Response Surface Models: To Reduce or Not to Reduce?

Maria L. Weese

MIAMI UNIVERSITY

Joint Work With…

Byran Smucker
Miami University

David Edwards
VCU

Background and Setting

High-level overview of RSM

- **Screening**
  - If necessary reduce the number of factors from many to just a few.
- **Initial Improvement**
  - If far from a process optimum, rapid improvement is likely possible using a simple design and first-order model
- **Optimization**
  - Once the experimenter nears a local optimum, an approximate optimum can be ascertained using a second-order model.

Motivation

What if the first two RSM steps are ignored?

- Many of the RSM experiments we examined from the literature (83 out of 129) failed to mention a prior screening experiment (Ockuly et al., 2017)
- Reasonable that not all of the factors are important
- Even more likely that certain interactions/quadratic terms will be unimportant
- Could overfitting reduce the quality of prediction and/or optimization?

Chemometrics and Intelligent Laboratory Systems 164 (2017) 64–75

Contents lists available at ScienceDirect

ELSEVIER

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemometrics

CHEMOMETRICS AND INTELLIGENT LABORATORY SYSTEMS

Response surface experiments: A meta-analysis

Rebecca A. Ockuly[a], Maria L. Weese[a,*], Byran J. Smucker[a], David J. Edwards[b], Le Chang[a]

[a] Miami University, Oxford, OH, USA
[b] Virginia Commonwealth University, Richmond, VA, USA

ARTICLE INFO

Keywords:
Box-Behnken designs
Central composite designs
Effect heredity
Effect hierarchy
Effect sparsity

ABSTRACT

Response Surface Methodology is a set of experimental design techniques for system and process optimization that is commonly employed as a tool in chemometrics. In the last twenty years, thousands of studies involving response surface experiments have been published. The goal of the present work is to study regularities observed among factor effects in these experiments. Using the Web of Science Application Program Interface, we searched for journal articles associated with response surface studies and extracted over 20,000 records from all Science Citation Index and Social Science Citation Index disciplines between 1990 and the end of 2014. We took a random sample of these papers, stratified by the number of factors, and ended up with a total of 129 experiments and 183 response variables. Extracting the data from each publication, we reanalyzed the experiments and combined the results together in a meta-analysis to reveal information about effect sparsity, heredity, and hierarchy. We empirically quantify these principles to provide a better understanding of response surface experiments, to calibrate experimenter expectations, and to guide researchers toward more realistic simulation scenarios and improved design construction.

To Reduce or Not to Reduce?

If your goal is prediction and/or optimization via a second-order model:

1. Should you use the full second-order model for response surface analysis or should you reduce it?

2. If you should reduce, what method should you use?

Analysis Methods to Compare

For prediction and optimization

1. The full second-order model
2. A reduced second-order model, based upon p-values and $\alpha = 0.05$
3. A reduced second-order model using p-values adjusted using the False Discovery Rate (FDR) (Benjamini and Hochberg 1995), with $\alpha = 0.05$
4. Forward selection using AICc as the selection criterion
5. LASSO-regularized regression (Tibshirani 1996)
6. The Gauss-LASSO (Rigollet and Tsybakov 2011)
7. Bayesian optimization using the posterior predictive distribution (Peterson 2004)

Response Surface Optimization

Formally Defined

$$[X_1^*, \ldots, X_m^*] = argmax_{[X_1,\ldots,X_m] \in R} f(X_1, \ldots, X_m)\hat{\beta}$$

For a given fitted model, numerical optimization methods can be used to find $[X_1^*, \ldots, X_m^*]$.

