NAVAL POSTGRADUATE SCHOOL
Monterey, California

DISSERTATION

EFFICIENT NEARLY ORTHOGONAL AND SPACE-FILLING EXPERIMENTAL DESIGNS FOR HIGH-DIMENSIONAL COMPLEX MODELS

by

Thomas M. Cioppa

September 2002

Dissertation Supervisor: Thomas W. Lucas

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2002	**3. REPORT TYPE AND DATES COVERED** Dissertation

4. TITLE AND SUBTITLE: Title (Mix case letters) Efficient Nearly Orthogonal and Space-Filling Experimental for High-Dimensional Complex Models	**5. FUNDING NUMBERS**
6. AUTHOR(S) Cioppa, Thomas M.	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	**12b. DISTRIBUTION CODE**

13. ABSTRACT *(maximum 200 words)*

The Department of Defense uses complex high-dimensional simulation models as an important tool in its decision-making process. To improve on the ability to efficiently explore larger subspaces of these models, this dissertation develops a set of experimental designs for searching over as many as 22 variables in as few as 129 runs. These new designs combine orthogonal Latin hypercubes and uniform designs to create designs having near orthogonality and excellent space-filling properties. Multiple measures are used to assess the quality of candidate designs and to identify the best one. For situations in which more than the minimum number of required runs are available, the designs can be permuted and appended to create additional design points that improve upon the design's orthogonality and space-filling.

The designs are used to explore two surfaces. For a known 11 dimensional stochastic response function containing nonlinear and interaction terms, it is shown that the near orthogonal Latin hypercube is substantially better than the orthogonal Latin hypercube in estimating model coefficients. The other exploration uses the agent-based simulation MANA to analyze 22 variables in a complex military peace enforcement operation. The need for maintaining the initiative and speed of execution during these peace enforcement operations is identified.

14. SUBJECT TERMS Experimental Design, Latin Hypercube, Uniform Design, Agent-Based Simulation, Military Peace Enforcement Operations	**15. NUMBER OF PAGES** 146
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited.

EFFICIENT NEARLY ORTHOGONAL AND SPACE-FILLING EXPERIMENTAL DESIGNS FOR HIGH-DIMENSIONAL COMPLEX MODELS

Thomas M. Cioppa
Lieutenant Colonel, United States Army
B.S., United States Military Academy, 1985
M.S., Naval Postgraduate School, 1993

Submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 2002

Author: ____
Thomas M. Cioppa

Approved by:

Thomas W. Lucas Associate Professor of Operations Research Dissertation Supervisor	Robert R. Read Professor of Operations Research Dissertation Committee Chair
Gordon H. Bradley Professor of Operations Research	Eugene P. Paulo Assistant Professor of Operations Research
Craig W. Rasmussen Associate Professor of Mathematics	Susan M. Sanchez Professor of Operations Research

Approved by: ____
James N. Eagle, Chair, Department of Operations Research

Approved by: ____
Carson K. Eoyang, Associate Provost for Academic Affairs

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The Department of Defense uses complex high-dimensional simulation models as an important tool in its decision-making process. To improve on the ability to efficiently explore larger subspaces of these models, this dissertation develops a set of experimental designs for searching over as many as 22 variables in as few as 129 runs. These new designs combine orthogonal Latin hypercubes and uniform designs to create designs having near orthogonality and excellent space-filling properties. Multiple measures are used to assess the quality of candidate designs and to identify the best one. For situations in which more than the minimum number of required runs are available, the designs can be permuted and appended to create additional design points that improve upon the design's orthogonality and space-filling.

The designs are used to explore two surfaces. For a known 11 dimensional stochastic response function containing nonlinear and interaction terms, it is shown that the near orthogonal Latin hypercube is substantially better than the orthogonal Latin hypercube in estimating model coefficients. The other exploration uses the agent-based simulation MANA to analyze 22 variables in a complex military peace enforcement operation. The need for maintaining the initiative and speed of execution during these peace enforcement operations is identified.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF SYMBOLS, ACRONYMS, AND ABBREVIATIONS

AO	Area of Operation
CAA	Center for Army Analysis
CASTFOREM	Combined Arms and Support Task Force Evaluation Model
DoD	Department of Defense
ER	Exchange Ratio
HQ	Headquarters
ISAAC	Irreducible Semi-Autonomous Adaptive Combat
JWARS	Joint Warfare System Simulation Model
k	Number of Variables
MANA	Map Aware Non-Uniform Automata
ML_2	Modified L_2 Discrepancy
Mm	Maximin
n	Number of Simulation Runs
$(N_O)_k^n$	Nearly Orthogonal Latin Hypercube with n Runs and k Variables
NOLHC	Nearly Orthogonal Latin Hypercube
$(O)_k^n$	Orthogonal Latin Hypercube with n Runs and k Variables
OLHC	Orthogonal Latin Hypercube
q	Number of Positive Levels of a Variable
SVD	Singular Value Decomposition
TRAC	Training and Doctrine Command Analysis Center
UN	United Nations
U.S.	United States
VIC	Vector-In-Commander Model

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The United States Department of Defense uses simulation models to support its decision-making process. Defense analysts need experimental designs capable of efficiently searching an intricate simulation model that has a high-dimensional input space characterized by a complex response surface (substantial non-linearities may be prevalent). To efficiently explore these simulations, the experimental designs should have the following desirable characteristics:

- approximate orthogonality of the input variables,
- space-filling, that is, the collection of experimental cases should be a representative subset of the points in the hypercube of explanatory variables,
- ability to examine many variables (20 or more) efficiently,
- flexibility in analyzing and estimating as many effects, interactions, and thresholds as possible,
- requiring minimal *a priori* assumptions on the response,
- ease in generating the design, and
- ability to gracefully handle premature experiment termination.

This dissertation develops experimental designs, satisfying each of the above characteristics, that provide the ability to search a high-dimensional (up to 22 variables) simulation model and reliably identify critical variables, important interactions, and the ranges of the variables where these effects occur. Furthermore, the number of runs required is small (e.g., a minimum of 129 runs for 22 variables) when compared to most existing experimental designs.

The two most important characteristics for these designs are orthogonality and space-filling. Two measures are used to assess the orthogonality of a design matrix. These measures are the maximum pairwise correlation and singular value decomposition condition number. The use of both measures provides a better ability to differentiate between the orthogonality of candidate designs. We also show how to improve upon the orthogonality of a design matrix.

There are two measures used to assess the space-filling of a design matrix. These measures are the Euclidian maximum minimum distance between design points and, from

uniform design theory, the modified L_2 discrepancy. The use of both measures provides a better ability to differentiate between the space-filling of candidate designs.

The designs are constructed by taking a current algorithm from Ye [1998] that creates orthogonal Latin hypercube designs and expanding on the number of variables that these designs can have. By doing this, one is able to significantly increase the number of variables that can be examined within a fixed number of runs (see Table E.1). While we are able to generate orthogonal Latin hypercubes for more variables, some of the orthogonality is deliberately sacrificed in order to get better space-filling. Designs for up to 22 variables are included in the dissertation, but the algorithm generalizes for an arbitrary number of variables.

Number of experiments	Number of variables examined in the orthogonal or nearly orthogonal designs	Number of variables examined in previous orthogonal designs	Percent increase in number of variables examined
17	7	6	17%
33	11	8	38%
65	16	10	60%
129	22	12	83%

Table E.1. The designs developed in this dissertation are able to examine a greater number of variables than similar previous designs in the same number of runs. These new designs still have excellent orthogonality and space-filling characteristics.

The experimental design for 11 variables is used on a known response function. The design is able to efficiently identify nonlinear terms and interactions in the associated regression equation. The advantages of this design over Latin hypercubes and orthogonal Latin hypercubes are shown.

The experimental design for 22 variables is used to analyze a complex military peace enforcement operation using an agent-based simulation. The subsequent data analysis, coupled with the author's military experience, identifies potential insights that may benefit senior military decision-makers in preparing for future peace enforcement operations. Furthermore, we identify a possible flaw in the agent-based simulation.

Two major United States Army analytical organizations (Center for Army Analysis and Training and Doctrine Command Analytical Center) are using or considering the use of these designs for studies that have multi-billion dollar implications. Furthermore, two Naval Postgraduate School Masters students are using these designs and the peace enforcement scenario in their research.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The goal of this dissertation is to provide new experimental designs that can enable analysts to conduct more thorough investigations of simulation models. A *computer simulation*[1] is a computerized model that attempts to imitate or characterize a real-world problem, scenario, or an abstraction of it. In this dissertation, the terms "simulation model" and "simulation" are used interchangeably. It is also assumed that the analyst can chose, or specify, the input variable values that are used to generate output from the simulation model. For stochastic simulation models, some of these input variables may represent distribution parameters. An *experimental design* is defined as a matrix of input variable[2] values ($\boldsymbol{X}$), where each column of $\boldsymbol{X}$ represents a variable and each row represents the combination of input variable values for a single run.

A. MOTIVATING PROBLEM

The United States (U.S.) Department of Defense (DoD) uses simulation models to support its decision-making process. These models are used to help test war plans against adversaries, decide what equipment to acquire, determine the best combination of forces, determine the best combination and use of weapons, and much more (e.g., Schmidt [1992], Rodgers and Prueitt [1993], Wilmer [1994], Appelget [1995], Barnes and Steffey [1995], Loerch et al. [1996], Shupenas and Armstrong [1998], Posadas [2001]). Since it is nearly impossible to conduct actual physical experiments to determine the effectiveness of war plans, force designs, or weapon system capabilities in actual conflict, the DoD relies on these simulation models to capture significant insights that enable senior leadership to make informed decisions.

Examples of simulation models used by the U.S. Army include the deterministic Vector-In-Commander (VIC) model, the stochastic Combined Arms and Support Task Force Evaluation Model (CASTFOREM), and the stochastic Joint Warfare System (JWARS).[3] VIC, developed by the Training and Doctrine Command Analysis Center

[1]Important terms and concepts will be italicized when they are defined.
[2]Unless otherwise specified, a variable is assumed to be continuous.
[3]Up-to-date information on these and other combat simulation models is available from http://www.dmso.mil/public and http://www.amso.army.mil.

(TRAC) in 1982, serves as the Army's principle Corps-level simulation. CASTFOREM was developed and is principally used by TRAC at White Sands, New Mexico for simulating force-on-force conflict between brigade and smaller forces. The DoD is sponsoring the development of JWARS, which will be a state-of-the art, object-oriented, stochastic, constructive simulation capable of modeling joint, theater-level warfare.

A new and stimulating area of combat models involves complex adaptive systems. The concept is to use multi-agent-based software tools to examine the relationship between numerous input variables and output measures. The self-adaptive nature of these models facilitates broad exploration and permits the possibility of gaining substantial insights into emergent behaviors on the battlefield (Horne and Leonardi [2001]). The major proponent of this current research is the Marine Corps Combat Development Command's Project Albert.[4]

A common characteristic of the above-mentioned models is the vast number (sometimes even greater than 100,000) of variables or data elements present—many of which are uncertain. Conducting a comprehensive experimental design on these numerous variables is prohibitive. Often, a small subset of the variables (usually no more than two or three) is chosen for experimentation. In such a case, the results are necessarily assumed to be invariant to the large number of uncertain variables held constant, but no empirical assessment is made. In addition, even a small, manageable subset does not guarantee that a detailed experimental design will be used. The problem is compounded since even if a manageable subset of input variables is chosen, determining the appropriate levels or settings of the variables remains an issue. Remembering that the main thrust of the experimentation is to identify significant insights, this goal may be jeopardized when a small subset of variables or inappropriate levels of the variables are used.

What is needed by the DoD to analyze simulation models in order to gain significant insights to make better, informed decisions? Defense analysts need experimental designs capable of efficiently searching an intricate simulation model that has a high-dimensional input space, characterized by a complicated response surface

[4] Additional information may be obtained from http://www.projectalbert.org.

(substantial non-linearities may be prevalent). The experimental designs developed in this dissertation provide the ability to search a comparatively high-dimensional (up to 22 variables[5]) subspace of a simulation model and reliably identify critical variables, important interactions, and the ranges of the variables where these effects occur. Furthermore, the number of runs required is small (e.g., a minimum of 129 runs for 22 variables) when compared to most existing experimental designs.

The following quote conveys a frank and simple message. Although, in theory, one may execute an astronomical number of runs, in reality and practicality it cannot be done. Other sound alternatives must be developed. Each of the designs proposed in this dissertation is one of these sound alternatives.

> "Forever" may sound overblown, but any length of time longer than that which we have available to us, because of nature or of orders from our superiors, is effectively forever. This fact has been delightfully dramatized by Major General Jasper Welch in the phrase, 10^{30} is forever. (Hoeber [1981])

B. DEFINITIONS AND TERMINOLOGY

A brief description of important definitions and terminology used in this dissertation is given in this section. Assume that a simulation model contains k input variables and generates a vector of output responses denoted as $\mathbf{y}$. Let the ith variable be denoted as x_i and let y_j be an individual output response from the simulation. To help us understand our simulation models, a metamodel to describe the relationship between the input variables ($x_1, x_2, \ldots, x_k$) and the output measure (y_j) is often used. A *metamodel* is a relatively simple[6] function g that is estimated given an experimental design and the corresponding responses. Mathematically this is modeled as

$$y_j = g(x_1, x_2, ..., x_k) + \varepsilon \, .^{7} \tag{1.1}$$

A good metamodel is one in which g makes parsimonious use of the variables available and the error term (ε) is small. One of the simplest metamodels is one in which g is a linear combination of the inputs. That is,

---

[5] Note: There is no theoretical limit on the number of variables that could be examined by the method developed in this dissertation, provided enough resources are available. However, in this dissertation, only designs for two to 22 variables are constructed.

[6] Here, the metamodel is "simple" when compared to the original simulation model.

[7] Here, an additive-error metamodel is assumed, but other error structures are possible.

$$g = \beta_0 + \sum_{i=1}^{k} \beta_i x_i . \tag{1.2}$$

In order to have sufficient degrees of freedom for estimating the $(k + 1)$ coefficients of (1.2), as well as the error term, the number of runs from the simulation, denoted by n, must satisfy

$$n > k + 1 . \tag{1.3}$$

When estimating the coefficients in (1.2), the precision of the estimates can be adversely affected by multicollinearity (or correlations) among the input variables (Myers [1986]). The *correlation* between two vectors $\boldsymbol{v}=[v_1,v_2,\ldots,v_n]^T$ and $\boldsymbol{w}=[w_1,w_2,\ldots,w_n]^T$, or two columns in a design matrix, is defined to be

$$\frac{\sum_{i=1}^{n}\left[(v_i - \bar{v})(\omega_i - \bar{\omega})\right]}{\sqrt{\sum_{i=1}^{n}(v_i - \bar{v})^2 \sum_{i=1}^{n}(\omega_i - \bar{\omega})^2}} . \tag{1.4}$$

If two columns have zero correlation, they are *orthogonal*. If the columns in the design matrix between input variables $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$ are orthogonal, then the regression estimates of β_i and β_j in (1.2) are uncorrelated. Of course, the two vectors are orthogonal if and only if the numerator of (1.4) is zero. However, the denominator in (1.4) limits the range to between –1 and 1, and allows for meaningful comparisons of the degree of nonorthogonality of pairs of vectors of different lengths (see, e.g., Iman and Conover [1980], Owen [1994], Tang [1998], Ye [1998]).

For many simulations, a linear metamodel may not sufficiently characterize the response surface. Unfortunately, it takes many more observations to estimate metamodels with curvilinear and interaction terms. For example, suppose that g includes quadratic and bilinear interaction effects, as well as the linear terms. That is,

$$g = \beta_0 + \sum_{i=1}^{k} \beta_{i\cdot} x_i + \sum_{j=1}^{k} \beta_{j\cdot} x_j^2 + \sum_{i=1}^{k} \sum_{j>i} \beta_{i,j} x_i x_j . \tag{1.5}$$

In order to have enough degrees of freedom to estimate the coefficients in (1.5) and the error term, the number n of simulation runs must now satisfy

$$n > \left[k + k + \binom{k}{2} + 1 \right]. \tag{1.6}$$

Thus, in this case, the sample size requirements for n grow on the order of k^2. More complicated metamodels require n to be even larger.

To help glean insights about relationships in simulations, an analyst desires experimental designs that allow one to fit a breadth of potential metamodels (perhaps quite complex) within a constrained number of runs, n. An *efficient experimental design* is referred to as one which (i) detects as many significant variables, nonlinear effects, interactions, and their associated ranges as possible, (ii) declares significant as few non-significant variables and interactions as possible, and (iii) accomplishes this with a minimal number of runs. This concept is used in the comparative sense.

A simulation model is considered to be *complex* if one of two conditions is satisfied. The first condition is a high-dimensional input space, defined as 20 or more variables in a model. Thus, in a simulation model, even if only a few variables out of 20 variables turn out to be important, and these important variables can represent the output in an additive fashion, the model will be considered complex. The second condition holds if, regardless of the number of variables, a large number of two-variable and higher interactions exist or the mathematical metamodel is sufficiently non-linear (e.g., the response surface is a high-degree polynomial, contains discontinuities, or has change-points). This encompassing statement permits models containing any number of variables to be considered complex, provided one of the two conditions is present. This allows for the possibility that even if a model only has three or four variables, it can be considered complex if its metamodel is defined by a high-degree polynomial or other complicated non-linear relationship. Examples of complex simulation models are models that simulate combat and include VIC, CASTFOREM, and JWARS.

C. EXPERIMENTAL DESIGNS AND THE ANALYTICAL DILEMMA

This section addresses the trade-offs made by an analyst when using experimental designs to analyze a simulation. Design and analysis are complementary activities. The design must support the desired analysis, and the analysis should derive as much information as possible from the allotted runs. The two should not be considered mutually exclusive constructs, but must be considered from the onset in tandem.

Many issues arise when designing a simulation experiment, such as: (i) what input variables will be varied?, (ii) what levels of the input variables should be investigated?, (iii) what is the plan for proceeding from one simulation run to another?, and (iv) how is analysis restricted by the proposed experimental design? (Wild and Pignatiello [1991]). The experimental designs in this dissertation provide substantial progress for the second and third issues.

Watson [1961] states that with experimental designs, there exists "a sort of uncertainty principle whereby if the number of runs is decreased, the number of assumptions is increased; and conversely." Furthermore, there is a relationship between the quantity and quality of information, I, that can be gained as the number of observations is increased and the resources required, R, to obtain this information. Included within I is what we call *discriminatory power*. This refers to both correctly identifying the important model terms and avoiding the inclusion of terms that do not significantly influence the response. Included in R are the resources required, such as time and computing power. Note that I and R together summarize the previously defined efficient experimental design. A gain in one causes the other to increase, thus establishing a generic relationship between the two denoted as

$$I \;\; \alpha \;\; R\,. \tag{1.7}$$

It is the analyst's objective (and dilemma) to determine which levels and configurations of variables to use, while simultaneously considering the effect of (1.7). Managing this relationship should not rest solely upon the shoulders of the technical expert (experimental designer) or solely upon the project manager, who is perhaps unskilled in some aspects of experimental design, but requires their joint consideration. The designs in this dissertation will greatly aid in addressing this dilemma by providing designs which sample across a representation of the entire experimental region in a reasonable number of runs.

The choice of an experimental design should depend not just on the discriminatory power and resource availability, but also on the analyst's goal in running the experiment. Sacks et al. [1989] list the three primary objectives of computer experiments as (i) *predicting* the response at untried inputs, (ii) *optimizing* a function of the input variables, or (iii) *tuning* the computer code to physical data (i.e., *calibration*). The purposes of our research require that a forth objective be added to this list, obtaining insight.

In simulations of multi-entity military conflict, due to a dearth of data, these models are such that users often cannot reliably predict, optimize, or calibrate. Rather, analysts typically use these models to develop insights into complicated relationships. This is done, in part, by identifying important variables and interactions. However, one may expect that many variables (and interactions) may be important over some range, so identifying those ranges is also of special interest. Thus, instead of endeavoring to make a specific prediction or optimization equation, the focus on simulating complicated military models is often centered on developing important "*golden nugget*" insights. These insights, coupled with other analytical results or experience, build a decision-maker's knowledge base to make a more informed decision. As Srivastava [1987] aptly states, "It often seems that to some statisticians, the goal behind an experiment is to use an optimal design, rather than to probe into the important unknown features of the experimental situation." This dissertation stresses the need for identifying these unknown features.

D. DISSERTATION ORGANIZATION

This section provides a roadmap on how the dissertation is organized to address the research questions posed. This dissertation presents experimental designs with the following capabilities.

- The ability to explore broad regions of a complex simulation model containing a relatively high-dimensional input space characterized by a response surface that may be non-linear.
- The ability to identify significant variables and first-order and second-order interactions and the ranges of the variables where these effects occur.
- The ability to gracefully handle premature experiment termination. That is, it is common in operational situations for the number of simulation runs to be unexpectedly cut short. Experimental designs that anticipate this contingency become the more valuable ones.

The flow of the dissertation is as follows. Chapter II discusses the desirable characteristics of an experimental design and builds the foundation for the subsequent development of the new designs. Chapter III specifies new experimental designs when there is either only one output measure of interest[8] or where each output measure has its own characterization. This chapter contains both the theory underlying these designs and the details necessary to construct them. Chapter IV contains an application of this methodology on a known non-linear response surface. A comparison is made between its performance and that of other designs that have appeared in recent literature. It is shown that the new design outperforms the existing designs to which it is compared. Chapter V details the results of applying a 22-variable experimental design, and a recommended analysis methodology, to an agent-based simulation of a peace enforcement operation. In this application, military judgment guides the construction and examination of alternative metamodels in order to obtain potential insights about peace enforcement operations. The last chapter, Chapter VI, concludes the dissertation with a summary of the contributions to the existing body of knowledge and suggestions for future research.

One final note is in order. Although the motivation for developing this methodology stems from defense analyses, the methodology can also be applied to simulations developed for other fields or other purposes.

