9-15-2016

Non-Linear Metamodeling Extensions to the Robust Parameter Design of Computer Simulations

Joseph P. Bellucci

**NON-LINEAR METAMODELING EXTENSIONS TO THE
ROBUST PARAMETER DESIGN OF COMPUTER SIMULATIONS**

DISSERTATION

Joseph P. Bellucci, Maj, USAF

AFIT-ENS-DS-16-S-026

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENS-DS-16-S-026

NON-LINEAR METAMODELING EXTENSIONS TO THE
ROBUST PARAMETER DESIGN OF COMPUTER SIMULATIONS

DISSERTATION

Presented to the Faculty

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy in Operations Research

Joseph P. Bellucci, BS, MS

Major, USAF

September 2016

AFIT-ENS-DS-16-S-026

NON-LINEAR METAMODELING EXTENSIONS TO THE
ROBUST PARAMETER DESIGN OF COMPUTER SIMULATIONS

Joseph P. Bellucci, BS, MS
Major, USAF

Committee Membership:

Kenneth W. Bauer, Jr., PhD
Chair

John O. Miller, PhD
Member

Lt Col Andrew J. Geyer, PhD
Member

ADEDEJI B. BADIRU, PhD
Dean, Graduate School of Engineering and Management

AFIT-ENS-DS-16-S-026

Abstract

Robust parameter design (RPD) is used to identify a system's control settings that offer a compromise between obtaining desired mean responses and minimizing the variability about those responses. Two popular combined-array strategies—the response surface model (RSM) approach and the emulator approach—are limited when applied to simulations. In the former case, the mean and variance models can be inadequate due to a high level of non-linearity within many simulations. In the latter case, precise mean and variance approximations are developed at the expense of extensive Monte Carlo sampling.

This research combines the RSM approach's efficiency with the emulator approach's accuracy. Non-linear metamodeling extensions, namely through Kriging and radial basis function neural networks, are made to the RSM approach. The mean and variance of second-order Taylor series approximations of these metamodels are generated via the Multivariate Delta Method and subsequent optimization problems employing these approximations are solved. Results show that improved prediction models can be attained through the proposed approach at a reduced computational cost. Additionally, a multi-response RPD problem solving technique based on desirability functions is presented to produce a solution that is mutually robust across all responses. Lastly, quality measures are developed to provide a holistic assessment of several competing RPD strategies.

To my wife and children for their undying love, sacrifice, and encouragement.

Thank you for completing this journey with me.

Acknowledgments

Although my name appears on the cover of this dissertation, it is only through the help and support of several people that this work reached fruition. Foremost, I wish to extend my sincerest gratitude to my advisor, Dr. Kenneth Bauer, for his remarkable guidance, incredible patience, and constant motivation. His enthusiasm for research and scholarship is infectious. I hope you can take as much pride in this is work as I do. I would also like to thank the members of my research committee, Dr. J. O. Miller and Lt Col Andrew Geyer, for their time and support throughout this arduous, but rewarding, process. Finally, I wish to express my appreciation to my officemates, Mike and Trevor, for their friendship over the last three years. Thank you for the conversations, both professional and juvenile. I wish you continued success in your careers and personal lives.

Joseph P. Bellucci

Table of Contents

List of Figures

List of Tables

NON-LINEAR METAMODELING EXTENSIONS TO THE

ROBUST PARAMETER DESIGN OF COMPUTER SIMULATIONS

I.   Introduction

1.1 Background

Computer simulations are mathematical models of complex real-world systems

for which it would be too expensive, too time consuming, too dangerous, or even

impossible to examine with physical experimentation. They have been applied across

such diverse disciplines as manufacturing, military applications, logistics and supply

chain management, and health care. Simulations, which are either deterministic or

stochastic in nature, can be used to gain insight into a system, examine "what-if"

scenarios, or perform system optimization [1]. Deterministic simulations yield the same

outputs when multiple replications are performed at a single design setting. Stochastic

simulations, on the other hand, generate different outputs when replications are

performed at a single design setting due to the presence of random variables within the

model [2].

This research focuses on stochastic simulations in which some parameters, called

control factors, can be set to specific values whereas other parameters, called *noise*

factors, are modeled by random variables with specific probability distributions. The

noise factors introduce undesirable variation in the system's outputs. When noise factors

are present, robust parameter design (RPD) principles can be applied as a cost-effective

strategy for identifying the ideal control factor settings that result in outputs that are robust, or insensitive, to the random fluctuation of the noise factors. Simulations are well-suited for RPD studies since the parameters, whether they are control factors or noise factors, can be manipulated and set to specific values. In fact, Kleijnen et al. [3] state that finding robust policies or decisions is one of the basic goals of simulations. RPD is a method for determining the control factor settings that reach a compromise between obtaining a desired mean response while minimizing the variability about that response [4].

In the single-response RPD problem, the objective is to identify the control factor setting that yields a desired mean response with minimum variance [4]. The literature offers several strategies that have stemmed from Taguchi's original RPD principles. However, this research is motivated by those strategies that utilize Welch et al.'s [5] combined-array design of experiment (DOE). Specifically, the methods of interest are the combined-array response surface model (RSM) approach [6–9] and the stochastic emulator approach [10–13]. Though each strategy has its own merits, they can be inappropriate or cumbersome when faced with the highly non-linear nature of typical simulations.

In the multi-response RPD problem, the objective is to find the optimal control parameter levels that return average responses close to their target values while minimizing the variance of each response. The literature proposes several methods for optimizing multi-response problems. These methods involve the use of desirability functions [14–18], loss functions [19–23], principal component analysis (PCA) [24–27], distance metrics [28, 29], and mean square error (MSE) criterion [30–32]. A majority of

these techniques transform the set of quality characteristics into new response variables in order to reduce the dimension of the optimization problem. These existing methods seek to find an optimal balance of means and variances across the set of responses. However, in some instances, the mean or variance of one response may influence the solution in such a way that the means and variances of the remaining responses are insignificant to the overall RPD problem. In this case, it can be difficult to attain a solution that is balanced across the set of responses.

Numerous organizations utilize RPD principles as an economical strategy for developing a product or process that is insensitive to a variety of operating conditions. Advantages of employing these principles are that the product or process will be *on target* while exhibiting less variability. This subsequently increases the end user's appeal for the product or process since it won't be as susceptible to deterioration and can be used in diverse situations. This research effort aims to advance the current RPD problem solving approaches.

1.2 Research Objectives

This dissertation strives to meet three key objectives. The first objective is to broaden the combined-array RSM approach that relies exclusively on low-order polynomial models. Since more accurate predictive response surface models result in better RPD solutions [33], a methodology will be developed that utilizes non-linear modeling efforts, such as Kriging and radial basis function neural networks (RBFNNs), in place of polynomial models. The second objective is to develop an approach for multi-response RPD problems that provides a collaborative solution that is balanced across the

3

means and variances of each response. Finally, the third objective is to generate a framework for comparing different RPD problem solving strategies via quality measures. Such measures can increase the understanding of each approach and allow the analyst to make a more knowledgeable evaluation of the competing procedures.

1.3 Chapter Overview

This dissertation is organized in the following manner. Chapter II establishes the foundation for the techniques utilized in this research by reviewing the pertinent literature in the areas of RPD and response modeling. Chapter III details the extension of the combined-array RSM approach to include the application of Kriging and RBFNN metamodels. Chapter IV proposes a multi-response RPD methodology based on desirability functions that generates solutions that are well-balanced across the means and variances of each response. Chapter V describes a method for comparing RPD problem solving strategies via quality measures. Chapter VI summarizes the research contributions and identifies opportunities for future research. Finally, supplemental mathematical derivations and figures are presented in the Appendices.

II. Pertinent Literature

2.1 Introduction

This chapter summarizes pertinent literature in the areas of RPD and response modeling. It is intended to establish the foundation for the techniques that are utilized and expanded upon in this dissertation.

2.2 Robust Parameter Design Approaches

This research focuses on the robust design of stochastic simulations in which some parameters, called *control factors*, can be set to specific values whereas other parameters, called *noise factors*, are modeled by random variables with specific probability distributions. The noise factors introduce undesirable variation in the system's output. When noise factors are present, RPD principles can be applied as a cost-effective strategy for identifying the ideal control factor settings that result in outputs that are robust, or insensitive, to the random fluctuation of the noise factors.

Simulations are well-suited for RPD studies since the parameters, whether they are control factors or noise factors, can be manipulated and set to specific values. RPD is a method for determining a system's ideal control factor setting that reaches a compromise between obtaining a desired target on the mean response while minimizing the variability of the system around that mean response. Three popular strategies for solving RPD problems are Taguchi's method [4], the response surface model (RSM) approach [6], and the stochastic emulator approach [11]. However, before discussing

these current strategies, this section will first illustrate the sources that cause variation in a system's output.

2.2.1 Sources of Output Variation

As seen in Figure 1a, a system can be modeled as a set of resources that transforms inputs into measureable outputs. This figure expands Figure 1-1 in Montgomery [9] to simulations. Noise causes variability in the outputs through any combination of four separate sources: uncontrollable factors, uncertainty in the control factors, uncertainty in the inputs, and internal system variability. Uncontrollable factors are system parameters that are difficult, costly, or impossible to control under normal operation of the system. In industrial scenarios, these factors are usually assumed controllable for the purposes of experimentation. Uncertainty in a control factor reflects variation about its nominal setting. Uncertainty in an input involves variability of the system's inputs. Finally, internal variability relates to intrinsic system variation. For example, in a discrete event simulation, intrinsic system variation occurs due to factors such as the initial state of the system, the warm-up period, the termination conditions, or the random number stream.

In this research, the term *noise factor* represents uncontrollable factors, uncertainty in the control factors, and uncertainty in the inputs. The set of noise factors is denoted by the vector **z** whereas the set of control factors is specified by the vector **x**. Typically, it is assumed that the noise factors are mutually independent and that each noise factor z_i is a normally distributed random variable with known mean μ_i and variance

σ_i^2. Therefore, $E[\mathbf{z}] = \boldsymbol{\mu}_{\mathbf{z}} = \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix}'$ and $Var[\mathbf{z}] = \Sigma_{\mathbf{z}} = diag\left(\sigma_1^2, ..., \sigma_n^2\right)$. Figure 1b illustrates the sources of variation for a scenario used in Chapter III.
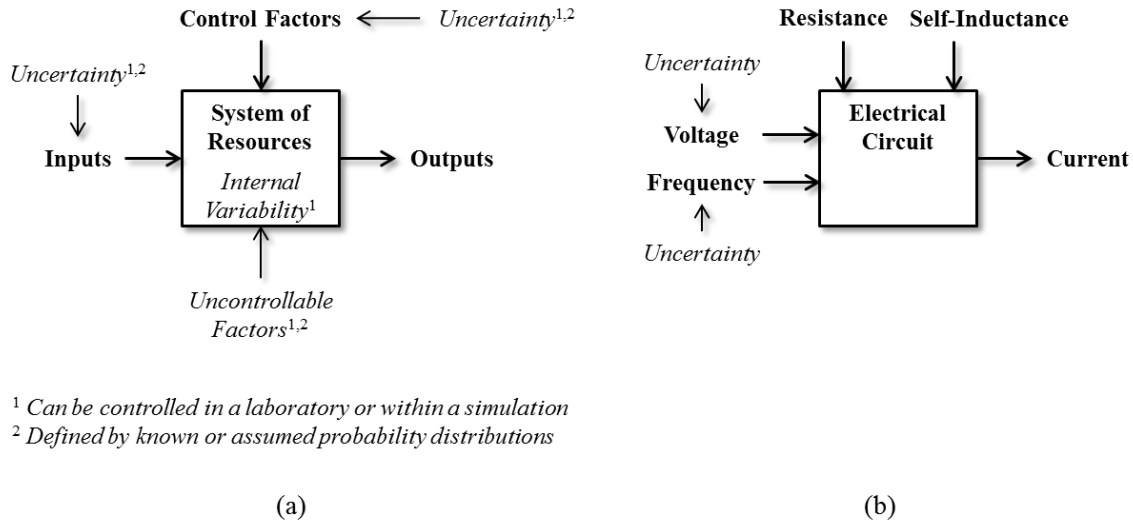
(a)   (b)

Figure 1. Illustrations of Sources of Output Variation for (a) a General System and (b) a Circuit Simulation

2.2.2 Taguchi's Method

Taguchi's approach to solving RPD problems crosses an inner orthogonal array of control factors with an outer orthogonal array of noise factors. That is, each combination of control factor settings within the inner array is performed over each combination of noise factor settings in the outer array. Taguchi then summarizes the observations of each inner array trial across the outer array settings via a statistic known as a signal-to-noise ratio (SNR). The SNR accounts for both the process mean and variance. It is then utilized as the response variable for statistical analysis purposes [4].

The goal of the overall experiment determines which of four SNR formulations to use. If the goal is to determine the values of the control factors that result in a minimum

7

response, then the *smaller-the-better* SNR should be utilized. If a maximum response is desired, then the *larger-the-better* SNR is used. Finally, in what is known as the *target-is-best* case, the goal may be to determine the control settings that yield a response near some desired target value. In this case, Taguchi first recommends using bias-eliminating tuning factors that result in an expected process response equal to the target. Then the form of the SNR is based on whether or not the response mean and variance are independent [9, 34]. These four SNRs are shown in Table 1 where $y_{i,j}$ is the response for inner array setting i and outer array setting j, $\overline{y}_i$ is the sample mean of inner array setting i, and s_i^2 is the sample variance of inner array setting i.

Table 1. Taguchi's Signal-to-Noise Ratios for Three Experimental Scenarios

Smaller-the-Better	Larger-the-Better	Target-is-Best
		Independent mean & variance
		$SNR_i = 10 \cdot \log\left(s_i^2\right)$
$SNR_i = -10 \cdot \log\left(\sum_{j=1}^{n} \dfrac{y_{i,j}^2}{n}\right)$	$SNR_i = -10 \cdot \log\left(\sum_{j=1}^{n} \dfrac{1/y_{i,j}^2}{n}\right)$	*Correlated mean & variance*
		$SNR_i = 10 \cdot \log\left(\dfrac{\overline{y}_i^2}{s_i^2}\right)$

Regardless of which SNR is utilized, modeling analysis of the system attempts to maximize the SNR while driving the mean response towards some preferred target value. Taguchi methods typically develop only main-effects models and are not concerned with control factor interactions. Analysis of Taguchi's approach is a two-stage method. First, the experimenter should choose the levels of those significant control factors that maximize the SNR. Second, the experimenter should choose the levels of the remaining

8

control factors that result in a mean response near the desired target value [34]. Standard ANOVA techniques can also help distinguish which control factors affect the average response and SNR. Control settings that are robust and insensitive to the variance caused by the noise factors are then determined [9].

If it can be assumed that there are no significant interactions among the control variables, then Taguchi's methodology is highly appropriate for identifying robust control parameter settings. However, Taguchi's critics claim that, more often than not, this assumption does not hold and that a main effects-only study may yield ambiguous results [7, 35]. Detractors also note that the use of Taguchi's SNRs, though they are concerned with the process mean and variance, don't allow for a complete understanding of which control factors affect the mean and which affect the variance [9, 34, 36]. Nair and Shoemaker [37] further argue that critical system information is lost by compressing the experimental responses into SNRs. Several authors claim that it is better to examine the mean and variance separately instead of combining them into a single SNR [38–40]. Even though the purpose behind SNRs is to uncouple the location and dispersion effects, Montgomery [9] contends that there is no assurance that this will occur. As an example, he shows how the use of the *smaller-the-better* SNR actually confounds location and dispersion effects. Pignatiello [41] states that, despite the criticism of Taguchi's tactics, his conceptual framework for planning a product or process design experiment is fundamentally sound.

2.2.3 Response Surface Model Approaches

The response surface model (RSM) approach solves an optimization problem involving models of the system's mean and variance. Several optimization schemes are discussed in Section 2.4. Two major RSM applications are found in the literature. One approach utilizes replications of an experimental design whereas the second approach uses a combined-array experimental design.

2.2.3.1 RSM Approach Using Replicated Experimental Designs

The first RSM application uses replications of a design of experiment (DOE) consisting of **x** only. This approach is used when the system's output variability is due to either random sampling from **z** or intrinsic system variation. Separate low-order polynomials [6], higher-order polynomials [33, 42], Kriging models [11, 13, 43], radial basis function approximations [43], or radial basis function neural networks [43] are fit to the sample means and variances of the design points to generate mean and variance models. This RSM application has been applied to simulations of a piston [11] and an optical profilometer [42]. Wild and Pignatiello [44] used a partitioning strategy to assess the system's mean and variance through crossed-array designs. They used a discrete event simulation for the robust design of a manufacturing facility. Several authors have also applied this RSM application to Box and Draper's [45] well-examined printing process study [33, 43, 46–48].

2.2.3.2 RSM Approach Using Combined-Array Experimental Designs

The second RSM application uses the combined-array DOE proposed by Welch et al. [5]. The combined-array is comprised of both **x** and **z**. A low-order polynomial

response model $\hat{y}(\mathbf{x}, \mathbf{z})$ is built to accommodate the linear control factor and noise factor

effects, the pure quadratic control factor effects, the control factor interaction effects, and

the control factor by noise factor interaction effects [7]. Mean and variance models are

then obtained by taking the expectation and variance of $\hat{y}(\mathbf{x}, \mathbf{z})$. Recent extensions have

been made to include pure quadratic noise factor effects and noise factor interaction

effects [49] as well as three-factor control by control by noise factor interactions and

control by noise by noise factor interactions [50]. The combined-array RSM approach

has been applied to a piston simulation [11] and an economic order quantity inventory

model [12, 13]. A textbook example of this approach for a physical experiment can be

found for a chemical production process in Montgomery [9].

2.2.4 Stochastic Emulator Approaches

The literature discusses two main applications of the stochastic emulator approach

for solving RPD problems for stochastic simulations. Regardless of the application, the

first step is to generate a model, or emulator, of the system. The emulator is then used to

generate large samples of data from the joint distribution of $\mathbf{z}$ at the design points in $\mathbf{x}$.

Mean and variance models are fit from the samples and a subsequent optimization

problem consisting of the mean and variance models is solved.

2.2.4.1 Emulator Approach with Control Factor Uncertainty

In the first emulator application, uncertainty in the control factors causes variation

in the system's output. In this case, control factor setting $\hat{x}_i$ can be decomposed into the

sum of the nominal control factor setting x_i and the normally distributed noise factor z_i.

11

That is, $\hat{x}_i = x_i + z_i$ where $E[\hat{x}_i] = x_i$ and $Var[\hat{x}_i] = \sigma_i^2$; hence $E[\hat{\mathbf{x}}] = \mathbf{x}$ and $Var[\hat{\mathbf{x}}] = \Sigma_{\mathbf{z}}$.

Initially, the noise factors are ignored by setting $\mathbf{z} = \mathbf{0}$ and the emulator $\hat{y}(\hat{\mathbf{x}} = \mathbf{x})$ is built from a DOE in the control factors $\mathbf{x}$. A second DOE in $\mathbf{x}$ is then constructed and the emulator is used to evaluate the noise factors' effect on the output. For each design point $\mathbf{x}_d$, a large sample of responses is generated from the normal distribution with mean $\mathbf{x}_d$ and covariance $\Sigma_{\mathbf{z}}$ as evaluated using the emulator. Separate models are fit to the sample means and variances of these responses. A robust control setting is determined by optimizing a problem involving the emulator's mean and variance models. Bates et al. [11] used this emulator strategy in an RPD study on a piston simulation.

2.2.4.2 Emulator Approach with Uncontrollable Factors or Input Uncertainty

In the second emulator application, variation in the system's output is due to uncertainty in the inputs and/or the existence of uncontrollable factors. The emulator is built from a combined-array DOE including $\mathbf{x}$ and $\mathbf{z}$. The high and low experimental levels for each z_i are set at $\mu_i \pm 3\sigma_i$ [13]. Once the emulator $\hat{y}(\mathbf{x}, \mathbf{z})$ is built, a second DOE in $\mathbf{x}$ is constructed. For each design point $\mathbf{x}_d$, a large sample is generated from the joint distribution of $\mathbf{z}$ and is evaluated using the emulator. As in the first emulator approach, mean and variance models are fit to the sample means and variances calculated at each design point. These models are then used in an optimization formulation to find a robust control setting. This RPD emulator strategy has been performed on simulations for a 2-bar truss design problem [10] and an economic order quantity inventory model [13].

2.3 Robust Parameter Design Modeling Strategies

The RPD approach developed in this research will be compared to the combined-array RSM approach and the combined-array stochastic emulator approach. Thus, this section will discuss the modeling efforts that have been utilized for those two approaches.

2.3.1 Low-Order Polynomial Models

There are four modeling strategies that have been applied to the combined-array RSM approach. Each builds low-order polynomial response models comprised of $r_\mathbf{x}$ control factors $\mathbf{x}$ and $r_\mathbf{z}$ noise factors $\mathbf{z}$. Each approach discussed in this subsection, expands upon the model that precedes it. They each assume that the model's error ε is normally distributed with a mean of zero and a constant variance σ_ε^2; hence any non-constant variance stemming from the process is attributed to the inability to control the noise factors [8]. They also assume that each noise factor z_i is a normally distributed random variable with a known mean μ_i and variance σ_i^2. For experimental purposes, the natural level of each noise factor is centered at its mean and its ± 1 levels are set at $\mu_i \pm \sigma_i$. Furthermore, the noise variables are assumed to be uncorrelated. Therefore, the mean and variance of $\mathbf{z}$ are defined as $E[\mathbf{z}] = \boldsymbol{\mu}_\mathbf{z}$ and $Var[\mathbf{z}] = \Sigma_\mathbf{z} = diag\left(\sigma_1^2, ..., \sigma_{r_z}^2\right)$, respectively. Thus, if the factors have been transformed to the coded variable space, $E[\mathbf{z}] = \mathbf{0}_{r_\mathbf{z}}$ and $Var[\mathbf{z}] = \mathbf{I}_{r_\mathbf{z}}$ [9].

13

2.3.1.1 The Standard RPD Response Model

The standard (*Std*) form of the RPD response model in Equation (1) considers the noise factor interactions and the pure quadratic noise factor effects negligible [8].

$$y_{Std}\left(\mathbf{x},\mathbf{z}\right)=\beta_0+\mathbf{x}'\boldsymbol{\beta}+\mathbf{x}'\mathbf{Bx}+\left(\boldsymbol{\gamma}'+\mathbf{x}'\boldsymbol{\Delta}\right)\mathbf{z}+\varepsilon \tag{1}$$

In Equation (1), β_0 represents the intercept, $\mathbf{x}$ is the $r_\mathbf{x}\times1$ vector of control factors, $\mathbf{z}$ is the $r_\mathbf{z}\times1$ vector of noise factors, $\boldsymbol{\beta}$ is the $r_\mathbf{x}\times1$ vector of linear control factor effects, $\mathbf{B}$ is the $r_\mathbf{x}\times r_\mathbf{x}$ matrix where the pure quadratic control factor effects are on the diagonal and one-half of the control factor interaction effects are on the off-diagonal, $\boldsymbol{\gamma}$ is the $r_\mathbf{z}\times1$ vector of linear noise factor effects, and $\boldsymbol{\Delta}$ is the $r_\mathbf{x}\times r_\mathbf{z}$ matrix of control factor by noise factor interaction effects.

Given the distributional assumptions of the noise variables and the model's error, the mean and variance of the standard response model in Equation (1) can be written respectively as

$$E\left[y_{Std}\left(\mathbf{x},\mathbf{z}\right)\right]=\beta_0+\mathbf{x}'\boldsymbol{\beta}+\mathbf{x}'\mathbf{Bx} \tag{2}$$

and

$$Var\left[y_{Std}\left(\mathbf{x},\mathbf{z}\right)\right]=\left(\boldsymbol{\gamma}'+\mathbf{x}'\boldsymbol{\Delta}\right)\Sigma_\mathbf{z}\left(\boldsymbol{\gamma}'+\mathbf{x}'\boldsymbol{\Delta}\right)'+\sigma_\varepsilon^2 \tag{3}$$

where σ_ε^2 is estimated using the MSE of the fitted response surface model. Note that Equations (2) and (3) are completely in terms of the control factors $\mathbf{x}$. These two response models can now be used to estimate the mean and variance of the process at any point within the control design space.

2.3.1.2 The Noise-by-Noise Response Model

Mindrup et al. [49] found that Equation (1) was inadequate in certain imaging applications. As a result, they removed the assumption that the noise factor interactions and the pure quadratic noise factor effects are insignificant and developed the noise-by-noise (*NN*) response model, mean model, and variance model in Equations (4)–(6).

$$y_{NN}(\mathbf{x}, \mathbf{z}) = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + (\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})\mathbf{z} + \mathbf{z}'\boldsymbol{\Phi}\mathbf{z} + \varepsilon \tag{4}$$

$$E\left[y_{NN}(\mathbf{x}, \mathbf{z})\right] = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + tr(\boldsymbol{\Phi}\boldsymbol{\Sigma}_{\mathbf{z}}) \tag{5}$$

$$Var\left[y_{NN}(\mathbf{x}, \mathbf{z})\right] = (\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})\boldsymbol{\Sigma}_{\mathbf{z}}(\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})' + 2tr(\boldsymbol{\Phi}\boldsymbol{\Sigma}_{\mathbf{z}}\boldsymbol{\Phi}\boldsymbol{\Sigma}_{\mathbf{z}}) + \sigma_\varepsilon^2 \tag{6}$$

In Equations (4)–(6), *tr* represents the trace of a square matrix and $\boldsymbol{\Phi}$ is the $r_{\mathbf{z}} \times r_{\mathbf{z}}$ matrix with pure quadratic noise factor effects on the diagonal and one-half of the noise factor interaction effects on the off-diagonal.

2.3.1.3 The Control-by-Noise-by-Noise Response Model

Williams et al. [50] further extended Mindrup et al.'s work in two successive expansions. First, they appended the three-factor control-by-noise-by-noise (*CNN*) interactions to generate the *CNN* RPD response model, mean model, and variance model in Equations (7)–(9).

$$y_{CNN}(\mathbf{x}, \mathbf{z}) = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + (\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})\mathbf{z} + \mathbf{z}'(\boldsymbol{\Phi} + \boldsymbol{\Psi}_{\mathbf{x}})\mathbf{z} + \varepsilon \tag{7}$$

$$E\left[y_{CNN}(\mathbf{x}, \mathbf{z})\right] = \beta_0 + \mathbf{x}'\boldsymbol{\beta} + \mathbf{x}'\mathbf{B}\mathbf{x} + tr((\boldsymbol{\Phi} + \boldsymbol{\Psi}_{\mathbf{x}})\boldsymbol{\Sigma}_{\mathbf{z}}) \tag{8}$$

$$Var\left[y_{CNN}(\mathbf{x}, \mathbf{z})\right] = (\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})\boldsymbol{\Sigma}_{\mathbf{z}}(\boldsymbol{\gamma}' + \mathbf{x}'\boldsymbol{\Delta})' + 2tr((\boldsymbol{\Phi} + \boldsymbol{\Psi}_{\mathbf{x}})\boldsymbol{\Sigma}_{\mathbf{z}})^2 + \sigma_\varepsilon^2 \tag{9}$$

The additional term in Equations (7)–(9) is $\mathbf{\Psi_x} = \sum_{i=1}^{r_x} \mathbf{\Psi}_i x_i$ where $\mathbf{\Psi}_i$ is the $r_z \times r_z$ matrix of control-by-noise-by-noise interaction effects corresponding to control factor x_i.

2.3.1.4 The Control-by-Control-by-Noise Response Model

Williams et al. [50] then further extended the *CNN* RPD models to include the three-factor control-by-control-by-noise (*CCN*) interaction terms. The *CCN* RPD response model, mean model, and variance model are shown in Equations (10)–(12).

$$y_{CCN}(\mathbf{x},\mathbf{z}) = \beta_0 + \mathbf{x'\beta} + \mathbf{x'Bx} + (\mathbf{\gamma'} + \mathbf{x'\Delta} + \tilde{\mathbf{x}})\mathbf{z} + \mathbf{z'}(\mathbf{\Phi} + \mathbf{\Psi_x})\mathbf{z} + \varepsilon \qquad (10)$$

$$E\left[y_{CCN}(\mathbf{x},\mathbf{z})\right] = \beta_0 + \mathbf{x'\beta} + \mathbf{x'Bx} + tr\left((\mathbf{\Phi} + \mathbf{\Psi_x})\Sigma_\mathbf{z}\right) \qquad (11)$$

$$Var\left[y_{CCN}(\mathbf{x},\mathbf{z})\right] = (\mathbf{\gamma'} + \mathbf{x'\Delta} + \tilde{\mathbf{x}})\Sigma_\mathbf{z}(\mathbf{\gamma'} + \mathbf{x'\Delta} + \tilde{\mathbf{x}})' + 2tr\left((\mathbf{\Phi} + \mathbf{\Psi_x})\Sigma_\mathbf{z}\right)^2 + \sigma_\varepsilon^2 \qquad (12)$$

The additional term in Equations (10)–(12) is the vector $\tilde{\mathbf{x}} = \left[\mathbf{x'\Omega}_1\mathbf{x},\ \mathbf{x'\Omega}_2\mathbf{x},\ \dots\ ,\ \mathbf{x'\Omega}_{r_z}\mathbf{x}\right]$ where $\mathbf{\Omega}_j$ is the $r_x \times r_x$ matrix of control-by-control-by-noise interaction effects corresponding to noise factor z_j.