Bayesian Optimization

(Peterson 2004)

» Noninformative prior

» Under this prior distribution, it is shown that the posterior predictive density follows a non-central t-distribution with n − p degrees of freedom

» Straightforward to compute the probability that a future response, at a new treatment combination, will conform to some desirable quality level:
$P\left(\tilde{Y} \in R \middle| Y, \tilde{X}_1, \dots, \tilde{X}_m\right) \rightarrow$ Bayesian Reliability

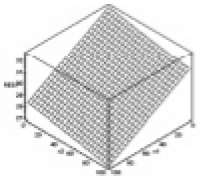
» Response surface optimization is conducted by maximizing Bayesian reliabilities over the experimental region

MIAMI UNIVERSITY

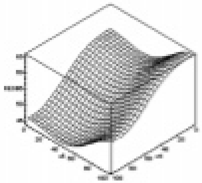
Simulation Study

RSM simulation testbed from McDaniel and Ankenman (2000)

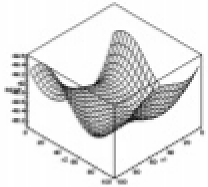
- Provides control over effect heredity, effect sparsity, and the “bumpiness” of the response surface.



S=0, T=0



S=4, T=0



S=6, T=1

QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL
Qual. Reliab. Engng. Int. 2000; **16**: 363–372

A RESPONSE SURFACE TEST BED

WILLIAM R. MCDANIEL AND BRUCE E. ANKENMAN*
Department of Industrial Engineering and Management Sciences, Northwestern University, 2145 Sheridan Road, Tech C210, Evanston, IL 60208-3119, USA

SUMMARY

A method is presented for creating randomly generated polynomial functions to be used as a test bed of simulated response surfaces. The need for the test bed to perform empirical comparisons of experimental design strategies is discussed and the methods used to create the surfaces are explained. An important feature of the test bed is that the user can control some of the characteristics of the surfaces without directly controlling the surface functions. This allows the user to choose the types of surfaces on which a simulation study is run while preserving the random nature of the surfaces needed for a valid simulation study.

KEY WORDS: experimental design; polynomial models; effect heredity; hierarchical model; random function; simulation

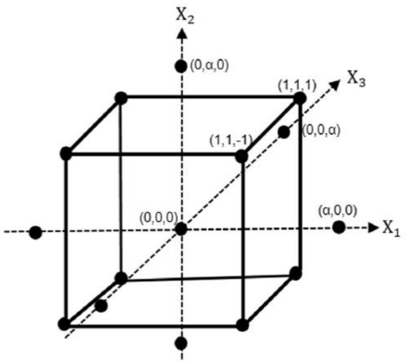
1. INTRODUCTION

The experimental study of a response surface for finding optimal or at least desirable settings for the factors is known as Response Surface Methodology (RSM) (see Myers and Montgomery [1]). Many classes of experimental designs have been developed for RSM, such as factorials, fractional factorials, Box–Behnken designs, and central composite designs. A natural question is *How well do these designs work?* That is, how well is the true response surface function modeled by the results of an experimental design? While this may seem like a simple question, in practice it is very difficult to answer because the true response composed of several stages of experimentation is used, as in traditional RSM (see Box and Wilson [3]). We are interested in using a simulation to determine how an experimental design or a design strategy performs when presented with different types of response surface functions. To this end, we have developed a test bed that will randomly generate polynomial functions to represent response surfaces in a simulation study. The characteristics of the surface functions created by the test bed are controllable, allowing researchers to specify the types of surface functions on which design strategies will be tested. When an experimental design is applied to a response

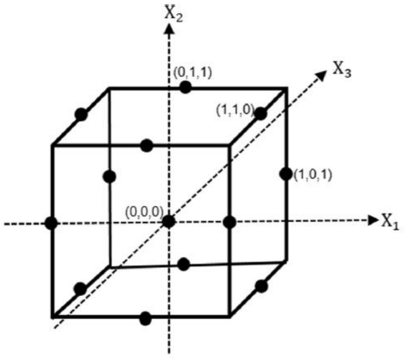
Simulation Study

RSM Designs

- Tested 15 designs (CCD (axial distance of 1 and $\sqrt{m}$) and BBD) ranging from 3 to 7 factors
- For each design, 27 settings of the testbed
- 1000 simulations each
- Error variance σ = 0.5