[8] Note that the measure could be a composite of several measures (e.g., a weighted sum).

II. EXPERIMENTAL DESIGNS FOR COMPLEX SIMULATIONS

This chapter contains the foundation for the subsequent development of the new designs (Sections A and B) and describes, in detail, the desirable characteristics of an experimental design (Section C).

The simulations that DoD analysts use are often quite large and almost unimaginably complex. Many models contain thousands of input variables, a vast number of which are potentially significant. Moreover, the response surface can be highly nonlinear. The complexity and uncertainty associated with these simulations makes utilizing strong prior knowledge (such as the distributional form of the error term) unreliable. To efficiently explore these simulations, experimental designs possessing the following desirable characteristics are needed:

- approximate orthogonality of the input variables,
- space-filling[9], that is, the collection of experimental cases should be a representative subset of the points in the hypercube of explanatory variables,
- ability to examine many variables (20 or more) efficiently,
- flexibility in analyzing and estimating as many effects, interactions, and thresholds as possible,
- requiring minimal *a priori* assumptions on the response,
- ease in generating the design, and
- ability to gracefully handle premature experiment termination.

A breadth of current design methods used in simulation was examined with respect to these desired characteristics, including group screening (e.g., Dorfman [1943], Patel [1962]), sequential bifurcation (e.g., Jacoby and Harrison [1962], Bettonvil [1995]), random balance (e.g., Satterthwaite [1959]) and Latin hypercubes (e.g., McKay et al. [1979], Ye [1998]), uniform designs (e.g., Hua and Wang [1981], Fang and Wang [1994]), robust designs (e.g., Taguchi [1988]), Bayes designs (e.g., Flournoy [1993],

---

[9] The principles of orthogonality and space-filling are described, in detail, in this chapter.

Chaloner and Verdinalli [1994]), search linear models (e.g., Srivastava [1975], Chatterjee et al. [2000]), and frequency domain (e.g., Schruben [1986], Morrice [1995]).[10]

The most promising of the current designs, in terms of satisfying the desirable characteristics, are the Latin hypercube designs and the uniform designs. These two types are explained in this chapter. The designs that are subsequently developed combine the strengths of these two types.

A. THE EVOLUTION OF ORTHOGONAL LATIN HYPERCUBES

This section traces the line of literature from random designs to Latin hypercube sampling to Latin hypercubes to orthogonal Latin hypercubes. The importance of orthogonality in experimental design matrices is stressed and examples are provided.

Satterthwaite [1959] proposed the use of a *random design*, "one for which a random sampling process [with replacement] is used to choose all or some of the elements of each variable in the design matrix." Significant correlations, as measured by (1.4), can exist between columns of the design matrix. Youden et al. [1959] present various criticisms of these designs. The principal criticisms are that the interpretation of the experimental results could not be sufficiently justified due to random confounding and that, for any variable setting, the estimators of the coefficients are biased.

McKay et al. [1979] show that one can improve upon random designs by using ideas from "quota sampling." They call their method *Latin hypercube sampling*, and state that the resulting design is a "first cousin" of the random design. In Latin hypercube sampling, the input variables are considered to be random variables with known distribution functions. For each input variable, $\boldsymbol{x_k}$, "all portions of its distribution [are] represented by input values" by dividing its range into "n strata of equal marginal probability $1/n$, and [sampling] once from [within] each strata." [11] (McKay et al. [1979]) For each $\boldsymbol{x_k}$, the n sampled input values are assigned at random to the n cases—with all $n!$ possible permutations being equally likely. This determines the column in the design matrix for $\boldsymbol{x_k}$. This is done independently for each of the k input variables. Therefore, for

[10] A comprehensive, but not complete, list of literature sources for these areas is included in the bibliography.

[11] In practice, many analysts take a fixed value within each strata (e.g., the median) rather than a random value.

each variable, $\boldsymbol{x_k}$, all of the n input values appear once and only once in the design. Also, for a given row in the design matrix, all of the n^k potential combinations of the input variable values (after the sampling) have an equal chance of occurring.

As an example, assume there are three input variables, each having a U[0,1] distribution, and that 10 simulation runs are to be made. Independently, for all three variables, one design value is chosen at random from within each of the 10 equal probable intervals [0,.1), [.1,.2), [.2,.3), [.3,.4), [.4,.5), [.5,.6), [.6,.7), [.7,.8), [.8,.9), and [.9,1]. For every input variable, the order in which the 10 sampled values appear in the design matrix is randomly determined, with all 10! possible orderings being equally likely. Table 2.1 shows one such realization of a design matrix obtained by this procedure. Note: As in this example, these design matrices will likely have correlations between columns.

Run	Variable 1	Variable 2	Variable 3
1	0.63	0.53	0.90
2	0.42	0.48	0.04
3	0.89	0.19	0.89
4	0.08	0.77	0.27
5	0.23	0.30	0.59
6	0.98	0.01	0.32
7	0.15	0.22	0.61
8	0.33	0.68	0.12
9	0.58	0.93	0.48
10	0.71	0.87	0.74

Table 2.1. An example of Latin hypercube sampling. The 10 run sample is taken from three independent U[0,1] input variables.

A common variant of the design obtained by Latin hypercube sampling is called a *Latin hypercube* (Tang [1993]). An $n \times k$ Latin hypercube consists of k permutations of the vector $\{1,2,\ldots,n\}^{\mathrm{T}}$. Therefore, the input values are predetermined and there is no sampling within strata. Each of the k columns contains the levels $1,2,\ldots,n$, randomly permuted, with each possible permutation being equally likely to appear in the design matrix. Each of these k columns is then randomly assigned, without replacement, to one of the k variables to create the Latin hypercube. The row vectors are design points in the

k-dimensional experimental region. All of the k one-dimensional projections of the Latin hypercube are evenly spaced; that is, the distance between any two adjacent levels is the same for all pairs of adjacent levels. This is known as the *equidistant* property. The Latin hypercubes that this dissertation addresses use a more general variant of the above. Specifically, the values of each of the variables may be any set of n evenly spaced values centered at the origin (Owen [1998]).

Since each variable has its predetermined values randomly ordered in the design matrix, Latin hypercubes are easy to generate. Moreover, as with Latin hypercube sampling, there are no restrictions on how the different variable columns are combined to form the design matrix. Table 2.2 gives an example of a Latin hypercube design for five variables, each at 11 levels, with the levels ranging from –1 to +1. Note that for each variable, the distance between adjacent levels is the same for each pair of adjacent levels, in this case a distance of 0.2. As in this example, Latin hypercube designs can have significant correlations—as measured by (1.4)—between the columns of the design matrix.

RUN	VARIABLE 1	VARIABLE 2	VARIABLE 3	VARIABLE 4	VARIABLE 5
1	0.2	-0.8	-0.4	-0.2	-0.8
2	0	0.6	0.2	0	-0.6
3	-0.8	1	-0.8	0.6	1
4	-1	-1	0.4	-0.8	0.2
5	0.4	-0.4	0	0.2	0.8
6	0.6	0	0.8	-1	0.4
7	-0.4	0.4	-0.2	-0.6	-1
8	-0.6	-0.6	-1	0.8	-0.4
9	0.8	-0.2	1	0.4	0.6
10	1	0.8	-0.6	1	0
11	-0.2	0.2	0.6	-0.4	-0.2

Table 2.2. A Latin hypercube having the equidistant property for each of its five variables. Each variable has 11 levels, with the levels ranging from –1 to +1 in increments of 0.2.

Ye [1998] constructs *orthogonal Latin hypercubes* in order to enhance the utility of Latin hypercube designs for regression analysis. Ye defines an orthogonal Latin hypercube (OLHC) as a Latin hypercube "for which every pair of columns has zero correlation." Furthermore, in Ye's OLHC construction, the elementwise square of each column has zero correlation with all other columns, and the elementwise product of every two columns has zero correlation with all other columns. These properties "ensure the independence of estimates of linear effects of each variable" and the "estimates of the quadratic effects and bilinear interaction effects are uncorrelated with the estimates of the linear effects." (Ye [1998])

As a simple example, assume two input variables each have the following five levels: –1.0, -0.5, 0.0, 0.5, and 1.0. A 5 × 2 OLHC for these two variables and five levels is shown in Table 2.3. The correlation between the two columns is 0.0.

Run	Variable A	Variable B
1	-1	-0.5
2	-0.5	1
3	0	0
4	0.5	-1
5	1	0.5

Table 2.3. A 5 × 2 orthogonal Latin hypercube with two variables, each at five levels.

Ye's [1998] method allows one to generate an OLHC when the number of runs is a power of 2 plus one (for a center point). Specifically, for any integer $m > 1$, Ye's (1998) technique builds OLHCs for k variables such that the number n of runs is related to k and m by

$$n = 2^m + 1, \tag{2.1}$$

$$k = 2m - 2. \tag{2.2}$$

Note that k must be even.

In the development of his orthogonal Latin hypercubes, Ye [1998] constructs three matrices. One matrix, **M**, has its columns composed of permutations of the variable levels. A second matrix, **S**, is similar to a two-level factorial design matrix on m–1 variables containing m–2 interaction terms; all entries are ±1. The third matrix, **T**, is created from the first two matrices. Succinctly, the columns of **M** correspond to permutations of the ordinal values of the positive levels of the variables (we assume there is an equal number of negative levels for the variables). The columns of **S** correspond to a subset of a two-level factorial design matrix consisting of –1's and 1's (with mutually orthogonal columns). The matrix **T** is created by the Hadamard product[12] of **M** and **S**. A mirror image of **T** and a row of 0's corresponding to the center point are then appended to the original **T** to create an OLHC.

1. Construction of the Matrix M for the OLHC

The matrix **M** from Ye (1998) is now considered in detail. The dimensions of **M** are $q \times k$, with $q = ((n-1)/2)$ being the number of positive levels of each variable. The first step in constructing **M** is to create a vector **e**, which is a random ordering of the first q natural numbers (1, 2, …, q). One column in **M** is **e**. Since the remaining columns of **M** depend on **e**, the choice of **e** is critical. A simple approach in choosing **e** is to use a simple 1, 2, …, q ordering. Although one may use the actual level values, it is easier to use ordinal values for the positive levels when constructing these matrices. For example, from Table 2.3, the value of 0.5 would correspond to 1 and the value of 1.0 would correspond to 2. Thus, if q represents the number of positive levels and a hierarchical ordering is used, then **e** is specified as

$$\mathbf{e} = [1, 2, \dots, q]^T . \tag{2.3}$$

Given an initial **e**, permutation matrices are used to generate the columns of **M**. Specifically, for L = 1, 2, …, m–1, create $q \times q$ permutation matrices, labeled $\mathbf{A_L}$, as follows. With **I** as the 2×2 identify matrix and

$$\mathbf{R} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{2.4}$$

each $\mathbf{A_L}$ is constructed by

$$\mathbf{A_L} = \underbrace{\mathbf{I} \otimes \dots \otimes \mathbf{I}}_{m\text{-}1\text{-}L} \otimes \underbrace{\mathbf{R} \otimes \dots \otimes \mathbf{R}}_{L}, \tag{2.5}$$

where $\otimes$ denotes the Kronecker product. There are m–1 of these permutation matrices created, each of size $q \times q$.

---

[12] A Hadamard product exists for two matrices that are conformable. The corresponding elements of the two matrices are multiplied together to yield the Hadamard product.

Additional permutation matrices, m–2 of them, are then created by multiplying any m–2 distinct pairs of the permutation matrices $\mathbf{A_1}$ through $\mathbf{A_{m-1}}$ by one another.[13] The k, where $k = 2m–2$, columns of $\mathbf{M}$ are composed of $\mathbf{e}$, $\mathbf{A_i e}$, for i = 1, 2, …, m–1, and $\mathbf{A_i A_j e}$, where there are m–2 distinct pairs of i and j, with i and j both $\in$ {1, 2, …, m–1}, with $i \neq j$.

For example, from (2.1), let m=4 and n=17. The six columns of $\mathbf{M}$ are formed from $\mathbf{e}$, $\mathbf{A_1e}$, $\mathbf{A_2e}$, $\mathbf{A_3e}$, $\mathbf{A_1A_2e}$, and $\mathbf{A_1A_3e}$. The matrix $\mathbf{M}$ that is generated by using $\mathbf{e}$=[1,2,3,4,5,6,7,8]^T is shown in Table 2.4.

e	$\mathbf{A_1e}$	$\mathbf{A_2e}$	$\mathbf{A_3e}$	$\mathbf{A_1A_2e}$	$\mathbf{A_1A_3e}$
1	2	4	8	3	7
2	1	3	7	4	8
3	4	2	6	1	5
4	3	1	5	2	6
5	6	8	4	7	3
6	5	7	3	8	4
7	8	6	2	5	1
8	7	5	1	6	2

Table 2.4. An example matrix M, which is used in the construction of an OLHC (Ye [1998]), having six variables and eight positive levels with e=[1,2,3,4,5,6,7,8]^T**. Note that not all possible pairwise combinations of the $\mathbf{A_L}$ are used.**

2. Construction of the Matrices S and T for the OLHC

The matrices $\mathbf{S}$ and $\mathbf{T}$ from Ye (1998) are now considered in detail. The dimensions of $\mathbf{S}$ are $q \times k$. The dimensions of $\mathbf{T}$ are also $q \times k$. The final OLHC is an $n \times k$ design matrix, with $n = 2q + 1$.

$\mathbf{S}$ is equivalent to a subset of k columns of an m–1 variable two-level full factorial design matrix, including the columns used to estimate interactions. The first column of $\mathbf{S}$ consists of q +1's. The next m–1 columns of $\mathbf{S}$ are identical to the columns used to estimate the main effects in an m–1 variable two-level full factorial design matrix. The remaining m–2 columns of $\mathbf{S}$ are identical to m–2 of the columns used to estimate

[13] Ye [1998] specifically used the m–2 pairs $\mathbf{A_1A_{m-1}}$, …, and $\mathbf{A_{m-2}A_{m-1}}$. However, any m–2 distinct pairs of permutation matrices are sufficient to generate orthogonal Latin hypercubes.

pairwise interactions in an m–1 variable two-level full factorial design matrix. They can be obtained by multiplying, element by element, the main effect columns together.

To illustrate this process, let us construct the matrix **S** for the case when $n = 17$ and $k = 6$ (i.e., $m = 4$). The six variables each have eight positive levels (similarly, they have eight negative levels). Thus, the construction requires eight rows ($q = 8$) and six columns (one column for each variable). The first column consists of +1's and the second, third, and fourth columns are orthogonal columns of +1's and –1's, and are identical to the main effects columns in a 2^3 full factorial design matrix (see, e.g., Box et al. [1978], Hicks [1993]). Columns five and six may consist of the product of (a) columns two and three, (b) columns two and four, or (c) columns three and four. In all cases, the columns are mutually orthogonal. Columns two, three, and four must not contain any confounding patterns because significant correlation will otherwise result. Because **M** can only accommodate six variables, as shown previously in Table 2.4, **S** has the same number of columns.

C_1	C_2	C_3	C_4	C_2C_3	C_2C_4
+1	-1	-1	-1	+1	+1
+1	+1	-1	-1	-1	-1
+1	-1	+1	-1	-1	+1
+1	+1	+1	-1	+1	-1
+1	-1	-1	+1	+1	-1
+1	+1	-1	+1	-1	+1
+1	-1	+1	+1	-1	-1
+1	+1	+1	+1	+1	+1

Table 2.5. An example of S for an OLHC (Ye [1998]) having six variables and eight positive levels, where C_i (i=1, 2, 3, 4) and C_iC_j (j=2, 3, 4 and $i \neq j$) indicate columns.

T is the Hadamard product of **M** and **S**. A mirror image of **T** and a row of 0's corresponding to the center point are appended to the original **T** to create an OLHC. The final OLHC, which has six variables and 17 runs, is shown in Table 2.6.

Run	Variable A	Variable B	Variable C	Variable D	Variable E	Variable F
1	1	-2	-4	-8	3	7
2	2	1	-3	-7	-4	-8
3	3	-4	2	-6	-1	5
4	4	3	1	-5	2	-6
5	5	-6	-8	4	7	-3
6	6	5	-7	3	-8	4
7	7	-8	6	2	-5	-1
8	8	7	5	1	6	2
9	0	0	0	0	0	0
10	-1	2	4	8	-3	-7
11	-2	-1	3	7	4	8
12	-3	4	-2	6	1	-5
13	-4	-3	-1	5	-2	6
14	-5	6	8	-4	-7	3
15	-6	-5	7	-3	8	-4
16	-7	8	-6	-2	5	1
17	-8	-7	-5	-1	-6	-2

Table 2.6. An OLHC with six variables and 17 levels using Ye's [1998] algorithm.

B. UNIFORM DESIGNS AND SPACE-FILLING

Uniform designs are introduced in this section. Fang et al. [2000] define a *uniform design* as a design "that allocates experimental points [which are] uniformly scattered on the domain." Uniform designs do not require orthogonality. Fang et al. [2000] classify uniform designs as space-filling designs. A good s*pace-filling design* is one in which the design points are scattered throughout the experimental region with minimal unsampled regions; that is, the voided regions are relatively small. This means that the design points are not concentrated in clusters or solely at corner points of the region, as can happen with two-level factorial designs.

Space-filling designs provide coverage of the entire experimental region, and this facilitates broad exploration of the model. They are particularly valuable when the experimenter is unsure of what the response surface might look like. Ye [1998] notes that good space-filling designs are "desirable for data analysis methods such as residual plots in regression diagnostics and nonparametric surface fitting."

To further clarify space-filling, this principle is illustrated with several figures. Figure 2.1 shows a traditional 2^3 factorial design, where each design point is at a corner of the cubical region. In Figure 2.1, it is assumed that the design points are at the endpoints of the variables, but this is not a requirement. Under this assumption, the interior of the cube does not have any design points, and is thus not sampled—although a center point is commonly added. Conversely, a uniform design (three variables with each variable having eight levels), as shown in Figure 2.2, has points distributed throughout the interior of the cube and is not limited to the corners or surfaces of the cube.

The key point is that the uniform design has design points scattered throughout the entire experimental domain in a somewhat uniformly distributed way. In this example, the uniform design has each variable at eight levels, but the factorial design has each variable at only two levels. If it turns out that only a small number of variables affect the response, then a uniform design allows an analyst more flexibility in fitting complex models, such as high-degree polynomials, to the essential variables. In the extreme case, in which only one variable turns out to be important, a Latin hypercube design contains *n* different (equally spaced) input values for the important variable.

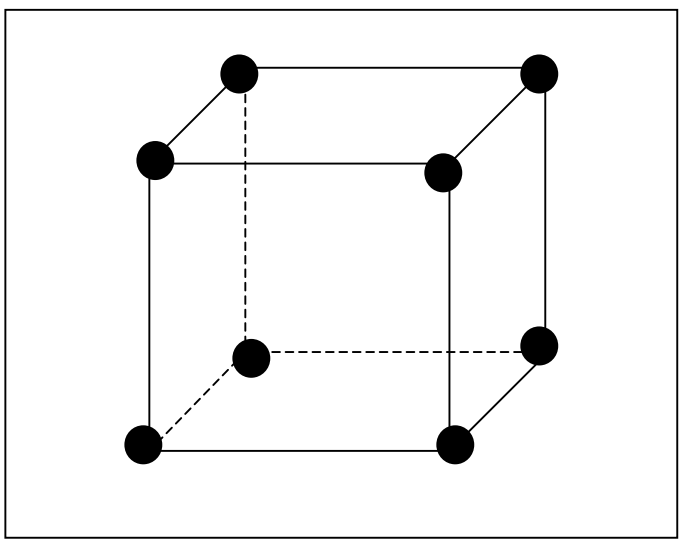

Figure 2.1. The design points of a 2^3 factorial design illustrating that only the corner points of the region are sampled.

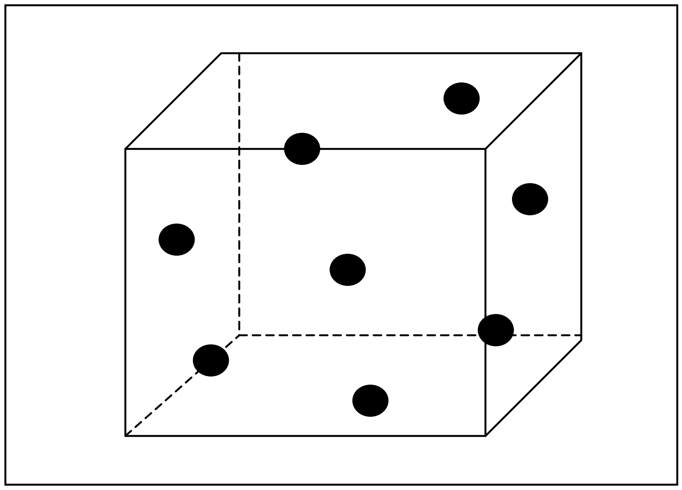

Figure 2.2. A uniform design illustrating the dispersion of points (space-filling) throughout the entire region.

Fang and Wang [1994] describe the goal of uniform designs as to find "design points which are uniformly scattered in the k-dimensional unit cube C^k," where uniformity, or space-filling, is measured by discrepancy. Using number-theoretic ideas, Fang and Wang [1994] define *discrepancy* as follows. Let $P = \{\mathbf{x_j}, \text{j=1,…,n}\}$ be a set of points on C^k and $v([\mathbf{0},\boldsymbol{\gamma}]) = \gamma_1\gamma_2\cdots\gamma_k$ the volume of the rectangle $[\mathbf{0},\boldsymbol{\gamma}]$. For any $\boldsymbol{\gamma} \in C^k$, let $N(\boldsymbol{\gamma}, P)$ be the number of points satisfying $\mathbf{x_j} \leq \boldsymbol{\gamma}$. Then the discrepancy is

$$L_\infty = \sup_{\gamma \in C^k} \left| \frac{N(\boldsymbol{\gamma}, P)}{n} - v([\mathbf{0},\boldsymbol{\gamma}]) \right|. \tag{2.6}$$

Equation (2.6) compares the proportion of points within rectangular subspaces to the volume of the rectangles. Discrepancy is the supremum of the absolute difference over all nested rectangles anchored at the origin. A large value (the theoretical maximum value is one) indicates that either a particular subregion has too many or too few design points in it. A smaller discrepancy measure (the theoretical minimum value is zero) indicates better space-filling.

