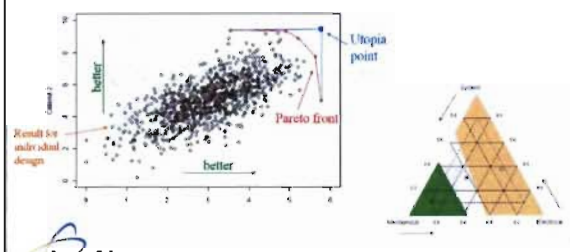
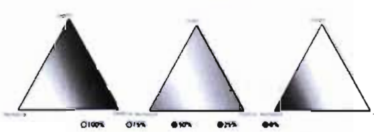
Abstract:

When decision-makers make important decisions, they are often forced to balance competing objectives that require weighing the importance of different alternatives and assessing the merits of different choices. We present a suite of graphical tools, based on the Pareto front multiple criteria optimization method, which allow the trade-offs between choices to be compared and assessed. The tools are presented in the context of two examples: a designed experiment where good estimation and protection against model misspecification are considered; and a resource allocation problem about what future data to collect when evaluating reliability for a population of systems based on several different data types.

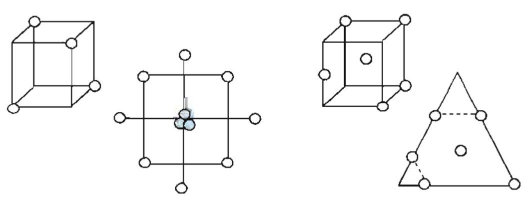
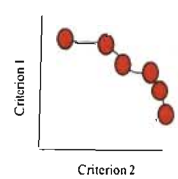
LA-UR 11-04665 (abstract)

Visualizing Trade-offs Between Multiple Objectives: Tools to Help Decision-Makers

Christine Anderson-Cook, PhD
Los Alamos National Laboratory
December 2011

A Very Brief History of Design of Experiments

Textbook designs
- only some N
- regular regions
- good general performance

"Optimal" designs
- flexible N, region, criterion

"Multiple Criteria" designs
- flexible N, region
- consider multiple objectives
- Pareto front based

Computer Power increasing dramatically

Two Non-Standard Design Problems

- Problem 1: Run a 14-run designed experiment to estimate the model:

$$Y = \beta_0 + \beta_A A + \beta_B B + \beta_C C + \beta_D D + \beta_E E + \beta_{AB} AB + \beta_{AC} AC + \beta_{BD} BD + \beta_{CE} CE + \varepsilon \qquad \varepsilon \sim N(0,\sigma^2)$$

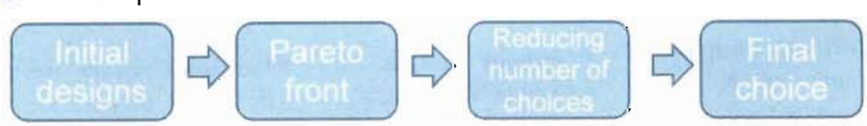
  We are worried that some of the other 2-factor interactions (AD,AE,BC,BE,CD,DE) might be active

- Problem 2: We have already collected data to estimate system reliability using component and system data. Now we can collect more data, and want to leverage current understanding of the system to guide this choice

- What design should we run?
- What basis should we use for choosing?

Outline

1. Motivation – why should we consider more than one objective during design construction and selection?
2. Pareto front approach (2 criteria)
3. Example 1 revisited



4. Example 2 revisited
5. Conclusions

UNCLASSIFIED

Metrics for Good Designs

1. Result in good fit of the model to the data
2. Provide good model parameter estimates
3. Provide good prediction throughout the design space.
4. Provide an estimate of "pure" experimental error.

Good estimation and prediction for chosen model

5. Give sufficient information to allow for lack of fit test.
6. Provide a check on the homogeneous variance assumption.

Ability to test various aspects of the model

7. Be insensitive (robust) to the presence of outliers in the data.
8. Be robust to to errors in the control of design levels.