2.3.2 Kriging

Kriging has been used to model the simulation's response, as well as its mean and variance, within the combined-array stochastic emulator strategy in Section 2.2.4. Kriging is a nonparametric, global, exact interpolation model. The term *nonparametric* implies that training points are used to both estimate the unknown model parameters and predict the response of new observations [43]. *Global* means that a Kriging model provides predictions across the entire experimental area and *exact* indicates that Kriging models predict the precise response of previously observed input combinations [51].

16

These properties are the reason as to why Kriging metamodels have become popular approaches for estimating the output of computer simulation models [2, 11, 43, 51–58]. The origins of Kriging are in geostatistics, or spatial statistics [59].

The Kriging model for the input vector $\mathbf{v}$ assumes the form

$$y(\mathbf{v}) = f(\mathbf{v}) + \delta(\mathbf{v}). \tag{13}$$

In Equation (13), $f(\mathbf{v})$ models the trend in the data and provides a global approximation of the design space [56]. *Ordinary Kriging* assumes that $f(\mathbf{v}) = \beta$ is the constant mean of the data in the experimental region of interest [2] whereas *Universal Kriging* uses a low-order polynomial to define $f(\mathbf{v})$ [43]. Also in Equation (13), $\delta(\mathbf{v})$ creates localized deviations [56] and is additive noise formed by a stationary covariance process with zero mean, variance σ^2, and covariance $\sigma^2\mathbf{R}$ in which the i,j^{th} element of the $N \times N$ spatial correlation matrix $\mathbf{R}$ is defined by

$$\mathbf{R}_{i,j} = \begin{cases} 1 & \text{for } i = j \\ \exp\left[-\sum_{k=1}^{K} \theta_k \left| v_k^{(i)} - v_k^{(j)} \right|^p \right] & \text{for } i \neq j \end{cases} \tag{14}$$

where $v_k^{(i)}$ is the kth feature of the ith training vector, $\theta_k > 0$ is the weight factor for the kth input vector feature, p is a parameter that defines the correlation between two training points, N is the number of training vectors, and K is the number of features in each training vector. This research uses Ordinary Kriging since that has been found to be sufficient for simulation models [43, 54, 60]. The Gaussian correlation function, where $p = 2$, is also utilized due to its widespread appeal [56].

Estimates for the parameters β, σ^2, and θ_k are determined through the maximized

likelihood estimation (MLE) approach. Interested readers can reference [43] for the

MLE derivations. As a result, given a set of training vectors, the Kriging model response

for a new observation $\mathbf{v}$ is

$$\hat{y}(\mathbf{v}) = \hat{\beta} + \mathbf{r}'(\mathbf{v})\mathbf{R}^{-1}(\mathbf{y} - \mathbf{1}\hat{\beta}) \tag{15}$$

where $\hat{\beta} = \dfrac{\mathbf{1}'\mathbf{R}^{-1}\mathbf{y}}{\mathbf{1}'\mathbf{R}^{-1}\mathbf{1}}$ is the mean parameter, $\mathbf{y}$ is the $N \times 1$ vector of training set responses, $\mathbf{1}$ is

a $N \times 1$ vector of ones, and $\mathbf{r}(\mathbf{v})$ is the $N \times 1$ correlation vector whose n^{th} element is the

correlation between $\mathbf{v}$ and the nth training vector $\mathbf{v}^{(n)}$ defined as

$$\mathbf{r}_n(\mathbf{v}) = \exp\left[-\sum_{k=1}^{K} \hat{\theta}_k \left| v_k - v_k^{(n)} \right|^p \right]. \tag{16}$$

2.3.3 Radial Basis Function Neural Networks

Artificial neural networks (ANNs), and in particular radial basis function neural

networks (RBFNNs), have also been used to provide a metamodel of a simulation's

response, mean, and variance within the stochastic emulator approach. ANNs are

mathematical models that update their parameters iteratively to learn the relationship

between a set of inputs and a set of outputs. The RBFNN is a special class of ANN [61–

64].

In a typical RBFNN framework, as illustrated in Figure 2, the number of input

layer neurons is equal to the number of input features and the number of hidden layer

neurons is equal to the number of training vectors. There is also an output layer neuron

for each system output. In this case, there are N training vectors with K features and a

single output y. The activation function of the nth neuron, $h_n(\cdot)$, is a radial basis function designed to *fire high* when a training vector is very close to the center vector $\boldsymbol{\mu}_n$ and give a diminishing response as the training vector moves away from the center vector's receptive field defined by the spread parameter σ_n. This research employs the Gaussian response function

$$h_n(\mathbf{v}) = \exp\left(\frac{-(\mathbf{v}-\boldsymbol{\mu}_n)'(\mathbf{v}-\boldsymbol{\mu}_n)}{2\sigma_n^2}\right). \tag{17}$$

The output layer simply computes a linear weighted sum of the hidden layer neuron activations.
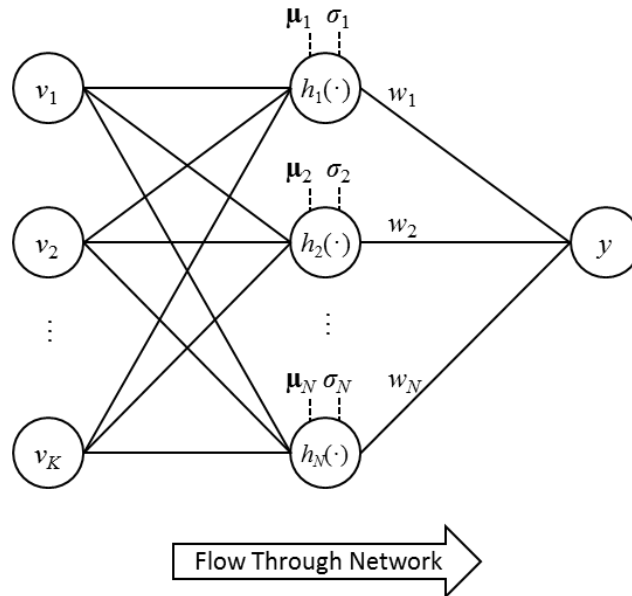
Figure 2. Structure of a Radial Basis Function Neural Network

The network's training phase occurs in a supervised manner. That is, a training set composed of input vectors $\mathbf{v}_1, ..., \mathbf{v}_N$ and their associated target values $\mathbf{t} = \begin{bmatrix} t_1 & \cdots & t_N \end{bmatrix}'$

19

must be available. To begin, the center vectors are placed at the location of the input vectors themselves. That is, $\boldsymbol{\mu}_i = \mathbf{v}_i$ for $i = 1, \dots, N$. The $N \times N$ matrix of hidden layer neuron activations $\mathbf{A}$ is then computed in which the elements of $\mathbf{A}$ are

$$a_{ij} = \exp\left(\frac{-\left(\mathbf{v}_i - \mathbf{v}_j\right)'\left(\mathbf{v}_i - \mathbf{v}_j\right)}{2\sigma_j^2}\right). \tag{18}$$

The weight vector $\mathbf{w} = \begin{bmatrix} w_1 & \cdots & w_N \end{bmatrix}'$ satisfies the equation

$$\mathbf{A}\mathbf{w} = \mathbf{t} \tag{19}$$

and can be written as

$$\mathbf{w} = \mathbf{A}^\dagger \mathbf{t} \tag{20}$$

where $\mathbf{A}^\dagger$ represents the pseudoinverse of $\mathbf{A}$.

Following the training phase, new observations can be presented to the network to generate outputs. Mathematically, an output of the RBFNN for a given K-feature vector $\mathbf{v}$ can be calculated using

$$\hat{y}(\mathbf{v}) = \sum_{n=1}^{N} w_n \exp\left(-\frac{1}{2\sigma_n^2} \sum_{k=1}^{K}\left(v_k - \mu_k^{(n)}\right)^2\right) \tag{21}$$

where w_n is the weight of the nth center vector, v_k is the kth feature of $\mathbf{v}$, and $\mu_k^{(n)}$ is the kth feature of the nth center vector.

There are many benefits for employing ANNs in a study. In fact, the power of utilizing multilayer ANNs comes from the fact that any continuous function can be implemented in a three-layer neural net provided that there is a sufficient number of hidden layer neurons and the proper nonlinear activation functions are chosen [61]. Haykin [63] outlines 9 useful properties and capabilities for ANNs:

1. The network's nonlinear nature is important if the underlying relationships between the components of the input signal are nonlinear.

2. The ANN utilizes *training* samples to learn an input-output mapping that it applies to *new* samples.

3. ANNs are flexible to changes in the operating environment.

4. As it relates to pattern classification, ANNs offer information about which pattern to choose as well as a level of *confidence* in making that choice.

5. The network manages contextual information naturally since each neuron is potentially affected by all of the other neurons in the network.

6. ANNs are considered *fault tolerant* in that when a network's performance degrades, it does so gracefully instead of catastrophically.

7. A neural network is suitable for real-time application in *very-large-scale-integrated* (VLSI) technology that necessitates describing complex behavior in a hierarchical manner.

8. Though there are many types of ANNs, their analysis and design are universal across domains and their commonalities greatly expand the ability to share theories.

9. ANN's correlation to the brain facilitates an expansion in the areas of neural computing as well as neurobiology.

Given all of their beneficial properties and capabilities, there are still many concerns regarding the use of ANNs. First, neural networks are normally used to make predictions about a system rather than to build models or develop any underlying knowledge about that system [9]. Second, there exists a risk of overfitting the data when using ANNs. Neural networks can provide a near-perfect fit to historical or training data, but can be poor at predicting new data [9, 35]. A third issue when working with ANNs lies in adjusting the network's complexity. The large number of free parameters, or weights, in the network creates difficulty in finding a balance between choosing too few and too many neurons to achieve the best generalization of the phenomenon of interest

[61]. Finally, although ANNs can be a powerful and fast tool, they should be used as a supplement, not a substitute, to standard regression and designed experiments statistical tools since they do not allow fundamental insights into the underlying system mechanism that produced the data. Neural networks cannot provide the solution on their own and should be integrated into a consistent system engineering approach [9, 35, 61, 63].

2.4 Optimization Approaches Using the Mean and Variance Models

Given that there are available mean and variance models, $\hat{\mu}$ and $\hat{\sigma}^2$ respectively, a number of dual response optimization approaches can be taken to identify a system's robust control factor setting. Vining and Myers' [6] determined robust operating conditions by optimizing the primary response model subject to a constraint on the secondary response model. In a *smaller-the-better* case, $\hat{\sigma}^2$ is restricted at some specified value σ_T^2 while $\hat{\mu}$ is minimized. A *larger-the-better* situation maximizes $\hat{\mu}$ while controlling $\hat{\sigma}^2$ at some specified value σ_T^2. Finally, in a *target-is-best* case, the concern is maintaining $\hat{\mu}$ at some specified value μ_T while $\hat{\sigma}^2$ is minimized. These three optimization problems are shown in Table 2.

Table 2. Vining and Myers' Dual Response Optimization Scenarios

Smaller-the-Better	Larger-the-Better	Target-is-Best
Minimize $\hat{\mu}$	Maximize $\hat{\mu}$	Minimize $\hat{\sigma}^2$
Subject to $\hat{\sigma}^2 = \sigma_T^2$	Subject to $\hat{\sigma}^2 = \sigma_T^2$	Subject to $\hat{\mu} = \mu_T$

Lin and Tu [46] contended that Vining and Myers' use of equality constraints most likely eliminates finding better global solutions. Lin and Tu make the case that a better solution can be found by allowing some deviation, or bias, of the mean around the target value μ_T while keeping the variance small. Their method solves for optimal control factor settings by estimating the MSE using a function of the process mean and variance. Their MSE criteria for the three experimental scenarios are formulated in Table 3. Optimal control factor settings can be found by solving the optimization problem

$$
\begin{aligned}
&\text{Minimize } MSE \\
&\text{Subject to } \mathbf{x} \in \mathbf{D}
\end{aligned}
\tag{22}
$$

where $\mathbf{D}$ is the experimental design space.

Table 3. Lin and Tu's Dual Response Optimization Scenarios

Smaller-the-Better	Larger-the-Better	Target-is-Best
$MSE = \hat{\sigma}^2 + \hat{\mu}^2$	$MSE = \hat{\sigma}^2 - \hat{\mu}^2$	$MSE = \hat{\sigma}^2 + \left(\hat{\mu} - \mu_T \right)^2$

One criticism of Lin and Tu's *target-is-best* method is that there is no restriction on how far the mean process response may deviate from the target value and, as a result, may be deficient if it is critical to maintain the mean close to the target [65]. In situations such as these, Copeland and Nelson [66] recommended obtaining a solution in which $\hat{\mu}$ is within a specified distance (Δ_μ) of μ_T. They endorsed minimizing $\hat{\sigma}$ subject to

$\left(\hat{\mu} - \mu_T \right)^2 \le \Delta_\mu^2$. The additional *smaller-the-better* and *larger-the-better* instances are shown in Table 4 where Δ_{σ^2} is a maximum allowable value for $\hat{\sigma}^2$.

Table 4. Copeland and Nelson's Dual Response Optimization Scenarios

Smaller-the-Better	Larger-the-Better	Target-is-Best
Minimize $\hat{\mu}$	Minimize $-\hat{\mu}$	Minimize $\hat{\sigma}^2$
Subject to $\hat{\sigma}^2 \leq \Delta_{\sigma^2}$	Subject to $\hat{\sigma}^2 \leq \Delta_{\sigma^2}$	Subject to $\left(\hat{\mu} - \mu_T\right)^2 \leq \Delta_{\mu}^2$

Ding et al. [67] suggested a weighted MSE (*WMSE*) approach by utilizing the convex combination of the mean and variance functions. Their proposal minimizes the *WMSE*s in Table 5 where $\lambda \in [0,1]$.

Table 5. Ding et al.'s Dual Response Optimization Scenarios

Smaller-the-Better	Larger-the-Better	Target-is-Best
$WMSE = \lambda\hat{\mu}^2 + \left(1-\lambda\right)\hat{\sigma}^2$	$WMSE = -\lambda\hat{\mu}^2 + \left(1-\lambda\right)\hat{\sigma}^2$	$WMSE = \lambda\left(\hat{\mu} - \mu_T\right)^2 + \left(1-\lambda\right)\hat{\sigma}^2$

The methodologies of Vining and Myers, Lin and Tu, Copeland and Nelson, and Ding et al. are useful when a single optimal solution is necessary. Köksoy and Doganaksoy [68] present a flexible nonlinear multi-objective approach by considering the secondary response as another primary response. They claim that the restriction placed upon the secondary response may exclude better conditions. Their method, which only focuses on the *smaller-the-better* and *larger-the-better* problem structures, allows further insight into the RPD problem by exploring trade-offs between the mean and variance responses. These formulations are shown in Table 6. Solving these problems results in finding a string of Pareto alternative solutions in some region of interest **R** that jointly optimize $\hat{\mu}$ and $\hat{\sigma}^2$. That is, it is impossible to improve $\hat{\mu}$ without making $\hat{\sigma}^2$ worse and vice versa.

Table 6. Köksoy and Doganaksoy's Dual Response Optimization Scenarios

Smaller-the-Better	Larger-the-Better
$\{$Minimize $\hat{\mu}$, Minimize $\hat{\sigma}^2\}$	$\{$Maximize $\hat{\mu}$, Minimize $\hat{\sigma}^2\}$
Subject to $\mathbf{x} \in \mathbf{R}$	Subject to $\mathbf{x} \in \mathbf{R}$

2.5 Multi-Response Robust Parameter Design Approaches

In the multi-response RPD problem, the objective is to find the optimal control parameter levels that return average responses close to their target values while minimizing the variance of each response. The literature offers several methods for optimizing multi-response problems. These methods involve the use of desirability functions [14–18], loss functions [19–23], principal component analysis (PCA) [24–27], distance metrics [28, 29], and MSE criterion [30–32]. A majority of these techniques transform the quality characteristics into new response variables in order to reduce the dimension of the optimization problem. Typically, the transformations convert the output responses to a single response.

2.5.1 Desirability Functions

Derringer and Suich [14] adopted Harrington's [69] use of desirability functions which map each of the estimated responses $\hat{y}_i$ into a desirability value d_i where $0 \leq d_i \leq 1$. An overall system desirability D is then generated by combining the J individual desirabilities via the geometric mean as in Equation (20).

$$D = \left(\prod_{j=1}^{J} d_j \right)^{1/J}$$

(23)

An optimal operating point for the set of responses is then found by maximizing D. Given the actual desire for the individual output responses in the experiment, such as a minimum or maximum value, the desirability values can be defined as in Table 7.

Table 7. Derringer and Suich's Desirability Functions

Smaller-the-Better	Larger-the-Better	Target-is-Best

$$d_i = \begin{cases} 1 & \hat{y}_i < L_i \\ \left(\dfrac{U_i - \hat{y}_i}{U_i - L_i}\right)^r & L_i \le \hat{y}_i \le U_i \\ 0 & \hat{y}_i > U_i \end{cases}$$

$$d_i = \begin{cases} 0 & \hat{y}_i < L_i \\ \left(\dfrac{\hat{y}_i - L_i}{U_i - L_i}\right)^r & L_i \le \hat{y}_i \le U_i \\ 1 & \hat{y}_i > U_i \end{cases}$$

$$d_i = \begin{cases} \left(\dfrac{\hat{y}_i - L_i}{T_i - L_i}\right)^r & L_i \le \hat{y}_i \le T_i \\ \left(\dfrac{U_i - \hat{y}_i}{U_i - T_i}\right)^s & T_i < \hat{y}_i \le U_i \\ 0 & \hat{y}_i < L_i \text{ or } \hat{y}_i > U_i \end{cases}$$

The minimum and maximum allowable values for $\hat{y}_i$ are denoted L_i and U_i, respectively. These points can also represent levels at which permitting either $\hat{y}_i < L_i$ or $\hat{y}_i > U_i$ adds very little value to the overall process. Also, T_i is the desired target value of $\hat{y}_i$ between L_i and U_i. The exponents r and s operate as shape parameters for the desirability function. A large value for r or s puts greater importance on the response values being closer to the respective target. Smaller values imply that the desirability value is large even if the response is far from its target value. The desirability functions for each of Taguchi's three experimental cases are illustrated in Figure 3. Finally, to find the optimal process settings, D is maximized with respect to the controllable factors.
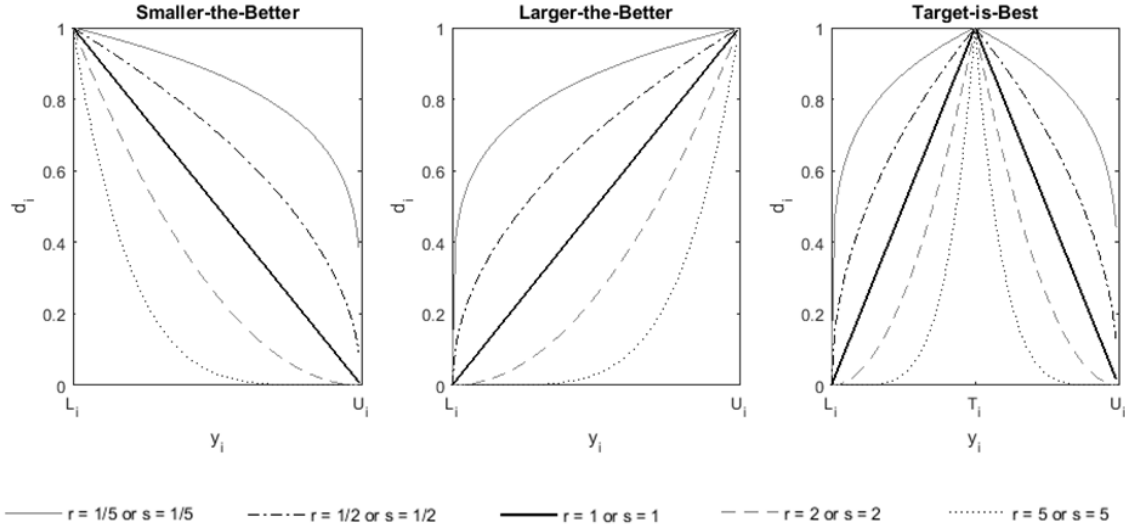
Figure 3. Desirability Functions for Taguchi's Three Experimental Cases

There are several variations to Derringer and Suich's methodology. Instead of the geometric mean, Park [15] advocated the use of the harmonic mean of the *J* desirabilities. Del Castillo et al. [16] presented modified desirability functions that are everywhere differentiable. Park and Park [17] introduced a weighted desirability function approach that allows for varying degrees of importance to be applied to the different responses. The weighted desirability formulation is

$$D_w = \left(\prod_{j=1}^{J} d_j^{w_j} \right)^{1/J} \tag{24}$$

where $w_j > 0$ is the weight of the jth response and $\sum_{j=1}^{J} w_j = J$. Wang et al. [18] proposed a robust desirability function.

27

2.5.2 Loss Functions

Pignatiello [19] focused on Taguchi's *target-is-best* experimental scenario and based his method on the multivariate quadratic loss function

$$L[\mathbf{y}(\mathbf{x})] = (\mathbf{y}(\mathbf{x}) - \boldsymbol{\tau})' \mathbf{C}(\mathbf{y}(\mathbf{x}) - \boldsymbol{\tau}) \tag{25}$$

where $\mathbf{y}(\mathbf{x})$ is a vector of responses for parameter setting $\mathbf{x}$, $\boldsymbol{\tau}$ is the target response vector, and $\mathbf{C}$ is a cost matrix. The optimal parameter setting is then be found by minimizing the expected loss defined by

$$E\big[L[\mathbf{y}(\mathbf{x})]\big] = \big(E[\mathbf{y}(\mathbf{x})] - \boldsymbol{\tau}\big)' \mathbf{C}\big(E[\mathbf{y}(\mathbf{x})] - \boldsymbol{\tau}\big) + trace\big(\mathbf{C} \cdot Cov[\mathbf{y}(\mathbf{x})]\big) \tag{26}$$

where $E[\mathbf{y}(\mathbf{x})]$ and $Cov[\mathbf{y}(\mathbf{x})]$ are the respective mean vector and covariance matrix of the responses at parameter setting $\mathbf{x}$. Many authors have modified Pignatiello's approach. Ames et al. [20] developed loss functions focusing on the individual responses being on target, but having no consideration for the correlation structure between the responses. Vining [21] modified Equation (26) to use the mean and variance-covariance structure of the predicted responses $\hat{\mathbf{y}}(\mathbf{x})$ instead of the mean and variance-covariance structure of the actual responses $\mathbf{y}(\mathbf{x})$. This approach considered the prediction quality as well as the correlation structure of the responses. Romano et al. [22] adopted Vining's multi-response quality loss function and minimized the expected total loss subject to lower and upper bound constraints placed on the means and variances of the individual responses. Ko et al. [23] integrated the strengths of Pignatiello and Vining's approaches to minimize the expected future loss. One drawback to the loss function approach is that it may be difficult to define the cost matrix $\mathbf{C}$ [70].

2.5.3 Principal Component Analysis

Su and Tong [24] grounded their strategy on transforming the normalized quality losses into a set of uncorrelated components via PCA. Salmasnia et al. [25] first transformed principal component models into a desirability function. They then found an optimal solution by maximizing the overall desirability of the selected principal components within the desired region of the normalized means and standard deviations of the original responses. Paiva et al. [26] combined the PCA and MSE approaches into a Multivariate Mean Square Error (MMSE) measure. Gomes et al. [27] expanded on Paiva et al.'s approach and presented the Weighted Multivariate Mean Square Error (WMMSE) to appropriately weight the individual responses in the MMSE approach.

2.5.4 Distance Metrics

Khuri and Conlon [28] considered the Mahalanobis distance between a vector of each response function and a corresponding vector of their optimum function values. The robust solution across the set of responses is the $\mathbf{x}^*$ in the experimental region that minimizes this distance. Govindaluri and Cho [29] decoupled the J individual response MSE functions from the expected total loss. They then found the setting that minimized the distance between the vector of individual response MSE functions and the vector of ideal MSE values. Chiao and Hamada [71] modeled the mean and covariance structure of the assumed multivariate normal responses and then maximized a "proportion of conformance" measure defined as the probability that the J responses jointly meet their respective specification limits.

2.5.5 Mean Square Error Criterion

Köksoy and Yalcinoz [30], Köksoy [31], and Köksoy [32] extended Lin and Tu's [46] MSE criterion to the multi-response robust design case. Köksoy and Yalcinoz [30] promoted the minimization of the weighted summation of the individual MSE functions

$$\text{Minimize } \sum_{j=1}^{J} W_j MSE_j$$
$$\text{Subject to } \mathbf{x} \in \mathbf{R} \tag{27}$$

where W_j is the weight of the jth MSE function, $\sum_{j=1}^{J} W_j = 1$, and $\mathbf{R}$ is the region of interest.

Köksoy [31] recommended solving the following multi-objective optimization problem:

$$\text{Minimize } \{MSE_1, MSE_2, ..., MSE_J\}$$
$$\text{Subject to } \mathbf{x} \in \mathbf{R} \tag{28}$$

In the non-trivial multi-objective optimization problem, a single solution does not exist that simultaneously optimizes each objective. Therefore, a list of Pareto optimal solutions in which an improvement to one objective causes degradation to at least one other objective is generated for the decision maker. All Pareto solutions are considered equally good. Köksoy [32] further proposed solving the optimization problem

$$\text{Minimize } MSE_j \qquad \text{for } 1 \le j \le J$$
$$\text{Subject to } MSE_i = MSE_{i0} \quad \text{for } i = 1, 2, ..., J; \ i \ne j \tag{29}$$
$$\mathbf{x} \in \mathbf{R}$$

where MSE_{i0} are specified values for each MSE function. By successively changing the specified constraint values, a string of solutions is generated rather than a single solution. This allows for an improved understanding of the problem by examining the trade-offs that must be considered to obtain a compromised solution.

2.6 The Delta Method

Situations surface where interest lies in the distribution of some nonlinear function of a random variable and not necessarily the distribution of the random variable itself. With that, the concern then turns to the properties of the function of the random variable. In particular, how can the variance of the function of the random variable be estimated? One such technique, known as the Delta Method, utilizes a Taylor series approximation to obtain reasonable estimates for the mean and variance of the function of a random variable.

Though the original author of the Delta Method is unknown, an article by Robert Dorfman in the 1938 journal *Biometric Bulletin* is credited as the earliest use of the "δ-method." According to Ver Hoef [72], Dorfman proposed the technique to approximate the variance of a nonlinear function of multiple random variables. Ver Hoef also reproduced Dorfman's contribution in which he comments that, if f is a linear function, then the δ-method is exact. He also states that, if f "does not deviate sharply from linear," the δ-method gives a good approximation.

Since its origin, scientists and statisticians across a variety of fields have utilized a version of Dorfman's original δ-method. Chapra and Di Toro [73] extended the Delta Method to modeling water quality and estimating stream reaeration, production, and respiration rates. Durbin et al. [74] utilized the Delta Method to derive a transformation of non-normally distributed DNA microarray data to stabilize the asymptotic variance over the full range of data. Powell [75] focused on providing variance approximations for common parameters used by avian ecologists such as annual bird population growth

and mean annual density of bird species. White [76] has also developed a Windows-based software called MARK which utilizes the Delta Method to aid in the parameter estimation of marked animals when they are re-encountered via dead recoveries, live recaptures, or radio tracking.