An illustrative example of discrepancy calculations from Fang and Wang [1994] for two dimensions is given. Assume that two variables are chosen for a simulation. A uniform design strives to uniformly scatter the design points in the two-dimensional experimental region. If, for a particular rectangle, the "absolute value for the ratio of the number of points lying in the rectangle $[\mathbf{0},\boldsymbol{\gamma}]$ and the total number of points of the set minus the volume of the rectangle $[\mathbf{0},\boldsymbol{\gamma}]$ is small," then the proportion of points within the rectangle is nearly proportional to the volume of the rectangle—indicating good uniformity. Figure 2.3 illustrates this principle. Only two of the infinite number of possible rectangles are shown. In this example, a disproportionate number of the total points fall into Rectangle 2. Thus, the discrepancy will be large—i.e., the design's space-filling is poor.

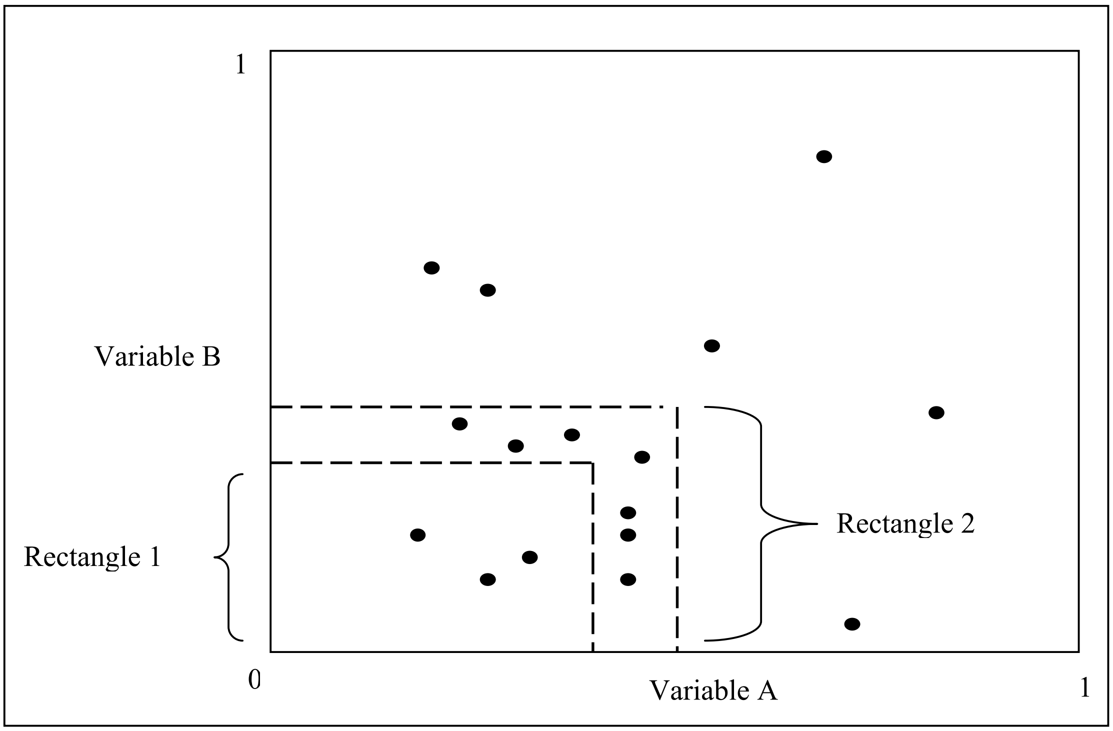

Figure 2.3. Example of discrepancy for two dimensions. An infinite number of nested rectangles exist. Two of these rectangles are shown with Rectangle 2 having a larger discrepancy (or poorer space-filling) than Rectangle 1.

The discrepancy measure of (2.6) provides the most accurate measure of the space-filling of the design points. Fang et al. [2000] state that "discrepancy has been universally accepted in quasi-Monte-Carlo methods and number theoretic methods." Unfortunately, as they note, "one disadvantage of [this] measure is that it is expensive to compute." Equation (2.6) has been used to assess the space-filling of designs having no more than two variables and 10 runs (Fang and Wang [1994]). For designs having more variables or runs or when the L_∞ discrepancy from (2.6) is too computationally burdensome to calculate (as is the case with our designs), the *modified L_2 discrepancy* (ML_2), shown in (2.7), can be used. The ML_2 is an approximation of the L_∞ discrepancy, and is easier to calculate numerically when there are either more than two variables or more than 10 runs (Fang et al. [2000]), and considers "projection uniformity over all subdimensions." (Fang et al. [1998]) Furthermore, (2.7) is considered to be an excellent alternative to (2.6) and is commonly used in assessing the space-filling of proposed experimental designs (see, e.g., Fang et al. [1998], Matousek [1998], Hickernell [1999], Okten [2001]). Consequently, since the designs developed in this dissertation have more than two variables and 10 runs, (2.7) is used when assessing the space-filling of a design.

$$ML_2 = \left(\frac{4}{3}\right)^k - \frac{2^{1-k}}{n}\sum_{d=1}^{n}\prod_{i=1}^{k}(3 - x_{di}^2) + \frac{1}{n^2}\sum_{d=1}^{n}\sum_{j=1}^{n}\prod_{i=1}^{k}[2 - \max(x_{di}, x_{ji})] \qquad (2.7)$$

Given two designs, the design with a smaller ML_2 discrepancy has better space-filling.

C. DESIRABLE CHARACTERISTICS

The desirable characteristics of an experimental design are described in this section. Furthermore, the measures that we use in assessing an experimental design's ability to achieve these characteristics are discussed. Orthogonality and space-filling are the primary characteristics of the designs developed in this dissertation.

1. Orthogonality Measures

An orthogonal design is desirable since it ensures independence among the coefficient estimates in a regression model. Orthogonality enhances our ability to analyze and estimate as many effects, interactions, and jump discontinuities as possible. Two measures are used to assess the degree of orthogonality. One measure is the maximum pairwise correlation of the columns of a design matrix. The maximum pairwise correlation, ρ, is found by calculating the absolute value of (1.4) for all pairs of column vectors in the design matrix, and then selecting the maximum of these values. A value of 0 is best (signaling orthogonality), and a value of 1 is worst (indicating that at least one column in the design matrix is a linear combination of the remaining columns).

The second measure of orthogonality is a condition number of $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$, where $\boldsymbol{X}$ is the design matrix. The condition number is commonly used in numerical linear algebra applications (e.g., Golub and Van Loan [1983], Demmel [1997], Leon [1998]) to examine the sensitivities of a linear system. Additionally, it can reveal the degree of orthogonality of the proposed design matrix. The author is unaware of any literature that uses the condition number to measure the orthogonality of a design matrix. An orthogonal design matrix has a condition number of 1. A non-orthogonal design matrix has a condition number greater than 1. A large condition number indicates that the candidate design matrix may be ill-conditioned (i.e., has substantial multicollinearity). The condition number (using the infinity norm) is defined by

$$cond_{\infty}(\phi) = \|\phi\|_{\infty} \, \|\phi^{-1}\|_{\infty} \ , \tag{2.8}$$

where ϕ represents the correlation matrix of the proposed design matrix. A companion condition number is generated from the singular value decomposition (SVD). This *SVD condition number* (using the 2-norm of the design matrix) is defined by

$$cond_2(X^T X) = \frac{\psi_1}{\psi_n} , \tag{2.9}$$

where ψ_1 is the largest singular value, and ψ_n is the smallest singular value of $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$. When a condition number is referenced in this dissertation, it corresponds to (2.9). This measure represents the degree of orthogonality of the design matrix, with a value of 1 indicating orthogonality and a value greater than 1 indicating the degree of non-orthogonality. Thus, a condition number as close to 1 as possible is desired.

There is not necessarily a one-to-one correspondence between ρ and the condition number, but the condition number is related to the number of the pairs of columns that are correlated and the magnitudes of the correlations. The author is unaware of any previous literature using both the maximum pairwise correlation and condition number to assess the degree of orthogonality of a design matrix. One measure, ρ, gives the worst case correlation between design matrix columns, while the other measure, the condition number, provides an assessment of the overall orthogonality of the proposed design matrix. A non-orthogonal design matrix has at least one non-zero correlation between two of its columns, and a condition number greater than 1. A design matrix will be classified as *nearly orthogonal* if it has a maximum pairwise correlation no greater than 0.03 and a condition number no greater than 1.13.[14]

2. Space-Filling Measures

A design matrix with good space-filling properties is desirable since design points are distributed throughout the entire experimental region. This permits a greater opportunity to identify contours that define regions where interesting behavior occurs. Two measures are used to assess the space-filling of a design matrix. The first measure is the previously described ML_2 discrepancy.

The second measure used in assessing the space-filling of a design is the Euclidean maximin (Mm) distance (Ye [1998], Johnson et al. [1990], Morris and Mitchell [1992], [1995]). For a given design, define a distance list $\mathbf{d}=(d_1,d_2,\ldots,d_{[n(n-1)]/2})$, where the elements of **d** are the Euclidean inter-site distances of the n design points, ordered from smallest to the largest. The *Euclidean Mm distance* is defined as d_1, where a larger value is better. A large value of d_1 means that no two points are close to (within d_1 of) each other. Other distance metrics that practitioners use include Mahalanobis, Euclidean, and rectangular, with the most common being rectangular and Euclidean (Johnson et al. [1990], Morris and Mitchell [1992], [1995]). This dissertation uses the Euclidean Mm distance since it emphasizes the shortest distance between points. Furthermore, when Mm distance is referenced here, it refers to the Euclidean Mm distance. The author is unaware of any literature that uses both ML_2 discrepancy and Mm distance to measure the space-filling of a design. Both measures are used in this dissertation because in some cases a single measure by itself does not provide sufficiently adequate discrimination between candidate designs.

3. Other Criteria

The ability to quickly and easily generate an experimental design is important. For example, one of the major disadvantages of uniform designs (Fang and Wang [1994]) is the difficulty in finding a design for many combinations of variables and runs, thus severely restricting the number of uniform designs readily available for use. If the goal

[14] Although these values are somewhat arbitrary, designs satisfying these criteria suffer minimal multicollinearity effects (see, e.g., Golub and Van Loan [1983], Pukelsheim [1993]). Furthermore, good space-filling designs exist with this degree of non-orthogonality.

of an analysis is to explore the experimental region, then expending an inordinate amount of time deriving the experimental design makes this goal harder to realize.

Constructing a design should not require substantial *a priori* distributional assumptions on the response and its relationship to the input variables. In most defense analyses, it is not unreasonable to ask the experts which variables they think *a priori* will be important. It is almost always unreasonable to ask experts to provide *a priori* distributions (including correlation structure) on the variables' effects on the outputs. Furthermore, even expert judgment concerning the appropriate variable levels can be erroneous. This concern is especially relevant with military models, where "surprises" are more the rule than the exception.

The designs should be relatively insensitive to the premature termination of the planned set of experimental runs. This is a common problem in defense analyses, where results can be required sooner than originally planned. If an experiment is terminated early, the subset of runs may not be orthogonal. The subsequent regression analysis can suffer from the effects of multicollinearity.

Finally, the designs should have the ability to examine high-dimensional input spaces (more than 20 variables) efficiently. The ability to search across a breadth of factors greatly enhances the opportunity to find significant effects, interactions, and interesting regions of behavior in the output response.

D. SUMMARY

This chapter focused on desirable design characteristics. The two most critical characteristics are (near) orthogonality and space-filling. Specifically, both the maximum pairwise correlation and the condition number measure the degree of orthogonality. Space-filling is assessed with both the ML_2 discrepancy and Mm distance measures. The OLHC designs provide orthogonal designs, while the uniform designs focus on space-filling. In the next chapter, these types are melded together to create new designs that perform well on both of these characteristics. Awareness of the other design characteristics mentioned in this section is also maintained.

III. DEVELOPMENT OF NEW EXPERIMENTAL DESIGNS

This chapter details an approach to designing Latin hypercubes that are orthogonal or nearly orthogonal and have good space-filling properties. Specifically, we present designs for two to 22 variables using an initial set of runs ranging from 17 to 129 in number. Although this dissertation limits itself to designs with, at most, 22 variables, the algorithm can apply directly to any number of variables; but of course, the computational resources required would grow rapidly.

The general plan is to extend the use of Ye's [1998] algorithm in order to construct additional designs. Some of these preserve the orthogonality property and some do not. Typically the ones that preserve the orthogonality property have poor space-filling capabilities. Algorithms that improve the space-filling capabilities may do so while compromising orthogonality. The goal is to provide a sequence of steps that lead to an effective trade-off between the concepts of near orthogonality and space-filling. This activity is computer intensive, but the steps provided lead to effective designs that achieve the goal.

In Section A, Ye's [1998] algorithm is extended to allow the examination of a greater number of variables. In Section B, some orthogonality is sacrificed in order to achieve improved space-filling. Section C provides the best designs found to date for up to 22 variables. Section D gives an approach for adding additional design runs that (at least) maintain the orthogonality measures, while simultaneously improving on the design's space-filling properties. The last section, Section E, summarizes the new approach, including the specific steps necessary to generate nearly orthogonal Latin hypercubes.

A. CONSTRUCTING ORTHOGONAL LATIN HYPERCUBES

This section describes the development of experimental designs that satisfy the desirable characteristics. These orthogonal designs build directly from Ye's [1998] OLHC construction. Specifically, his three matrices (**M**, **S**, and **T**) are augmented with additional columns, thus permitting the analyst to examine a greater number of variables in the same number of runs. The roles played by these matrices are the same as before.

The matrix **M** contains permutations of the values of the variables and **S** attaches signs to these values. The output matrix **T** is the Hadamard product of **M** and **S**.

1. Incorporating Additional Variables into OLHC Designs

This section describes how to extend Ye's [1998] OLHC designs so that additional variables can be examined in the same number of runs. In his construction, Ye uses only m–2 of the $\binom{m-1}{2}$ possible pairwise combinations of the permutation matrices, denoted $\mathbf{A_L}$**,** in the creation of **M**. This is the starting point for the new designs. A similar matrix **M** is constructed, but all of the pairwise combinations of the matrices $\mathbf{A_L}$ (Ye [1998]) are used. The number of variables that can be examined by using all pairwise combinations of the $\mathbf{A_L}$'s in **M** is found using our following theorem.

Theorem 3.1: Within n runs, where $n = 2^m + 1$, with m an integer greater than 1, the maximum number of variables that can be examined in a Latin hypercube, using all original and pairwise combinations of Ye's [1998] matrices $\mathbf{A_L}$, is

$$m + \binom{m-1}{2}. \tag{3.1}$$

Proof: This follows by construction. The vector **e** constitutes one variable. Each $\mathbf{A_L}$, up to a maximum of m–1, corresponds to a column in the design matrix. Finally, each of the $\binom{m-1}{2}$ pairwise combinations of the $\mathbf{A_L}$'s also corresponds to a column in the design matrix. Recall from Chapter II that the vector **e** determines the subsequent matrices $\mathbf{A_L}$. Note that different vectors of **e** may result in the same overall design matrix, but (3.1) holds under each specification of **e**. □

The matrix **M** is constructed using (2.3), (2.4), and (2.5). The matrix **S**, which must match the dimensions of **M**, is similarly augmented with additional columns. The additional columns are equivalent to the (previously unused) columns used in estimating pairwise interactions in an m–1 two-level full-factorial design. The matrix **T**, which is the Hadamard product of **M** and **S**, is calculated as before.

If there are eight positive levels (and correspondingly eight negative levels and a center point), for a total of 17 levels, the maximum number of variables that we can examine is $1 + 3 + \binom{3}{2} = 7$. Similarly, if there are 64 positive levels (and correspondingly 64 negative levels) for a total of 129 levels, including the center point, the maximum number of variables which may be examined is $1 + 6 + \binom{6}{2} = 22$.

Under Ye's OLHC construction, he only guarantees orthogonal designs as specified by (2.1) and (2.2). The OLHC's can be constructed for the number of variables specified in Theorem 3.1. For example, although an OLHC can be created for eight variables with each variable at 33 levels as specified by Ye, given the same 33 levels, one can construct an OLHC with 11 variables. The key in designing this OLHC is that the first column in **M** from Section II.A.1 must be

$$\mathbf{e} = (1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16)^T . \tag{3.2}$$

Theorem 3.2 generalizes this finding.

> Theorem 3.2: If $\mathbf{e} = [1, 2, \ldots, q]^T$, where q represents the number of positive levels, is used to generate a Latin hypercube as specified in Theorem 3.1 (for up to m=10), the resulting Latin hypercube is orthogonal.

Proof: The proof is by computational verification. That is, the author has used this method to construct an OLHC for all choices between two and 46 variables. Note that in every case examined, this approach has found an OLHC.[15] □

A comparison between the number of variables that can be examined using Ye's [1998] designs and the extended orthogonal designs is shown in Table 3.1.

---

[15] It is conjectured that Theorem 3.2 applies for any value of m more than 10.

Total number of levels for each variable	m	Maximum number of variables by extending Ye's OLHC	Maximum number of variables for Ye's OLHC
17	4	7	6
33	5	11	8
65	6	16	10
129	7	22	12

Table 3.1. A comparison illustrating the increased number of variables that can be examined by extending Ye's [1998] construction algorithm for OLHC's.

It is readily apparent from Table 3.1 that as the number of levels doubles (less one, for the center point), Ye's OLHC designs are able to accommodate exactly two more variables. In the new designs, the corresponding maximum number of variables increases by the previous *m*. This difference grows dramatically as the number of variables to be explored increases. For example, Ye's approach requires 4,097 runs to build an OLHC for 22 variables. The difference gets even more dramatic when there are more variables in the design. Thus, the new designs (for up to 22 variables from Table 3.1) are capable of examining many more variables than Ye's [1998] designs while maintaining orthogonality.

2. An Example OLHC with Seven Variables and 17 Levels

An OLHC which has more columns than Ye's [1998] OLHC is constructed using Theorems 3.1 and 3.2. S-Plus [1991] is employed for this endeavor. Assume one constructs an OLHC with seven variables and 17 levels (including the 0.0 center point) using Theorem 3.2, where

$$\mathbf{e} = [1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8]^T. \tag{3.3}$$

The matrix $\mathbf{M}$ is constructed using (2.3), (2.4), (2.5), and Theorem 3.1, and is shown in Table 3.2. The difference between this design and that in Table 2.4—using Ye's construction—is that all three of the pairwise combinations of the $\mathbf{A_L}$'s are used. That is, $\mathbf{A}_1\mathbf{A}_2\mathbf{e}$, $\mathbf{A}_1\mathbf{A}_3\mathbf{e}$, and $\mathbf{A}_2\mathbf{A}_3\mathbf{e}$ are all included in $\mathbf{M}$.

e	A_1e	A_2e	A_3e	A_1A_2e	A_1A_3e	A_2A_3e
1	2	4	8	3	7	5
2	1	3	7	4	8	6
3	4	2	6	1	5	7
4	3	1	5	2	6	8
5	6	8	4	7	3	1
6	5	7	3	8	4	2
7	8	6	2	5	1	3
8	7	5	1	6	2	4

Table 3.2. The matrix M for a seven-variable, 17-level OLHC.

The matrix **S** is constructed using the two-level factorial design shown in Table 3.3. Recall that any version of this two-level factorial design may be used without jeopardizing the orthogonality of the final design matrix.

C_1	C_2	C_3	C_4	C_2C_3	C_2C_4	C_3C_4
+1	-1	-1	-1	+1	+1	+1
+1	+1	-1	-1	-1	-1	+1
+1	-1	+1	-1	-1	+1	-1
+1	+1	+1	-1	+1	-1	-1
+1	-1	-1	+1	+1	-1	-1
+1	+1	-1	+1	-1	+1	-1
+1	-1	+1	+1	-1	-1	+1
+1	+1	+1	+1	+1	+1	+1

Table 3.3. The matrix S for a seven-variable, 17-level OLHC.

The matrix **T** is then constructed using the Hadamard product of **M** and **S**. The design matrix is completed by augmenting **T** with its mirror image and the center point, resulting in the 17×7 OLHC.

We will represent an OLHC by the notation of $(O)_k^n$, where n represents the number of runs or experiments and k represents the number of variables. An $(O)_7^{17}$ design is shown in Table 3.4. Each column represents an individual variable and its associated values, while each row corresponds to the variable settings for a particular run or observation.

Run	Variable A	Variable B	Variable C	Variable D	Variable E	Variable F	Variable G
1	1	-2	-4	-8	3	7	5
2	2	1	-3	-7	-4	-8	6
3	3	-4	2	-6	-1	5	-7
4	4	3	1	-5	2	-6	-8
5	5	-6	-8	4	7	-3	-1
6	6	5	-7	3	-8	4	-2
7	7	-8	6	2	-5	-1	3
8	8	7	5	1	6	2	4
9	0	0	0	0	0	0	0
10	-1	2	4	8	-3	-7	-5
11	-2	-1	3	7	4	8	-6
12	-3	4	-2	6	1	-5	7
13	-4	-3	-1	5	-2	6	8
14	-5	6	8	-4	-7	3	1
15	-6	-5	7	-3	8	-4	2
16	-7	8	-6	-2	5	1	-3
17	-8	-7	-5	-1	-6	-2	-4

Table 3.4. An OLHC for seven variables where each variable has 17 levels.

The variables in Table 3.4 all range from –8 to 8. Of course they can be scaled as necessary. For example, if for the analyses one wants to vary each of the variables in Table 3.4 from –1 to 1, one can use the design matrix in Table 3.5.