Protection if things go wrong

9. Allow models of increasing order to be constructed sequentially.
10. Allow for experiments to be done in blocks.

Flexibility to run and expand experiment

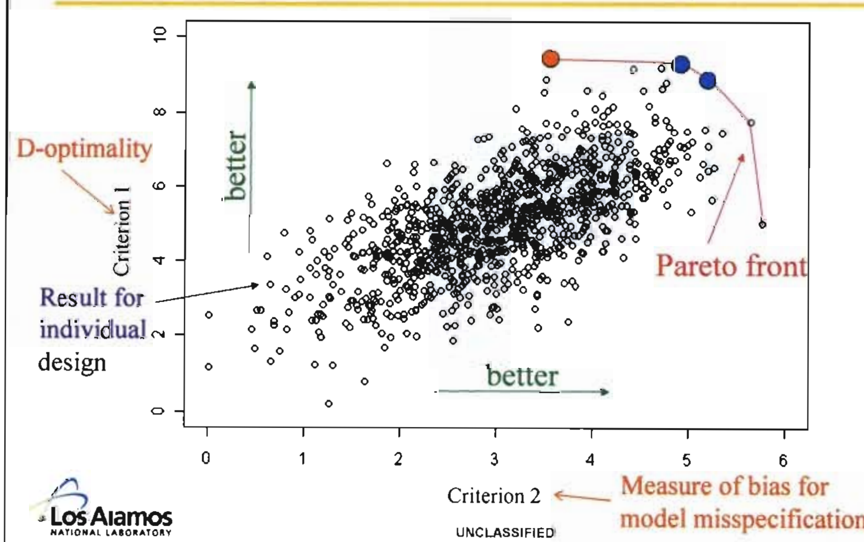
11. Be cost-effective.

Cost

Myers, Montgomery, Anderson-Cook RSM (2009) p. 282

UNCLASSIFIED

The Weakness of Single Criterion Optimization



Los Alamos NATIONAL LABORATORY EST. 1943

UNCLASSIFIED

Slide 5

Operated by the Los Alamos National Security, LLC for the DOE/NNSA

NNSA

For Example 1Design

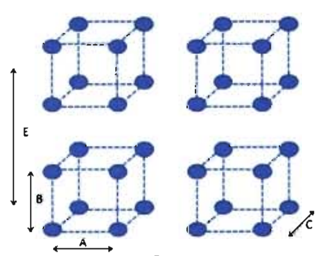
- If the model is correct:
  - Good estimation of model parameters
- If the model is incorrect (some of AD,AE,BC,BE,CD,DE active)
  - Estimates for terms in model minimally affected
  - Estimation of variance minimally affected

How do we quantify this?

Example – Choosing a "Best" Screening Design Based on Multiple Criteria

Design requirements:
- N = 14 runs
- 5 factors
- possible factor levels (-1, +1)
- estimate all 5 main effects (A – E)
- estimate the following interactions: AB, AC, BD, CE
  experts suggest that remaining interactions unlikely

32 possible design points:
$(\pm1, \pm1, \pm1, \pm1, \pm1)$



Criterion to Consider – (1) D-Optimality

Quantifies how well model parameters are estimated for the model

$$Y_i = \beta_0 + \beta_A A + \beta_B B + \beta_C C + \beta_D D + \beta_E E + \beta_{AB} AB + \beta_{AC} AC + \beta_{BD} BD + \beta_{CE} CE + \varepsilon \qquad \varepsilon \overset{i.i.d.}{\sim} N(0,\sigma^2)$$

D-criterion

$$\text{maximize} \quad |\mathbf{M}| = |\mathbf{X}'\mathbf{X}| / N^p \qquad \text{p = \# parameters}$$

design matrix expanded to model form

*inversely proportional to the square of the volume of the confidence region on the regression coefficients

Criterion to Consider – (2) Bias on Model Terms

Assumed model: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}$

$X_1 \in \{A, B, C, D, E, AB, AC, BD, CE\}$

Model to protect against: $\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$