CCD



BBD

Simulation Study

Models and Surfaces

- Simulation assumed no previous screening therefore:
  - its possible that all m factors are active
  - its also possible that only a subset factors are active
- Testbed inputs allow:
  - Just a proportion of terms of each type active (ME's, 2fi's, quadratic terms)
    - Heredity based on findings from Ouckly et al. (2017)
  - All 2fi's and/or all quadratic terms are active
  - True response surface more complex than quadratic

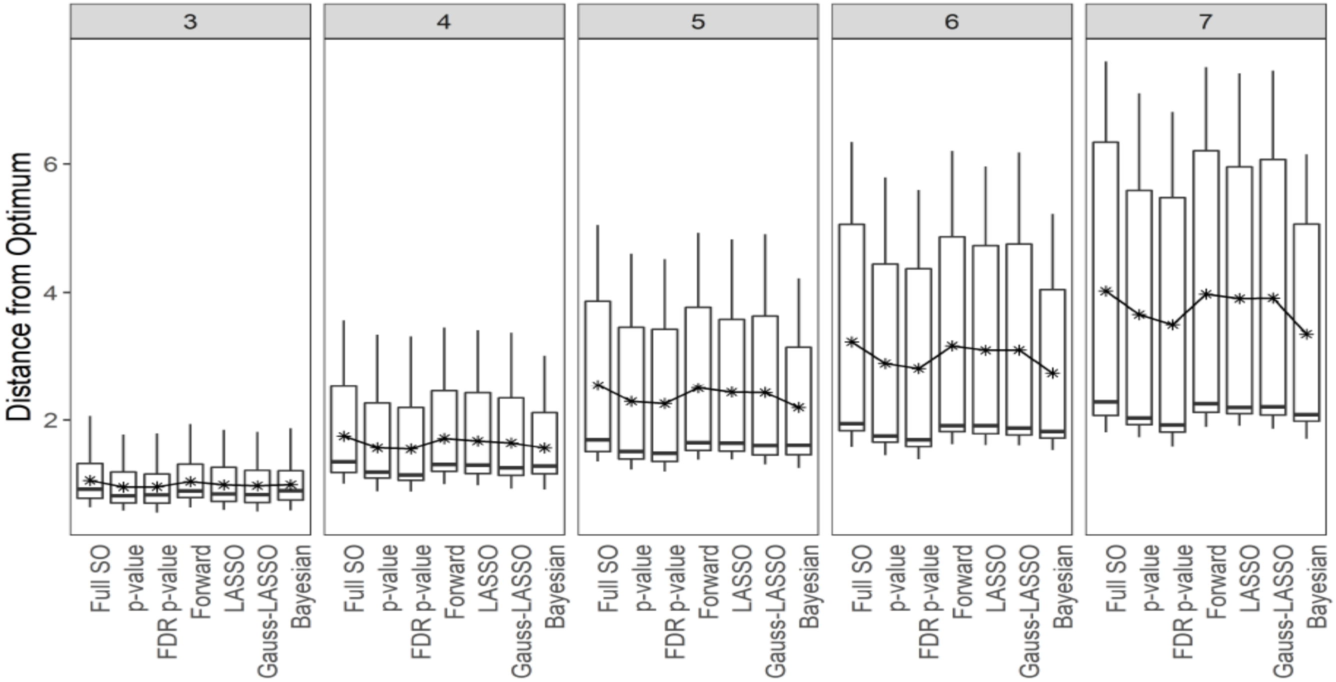
Simulation Study

Validation Points (Confirmation Runs)

- Validation points:
  - 1 center run
  - 4 randomly selected non-center run design points
  - 4 randomly selected design points from the design space.
- We compare the different analysis methods based on their predictive performance on the 9 validation points as well as their ability to locate the true optimum.
- Scenarios are omitted for which more than 10% of the simulations results exhibit lack-of-fit.