Run	Variable A	Variable B	Variable C	Variable D	Variable E	Variable F	Variable G
1	0.125	-0.25	-0.5	-1	0.375	0.875	0.625
2	0.25	0.125	-0.375	-0.875	-0.5	-1	0.75
3	0.375	-0.5	0.25	-0.75	-0.125	0.625	-0.875
4	0.5	0.375	0.125	-0.625	0.25	-0.75	-1
5	0.625	-0.75	-1	0.5	0.875	-0.375	-0.125
6	0.75	0.625	-0.875	0.375	-1	0.5	-0.25
7	0.875	-1	0.75	0.25	-0.625	-0.125	0.375
8	1	0.875	0.625	0.125	0.75	0.25	0.5
9	0	0	0	0	0	0	0
10	-0.125	0.25	0.5	1	-0.375	-0.875	-0.625
11	-0.25	-0.125	0.375	0.875	0.5	1	-0.75
12	-0.375	0.5	-0.25	0.75	0.125	-0.625	0.875
13	-0.5	-0.375	-0.125	0.625	-0.25	0.75	1
14	-0.625	0.75	1	-0.5	-0.875	0.375	0.125
15	-0.75	-0.625	0.875	-0.375	1	-0.5	0.25
16	-0.875	1	-0.75	-0.25	0.625	0.125	-0.375
17	-1	-0.875	-0.625	-0.125	-0.75	-0.25	-0.5

Table 3.5. An OLHC for seven variables where each variable has a range of –1 to 1.

3. Space-Filling of the OLHC Example

Orthogonality (or near orthogonality) is a critical design characteristic. Space-filling is another critical design characteristic, and Ye [1998] notes that "an OLHC design…does not necessarily have a good space-filling property." Indeed, although orthogonal, generally the space-filling properties of the designs generated using Theorems 3.1 and 3.2 is poor. The goal is to improve upon the space-filling of these $(O)_7^{17}$ designs.

To visually display the space-filling of a design, it is typical to project the design points into two dimensions (e.g., Johnson et al. [1990], Morris and Mitchell [1995], Ye [1998]). Figure 3.1 presents the two-dimensional projections of variable pairs from Table 3.5. In two dimensions, the design points exhibit systematic patterns that concentrate on specific regions instead of across the entire region. Note that the three two-dimensional projection of variables A and B, C and E, and D and F make an approximate "X" figure and do not adequately sample the region. Specifically, there are substantial regions in the two-dimensional subspaces with no points in them. Thus, any effects that may occur in those regions will be missed by the design. Considering Figure 3.1, the only two-dimensional projections which visually present adequate space-filling are the three pairs of variables B and G, C and F, and D and E.

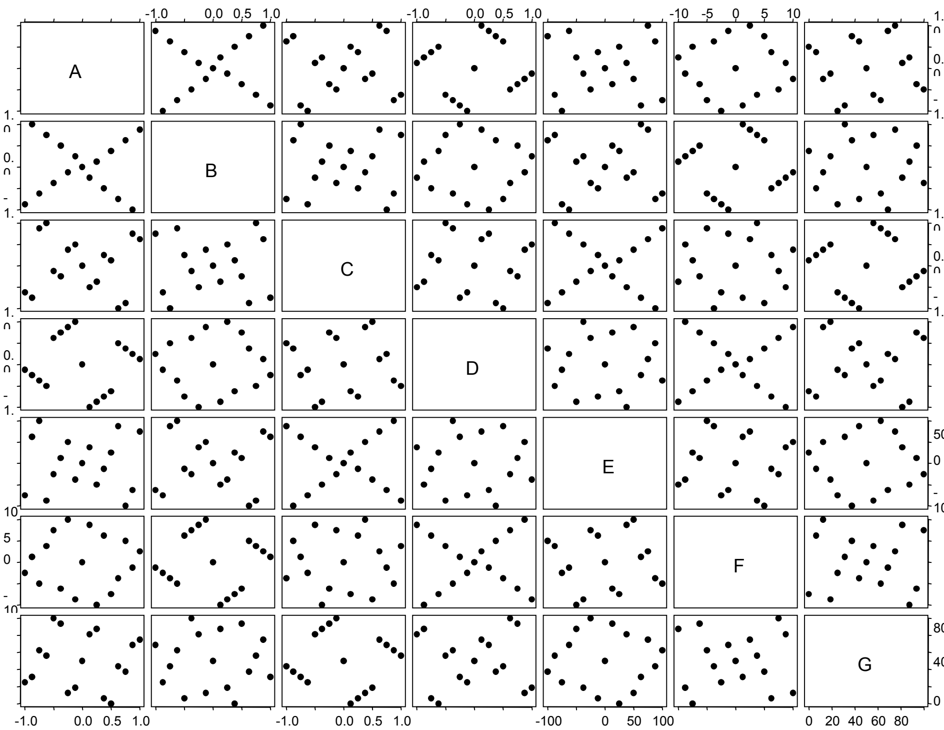

Figure 3.1. Two-dimensional projections for the variable pairs from Table 3.5.

Although the design matrix generated from Theorems 3.1 and 3.2 is orthogonal, the space-filling of the design is poor. Similarly, poor space-filling (i.e., systematic patterns in the two-dimensional projections and substantial regions in the two-dimensional subspace with no design points) regions are found in the $(O)_{11}^{33}$, $(O)_{16}^{65}$, and $(O)_{22}^{129}$ designs.

4. Finding the Best Space-Filling OLHC with Seven Variables and 17 Levels

Following Theorem 3.2, the $(O)_7^{17}$ design in Figure 3.1 was generated using $\mathbf{e}=[1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8]^T$. Recall that **e** uniquely specifies the subsequent development of **M** (and thus the final design matrix), and that not all candidate vectors **e** produce an OLHC. The number of possible orderings of the first column (**e**) of **M** is $q!$.

In the $(O)_7^{17}$ example, there are 40,320 possible permutations of **e**. The reader should note the combinatorial problem associated with constructing **M** as the number of levels increases. Enumerating all permutations of **e** is feasible for the design matrices with seven or fewer variables, but is computationally difficult for more than seven variables.

From the 40,320 possible different $(O)_7^{17}$ designs, there are 143 distinct designs that are orthogonal. From these designs, the designer seeks a design with good space-filling properties. Unfortunately, each of these 143 $(O)_7^{17}$ designs has an Mm distance of 1.47902. Thus, if the previous literature is followed, (e.g., Johnson et al. [1990], Morris and Mitchell [1992], Ye [1998]), there is no space-filling distinction between these 143 $(O)_7^{17}$ designs. This fact is one of the reasons that a second measure of space-filling is used for comparing designs.

Next consider the ML_2 discrepancies for the 143 distinct $(O)_7^{17}$ designs. The ML_2 discrepancies range from .151854 to .173952. The $(O)_7^{17}$ design generated from Theorems 3.1 and 3.2 has an ML_2 discrepancy of .173223 (almost, but not quite, the worst ML_2 discrepancy). The choice of **e** corresponding to the minimum (i.e., preferred) ML_2 discrepancy is $\mathbf{e}=[1,2,8,4,5,6,7,3]^T$. The choice of **e** corresponding to the maximum ML_2 discrepancy is $\mathbf{e}=[2,7,1,8,4,5,3,6]^T$. The $(O)_7^{17}$ design having the minimum ML_2 discrepancy is shown in Table 3.6. The two-dimensional projections of the variables of Table 3.6 are shown in Figure 3.2. From a visual inspection, it is evident that the two-dimensional projections of the best $(O)_7^{17}$ design have better space-filling than the $(O)_7^{17}$ design constructed using Theorems 3.1 and 3.2 and illustrated in Figure 3.1.

Run	Variable A	Variable B	Variable C	Variable D	Variable E	Variable F	Variable G
1	0.125	-0.25	-0.5	-0.375	1	0.875	0.625
2	0.25	0.125	-1	-0.875	-0.5	-0.375	0.75
3	1	-0.5	0.25	-0.75	-0.125	0.625	-0.875
4	0.5	1	0.125	-0.625	0.25	-0.75	-0.375
5	0.625	-0.75	-0.375	0.5	0.875	-1	-0.125
6	0.75	0.625	-0.875	1	-0.375	0.5	-0.25
7	0.875	-0.375	0.75	0.25	-0.625	-0.125	1
8	0.375	0.875	0.625	0.125	0.75	0.25	0.5
9	0	0	0	0	0	0	0
10	-0.125	0.25	0.5	0.375	-1	-0.875	-0.625
11	-0.25	-0.125	1	0.875	0.5	0.375	-0.75
12	-1	0.5	-0.25	0.75	0.125	-0.625	0.875
13	-0.5	-1	-0.125	0.625	-0.25	0.75	0.375
14	-0.625	0.75	0.375	-0.5	-0.875	1	0.125
15	-0.75	-0.625	0.875	-1	0.375	-0.5	0.25
16	-0.875	0.375	-0.75	-0.25	0.625	0.125	-1
17	-0.375	-0.875	-0.625	-0.125	-0.75	-0.25	-0.5

Table 3.6. The $(O)_7^{17}$ design with the minimum ML_2 discrepancy.

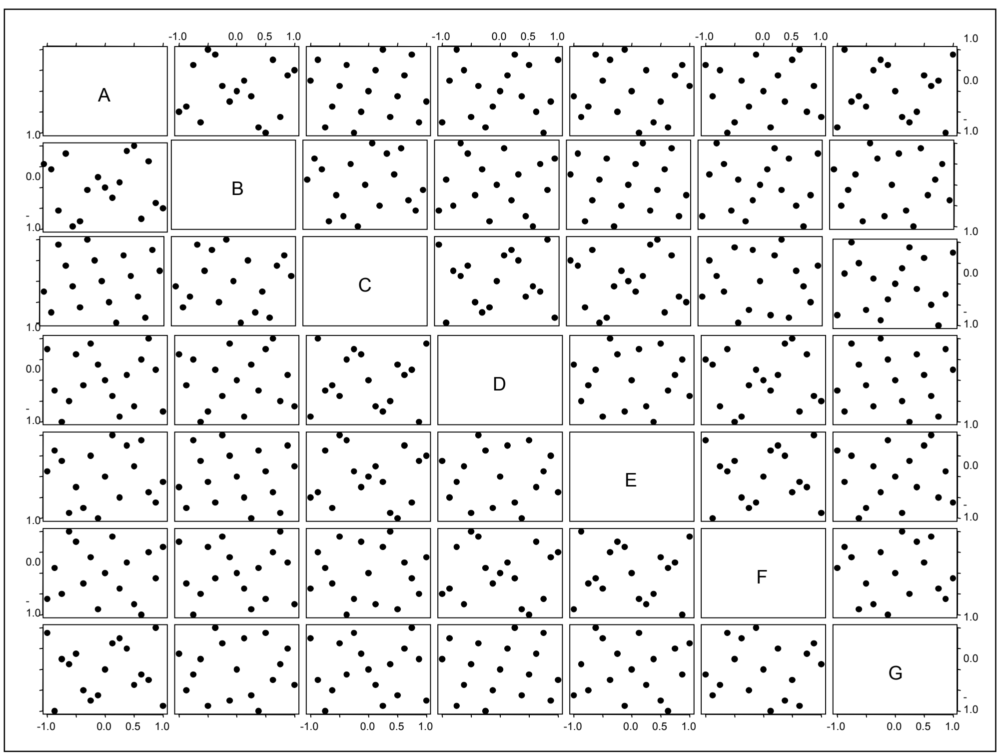

Figure 3.2. Two-dimensional projections of the best space-filling $(O)_7^{17}$ design.

The proposed $(O)_7^{17}$ design in Table 3.6 is orthogonal. Next let us ask how does the proposed design's space-filling measures (ML_2 discrepancy and Mm distance) compare with the optimal uniform design? A uniform design having seven variables and 17 levels is one of the few published optimal uniform designs (Fang et al. [2000]). It is expected that the uniform design will have a better Mm distance and ML_2 discrepancy since this is the major goal in their construction. A summary of the comparison between these designs is shown in Table 3.7.

	Max Pairwise Correlation	Condition Number	ML_2	Mm Distance
OLHC	0	1	0.151854	1.47902
Optimal Uniform	0.08088	1.35966	0.144309	1.61051

Table 3.7. Comparison of the orthogonality and space-filling properties of the OLHC and uniform 17-run, seven-variable designs.

Although the optimal uniform design enjoys an approximate five percent advantage in ML_2 discrepancy and an approximate eight percent advantage in Mm distance, the $(O)_7^{17}$ design has better orthogonality measures. Most notably, the condition number is 36 percent higher for the optimal uniform design. Furthermore, the $(O)_7^{17}$ design satisfies the desired characteristics and assumptions, but the uniform design fails to satisfy even the near orthogonality requirement.

B. CONSTRUCTING NEARLY ORTHOGONAL LATIN HYPERCUBES

This section describes the relaxation of strict orthogonality in order to achieve designs with improved space-filling properties. While one can find orthogonal Latin hypercubes for more than seven variables, the space-filling properties of these designs are quite poor. Therefore, for a specified combination of variables (more than seven) and runs, millions of candidate designs, which sacrifice some of their orthogonality, are generated by the computer and explored. For the most promising of these, a method (from Florian [1992]) to improve on their measures of near orthogonality is applied. From among a subset of those designs that are nearly orthogonal (i.e., have a maximum pairwise correlation no greater than 0.03 and a condition number no greater than 1.13), the design with the best combination of ML_2 discrepancy and Mm distance is chosen.

1. Achieving Near Orthogonality for Latin Hypercubes

Although $(O)_{11}^{33}$, $(O)_{16}^{65}$, and $(O)_{22}^{129}$ designs exist, their space-filling is poor. All permutations of the components of **e** for the $(O)_{7}^{17}$ design were generated in under 10 hours using a 1.0 GHz Pentium© 4 processor computer. Unfortunately, this enumerative approach is computationally difficult for the $(O)_{11}^{33}$, $(O)_{16}^{65}$, and $(O)_{22}^{129}$ designs. There are 16! permutations of **e** for the $(O)_{11}^{33}$ design, 32! permutations of **e** for the $(O)_{16}^{65}$ design, and 64! permutations of **e** for the $(O)_{22}^{129}$ design. To date, no other $(O)_{11}^{33}$, $(O)_{16}^{65}$, and $(O)_{22}^{129}$ designs (except for the ones constructed using Theorems 3.1 and 3.2) have been found.

After generating over one million random permutations of the elements of **e** in an attempt to find an $(O)_{11}^{33}$ design, over two million random permutations to find an $(O)_{16}^{65}$ design, and over three million random permutations to find an $(O)_{22}^{129}$ design, none of the generated designs even satisfied the requirements for near orthogonality. Table 3.8 shows the best maximum pairwise correlation and condition number found from these permutations. Note that the values in Table 3.8 do not occur for one single design matrix for the specified variables and levels.

Variables	Levels	Maximum Pairwise Correlation	Condition Number
11	33	0.033	1.11
16	65	0.146	1.85
22	129	0.159	2.38

Table 3.8. Best measures for designs, in terms of maximum pairwise correlation (a value of 0 is best) and condition number (a value of 1 is best), for selected variable and level combinations.

For more than seven variables (specifically 33 runs and 11 variables, 65 runs and 16 variables, and 129 runs and 22 variables), the designs generated by adding additional columns are either orthogonal (using Theorems 3.1 and 3.2) with poor space-filling or non-orthogonal. However, some of the non-orthogonal designs have good space-filling properties. Techniques for improving on the near orthogonality measures can be applied. Iman and Conover [1980] present a method that can reduce the correlation between input variables. Florian [1992] uses this same method to reduce the pairwise correlations between variables in a design matrix. Florian's procedure is adopted in order to decrease the maximum pairwise correlation. One minor weakness with this scheme is that it is possible that an original orthogonal variable pair can have small correlations induced by the computations. Although the maximum pairwise correlation is decreased, the number of orthogonal variable pairs may decrease as well. Since the correlations introduced to the original orthogonal variable pairs are typically small (e.g., less than .01), this trade-off is advantageous to the overall properties of the design matrix.

The net effect of Florian's procedure is that within one or more of the columns of the design matrix, the levels are permuted. This can result in a decreased maximum pairwise correlation without altering the actual levels.[16] There is a major distinction in how Florian's procedure is used. The procedures of both Iman and Conover and Florian examine only the correlations between pairs of variables. The present work includes the condition number as well.

Florian's [1992] method is now described. Each column element of the design matrix is replaced with the element's rank, $(1,2,\ldots,n)$, within the column. This $n \times k$ matrix is denoted by $\mathbf{W}$. Let $\mathbf{C}$ (a $k \times k$ matrix) represent the rank correlation matrix of $\mathbf{W}$. If each pair of columns in $\mathbf{W}$ is uncorrelated, then $\mathbf{C}$ is equal to the unit matrix $\mathbf{I}$ ($k \times k$ matrix). Only those realizations of $\mathbf{W}$ for which matrix $\mathbf{C}$ is positive definite are considered. The basic idea is to transform $\mathbf{W}$ into a set of uncorrelated variates. A Cholesky factorization scheme is used (since $\mathbf{C}$ is positive definite) to determine a lower triangular matrix, $\mathbf{Q}$, which is $k \times k$. Then, let $\mathbf{D}=\mathbf{Q}^{-1}$ and $\mathbf{C}=\mathbf{Q}*\mathbf{Q}^{T}$ such that $\mathbf{D}$ has the property

$$\mathbf{D}*\mathbf{C}*\mathbf{D}^{T} = \mathbf{I}. \tag{3.4}$$

The original $\mathbf{W}$ is then transformed into a new matrix, $\mathbf{W}_{B}$ ($n \times k$ matrix), using

$$\mathbf{W}_{B} = \mathbf{W}*\mathbf{D}^{T}. \tag{3.5}$$

---

[16] Other methods (i.e., cosine-sine decomposition and Gram-Schmidt orthogonalization) can alter the levels.

Since the elements of the matrix $\mathbf{W}_B$ are not necessarily integral, the elements in each column are replaced by their rank order (1,2,…,n).

As proven by Iman and Conover [1980], the difference between appropriate elements in the rank correlation matrix of $\mathbf{W}_B$ and $\mathbf{I}$ is lower than in the case of matrix $\mathbf{W}$ and $\mathbf{I}$. Since the elements of $\mathbf{W}_B$ are replaced by ranks, this process can be repeated. We do so until there is no further decrease in the maximum pairwise correlation. Finally, to reconstruct the Latin hypercube design matrix, the ordered ranks in the final $\mathbf{W}_B$ are then mapped back into the original input variable values. Appendix A contains an example of these calculations.

As previously noted, Iman and Conover [1980] and Florian [1992] used this scheme and focused only on a correlation measure. The condition number serves to improve the process for the following reason. As this procedure is performed on numerous matrices, it is quite common that although the maximum pairwise correlation value does not change, the condition number continues to decrease. Thus, if the procedure uses only the maximum pairwise correlation value, then this iteration process may stop too soon, even though a better design matrix (in terms of both maximum pairwise correlation and condition number) may exist. Additionally, this procedure can only provide limited improvement for the maximum pairwise correlation and condition number. Initialization using a screening value (found by exploratory trial and error) for the maximum pairwise correlation and condition number speeds the process and dramatically enhances the non-orthogonality measures of the final design matrix. Florian's method is applied to only those Latin hypercubes that achieve the screening non-orthogonality measures.

2. An Algorithm for Constructing Nearly Orthogonal Latin Hypercubes

This section contains a method for constructing nearly orthogonal Latin hypercubes for $k > 7$ that satisfy the desirable design characteristics. Specifically, this method is appropriate for designs having eight to 11 variables and 33 levels, 12 to 16 variables and 65 levels, or 17 to 22 variables and 129 levels.

The proposed experimental designs with near orthogonality will be denoted by $(N_O)_k^n$, where N_O represents near orthogonality, n represents the number of runs or experiments, and k represents the number of variables. Recall that these designs must have a maximum pairwise correlation no greater than 0.03 and condition number no greater than 1.13.

Designs are generated using the extension of Ye's [1998] algorithm discussed in this chapter. Since no orthogonal designs (except for those generated using Theorems 3.1 and 3.2) have been found, the strict orthogonality requirement for initializing the process is removed. Instead, near orthogonality is the goal. Random permutations of **e** are used to generate proposed designs. Since Florian's [1992] procedure can provide limited improvement, only those designs satisfying a pre-set maximum threshold pairwise correlation, ρ, and condition number are retained. Later in the chapter, guidance on the pre-set values to choose is given. Florian's [1992] method is applied to those designs achieving the pre-set values. The values specified are such that after the designs are subjected to Florian's [1992] procedure, the resulting designs are nearly orthogonal. Of the nearly orthogonal designs, their space-filling properties are compared. The candidate design with the most desirable combination of Mm distance and ML_2 discrepancy is chosen.

The algorithm for finding a nearly orthogonal Latin Hypercube (NOLHC) experimental design having eight to 22 variables is summarized.

- Step 1. Determine the number of variables (k>7) required for experimentation. If the number of variables is other than 11, 16, or 22, round the required number of variables up to the nearest one of these numbers.
- Step 2. Establish a maximum threshold pairwise correlation value, ρ, and a maximum threshold condition number.
- Step 3. Using a randomly permuted **e**, construct a design matrix as previously described in this chapter.
- Step 4. Calculate the pairwise correlations and the condition number.
- Step 5. If any of the values in Step 4 exceed the thresholds in Step 2, discard the design and go to Step 3 with a randomly permuted **e** (with replacement). Otherwise, keep the design and proceed to Step 6. Repeat Steps 3-5 until a desired pre-set number of candidate designs are found.

- Step 6. Subject each of the candidate designs to Florian's [1992] method of factorization to decrease the maximum pairwise correlation and condition number.

- Step 7. Calculate the Mm distance and ML_2 discrepancy for each of the Step 6 designs. Rank the designs according to these measures. Choose the design with the minimum rank sum over the two measures.