$X_2 \in \{AD, AE, BC, BE, CD, DE\}$

Bias if model incorrect: $E(\hat{\boldsymbol{\beta}}_1) - E(\boldsymbol{\beta}_1) = [\boldsymbol{\beta}_1 + (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2\boldsymbol{\beta}_2] - \boldsymbol{\beta}_1$

$= \mathbf{A}\boldsymbol{\beta}_2$

If these exist, then size unknown

$$\begin{aligned} \mathrm{E}(SS_{bias}) &= \mathrm{E}(\boldsymbol{\beta}_2'\mathbf{A}'\mathbf{A}\boldsymbol{\beta}_2) \\ &= \mathrm{E}(\mathrm{tr}(\mathbf{A}'\mathbf{A}\boldsymbol{\beta}_2\boldsymbol{\beta}_2')) \\ &= \sigma^2_{\beta_2}\,\mathrm{tr}(\mathbf{A}\mathbf{A}'). \end{aligned}$$

Therefore, minimize $\mathrm{tr}(\mathbf{AA}')$

Criterion to Consider – (3) Bias on Error Estimate

For same $X_1 \in \{A, B, C, D, E, AB, AC, BD, CE\}$

$X_2 \in \{AD, AE, BC, BE, CD, DE\}$

Bias on estimate of error,

$$E(\text{MSE}_{\text{user}}) - \sigma^2 = \boldsymbol{\beta}_2'[\mathbf{X}_1\mathbf{A} - \mathbf{X}_2]'[\mathbf{X}_1\mathbf{A} - \mathbf{X}_2]\boldsymbol{\beta}_2 / p_1$$
$$= \boldsymbol{\beta}_2'\mathbf{R}'\mathbf{R}\boldsymbol{\beta}_2 / p_1$$

