PRACTICAL NECESSITY AS THE MOTHER OF STATISTICAL INNOVATION

CHRISTINE M. ANDERSON-COOK LOS ALAMOS, NM candcook@gmail.com



2021 Fall Technical Conference Gerald Hahn Talk

GERALD J. HAHN

- 46 years at GE Global Research Center
- 26 years as Manager of the Applied Statistics Program
- First Chair of ASA Quality and Productivity Section

Role model for "leadership in developing, promoting, and successfully improving the quality and productivity of products and organizational performance using statistical concepts and methods"





Some of Gerry's hallmarks: Seeing what is important, articulating it clearly and leading to way to the right solution

PREFLECTIONS ON MY CAREER

1994 – Ph.D. University of Waterloo (Advisor: C.F.J. Wu)
 1994-1996 – Assistant Professor, Department of Statistics and Actuarial Science, Western University
 1996-2004 – Assistant/Associate Professor, Department of Statistics, Virginia Tech
 2002 – Visiting Professor at Arizona State University
 2004-2021 – Research Scientist, Statistical Sciences Group, Los Alamos National Laboratory

Privileged to have my career span some exciting times in statistics:

Innovation in graphical capabilities

Renaissance of design of experiments

Dramatic increase in computational power for algorithmic searches

Dawn of "data science" and "big data"

Statistical Engineering (Hoerl and Snee, 2010) \rightarrow ISEA (isea-change.org)

David Bellhouse (Western Ontario): "If you are lucky, you might have a handful of ideas in your career that really have a broader impact. The profession needs both explorers and settlers."

[©] MY ROLE AS STATISTICIAN IN COLLABORATIONS

How my role has evolved throughout my career:

Early on: Sales clerk approach

Customer comes in with a problem (want to collect some data)

Take their description and find the right product to solve the problem (create/find the right design to match needs)

Send them on their way, until they have another problem (often, help with the analysis)



In this talk, illustrations will focus on design of Experiments

MY ROLE AS STATISTICIAN IN COLLABORATIONS

Later:

Interior designer approach

Meet customer in their space (build context for data collection)

After understanding what is important to them, develop several possibilities (identify several options to compare)

Discuss what they like/dislike about options (deeper understanding of their priorities and context)

Iterate until they are satisfied with the final choice (build a deeper level of comfort for customer with design)



Key differences:

- Collaborative as a team member
- Deeper understanding of what is the right solution
- Multiple options discussed and compared
- Iterative creating new designs is easy compared to executing the experiment
- Experimenter owns the final decision

SALES CLERK VS INTERIOR DESIGNER

Advantages of Designer Approach

For organization:

- Class of problems that can be solved is much bigger
- Better definition of problem to be solved
- Better solutions
- Collaborators feel more ownership of statistical part of solution
- Collaborators gain insights into statistical expertise that they can use for other projects

For statistician:

- Contributions have potential for much larger impact
- Creates pull for involvement in subsequent projects
- Full team member status (not technical support)
- Greater professional fulfillment



• HOW DID WE GET FROM THERE TO HERE?

- Need #1: Richer set of tools to compare designs Innovation: Fraction of Design Space (FDS) Plot
- Need #2: Ability to handle multi-faceted goals and protection from the unexpected

Pareto Fronts with flexible criteria

 Need #3: Ability to hedge risk from lack of knowledge with a "learn and leverage" approach



Sequential Design of Experiments (SDoE)

• Need #4: Improved structure for team discussions and decision-

making DMRCS, Collaboration & Graphical Tools



NEED #1 – TOOLS FOR COMPARING DESIGNS

• Need #1: Richer set of tools to compare designs

We make better decisions if we have a variety of choices:



We are much more engaged and discriminating about the fit of the solution to our problem if we have choices and can evaluate/compare their differences

Goal: Have a graphical and/or numerical summary that matches the priority of the study to compare the designs



- Time travel back to 2000
- <u>Goal</u>: compare how well different designs predict throughout input space

De	sign Diag	nostics	
		Efficiency of Central Composite Design Relative to Box-Behnken	
	D-efficiency	1.220	
	G-efficiency	1.751	Worst case prediction variant
	A-efficiency	1.066	
	I-efficiency	1.047	Average prediction variance
Additional Run Size		0	

Problems:

- Too simplistic a summary for performance across a higher dimensional space
- We are going to use our model to predict at multiple input locations not "average" or "worst"



NEED #1 – COMPARING DESIGNS (CONTINUED)