- Step 8: If a number of variables other than seven, 11, 16, or 22 is required, construct each of the possible combination of columns (having the appropriate number of desired variables) from the Step 7 design and calculate the Mm distance and ML_2 discrepancy. Choose the design with the minimal rank sum over the two measures.

The reader is reminded that except for the $(O)_7^{17}$ design, there is no guarantee that the designs generated from this algorithm are globally optimal. Conversely, the designs do have near orthogonality and excellent space-filling properties. The designs are easy to generate (recommended designs for up to 22 variables are provided later in this chapter). The statistical analysis of results is facilitated since the estimates of linear effects of each variable are nearly uncorrelated and the cases are well scattered throughout the experimental region. Finally, prior to the experiment, there are no assumptions made concerning which variables may be correlated (e.g., Iman and Conover [1980]) or what distribution the response function will have from the variable's settings (e.g., Currin et al. [1998], Clyde et al. [1996]). In essence, the desirable design characteristics are satisfied save the issue of promoting insensitivity to premature experiment termination. This issue is discussed later in this chapter.

C. ORTHOGONAL AND NEARLY ORTHOGONAL LATIN HYPERCUBE DESIGNS FOR UP TO 22 VARIABLES

This section presents the best designs that have been generated using the algorithm from the previous section. This provides the reader with ready-to-use orthogonal or nearly orthogonal Latin hypercube designs for two to 22 variables.

1. Orthogonal Latin Hypercubes for Two to Seven Variables

This section provides the best space-filling $(O)_7^{17}$ design and the best designs derived from this $(O)_7^{17}$ design having fewer than seven variables. The $(O)_7^{17}$ design was extensively covered earlier in this chapter. Table 3.6 and Figure 3.2 summarize these findings. Table 3.9 generalizes Table 3.6 in that the entries of Table 3.9 indicate the ordinal level of that particular variable.

If fewer than seven variables are required, then selected columns can be removed from the original seven variable design matrix (Table 3.9) to correspond to the desired number of variables, while still maintaining good space-filling properties (e.g., if only five variables are required, then two columns are removed, such that the remaining 17-run, five-variable design matrix has good space-filling properties). As stated in the algorithm, all possible combinations of columns are examined from Table 3.9 by calculating the Mm distance and ML_2 discrepancy. The design with the minimal rank sum over the two measures is chosen.[17] Table 3.10 summarizes the results for the 17-run case when two to six variables are desired.

Run	Variable A	Variable B	Variable C	Variable D	Variable E	Variable F	Variable G
1	10	7	5	6	17	16	14
2	11	10	1	2	5	6	15
3	17	5	11	3	8	14	2
4	13	17	10	4	11	3	6
5	14	3	6	13	16	1	8
6	15	14	2	17	6	13	7
7	16	6	15	11	4	8	17
8	12	16	14	10	15	11	13
9	9	9	9	9	9	9	9
10	8	11	13	12	1	2	4
11	7	8	17	16	13	12	3
12	1	13	7	15	10	4	16
13	5	1	8	14	7	15	12
14	4	15	12	5	2	17	10
15	3	4	16	1	12	5	11
16	2	12	3	7	14	10	1
17	6	2	4	8	3	7	5

Table 3.9. The $(O)_7^{17}$ design with ordinal levels for the variables.

[17] Of course, the reader can use other criteria to select between competing designs.

Desired Variables	Deleted Columns	Maximum Pairwise Correlation	Condition Number	Mm Distance	ML_2
6	1	0	1	1.43069	0.078914
5	1, 6	0	1	1.26861	0.038799
4	1, 3, 6	0	1	1.03078	0.01725
3	1, 2, 3, 6	0	1	0.57282	0.007273
2	1, 3, 4, 6, 7	0	1	0.51539	0.002525

Table 3.10. Orthogonal designs for fewer than seven variables derived from the $(O)_7^{17}$ design.

The assumption is that using the $(O)_7^{17}$ design to construct designs with fewer variables will result in acceptable designs that are nearly orthogonal and have acceptable space-filling properties. The validity of this assumption is illustrated in the case of a design with two variables and 17 levels. Specifically, comparisons between the $(O)_2^{17}$ design, the published uniform design of Fang and Wang [1994], and the design with the best Mm distance measure (Morris and Mitchell [1992], [1995]) are made. The $(O)_2^{17}$ design fares extremely well against the two optimal designs with respect to their optimality criteria, as shown in Table 3.11.

	Maximum Correlation	Condition Number	Mm Dist	ML_2
$(O)_2^{17}$ design	0	1	0.51539	0.002525
Uniform design	0	1	0.27905	0.002201
Best Mm distance design	0.0588	1.125	0.53033	0.002354

Table 3.11. Comparison of the proposed, uniform, and best Mm distance designs for the 17-run and two-variable case.

For orthogonality measures, a maximum pairwise correlation of 0 and condition number of 1 are the best measures. The $(O)_2^{17}$ design and uniform designs from Table 3.11 are orthogonal, but the best Mm distance design is not orthogonal. For the space-filling measures, a larger value for Mm distance is better (in this case, the measures can range from 0 to 0.53033) and a smaller value for ML_2 discrepancy is better (in this case, the measures can range from 0.002201 to 0.7778). Although the best Mm distance design has approximately a three percent better Mm distance and approximately a seven percent better ML_2 discrepancy than the $(O)_2^{17}$ design, the $(O)_2^{17}$ design is orthogonal, but the best Mm distance design fails to satisfy near orthogonality. Furthermore, the $(O)_2^{17}$ design has a 46 percent better Mm distance than the uniform design, while only a 13 percent poorer ML_2 discrepancy.

2. Nearly Orthogonal Latin Hypercubes for Eight to 11 Variables

This section describes the construction of the best $(N_O)_{11}^{33}$ design and the best associated designs with fewer variables. An exhaustive search of the 16! designs was not attempted. Instead, using the design construction discussed previously, approximately one million randomly selected vectors **e** were used to find 15 (a pre-set number) designs satisfying a maximum threshold ρ value of .05 and maximum threshold condition number of 1.15 (these threshold values were chosen using exploratory trial and error). These 15 designs were then subjected to Florian's [1992] procedure to reduce the maximum pairwise correlation and condition number. These designs achieved a maximum pairwise correlation no greater than 0.03 and a condition number no greater than 1.13, satisfying the near orthogonality criteria. These 15 designs were than compared using Mm distance and the ML_2 discrepancy and are shown in Table 3.12. Note that all of these designs are practically indistinguishable in terms of correlations and condition numbers.

Design 15 corresponds to the orthogonal design using Theorems 3.1 and 3.2. Although this design is orthogonal, it has the worst ML_2 discrepancy. Design 6 is chosen as the best design since it has the minimal rank sum (best Mm distance and second-best ML_2 discrepancy). Its maximum correlation is 0.0234 and condition number is 1.123. The appropriate levels for this design are shown in Appendix B. Figure 3.3 displays the two-dimensional projections of this nearly orthogonal design. Since the author is unaware of any published literature on uniform designs with this number of variables and levels, no comparison can be made, but the proposed design does exhibit excellent orthogonality and space-filling properties.

As a means of comparison, 1,000 Latin hypercubes with 11 variables, each with 33 levels, are generated. These 1,000 designs have an average maximum pairwise correlation of 0.4015, average condition number of 8.315, average Mm distance of 1.105, and average ML_2 discrepancy of 0.8117. The nearly orthogonal design is considerably better in all measures than an average Latin hypercube.

Design Number	Mm Distance	ML_2	Mm Distance Rank	ML_2 Rank	Rank Sum
1	1.6262	0.74	7	3	10
2	1.317	0.77	14	7	21
3	1.6724	0.77	3	10	13
4	1.3793	0.78	13	11	24
5	1.7139	0.75	2	4	6
6	**1.7578**	**0.73**	**1**	**2**	**3**
7	1.6618	0.75	5	5	10
8	1.6117	0.73	9	1	10
9	1.2885	0.77	15	8	23
10	1.513	0.76	12	6	18
11	1.6441	0.92	6	14	20
12	1.6154	0.77	8	9	17
13	1.5487	0.8	11	13	24
14	1.5737	0.79	10	12	22
15	1.6713	0.95	4	15	19

Table 3.12. Candidate $(N_O)_{11}^{33}$ designs showing the corresponding space-filling measures and ranks. Each of the designs has a maximum pairwise correlation less than 0.03 and condition number less than 1.13.

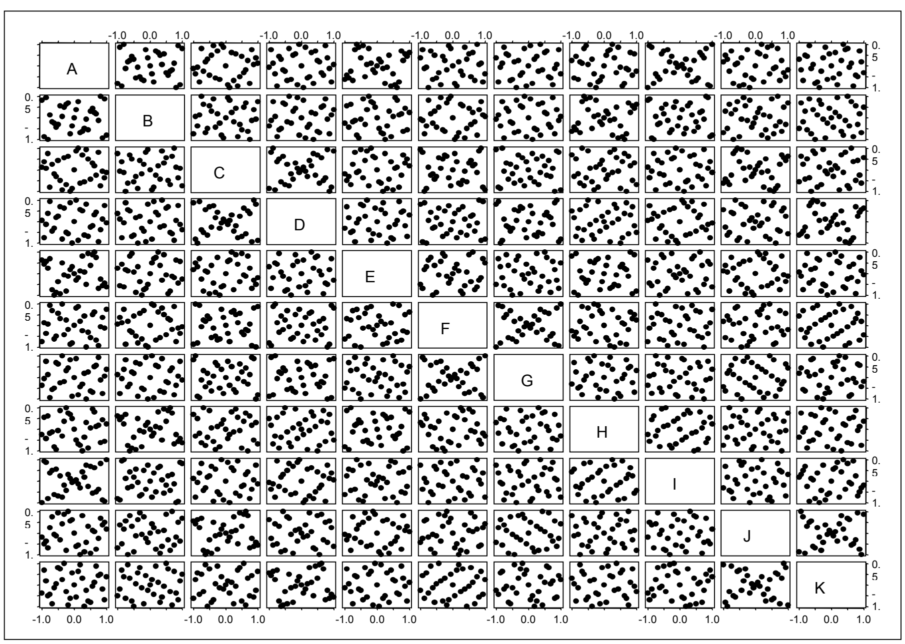

Figure 3.3. Two-dimensional projections of columns for the best $(N_O)_{11}^{33}$ design.

Designs containing between eight and 10 variables are now considered. Each of the possible combinations of columns from Appendix B is examined by calculating the Mm distance and ML_2 discrepancy as columns are deleted. Table 3.13 summarizes the results for the 33-run case for between eight to 10 variables. Although Ye [1998] states that an orthogonal design exists for 33 runs and eight variables, a good space-filling design has not been found, and none was shown by Ye. Table 3.13 provides a readily available alternative that has good orthogonality and space-filling properties.

Desired Variables	Deleted Columns	Maximum Pairwise Correlation	Condition Number	Mm Distance	ML_2
10	1	0.0234	1.112	1.70478	0.412687
9	8, 10	0.0234	1.1	1.51167	0.229329
8	1, 2, 10	0.0234	1.089	1.42522	0.124826

Table 3.13. Nearly orthogonal designs for fewer than 11 variables derived from the $(N_O)_{11}^{33}$ design.

3. Nearly Orthogonal Latin Hypercubes for 12 to 16 Variables

The construction of the best $(N_O)_{16}^{65}$ design and the best associated designs with fewer variables is described. An exhaustive search of the 32! designs was not attempted. Instead, using the design construction discussed previously, approximately two million randomly selected vectors of **e** were used to find 15 designs satisfying a maximum threshold ρ value of 0.17 and maximum threshold condition number of 2.4 (these threshold values were chosen by exploratory trial and error). These 15 (a pre-set number) designs were subjected to Florian's [1992] procedure to reduce the maximum pairwise correlation and condition number. These designs achieved a maximum pairwise correlation no greater than 0.022 and a condition number no greater than 1.11, satisfying the near orthogonality criteria. These 15 designs were then compared using the Mm distance and the ML_2 discrepancy and are shown in Table 3.14.

Design Number	Mm Distance	ML_2	Mm Distance Rank	ML_2 Rank	Rank Sum
1	1.7941	7.98	8	15	23
2	1.6759	5.4	14	14	28
3	1.6247	4.6	15	5	20
4	1.7741	4.64	9	8	17
5	1.8408	4.71	6	10	16
6	1.8949	4.99	4	13	17
7	1.7402	4.52	12	3	15
8	1.7727	4.87	10	12	22
9	1.8496	4.64	5	7	12
10	2.0146	4.59	2	4	6
11	1.7675	4.81	11	11	22
12	**2.0353**	**4.46**	**1**	**1**	**2**
13	1.7205	4.7	13	9	22
14	1.8219	4.63	7	6	13
15	1.9939	4.48	3	2	5

Table 3.14. Candidate $(N_O)_{16}^{65}$ designs showing the corresponding space-filling measures and ranks. Each of the designs has a maximum pairwise correlation less than 0.022 and condition number less than 1.11.

Design 1 corresponds to the orthogonal design using Theorems 3.1 and 3.2. Although this design is orthogonal, it has the worst ML_2 discrepancy. Design 12 is chosen as the best design since it has the best Mm distance and best ML_2 discrepancy. Its maximum correlation is 0.0219 and condition number is 1.103. The appropriate levels for this design are shown in Appendix C. Since the author is unaware of any published literature on uniform designs with this number of variables and levels, no comparison can be made, but the proposed design does exhibit excellent orthogonality and space-filling properties.

As a means of comparison, 1,000 Latin hypercubes with 16 variables, each with 65 levels, are generated. These 1,000 Latin hypercubes have an average maximum pairwise correlation of 0.3194, average condition number of 6.103, average Mm distance of 1.647, and average ML_2 discrepancy of 5.372. The nearly orthogonal design is substantially better in all measures. The cases where fewer than 16, but more than 11 variables are required is considered. Each of the possible combination of variables from Appendix C is examined by calculating the Mm distance and the ML_2 discrepancy as variable columns are deleted. Table 3.15 summarizes the results for the 65-run case when 12 to 15 variables are desired. Although Ye [1998] states that an orthogonal design exists for 65 runs and 10 variables, a good space-filling design has not been found, and none was shown by Ye. Table 3.15 provides a readily available alternative that has good orthogonality and space-filling properties.

Desired Variables	Deleted Columns	Maximum Pairwise Correlation	Condition Number	Mm Distance	ML_2
15	2	0.02194	1.097	2.03149	2.69304
14	7, 10	0.01844	1.0838	1.95456	1.59995
13	9, 10, 13	0.02194	1.0889	1.90497	0.95337
12	4, 7, 9, 10	0.01809	1.079	1.83259	0.56767

Table 3.15. Nearly orthogonal designs for fewer than 16 variables derived from the $(N_O)_{16}^{65}$ design.

4. Nearly Orthogonal Latin Hypercubes for 17 To 22 Variables

This section describes the construction of the $(N_O)_{22}^{129}$ design and associated designs with fewer variables. An exhaustive search of the 64! designs was not attempted. Instead, using the design construction discussed previously, approximately three million randomly selected vectors of **e** were used to find 15 designs satisfying a maximum threshold ρ value of 0.16 and maximum threshold condition number of 2.8 (these threshold values were found by trial and error). These 15 (a pre-set number) designs were then subjected to Florian's [1992] procedure to reduce the maximum pairwise correlation and condition number. These designs achieved a maximum pairwise correlation no greater than 0.01 and a condition number no greater than 1.04, satisfying the near orthogonality criteria. These 15 designs were then compared using Mm distance and ML_2 discrepancy and are shown in Table 3.16.

Design Number	Mm Distance	ML_2	Mm Distance Rank	ML_2 Rank	Rank Sum
1	2.2386	38.4	2	4	6
2	1.8132	45.2	10	12	22
3	1.6386	38.6	14	5	19
4	2.0433	39	6	6	12
5	1.866	41.6	9	9	18
6	2.075	35.8	5	1	6
7	1.8899	47.9	8	14	22
8	**2.2655**	**37.8**	**1**	**2**	**3**
9	1.6129	43.7	15	10	25
10	2.1184	39.6	4	7	11
11	1.7885	96.6	12	15	27
12	1.9265	45.4	7	13	20
13	2.1907	38.1	3	3	6
14	1.8	40	11	8	19
15	1.6796	44	13	11	23

Table 3.16. Candidate $(N_O)_{22}^{129}$ designs showing the corresponding space-filling measures and ranks. Each of the designs has a maximum pairwise correlation less than 0.01 and condition number less than 1.04.[18]

Design 11 corresponds to the orthogonal design using Theorems 3.1 and 3.2. Although this design is orthogonal, it has the worst ML_2 discrepancy. Design 8 is chosen as the best design since it has the best Mm distance and the second best ML_2 discrepancy. Its maximum correlation is 0.0074 and condition number is 1.039. The appropriate levels for this design are shown in Appendix D. Since the author is unaware of any published literature on uniform designs with this number of variables and levels, no comparison can be made, but the proposed design does exhibit excellent orthogonality and space-filling properties.

As a means of comparison, 1,000 Latin hypercubes with 22 variables, each with 129 levels, are generated. These 1,000 Latin hypercubes have an average maximum pairwise correlation of 0.2332, average condition number of 4.073, average Mm distance

[18] Note that the ML_2 discrepancy measures are much larger than those exhibited earlier. Fang and Wang [1994] find similar high discrepancy measures when attempting to find designs with 20 or more variables and attribute it to the sparseness of design points in high-dimensional regions.

of 1.899, and average ML_2 discrepancy of 59.773. The nearly orthogonal design is better in all measures than the average Latin hypercube.

The cases where fewer than 22, but more than 16 variables are required is considered next. Each of the possible combination of columns from Appendix D are examined by calculating the Mm distance and the ML_2 discrepancy as the columns are deleted. Table 3.17 summarizes the results for the 129-run case when 17 to 21 variables are desired. Although Ye [1998] states that an orthogonal design exists for 129 runs and 12 variables, a good space-filling design has not been found, and none was shown by Ye. Table 3.17 provides an alternative that has good orthogonality and space-filling properties.

Desired Variables	Deleted Columns	Maximum Pairwise Correlation	Condition Number	Mm Distance	ML_2
21	1	0.0074	1.0376	2.22446	23.17738
20	1, 5	0.0074	1.0372	2.20689	14.35779
19	1, 5, 20	0.0074	1.035	2.13806	8.86844
18	1, 5, 20, 21	0.0074	1.0345	2.09358	5.42232
17	1, 5, 7, 16, 20	0.0074	1.0326	2.01065	3.38073

Table 3.17. Nearly orthogonal designs for fewer than 22 variables derived from the $(N_O)_{22}^{129}$ design.

D. GENERATING ADDITIONAL DESIGN POINTS

Section C contains a set of orthogonal and nearly orthogonal Latin hypercubes that allow one to explore from two to 22 variables in a given number of runs (17, 33, 65, or 129). In this section, the following question is addressed: If an analyst can take more runs, how should one do so? This question is also related to the issue of premature experiment termination. The assumption here is that the termination cannot occur after an arbitrary number of runs, but rather at epochs in the number of runs marking the completion of specified blocks of runs

1. Sequential Approach to Selecting Run Blocks

This section discusses why a sequential approach is used in selecting the blocks of runs. Specifically, the algorithm selects blocks of additional runs (of sizes 16, 32, 64, and 128), such that the near orthogonality is retained, while the space-filling properties are improved. The algorithm is presented in the context of a sequential analysis, though it applies equally well if all of the runs are made at once. This is done, in part, because this is how the algorithm is used in Chapters IV and V. Specifically, an experiment is conducted, and then the results are analyzed. Another experiment is completed, and the results are analyzed to see if the hypotheses generated from the first experiment are supported by the second experiment, and so on. This procedure is similar to a cross-validation procedure. When the analyst is satisfied with the results, no further experiments need be conducted.

For example, assume that a $(O)_7^{17}$ design is executed, the entire experimental region is examined, and interim results obtained. An additional 16 runs might then be identified, executed, and cross-validated with the first 17 runs. This sequence permits sound, interim results to be obtained if premature termination (compatible with these constraints) occurs. That is, if the second set of 16 runs cannot be made, the initial runs are orthogonal. This approach also allows for a systematic, sequential approach to analyzing the relationship between the variables and the output measure of interest of the model.

There is another advantage to this sequential approach—region reduction. This permits the experimenter to adjust, if necessary, the levels of a particular variable after the first set of runs. Since the variables are continuous, a variable found to have no effect on the measure of interest may be finely partitioned into a narrower range of values, provided the new values maintain the equidistant property. Thus, it is not possible to use this approach to reduce the region of a variable that has an effect on the measure of interest at the variable's lower and upper values, but not at its middle values.

As an example, assume an initial $(O)_7^{17}$ design is executed where each of the variables are continuous from -1 to 1, with 17 distinct values (-1,-0.875,-0.75,…,0.875,1). Suppose that during the analysis, it is found that the measure of interest is stable for the largest 11 values (-0.25 to 1) of one of the variables. The experimenter has a choice. He/she may decide to keep that variable at all of the original 17 levels (less the 9^{th} level which corresponds to the center point) for the next set of 16 runs; or he/she may opt to not sample the ineffective region, and instead use a finer partition to explore the region from -1 to -0.25 and rescaled[19], being careful to maintain the equidistant property. In this case, the 16 new set of levels would range from –1 to -0.25 in increments of 0.05. Thus, in addition to information being gained concerning the relationship between the variables and measure of interest, the experimental region has been reduced in order to focus on those areas of importance that were suggested by the first set of runs.

2. Column Permuting and Appending Heuristic

The major issue is how to generate additional design points from the original design matrix such that orthogonality (in the case of seven or fewer variables) or near orthogonality (for more than seven variables) is maintained and space-filling improved. This section describes the implementation of a permuting and appending procedure on the columns.