Therefore, minimize $\text{tr}(R'R)$

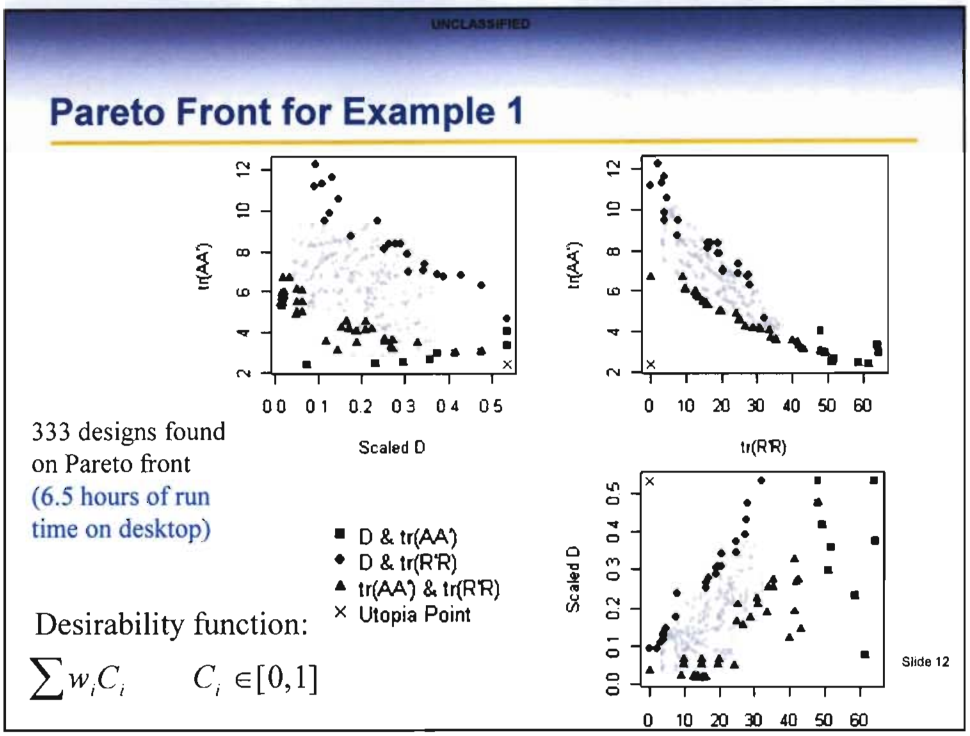
Process for Selecting a Best Design

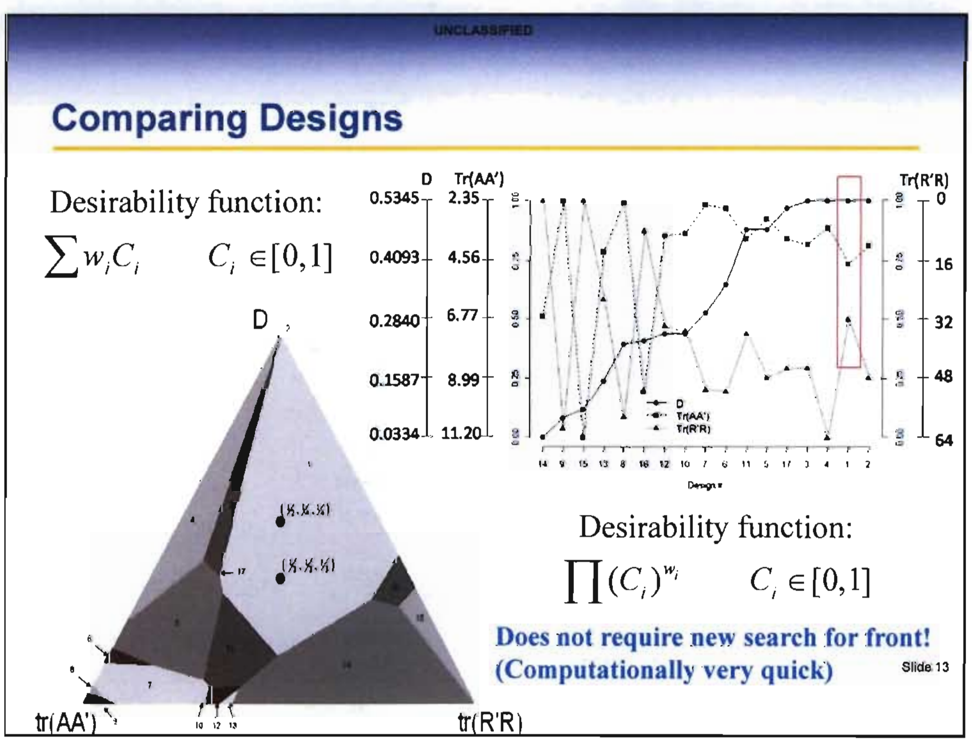
- The process for finding a best design for our specific goals can be summarized by a multi-stage algorithm:

1. Create designs, and measure the criteria for all designs.
2. Construct the Pareto front, which consists of all designs which are not inferior to (*Pareto dominated* by) any other designs [OBJECTIVE]
3. Select a best design from the Pareto front which best suits the needs of the experimenter [SUBJECTIVE].

Pareto Aggregating Point Exchange (PAPE) Algorithm: efficiently creates designs and builds Pareto front

Lu, Anderson-Cook, Robinson (2011 Technometrics)

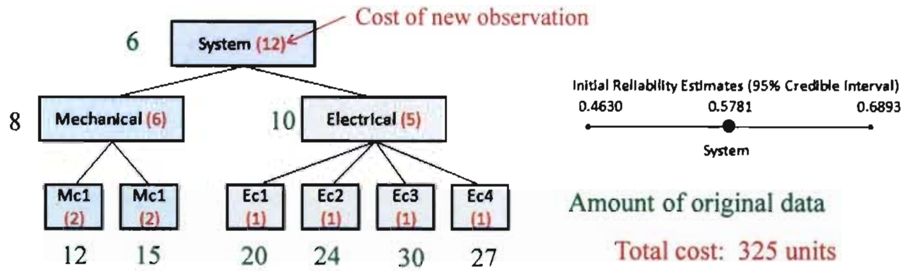

UNCLASSIFIED
Pareto Front for Example 1
tr(AA')
Scaled D
tr(R'R)
333 designs found on Pareto front
(6.5 hours of run time on desktop)
D & tr(AA')
D & tr(R'R)
tr(AA') & tr(R'R)
Utopia Point
Desirability function:
$\sum w_i C_i \qquad C_i \in [0,1]$
Slide 12