2.6.1 Univariate Case

Recall from calculus [77] that if a function $f : \mathbb{R} \to \mathbb{R}$ has derivatives of order n, that is $f^{(n)}(x) = \dfrac{d^n f(x)}{dx^n}$ exists, then the Taylor series expansion of f centered at some constant a is defined as

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n . \tag{30}$$

Consequently, the second-order Taylor series approximation of f centered at a is

$$f(x) \approx f(a) + f'(a)(x-a) + \tfrac{1}{2}f''(a)(x-a)^2 . \tag{31}$$

Now consider $Y = f(X)$ as a function of the normally distributed random variable X. Casella and Berger [78] consider estimating the mean and variance of Y when the mean and variance of X are known parameters. That is, the interest is in finding $E[Y]$ and $Var[Y]$ given that $E[X] = \mu_X$ and $Var(X) = \sigma_X^2$. Following from Equation (31), the second-order Taylor series approximation of Y centered at the point $a = \mu_X$ is defined as

$$Y = f(X) \approx f(\mu_X) + f'(\mu_X)(X - \mu_X) + \tfrac{1}{2}f''(\mu_X)(X - \mu_X)^2 . \tag{32}$$

Therefore, applying the expectation operator to Equation (32) generates an estimate for the mean of Y:

$$E[Y] \approx f(\mu_X) + \tfrac{1}{2} f''(\mu_X) \sigma_X^2 \tag{33}$$

Similarly, applying the variance operator to Equation (32) generates an estimate for the variance of Y:

$$Var[Y] \approx f'(\mu_X)^2 \sigma_X^2 + \tfrac{1}{2} f''(\mu_X)^2 \sigma_X^4 \tag{34}$$

Derivations of Equations (33) and (34) are shown in Appendix A.

2.6.2 Multivariate Case

In the multivariate case [79], let S be a nonempty set in $\mathbb{R}^n$ and let $f: S \to \mathbb{R}$. For $\mathbf{x} \in S$, if the gradient vector $\nabla f(\mathbf{x})$ and hessian matrix $\mathbf{H}(\mathbf{x})$ exist, then the multivariate second-order Taylor series approximation of f centered at some constant vector $\mathbf{a}$ is

$$f(\mathbf{x}) \approx f(\mathbf{a}) + \nabla f(\mathbf{a})'(\mathbf{x} - \mathbf{a}) + \tfrac{1}{2}(\mathbf{x} - \mathbf{a})'\mathbf{H}(\mathbf{a})(\mathbf{x} - \mathbf{a}). \tag{35}$$

Now consider a function $f: \mathbb{R}^n \to \mathbb{R}$ and a n-dimensional normally distributed random vector $\mathbf{X}$ with mean vector $\boldsymbol{\mu}_\mathbf{X}$ and covariance matrix $\Sigma_\mathbf{X}$. The second-order Taylor series approximation of the function $Y = f(\mathbf{X})$ centered at the vector $\mathbf{a} = \boldsymbol{\mu}_\mathbf{X}$ is

$$f(\mathbf{X}) \approx f(\boldsymbol{\mu}_\mathbf{X}) + \nabla f(\boldsymbol{\mu}_\mathbf{X})'(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) + \tfrac{1}{2}(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})'\mathbf{H}(\boldsymbol{\mu}_\mathbf{X})(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}). \tag{36}$$

The estimated mean and variance of Y are then defined respectively as

$$E[Y] \approx f(\boldsymbol{\mu}_\mathbf{X}) + \tfrac{1}{2} tr\left(\mathbf{H}(\boldsymbol{\mu}_\mathbf{X})\Sigma_\mathbf{X}\right) \tag{37}$$

and

$$Var[Y] \approx \nabla f(\boldsymbol{\mu}_\mathbf{X})' \Sigma_\mathbf{X} \nabla f(\boldsymbol{\mu}_\mathbf{X}) + \tfrac{1}{2} tr\left(\mathbf{H}(\boldsymbol{\mu}_\mathbf{X})\Sigma_\mathbf{X}\mathbf{H}(\boldsymbol{\mu}_\mathbf{X})\Sigma_\mathbf{X}\right). \tag{38}$$

Derivations of Equations (37) and (38) are shown in the Appendix A.

III. Extending the Combined-Array Response Surface Model Approach

3.1 Introduction

The first objective of this dissertation is to extend the combined-array RSM approach that relies exclusively on the low-order polynomial models discussed in Section 2.3.1. Since more accurate predictive response surface models result in better RPD solutions [33], a methodology will be developed that utilizes the non-linear Kriging and RBFNN models in place of the polynomial models. From there, the mean and variance of a second-order Taylor series approximation of the Kriging and RBFNN models will be calculated via the Multivariate Delta Method. Finally, an existing optimization problem that employs these approximations will be solved to identify the robust control parameter setting. Henceforth, this procedure is referred to as the combined-array Multivariate Delta Method approach, or simply MDM.

The rest of Chapter III is organized as follows. Section 3.2 outlines the proposed MDM methodology. Section 3.3 uses two case studies to compare the combined-array MDM approach to the combined-array RSM approach and the combined-array stochastic emulator approach. Finally, Section 3.4 summarizes this chapter.

3.2 The Combined-Array Multivariate Delta Method (MDM) Approach

This section outlines the methodology behind the proposed MDM approach. In Section 3.2.1, the mean and variance models for the second-order Taylor series approximations to the Kriging and RBFNN models are developed. In Section 3.2.2, the combined-array MDM approach is outlined against the combined-array RSM approach

and the combined-array emulator approach. A cross-validation procedure for determining an appropriate network structure for the RBFNN is highlighted in Section 3.2.3.

3.2.1 Mean and Variance Models via Taylor Series Approximation

The MDM approach uses the same distributional assumptions regarding the noise variables and the model's error as the RSM approach does. To recap, the model's error ε is normally distributed with a zero mean and a constant variance σ_ε^2. Also, it is assumed that the noise factors are uncorrelated and that $\mathbf{z} \sim N(\boldsymbol{\mu}_\mathbf{z}, \Sigma_\mathbf{z})$.

First, let $\hat{y}(\mathbf{v})$, where $\mathbf{v} = \begin{bmatrix} \mathbf{x} \\ \mathbf{z} \end{bmatrix}$, represent the Kriging or RBFNN model of the system. The second-order Taylor series approximation of $\hat{y}(\mathbf{v})$ about the vector $\mathbf{a} = \boldsymbol{\mu}_\mathbf{v} = \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\mu}_\mathbf{z} \end{bmatrix}$ is defined as

$$T_2\left(\hat{y}(\mathbf{v})\right) = \hat{y}(\boldsymbol{\mu}_\mathbf{v}) + \nabla\hat{y}(\boldsymbol{\mu}_\mathbf{v})'\left(\mathbf{v} - \boldsymbol{\mu}_\mathbf{v}\right) + \tfrac{1}{2}\left(\mathbf{v} - \boldsymbol{\mu}_\mathbf{v}\right)'\hat{\mathbf{H}}(\boldsymbol{\mu}_\mathbf{v})\left(\mathbf{v} - \boldsymbol{\mu}_\mathbf{v}\right) + \varepsilon \tag{39}$$

where $\nabla\hat{y}(\boldsymbol{\mu}_\mathbf{v})$ and $\hat{\mathbf{H}}(\boldsymbol{\mu}_\mathbf{v})$ are the gradient vector and Hessian matrix of $\hat{y}(\mathbf{v})$ evaluated at $\boldsymbol{\mu}_\mathbf{v}$. By way of the Multivariate Delta Method [78], the estimated mean and variance of $T_2\left(\hat{y}(\mathbf{v})\right)$ are then calculated as

$$E\left[T_2\left(\hat{y}(\mathbf{v})\right)\right] = \hat{y}(\boldsymbol{\mu}_\mathbf{v}) + \tfrac{1}{2}tr\left(\hat{\mathbf{H}}(\boldsymbol{\mu}_\mathbf{v})\Sigma_\mathbf{v}\right) \tag{40}$$

and

$$Var\left[T_2\left(\hat{y}(\mathbf{v})\right)\right] = \nabla\hat{y}(\boldsymbol{\mu}_\mathbf{v})'\Sigma_\mathbf{v}\nabla\hat{y}(\boldsymbol{\mu}_\mathbf{v}) + \tfrac{1}{2}tr\left(\hat{\mathbf{H}}(\boldsymbol{\mu}_\mathbf{v})\Sigma_\mathbf{v}\hat{\mathbf{H}}(\boldsymbol{\mu}_\mathbf{v})\Sigma_\mathbf{v}\right) + \sigma_\varepsilon^2, \tag{41}$$

respectively. Since $\Sigma_{\mathbf{v}} = \begin{bmatrix} \mathbf{0}_{r_{\mathbf{x}} \times r_{\mathbf{x}}} & \mathbf{0}_{r_{\mathbf{x}} \times r_{\mathbf{z}}} \\ \hline \mathbf{0}_{r_{\mathbf{z}} \times r_{\mathbf{x}}} & \Sigma_{\mathbf{z}} \end{bmatrix}$, the mean and variance models in Equations (40)

and (41) can be further reduced to

$$E\left[T_2\left(\hat{y}(\mathbf{v})\right)\right] = \hat{y}(\mathbf{\mu}_{\mathbf{v}}) + \tfrac{1}{2} tr\left(\hat{y}_{\mathbf{zz}}(\mathbf{\mu}_{\mathbf{v}})\Sigma_{\mathbf{z}}\right) \tag{42}$$

and

$$Var\left[T_2\left(\hat{y}(\mathbf{v})\right)\right] = \hat{y}_{\mathbf{z}}(\mathbf{\mu}_{\mathbf{v}})'\Sigma_{\mathbf{z}}\hat{y}_{\mathbf{z}}(\mathbf{\mu}_{\mathbf{v}}) + \tfrac{1}{2} tr\left(\hat{y}_{\mathbf{zz}}(\mathbf{\mu}_{\mathbf{v}})\Sigma_{\mathbf{z}}\hat{y}_{\mathbf{zz}}(\mathbf{\mu}_{\mathbf{v}})\Sigma_{\mathbf{z}}\right) + \sigma_{\varepsilon}^2, \tag{43}$$

where $\hat{y}_{\mathbf{z}}(\mathbf{\mu}_{\mathbf{v}})$ is the vector of first-order partial derivatives of $\hat{y}(\mathbf{v})$ with respect to $\mathbf{z}$

evaluated at $\mathbf{\mu}_{\mathbf{v}}$ and $\hat{y}_{\mathbf{zz}}(\mathbf{\mu}_{\mathbf{v}})$ is the matrix of second-order partial derivatives of $\hat{y}(\mathbf{v})$ with

respect to $\mathbf{z}$ evaluated at $\mathbf{\mu}_{\mathbf{v}}$. The MSE of the fitted response model is used to estimate σ_{ε}^2.

Since the Kriging model predicts the exact response for observed training vectors, its

MSE is equal to zero. The MSE of the RBFNN model is discussed in Section 3.2.3.

Finally, in a manner similar to the expressions for the mean and variance of the

standard and extended quadratic models, Equations (42) and (43) are in terms of only the

control factors $\mathbf{x}$ and, as such, they can be used to approximate the mean and variance of

the system anywhere in the control design space. Given these mean and variance

estimates, a dual response optimization approach is then solved to locate a robust control

parameter setting. A subset of the optimization formulations is discussed in Section 2.4.

Below, Sections 3.2.1.1 and 3.2.1.2 derive the gradient vector and Hessian matrix

of the Kriging and RBFNN metamodels.

3.2.1.1 Gradient Vector and Hessian Matrix of the Kriging Model

Let $K = r_{\mathbf{x}} + r_{\mathbf{z}}$. The kth element of the gradient vector for the Kriging model output in Equation (15) for $p = 2$ is defined as

$$\frac{\partial \hat{y}}{\partial v_k} = \frac{\partial \mathbf{r}'(\mathbf{v})}{\partial v_k} \mathbf{R}^{-1} (\mathbf{y} - \mathbf{1}\beta) \text{ for } k = 1,...,K \tag{44}$$

where

$$\frac{\partial \mathbf{r}(\mathbf{v})}{\partial v_k} = -2\theta_k \begin{bmatrix} \left(v_k - v_k^{(1)}\right)\exp\left[-\sum_{m=1}^{K}\theta_m\left(v_m - v_m^{(1)}\right)^2\right] \\ \vdots \\ \left(v_k - v_k^{(N)}\right)\exp\left[-\sum_{m=1}^{K}\theta_m\left(v_m - v_m^{(N)}\right)^2\right] \end{bmatrix} \tag{45}$$

is the first-order partial derivative of $\mathbf{r}(\mathbf{v})$ with respect to input feature k.

The i,j^{th} element of the Hessian matrix for the Kriging output is defined as

$$\frac{\partial^2 \hat{y}}{\partial v_i \partial v_j} = \frac{\partial^2 \mathbf{r}'(\mathbf{v})}{\partial v_i \partial v_j} \mathbf{R}^{-1}(\mathbf{y} - \mathbf{1}\beta) \text{ for } i, j = 1,...,K \tag{46}$$

where

$$\frac{\partial^2 \mathbf{r}(\mathbf{v})}{\partial v_i \partial v_j} = \begin{cases} 2\theta_i \begin{bmatrix} \left(2\theta_i\left(v_i - v_i^{(1)}\right)^2 - 1\right)\exp\left[-\sum_{m=1}^{K}\theta_m\left(v_m - v_m^{(1)}\right)^2\right] \\ \vdots \\ \left(2\theta_i\left(v_i - v_i^{(N)}\right)^2 - 1\right)\exp\left[-\sum_{m=1}^{K}\theta_m\left(v_m - v_m^{(N)}\right)^2\right] \end{bmatrix} & \text{for } i = j \\[4em] 4\theta_i\theta_j \begin{bmatrix} \left(v_i - v_i^{(1)}\right)\left(v_j - v_j^{(1)}\right)\exp\left[-\sum_{m=1}^{K}\theta_m\left(v_m - v_m^{(1)}\right)^2\right] \\ \vdots \\ \left(v_i - v_i^{(N)}\right)\left(v_j - v_j^{(N)}\right)\exp\left[-\sum_{m=1}^{K}\theta_m\left(v_m - v_m^{(N)}\right)^2\right] \end{bmatrix} & \text{for } i \neq j \end{cases} \tag{47}$$

is the second-order partial derivative of $\mathbf{r}(\mathbf{v})$ with respect to input features i and j.

3.2.1.2 Gradient Vector and Hessian Matrix of the RBFNN Model

The kth element of the gradient vector for the RBFNN output in Equation (21) is

defined as

$$\frac{\partial \hat{y}}{\partial v_k} = -\sum_{n=1}^{N} \frac{w_n \left(v_k - \mu_k^{(n)} \right)}{\sigma_n^2} \exp\left[-\frac{1}{2\sigma_n^2} \sum_{m=1}^{K} \left(v_m - \mu_m^{(n)} \right)^2 \right] \text{ for } k = 1,...,K. \tag{48}$$

The i,j^{th} element (for $i, j = 1,..., K$) of the Hessian matrix for the RBFNN output is

defined as

$$\frac{\partial^2 \hat{y}}{\partial v_i \partial v_j} = \begin{cases} \displaystyle\sum_{n=1}^{N} \frac{w_n \left[\left(v_i - \mu_i^{(n)} \right)^2 - \sigma_n^2 \right]}{\sigma_n^4} \exp\left[-\frac{1}{2\sigma_n^2} \sum_{m=1}^{K} \left(v_m - \mu_m^{(n)} \right)^2 \right] & \text{for } i = j \\[2em] \displaystyle\sum_{n=1}^{N} \frac{w_n \left[\left(v_i - \mu_i^{(n)} \right)\left(v_j - \mu_j^{(n)} \right) \right]}{\sigma_n^4} \exp\left[-\frac{1}{2\sigma_n^2} \sum_{m=1}^{K} \left(v_m - \mu_m^{(n)} \right)^2 \right] & \text{for } i \neq j \end{cases} \tag{49}$$

3.2.2 RSM Approach vs. Emulator Approach vs. MDM Approach

The main steps for the RSM approach, the emulator approach, and the MDM

approach are outlined in Figure 4. This figure also highlights the experimental designs

(DOE) and the modeling efforts associated with each approach. The RSM and MDM

approaches each require the development of one DOE and one modeling effort. The

emulator approach, on the other hand, necessitates the generation of two DOEs and three

individual modeling efforts.

In order to compare the results across the three different approaches, each $\hat{y}(\mathbf{x}, \mathbf{z})$

was built from a combined-array DOE in which the coded ±1 levels of each noise factor

z_i were set to $\mu_i \pm 3\sigma_i$. This ensures that the emulator remains valid during the sampling

phase [13]. Consequently, setting the levels of the noise factors in this manner implies that $\Sigma_{\mathbf{z}} = \left(\frac{1}{3}\right)^2 \mathbf{I}_{r_{\mathbf{z}}}$.

The last step of each approach is to identify the robust control setting by utilizing the appropriate mean and variance models within an optimization problem. Due to its simplicity and the manner in which it balances being *on target* with minimal variance, Lin and Tu's MSE approach in Table 3 was used in this research. The optimal settings $\mathbf{x}^{*}$ were found using MATLAB's *fmincon* function was used within a greedy randomized adaptive search procedure (GRASP) heuristic that solves the problem from a number of different starting points and chooses the best overall solution [80]. This procedure can be slow because it seeks to avoid getting trapped at a local minimum. This is a concern since *fmincon* does not guarantee convergence to a global minimum for a potentially nonconvex problem [81].
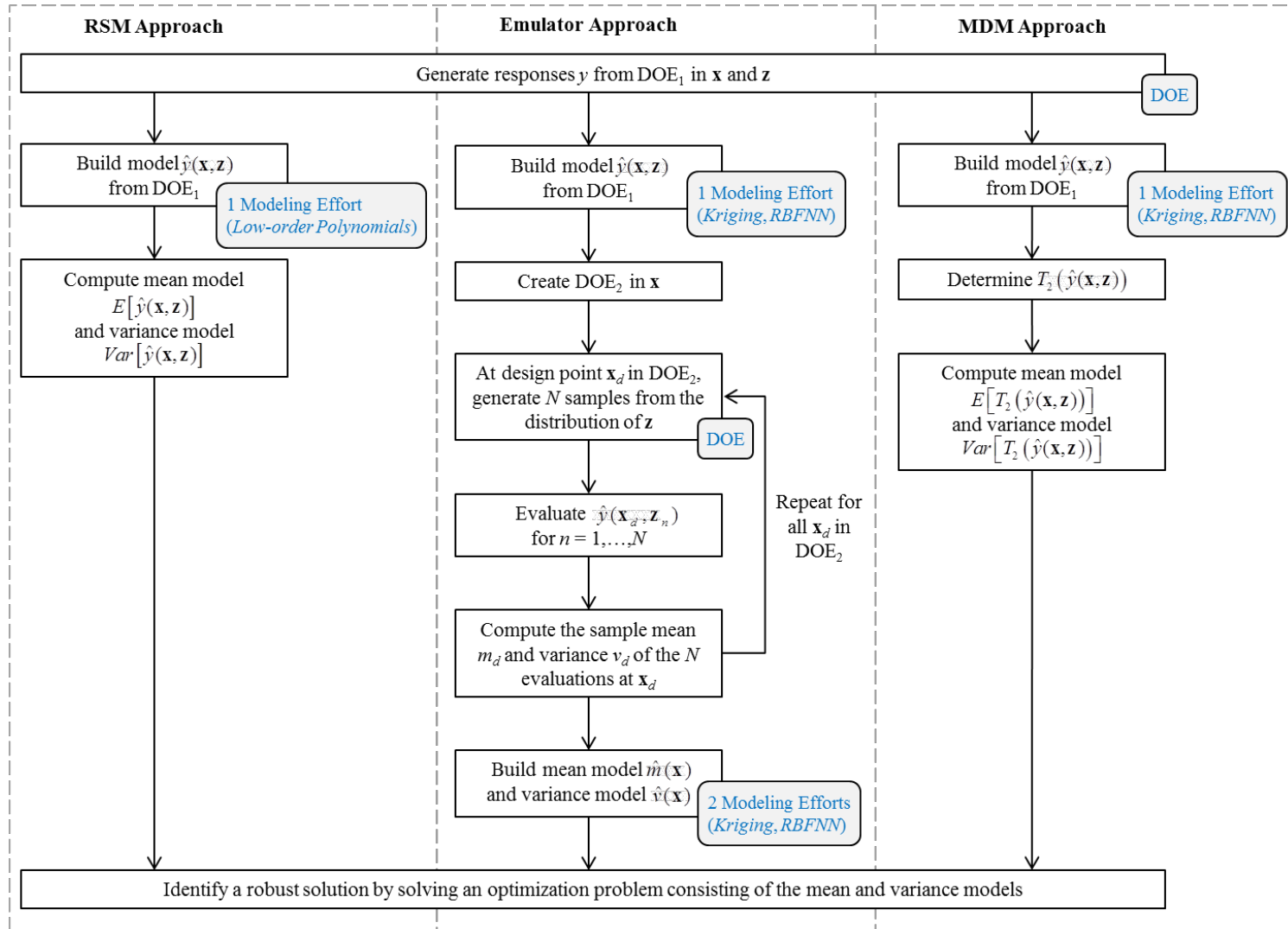
Figure 4. Combined-Array RPD Approaches: RSM vs. Emulator vs. MDM

3.2.3 Choosing the RBFNN Structure

Although RBFNNs have many advantages, one disadvantage they have is that, as the size of the training set increases, the number of hidden neurons that are needed also increases. This subsequently increases the time necessary to train the network. Another drawback to their use involves the selection of the network's architecture [61]. If too many neurons are used, then the overall generalization of the network will be deficient. On the other hand, the network will not be able to sufficiently learn the training data if too few neurons are selected. Therefore, a cross-validation (CV) procedure was used to determine the structure of the RBFNN so that the resulting function was well-generalized with minimal risk of over-fitting the experimental data. The structural parameters of concern here were the number of hidden layer neurons used in the network and the spread parameter σ of the hidden layer neurons.

Given an experimental design with N runs, a set of m_n hidden layer sizes and a set of m_σ spread values were first defined. Then, $M = m_n \times m_\sigma$ combinations of RBFNN structure pairs were generated such that each pairing was composed of a hidden layer size (n) and a spread value (σ). Next, a CV procedure was performed across the set of structures and the MSE was recorded for each (n, σ) pair. The size of the experimental design—this research used designs of 25, 81, and 256 runs—determined which CV method was used. For $N \leq 25$, leave-one-out CV was used. For computational efficiency, ten rounds of 10-fold CV were used for $N > 25$. The final neural network was trained on the complete design and structured via the (n, σ) pairing that produced the minimum MSE (or minimum average MSE for larger datasets). For the examples used in

this research, this procedure led to well-generalized networks with a reduced risk of over-fitting the data since only 20–50% of the total number of training vectors were routinely chosen as neuron centers. The chosen spread parameter ranged from 2 to 10. Since the design space was limited between -1 and 1, these values for the spread parameter led to smooth functions. The RBFNNs were trained using MATLAB's *newrb* function.

The MSE of the RBFNN that is used to estimate σ_ε^2 in Equation (43) is

$$\hat{\sigma}_\varepsilon^2 = \sum_{i=1}^{N} e_i^2 \bigg/ (N-n) \tag{50}$$

where e_i is the prediction error of the ith design point. The value n, which is the number of neurons in the trained network's hidden layer, corresponds to the number of estimated weight parameters in the RBFNN.

3.3 Application and Results

In this section, the MDM approach is applied to two case studies. Section 3.3.1 provides a proof-of-concept demonstration of the MDM approach using a synthetic case study. Section 3.3.2 applies the MDM approach to a computer simulation.

A popular method for comparing different analysis techniques for simulation studies is through the use of the Root Mean Squared Error (*RMSE*) [43, 51, 53, 54, 60]. Since the analysis is of known models, the model predictions $\hat{y}_m$ can be compared to the known true values y_m for a test set of $M = 200$ random validation points via

$$RMSE = \sqrt{\frac{1}{M} \sum_{m=1}^{M} (y_m - \hat{y}_m)^2} \ . \tag{51}$$

42

A model's *RMSE* can then be compared to the range of the true responses to gain insight into its accuracy [11].

3.3.1 A Synthetic Case Study

In this case study, data was generated from the truth model in Equation (51). By knowing the truth model, the system's true mean and variance—shown in Equations (52) and (53)—can be used as reference points for our modeling efforts. The truth model is employed with the control factor $x_1 \in [-1,1]$ and the noise factor $z_1 \sim N(\mu = 0, \sigma = 1)$. The goal of this RPD study was to locate the appropriate operating point that resulted in a mean response of 8 with a minimal variance. A 5^2 full factorial design was used to generate experimental data.

$$y = 9/(1+\exp(1-5x_1)) + 3z_1^2 - 4x_1^2 z_1 \tag{52}$$

$$E[y] = 9/(1+\exp(1-5x_1)) + 3 \tag{53}$$

$$Var[y] = 16x_1^4 + 18 \tag{54}$$

The MDM approach using Kriging (*KR*) and RBFNN (*RBF*) models was compared to the RSM approach that uses the four polynomial models in Section 2.3.1: $y_{Std}(\mathbf{x}, \mathbf{z})$, $y_{NN}(\mathbf{x}, \mathbf{z})$, $y_{CNN}(\mathbf{x}, \mathbf{z})$, and $y_{CCN}(\mathbf{x}, \mathbf{z})$. Henceforth, these models will simply be referred to as *Std*, *NN*, *CNN*, and *CCN*.