The original design matrix has its columns permuted.[20] This permuted design matrix is then appended vertically to the original design matrix. The center point run is redundant and not repeated. If n was the initial number of runs in the design matrix, then the number of runs is increased by $n–1$ (the original center point is omitted from the additional points) in the subsequent set. The encouraging result, which is summarized in Theorem 3.3, is the likely reduction in the maximum pairwise correlation. In practice, the condition number is also non-increasing. Although the theorem indicates non-increasing values instead of decreasing values, in practice, the values are typically decreasing.

> Theorem 3.3. By permuting the columns of the original NOLHC containing n runs and appending these columns to the original NOLHC, the number of runs is increased to $(2n–1)$, and the maximum pairwise correlation is non-increasing.

Proof: Recall from (1.4) that the correlation between two columns in a design matrix, $\boldsymbol{v}=[v_1,v_2,\ldots,v_n]^T$ and $\boldsymbol{w}=[w_1,w_2,\ldots,w_n]^T$, is defined to be

[19] Although the level of -0.375 was found to be influential on the measure of interest and -0.25 was found not to be influential, the new partition should include the region from -0.375 to -0.25 to ensure better exploration.

[20] The permutation of the columns of a design matrix does not affect its space-filling.

$$r(\mathbf{v},\mathbf{w}) = \frac{\sum_{i=1}^{n}\left[(\nu_i - \overline{\nu})(\omega_i - \overline{\omega})\right]}{\sqrt{\sum_{i=1}^{n}(\nu_i - \overline{\nu})^2 \sum_{i=1}^{n}(\omega_i - \overline{\omega})^2}}. \tag{3.6}$$

Furthermore, without loss of generality, we consider the absolute value of (1.4) and (3.6) in determining the maximum pairwise correlation. For a sample size of n, the values in the columns of our Latin hypercubes take the integer values from $(-n+1)/2$ to $(n-1)/2$. Thus, for any column $\mathbf{v}$, $\overline{v} = 0$ and $\sum_{i=1}^{n} v_i^2 = \frac{(n-1)n(n+1)}{12}$. Therefore, for any two columns of $\mathbf{v}$ and $\mathbf{w}$,

$$r(\mathbf{v},\mathbf{w}) = \frac{\sum_{i=1}^{n} \nu_i \omega_i}{\frac{(n-1)n(n+1)}{12}}. \tag{3.7}$$

Now, assume that the columns of the design matrix are permuted and append the permuted matrix to the bottom of the initial design matrix to create the new, expanded design matrix. The new columns consist of $n + (n-1)$ entries (we do not include a replicate center point in the permuted matrix). Suppose columns $\mathbf{x}$ and $\mathbf{y}$ are appended to $\mathbf{v}$ and $\mathbf{w}$, respectively. Then, the new correlation between the two columns is

$$r_{new}(\mathbf{v{:}x},\mathbf{w{:}y}) = \frac{\sum_{i=1}^{n} \nu_i \omega_i + \sum_{i=1}^{n-1} x_i y_i}{\frac{(n-1)n(n+1)}{6}}. \tag{3.8}$$

Note that the denominator of $r_{new}(\mathbf{v{:}x},\mathbf{w{:}y})$ is twice that of $r(\mathbf{v},\mathbf{w})$. Without loss of generality, suppose that maximum pairwise correlation is greater than or equal to the negative of the minimum pairwise correlation. Also, suppose that $r(\mathbf{v},\mathbf{w}) = \rho$, where ρ is the maximum pairwise correlation. Then, $r(\mathbf{x}, \mathbf{y}) \leq r(\mathbf{v},\mathbf{w})$, and therefore, $r_{new}(\mathbf{v{:}x},\mathbf{w{:}y}) \leq r(\mathbf{v},\mathbf{w})$. □

Since the original experimental design is nearly orthogonal, the maximum pairwise correlation value and condition number are generally only marginally improved. Thus, when selecting columns to permute it seems wise to emphasize space-filling. Although other nearly orthogonal designs could be appended to the original design matrix, the choice here is to permute and append the columns of the original design matrix based upon their space-filling properties.

In the $(O)_7^{17}$ design, an exhaustive enumeration of the column permutations (7!) is possible. In finding the best permutation of columns to be appended, the rank sum of the Mm distance and the ML_2 discrepancy are used in the same way that is done (see Section C of this chapter) when seeking columns to delete.

An exhaustive enumeration of the column permutations for the $(N_O)_{11}^{33}$, $(N_O)_{16}^{65}$, and $(N_O)_{22}^{129}$ designs is not feasible. One possibility is to sample randomly from the possible permutations, rank order the resulting designs for their Mm distances and ML_2 discrepancies, and choose the permutation design with the smallest rank sum. To do this more efficiently, a heuristic is used to narrow the possible permutations for the random sampling.[21] This is achieved as follows.

The ML_2 discrepancy is calculated for each combination of three variables (e.g., in the $(N_O)_{11}^{33}$ design, there are $\binom{11}{3}$ = 165 combinations). The ML_2 discrepancies are then rank ordered from highest (worst space-filling) to lowest (best space-filling). The number of times each variable appears in a combination having a high ML_2 discrepancy (e.g., in the $(N_O)_{11}^{33}$ design, this is the upper half of the 165 measures, which corresponds to 82 measures, since the midpoint is omitted) is compared to the number of times each variable appears in a combination having a low ML_2 discrepancy (e.g., in the $(N_O)_{11}^{33}$ design, this is the lower half of the 165 measures, which corresponds to 82 measures, since the midpoint is omitted). Under the assumption that a variable has an equal probability of appearing in either the upper half or lower half, an exact binomial test (Conover [1999]) at the 0.10 significance level is performed to identify those variables which are more likely to appear in the better combinations and those variables which are more likely to appear in the poorer combinations. The good variables are then

[21] Other heuristics are possible. This one is used because it performs well in the cases examined.

restricted to being appended to those variables that are the poorest performing.[22] The use of this heuristic appears to provide additional design points that improve both near orthogonality and space-filling.

Three variable combinations are chosen since three-way interactions of this type in regression analysis are somewhat possible to explain. Higher order interactions are more difficult to interpret. A significance level of 0.10 is chosen (over, say 0.05) to permit a greater number of variables to be identified as good and poor performers and to reduce the total number of required permutations. Of course, others can choose their own levels. Finally, in all of the cases detailed below, the heuristic has been able to identify a best (though not necessarily globally optimal) permutation, whereas random sampling has not found a better permutation in a like (or greater) number of attempts.

3. Application of the Column Permuting and Appending Heuristic to Selected Designs

This section provides the suggested column permuting and appending schemes for the $(O)_7^{17}$, $(N_O)_{11}^{33}$, $(N_O)_{16}^{65}$, and $(N_O)_{22}^{129}$ designs from Section C. The heuristic may be repeated to generate additional blocks of runs.

a. The $(O)_7^{17}$ Design

For the $(O)_7^{17}$ design, a complete enumeration is possible. The best possible permutation of the original columns (variables) from Table 3.13 is 2, 6, 4, 7, 1, 5, and 3. For example, the first column of Table 3.9 is appended with the second column of Table 3.9 (less the center point corresponding to level 9), the second column of Table 3.9 is appended with the sixth column of Table 3.9, and so on. This permutation achieves the best rank sum for Mm distance and *ML_2* discrepancy.

When the columns are appended, the resulting design is an $(O)_7^{33}$ design. The design matrix has an Mm distance of approximately 1.2, as compared to the original 1.479. This follows since additional design points are being added to the region, so the decrease is expected. Conversely, the *ML_2* discrepancy decreases from 0.15184 to

[22] Computational experiments indicate that additionally restricting the columns to which the poor performing variables are appended is not beneficial. Combining these additional design points with the original design points does not yield the best space-filling design.

0.09149, indicating that the design points achieve a greater degree of space-filling over the region.

b. *The* $(N_O)_{11}^{33}$ *Design*

For the $(N_O)_{11}^{33}$ design in Appendix B, the seventh column is identified as a good performing variable since the *p*-value associated with its binomial tests is less than 0.001. There is no variable identified as a poor performer having a *p*-value less than 0.10. The poorest performing variables are the first and eighth columns since they each appear 11 more times in poor performing combinations than in good performing combinations (*p*-value = 0.135). Thus, to alleviate some computational burden, the seventh column is restricted to appending to either the first or eighth columns.

With this restriction, there are 11! possible permutations of the columns. By restricting where the seventh column is appended, the required permutations decreases from almost 40 million to approximately 7.2 million (a decrease of over 81 percent). Two million permutations were done for the unrestricted case and one million permutations were done for the restricted case. The best (not necessarily globally optimal) permutation was found from the restricted permutations and had the permuted column ordering of 11, 1, 6, 8, 2, 9, 10, 7, 3, 4, and 5.

The resulting $(N_O)_{11}^{65}$ design has a Mm distance of 1.363 (compared to the original 1.758) and improved ML_2 discrepancy of 0.36905 (compared to the original 0.73182). The design has a non-increasing maximum pairwise correlation (0.0234) and condition number (1.13). Thus, the additional design points are added in such a way that the near orthogonality is not jeopardized, but space-filling is improved.

c. *The* $(N_O)_{16}^{65}$ *Design*

For the $(N_O)_{16}^{65}$ design in Appendix C, the twelfth column is identified as a good performing variable since its *p*-value is less than 0.032 from the exact binomial test. The seventh column is the poorest performer since it appears 19 more times in poor performing combinations than in good performing combinations and has a *p*-value less than 0.079. Thus, the twelfth column is restricted to appending to the seventh column.

With this restriction, there are 16! possible permutations of the columns. By restricting where the twelfth column is appended, the required number of permutations decreases approximately 94 percent. Three million permutations were done for the unrestricted case and 1.5 million permutations were done for the restricted case. The best (not necessarily globally optimal) permutation was found from the restricted permutations and had the permuted column ordering of 2, 3, 8, 13, 16, 5, 12, 7, 1, 14, 9, 15, 11, 10, 6, and 4.

The resulting $(N_O)_{16}^{129}$ design has a Mm distance of 1.91 (compared to the original 2.035) and improved ML_2 discrepancy of 2.282 (compared to the original 4.465). The design has a non-increasing maximum pairwise correlation (0.0291) and condition number (1.103). Thus, the additional design points are added in such a way that the near orthogonality is not jeopardized, but space-filling is improved.

d. The $(N_O)_{22}^{129}$ Design

For the $(N_O)_{22}^{129}$ design in Appendix D, the third and fifteenth columns are identified as the best performing variables with *p*-values less than 0.023 from the exact binomial test. The first, seventh, tenth, and nineteenth columns are the poorest performers as they all have *p*-values less than 0.085. Thus, the third and fifteenth columns are restricted to appending to one of these four poor performing variables. Four million permutations were done for the unrestricted case and two million permutations were done for the restricted case. The best (not necessarily globally optimal) permutation was found from the restricted permutations and had the permuted column ordering of 3, 16, 20, 11, 9, 19, 4, 14, 12, 15, 22, 8, 1, 5, 6, 21, 2, 17, 13, 10, 18, and 7.

The resulting $(N_O)_{22}^{257}$ design has a Mm distance of 2.246 (compared to the original 2.265) and improved ML_2 discrepancy of 19.032 (compared to the original 37.777). The design has a non-increasing maximum pairwise correlation (0.0074) and condition number (1.039). Thus, the additional design points are added in such a way that the near orthogonality is not jeopardized, but space-filling is improved.

e. Subsequent Column Permuting and Appending

Although this heuristic may be repeated to generate additional run blocks, a minor modification is necessary. Subsequent permutations must take into account that the columns have (2*n*–1) points instead of the original *n* points. For example, in the $(N_O)_{22}^{129}$ design, after the first iteration, the first column of the expanded design is composed of variables 1 and 3 and has 257 points. These hybrid columns are used to identify which of these columns are good and poor performers.

Thus, when an additional permutation is identified using the same heuristic previously described, the subsequent appending yields 256 design points (no replications of the center point). Since only 128 design points are necessary for the third set of runs, the user can choose whether the first 128 or second 128 design points of the new design matrix are appropriate, depending on the Mm distance and ML_2 discrepancy.

E. SUMMARY

The development of the new experimental designs is complete. Each of the desirable design characteristics is satisfied. These designs are either orthogonal or nearly orthogonal and have good space-filling properties. The measures of maximum pairwise correlations and condition numbers are used to assess near orthogonality, and the measures of Mm distances and ML_2 discrepancies are used to assess space-filling. The combination of these measures allows for an excellent blend of orthogonality and space-filling. The end result is a design matrix that offers the means to conduct a systematic and comprehensive exploration of a representative sample of the entire experimental region.

The $(N_O)_{11}^{33}$ and $(N_O)_{22}^{129}$ designs are used in Chapters IV and V, respectively, to illustrate their applicability and strengths. The previous construction algorithm for our designs is augmented with the shifting procedure to provide a complete procedure.

- Step 1. Determine the number of variables (k>7) required for experimentation. If the number of variables is other than 11, 16, or 22, round up the required number of variables up to the nearest one of these numbers.
- Step 2. Establish a maximum threshold pairwise correlation value, ρ, and a maximum threshold condition number.

- Step 3. Using a randomly permuted **e**, construct a design matrix as previously described in this chapter.
- Step 4. Calculate the pairwise correlations and the condition number.
- Step 5. If any of the values in Step 4 exceed the thresholds in Step 2, discard the design and go to Step 3 with a randomly permuted **e** (with replacement). Otherwise, keep the design and proceed to Step 6. Repeat Steps 3-5 until a desired pre-set number of candidate designs are found.
- Step 6. Subject each of the candidate designs to Florian's [1992] method of factorization to decrease the maximum pairwise correlation and condition number.
- Step 7. Calculate the Mm distance and ML_2 discrepancy for each of the Step 6 designs. Rank the designs according to these measures. Choose the design with the minimum rank sum over the two measures.
- Step 8: If a number of variables other than seven, 11, 16, or 22 is required, construct each of the possible combination of columns (having the appropriate number of desired variables) from the Step 7 design and calculate the Mm distance and ML_2 discrepancy. Choose the design with the minimal rank sum over the two measures.
- Step 9: Conduct the experiment and associated data analysis.
- Step 10: Calculate the ML_2 discrepancy for each three-variable combination in the design matrix. Order the ML_2 discrepancies from highest to lowest.
- Step 11: Identify the best and poorest performing variables by comparing how often the individual variables appear in the three-variable combinations in the better half of the combinations versus the poorer half of the combinations. An exact binomial test with a significance level of α (the author chose 0.10) is used to identify the acceptable and the unacceptable performing variables.
- Step 12: Restrict the best performing variables by appending these variables to one of the poorer performing variables. Identify the best permutation of columns yielding the additional design points by conducting various column permutations and comparing the Mm distances and ML_2 discrepancies.

IV. APPLICATION OF A 33-RUN, 11-VARIABLE NEARLY ORTHOGONAL LATIN HYPERCUBE

This chapter details the application of the $(N_O)_{11}^{33}$ design of Appendix B to a known complicated response function that is specified. The experimental domain is $[-1, 1]^{11}$. Each of the 11 variables ranges from –1 to +1. Its performance is compared against both a $(O)_{11}^{33}$ design and a Latin hypercube. The $(N_O)_{11}^{33}$ design offers advantages over a two-level full-factorial design by being able to identify and estimate nonlinear terms. Since the design matrix is nearly orthogonal (not a requirement for uniform designs), there is minimal multicollinearity and coefficient estimates are sharp. Although regression analysis is done to analyze the results of the proposed experiment, this does not imply that the analysis need be restricted to regression analysis.[23]

To illustrate a sequential approach to using the nearly orthogonal designs, the analysis is as follows. An initial experiment is done using the $(N_O)_{11}^{33}$ design of Appendix B. A predictive equation is formulated for the permuted design. A second experiment is conducted, and the predictive results are compared against the actual results. In this example, the second experiment corroborates the first experiment's results, and the experimentation sequence is terminated.

A. KNOWN RESPONSE FUNCTION

The known response function for the example is explicitly defined in this section. There are 11 variables or combinations of these variables that, as far as the analyst knows, may contribute to the response function. If common group screening assumptions are used (e.g., Dorfman [1943] and Watson [1961]), one would expect no more than two variables to be significant. Furthermore, a variable not declared as significant would not be expected to appear in a significant interaction.

The response, denoted as *Y*, expressed in terms of the input variables labeled from *A* to *K*, is shown in (4.1). With two quadratic terms, two two-variable interactions, and

[23] As an example, Ipekci [2002] uses four replications of a $(N_O)_{22}^{129}$ design and applies neural nets, classification trees, and Bayesian nets to analyze the data.

one three-variable interaction, this meets our definition of a high-dimensional complex model from Chapter I. Note that a full-factorial design requiring 2^{11} experiments would be incapable of estimating the coefficients of the quadratic terms, and a 3^{11} design would require over 177,000 runs per replication. To further complicate the proposed experiment, (4.1) also includes an error term (noise) of independent N(0, 1) values.

$$Y = 2A^2 + 2B^2 - AB + 3CF - 3DEF + \varepsilon \tag{4.1}$$

The error term can have a large effect on the observed output, as compared to the true output. As a means of comparison, the $(O)_{11}^{33}$ design generated from Theorems 3.1 and 3.2 (the two-dimensional projections of this design is shown in Figure 4.1) is also subjected to (4.1). Both the $(O)_{11}^{33}$ and $(N_O)_{11}^{33}$ designs have an experimental domain of $[-1, 1]^{11}$.

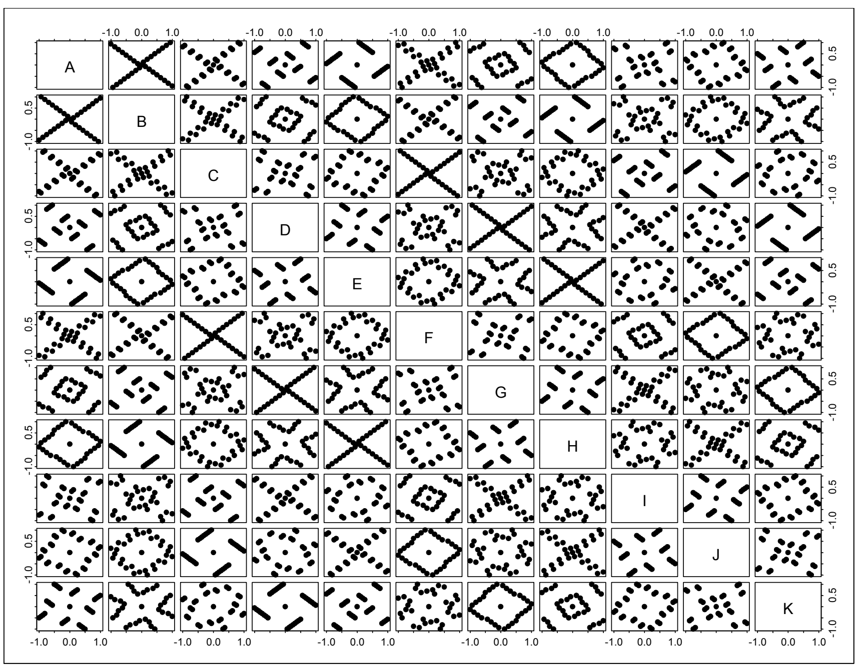

Figure 4.1. Two-dimensional projections of the $(O)_{11}^{33}$ design constructed using Theorems 3.1 and 3.2. Although this design is orthogonal, its space-filling is poor.

The space-filling seen in Figure 4.1 suggests that there might be difficulty in accurately identifying the terms in (4.1) when using the $(O)_{11}^{33}$ design. The patterns associated with variables A and B, variables C and F, variables D and G, and variables E and H suggest that possible interactions or quadratic terms might be difficult to assess.

Upon further investigation, we find that if there is more than one quadratic term in the true response function (note (4.1) has two quadratic terms), then significant pairwise correlations can exist between the quadratic terms, resulting in highly variable regression coefficient estimates for the quadratic terms when using the $(O)_{11}^{33}$ design.

A further comparison of the $(O)_{11}^{33}$ and $(N_O)_{11}^{33}$ designs, using (4.1), gives additional evidence of the nearly orthogonal design's capability. The $(O)_{11}^{33}$ and $(N_O)_{11}^{33}$ designs each have 33 separate design points or input variable settings. A new independent N(0,1) error term is added to each of the 33 responses (for each of the $(O)_{11}^{33}$ and $(N_O)_{11}^{33}$ designs). The corresponding regression analysis is done in S-Plus by using forward and backward stepwise regression with the Akaike information criterion [S-Plus, 1991]. This automatic process is repeated 1,000 times with the same stepwise regression implementation (i.e., nothing other than the noise is changed).

The nearly orthogonal design is closer than the orthogonal design to the true A^2 coefficient a total of 950 times out of the 1,000 different experiments. The nearly orthogonal design is closer than the orthogonal design to the true B^2 coefficient a total of 952 times out of the 1,000 different experiments. The nearly orthogonal design is closer than the orthogonal design to the true *AB* coefficient a total of 808 times out of the 1,000 different experiments. The nearly orthogonal design is closer than the orthogonal design to the true *CF* coefficient a total of 797 times out of the 1,000 different experiments. The nearly orthogonal design is closer than the orthogonal design to the true *DEF* coefficient a total of 620 times out of the 1,000 different experiments. All of these are statistically significant using the exact binomial test.

In 401 of the 1,000 cases, the nearly orthogonal design has closer estimates to all five coefficients than the orthogonal design. In 811 of the 1,000 cases, the nearly orthogonal design has closer estimates to at least four of the five coefficients than the orthogonal design. In 971 of the 1,000 cases, the nearly orthogonal design has closer estimates to at least three of the five coefficients than the orthogonal design. Finally, the mean and standard deviation of each of the 1,000 cases reveals that, while both designs give unbiased estimates, the nearly orthogonal coefficient estimates are much less variable. These mean and standard deviation results are summarized in Table 4.1.