UNCLASSIFIED
Comparing Designs
Desirability function:
$\sum w_i C_i \qquad C_i \in [0,1]$
D
tr(AA')
tr(R'R)
Tr(AA')
Tr(R'R)
0.5345
0.4093
0.2840
0.1587
0.0334
2.35
4.56
6.77
8.99
11.20
0
16
32
48
64
Desirability function:
$\prod (C_i)^{w_i} \qquad C_i \in [0,1]$
Does not require new search for front!
(Computationally very quick)
Slide 13

UNCLASSIFIED

System Reliability Case Study: What new data to collect?

- Given the results of an existing reliability analysis based on multiple sources of data, what new data should we collect to **maximally improve** our estimation?

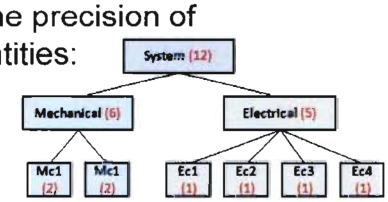


- What new data should we collect?
- What basis should we use for choosing?

UNCLASSIFIED

Goal of New Data Collection for our Example

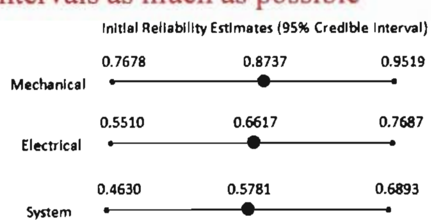
- Engineers would like to improve the precision of estimation for the following 3 quantities:
  - System reliability estimate
  - Mechanical Sub-system
  - Electrical Sub-system



- Focus on the width of the credible interval:

Goal: Reduce the width of each of these 3 intervals as much as possible

Baseline:



Width: 0.1841

Width: 0.2177

Width: 0.2263

Los Alamos NATIONAL LABORATORY EST. 1943

Slide 15

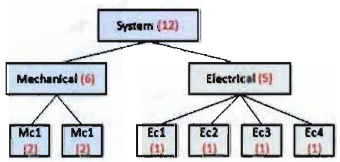
Operated by the Los Alamos National Security, LLC for the DOE/NNSA

NNSA

UNCLASSIFIED

Allocations Possible

25 possible allocations:
- All have same total cost
- Good variety of where data are collected



Results

Mc1	Mc2	Ec1	Ec2	Ec3	Ec4	Mechanical	Electronic	System	Alloc. #
0	0	0	0	0	0	0	0	10	1
0	0	0	0	0	0	20	0	0	2
0	0	0	0	0	0	0	24	0	3
0	0	0	0	0	0	10	0	5	4
0	0	0	0	0	0	0	12	5	5
0	0	0	0	0	0	10	12	0	6
15	15	0	0	0	0	0	0	5	7
0	0	15	15	15	15	0	0	5	8
8	7	0	0	0	0	5	0	5	9
0	0	9	8	6	7	5	0	5	10
0	0	9	8	6	7	0	6	5	11
8	7	0	0	0	0	0	6	5	12
8	7	9	8	6	7	0	0	5	13
16	14	0	0	0	0	10	0	0	14
8	7	9	8	6	7	10	0	0	15
0	0	18	16	12	14	0	12	0	16
8	7	9	8	6	7	0	12	0	17
8	7	0	0	0	0	5	12	0	18
0	0	9	8	6	7	5	12	0	19
0	0	9	8	6	7	10	6	0	20
8	7	0	0	0	0	10	6	0	21
30	30	0	0	0	0	0	0	0	22
0	0	30	30	30	30	0	0	0	23
15	15	15	15	15	15	0	0	0	24
17	13	18	16	12	14	0	0	0	25