- Single line for each design
- Same interpretation regardless of dimension, shape of region
- Comparison points easier to identify (best, worst, median)
- (Easy to generate)

Versatile:

- Any shaped input region
- Compare different sizes
- Robustness to model choice
- Generalized linear model
 parameter choices



NEED #1 – COMPARING DESIGNS (CONTINUED)



$^{\circ}$ NEED #1 – COMPARING DESIGNS (CONTINUED)

Process for data collection planning with collaborators:

- Begin with discussion about priorities what do success and failure look like?
- Creating designs is an iterative process be responsive to new information, emphasize that creating a new design is easy
- Present multiple alternatives with their pros and cons
- Use graphical methods (like FDS plot, if appropriate) to demonstrate differences between choices
- Let experimenters make final choice of design between sensible alternatives

Greater comfort for experimenter about why this is the right design Better chance for following implementation



- Need #1: Richer set of tools to compare designs Fraction of Design Space (FDS) Plot
- Need #2: Ability to handle multi-faceted goals and protection from the unexpected

Pareto Fronts with flexible criteria

 Need #3: Ability to hedge risk from lack of knowledge with a "learn and leverage" approach

Sequential Design of Experiments (SDoE)

• Need #4: Improved structure for team discussions and decision-

making DMRCS, Collaboration & Graphical Tools

NEED 2: ABILITY TO HANDLE MULTI-FACETED GOALS AND PROTECTION FROM THE UNEXPECTED

There are many potential priorities to consider when designing an experiment



NEED 2: ABILITY TO HANDLE MULTI-FACETED GOALS AND PROTECTION FROM THE UNEXPECTED

There are many potential priorities to consider when designing an experiment





NEED 2 (CONTINUED)

Some Examples in the Literature

- Screening Experiment: 1.
 - D-optimality [maximize |X'X|] Good estimation of pure error
 - [maximize df_{PF}]
 - Good estimation of lack of fit [maximize tr(R'R) /(m-p)]

Robust Parameter Design Experiment 2.

- 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]

3. Split Plot Design

- Good estimation of terms when WP to SP variance ratio is unknown
- Size of experiment
- Number of Whole Plots

4. Sequential Experimentation for Estimating System Reliability

- Good prediction of
 - system reliability [min uncertainty interval]
 - mechanical sub-system reliability

[min uncertainty interval] • electrical sub-system reliability [min uncertainty interval]

[maximin in X]

Space-Filling Design 5.

- Good spacing in input space
- Good spacing in predicted response range. [maximin in Y]

[max D(0.1), max D(10)] [min N] [min #WP]

NEED 2: ABILITY TO HANDLE MULTI-FACETED GOALS AND PROTECTION FROM THE UNEXPECTED SUMMARY

Experience has taught me that:

1. Getting the model right before you have collected the data is not common or easy.

2. Different stakeholders have different priorities for the experiment. Don't require everyone to come to consensus on one goal.

3. A close-to-optimal design that can survive some surprises is better than a non-robust optimal design.

The Pareto front approach allows for:

- Multiple user-selected priorities to be considered
- Multiple competitive designs to be created and compared
- Graphical tools to facilitate discussion about alternatives
- Understanding of how severe the trade-offs are for different priorities
- Intentional protection from things not going as planned (imperfect knowledge and/or problems running experiment)





NEED 3: HEDGE RISK WITH A "LEARN AND LEVERAGE" APPROACH

Some of my favorite solutions for incorporating knowledge into experiments:

1. Sequential Design of Experiments

If you are going on a long trip, would you rather:

- 1. Commit to exactly the route, street-by-street at the start of your journey
- 2. Plan the first part of your route, and then each evening adapt the next step of the journey using updated information

Same logic applies to collecting data – why commit the entire experimental

budget at the beginning? Why not learn as we go?

2. Non-Uniform Space-Filling Designs

Space-filling designs that provide the experimenter with control to change the density of points to match their level of interest in a region



SEQUENTIAL DESIGN OF EXPERIMENTS

- 1. If it is possible to collect, measure and analyze the data in a **timely** fashion.
 - Makes implementation possible.
- 2. If the process is **stable and consistent** over time.
 - Makes combining data from various stages straightforward
- 3. If there is **uncertainty about the input region** of interest.
 - Avoid wasting resources in undesirable regions
- 4. If the experiment has several objectives.
 - Easier to tackle objectives individually and leverage what is already known.
- 5. If **little is known** about the process.
 - Learn as you go.
- 6. If logistics impose **constraints on the size of the experiments** that could be run at any time window.
 - -Natural if blocking inherently imposed



SEQUENTIAL DESIGN OF EXPERIMENTS



Can we get quality data and measure what we need?