Term	Actual Coefficient	Nearly Orthogonal Design	Standard Deviation	Orthogonal Design	Standard Deviation
A^2	2	2.007	0.627	2.204	9.685
B^2	2	2.003	0.634	1.812	9.694
AB	-1	-1.001	0.416	-0.982	1.663
CF	3	2.991	0.486	2.878	1.239
DEF	-3	-2.997	0.808	-2.899	1.167

Table 4.1. Comparison of regression coefficients for nearly orthogonal (columns 3 and 4) and orthogonal designs (columns 5 and 6) using 1,000 replications of the $(O)_{11}^{33}$ and $(N_O)_{11}^{33}$ designs with (4.1), including error terms. The nearly orthogonal design is closer than the orthogonal design for each of the five coefficients. The standard deviations for these coefficients are also considerably smaller for the nearly orthogonal design.

The $(N_O)_{11}^{33}$ design is compared to a Latin hypercube (again using the experimental domain of $[-1,1]^{11}$). One thousand different Latin hypercubes are used with error terms as specified previously. The Latin hypercubes are competitive with the nearly orthogonal design, but the nearly orthogonal design has uniformly closer coefficient estimates with smaller standard deviations (over the 1,000 replications). The nearly orthogonal design appears to have the best chance of accurately estimating the true regression coefficients and predicting future outcomes. We also expect that as more terms appear in the regression equation, the nearly orthogonal designs will perform even better against Latin hypercubes (Latin hypercubes will be more affected by multicollinearity).

Term	Actual Coefficient	Nearly Orthogonal Design	Standard Deviation	Orthogonal Design	Standard Deviation
A^2	2	2.007	0.627	2.019	0.688
B^2	2	2.003	0.634	2.011	0.691
AB	-1	-1.001	0.416	-0.981	0.585
CF	3	2.991	0.486	2.933	0.567
DEF	-3	-2.997	0.808	-2.951	1.001

Table 4.2. Comparison of regression coefficients for nearly orthogonal (columns 3 and 4) and Latin hypercubes (columns 5 and 6) using 1,000 replications of the $(N_O)_{11}^{33}$ and Latin hypercube designs with (4.1), including error terms. The nearly orthogonal design is closer than the Latin hypercubes for each of the five coefficients. The standard deviations for these coefficients are also smaller for the nearly orthogonal design.

B. REGRESSION ANALYSIS FOR THE FIRST EXPERIMENT

In this section, the analysis performed after the first experiment is explained, and the recommended sequential approach for using the designs is illustrated. Since an analyst would not actually conduct 1,000 experiments, as was done previously for comparative purposes, a single random experiment of 33 runs is performed. As before, a separate independent N(0, 1) error is added to each of the 33 runs. After the first experiment is conducted, a regression analysis is done with forward and backward stepwise selection using the Akaike information criterion and sum of squares to identify significant terms. The fitted model achieves an R^2 of 0.80, and has a residual standard error of 0.966 with 27 degrees of freedom. The regression equation is shown in (4.2).

$$\hat{Y} = 1.905A^2 + 2.091B^2 - .936AB + 2.711CF - 3.04DEF \qquad (4.2)$$

Table 4.3 shows the percentage of the additive error term when divided by the response function (4.1) without the additive error term for each of the 33 runs. These percentages range from –1163 percent to 565 percent, indicating that the error term can be substantial. The NA in the table corresponds to the center point, which has a true response value of 0.0. The quantile-normal plot of the residuals, shown in Figure 4.2, reveals that the residuals are normally distributed. The plot of the residuals versus the predicted values in Figure 4.3 shows a slight curvilinear relation, but is reasonable based on (4.1) and the large error terms.[24]

Run	Percentage	Run	Percentage
1	4.7	18	-6.3
2	25.1	19	-0.4
3	130.0	20	12.2
4	38.0	21	-55.1
5	25.2	22	-21.4
6	16.4	23	-23.2
7	19.5	24	-85.0
8	-27.6	25	78.1
9	564.7	26	-1163.2
10	34.4	27	-2.8
11	-256.4	28	-99.0
12	5.3	29	-139.2
13	19.9	30	7.6
14	-60.1	31	86.8
15	-61.8	32	35.6
16	-146.1	33	-15.1
17	NA		

Table 4.3. The percentage of the error term divided by the mean response for the first experiment involving 33 runs shows the large effect of the error term.

[24] Although forecasting or inference is not done using the results of the regression analysis, Chapters IV and V provide an exploration of the residuals in order to give the interested reader a more complete analysis.

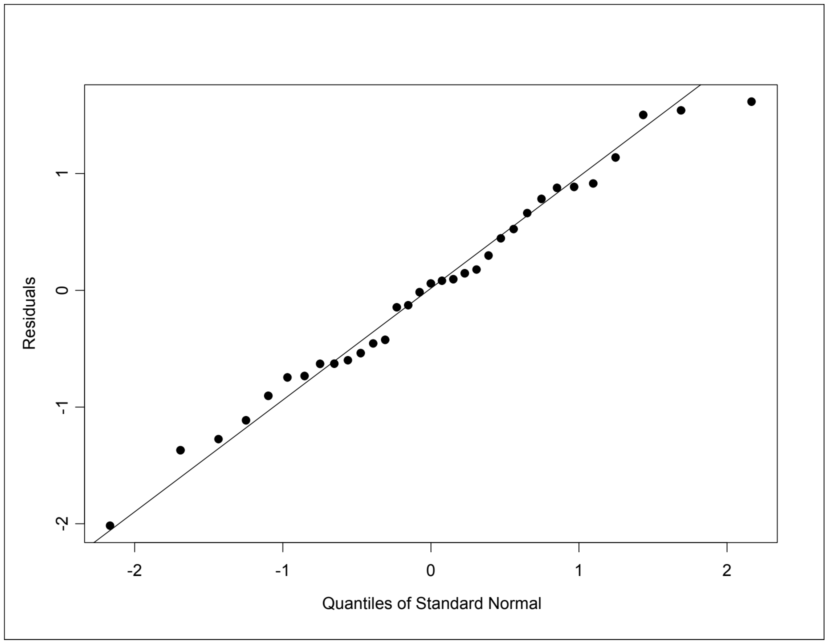

Figure 4.2. Quantile-normal plot of residuals (for the first experiment with 33 runs).

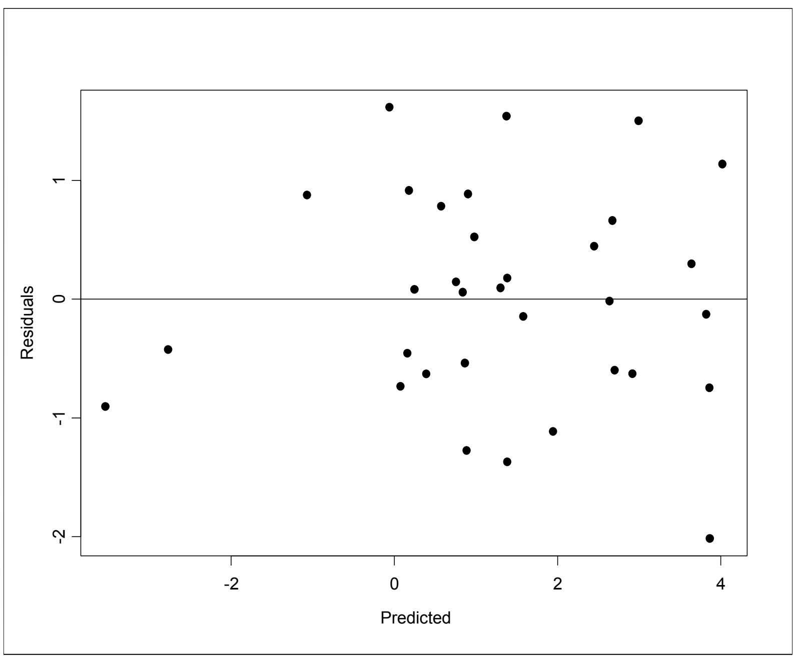

Figure 4.3. Residuals versus predicted value plot (for the first experiment with 33 runs).

From the analysis, (4.2) appears to be a reasonable regression equation for the experimental results. If the experimentation were terminated at this point, the correct terms of the model would be identified. Although the coefficients would not be entirely accurate, their estimates are reasonably correct.

C. REGRESSION ANALYSIS FOR THE SECOND EXPERIMENT

This section describes how the results from the first experiment can be used to assist in the analysis of the second experiment. The design matrix of Appendix B has its columns permuted, as described in the previous chapter, to generate an additional 32 design points. Using this design matrix, the response for these runs is predicted using (4.2). The experiment, consisting of the new 32 design points, is conducted (which includes the additive noise). Table 4.4 shows the percentage of the error term divided by the mean of the response function. Again, the error term significantly influences the response.

Run	Percentage	Run	Percentage
1	-12.1	17	-12.1
2	0.8	18	-185.1
3	-39.8	19	-33.0
4	-151.0	20	-433.2
5	-233.1	21	33.6
6	-28.1	22	1.4
7	80.4	23	47.7
8	95.6	24	19.8
9	61.4	25	-25.1
10	-52.1	26	-135.4
11	-26.7	27	30.3
12	186.2	28	13.1
13	-59.3	29	31.7
14	-9.8	30	3.0
15	95.3	31	88.3
16	-66.7	32	-0.5

Table 4.4. The percentage of the error term divided by the mean response for the second experiment involving 32 runs.

The next two figures compare the predicted values of the experiment with the actual values. Figure 4.4 plots the predicted values of the permuted design matrix using (4.2) against the true values obtained from (4.1). Figure 4.5 plots the predicted values of the permuted design matrix using (4.2) against actual values obtained from the experiment (which includes noise).

Figures 4.4 and 4.5 indicate that the proposed regression of (4.2) does capture the relationships between variables. Furthermore, even with extensive random noise, the predicted values are relatively accurate. The regression equation from the second experiment is

$$\hat{Y} = 2.069A^2 + 2.282B^2 - 1.160AB + 3.060CF - 3.126DEF\,. \qquad (4.3)$$

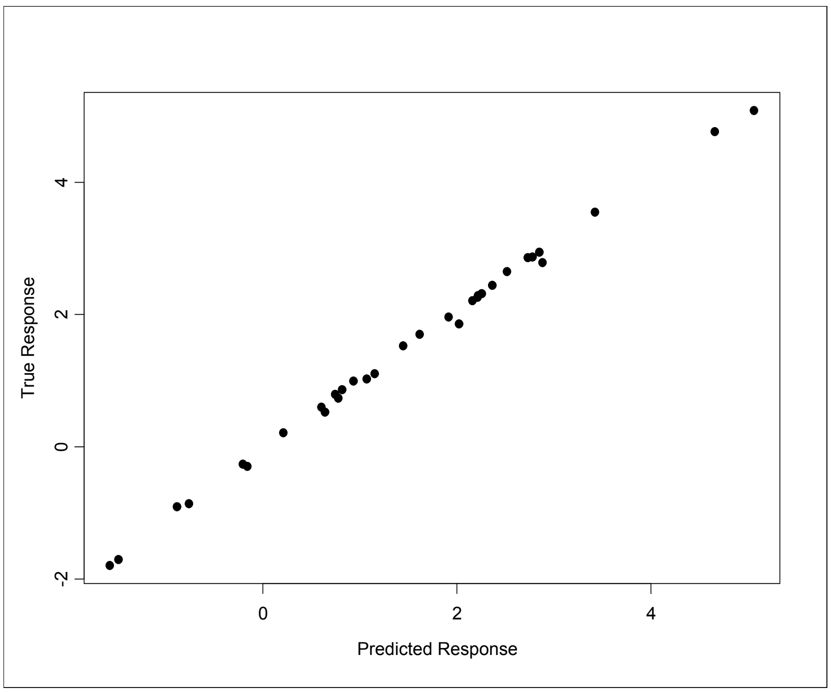

Figure 4.4. Second experiment predicted values versus true values.

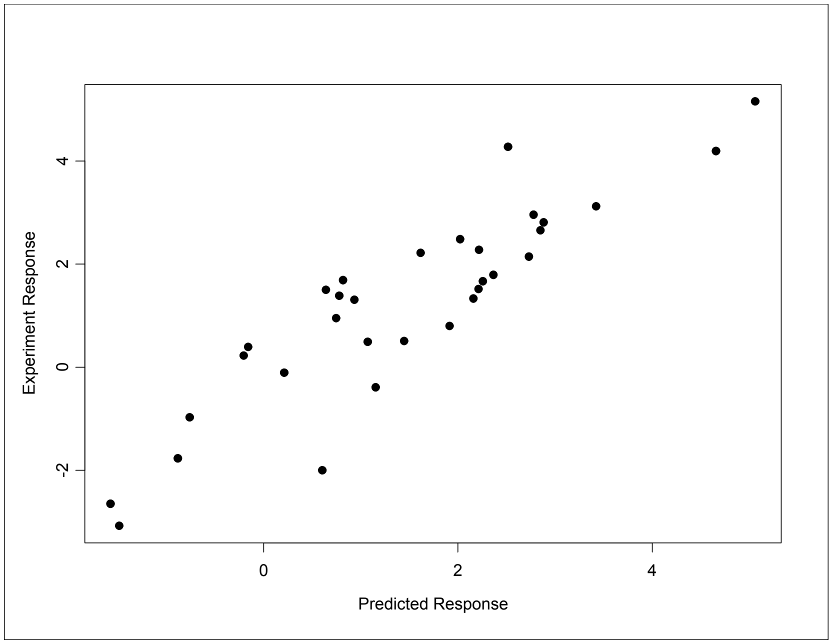

Figure 4.5. Second experiment predicted values versus actual experiment values.

The fitted model achieves an R^2 of 0.81, and has a residual standard error of 0.923 with 26 degrees of freedom. An analysis of the residuals (from Figures 4.6 and 4.7) indicates that the assumption of normally distributed errors is reasonable. The model fit does suggest that the correct terms have been identified, although ascertaining the exact coefficient values is difficult due to the extensive noise.

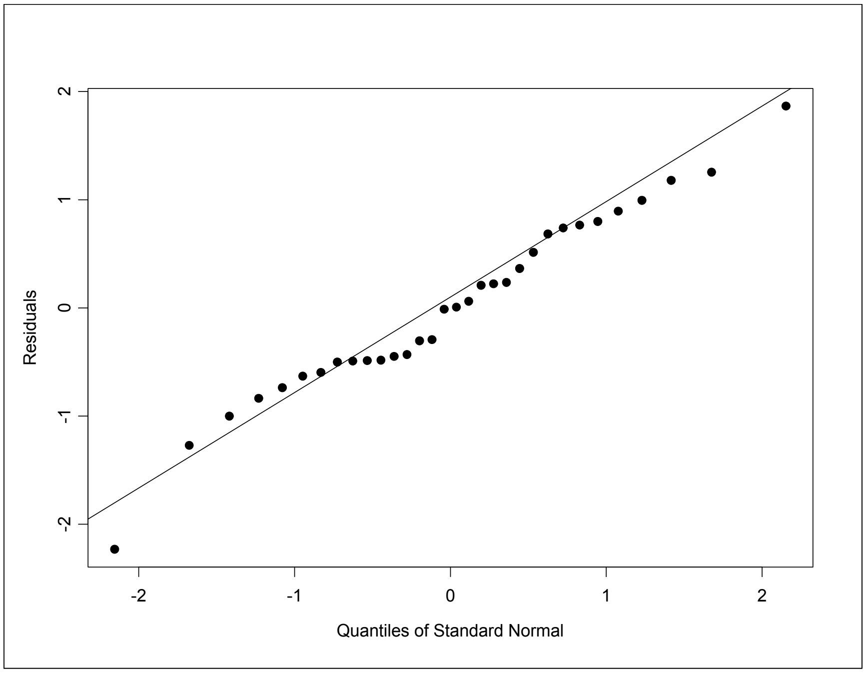

Figure 4.6. Quantile-normal plot of residuals (for the second experiment with 32 runs).

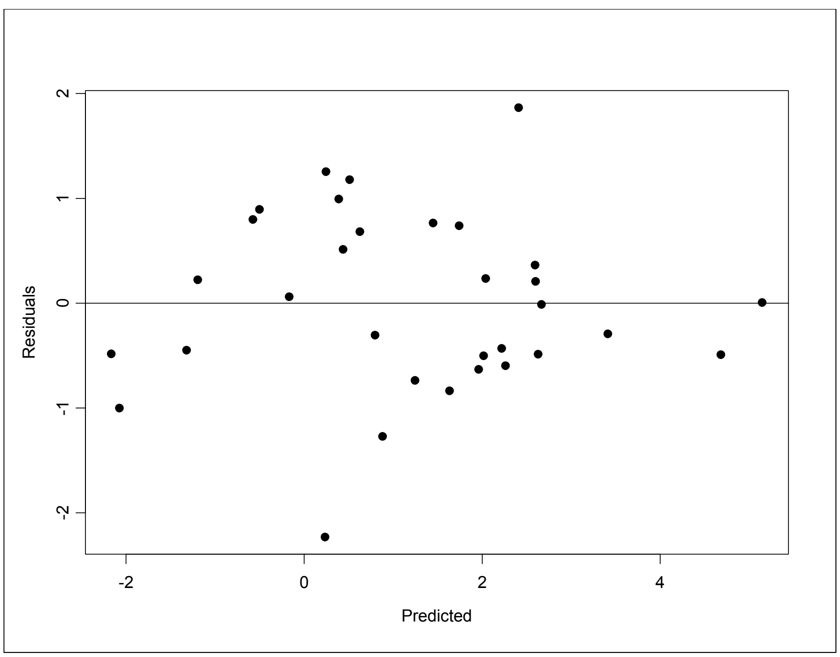

Figure 4.7. Residuals versus predicted value plot (for the second experiment with 32 runs).

To further refine the coefficient estimates, both sets of experiments may be combined to give 65 runs; that is, the 32-run experiment is appended to the 33-run experiment. This increases the associated degrees of freedom and should provide greater model fidelity. The resulting regression equation from combining the two experiments is

$$\hat{Y} = 1.983A^2 + 2.173B^2 - 1.000AB + 2.884CF - 3.085DEF. \quad (4.4)$$

The fitted model achieves an R^2 of 0.80, and has a residual standard error of 0.904 with 59 degrees of freedom. The analysis of residuals, shown in Figures 4.8 and 4.9, indicate that the residuals are reasonably normally distributed. Although the coefficient estimates are not exact due to the extensive noise, they are substantially correct. More importantly, the two quadratic terms, two two-way interactions, and one three-way interaction are accurately identified. It is important to note that this was an illustrative example (as opposed to the 1,000 samples which were used for comparisons) to show how one can apply a sequential approach with these nearly orthogonal designs.

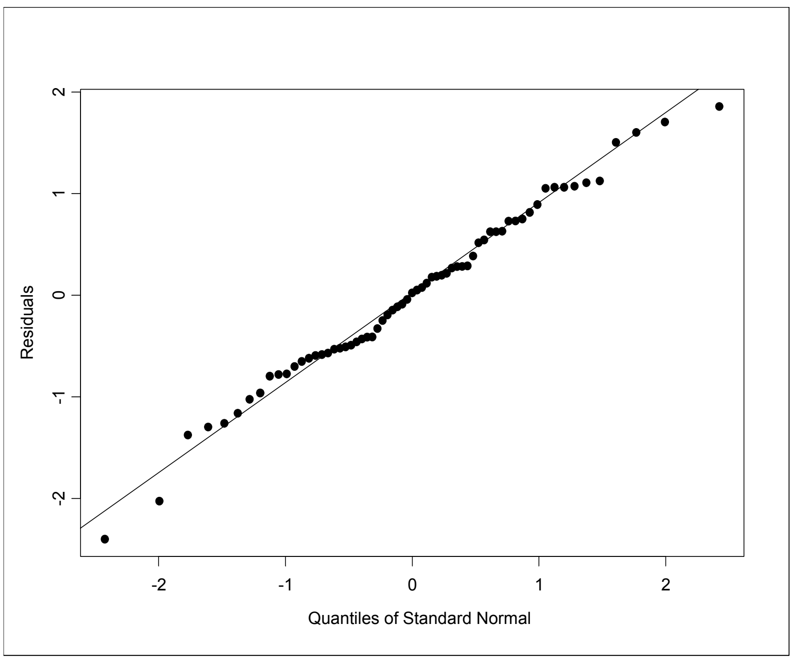

Figure 4.8. Quantile-normal plot of residuals (for the combined experiment with 65 runs).

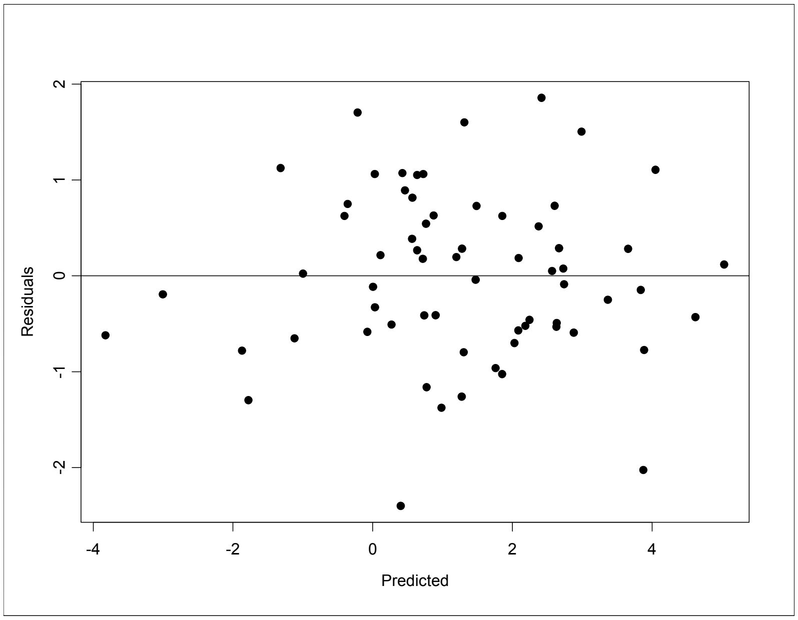

Figure 4.9. Residuals versus predicted value plot (for the combined experiment with 65 runs).