UNCLASSIFIED

Methodology to Predict Improvement for Allocations before Data are Collected

- Bayesian Analysis $$f(\theta \mid x, x_{new}) = \frac{p(x_{new} \mid \theta) f(\theta \mid x)}{p(x_{new} \mid x)}$$

	θ	$P(x_2 = 0 \mid \theta, x_1)$	......	$P(x_2 = n_2 \mid \theta, x_1)$
$\tilde{\theta}_1$	$\theta_{(1)}$	$P(x_2 = 0 \mid \theta_{(1)}, x_1)$	......	$P(x_2 = n_2 \mid \theta_{(1)}, x_1)$
	$\theta_{(2)}$	$P(x_2 = 0 \mid \theta_{(2)}, x_1)$	......	$P(x_2 = n_2 \mid \theta_{(2)}, x_1)$
$\tilde{\theta}_2$	$\theta_{(3)}$	$P(x_2 = 0 \mid \theta_{(3)}, x_1)$	......	$P(x_2 = n_2 \mid \theta_{(3)}, x_1)$
	$\theta_{(4)}$	$P(x_2 = 0 \mid \theta_{(4)}, x_1)$	......	$P(x_2 = n_2 \mid \theta_{(4)}, x_1)$
	$\theta_{(5)}$	$P(x_2 = 0 \mid \theta_{(5)}, x_1)$	......	$P(x_2 = n_2$
$\vdots$	$\vdots$	$\vdots$	......	$\vdots$
$\tilde{\theta}_N$	$\theta_{(M-3)}$	$P(x_2 = 0 \mid \theta_{(M-3)}, x_1)$	......	$P(x_2 = n_2 \mid \theta_{(M-3)}, x_1)$
	$\theta_{(M-2)}$	$P(x_2 = 0 \mid \theta_{(M-2)}, x_1)$	......	$P(x_2 = n_2 \mid \theta_{(M-2)}, x_1)$
	$\theta_{(M-1)}$	$P(x_2 = 0 \mid \theta_{(M-1)}, x_1)$	......	$P(x_2 = n_2 \mid \theta_{(M-1)}, x_1)$
	$\theta_{(M)}$	$P(x_2 = 0 \mid \theta_{(M)}, x_1)$	......	$P(x_2 = n_2 \mid \theta_{(M)}, x_1)$

- Approximate $p(\theta \mid x_1)$ by treating θ as discrete and "making a histogram"
- For each (discrete) value of θ, approximate $p(x_2 \mid \theta)$ by averaging over the rows that correspond to that value of θ
- Approximate $p(x_2 \mid x_1)$ by averaging over all rows

Chapman, Morris, Anderson-Cook, 2011

UNCLASSIFIED

Initial:
0.1841 0.2177 0.2263

Results of Analysis for 25 Allocations

Mc1	Mc2	Ec1	Ec2	Ec3	Ec4	Mechanical	Electronic	System	Alloc. #	Mechanical	Electrical	System
0	0	0	0	0	0	0	0	10	1	0.182896	0.208562	0.212798
0	0	0	0	0	0	20	0	0	2	0.155939	0.217593	0.216135
0	0	0	0	0	0	0	24	0	3	0.186184	0.188425	0.205978
0	0	0	0	0	0	10	0	5	4	0.167084	0.212652	0.213953
0	0	0	0	0	0	0	12	5	5	0.183102	0.196996	0.208728
0	0	0	0	0	0	10	12	0	6	0.169521	0.201595	0.208982
15	15	0	0	0	0	0	0	5	7	0.160029	0.212583	0.211654
0	0	15	15	15	15	0	0	5	8	0.184335	0.193014	0.20547
8	7	0	0	0	0	5	0	5	9	0.163004	0.212495	0.212505
0	0	9	8	6	7	5	0	5	10	0.176321	0.202387	0.209188
0	0	9	8	6	7	0	6	5	11	0.184579	0.196233	0.207769
8	7	0	0	0	0	0	6	5	12	0.171135	0.204599	0.20955
8	7	9	8	6	7	0	0	5	13	0.172838	0.202965	0.208023
16	14	0	0	0	0	10	0	0	14	0.150264	0.217789	0.214795
8	7	9	8	6	7	10	0	0	15	0.158434	0.206712	0.20864
0	0	18	16	12	14	0	12	0	16	0.186341	0.185205	0.202329
8	7	9	8	6	7	0	12	0	17	0.17274	0.192642	0.203601
8	7	0	0	0	0	5	12	0	18	0.166135	0.201575	0.207661
0	0	9	8	6	7	5	12	0	19	0.178806	0.19274	0.205274
0	0	9	8	6	7	10	6	0	20	0.169961	0.199462	0.207632
8	7	0	0	0	0	10	6	0	21	0.159256	0.209364	0.211409
30	30	0	0	0	0	0	0	0	22	0.143072	0.217661	0.213209
0	0	30	30	30	30	0	0	0	23	0.186722	0.18047	0.19775
15	15	15	15	15	15	0	0	0	24	0.163415	0.197407	0.203219
17	13	18	16	12	14	0	0	0	25	0.161928	0.196507	0.20274