Response surface graphs for the six modeling efforts are shown alongside the truth model in Figure 5. Visual inspection shows that the *Std*, *NN*, and *CNN* models only provide the general trend of the data. This was expected due to the highly non-linear nature of Equation (52). The *CCN* model offers a markedly improved representation of

the non-linearity of the system over its predecessors. Finally, the *KR* and *RBF* models represent the true input-output relationship of the system very well. This information is summarized in Table 8.
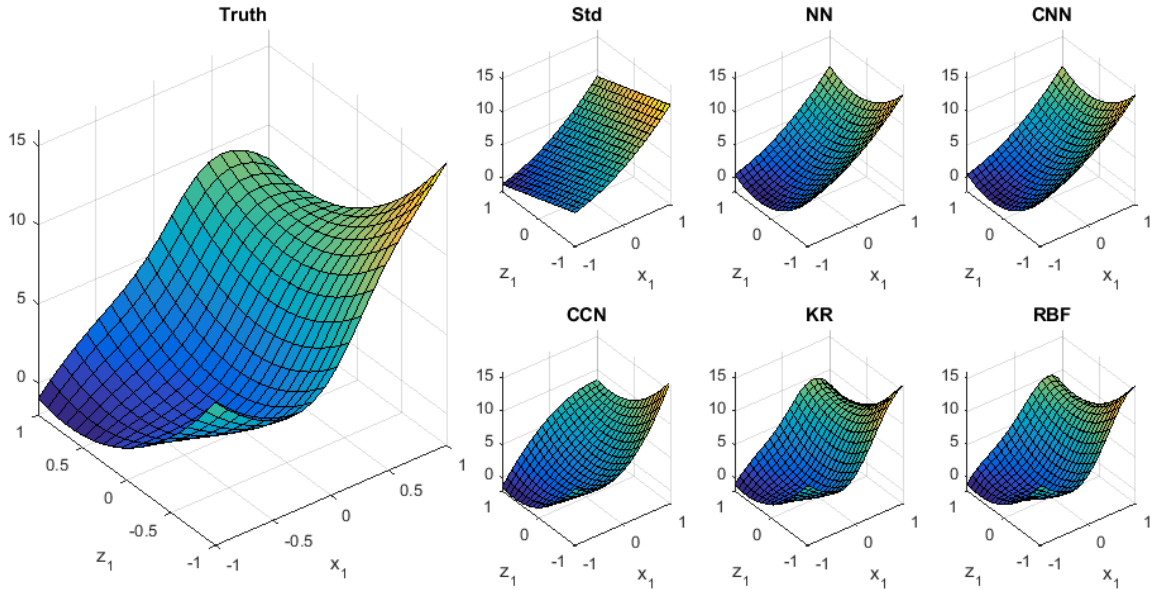
Figure 5. Response Surface Comparison for the Synthetic Case Study

Table 8. Quality of Models for the Synthetic Case Study

Model	RMSE	% of Range
Std	1.62	11.39%
NN	1.13	7.97%
CNN	1.13	7.97%
CCN	0.79	5.58%
RBF	0.31	2.18%
KR	0.19	1.37%

The mean, variance, and MSE models that correspond to the *Truth*, *Std*, *NN*, *CNN*, *CCN*, *KR*, and *RBF* response models are depicted in Figure 6. The locations of the estimated robust points are also plotted. In this case, the *NN*, *CNN*, and *CCN* mean

models are nearly identical and are plotted on top of one another. Similarly, the *NN* and

CNN variance models are graphically the same. Of the six modeling efforts, the *KR* and

RBF mean models are very good representations of the true mean response. Also, the

CCN, *KR*, and *RBF* variance models provide the best depictions of the system's true

variance. In fact, the *RBF* variance model is nearly identical to the system's true variance

model. The *Std*, *NN*, and *CNN* mean and variance models are poor relative to their

competitors.
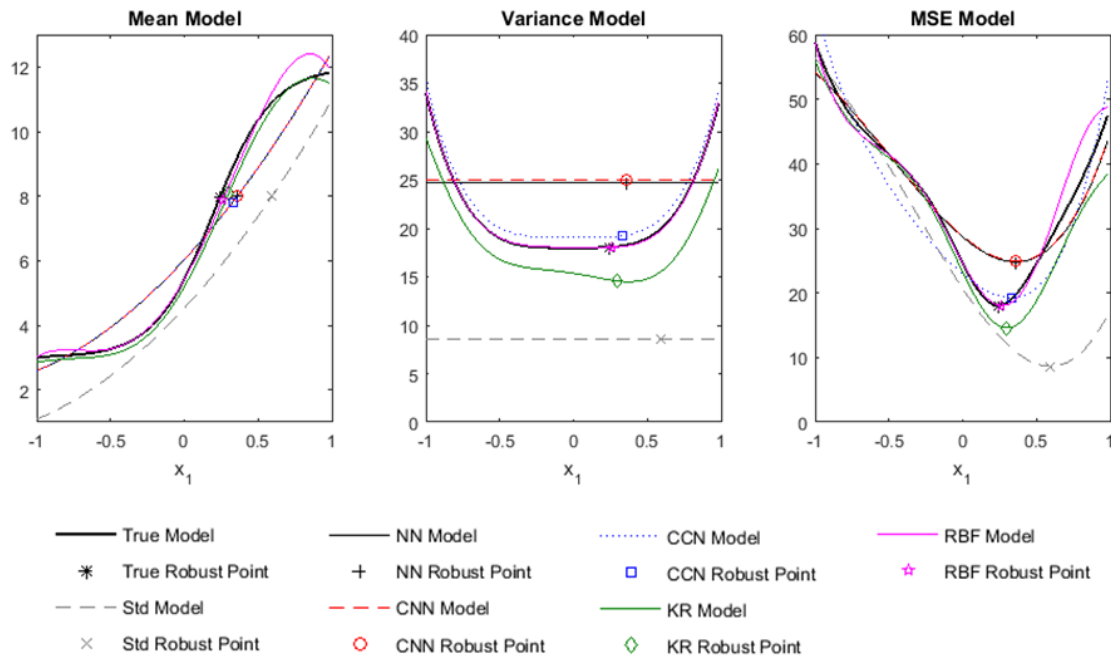
Figure 6. Mean, Variance, and MSE Models for the Synthetic Case Study

Table 9 summarizes the predicted and realized means, variances, and MSEs at

each robust point. Confirmation experiments were performed in which 2,000 Monte

Carlo simulations of the truth model were run at the estimated robust points. Common

random numbers were used to allow an *apples-to-apples* comparison. It can be observed

that, not only is the *RBF* robust point the closest to the true robust point, but its predictions are indicative of the system's actual performance. An interesting note is that, without knowledge of the system's truth model—which is typically never known for a simulation—and based solely on the predictions provided in Table 9, initial conclusions would have chosen the *Std* robust point as the best operating point since the prediction of the system's mean response is exactly on target and the prediction of the system's variance is the smallest. This highlights the importance of performing confirmation experiments.

Table 9. RPD Results for the Synthetic Case Study

Model	Robust Point x_1	Goal →	Mean 8	Variance Min	MSE Min
Truth	**0.24**	**Actual**	**7.96**	**18.05**	**18.05**
		Realized	**7.90**	**18.63**	**18.64**
Std	0.59	Predicted	8.00	8.62	8.62
		Realized	10.82	19.75	27.70
NN	0.36	Predicted	8.00	24.79	24.79
		Realized	9.14	18.65	19.95
CNN	0.36	Predicted	8.00	24.95	24.95
		Realized	9.14	18.65	19.95
CCN	0.33	Predicted	7.81	19.25	19.29
		Realized	8.82	18.63	19.30
KR	0.30	Predicted	8.08	14.57	14.58
		Realized	8.54	18.62	18.91
RBF	0.27	Predicted	7.86	18.03	18.05
		Realized	8.10	18.62	18.63

3.3.2 A Resistor-Inductor (RL) Circuit Simulation

An RPD study was performed for a simulation of an RL electrical circuit described by Kenett and Zacks [82]. The response of interest is the output current (in amperes) of the circuit defined by

$$Y = V \Big/ \sqrt{R^2 + (2\pi fL)^2} \,. \tag{55}$$

The four factors that influence the output current are listed in Table 10. The factors R

and L are controllable whereas the factors V and f are assumed to be normally distributed

random variables with known means and standard deviations. The goal of this RPD

study was to identify the robust setting of R and L that yields a mean output current of 10

amperes with minimum variation. A simulation of Equation (55) was created in

MATLAB to approximate the true mean and variance of the system. To generate these

approximations, Monte Carlo simulations were performed 5,000 times at 900 uniformly-

spaced control design points. This provided a reference point to compare each modeling

effort.

Table 10. Factors for the *RL* Electrical Circuit Simulation

Factor	Description (units)	Min	Max	Mean	Standard Deviation
R	Resistance (Ω)	0.05	9.5		
L	Self-inductance (H)	0.01	0.03		
V	Input voltage (V)			100	3
f	Input frequency (Hz)			55	5/3

In this section, the MDM approach is demonstrated against two popular RPD

strategies. Section 3.3.2.1 compares the combined-array MDM approach using *KR* and

RBF models to the combined-array RSM approach that employs the *NN* model in

Equation (4). The *NN* model was chosen as opposed to the *Std* model in Equation (1) due

to the high degree of non-linearity in the simulation. Section 3.3.2.2 compares the

combined-array RSM and MDM approaches to the combined-array stochastic emulator

strategy to show that equivalent results can be achieved via the former approach at a greatly reduced computational cost and without the need for secondary modeling efforts.

3.3.2.1 Combined-Array MDM Approach vs. Combined-Array RSM Approach

The *NN* model was built from a 25-run face-centered cube design which can be used within a cuboidal region to estimate the quadratic effects necessary for the response model in Equation (4) [9]. Space-filling designs, such as LHS designs, are commonly used for developing Kriging and RBFNN models of simulations [83]. Hence, to ensure a fair competition between each model, the *KR* and *RBF* models were each produced from the same 25-run Latin Hypercube Sampling (LHS) design. Inspection of the RMSEs of each model in Table 11 reveals that the *RBF* model provides a slightly better representation of the simulation than the *NN* and *KR* models.

Since there are only two control factors in this RPD problem, the resulting mean, variance, and MSE models can be compared visually. These models are shown in Figure 7. Approximations of the circuit simulation's true mean, variance, and MSE are also shown for comparison. The dots represent the location of the associated model's robust point. Figure 7 shows that the individual mean models provide the general trend of the simulation's true mean response. The variance models, on the other hand, are not indicative of the simulation's true variance. The *NN* variance model over-estimates the true variance throughout the entire design region. The *KR* and *RBF* variance models are better estimates, though they still do not closely model the non-linearity of the variance surface.

In terms of robust points, *NN* and *RBF* provide similar points whereas *KR* most closely identifies the location of the true robust point. The experiment is summarized in Table 11. Each model's robust point was simulated 5,000 times using common random numbers for *V* and *f* in order to assess the quality of the predicted values. All three models under-estimate the simulation's mean at their robust point, though the *RBF* provides the closer prediction. The *KR* model offers the best prediction of the simulation's true variance.
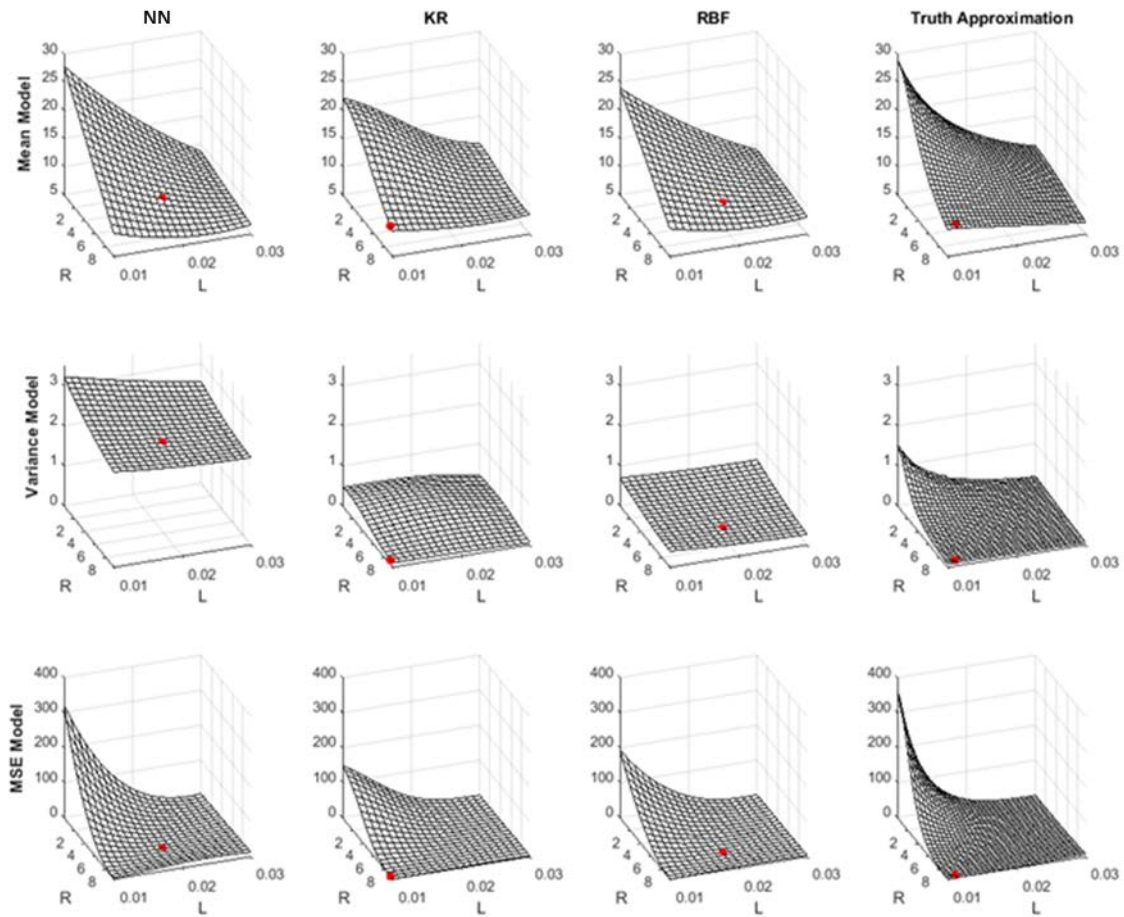
Figure 7. Circuit Simulation Mean, Variance, and MSE Models Using the 25-run Designs

Table 11. Circuit Simulation RPD Results Using the 25-run Designs

Model	RMSE (% Range)	Robust Point R	L		Mean	Variance	MSE
NN	1.13 (5.5%)	6.48	0.019	Predicted	9.98	2.36	2.36
				Realized	10.73	0.13	0.66
KR	0.76 (4.4%)	9.10	0.010	Predicted	9.99	0.12	0.12
				Realized	10.28	0.10	0.18
RBF	0.68 (3.9%)	7.15	0.020	Predicted	9.99	0.35	0.35
				Realized	10.15	0.12	0.14
Truth		**9.17**	**0.011**		**10.02**	**0.09**	**0.09**

At this point in the analysis, it can be concluded that after only 25 experimental runs, the RSM and MDM approaches have provided adequate, but not highly accurate, representations of the simulation's true mean and variance. However, since computer simulations allow experimenters to explore many more factor levels and combinations of factor levels than typically allowed in physical experiments, the number of experimental runs in the DOE was increased from 25 to 81. The *NN*, *KR*, and *RBF* models were then build from the same 81-run LHS design. Their corresponding mean, variance, and MSE models are depicted in Figure 8. This new experiment is summarized in Table 12. By comparing Figure 7 and Figure 8, improvement in the models' mean and variance representations can be observed. The *KR* and *RBF* mean models are capturing the curvature of the simulation while the *NN* mean model still only represents its general trend. There is also significant improvement in the *KR* and *RBF* variance models. These improvements can also be seen in Table 12 as the *KR* and *RBF* models' mean and variance predictions at the determined robust points are improved.

Figure 8. Circuit Simulation Mean, Variance, and MSE Models Using the 81-run LHS Design

Table 12. Circuit Simulation RPD Results Using the 81-run LHS Design

Model	RMSE (% Range)	Robust Point R	L		Mean	Variance	MSE
NN	0.61 (2.8%)	6.60	0.021	Predicted	9.97	0.52	0.52
				Realized	10.29	0.12	0.20
KR	0.37 (1.7%)	2.36	0.028	Predicted	9.98	0.14	0.14
				Realized	10.00	0.17	0.17
RBF	0.34 (1.6%)	8.66	0.013	Predicted	9.98	0.11	0.11
				Realized	10.21	0.10	0.14
Truth		**9.17**	**0.011**		**10.02**	**0.09**	**0.09**

Next, the number of experimental design runs was increased further to 256. The

NN, *KR*, and *RBF* models were now built from the same 256-run LHS design. Figure 9

and Table 13 show additional evidence that the MDM approach using either the *KR* or

RBF models is superior to the RSM approach using the *NN* model. The mean and

variance surfaces provide close approximations of the simulation's true mean and

variance. The predicted means and variances for each model at their robust points are

also nearly identical to their realized values. The *NN* model never truly captures the non-

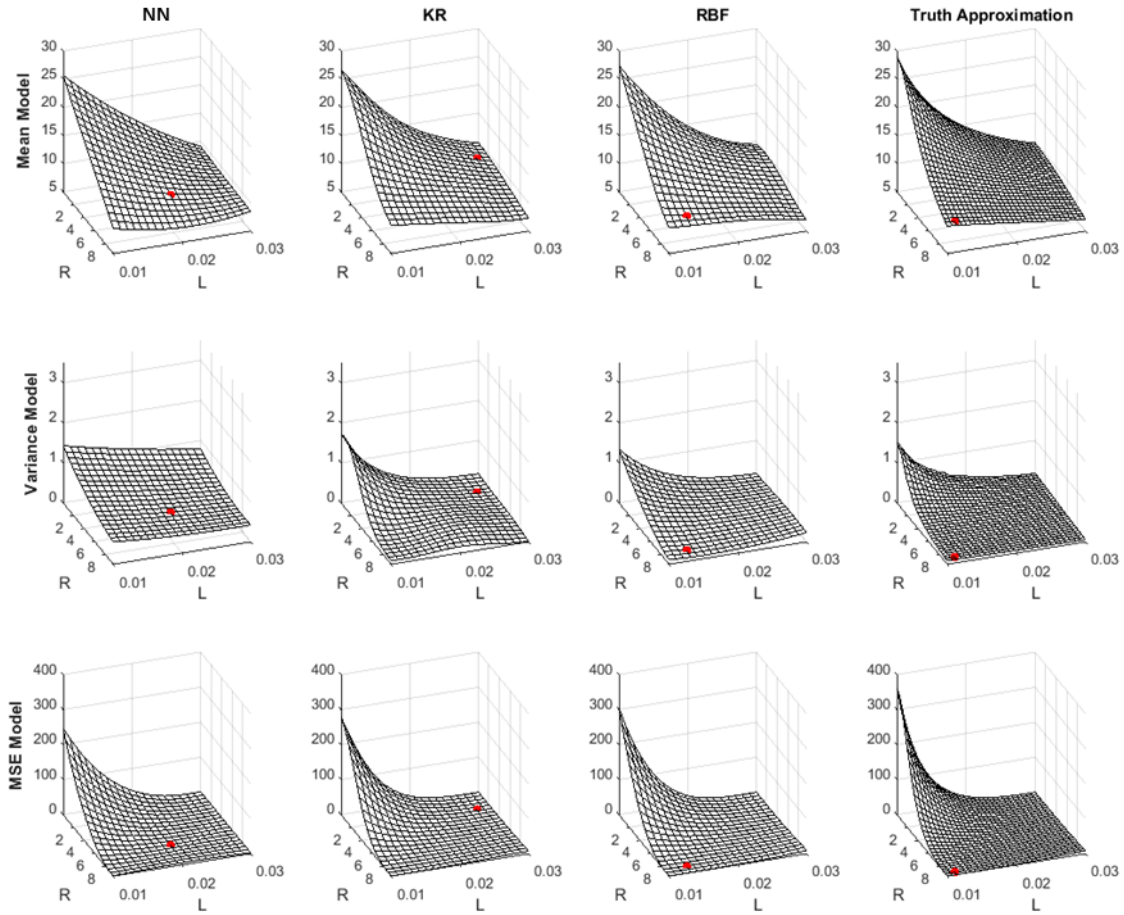linearity of either the mean or variance of the simulation.
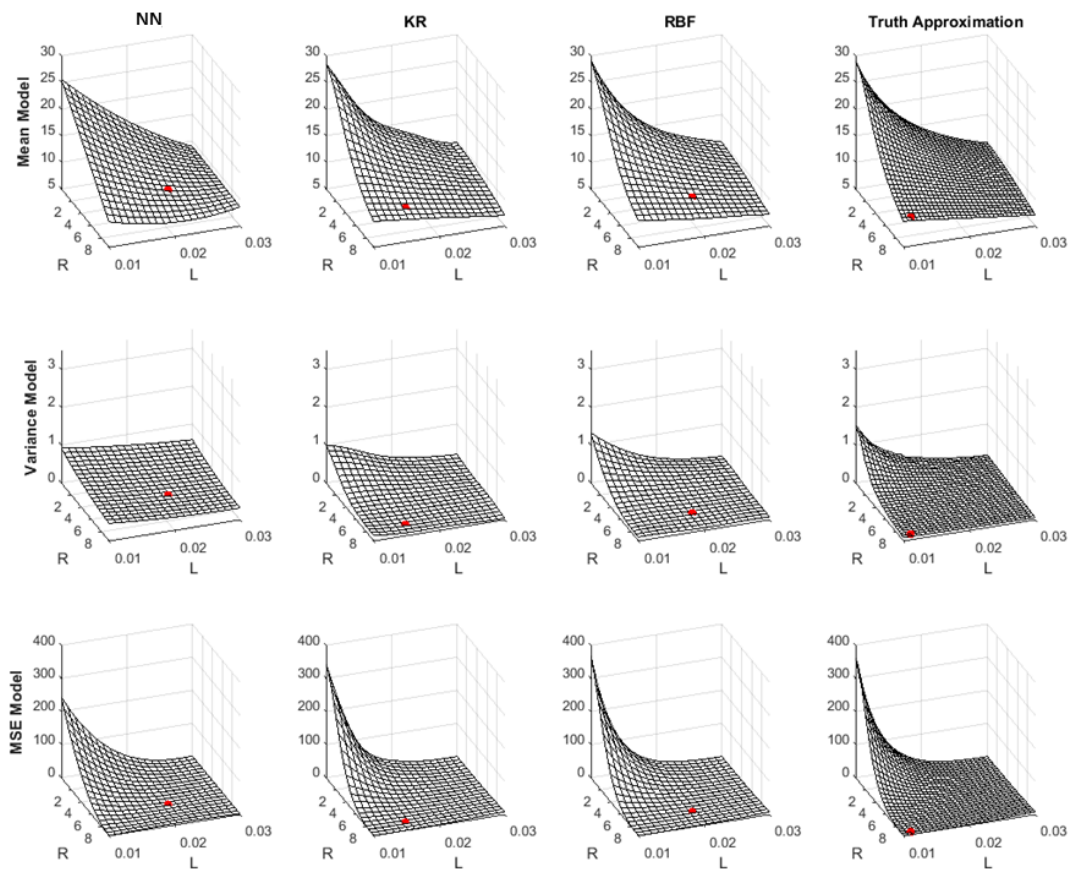
Figure 9. Circuit Simulation Mean, Variance, and MSE Models Using the 256-run LHS Design

Table 13. Circuit Simulation RPD Results Using the 256-run LHS Design

Model	RMSE (% Range)	Robust Point R	L		Mean	Variance	MSE
NN	0.49 (2.8%)	6.52	0.021	Predicted	9.98	0.43	0.43
				Realized	10.24	0.12	0.18
KR	0.05 (0.3%)	8.46	0.016	Predicted	9.99	0.08	0.08
				Realized	9.99	0.10	0.10
RBF	0.08 (0.4%)	7.34	0.020	Predicted	9.99	0.11	0.11
				Realized	9.99	0.11	0.11
Truth		**9.17**	**0.011**		**10.02**	**0.09**	**0.09**

3.3.2.2 Combined-Array RSM/MDM Approaches vs. Combined-Array

Emulator Approach

The combined-array RSM and MDM approaches were compared to the
combined-array stochastic emulator approach using the *NN*, *KR*, and *RBF* models
generated via the 256-run LHS design as the emulators. Within the emulator approach,
the individual mean and variance models were built from a 5^2 factorial design (labeled
DOE$_2$ in Figure 4). These models were constructed using the same modeling approach
that created the emulator itself. That is, if Kriging was used to build the emulator, then
Kriging was also used to build the associated mean and variance models.

The emulator strategy requires a large number of replications to be performed at
each design point. To determine an appropriate number of replications, the long-run
mean and variance of each model was examined at the center point and the four corner
points of the design (Figure 10). Based on these plots, the choice was made to run 500
replications at each of the 25 design points of DOE$_2$ since both the mean and variance of
each model tend to reach a steady-state at that point. Therefore, the emulator approach

required $25 \times 500 = 12,500$ evaluations of the emulator in order to build the mean and variance models. This is in addition to the original 256 experiments necessary to build the emulator in the first place.
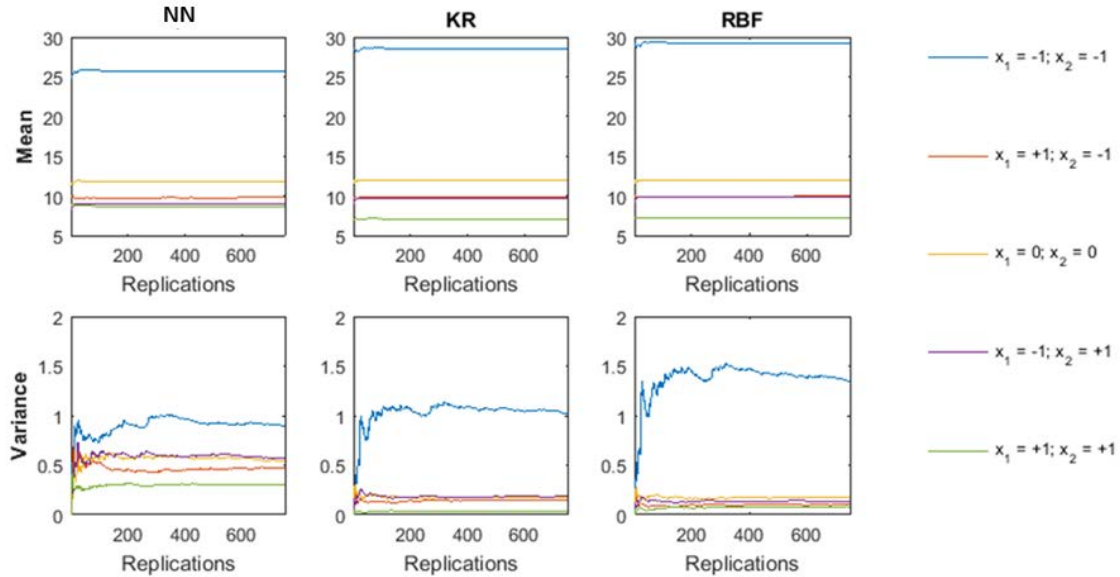
Figure 10. Long-Run Analysis of *NN*, *KR*, and *RBF* Circuit Simulation Emulators

A comparison of the RSM and MDM's mean, variance, and MSE models in Figure 9 to the emulator's corresponding models in Figure 11 reveals very similar response surfaces. A further examination of the robust point summaries for the RSM/MDM (Table 13) and emulator (Table 14) approaches also shows nearly equivalent results. The robust points, as well as their predicted/realized means and variances, for each model type are approximately equal. Consequently, it can be concluded that the combined-array RSM and MDM approaches and the stochastic emulator strategy yielded similar results. However, the emulator strategy required the development of two experimental designs, three response modeling efforts (one each for the emulator, the

mean, and the variance), one optimization procedure, and 12,500 additional function

evaluations. The RSM/MDM approach, on the other hand, only required one

experimental design, one response modeling effort, and one optimization procedure.
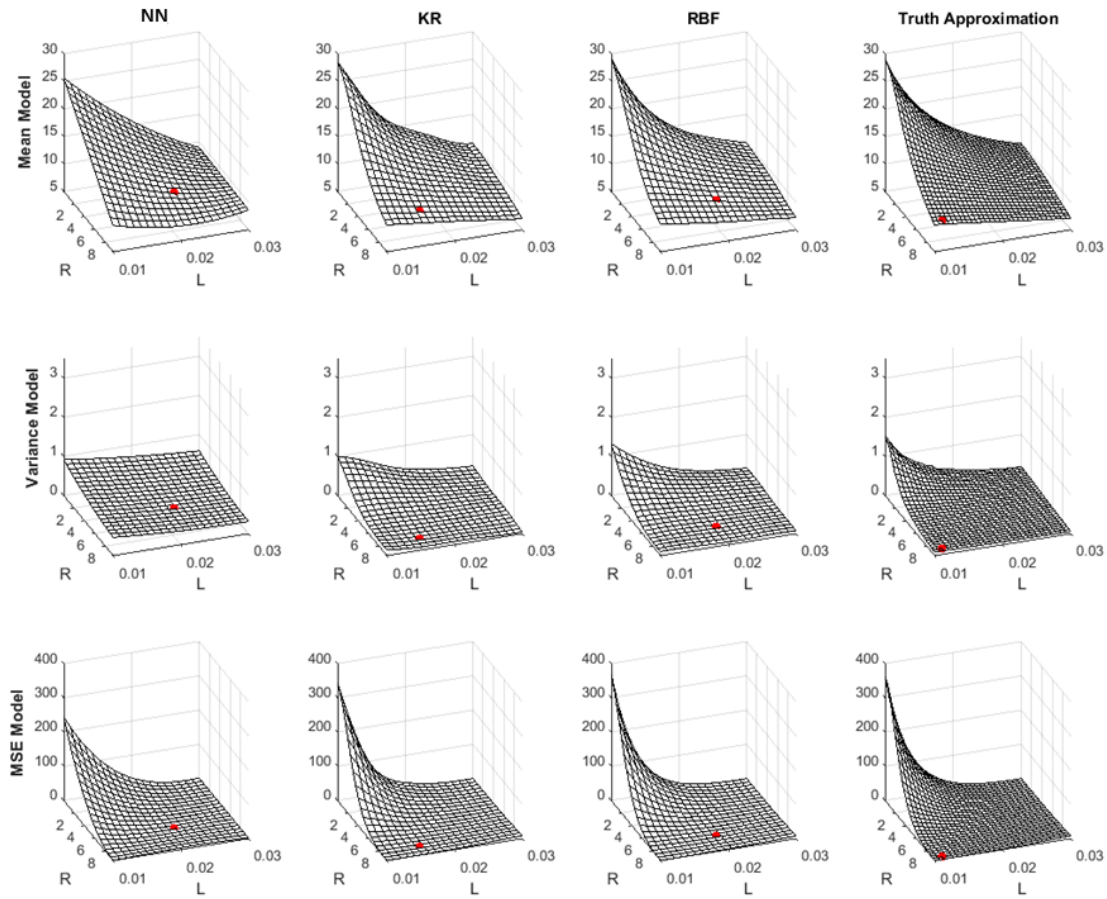
Figure 11. Emulator Approach Mean, Variance, and MSE Models for the Circuit Simulation

Table 14. Emulator Approach Results for the Circuit Simulation

Model	Robust Point R	L		Mean	Variance	MSE
NN	6.45	0.021	Predicted	9.98	0.39	0.39
			Realized	10.22	0.12	0.17
KR	8.32	0.016	Predicted	9.99	0.08	0.08
			Realized	10.06	0.10	0.10
RBF	7.33	0.020	Predicted	9.99	0.10	0.10
			Realized	9.97	0.11	0.11
Truth	**9.17**	**0.011**		**10.02**	**0.09**	**0.09**

3.4 Summary

Chapter III has extended the combined-array RSM approach that relies upon low-order polynomial models. It was demonstrated that improved models of a computer simulation's mean and variance can be attained through the MDM approach that employs non-linear response modeling techniques such as Kriging or RBFNN models. It was further shown that the combined-array MDM approach generates results that are approximately equivalent to the stochastic emulator approach. However, these results can be achieved at a greatly reduced computational cost by utilizing the MDM approach.