Understand basic relationship between inputs and response

Verify model adequately characterizes relationship with acceptable level of uncertainty

Focus on portion of region with best values of the response

Verify results for production or operational use



Conceptual Space-Filling Example – 2 factors 18 run budget





[©] NON-UNIFORM SPACE-FILLING (NUSF) DESIGNS

Key to NUSF: define weights for each location in the input space that reflect experimenter interest

Assign higher weights to regions where:

- Response values are better (optimize)
- More interesting features (exploration)
- Response changing more quickly (exploration)
- Prediction has larger uncertainty (model refinement)
- Discrepancy is larger between computer model and observed data (calibration)



NON-UNIFORM SPACE-FILLING (NUSF) DESIGNS

Idea:

- User specifies a weight function for all points in their candidate set (larger weights for regions with higher density of points)
- Distance function for maximin criterion is adjusted from

$$d(\mathbf{x}_1, \mathbf{x}_2) = \sqrt{\sum_{j=1}^{p} (\mathbf{x}_{1j} - \mathbf{x}_{2j})^2}$$
 to

$$d^{w}(\boldsymbol{x}_{1}, \boldsymbol{x}_{2}) = \sqrt{w(\boldsymbol{x}_{1})w(\boldsymbol{x}_{2})d(\boldsymbol{x}_{1}, \boldsymbol{x}_{2})^{2}}$$

Weights associated with each candidate point.

To tune the degree of non-uniformity: scale raw weights from [Min, Max] \rightarrow [1, MWR]

If MWR = 1 $\leftarrow \rightarrow$ uniform

as MWR increases, degree of non-uniformity increases

Original Runs Augmented Runs Pairwise Distance Required to Evalu





Q

NON-UNIFORM SPACE-FILLING (NUSF) DESIGNS

Idea:

- User specifies a weight function for all points in their candidate set (larger weights for regions with higher density of points)
- Distance function for maximin criterion is adjusted from

$$d(\mathbf{x}_1, \mathbf{x}_2) = \sqrt{\sum_{j=1}^{p} (\mathbf{x}_{1j} - \mathbf{x}_{2j})^2}$$
 to

$$d^{w}(\boldsymbol{x}_{1}, \boldsymbol{x}_{2}) = \sqrt{w(\boldsymbol{x}_{1})w(\boldsymbol{x}_{2})d(\boldsymbol{x}_{1}, \boldsymbol{x}_{2})^{2}}$$

Weights associated with each candidate point.

To tune the degree of non-uniformity: scale raw weights from [Min, Max] \rightarrow [1, MWR]

If MWR = 1 $\leftarrow \rightarrow$ uniform

as MWR increases, degree of non-uniformity increases











- Need #1: Richer set of tools to compare designs Fraction of Design Space (FDS) Plot
- Need #2: Ability to handle multi-faceted goals and protection from the unexpected

Pareto Fronts with flexible criteria

 Need #3: Ability to hedge risk from lack of knowledge with a "learn and leverage" approach

Sequential DoE and NUSF

• Need #4: Improved structure for team discussions and decision-

making DMRCS, Collaboration & Graphical Tools



We are only effective as statisticians when our voices and message are heard.

Some of ideas:

- DMAIC: Define-Measure-Analyze-Improve-Control for Problem-Solving
- DMRCS: Define-Measure-Reduce-Combine-Select for Decision-Making
- Facilitating Discussion and Decision-Making
- Collaboration Strategies
- STATISTICAL ENGINEERING

COMMON PROBLEMS WITH DECISION-MAKING

- 1. Narrow framing: limiting choices and options
- 2. Confirmation bias: seeing and evaluating data based on current leading choice
- 3. Thinking too short-term: not including longer term consequences into decision
- 4. Overconfidence: not building in sufficient uncertainty about future
- 5. Team-dynamics: personalities, process, competing priorities





STEVEN

OHNS

NEED 4: IMPROVED STRUCTURE FOR TEAM DISCUSSIONS AND
DECISION-MAKINGExample: Choosing a Design

31

Define	What is important for this experiment?What designs are available to choose between?
Measure	 What is the right criterion to measure the important attribute? High quality data to assess designs for identified criteria? Gather relevant information
Reduce	 Are some aspects secondary in importance? Remove or defer? Can some designs be eliminated as implausible or non-contenders?
Combine	 How can I evaluate trade-offs between criteria (that are probably on different scales)? How much do I value the different criteria?
Select	 Which solution is best given my priorities? How does this solution compare to other options? Can I defend my choice?
11	