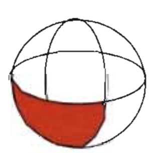
Slide 18

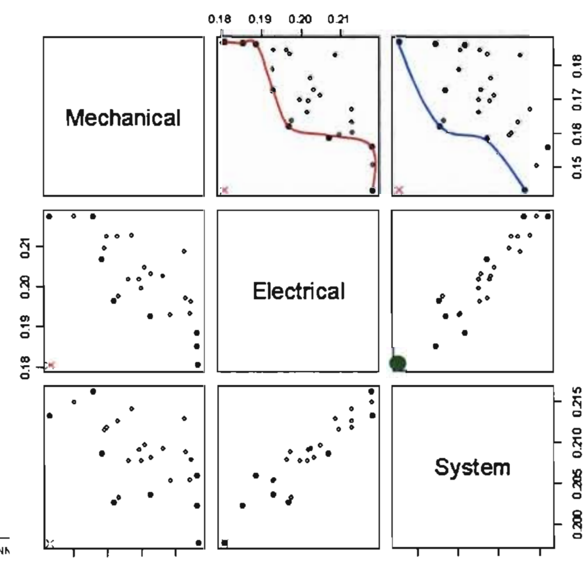
UNCLASSIFIED

Pareto Front for 25 allocations

Pareto Front (9 pts) – consider further

Not on front (16 pts) – discard

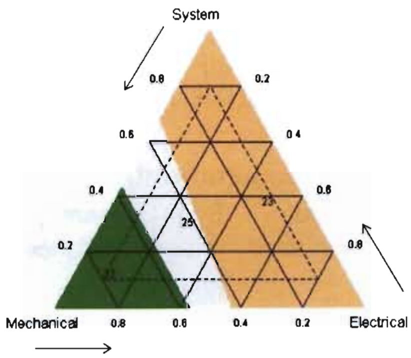


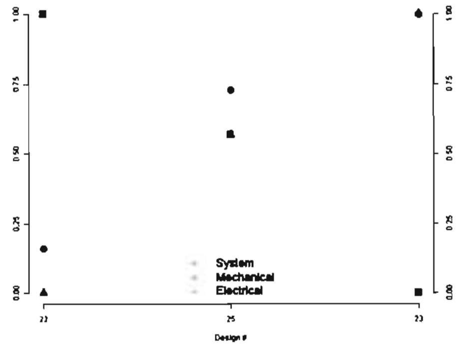


Los Alamos
NATIONAL LABORATORY
EST 1943
Operated by the Los Alamos National Security, LLC for the DOE/NN

Finding Best Allocation for All Weightings of Criteria (Additive)

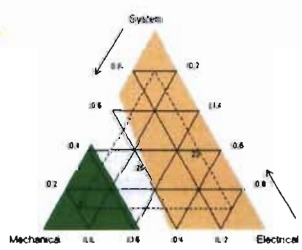
$$Desirability = w_S CW_S + w_M CW_M + w_E CW_E$$