Throughout this chapter, the concern was more with examining the individual mean and variance models and less with the actual location of the robust points. In fact, each of the 12 robust points that were identified for the circuit simulation are very good solutions. This is illustrated in Figure 12. The points represented with "•" are design space locations in which the true MSE of the simulation is less than or equal to 0.25. Actually, these points represent a region of points and not just individual points. The minimum MSE, denoted with "×", is 0.09. By comparison, the maximum MSE in the region is 361.18. The 12 individual robust points are denoted with "○". It is shown that

they each fall in or near a "robust region" which implies that they are, in their own right, robust solutions for this RPD problem. The actual interest, in this case, lies with the predicted means and variances at these robust points. It was shown that the MDM approach was superior to the RSM approach in providing improved predictions of the system.
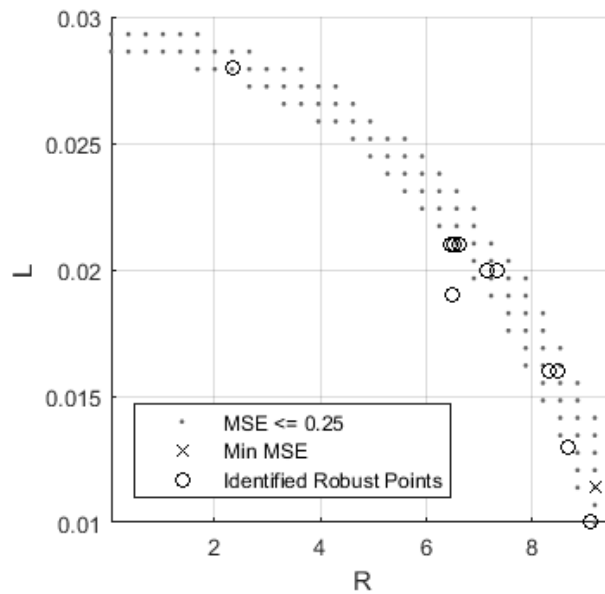
Figure 12. Illustration of the Circuit Simulation's "Robust Region"

IV. A Nested Desirability Function-Based Approach to Multi-Response RPD

4.1 Introduction

This chapter considers the situation in which an experimenter seeks to examine the influence a set of independent variables has on several system responses simultaneously. For example, a situation may require finding the optimal set of conditions that reduce cost while also increasing yield. Unfortunately, the solution that minimizes cost is most likely not the same as the solution that maximizes yield. In fact, the two solutions may diametrically oppose each other. In a case such as this, trade-offs between the responses must be explored to find a collaborative solution.

Chapter IV focuses on the second objective of this dissertation: to develop an approach for multi-response RPD problems that provides a collaborative solution that is balanced across the means and variances of each response. Existing techniques seek to find an optimal balance across the set of responses. However, in some cases the mean or variance of one response influences the solution in such a way that the means and variances of the remaining responses are insignificant to the overall RPD problem. In these situations, it is difficult to attain a solution that is balanced across all responses.

This chapter presents a technique based on well-known desirability functions that places the means and variances of all responses on a level playing field. The proposal also allows decision makers to integrate their personal preferences for the individual response means and variances. For example, they may be willing to accept sacrificing being slightly "off-target" in one response, but be reluctant to allow the variation of another response to get too large.

The rest of this chapter continues in the following manner. Section 4.2 describes the proposed desirability function-based methodology for solving multi-response RPD problems. Section 4.3 utilizes two case studies to compare the proposed approach to a popular MSE-based approach. Finally, Section 4.4 summarizes this chapter.

4.2 The Nested Desirability Function Approach

The proposed approach to solving the multi-response RPD problem employs the desirability functions established by Derringer and Suich [14]. However, instead of using them to transform each response into a corresponding desirability level, they were used to transform the individual response means and variances into desirability levels. This approach allows a decision maker to state their preferences regarding what is, and is not, acceptable for each response's mean and variance. Park and Park's [17] weighted desirability function approach for optimization problems was also implemented to apply distinct degrees of importance to the different responses, response means, and response variances.

Let $\hat{y}_j$ denote an estimated relationship between response variable j and the vector of independent variables **x**. Also, suppose that $\hat{\mu}_j$ and $\hat{\sigma}_j^2$ are the respective mean and variance of $\hat{y}_j$. The desirability of response j is defined as

$$D_j = \left(\left(d_{j,\mu} \right)^{w_{j,\mu}} \times \left(d_{j,\sigma^2} \right)^{w_{j,\sigma^2}} \right)^{\frac{1}{2}} \tag{56}$$

where $0 \leq d_{j,\mu} \leq 1$ and $0 \leq d_{j,\sigma^2} \leq 1$ are the respective desirability transformations for $\hat{\mu}_j$ and $\hat{\sigma}_j^2$. Also, $w_{j,\mu} > 0$ and $w_{j,\sigma^2} > 0$ are the preferred weights for $\hat{\mu}_j$ and $\hat{\sigma}_j^2$, respectively. Since there are two desirabilities—one for the mean and one for the variance—the

59

relationship $w_{j,\mu} + w_{j,\sigma^2} = 2$ must hold. D_j, which is contained in the interval [0,1], will

be 0 if either $d_{j,\mu}$ or d_{j,σ^2} are 0. This implies that response j is unacceptable if either its

mean or variance is unacceptable. Finally, the overall desirability of the combined mean

and variance levels across all J responses is

$$D = \left(\prod_{j=1}^{J} \left(D_j \right)^{W_j} \right)^{1/J} \tag{57}$$

where $W_j > 0$ is the preferred weight for response j and $\sum_{j=1}^{J} W_j = J$. Again, if $D_j = 0$ for

any j, then $D = 0$ and the whole product is unacceptable. The robust point is the $\mathbf{x}^*$ that

maximizes D. Since D is composed of the desirabilities $D_1, D_2, ..., D_J$ and each D_j is itself

composed of the desirability functions $d_{j,\mu}$ and d_{j,σ^2}, the proposed procedure will be

referred to as the Nested Desirability (*ND*) approach.

4.2.1 Desirability Transformations for $\hat{\mu}_j$

Taguchi [4] specified three experimental cases for managing a system's mean

response in a RPD scenario: *smaller-the-better*, *larger-the-better*, and *target-is-best*. Let

$L_{j,\mu}$ and $U_{j,\mu}$ be the respective minimum and maximum values of $\hat{\mu}_j$. If response j is a

smaller-the-better case, then the desirability transformation for $\hat{\mu}_j$ is

$$d_{j,\mu} = \begin{cases} \left(\dfrac{U_{j,\mu} - \hat{\mu}_j}{U_{j,\mu} - L_{j,\mu}} \right)^r & L_{j,\mu} \leq \hat{\mu}_j \leq U_{j,\mu} \\ 0 & \textit{otherwise} \end{cases} \tag{58}$$

If response j is a *larger-the-better* case, then the desirability transformation for $\hat{\mu}_j$ is

$$d_{j,\mu} = \begin{cases} \left(\dfrac{\hat{\mu}_j - L_{j,\mu}}{U_{j,\mu} - L_{j,\mu}} \right)^r & L_{j,\mu} \le \hat{\mu}_j \le U_{j,\mu} \\ 0 & otherwise \end{cases} \tag{59}$$

Finally, if response j is a *target-is-best* case with a desired target of τ_j where

$L_{j,\mu} \le \tau_j \le U_{j,\mu}$, then the desirability transformation for $\hat{\mu}_j$ is

$$d_{j,\mu} = \begin{cases} \left(\dfrac{\hat{\mu}_j - L_{j,\mu}}{\tau_j - L_{j,\mu}} \right)^r & L_{j,\mu} \le \hat{\mu}_j \le \tau_j \\ \left(\dfrac{U_{j,\mu} - \hat{\mu}_j}{U_{j,\mu} - \tau_j} \right)^s & \tau_j < \hat{\mu}_j \le U_{j,\mu} \\ 0 & otherwise \end{cases} \tag{60}$$

The exponents r and s in Equations (58)–(60) are shape parameters for the desirability

function. A value for r or s that is greater than 1 puts greater importance on the mean

response values being closer to the respective minimum, maximum, or target value. On

the other hand, a value for r or s that is between 0 and 1 implies that the desirability value

is large even if the mean response is far from its goal value.

4.2.2 Desirability Transformations for $\hat{\sigma}_j^2$

Whereas there are three distinct cases for managing $\hat{\mu}_j$, there is only one such

case for $\hat{\sigma}_j^2$ since a minimum response variance is always desired. Therefore, let L_{j,σ^2} and

U_{j,σ^2} be the respective minimum and maximum values of $\hat{\sigma}_j^2$. Then, the desirability

transformation for $\hat{\sigma}_j^2$ is

$$d_{j,\sigma^2} = \begin{cases} \left(\dfrac{U_{j,\sigma^2} - \hat{\sigma}_j^2}{U_{j,\sigma^2} - L_{j,\sigma^2}} \right)^r & L_{j,\sigma^2} \leq \hat{\sigma}_j^2 \leq U_{j,\sigma^2} \\ 0 & \textit{otherwise} \end{cases}$$ (61)

The value r is defined as it was in Section 4.2.1.

4.3 Application and Results

This section demonstrates Köksoy and Yalcinoz's (*KY*) MSE procedure [30] and the Nested Desirability (*ND*) procedure on two case studies: a synthetic case consisting of two known functions and a textbook example for a physical experiment.

4.3.1 A Synthetic Case Study

The response functions in Equations (61) and (62) were used to demonstrate the *ND* procedure alongside the *KY* procedure in Equation (27). The responses, which were considered to be equally important, were influenced by the control factor $x \in [-1,1]$ and the noise factor $z \sim N\left(\mu_z = 0,\ \sigma_z^2 = 1 \right)$.

$$y_1 = -4x^2 - 3x + 4z^2 + 3z + 4xz + 15$$ (62)

$$y_2 = -\tfrac{4}{5}x - \tfrac{1}{2}z^2 - \tfrac{1}{5}z + \tfrac{2}{5}xz + 6$$ (63)

Given the known distributional parameters for z, the means and variances of Equations (62) and (63) can be determined analytically. These functions are shown in Equations (63)–(66).

$$\mu_1 = -4x^2 - 3x + 19$$ (64)

$$\sigma_1^2 = (4x+3)^2 + 32$$ (65)

$$\mu_2 = -\tfrac{4}{5}x + \tfrac{11}{2} \tag{66}$$

$$\sigma_2^2 = \left(\tfrac{2}{5}x - \tfrac{1}{5}\right)^2 + \tfrac{1}{2} \tag{67}$$

In this scenario, y_1 was a *larger-the-better* response and y_2 was a *smaller-the-better* response.

4.3.1.1 The Köksoy and Yalcinoz (KY) Procedure

Since the responses were deemed equally important, the weights in the KY procedure were $W_1 = W_2 = \tfrac{1}{2}$. Thus, the jointly robust point x^{KY} was found by minimizing $KY = \tfrac{1}{2}MSE_1 + \tfrac{1}{2}MSE_2$ where $MSE_1 = \sigma_1^2 - \mu_1^2$ and $MSE_2 = \sigma_2^2 + \mu_2^2$. Also, in order to examine any trade-offs made among the responses in determining a jointly robust operating point, x^{KY} was compared to the results of the two single response RPD problems suggested by Equations (62) and (63). That is, the robust point for y_1 was identified without any consideration of y_2 by minimizing MSE_1 to yield the solution x_1^{KY}. Similarly, MSE_2 was minimized to find the robust point for y_2 only. This solution is identified as x_2^{KY}.

Figure 13 illustrates the results of solving the three RPD problems. The locations of the single response robust solutions are shown in Figure 13a and Figure 13b. Clearly, these individual solutions occur at opposite ends of the range of x suggesting that some compromise will need to be made to identify an operating point that is mutually robust for y_1 and y_2. Finally, the jointly robust point x^{KY} is shown in Figure 13c. It can be observed that, even though the responses were given equal weights, x^{KY} is nearly

identical to x_1^{KY}. It appears as if there was no compromise amongst the responses. This can be further examined by decomposing *KY*.
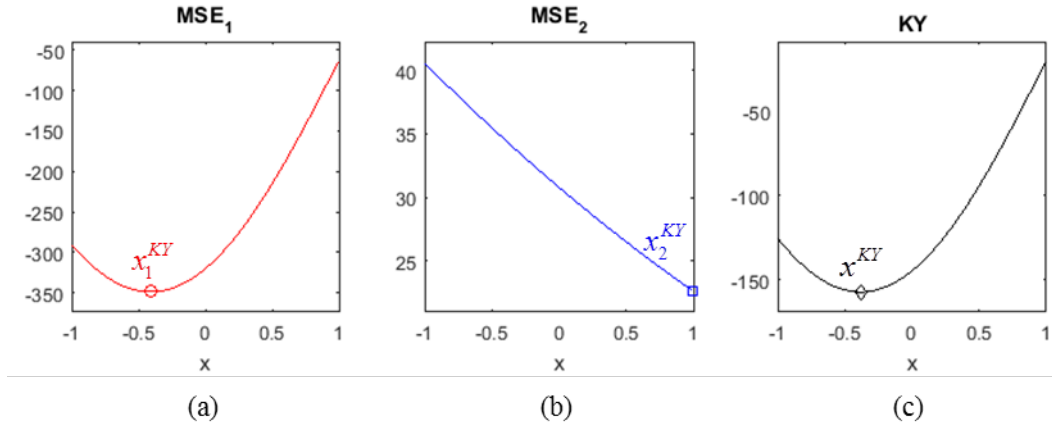
Figure 13. Robust Point Locations for the Synthetic Case Using the *KY* Procedure

Figure 14 plots *KY* with its decomposed functions, namely $\frac{1}{2}MSE_1$ and $\frac{1}{2}MSE_2$, on the same graph. It is now clear that x^{KY} is indistinguishable from x_1^{KY} due to the fact that $\frac{1}{2}MSE_2$ is relatively unchanging when compared to $\frac{1}{2}MSE_1$ and has little to no effect on *KY*. The range of $\frac{1}{2}MSE_1$ is 142.8 while the range of $\frac{1}{2}MSE_2$ is only 8.95. Based on this simple decomposition, it can be concluded that the jointly robust solution is greatly influenced by y_1 whereas y_2 has a negligible effect on the joint solution.
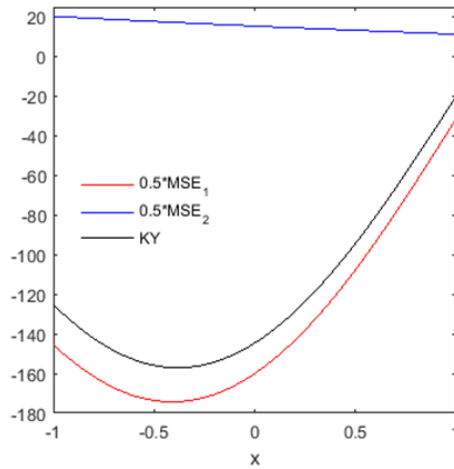
Figure 14. Decomposition of *KY* for the Synthetic Case Study

Each response's MSE function can be further examined by decomposing them into their variance and squared bias terms. In the case of MSE_1, its variance and negative squared bias term $(-\mu_1^2)$ will actually be investigated. Again, it is shown in Figure 15a and Figure 15b that the MSE functions are driven by a single term—in each case, the squared bias. The variance terms have minimal influence on the MSE functions and, as a result, minimal influence on determining x_1^{KY}, x_2^{KY}, or x^{KY}. It can be concluded that, for this example, the jointly robust solution is influenced significantly by μ_1. Conversely, μ_2, σ_1^2, and σ_2^2 have limited, if any, impact on the overall RPD solution derived via the *KY* procedure.
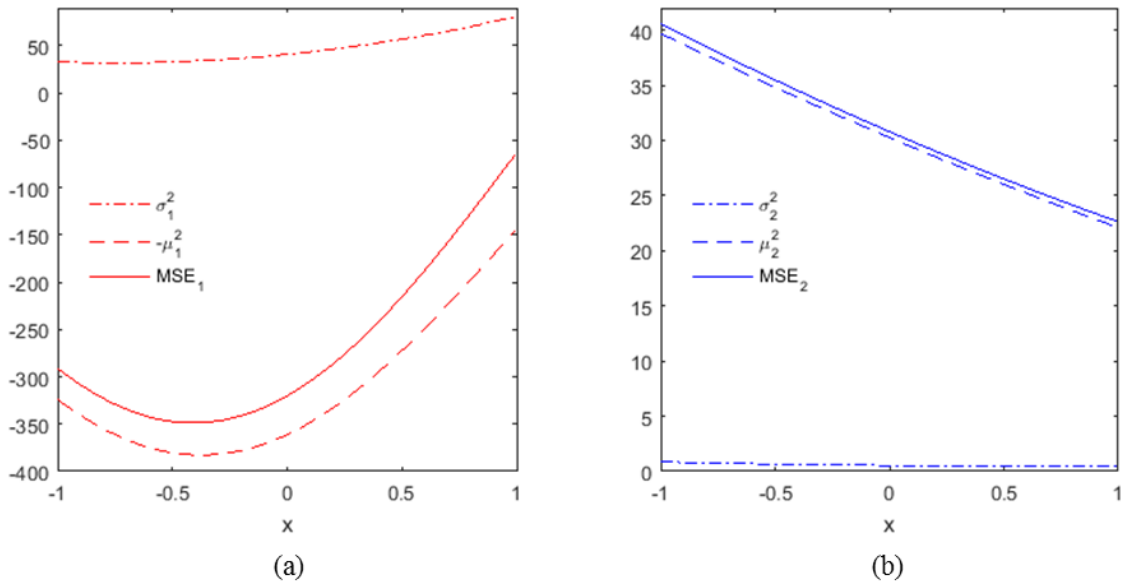
Figure 15. Decomposition of Individual MSE Functions for the Synthetic Case Study

4.3.1.2 The Nested Desirability (ND) Procedure

Linear $(r = 1)$ desirability functions were utilized for the *ND* procedure. Again, since the responses hold equal importance, the weights in the *ND* formulation of the problem were $W_1 = W_2 = 1$. Now, the *KY* procedure does not consider weighting the contributions of the individual means and variances; here, the *ND* procedure weighted them equally. Specifically, $w_{j,\mu} = w_{j,\sigma^2} = 1$ for $j = 1, 2$. The jointly robust point x^{ND}

maximizes $D = \left(\left(D_1 \right)^1 \times \left(D_2 \right)^1 \right)^{1/2}$ where $D_1 = \left(\left(d_{1,\mu} \right)^1 \times \left(d_{1,\sigma^2} \right)^1 \right)^{1/2}$ and

$D_2 = \left(\left(d_{2,\mu} \right)^1 \times \left(d_{2,\sigma^2} \right)^1 \right)^{1/2}$. Even though the weights are equal to 1, they were included here as the exponents for completeness. Again, the results of the two single response RPD problems are presented. That is, the robust point for y_1, denoted x_1^{ND}, was found by

66

maximizing D_1. Similarly, D_2 was maximized to find the robust point for y_2 only. This solution is identified as x_2^{ND}.

Figure 16 illustrates the results of solving the three RPD problems using the *ND* approach. The locations of the single response robust solutions are shown in Figure 16a and Figure 16b. Again, as the *KY* procedure showed, these solutions are conflicting. Finally, the jointly robust point x^{ND} is shown in Figure 16c. As opposed to the results of the *KY* procedure, some compromise can now be observed between the two responses in order to generate a jointly robust point. Table 15 and Table 16 can be examined further in Section 4.3.1.3 to see where these compromises were made.
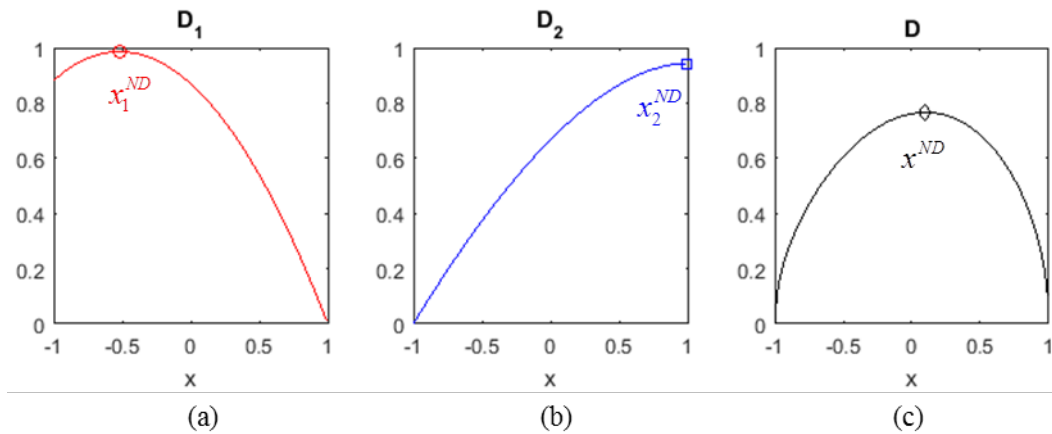
Figure 16. Robust Point Locations for the Synthetic Case Using the *ND* Procedure

4.3.1.3 Comparison of the KY Procedure and the ND Procedure

The results of the *KY* procedure and the *ND* procedure can be observed through two different points-of-view (POV). Table 15 summarizes the means, variances, and MSEs of the three robust points for the two RPD strategies by taking a "MSE POV." The italicized values are the corresponding results of the secondary response when each single

response RPD problem is solved. The bold values are the results of solving the multi-response RPD problems. As noted in Figure 13a and Figure 13c, x_1^{KY} and x^{KY} are nearly equal. Therefore, when an equally important y_2 is added into the problem formulation, there exists a slight trade-off by increasing σ_1^2 for marginal improvements in μ_2 and σ_2^2. For all intents and purposes, there is practically zero compromise between the responses in the joint RPD case when using the *KY* procedure. This is not, however, the situation when using the *ND* procedure where noticeable trade-offs occur amongst the means and variances of the two responses. Now, when y_2 is considered as important as y_1, larger degradations in μ_1 and σ_1^2 are incurred in exchange for substantial improvements in μ_2 and σ_2^2. Based on the decision maker's original preferences, this is a solution they are willing to accept. However, by observing the results from an MSE POV, x^{ND} represents a 12.51% decrease in overall "*KY* value" when compared to x^{KY}. This is understandable since x^{KY} is the best solution in terms of MSE.

Table 15. Synthetic Multi-Response Case Study from a MSE Point-of-View

Robust Point	μ_1	σ_1^2	MSE_1	μ_2	σ_2^2	MSE_2	KY
RPD Goal $\rightarrow$	*Max*	*Min*	*Min*	*Min*	*Min*	*Min*	*Min*
$x_1^{KY} = -0.41$	19.56	33.85	-348.74	*5.83*	*0.63*	*34.62*	-157.06
$x_2^{KY} = 1.00$	*12.00*	*81.00*	*-63.00*	4.70	0.54	22.63	-20.19
$x^{KY} = -0.38$	**19.56**	**34.17**	**-348.42**	**5.81**	**0.62**	**34.38**	**-157.02**
$x_1^{ND} = -0.52$	19.48	32.86	-346.61	*5.91*	*0.67*	*35.60*	-155.51
$x_2^{ND} = 1.00$	*12.01*	*80.93*	*-63.31*	4.70	0.54	22.63	-20.34
$x^{ND} = 0.10$	**18.66**	**43.53**	**-304.67**	**5.42**	**0.53**	**29.91**	**-137.38**
% Improvement from x^{KY} to x^{ND}	-4.60%	-27.39%	-12.56%	6.71%	14.52%	13.02%	**-12.51%**

Additionally, the results can also be examined from a "desirability POV." Table 16 summarizes the desirabilities for the means, variances, and responses of the three robust points for the two RPD strategies. The trade-offs in the desirability POV are similar to the trade-offs in the MSE POV. However, when looking at the same results through a desirability lens, utilizing x^{ND} as the jointly robust point instead of x^{KY} results in 12.00% and 20.83% decreases in desirability for μ_1 and σ_1^2, respectively. In exchange for this, however, the desirability levels increase significantly—77.42% and 43.08%—for μ_2 and σ_2^2. This ultimately results in a 16.67% increase in the overall system desirability at x^{ND}. Again, this is expected since x^{ND} is the best solution in terms of desirability.

Table 16. Synthetic Multi-Response Case Study from a Desirability Point-of-View

Robust Point	$d_{1,\mu}$	d_{1,σ^2}	D_1	$d_{2,\mu}$	d_{2,σ^2}	D_2	D
$x_1^{KY} = -0.41$	1.00	0.96	0.98	*0.30*	*0.63*	*0.43*	0.65
$x_2^{KY} = 1.00$	*0.00*	*0.00*	*0.00*	1.00	0.89	0.94	0.00
$x^{KY} = -0.38$	**1.00**	**0.96**	**0.98**	**0.31**	**0.65**	**0.45**	**0.66**
$x_1^{ND} = -0.52$	0.99	0.98	0.99	*0.24*	*0.54*	*0.36*	0.60
$x_2^{ND} = 1.00$	*0.00*	*0.00*	*0.00*	1.00	0.89	0.94	0.04
$x^{ND} = 0.10$	**0.88**	**0.76**	**0.82**	**0.55**	**0.93**	**0.71**	**0.77**
% Improvement from x^{KY} to x^{ND}	-12.00%	-20.83%	-16.33%	77.42%	43.08%	57.78%	**16.67%**

4.3.2 The Force Transducer Experiment

Details for the following example can be found in Romano et al. [22]. Several authors have also used this example as a case study for their multi-response RPD methods [30–32]. In short, the problem consists of two response variables (y_1 and y_2), three control variables (x_1, x_2 and x_3), and two noise variables (z_1 and z_2). The experimental results are displayed in Table 17. The noise factors were assumed to be independent with zero mean and variances σ_1^2 and σ_2^2. Thus, the ± 1 experimental levels for z_j were set to $\pm \sigma_j$. Per Romano et al. [22], the fitted response surface functions for y_1 and y_2 are

$$\hat{y}_1(\mathbf{x}, \mathbf{z}) = 1.38 - 0.361x_1 - 0.155x_2 + 0.0771x_3 - 0.148x_1x_2 + 0.0218x_1x_3 \\ + 0.0130x_2x_3 + 0.0481x_1^2 - 0.0588z_1 - 0.0116z_2 + 0.0100x_1z_1 \tag{68}$$

and

$$\hat{y}_2(\mathbf{x}, \mathbf{z}) = 1.64 - 0.592x_1 + 0.438x_2 - 0.0950x_3 + 0.301x_1x_2 - 0.143x_1x_3$$
$$+0.201x_1^2 - 0.0844x_1x_2x_3 + 0.0794x_1z_1 \tag{69}$$

The authors reported that no lack of fit was detected and the MSEs of the models were

$\hat{\sigma}_{\varepsilon,1}^2 = 0.0003253$ and $\hat{\sigma}_{\varepsilon,2}^2 = 0.024$.