Anderson-Cook, C.M., Lu, L. (2018) "Graphics to Facilitate Informative Discussion and Team Decision-Making" Applied Stochastic Models in Business and Industry (with discussion and rejoinder)

General

- G1 Use a process to find common ground
- G2 Clarify the difference between "right and wrong" and "choosing differently based on priorities" **Define**
- **D1** Define and use an appropriate summary that directly connects to the decision
- **D2** Build the decision space to include diverse alternatives
- D3 Push the boundaries on assumptions
- **D4** Incorporate cost into the comparisons of alternatives

Measure

- **M1** For realistic decision making, it is important to provide uncertainty quantification to inform the decision makers about the uncertainty and potential risks associated with a specific decision
- M2 Devise analysis summaries that reduce the uncertainty can improve the decision-making process

Reduce

R1 It is helpful to take strategic steps to reduce the number of choices on which to do a detailed comparison to a manageable number

Combine

- C1 Allow common visualization and discussion about results, instead of keeping subjective elements of the decision unshared
- **C2** When comparing the smaller set of contenders, choose plots that highlight the impact of subjective choices to facilitate discussion
- C3 Think globally and locally
- C4 Consider dynamic graphics when dimensionality of problem suggests it

Select

- **S1** Create or use a graphical summary that appropriately captures the needed level of detail
- S2 Include graphics to formalize conclusions

Collaboration Strategies

- Anderson-Cook, C.M., Lu, L., Parker, P.A. (2019) "Effective Interdisciplinary Collaboration Between Statisticians and Other Subject Matter Experts" Quality Engineering (with discussion and rejoinder)
- 2020 FTC talk: https://www.youtube.com/watch?v=9PUqc2puBag

STATISTICAL ENGINEERING

- International Statistical Engineering Association (free membership isea-change.org)
- ISEA Workshop: Nov 18-19
- 2012 Quality Engineering Special Issue

(https://www.tandfonline.com/toc/lqen20/24/2?nav=tocList)



CONCLUSIONS



We needed several key elements:

- Ability to generate multiple designs that allow for comparisons
- Ability to have the right tools/graphics to make specific and meaningful comparisons
- Ability to collect data, learn from it, pivot using what we learn, collect the next set of data
- Develop "human skills", confidence and strategies to lead and/or be full team members on collaborative teams

We are in exciting times for statistics – many new areas emerging, where we can make important contributions.

Consider new ideas to help improve problem-solving and decision-making. Let the needs of collaborators, practitioners guide our innovation.

*Simon Sinek

© REFERENCES FOR FRACTION OF DESIGN SPACE PLOTS

- 1. Zahran, A., Anderson-Cook, C.M. and Myers, R.H. (2003) "Fraction of Design Space to Assess the Prediction Capability of Response Surface Designs" Journal of Quality Technology 35 377-386.
- Goldfarb, H.B., Anderson-Cook, C.M., Borror, C.M., Montgomery, D.C. (2004) "Fraction of Design Space to Assess the Prediction Capability of Mixture and Mixture-Process Designs" Journal of Quality Technology 36 169-179.
- **3.** Ozol-Godfrey, A., Anderson-Cook, C.M., Montgomery, D.C. (2005) "Fraction of Design Space Plots for Examining Model Robustness" Journal of Quality Technology 37 223-235.
- 4. Liang, L., Anderson-Cook, C.M., Robinson, T.J. (2006) "Fraction of Design Space Plots for Split-Plot Designs" Quality and Reliability Engineering International 22 275-289.
- 5. Ozol-Godfrey, A., Anderson-Cook C. M. (2007) "Fraction of Design Space Plots for Examining Mixture Design Robustness to Measurement Errors" Journal of Statistics and Applications 1 171-183.
- 6. Ozol-Godfrey, Anderson-Cook, C.M., Robinson, T.J. (2008) "Fraction of Design Space Plots for Generalized Linear Models" Journal of Statistical Planning and Inference 138 203-219.
- 7. Jang, D.-H., Anderson-Cook, C.M. (2011) "Fraction of Design Space Plots for Evaluating Ridge Estimators in Mixture Experiments" Quality and Reliability Engineering International 27 27-34.