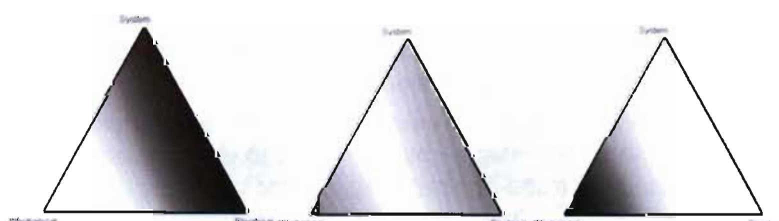




Mc1	Mc2	Ec1	Ec2	Ec3	Ec4	Mechanical	Electronic	System	Alloc. #
30	30	0	0	0	0	0	0	0	22
0	0	30	30	30	30	0	0	0	23
15	15	15	15	15	15	0	0	0	24
17	19	18	16	12	14	0	0	0	25

Synthesized Efficiency to Evaluate the Chosen Allocation





O100% O75% ●50% ●25% ●0%

UNCLASSIFIED

Other Applications of Pareto Front Approach for Design (in various publications)

- Example 3 (Screening Experiment):
  - D-optimality [maximize |X'X|]
  - Good estimation of pure error [maximize df_{PE}]
  - Good estimation of lack of fit [maximize tr(R'R) /(m-p)]
- Example 4 (Robust Parameter Design Experiment):
  - Good estimation of terms affecting the mean [max D_s-mean]
  - Good estimation of terms affecting the variance [max D_s-variance]
  - Size of experiment [min N]
- Example 5 (Split Plot Design):
  - Good estimation of terms when WP to SP variance ratio is unknown [max D(1), max D(10)]
  - Size of experiment [min N]
  - Number of Whole Plots [min #WP]

Slide 22

UNCLASSIFIED

Conclusions

- Looking at multiple characteristics can lead to better choices of which design to run (do well for several priorities – not just one!)
- Different designs have different advantages and risks – select criteria to consider which best capture the important considerations for your experiment. It is now possible to focus on what is most important to the experimenter – and do well on those objectives.
- The Pareto front approach can divide possible designs into (1) those consider further and (2) those to eliminate, because they are dominated by other better choices. This objective step selects which designs are sensible to consider.
- Once the Pareto front has been selected, there are multiple ways of selecting the final design – but the key is to examine and understand the trade-offs between the choices. This subjective phase allows experimenter needs to be emphasized

Los Alamos NATIONAL LABORATORY EST. 1943

UNCLASSIFIED

Operated by the Los Alamos National Security, LLC for the DOE/NNSA

NNSA

UNCLASSIFIED

References

- Anderson-Cook, C.M., Chapman, J., Lu, L. (2011) "Selecting a Best Allocation of New Data for Improving Estimation Precision of System and Sub-System Reliability using Pareto Fronts: A Demonstration" LANL Technical Report, LAUR 11-05479
- Chapman, J., Morris, M., Anderson-Cook, C.M. (2011) "A Computationally Efficient Strategy for Evaluating the Estimation Improvement for Candidates in a Resource Allocation Study" LANL Technical Report LA-UR 10-04501
- Lu L., Anderson-Cook, C.M. (2011) "Rethinking the Optimal Response Surface Design for a First-Order Model with Two-Factor Interactions, when Protecting against Curvature" **Quality Engineering** (to appear)
- Lu, L., Anderson-Cook, C.M., Robinson, T.J. (2011) "Optimization of Designed Experiments Based on Multiple Criteria Utilizing a Pareto Frontier" **Technometrics** 53 353-365.
- Lu, L., Anderson-Cook, C.M., Robinson, T.J. (2011) "A Case Study to Demonstrate Pareto Frontiers for Selecting a Best Response Surface Design While Simultaneously Optimizing Multiple Criteria" Applied Stochastic Models in Business and Industry (to appear)
- Lu, L., Anderson-Cook, C.M., Robinson, T.J. (2011) "Using a Pareto Frontier to Select an Optimal Split-Plot Design for a Mixture-Process Experiment" LANL Technical Report, LAUR 11-

Slide 24