Table 17. Experimental Results for the Force Transducer Experiment

Run	x_1	x_2	x_3	z_1	z_2	y_1	y_2
1	-1	-1	-1	-1	1	1.81	1.10
2	-1	-1	-1	1	-1	1.69	1.11
3	-1	-1	1	-1	-1	1.90	1.07
4	-1	-1	1	1	1	1.78	1.07
5	-1	1	-1	-1	-1	1.80	1.47
6	-1	1	-1	1	1	1.63	1.18
7	-1	1	1	-1	1	1.92	1.41
8	-1	1	1	1	-1	1.78	1.58
9	1	-1	-1	-1	-1	1.36	1.57
10	1	-1	-1	1	1	1.22	2.03
11	1	-1	1	-1	1	1.48	1.38
12	1	-1	1	1	-1	1.44	1.68
13	1	1	-1	-1	1	0.693	3.37
14	1	1	-1	1	-1	0.616	3.75
15	1	1	1	-1	-1	0.950	2.81
16	1	1	1	1	1	0.817	2.83
17	-1	0	0	0	0	1.79	1.24
18	1	0	0	0	0	1.03	2.46
19	0	-1	0	0	0	1.53	1.23
20	0	1	0	0	0	1.22	1.73
21	0	0	-1	0	0	1.30	1.63
22	0	0	1	0	0	1.44	1.67
23	0	0	0	0	0	1.38	1.73
24	0	0	0	0	0	1.39	1.74
25	0	0	0	0	0	1.40	1.74

The goal of this RPD study was to find the robust settings of x_1, x_2 and x_3 that

minimized y_1 and y_2. Similar to the synthetic case study in Section 4.3.1, this problem

71

was solved using the *KY* procedure and the *ND* procedure. The estimated mean models

are

$$\hat{\mu}_1 = 1.38 - 0.361x_1 - 0.155x_2 + 0.0771x_3 - 0.148x_1x_2 + 0.0218x_1x_3$$
$$+ 0.0130x_2x_3 + 0.0481x_1^2 \tag{70}$$

and

$$\hat{\mu}_2 = 1.64 - 0.592x_1 + 0.438x_2 - 0.0950x_3 + 0.301x_1x_2 - 0.143x_1x_3$$
$$+ 0.201x_1^2 - 0.0844x_1x_2x_3 \tag{71}$$

The estimated variance models are

$$\hat{\sigma}_1^2 = \left(-0.0588 + 0.01x_1\right)^2 + \left(-0.0116\right)^2 + \hat{\sigma}_{\varepsilon,1}^2 \tag{72}$$

and

$$\hat{\sigma}_2^2 = \left(0.0794x_1\right)^2 + \hat{\sigma}_{\varepsilon,2}^2 . \tag{73}$$

Table 18 summarizes the means, variances, and MSEs of the three robust points

for the two RPD strategies from a "MSE POV." In this case, as opposed to the synthetic

case study, both the *KY* procedure and the *ND* procedure generated solutions in which

compromises were made amongst the means and variances. However, x^{ND} represents a

33.29% decrease in overall "*KY* value" when compared to x^{KY}. Similar conclusions

regarding trade-offs between the response means and variances can be made by

examining Table 19. Now, from a "desirability POV," utilizing x^{ND} as the jointly robust

point instead of x^{KY} yields a 15.25% increase in the overall system desirability.

Table 18. Force Transducer Experiment Results from a MSE Point-of-View

Robust Point	μ_1	σ_1^2	MSE_1	μ_2	σ_2^2	MSE_2	KY
RPD Goal $\rightarrow$	*Min*	*Min*	*Min*	*Min*	*Min*	*Min*	*Min*
$x_1^{KY} = (1.00, 1.00, -1.00)$	0.65	0.0028	0.43	*3.49*	*0.0302*	*12.21*	6.32
$x_2^{KY} = (-0.57, -1.00, 1.00)$	*1.72*	*0.0046*	*2.96*	1.04	0.0260	1.11	2.04
$\mathbf{x^{KY} = (0.07, -1.00, 1.00)}$	**1.59**	**0.0038**	**2.53**	**1.12**	**0.0240**	**1.28**	**1.91**
$x_1^{ND} = (1.00, 1.00, -1.00)$	0.65	0.0028	0.43	*3.49*	*0.0302*	*12.21*	6.32
$x_2^{ND} = (-0.04, -1.00, 1.00)$	*1.61*	*0.0040*	*2.60*	1.10	0.0240	1.23	1.92
$\mathbf{x^{ND} = (0.32, -0.20, -1.00)}$	**1.23**	**0.0036**	**1.52**	**1.88**	**0.0246**	**3.56**	**2.54**
% Improvement from x^{KY} to x^{ND}	22.64%	5.26%	40.29%	-67.86%	-2.50%	-176.79%	**-33.29%**

Table 19. Force Transducer Experiment Results from a Desirability Point-of-View

Robust Point	$d_{1,\mu}$	d_{1,σ^2}	D_1	$d_{2,\mu}$	d_{2,σ^2}	D_1	D
$x_1^{KY} = (1.00, 1.00, -1.00)$	1.00	1.00	1.00	*0.00*	*0.00*	*0.00*	0.00
$x_1^{KY} = (-0.57, -1.00, 1.00)$	*0.11*	*0.24*	*0.16*	1.00	0.68	0.82	0.36
$\mathbf{x^{KY} = (0.07, -1.00, 1.00)}$	**0.22**	**0.58**	**0.36**	**0.97**	**1.00**	**0.98**	**0.59**
$x_1^{ND} = (1.00, 1.00, -1.00)$	1.00	1.00	1.00	*0.00*	*0.00*	*0.00*	0.00
$x_2^{ND} = (-0.04, -1.00, 1.00)$	*0.20*	*0.52*	*0.32*	0.98	1.00	0.99	0.57
$\mathbf{x^{ND} = (0.32, -0.20, -1.00)}$	**0.52**	**0.70**	**0.60**	**0.66**	**0.90**	**0.77**	**0.68**
% Improvement from x^{KY} to x^{ND}	136.36%	20.69%	66.67%	-31.96%	-10.00%	-21.43%	**15.25%**

4.4 Summary

MSE-based strategies for solving multi-response RPD problems are popular

methods. However, as the synthetic case study showed, sometimes the mean or variance

of one response can render the means and variances of the remaining responses

insignificant to the overall RPD problem. This can diminish the chance of finding a

compromised solution. Also, trade-offs among the means and variances of the individual

responses are typically ignored. Chapter IV presented an approach, based on well-known

desirability functions, that is beneficial in two ways. First, it places the responses, as well

as their means and variances, on equal footing. Second, it allows a decision maker to

declare their personal preferences for the responses' means and variances from the outset.

The resultant operating point is a system setting that, whether one is looking at the

problem from a MSE POV or a desirability POV, produces a mutually robust set of

responses the decision maker considers acceptable.

V. Quality Measures for Comparing Multiple RPD Strategies

5.1 Introduction

There is a growing literature in which multiple RPD problem solving strategies are contrasted. Articles typically report a system's predicted mean and variance found using a number of approaches; however, some neglect to demonstrate the quality of those predictions through confirmation experiments [22, 24, 25, 29–32, 46, 71]. There are two operative questions. Are the predictions good estimates of the system's actual performance at the estimated robust settings? Also, which method's robust point estimate actually realizes the best combination of mean and variance? These questions become more challenging when there are multiple responses of interest. To provide a framework for addressing these questions, this research first posits that any RPD problem solving methodology can be evaluated on the basis of the accuracy of its predictions and its realized robustness relative to its competitors.

Chapter V addresses the third objective of this dissertation which is aimed at generating measures for comparing different RPD problem solving strategies. Such measures can increase the understanding of each strategy and allow the analyst to make a more knowledgeable evaluation of the competing procedures. The rest of Chapter V is organized as follows. Section 5.2 introduces a methodology for meeting this research objective. Section 5.3 conducts a case study using a discrete event simulation that compares the results of 12 competing robust design strategies. Finally, Section 5.4 summarizes this chapter.

5.2 Quality Indices for Robust Parameter Design

As mentioned, the RPD literature generally compares the system's predicted mean and variance at the robust setting determined through one approach to the system's predicted mean and variance at the robust setting found though a competing method. Authors then make a choice about which method is superior based upon the closeness of these predictions to ideal mean and variance targets. For example, in Taguchi's *target-is-best* scenario, one would prefer a robust point with a predicted mean close to the target value with a small predicted variance over another setting with a predicted mean further from the target having a higher predicted variance. Based on the predictions, it is expected that the former operating point is better than the latter. It is this comparison that drives some authors to conclude the dominance of one methodology over another. However, if the methodologies provide poor predictions, then their conclusions may be inaccurate.

To make a more comprehensive assessment of the competing strategies, this research recommends expanding the analysis to include confirmation experiments at each robust setting. First, it must be stated that this is obviously not feasible in all cases. It is highly unlikely that a capital-producing manufacturing line would shut down in order to make some trial runs. It may also be extremely costly or dangerous to make these test runs in other situations. However, these confirmatory runs may be easy and inexpensive to perform when simulations are used. By performing confirmation experiments at each robust setting, further insight can be gained into each RPD approach's *accuracy* and *robustness* qualities. Accuracy is a quality measure of how close the predicted mean and variance of the system are to the realized mean and variance of the system when it is

repeatedly executed at a specific operating point. Robustness, on the other hand, assesses how close the system's realized mean and variance are to their desired target values. A minimum response variance is always preferred; however, Taguchi's overall experimental goal governs whether a minimum, maximum, or specific mean response is required. Combining the qualities of accuracy and robustness can strengthen the understanding of each RPD approach and a more informed evaluation of the competing procedures can be made.

The proposed approach for using accuracy and robustness to compare multiple RPD strategies is now presented for a general problem in which there exist K strategies and J responses. For $k = 1, 2, ..., K$ and $j = 1, 2, ..., J$, define the following:

- $\mathbf{x}_k$ is the robust point found using RPD strategy k

- $m_{k,j}^{(p)}$ and $v_{k,j}^{(p)}$ are the system's predicted mean and variance, respectively, at $\mathbf{x}_k$ for response j

- $m_{k,j}^{(r)}$ and $v_{k,j}^{(r)}$ are the system's realized mean and variance, respectively, at $\mathbf{x}_k$ for response j

- t_j is the desired target value for response j if it is a *target-is-best* experimental scenario

These values can be visualized in tabular form as shown in Table 20. The last row of Table 20 is labeled "Ideal." The ideal values are the best possible means and variances that can be attained for each response of the system. These can be gained analytically or through experimentation when known functions or simulations are being used. Let y_j represent the true response j. The ideal mean for response j is then defined as:

$$m_j^* = \begin{cases} \min E\left[y_j\right] & \text{if response } j \text{ is a "smaller-the-better" scenario} \\ t_j & \text{if response } j \text{ is a "target-is-best" scenario} \\ \max E\left[y_j\right] & \text{if response } j \text{ is a "larger-the-better" scenario} \end{cases} \qquad (74)$$

The ideal variance for response j is defined as

$$v_j^* = \min Var\left[y_j\right]. \qquad (75)$$

Table 20. Tabular Visualization of RPD Results for K Strategies and J Responses

Robust Point		μ_1	σ_1^2		μ_J	σ_J^2
	RPD Goal $\rightarrow$	*t_1/Min/Max*	*Min*		*t_J/Min/Max*	*Min*
$\mathbf{x}_1$	Predicted	$m_{1,1}^{(p)}$	$v_{1,1}^{(p)}$		$m_{1,J}^{(p)}$	$v_{1,J}^{(p)}$
	Realized	$m_{1,1}^{(r)}$	$v_{1,1}^{(r)}$		$m_{1,J}^{(r)}$	$v_{1,J}^{(r)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$\mathbf{x}_K$	Predicted	$m_{K,1}^{(p)}$	$v_{K,1}^{(p)}$		$m_{K,J}^{(p)}$	$v_{K,J}^{(p)}$
	Realized	$m_{K,1}^{(r)}$	$v_{K,1}^{(r)}$		$m_{K,J}^{(r)}$	$v_{K,J}^{(r)}$
	Ideal	m_1^*	v_1^*		m_J^*	v_J^*

5.2.1 Accuracy Quality Index

Previously, the accuracy of a RPD strategy was defined as how close the predictions of the system are to the actual realizations of the system when it is continually executed at a specific operating point. Therefore, the row in Table 20 labeled "Ideal" can be ignored at this time. Since the values in Table 20 are most likely in different units or scales, each column of the table containing the "Predicted" and "Realized" rows must first be normalized to the range [0,1] to eliminate these effects. Normalized values are denoted with "~". For example, $\tilde{m}_{k,j}^{(p)}$ or $\tilde{v}_{k,j}^{(r)}$. Now define the following vectors:

- $\mathbf{p}_k = \left[\tilde{m}_{k,1}^{(p)}, \tilde{v}_{k,1}^{(p)}, ..., \tilde{m}_{k,J}^{(p)}, \tilde{v}_{k,J}^{(p)} \right]$ is the vector of normalized predicted means and variances for RPD strategy k

- $\mathbf{r}_k = \left[\tilde{m}_{k,1}^{(r)}, \tilde{v}_{k,1}^{(r)}, ..., \tilde{m}_{k,J}^{(r)}, \tilde{v}_{k,J}^{(r)} \right]$ is the vector of normalized realized means and variances for RPD strategy k

The *Accuracy Quality Index* for strategy k across all J responses, denoted A_k, is defined as a measure of the distance between its vector of normalized predicted values $\mathbf{p}_k$ and its vector of normalized realized values $\mathbf{r}_k$. It is computed using Equation (75).

$$A_k = 1 - \left\| \mathbf{p}_k - \mathbf{r}_k \right\| \Big/ \sqrt{2J} \tag{76}$$

The distance measure in Equation (76) is first scaled by the maximum possible distance between any two points in the 2J-dimensional unit hypercube and is then subtracted from 1. This procedure results in A_k being contained in the interval [0,1] in which larger values are preferred.

Accuracy can be illustrated using the single response synthetic case study from Section 3.3.1. To recap, the combined-array RSM approach using the *Std*, *NN*, *CNN*, and *CCN* models were tested against the combined-array MDM approach using the *KR* and *RBF* models. The predicted and realized means and variances for each of the 6 RPD strategies are shown in Table 9. Figure 17 plots the normalized predicted means and variances for each RPD procedure against their normalized realized values. The accuracy is a measure of the length of the dashed lines connecting the predicted and realized values. It is obvious that the MDM approach using the RBFNN model is the most accurate whereas the RSM approach using the standard model is the least accurate.
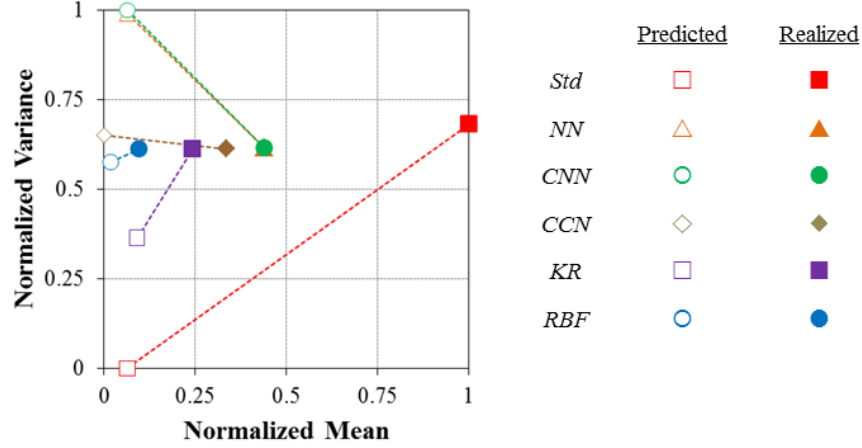
79

Figure 17. Accuracy of Each Strategy for the Single Response Synthetic Case

5.2.2 Robustness Quality Index

The robustness of a RPD strategy relative to its competitors, on the other hand, is defined as how close the system's actual realizations are to the desired mean and variance targets. Thus, the rows in Table 20 labeled "Predicted" can now be ignored. Again, each column of the table containing the "Realized" and "Ideal" rows must be normalized to the range [0,1]. Now, with "~" once more denoting normalized values, define the following vectors:

- $\mathbf{r}_k = \left[\tilde{m}_{k,1}^{(r)}, \tilde{v}_{k,1}^{(r)}, ..., \tilde{m}_{k,J}^{(r)}, \tilde{v}_{k,J}^{(r)} \right]$ is the vector of normalized realized means and variances for RPD strategy k

- $I^* = \left[\tilde{m}_1^*, \tilde{v}_1^*, ..., \tilde{m}_J^*, \tilde{v}_J^* \right]$ is the vector of normalized ideal means and variances

It should be noted that I^* may not represent a combination of system means and variances that can actually be achieved by any control factor setting in the design space. This is why I^* is treated as an ideal, or utopian, vector of system means and variances.

80

The *Robustness Quality Index* for strategy k across all J responses, denoted R_k, is defined as a measure of the distance between its vector of normalized realized values $\mathbf{r}_k$ and the vector of normalized ideal values I^*. It is computed using Equation (76).

$$R_k = 1 - \left\| \mathbf{r}_k - I^* \right\| / \sqrt{2J} \tag{77}$$

Similar to A_k, larger values of R_k in the interval [0,1] are preferred. The synthetic case study from Section 3.3.1 can further illustrate the robustness quality. The goal of the study was to locate the operating point that resulted in a response of 8 with minimal variance. In this case, the smallest achievable variance in the design space is 18. This is found by minimizing Equation (54). Therefore, the ideal vector was $I^* = [8,18]$. Figure 18 plots the normalized realized means and variances for each RPD procedure against their normalized ideal values. Robustness is a measure of the length of the dashed lines connecting the realized and ideal values. The operating point found via the MDM approach using the RBFNN model stands out as the most robust of the 6 solutions.

As a matter of fact, Figure 18 also shows that points labeled "*d*" are solutions that are completely dominated by another solution. A solution is dominated if another solution resulted in a better realized mean and a better realized variance. In this case, *Std* is dominated by the other 5 solutions. Also, *NN* and *CNN* are dominated by *CCN, KR*, and *RBF*. *CCN* is dominated by *KR* and *RBF*. Finally, *KR* and *RBF* are considered non-dominated solutions.
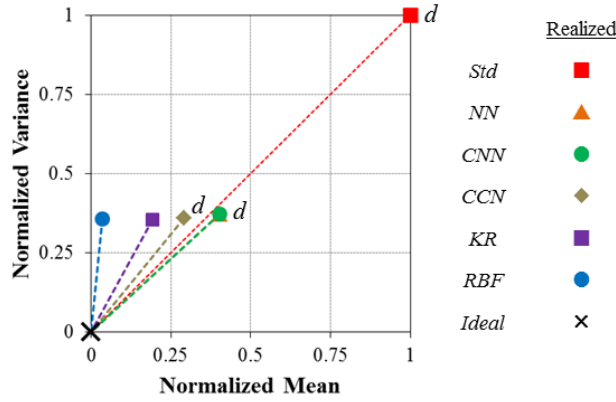
Figure 18. Robustness of Each Strategy for the Single Response Synthetic Case

5.2.3 Joint Quality Index

The perfect RPD strategy has two characteristics. First, it precisely predicts the realized means and variances for each response. Second, it realizes the ideal means and variances. These characteristics equate to having an Accuracy Index and a Robustness Index equal to 1. The accuracy and robustness indices can now be used to determine the overall *Joint Quality Index* for each strategy k by utilizing the equation

$$Q_k = 1 - \left\| \mathbf{q}_k - \mathbf{1} \right\| / \sqrt{2} \tag{78}$$

where $\mathbf{q}_k = [A_k, R_k]$ and $\mathbf{1} = [1,1]$. Again, the distance measure in Equation (78) is scaled by the maximum possible distance between any two points in the unit square and is then subtracted from 1. The strategy resulting in the largest value for Q_k is then determined to be the best complete procedure—out of the K competing strategies—for solving the RPD problem across the J responses concurrently. Figure 19 plots the Accuracy Index and the Robustness Index for each RPD strategy in the synthetic case study from Section 3.3.1. The overall joint quality index is simply a measure of the distance between each

82

strategy's accuracy-robustness pairing and the point (1,1). Again, the operating point found via the MDM approach using the RBFNN model shows to have the best overall quality.
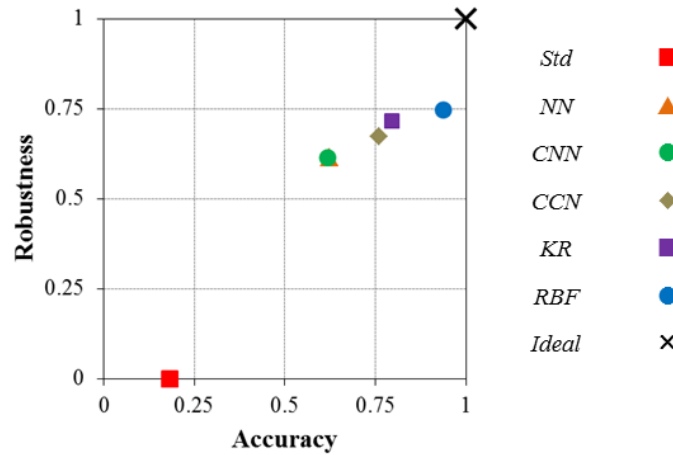
Figure 19. Overall Joint Quality of Each Strategy for the Single Response Synthetic Case

The accuracy, robustness, and overall quality indices that correspond to Figure 17, Figure 18, and Figure 19 are displayed in Table 21. It is clear that for this case study, of the 6 RPD procedures, the MDM approach that utilizes the *RBF* model is superior. A comment must also be made in regards to the *Std* model's robustness score of 0. The RSM approach that utilizes the standard response surface model has been proven to be a very successful strategy. However, the accuracy, robustness, and overall quality indices as defined here are relative to the *K* strategies being compared in the study. Therefore, as it relates to the other five competitors in this particular problem, the RSM approach using the *Std* model does not produce a very robust solution.

Table 21. Quality Indices for Each Strategy in the Single Response Synthetic Case

RPD Strategy	k	Accuracy A_k	Robustness R_k	Overall Quality Q_k
RSM w/ *Std*	1	0.18 (*6*)	0.00 (*6*)	0.09 (*6*)
RSM w/ *NN*	2	0.62 (*4*)	0.61 (*4*)	0.62 (*4*)
RSM w/ *CNN*	3	0.62 (4)	0.61 (*4*)	0.62 (4)
RSM w/ *CCN*	4	0.76 (*3*)	0.67 (*3*)	0.71 (*3*)
MDM w/ *KR*	5	0.79 (*2*)	0.72 (*2*)	0.75 (*2*)
MDM w/ *RBF*	6	0.94 (*1*)	0.75 (*1*)	0.82 (*1*)

Rank in Parenthesis ( )

5.3 Application and Results

The benefits of a more holistic RPD strategy assessment using the quality index will be demonstrated with a discrete event simulation. The *ND* procedure from Chapter IV was assessed against the *KY* procedure using six different RPD strategies: the combined-array RSM approach that employs the *Std*, *NN*, *CNN*, and *CCN* models and the MDM approach that employs the *KR* and *RBF* models. The proposed quality indices were then used to compare the 12 different multi-response strategies listed in Table 22.

Table 22. Multi-Response RPD Strategies Used for Comparison

Label	Strategy	Modeling Approach
KY^{Std}	Minimize *KY*	RSM w/ Standard Model
ND^{Std}	Maximize *D*	RSM w/ Standard Model
KY^{NN}	Minimize *KY*	RSM w/ Noise-by-Noise Model
ND^{NN}	Maximize *D*	RSM w/ Noise-by-Noise Model
KY^{CNN}	Minimize *KY*	RSM w/ Control-by-Noise-by-Noise Model
ND^{CNN}	Maximize *D*	RSM w/ Control-by-Noise-by-Noise Model
KY^{CCN}	Minimize *KY*	RSM w/ Control-by-Control-by-Noise Model
ND^{CCN}	Maximize *D*	RSM w/ Control-by-Control -by-Noise Model
KY^{KR}	Minimize *KY*	MDM w/ Kriging Model
ND^{KR}	Maximize *D*	MDM w/ Kriging Model
KY^{RBF}	Minimize *KY*	MDM w/ RBFNN Model
ND^{RBF}	Maximize *D*	MDM w/ RBFNN Model

A multi-response RPD study was performed using Kelton et al.'s [84] automotive maintenance and repair shop (AMRS) simulation (Model 6-1) as it was developed in Arena®. The two performance measures of interest were the average daily profit and the average daily number of late wait jobs, labeled y_1 and y_2 respectively. The responses were averages of 100 independent replicates at a specified design setting. The input factors that influence y_1 and y_2 are listed in Table 23. The factors x_1 and x_2 were controllable within their minimum and maximum values. Also, since the desire was to find the control factor settings that were robust to uncertain demand, the number of calls that arrive to the shop each day was a Poisson random variable with mean z_1. The mean of the Poisson distribution itself was assumed to be normally distributed with a known mean and standard deviation. This is comparable to the approach taken by Wild and Pignatiello [44] in which they found a job shop design that was robust to uncontrollable environmental factors such as the mean inter-arrival times for parts. Finally, in order to structure the RBFNN metamodel so it was well-generalized with a minimal risk of over-fitting the experimental data, 10 rounds of the 10-fold cross validation procedure were used.

Table 23. Input Factors for the AMRS Problem

Label	Factor	Min	Max	Mean	Std Dev
x_1	Maximum work hours available per day	20	40		
x_2	Service buffer hours allowed for waiting customers	0.5	2		
z_1	Mean number of calls per day			25	3

The goal of this RPD study was to identify the robust setting of x_1 and x_2 that jointly maximized the average daily profit (a *larger-the-better* case) and minimized the average daily number of late wait jobs (a *smaller-the-better* case) with minimum variability for each response. Each response was treated with equal weighting. Each mean and its associated variance were also treated with equal weighting. A 5^3 full-factorial combined array design was used to build each model. The mean and variance models for the *Std*, *NN*, *CNN*, *CCN*, *KR*, and *RBF* response models can be viewed in Appendix B. Table 24 summarizes the predicted, realized, and ideal means and variances of y_1 and y_2 at the 12 robust points. The realized values are the results of 100 simulations of each robust operating point using common random numbers for the noise factor.

Table 24. Multi-Response RPD Results for the AMRS Simulation

RPD Strategy	Robust Point		Goal →	μ_1 Max	σ_1^2 Min	μ_2 Min	σ_2^2 Min
	x_1	x_2					
KY^{Std}	0.21	1.00	Predicted	576.68	637.39	0.43	8.03×10^{-4}
			Realized	565.11	375.86	0.43	4.05×10^{-4}
ND^{Std}	-0.10	1.00	Predicted	561.95	601.28	0.41	7.56×10^{-4}
			Realized	578.70	224.51	0.39	2.86×10^{-4}
KY^{NN}	0.21	1.00	Predicted	576.60	641.85	0.42	1.35×10^{-3}
			Realized	565.11	375.86	0.43	4.05×10^{-4}
ND^{NN}	-0.10	1.00	Predicted	559.87	605.74	0.40	1.30×10^{-3}
			Realized	578.70	224.51	0.39	2.86×10^{-4}
KY^{CNN}	0.20	1.00	Predicted	574.47	641.76	0.42	1.60×10^{-3}
			Realized	566.22	349.73	0.43	5.58×10^{-4}
ND^{CNN}	-0.18	1.00	Predicted	550.86	599.38	0.39	1.08×10^{-3}
			Realized	573.68	172.56	0.38	3.81×10^{-4}
KY^{CCN}	0.20	1.00	Predicted	574.47	563.20	0.42	1.56×10^{-3}
			Realized	566.22	349.73	0.43	5.58×10^{-4}
ND^{CCN}	0.14	1.00	Predicted	573.85	566.06	0.42	1.50×10^{-3}
			Realized	570.48	410.09	0.42	3.01×10^{-4}
KY^{KR}	-0.06	0.83	Predicted	583.46	52.51	0.43	6.44×10^{-3}
			Realized	581.85	357.48	0.44	2.54×10^{-4}
ND^{KR}	-0.29	1.00	Predicted	565.37	37.58	0.36	2.47×10^{-3}
			Realized	561.05	124.41	0.37	3.63×10^{-4}
KY^{RBF}	-0.04	1.00	Predicted	570.77	122.08	0.39	5.60×10^{-4}
			Realized	576.11	304.20	0.40	2.91×10^{-4}
ND^{RBF}	-0.12	1.00	Predicted	569.78	122.48	0.38	5.47×10^{-4}
			Realized	580.86	209.28	0.39	2.43×10^{-4}
			Ideal	**582.04**	**16.70**	**0.28**	$\mathbf{1.59\times10^{-4}}$

The first thing to note while examining Table 24 is that it is very cumbersome.

This study compares the results of 12 procedures across only two responses. It is difficult

to inspect this table and make strong conclusions regarding the study. However, utilizing

the accuracy, robustness, and joint quality measures allows for a more comprehensive

assessment of the competing strategies. Illustrations of these measures for each response

are shown in Figure 20 while the indices across both responses are displayed in Table 25.

Based on their quality indices, it is clear that ND^{RBF} and KY^{RBF} are superior to the other

10 strategies. This has two implications. First, the mean and variance predictions at their robust points are comparable to how the system actually performs at their robust points. This implies that the mean and variance models generated via the MDM approach using a RBFNN model are accurate. Second, the system's actual performance across the two responses at their robust points is more robust than the other strategies. This illustrates that, in this case, utilizing the MDM approach with a RBFNN model for an RPD study of a simulation is superior to the RSM approach that uses the polynomial regression models.