REFERENCES FOR PARETO FRONT DESIGNS

- 1. 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. (2012) "A Case Study to Demonstrate Pareto Frontiers for Selecting a Best Response Surface Design with Simultaneously Optimizing Multiple Criteria" Applied Stochastic Models in Business and Industry 28 206-221.
- **3.** Lu L., Anderson-Cook, C.M. (2012) "Rethinking the Optimal Response Surface Design for a First-Order Model with Two-Factor Interactions, when Protecting against Curvature" **Quality Engineering** 24 404-422.
- Lu, L., Chapman, J., Anderson-Cook, C.M. (2013) "A Case Study on Selecting A Best Allocation of New Data for Improving the Estimation Precision of System and Sub-System Reliability using Pareto Fronts" Technometrics 55 473-487.
- Lu, L., Anderson-Cook, C.M. (2013) "Adapting the Hypervolume Quality Indicator to Quantify Trade-Offs and Search Efficiency for Multiple Criteria Decision-Making Using Pareto Fronts" Quality and Reliability Engineering International 29 1117-1133
- 6. Lu, L., Anderson-Cook, C.M. (2014) "Balancing Multiple Criteria Incorporating Cost using Pareto Front Optimization for Split-Plot Designed Experiments" Quality and Reliability Engineering International 30 37-55
- 7. Lu, L., Anderson-Cook, C.M., Lin, D. (2014) "Optimal Designed Experiments Using a Pareto Front Search for Focused Preference of Multiple Objectives" Computational Statistics and Data Analysis 71 1178-1192
- 8. Chapman, J., Lu, L., Anderson-Cook, C.M. (2014) "Process Optimization for Multiple Responses Utilizing the Pareto Front Approach" Quality Engineering 26 253–268.
- 9. Chapman, J.L., Lu, L., Anderson-Cook, C.M. (2014) "Incorporating Response Variability and Estimation Uncertainty into Pareto Front Optimization" Computers & Industrial Engineering 76 253-267

© REFERENCES FOR SEQUENTIAL DESIGN OF EXPERIMENTS

- Lu, L., Anderson-Cook, C.M. (2021) "Strategies for Sequential Design of Experiments and Augmentation" Quality and Reliability Engineering International 37(5) 1740-1757.
- 2. Lu, L., Anderson-Cook, C.M., Ahmed, T. (2020) "Non-Uniform Space-Filling Designs" Journal of Quality Technology 53(3) 309-330.
- 3. Lu, L., Anderson-Cook, C.M. (2021) "Input-Response Space-Filling Designs" Quality and Reliability Engineering International http://doi.org/10.1002/qre.2931
- 4. Anderson-Cook, C.M. Lu, L. (2020) "Why Not Sequential DoEs?" Quality Progress 52(11) 44-47.
- Anderson-Cook, C.M., Graves, T.G., Hamada, M. (2008) "Resource Allocation: Sequential Design for Analyses Involving Several Types of Data" IEEE Xplore Refereed Proceedings of IEEM Singapore

REFERENCES FOR STATISTICAL ENGINEERING & COLLABORATION

- 1. Anderson-Cook, C.M., Lu, L. (2015) "Much-Needed Structure: A New 5-Step Decision-Making Process Helps You Evaluate, Balance Competing Objectives" **Quality Progress** 48(10) 42-50.
- Anderson-Cook, C.M., Lu, L., Parker, P.A. (2019) "Effective Interdisciplinary Collaboration Between Statisticians and Other Subject Matter Experts" Quality Engineering (with discussion and rejoinder) 31(1) 164-204.
- Anderson-Cook, C.M., Lu, L. (2018) "Graphics to Facilitate Informative Discussion and Team Decision-Making" Applied Stochastic Models in Business and Industry (with discussion and rejoinder) 34 (6) 963-991.
- 4. Anderson-Cook, C.M., Lu, L., Clark, G., DeHart, S.P., Hoerl, R., Jones, B., MacKay, R.J., Montgomery, D.C., Parker, P.A., Simpson, J., Snee, R., Steiner, S., Van Mullekom, J., Vining, G.G., Wilson, A.G. (2012) "Statistical Engineering – Forming the Foundations" Quality Engineering 24 110-132.
- Anderson-Cook, C.M., Lu, L., Clark, G., DeHart, S.P., Hoerl, R., Jones, B., MacKay, R.J., Montgomery, D.C., Parker, P.A., Simpson, J., Snee, R., Steiner, S., Van Mullekom, J., Vining, G.G., Wilson, A.G. (2012) "Statistical Engineering – Roles for Statisticians and the Path Forward" Quality Engineering 24 133-152.