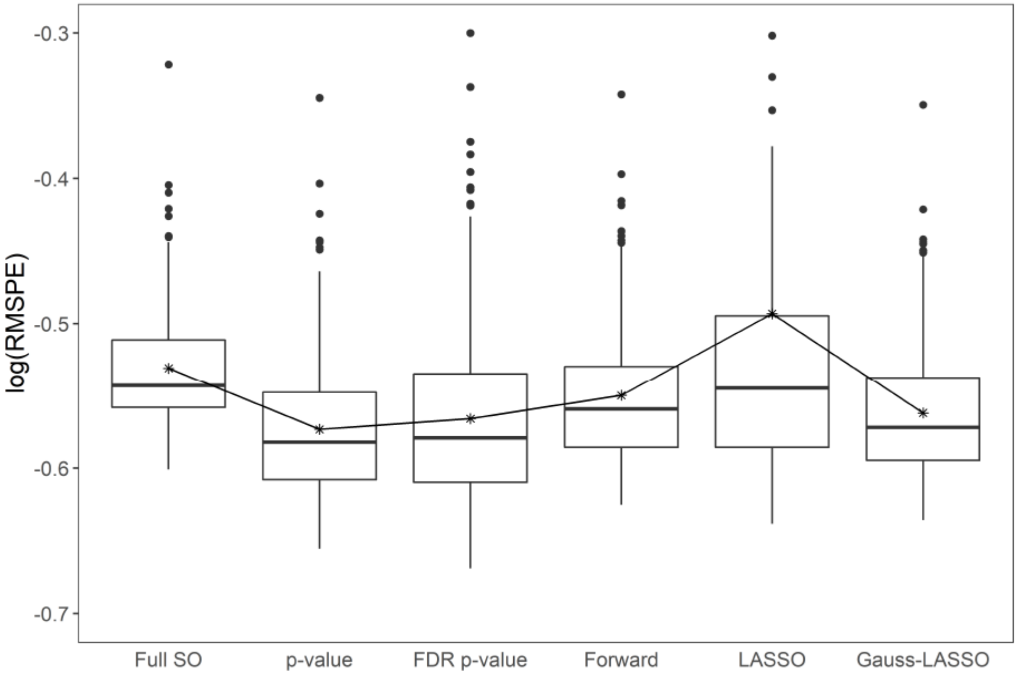
Graphical Results

Distance from Optimum



Graphical Results

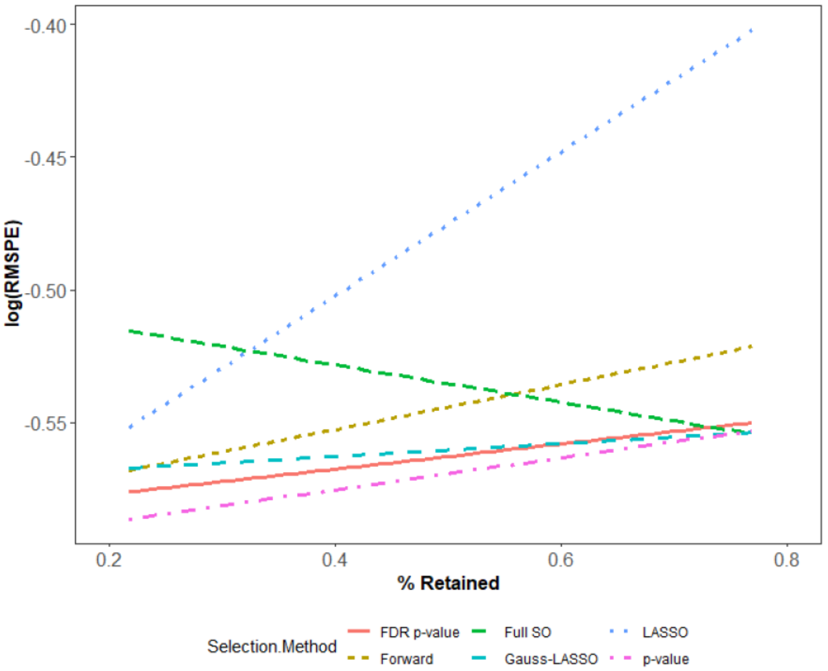
Prediction Performance



Model Results

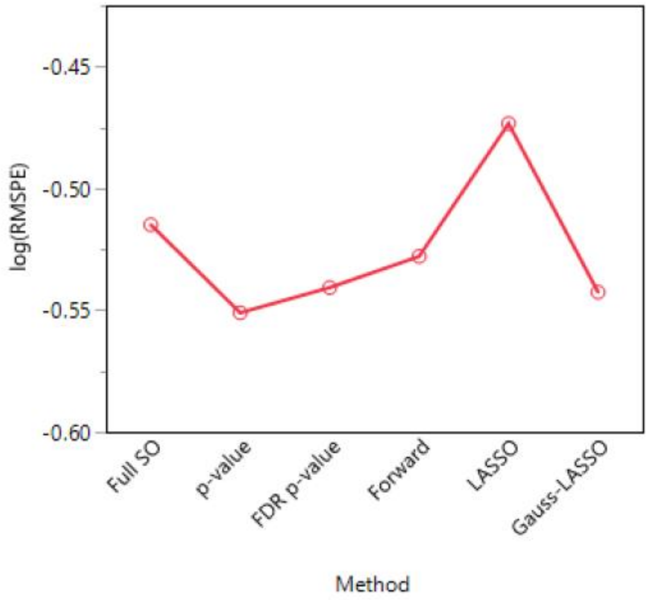
Formal Analysis: % Retained (model size) by Analysis Method Interaction

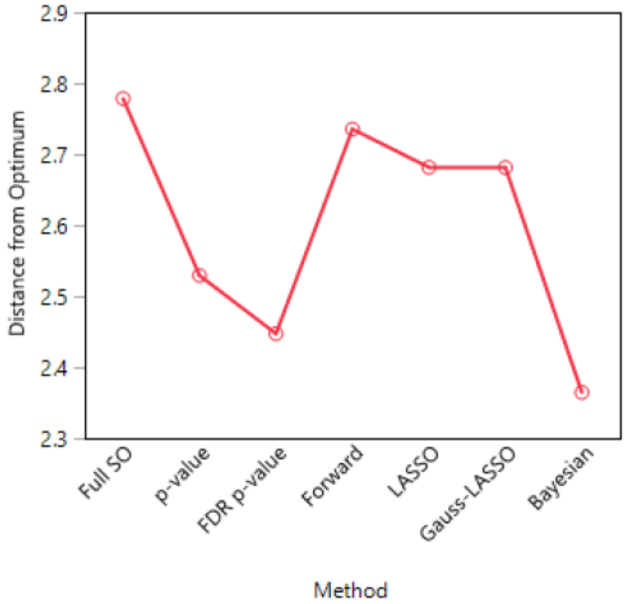
- When the average percent retained is less than roughly 35% all five methods outperform the full second order model.
- The competitive advantage of the reducing methods relative to the full model is diminished as the average percent retained increases.



Model Results

Formal Analysis: Main Effect Plots from models





Conclusions

Optimization

- Optimization is best done using the Bayesian method, or by first reducing using the FDR-adjusted p-value method.

- Using the full second-order model is not recommended for optimization, if there are many inactive terms.

Conclusions

Prediction

- For more general prediction of out-of-sample points, using the unadjusted p-value method will be effective.
- The FDR-adjusted p-value method and Gauss-LASSO perform well also.

Conclusions

General

- The full second-order model is not recommended when many terms are inactive.

- As the underlying true models get larger, the full second-order model eventually predicts as well as the best of the other methods.

One step RSM

A few comments

- A one-shot experiment not only means that screening hasn't been done, but also first-order line searches have been neglected and the quality of the model and the optimum is likely degraded because the design region is necessarily larger.

- When the ideal (sequential experimentation) path cannot be followed we hope this work provides some insight.

Thank you!

Maria L. Weese
weeseml@miamioh.edu
Twitter: @MFWeese

Smucker, B.J., Edwards, D.J., and Weese, M.L. (2020). "Response surface models: To reduce or not to reduce?", Journal of Quality Technology, DOI: 10.1080/00224065.2019.1705208