It is also interesting to note the comparison between the *ND* procedure and the *KY* procedure. The *ND* procedure outperforms the *KY* procedure for every modeling strategy in terms of robustness. That is, $R_2 > R_1$, $R_4 > R_3$, $R_6 > R_5$, $R_8 > R_7$, $R_{10} > R_9$, and $R_{12} > R_{11}$. Five of the top six approaches in terms of robustness utilize the *ND* procedure. Also, ND^{KR} and ND^{RBF} are the only non-dominated solutions across both responses. Similarly, the *ND* procedure surpasses the *KY* procedure for every modeling strategy in terms of overall joint quality. This further demonstrates that better joint robust points can be located by first putting the means and variances of all responses on a level playing field.
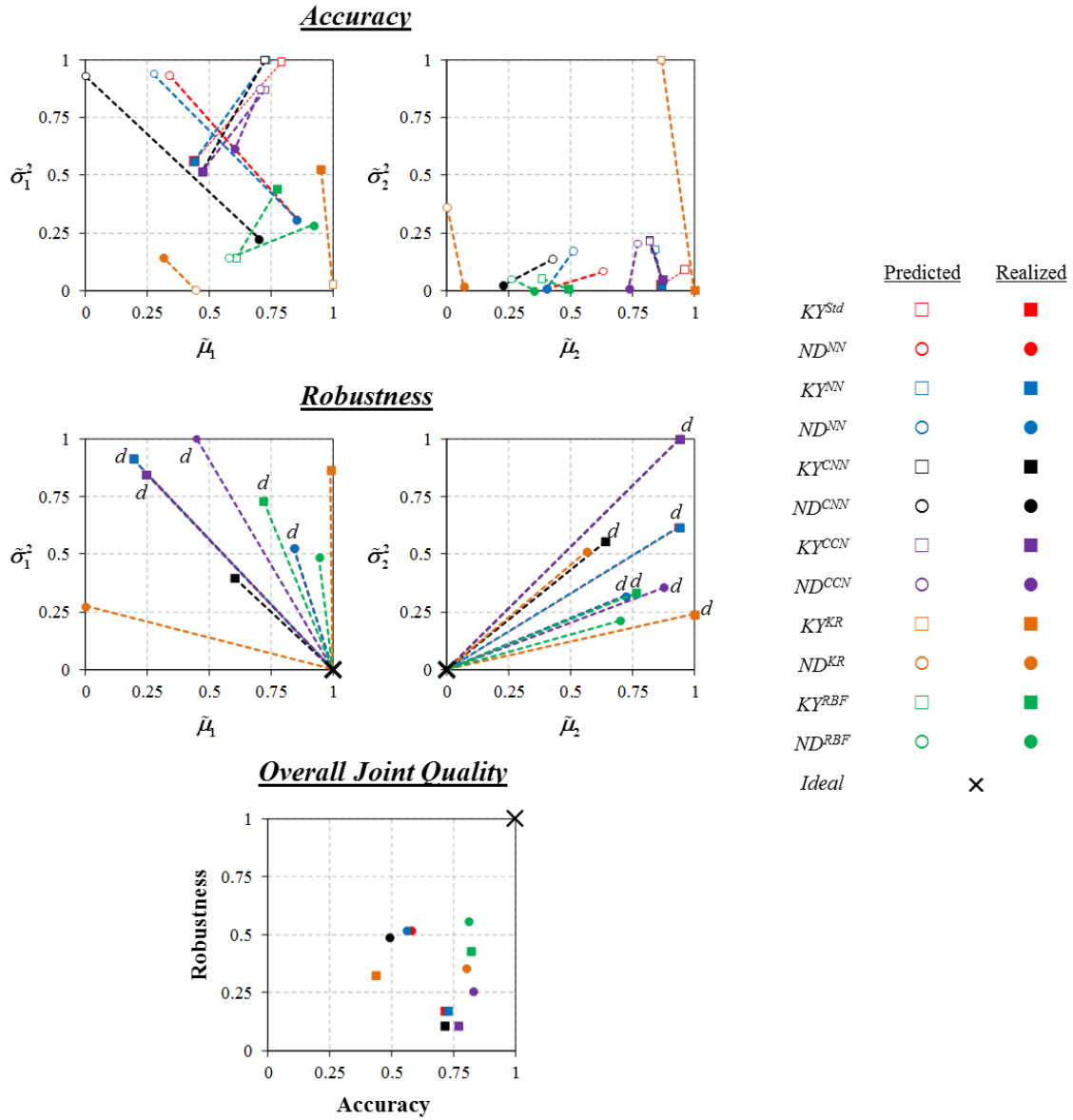
Figure 20. Accuracy, Robustness, and Overall Joint Quality of Each Strategy for the AMRS Simulation

Table 25. Quality Indices of Each RPD Strategy for the AMRS Simulation

RPD Strategy	k	A_k		R_k		Q_k	
KY^{Std}	1	0.71	*(7)*	0.17	*(9)*	0.38	*(9)*
ND^{Std}	2	0.58	*(9)*	0.52	*(2)*	0.55	*(3)*
KY^{NN}	3	0.73	*(6)*	0.17	*(9)*	0.38	*(8)*
ND^{NN}	4	0.56	*(10)*	0.52	*(2)*	0.54	*(4)*
KY^{CNN}	5	0.71	*(8)*	0.11	*(11)*	0.34	*(12)*
ND^{CNN}	6	0.49	*(11)*	0.49	*(4)*	0.49	*(6)*
KY^{CCN}	7	0.77	*(5)*	0.11	*(11)*	0.35	*(11)*
ND^{CCN}	8	0.83	*(1)*	0.26	*(8)*	0.46	*(7)*
KY^{KR}	9	0.44	*(12)*	0.33	*(7)*	0.38	*(10)*
ND^{KR}	10	0.80	*(4)*	0.36	*(6)*	0.52	*(5)*
KY^{RBF}	11	0.82	*(2)*	0.43	*(5)*	0.58	*(2)*
ND^{RBF}	12	0.81	*(3)*	0.56	*(1)*	0.66	*(1)*

Rank in Parenthesis ( )

5.4 Summary

Chapter V presented a comprehensive assessment framework for comparing the results of several RPD problem solving strategies. This approach expanded beyond the typical tactic of simply investigating how close the system's predicted means and variances at each of the established robust points are to ideal target values. When it is feasible, the use of confirmation experiments permits further examination into the accuracy and robustness qualities of each method. This approach allows for a more knowledgeable evaluation of the competing procedures.

VI. Conclusions

This dissertation addressed three fundamental objectives. The first objective was to broaden the combined-array RDM approach that relies exclusively on low-order polynomial models. The second objective was to develop an approach for multi-response RPD problems that provides a collaborative solution that is balanced across the means and variances of each response. Finally, the third objective was to generate a framework for evaluating competing RPD problem solving strategies. Chapters III–V detailed the methodology for achieving these objectives.

6.1 Original Contributions

Chapter III extended the combined-array RSM approach to include the application of non-linear metamodels. The proposed MDM approach replaced the low-order polynomial models with Kriging and RBFNN models. Then, via the Multivariate Delta Method, mean and variance models were generated from second-order Taylor series approximations of the Kriging and RBFNN models. Finally, an existing optimization problem that employed these approximations was solved to identify the robust control parameter setting. The combined-array MDM approach was compared with two current RPD strategies. First, when compared to the combined-array RSM approach that uses polynomial models, the MDM approach demonstrated improved predictive models of a computer simulation's mean and variance. Second, the MDM approach was shown to generate results that are approximately equivalent to the stochastic emulator approach at a significantly reduced computational cost.

Chapter IV proposed a multi-response RPD procedure based on well-known desirability functions that showed two benefits. First, it placed each response, as well as their means and variances, on common ground. This increased the opportunity to identify a solution that is well-balanced across the means and variances of each response. Second, it allowed a decision maker to state their personal preferences for the responses' means and variances. The resultant operating point is a system setting that, whether one is examining the problem from a MSE POV or a desirability POV, produces a mutually robust set of responses the decision maker considers acceptable.

Chapter V describes a framework for comparing and contrasting competing RPD problem solving strategies via several quality measures. This approach expanded beyond the typical tactic of only investigating how close the system's predicted means and variances at each of the established robust points are to ideal target values. By performing confirmation experiments, further insight can be gained into each RPD approach's accuracy and robustness qualities. Accuracy measures of how close the predicted mean and variance of the system are to the realized mean and variance of the system when it is repeatedly executed at a specific operating point. Robustness, on the other hand, assesses how close the system's realized mean and variance are to their desired target values. The combination of accuracy and robustness can increase an analyst's understanding of each RPD approach and allow them to make a more informed evaluation of the competing procedures.

6.2 Future Research

This research focused on the utilization of Kriging and RBFNN metamodels. However, there are numerous other non-linear modeling techniques. There is an obvious opportunity to extend this research to examine the effect other metamodeling techniques, such as spline regression or other neural networks, have on the overall robust solution quality.

A second opportunity for research focuses on the generalization of radial basis function neural networks. If too many neurons are used, then the overall generalization of the network will be deficient. On the other hand, if too few neurons are used, the network will not be able to sufficiently learn the training data. This research employed a cross-validation procedure to determine the structure of the RBFNN so that the resulting function was well-generalized with minimal risk of over-fitting the experimental data. Opportunities exist for further heuristic development that identifies a "robust neural network structure.

A third area of future research is concerned with how critics view ANNs in general. This dissertation showed that, by using RBFNNs, a robust solution could be generated that is as good as, and in some cases significantly better than, those produced from strategies using standard polynomial regression models. However, critics of neural networks typically have points of view similar to that of Myers and Montgomery:

> Our view is that neural networks are a complement to the familiar statistical tools of regression analysis, RSM, and designed experiments, but certainly not a replacement for them, because a neural network can at best only give a prediction model and not fundamental insight into the underlying process mechanism that produced the data. [8]

Although there is already abundant literature on identifying salient features using ANNs, perhaps there is some potential for translating the weights and interconnected neurons of an ANN to the standard regression coefficients that the experimental community is most familiar with.

Appendix A. Mean and Variance Derivations for a Function of a Random Variable

A.1 Univariate Case

Let $Y = f(X)$ be a function of the normally distributed random variable X where

$E[X] = \mu_X$ and $Var(X) = \sigma_X^2$. Also, let $\varepsilon \sim N\left(0, \sigma_\varepsilon^2\right)$. Finally, X is independent of ε. If

$$Y = f(X) = f(\mu_X) + f'(\mu_X)(X - \mu_X) + \tfrac{1}{2} f''(\mu_X)(X - \mu_X)^2 + \varepsilon \tag{79}$$

is a second-order Taylor series approximation of Y centered at the point $a = \mu_X$, then an

estimate for the mean of Y is

$$
\begin{aligned}
E[Y] &= E\left[f(\mu_X) + f'(\mu_X)(X - \mu_X) + \tfrac{1}{2} f''(\mu_X)(X - \mu_X)^2 + \varepsilon \right] \\[2mm]
&= f(\mu_X) + f'(\mu_X)E\left[X - \mu_X\right] + \tfrac{1}{2} f''(\mu_X)E\left[(X - \mu_X)^2\right] + E[\varepsilon] \\[2mm]
&= f(\mu_X) + \tfrac{1}{2} f''(\mu_X)\sigma_X^2
\end{aligned}
\tag{80}
$$

Similarly, an estimate for the variance of Y is

$$
\begin{aligned}
Var[Y] &= Var\left[f(\mu_X) + f'(\mu_X)(X - \mu_X) + \tfrac{1}{2} f''(\mu_X)(X - \mu_X)^2 + \varepsilon \right] \\[2mm]
&= Var\left[f'(\mu_X)(X - \mu_X) + \tfrac{1}{2} f''(\mu_X)(X - \mu_X)^2 \right] + Var[\varepsilon] \\[2mm]
&= E\left[\left(f'(\mu_X)(X - \mu_X) + \tfrac{1}{2} f''(\mu_X)(X - \mu_X)^2 \right)^2 \right] \\
&\quad - E\left[f'(\mu_X)(X - \mu_X) + \tfrac{1}{2} f''(\mu_X)(X - \mu_X)^2 \right]^2 + \sigma_\varepsilon^2 \\[2mm]
&= f'(\mu_X)^2 E\left[(X - \mu_X)^2\right] + f'(\mu_X)f''(\mu_X)E\left[(X - \mu_X)^3\right] \\
&\quad + \tfrac{1}{4} f''(\mu_X)^2 E\left[(X - \mu_X)^4\right] - \left[\tfrac{1}{2} f''(\mu_X)\sigma_X^2 \right]^2 + \sigma_\varepsilon^2 \\[2mm]
&= f'(\mu_X)^2 \sigma_X^2 + \tfrac{1}{2} f''(\mu_X)^2 \sigma_X^4 + \sigma_\varepsilon^2
\end{aligned}
\tag{81}
$$

The following central moments for X were used to derive Equations (80) and (81):

$$E[X - \mu_X] = 0 \tag{82}$$

$$E\left[(X - \mu_X)^2\right] = \sigma_X^2 \tag{83}$$

$$E\left[(X - \mu_X)^3\right] = 0 \tag{84}$$

$$E\left[(X - \mu_X)^4\right] = 3\sigma_X^4 \tag{85}$$

A.2 Multivariate Case

Let $Y = f(\mathbf{X})$ be a function of the normally distributed random vector $\mathbf{X}$ where

$E[\mathbf{X}] = \boldsymbol{\mu}_\mathbf{X}$ and $Var(\mathbf{X}) = \Sigma_\mathbf{X}$. Also, let $\varepsilon \sim N\left(0, \sigma_\varepsilon^2\right)$. Finally, $\mathbf{X}$ is independent of ε. If

$$Y = f(\mathbf{X}) = f(\boldsymbol{\mu}_\mathbf{X}) + \nabla f(\boldsymbol{\mu}_\mathbf{X})'(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) + \tfrac{1}{2}(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})'\mathbf{H}(\boldsymbol{\mu}_\mathbf{X})(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) + \varepsilon \tag{86}$$

is a second-order Taylor series approximation of Y centered at the vector $\mathbf{a} = \boldsymbol{\mu}_\mathbf{X}$, then an

estimate for the mean of Y is

$$E[Y] = E\left[f(\boldsymbol{\mu}_\mathbf{X}) + \nabla f(\boldsymbol{\mu}_\mathbf{X})'(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) + \tfrac{1}{2}(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})'\mathbf{H}(\boldsymbol{\mu}_\mathbf{X})(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}) + \varepsilon\right]$$

$$= f(\boldsymbol{\mu}_\mathbf{X}) + \nabla f(\boldsymbol{\mu}_\mathbf{X})'E\left[\mathbf{X} - \boldsymbol{\mu}_\mathbf{X}\right] + \tfrac{1}{2}E\left[(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})'\mathbf{H}(\boldsymbol{\mu}_\mathbf{X})(\mathbf{X} - \boldsymbol{\mu}_\mathbf{X})\right] + E[\varepsilon] \tag{87}$$

$$= f(\boldsymbol{\mu}_\mathbf{X}) + \tfrac{1}{2}tr\left(\mathbf{H}(\boldsymbol{\mu}_\mathbf{X})\Sigma_\mathbf{X}\right)$$

Similarly, an estimate for the variance of Y is

$$Var[Y] = Var\left[f(\boldsymbol{\mu_X}) + \nabla f(\boldsymbol{\mu_X})'(\mathbf{X}-\boldsymbol{\mu_X}) + \tfrac{1}{2}(\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{H}(\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X}) + \varepsilon \right]$$

$$= Var\left[\nabla f(\boldsymbol{\mu_X})'(\mathbf{X}-\boldsymbol{\mu_X}) + \tfrac{1}{2}(\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{H}(\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X}) \right] + Var[\varepsilon]$$

$$= E\left[\left(\nabla f(\boldsymbol{\mu_X})'(\mathbf{X}-\boldsymbol{\mu_X}) + \tfrac{1}{2}(\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{H}(\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X}) \right)^2 \right]$$
$$- E\left[\nabla f(\boldsymbol{\mu_X})'(\mathbf{X}-\boldsymbol{\mu_X}) + \tfrac{1}{2}(\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{H}(\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X}) \right]^2 + \sigma_\varepsilon^2$$

$$= E\big[\nabla f(\boldsymbol{\mu_X})'(\mathbf{X}-\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X})'\nabla f(\boldsymbol{\mu_X})$$
$$+ \nabla f(\boldsymbol{\mu_X})'(\mathbf{X}-\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{H}(\boldsymbol{\mu_X})'(\mathbf{X}-\boldsymbol{\mu_X})$$
$$+ \tfrac{1}{4}(\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{H}(\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{H}(\boldsymbol{\mu_X})'(\mathbf{X}-\boldsymbol{\mu_X})\big]$$
$$- \left[\tfrac{1}{2} tr\left(\mathbf{H}(\boldsymbol{\mu_X})\Sigma_\mathbf{X} \right) \right]^2 + \sigma_\varepsilon^2$$

$$= \nabla f(\boldsymbol{\mu_X})'E\left[(\mathbf{X}-\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X})' \right]\nabla f(\boldsymbol{\mu_X})$$
$$+ \nabla f(\boldsymbol{\mu_X})'E\left[(\mathbf{X}-\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{H}(\boldsymbol{\mu_X})'(\mathbf{X}-\boldsymbol{\mu_X}) \right]$$
$$+ \tfrac{1}{4}E\left[\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{H}(\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X})(\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{H}(\boldsymbol{\mu_X})'(\mathbf{X}-\boldsymbol{\mu_X}) \right]$$
$$- \tfrac{1}{4}tr\left(\mathbf{H}(\boldsymbol{\mu_X})\Sigma_\mathbf{X} \right)^2 + \sigma_\varepsilon^2$$

$$= \nabla f(\boldsymbol{\mu_X})'\Sigma_\mathbf{X}\nabla f(\boldsymbol{\mu_X}) + \tfrac{1}{4}\left[2tr\left(\mathbf{H}(\boldsymbol{\mu_X})\Sigma_\mathbf{X}\mathbf{H}(\boldsymbol{\mu_X})\Sigma_\mathbf{X} \right) + tr\left(\mathbf{H}(\boldsymbol{\mu_X})\Sigma_\mathbf{X} \right)^2 \right]$$
$$- \tfrac{1}{4}tr\left(\mathbf{H}(\boldsymbol{\mu_X})\Sigma_\mathbf{X} \right)^2 + \sigma_\varepsilon^2$$

$$= \nabla f(\boldsymbol{\mu_X})'\Sigma_\mathbf{X}\nabla f(\boldsymbol{\mu_X}) + \tfrac{1}{2}tr\left(\mathbf{H}(\boldsymbol{\mu_X})\Sigma_\mathbf{X}\mathbf{H}(\boldsymbol{\mu_X})\Sigma_\mathbf{X} \right) + \sigma_\varepsilon^2$$

$$(88)$$

The following central moments for $\mathbf{X}$, obtained from Mathai and Provost [85] and Brookes [86], were used to derive Equations (87) and (88). If $\mathbf{A}$ is a symmetric matrix, then:

$$E\left[\mathbf{X}-\boldsymbol{\mu_X} \right] = \mathbf{0} \tag{89}$$

$$E\left[(\mathbf{X}-\boldsymbol{\mu_X})'\mathbf{A}(\mathbf{X}-\boldsymbol{\mu_X}) \right] = tr\left(\mathbf{A}\Sigma_\mathbf{X} \right) \tag{90}$$

$$E\left[(\mathbf{X}-\boldsymbol{\mu_X})\mathbf{A}(\mathbf{X}-\boldsymbol{\mu_X})' \right] = \mathbf{A}\Sigma_\mathbf{X} \tag{91}$$

97

$$E\left[(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})'\mathbf{A}(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})\right] = E\left[(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})'\mathbf{A}'(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})'\right]' = \mathbf{0} \qquad (92)$$

$$E\left[(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})'\mathbf{A}(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})'\mathbf{A}(\mathbf{X}-\boldsymbol{\mu}_\mathbf{X})\right] = 2tr\left(\mathbf{A}\Sigma_\mathbf{X}\mathbf{A}\Sigma_\mathbf{X}\right) + tr\left(\mathbf{A}\Sigma_\mathbf{X}\right)^2 \qquad (93)$$

Appendix B. Mean and Variance Models for the Automotive Maintenance and Repair Shop Simulation

The dots in the figures represent approximations of the Automotive Maintenance and Repair Shop (AMRS) simulation's true response mean or variance.
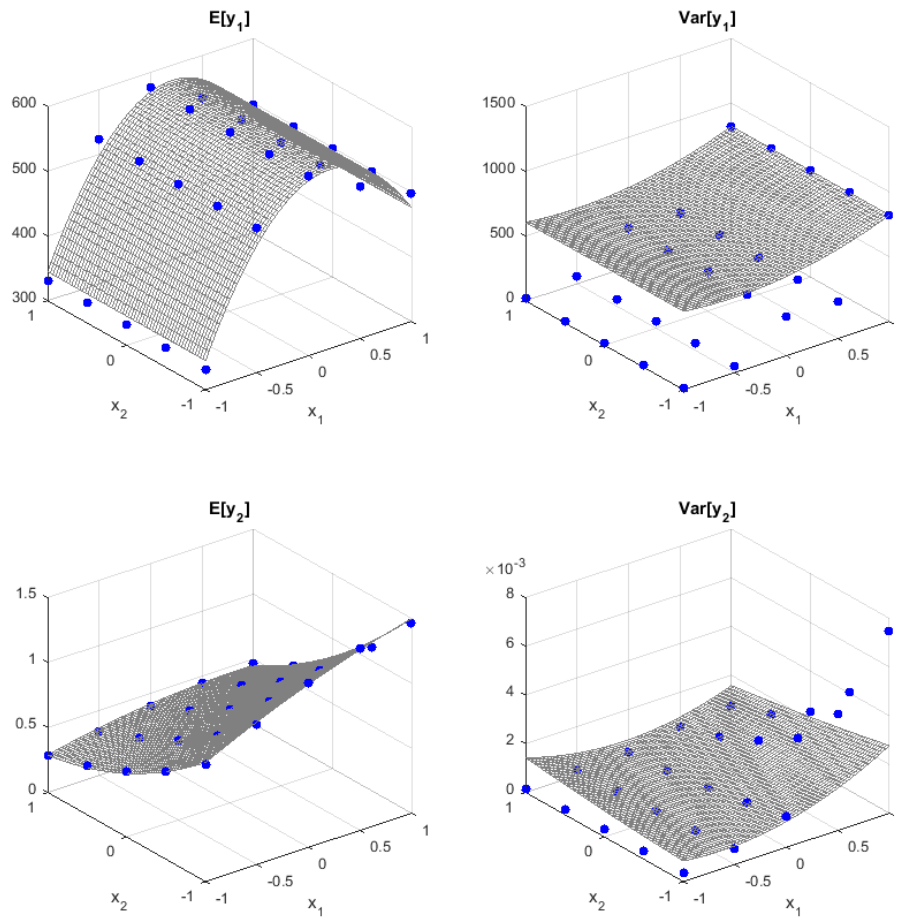
B.1 *Std* Models

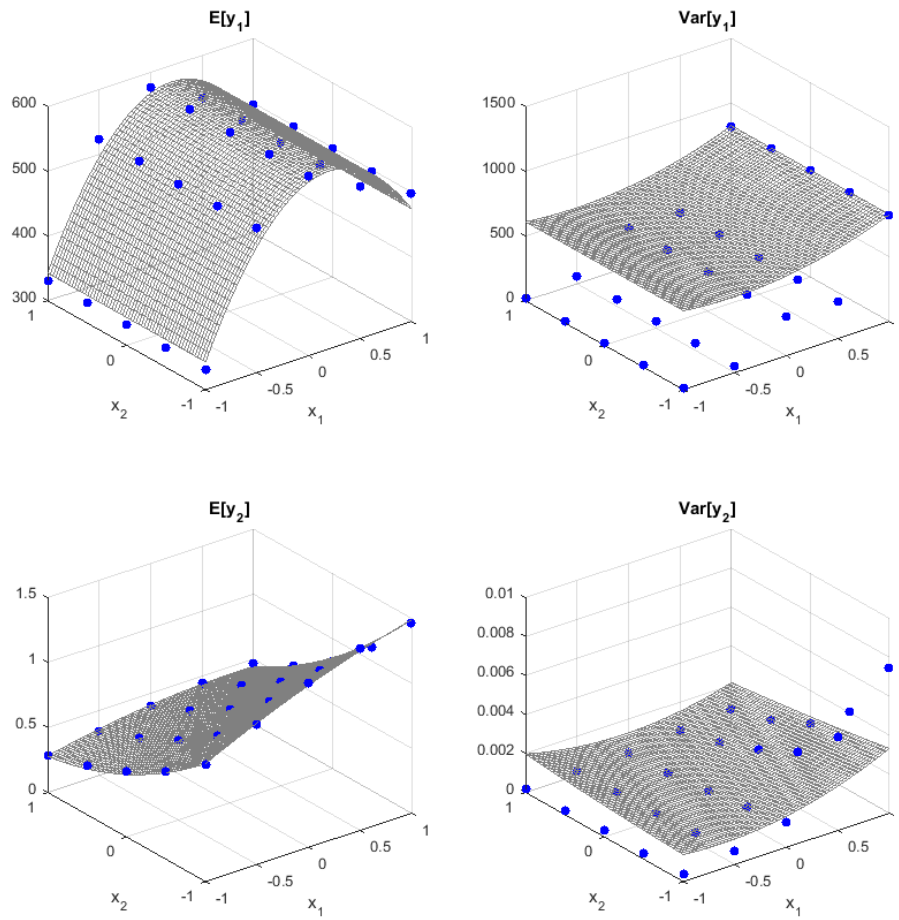**Figure 21. Mean and Variance of the *Std* Model for the AMRS Simulation**

B.2 *NN* **Models**

Figure 22. Mean and Variance of the *NN* Models for the AMRS Simulation
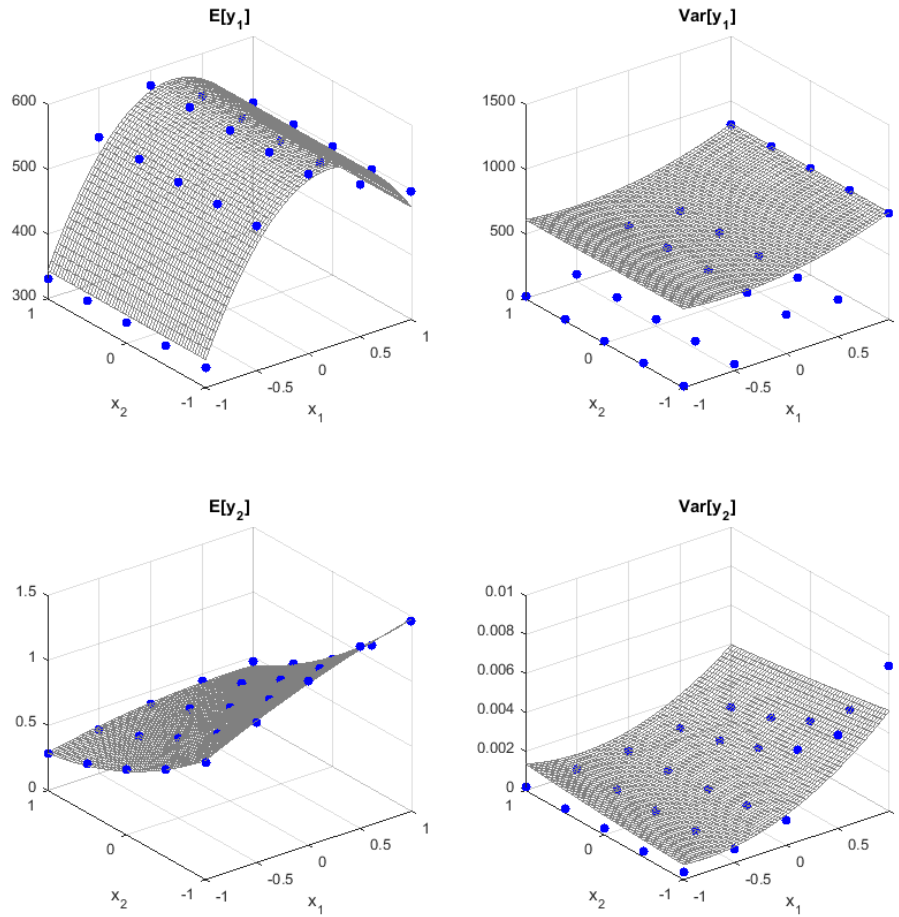
B.3 *CNN* Models

Figure 23. Mean and Variance of the *CNN* Models for the AMRS Simulation
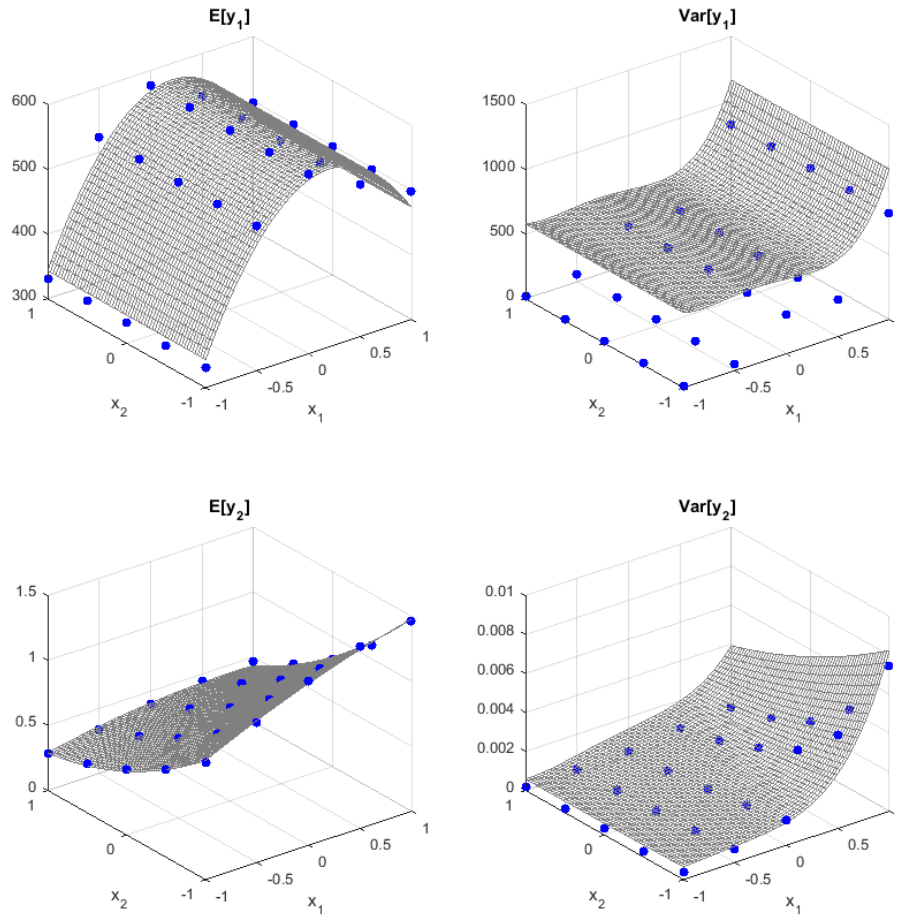
B.4 *CCN* Models

Figure 24. Mean and Variance of the *CCN* Models for the AMRS Simulation
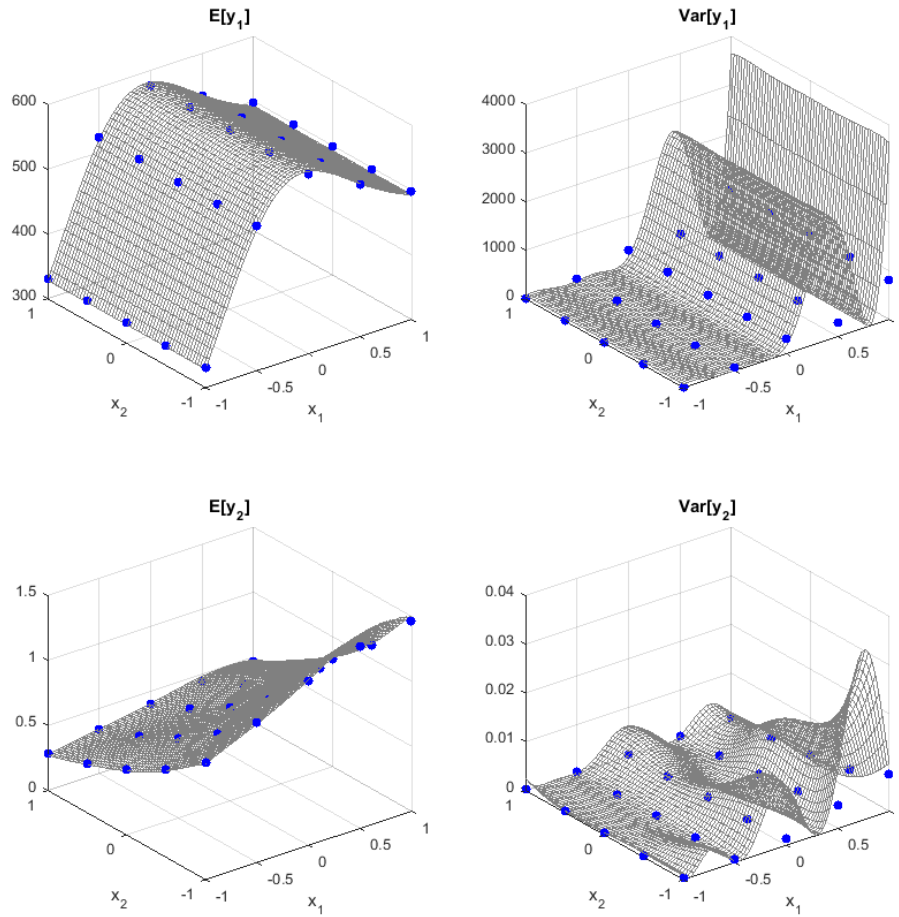
B.5 *KR* Models

Figure 25. Mean and Variance of the *KR* Models for the AMRS Simulation
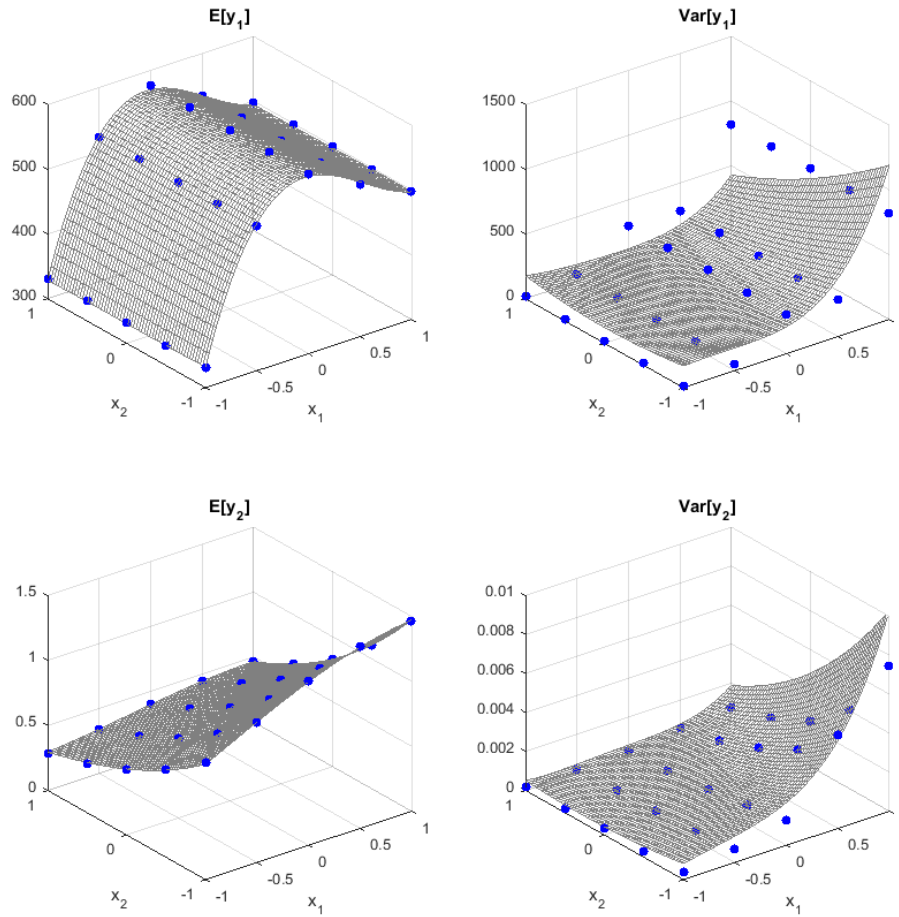
B.6 *RBF* Models

Figure 26. Mean and Variance of the *RBF* Models for the AMRS Simulation

Bibliography

[1]   J. Banks, J. S. Carson, B. L. Nelson and D. M. Nicol, Discrete Event Simulation, Upper Saddle River, NJ: Pearson Prentice Hall, 2005.

[2]   J. P. C. Kleijnen, "Kriging metamodeling in simulation: a review," *European Journal of Operational Research,* vol. 192, no. 3, pp. 707-716, 2009.

[3]   J. P. C. Kleijnen, S. M. Sanchez, T. W. Lucas and T. M. Cioppa, "State-of-the-art review: a user's guide to the brave new world of designing simulation experiments," *INFORMS Journal on Computing,* vol. 17, no. 3, pp. 263-289, 2005.

[4]   G. Taguchi, Introduction to Quality Engineering: Designing Quality Into Products and Processes, White Plains, NY: UNIPUB/Kraus International, 1986.

[5]   W. Welch, T.-K. Yu, S. M. Kang and J. Sacks, "Computer experiments for quality control by parameter design," *Journal of Quality Technology,* vol. 22, no. 1, pp. 15-22, 1990.

[6]   G. G. Vining and R. H. Myers, "Combining Taguchi and response surface philosophies: a dual response approach," *Journal of Quality Technology,* vol. 22, no. 1, pp. 38-45, 1990.

[7]   R. H. Myers, A. I. Khuri and G. Vining, "Response surface alternatives to the Taguchi robust rarameter design approach," *The American Statistician,* vol. 46, no. 2, pp. 131-139, 1992.

[8]   R. H. Myers and D. C. Montgomery, Response Surface Methodology: Process and Product Optimization Using Designed Experiments (2nd Edition), New York, NY: John Wiley & Sons, Inc., 2002.

[9]   D. C. Montgomery, Design and Analysis of Experiments (6th Edition), Hoboken, NJ: John Wiley & Sons, Inc., 2005.

[10] R. Jin, X. Du and W. Chen, "The use of metamodeling techniques for optimization under uncertainty," *Structural and Multidisciplinary Optimization,* vol. 25, no. 2, pp. 99-116, 2003.

[11] R. A. Bates, R. S. Kenett, D. M. Steinberg and H. P. Wynn, "Achieving robust

design from computer simulations," *Quality Technology and Quantitative Management,* vol. 3, no. 2, pp. 161-177, 2006.

[12] G. Dellino, J. P. C. Kleijnen and C. Meloni, "Robust optimization in simulation: Taguchi and response surface methodology," *International Journal of Production Economics,* vol. 125, no. 1, pp. 52-59, 2010.

[13] G. Dellino, J. P. C. Kleijnen and C. Meloni, "Robust optimization in simulation: Taguchi and Krige combined," *INFORMS Journal on Computing,* vol. 24, no. 3, pp. 471-484, 2012.

[14] G. Derringer and D. Suich, "Simultaneous optimization of several response variables," *Journal of Quality Technology,* vol. 12, no. 4, pp. 214-219, 1980.

[15] S. H. Park, "Simultaneous optimization techniques for multi-purpose response functions," *Journal of Military Operations Research Society of Korea,* vol. 7, pp. 118-138, 1981.

[16] E. Del Castillo, D. C. Montgomery and D. R. McCarville, "Modified desirability functions for multiple response optimization," *Journal of Quality Technology,* vol. 28, no. 3, pp. 337-345, 1996.

[17] S. H. Park and J. O. Park, "Simultaneous optimization of multiple responses using a weighted desirability function," in *Quality Improvement Through Statistical Methods*, Boston, MA, Birkhäuser, 1998, pp. 299-311.

[18] J. Wang, Z. He, J. H. Oh and S. H. Park, "Multi-response robust optimization using desirability function," *Advanced Management of Information for Globalized Enterprises, 2008. AMIGE 2008. IEEE Symposium on,* pp. 1-3, 28 Sep 2008.

[19] J. J. Pignatiello, "Strategies for robust multiresponse quality engineering," *IIE Transactions,* vol. 25, no. 3, pp. 5-15, 1993.

[20] A. E. Ames, N. Mattucci, S. MacDonald, G. Szonyi and D. M. Hawkins, "Quality loss functions for optimization across multiple response surfaces," *Journal of Quality Technology,* vol. 29, no. 3, pp. 339-346, 1997.

[21] G. G. Vining, "A compromise approach To multiresponse optimization," *Journal of Quality Technology,* vol. 30, no. 4, pp. 309-313, 1998.

[22] D. Romano, M. Varetto and G. Vicario, "Multiresponse robust design: a general framework based on combined array," *Journal of Quality Technology,* vol. 36, no. 1, pp. 27-37, 2004.

[23] Y.-H. Ko, K.-J. Kim and C.-H. Jun, "A new loss function-based method for multiresponse optimization," *Journal of Quality Technology,* vol. 37, no. 1, pp. 50-59, 2005.

[24] C.-T. Su and L.-I. Tong, "Multi-response robust design By principal component analysis," *Total Quality Management,* vol. 8, no. 6, pp. 409-416, 1997.

[25] A. Salmasnia, R. B. Kazemzadeh and S. T. A. Niaki, "An approach to optimize correlated multiple responses using principal component analysis and desirability function," *International Journal of Advanced Manufacturing Technology,* vol. 62, no. 5, pp. 835-846, 2012.

[26] A. P. Paiva, P. H. Campos, J. R. Ferriera, L. G. D. Lopes, E. J. Paiva and P. P. Balestrassi, "A multivariate robust parameter design approach for optimization of AISI 52100 hardened steel turning with wiper mixed ceramic tool," *International Journal of Refractory Metals and Hard Materials,* vol. 30, no. 1, pp. 152-163, 2012.

[27] J. H. F. Gomes, A. P. Paiva, S. C. Costa, P. P. Balestrassi and E. J. Paiva, "Weighted multivariate mean square error for process optimization: a case study of flux-cored arc welding for stainless steel claddings," *European Journal of Operational Research,* vol. 226, no. 3, pp. 522-535, 2013.

[28] A. I. Khuri and M. Conlon, "Simultaneous optimization of multiple responses represented by polynomial regression functions," *Technometrics,* vol. 23, no. 4, pp. 363-375, 1981.

[29] S. M. Govindaluri and B. R. Cho, "Robust design modeling with correlated quality characteristics using a multicriteria decision framework," *International Journal of Advanced Manufacturing Technology,* vol. 32, no. 5-6, pp. 423-433, 2007.

[30] O. Koksoy and T. Yalcinoz, "Mean square error criteria to multiresponse process optimization by a new genetic algorithm," *Applied Mathematics and Computation,* vol. 175, no. 2, pp. 1657-1674, 2006.

[31] O. Koksoy, "Multiresponse robust design: mean square error (MSE) criterion," *Applied Mathematics and Computation,* vol. 175, no. 2, pp. 1716-1729, 2006.

[32] O. Koksoy, "A nonlinear programming solution to robust multi-response quality problems," *Applied Mathematics and Computation,* vol. 196, no. 2, pp. 603-612, 2008.

[33] A. Shaibu and B. R. Cho, "Another view of dual response surface modeling and optimization in robust parameter design," *International Journal of Advanced Manufacturing Technology,* vol. 41, no. 7-8, pp. 631-641, 2009.

[34] C. F. J. Wu and M. Hamada, Experiments: Planning, Analysis, and Parameter Design Optimization, New York, NY: John Wiley & Sons, Inc., 2000.

[35] D. C. Montgomery, E. A. Peck and G. G. Vining, Introduction to Linear Regression Analysis (3rd Edition), New York, NY: John Wiley & Sons, Inc., 2001.

[36] R. V. Leon, A. C. Shoemaker and R. N. Kackar, "Performance measures independent of adjustment: an explanation and extension of Taguchi's signal-to-noise ratios," *Technometrics,* vol. 29, no. 3, pp. 253-265, 1987.

[37] V. N. Nair and A. C. Shoemaker, "The role of experimentation in quality engineering: a review of Taguchi's contribution," in *Statistical Design and Analysis of Industrial Experiments*, New York, NY, Marcel Dekker, Inc., 1990, pp. 247-277.

[38] G. E. P. Box, "Signal to noise ratios, performance criteria and transformations," *Technometrics,* vol. 30, no. 1, pp. 1-17, 1988.

[39] R. N. Kackar, "Taguchi's quality philosophy: analysis and commentary," *Quality Progress,* vol. 19, no. 12, pp. 21-29, 1986.

[40] J. J. Pignatiello and J. S. Ramberg, "Discussion: off-line quality control, parameter design, and the Taguchi method," *Journal of Quality Technology,* vol. 17, no. 4, pp. 198-206, 1985.

[41] J. J. Pignatiello, "An overview of the strategy and tactics of Taguchi," *IIE Transactions,* vol. 20, no. 3, pp. 247-254, 1988.

[42] A. Giovagnoli and D. Romano, "Robust design via simulation experiments: a modified dual response surface approach," *Quality and Reliability via Simulation Experiments: A Modified Dual Response Surface Approach,* vol. 24, no. 4, pp. 401-416, 2008.

[43] K. Elsayed and C. Lacor, "Robust parameter design optimization using Kriging, RBF, and RBFNN with gradient-based and evolutionary optimization techniques," *Applied Mathematics and Computation,* vol. 236, pp. 325-344, 2014.

[44] R. H. Wild and J. J. Pignatiello, "An experimental design strategy for designing robust systems using discrete-event simulation," *Simulation,* vol. 57, no. 6, pp. 358-368, 1991.

[45] G. E. P. Box and N. R. Draper, Empirical model-building and response surfaces, New York, NY: John Wiley and Sons, Inc, 1987.

[46] D. K. J. Lin and W. Tu, "Dual response surface optimization," *Journal of Quality Technology,* vol. 27, no. 1, pp. 34-39, 1995.

[47] I.-J. Jeong, K.-J. Kim and S. Y. Chang, "Optimal weighting of bias and variance in dual response surface optimization," *Journal of Quality Technology,* vol. 37, no. 3, pp. 236-247, 2005.

[48] O. Koksoy and T. Yalcinoz, "Robust design using pareto type optimization: a genetic algorithm with arithmetic crossover," *Computers & Industrial Engineering,* vol. 55, no. 1, pp. 208-218, 2008.

[49] F. M. Mindrup, K. W. Bauer and M. A. Friend, "Extending robust parameter design to noise by noise interactions with an application to hyperspectral imagery," *International Journal of Quality Engineering and Technology,* vol. 3, no. 1, pp. 1-19, 2012.

[50] J. P. Williams, K. W. Bauer and M. A. Friend, "Further extensions to robust parameter design: three factor interactions with an application to hyperspectral imagery," *International Journal of Quality Engineering and Technology,* vol. 3, no. 3, pp. 204-218, 2013.

[51] W. C. M. van Beers and J. P. C. Kleijnen, "Kriging interpolation in simulation: a survey," in *Proceedings of the 2004 Winter Simulation Conference*, Washington, D.C., 2004.

[52] J. Sacks, W. J. Welch, T. J. Mitchell and H. P. Wynn, "Design and analysis of computer experiments," *Statistical Science,* vol. 4, no. 4, pp. 409-423, 1989.

[53] T. J. Santner, B. J. Williams and W. I. Notz, The Design and Analysis of Computer

Experiments, New York, NY: Springer Publishing Company, Inc., 2003.

[54] J. D. Martin and T. W. Simpson, "Use of Kriging models to approximate deterministic computer models," *AIAA Journal,* vol. 43, no. 4, pp. 853-863, 2005.

[55] B. Ankenman, B. L. Nelson and J. Staum, "Stochastic Kriging for simulation metamodeling," in *Proceedings of the 2008 Winter Simulation Conference*, Miami, FL, 2008.

[56] T. W. Simpson, P. J. D, P. N. Koch and J. K. Allen, "Metamodels for computer-based engineering design: survey and recommendations," *Engineering with Computers,* vol. 17, no. 2, pp. 129-150, 2001.

[57] R. R. Barton, "Metamodels for simulation input-output relations," in *Proceedings of the 24th Conference on Winter Simulation*, Arlington, VA, 1992.

[58] J. Staum, "Better simulation metamodeling: the why, what, and how of stochastic Kriging," in *Proceedings of the 2009 Winter Simulation Conference*, Austin, TX, 2009.

[59] N. A. C. Cressie, Statistics for Spatial Data, New York: Wiley, 1993.

[60] J. D. Martin and T. W. Simpson, "On using Kriging models as probabilistic models in design," *SAE Transactions Journal of Materials & Manufacturing,* vol. 5, pp. 129-139, 2004.

[61] R. O. Duda, P. E. Hart and D. G. Stork, Pattern Classification (2nd Edition), New York, NY: John Wiley & Sons, Inc., 2001.

[62] P. D. Wasserman, Advanced Methods in Neural Computing, New York, NY: Van Nostrand Reinhold, 1993.

[63] S. Haykin, Neural Networks: A Comprehensive Foundation, New York, NY: Macmillan College Publishing Company, Inc., 1994.

[64] C. G. Looney, Pattern Recognition Using Neural Networks: Theory and Algorithms for Engineers and Scientists, New York, NY: Oxford University Press, Inc., 1997.

[65] T. J. Robinson, C. M. Borror and R. H. Myers, "Robust parameter design: a review," *Quality and Reliability Engineering International,* vol. 20, no. 1, pp. 81-101, 2004.

[66] K. A. F. Copeland and P. R. Nelson, "Dual response optimization via direct function minimization," *Journal of Quality Technology,* vol. 28, no. 3, pp. 331-336, 1996.

[67] R. Ding, D. K. Lin and D. Wei, "Dual-response surface optimization: a weighted MSE approach," *Quality Engineering,* vol. 16, no. 3, pp. 377-385, 2004.

[68] O. Koksoy and N. Doganaksoy, "Joint optimization of mean and standard deviation using response surface methods," *Journal of Quality Technology,* vol. 35, no. 3, pp. 239-252, 2003.

[69] E. C. Harrington, "The desirability function," *Industrial Quality Control,* vol. 21, no. 10, pp. 494-498, 1965.

[70] C. Auer, M. Erdbrugge and R. Gobel, "Comparison of multivariate methods for robust parameter design in sheet metal," *Applied Stochastic Models in Business and Industry,* vol. 20, no. 3, pp. 201-218, 2004.

[71] C.-H. Chiao and M. Hamada, "Analyzing experiments with correlated multiple responses," *Journal of Quality Technology,* vol. 33, no. 4, pp. 451-465, 2001.

[72] J. M. Ver Hoef, "Who invented the Delta method?," *The American Statistician,* vol. 66, no. 2, pp. 124-127, 2012.

[73] S. C. Chapra and D. M. Di Toro, "Delta method for estimating primary production, respiration, and reaeration in streams," *Journal of Environmental Engineering,* vol. 117, no. 5, pp. 640-655, 1991.

[74] B. P. Durbin, J. S. Hardin, D. M. Hawkins and D. M. Rocke, "A variance-stabilizing transformation for gene-expression microaray data," *Bioinformatics,* vol. 18, no. suppl 1, pp. S105-S110, 2002.

[75] L. A. Powell, "Approximating variance of demographic parameters using the Delta method: a reference for avian biologists," *The Condor,* vol. 109, no. 4, pp. 949-954, 2007.

[76] E. G. Cooch and G. C. White, "Program MARK: a gentle introduction (13th Edition)," 13 September 2013. [Online]. Available: http://www.fwspubs.org/doi/suppl/10.3996/122012-JFWM-110R1/suppl_file/10.3996_122012-jfwm-110r1.s8.pdf. [Accessed 12 January 2015].

[77] J. Stewart, Calculus (3rd Edition), Pacific Grove, CA: Brooks/Cole Publishing Company, 1995.

[78] G. Casella and R. L. Berger, Statistical Inference (2nd Edition), Pacific Grove, CA: Duxbury, 2002.

[79] M. S. Bazaraa, H. D. Sherali and C. M. Shetty, Nonlinear Programming: Theory and Algorithms (3rd Edition), Hoboken, NJ: John Wiley & Sons, Inc., 2006.

[80] T. A. Feo and M. G. C. Resende, "Greedy randomized adaptive search procedures," *Journal of Global Optimization,* vol. 6, no. 2, pp. 109-134, 1995.

[81] The Mathworks, "Optimization Toolbox: User's Guide (Version 2)," The Mathworks, Inc., Natick, MA, 2003.

[82] R. S. Kenett and S. Zacks, Modern Industrial Statistics: Design and Control of Quality and Reliability, Pacific Grove, CA: Duxbury Press, 1998.

[83] R. R. Barton, "Simulation optimization using metamodels," in *Winter Simulation Conference*, Austin, TX, 2009.

[84] W. D. Kelton, R. P. Sadowski and D. T. Sturrock, Simulation with Arena (3rd ed.), New York, NY: McGraw-Hill, 2004.

[85] A. M. Mathai and S. B. Provost, Quadratic Forms in Random Variables: Theory and Applications, New York, New York: Marcel Dekker, Inc., 1992.

[86] M. Brookes, "The Matrix Reference Manual: Stochastic Matrices," 11 Aug 2015. [Online]. Available: http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/expect.html. [Accessed 25 Nov 2015].

Vita

Major Joseph P. Bellucci graduated from Spring Hill High School in Columbia, Tennessee in May 1997. He entered undergraduate studies at Cumberland University in Lebanon, Tennessee where he graduated with Bachelor of Science degrees in Mathematics and Physics in May 2001. He earned his commission through the Air Force Officer Training School at Maxwell AFB Alabama on 28 June 2002.

Maj Bellucci's first assignment was to the Space Vehicles Directorate, Air Force Research Laboratory at Hanscom AFB Massachusetts where he served as a hyperspectral data analysis manager and acquisitions officer. From August 2005 to March 2007, he earned his Master of Science degree in Operations Research from the Graduate School of Engineering and Management, Air Force Institute of Technology (AFIT), Wright-Patterson AFB Ohio. In April 2007, he was assigned to the Analysis, Lessons Learned, and AFSO21 Directorate, Air Combat Command at Langley AFB Virginia. In March 2010, Maj Bellucci deployed as an operations research analyst to Camp Victory, Iraq in support of Operation Iraqi Freedom. In February 2011, he was assigned to Headquarters, Air Force Recruiting Service at Randolph AFB Texas where he served as the Operations Division, Analysis Branch Chief. In August 2013, he returned to AFIT for a PhD in Operations Research. Upon graduation, Maj Bellucci will serve as the Chief Scientist at the Air Education and Training Command Studies and Analysis Squadron at Randolph AFB Texas.

REPORT DOCUMENTATION PAGE

Public reporting burden for this 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 this 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 Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 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 any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)* 15-09-2016	2. REPORT TYPE Doctoral Dissertation	3. DATES COVERED *(From - To)* Aug 2013 – Sept 2016

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Non-Linear Metamodeling Extensions to the Robust Parameter Design of Computer Simulations	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Bellucci, Joseph P., Major, USAF	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way Wright-Patterson AFB OH 45433-7765	AFIT-ENS-DS-16-S-026

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) AF/A9FC
Maj Todd Paciencia, PhD Headquarters Air Force, Studies, Assessments, and Analyses, Campaign Analyses Division 1570 Air Force Pentagon Washington D.C. 20330-1570	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Statement A. Approved for Public Release; Distribution Unlimited.

13. SUPPLEMENTARY NOTES
This material is declared a work of the U. S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT
    Robust parameter design (RPD) is used to identify a system's control settings that offer a compromise between obtaining desired mean responses and minimizing the variability about those responses. Two popular combined-array strategies—the response surface model (RSM) approach and the emulator approach—are limited when applied to simulations. In the former case, the mean and variance models can be inadequate due to a high level of non-linearity within many simulations. In the latter case, precise mean and variance approximations are developed at the expense of extensive Monte Carlo sampling.

    This research combines the RSM approach's efficiency with the emulator approach's accuracy. Non-linear metamodeling extensions, namely through Kriging and radial basis function neural networks, are made to the RSM approach. The mean and variance of second-order Taylor series approximations of these metamodels are generated via the Multivariate Delta Method and subsequent optimization problems employing these approximations are solved. Results show that improved prediction models can be attained through the proposed approach at a reduced computational cost. Additionally, a multi-response RPD problem solving technique based on desirability functions is presented to produce a solution that is mutually robust across all responses. Lastly, quality measures are developed to provide a holistic assessment of several competing RPD strategies.

15. SUBJECT TERMS
Robust Parameter Design; Delta Method; Radial Basis Function Neural Networks; Kriging; Desirability Functions

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Dr. Kenneth W. Bauer, Jr., AFIT/ENS
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	128	19b. TELEPHONE NUMBER *(include area code)* (937) 255-3636, x4328 kenneth.bauer@afit.edu