D. SUMMARY

The application of the $(N_O)_{11}^{33}$ design (and its permuted and appended version) illustrates its capacity to capture the non-linear effects and interactions of a sufficiently complex model. The inclusion of a noise variable did not significantly degrade this ability. In this example, the $(N_O)_{11}^{33}$ design provides more accurate regression coefficients than the $(O)_{11}^{33}$ and Latin hypercube designs, the designs we are striving to improve. Ye's [1998] 33-run OLHC is capable of examining only eight variables, but our proposed experimental design examines 11 variables. Ye's [1998] algorithm requires 129 runs to examine 11 variables. Finally, the advantage of the sequential experimentation approach as a means of cross-validation and providing interim results is shown.

THIS PAGE INTENTIONALLY LEFT BLANK

V. APPLICATION OF A 129-RUN, 22-VARIABLE NEARLY ORTHOGONAL LATIN HYPERCUBE

This chapter describes the application of the $(N_O)_{22}^{129}$ design from Appendix D to an agent-based simulation of a military peace enforcement operation. A key feature is the ability of the proposed designs to efficiently handle many variables, in this case 22. The insights that are gleaned from the author's military experience and the data analysis are summarized.

Agent-based simulations, such as ISAAC and MANA,[25] are examples of complex models that may shed light on the nature of combat (e.g., Illachinski [1997], Brown [2000], Graves et al. [2000], Unrath [2000]). In these models, agents are guided by rule sets, and emergent behavior is identified. Agent-based models are an important facet of Project Albert, which is an effort by the U.S. Marine Corps Combat Development Command to provide quantitative answers to important combat questions. These models are called *distillations*—"simulations that attempt to model warfare scenarios by implementing a small set of rules and parameters that allows focus on specific questions." (Horne and Leonardi [2001])

Although agent-based simulations are used here, this does not mean that the designs are only appropriate for such models. The rationale for choosing an agent-based simulation is that most users of warfare models typically change only one or two variables at a time when running computational experiments. To the best of our knowledge, this is the first systematic and comprehensive exploration of such a higher-dimensional region in an agent-based simulation.

The scenario involves a peace enforcement operation. Peace enforcement is defined later in this chapter; here it is important to note that operations of this nature are becoming common for the U.S. military. Furthermore, senior decision-makers have set a high priority on attaining critical information and insights about peace enforcement

[25] ISAAC is an acronym for irreducible semi-autonomous adaptive combat. Information about ISAAC can be found at http://www.cna.org/isaac/isaac_page.htm. MANA is an acronym for map aware non-uniform automata. MANA is a Maori word signifying aura or respect and authority, which is how the New Zealand Army operates (Lauren and Stephen [2001]). Additional information about MANA can be found at http://www.projectalbert.org.

operations in order to reduce the risk to our forces and set the conditions for mission success.

A. MAP AWARE NON-UNIFORM AUTOMATA (MANA) OVERVIEW

This section contains a description of the agent-based simulation used in the experimentation. MANA was developed by the New Zealand Defence Technology Agency to analyze the effect of chaos and complexity theory in armed conflict. The intent is to identify nonlinearities between variables and the co-evolution and emergence of behavior in agents. The two central ideas of MANA are that the behavior of entities within a combat model is critical and highly detailed models are not effective (Lauren and Stephen [2001]). MANA is considered a distillation since it has the characteristics of transparency, speed, ease of answering specific questions, and requires little training to use (Horne and Leonardi [2001]). This dissertation does not enter into the debate of the usefulness of these model types. Instead, the focus is on employing the new experimental designs in a high-dimensional complex model.

One of the major advantages of MANA is that it runs very quickly—the scenario used took approximately seven seconds per iteration on a 1.0 GHz Pentium© 4 processor computer. This permits extensive experimentation to occur, but executing many thousands of runs may still not be an option. Another major advantage is that due to the agent-based and cellular automaton model of MANA, the entities are not controlled by central, predetermined, decision-making algorithms, but make their own decisions as they adapt to the environment. Thus, MANA is a good tool for exploration.

There are numerous variables that can be considered in any of the proposed scenarios of MANA. Figures 5.1-5.3 (best viewed in color) show samples and explanations of possible variables for squad-sized elements and how they may be defined as agents. The characteristics of how the agents react to other friendly and enemy agents in different environments and their weapon and sensor ranges can be modified. It is important to note not only the large number of variables that can be investigated for a particular scenario, but also the large selection of levels each variable can have. A complete overview and explanation of variables can be found in the MANA user's guide (Lauren and Stephen [2001]).

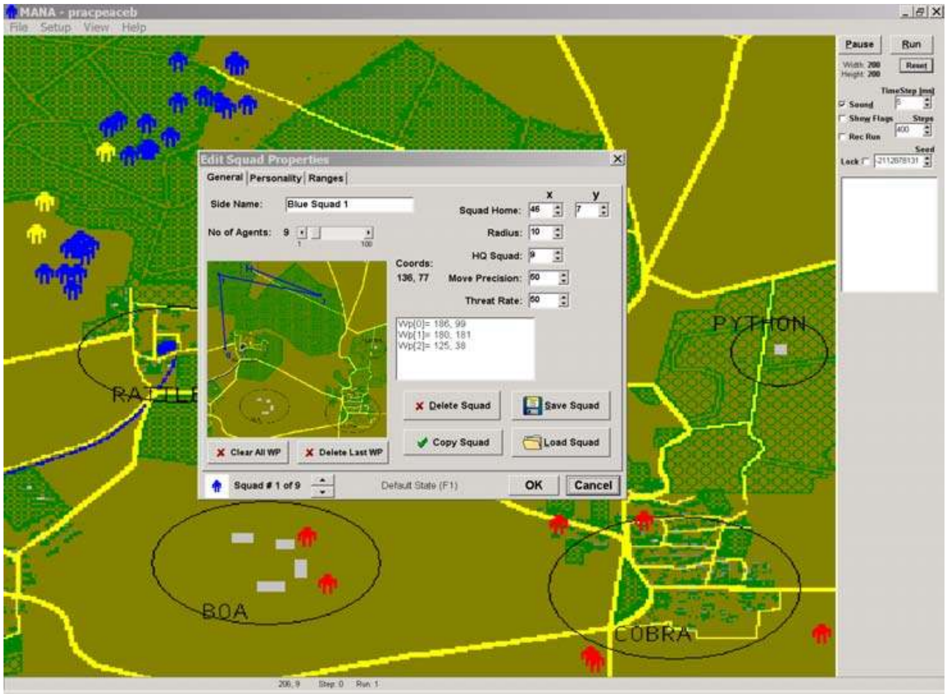

Figure 5.1. The MANA screen for general squad properties. Attributes such as the number of agents in the squad and the squad's location can be modified.

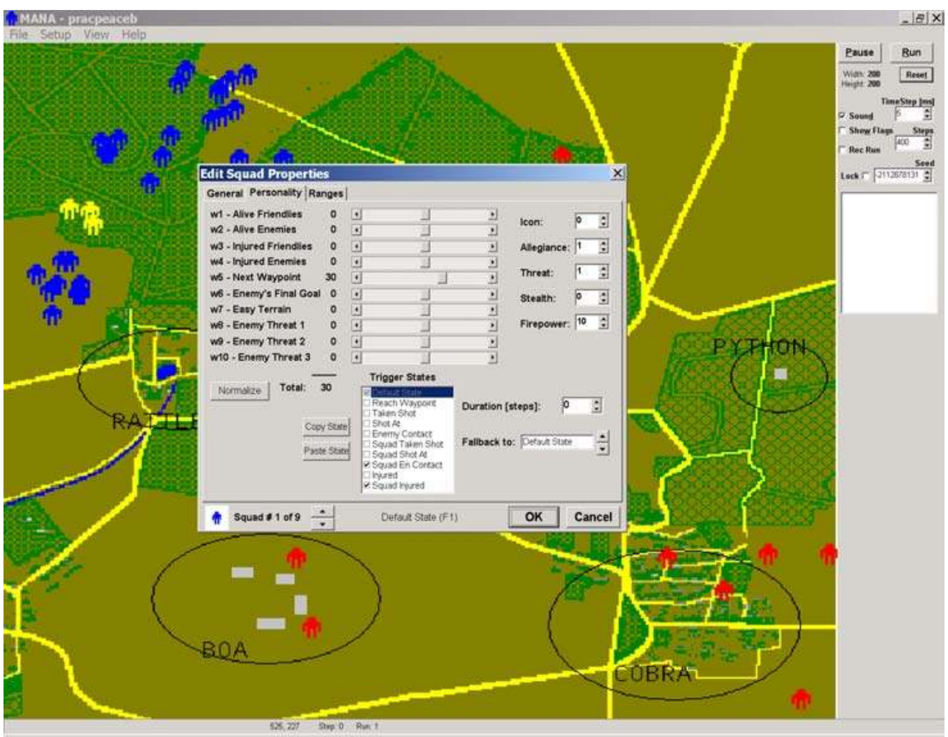

Figure 5.2. The MANA screen for defining the squad's personality. Attributes such as firepower, stealth, and how the agents react to other friendly and enemy agents in different states (i.e., in contact, shot at, injured) can be modified.

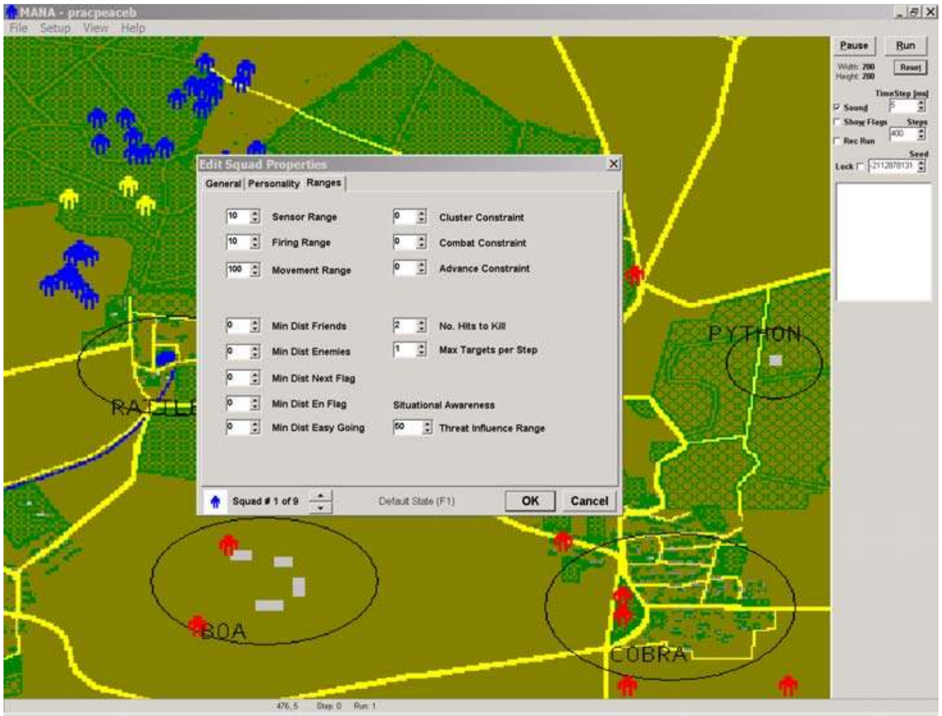

Figure 5.3. The MANA screen for defining the squad's ranges. Attributes such as sensor and weapon ranges and distances from other agents can be modified.

An interesting aspect of this model is shown in Figure 5.2. The model permits entities to have different personalities for different circumstances. For example, how an entity reacts when shot at can be defined differently than how an entity reacts when injured. Furthermore, a squad composed of different entities may have the same definition for each entity, or each entity may be uniquely defined. Thus, a squad of nine entities where each entity has 10 different properties in nine possible states can quickly make comprehensive exploration difficult, even if each simulation lasts approximately seven seconds.

MANA was an appealing candidate for use with the experimental designs since it did meet the distillation criteria. Since an expansive attempt at exploring a high-dimensional region in a model of this type has not previously been done, there is the added benefit of assessing MANA's suitability for addressing complex military issues.

B. SCENARIO OVERVIEW

This section describes the scenario used for experimentation in MANA. Peace enforcement is a critical component of current and future military operations. The U.S.

Army Field Manual 100-23 describes *peace enforcement* as "the application of military force or the threat of its use, normally pursuant to international authorization, to compel compliance with generally accepted resolutions or sanctions. The purpose of peace enforcement is to maintain or restore peace and support diplomatic efforts to reach a long-term political settlement."

The devised scenario is a challenging one since the Blue force is subjected to a series of encounters with the Red force and an originally non-hostile force (Yellow) turns hostile as the scenario progresses. Blue's mission is to clear area of operation (AO) Cobra (see Figure 5.4) within the next two hours in order to facilitate United Nations (UN) food distribution and military convoy operations. Blue uses a light infantry platoon composed of three nine-man rifle squads and a platoon headquarters (HQ) of seven soldiers containing two machine gun teams. Their movement scheme is one squad up and two squads back with the platoon HQ following the lead squad (2nd squad). The 1st squad's task is to follow and support 2nd squad with the purpose of clearing AO Cobra. Their follow-on task is to clear AO Python for subsequent UN food distribution and military convoy operations. The 2nd squad's task is to conduct a movement to contact with the purpose of clearing AO Cobra. Their follow-on task is to clear AO Cobra for subsequent UN food distribution and military convoy operations. The 3rd squad's task is to follow and support 2nd squad with the purpose of clearing AO Cobra. Their follow-on task is to clear AO Boa (a small urban area with four building structures) for subsequent UN food distribution and military convoy operations. After 2nd squad clears AO Cobra, the platoon HQ moves to AO Boa to provide supporting fires for 3rd squad.

Red has a five-member element located in the vicinity of AO Cobra and two two-member elements patrolling along the movement routes of Blue squads 1 and 2. Additionally, Red has a two-member element in the vicinity of AO Boa. An originally non-hostile Yellow three-member element is initially in Blue's starting location. After discovering no potable water in vicinity of AO Rattler, Yellow becomes hostile against Blue, seeks small arms from the vicinity of AO Boa, and moves to the vicinity of AO Python. The overall scenario is deemed doctrinally correct and plausible by the U.S. Army Infantry Simulation Center at Fort Benning, Georgia (McGuire [2001]). Figure 5.4

(best viewed in color) provides an initial graphical depiction of the proposed scheme of maneuver.

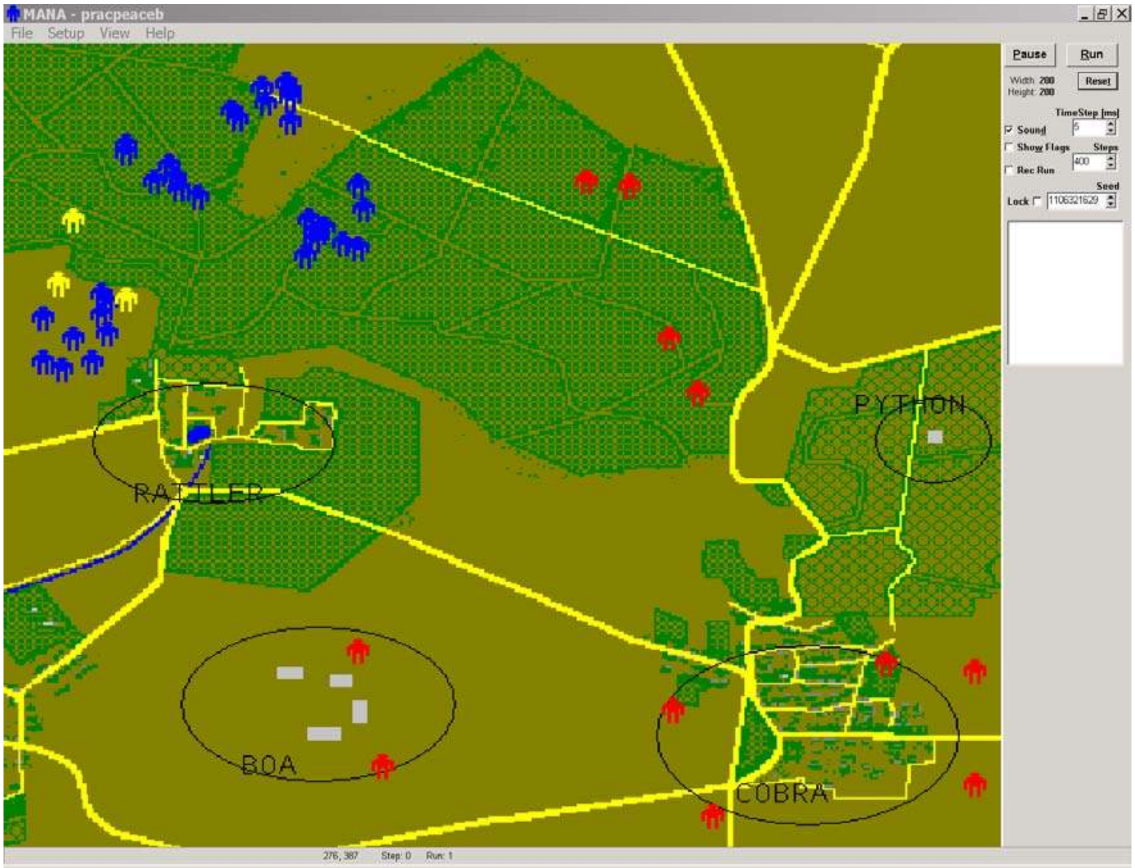

Figure 5.4. Initial graphical depiction of proposed scheme of maneuver for the MANA peace enforcement scenario.

There are 22 variables identified for experimentation. Choosing these 22 from among the many available variables and their levels was done using the author's military experience and judgment and from hundreds of small, interactive experiments of changing one or two variables and determining if a significant event occurred. For example, it was found that if Blue is given too high of a weapon and sensor range, upon initiation of the scenario, Blue immediately kills all of the threat agents. Thus, it was decided that although these variables are critical components of military conflict, in order to focus on entity personalities, these variables would not be candidates for experimentation. Although the primary emphasis is on testing the experimental designs, secondary criteria did include searching for important variables, interactions, and insights for peacekeeping operations and determining the appropriateness of MANA in modeling these operations. The variables identified for experimentation and a brief description follows. These variables are shown in Figures 5.1-5.3.

A. Blue Platoon HQ move precision: amount of randomness in blue movement

B. Blue Squad 1 move precision: amount of randomness in blue movement

C. Blue Squad 2 move precision: amount of randomness in blue movement

D. Blue Squad 3 move precision: amount of randomness in blue movement

E. Blue Platoon HQ in contact personality element w1: controls propensity to move towards agents of same allegiance

F. Blue Squad 1 in contact personality element w1: controls propensity to move towards agents of same allegiance

G. Blue Squad 2 in contact personality element w1: controls propensity to move towards agents of same allegiance

H. Blue Squad 3 in contact personality element w1: controls propensity to move towards agents of same allegiance

I. Blue Platoon HQ in contact personality element w2: controls propensity to move towards agents of enemy allegiance

J. Blue Squad 1 in contact personality element w2: controls propensity to move towards agents of enemy allegiance

K. Blue Squad 2 in contact personality element w2: controls propensity to move towards agents of enemy allegiance

L. Blue Squad 3 in contact personality element w2: controls propensity to move towards agents of enemy allegiance

M. Blue Platoon HQ injured personality element w1: controls propensity to move towards agents of same allegiance

N. Blue Squad 1 injured personality element w1: controls propensity to move towards agents of same allegiance

O. Blue Squad 2 injured personality element w1: controls propensity to move towards agents of same allegiance

P. Blue Squad 3 injured personality element w1: controls propensity to move towards agents of same allegiance

Q. Blue Platoon HQ injured personality element w2: controls propensity to move towards agents of enemy allegiance

R. Blue Squad 1 injured personality element w2: controls propensity to move towards agents of enemy allegiance

S. Blue Squad 2 injured personality element w2: controls propensity to move towards agents of enemy allegiance

T. Blue Squad 3 injured personality element w2: controls propensity to move towards agents of enemy allegiance

U. Blue movement range for all squads: controls movement speed of agents

V. Red personality element w8: controls propensity to move towards enemies (Blue) in situational awareness map which are of threat level 1

There is a requirement for 129 different levels for each input variable. This is done as follows. Variables A-D have settings of 1-513 in increments of 4, for a total of 129 levels. Variables E-T and V have settings of -64 to 64 in increments of 1. Variable U has settings of 72 to 200 in increments of 1. The firepower and sensor ranges of all allegiances are equal in order to amplify personalities.

The simulation has a duration of 1,000 time steps. For each run, 100 iterations are conducted with different random seeds. MANA is limited in its output measures. The key measure extracted is the exchange ratio (ER), defined as the quotient of the number of Red killed divided by the number of Blue killed. The other measure to investigate is whether Blue occupies each of the three AO's by time step 1,000. Due to the high variability of the ER, 100 replications are done for each of the 129 input combinations. In many cases, the standard deviation is almost one-half of the mean value—even after 100 runs. This is an appealing feature to members of Project Albert since it illustrates the variable and, perhaps, complex nonlinear nature of military conflict. Furthermore, it underscores the argument that attempting to predict, optimize, or calibrate results of this nature via regression analysis might be futile. A better alternative is to provide decision-makers with the insights obtained on the important variables, interactions, nonlinearities, and where they occur. These insights are gained from a systematic and comprehensive exploration of the high-dimensional region.

C. DATA ANALYSIS FOR THE FIRST EXPERIMENT

This section summarizes the data analysis associated with the experiment using the $(N_O)_{22}^{129}$ design from Appendix D and examining the resulting ER's. For each of the

129 input variable combinations, the response (ER) is the mean of the 100 runs for the regression analysis and is denoted as the mean ER (thus, there are a total of 12,900 runs). If the regression equation is fit to all of the raw data, the coefficients will be the same. However, the associated *p*-values and the R^2 will be different. An initial regression equation is constructed using the linear effects of all variables to identify the dominant main effects of variables C, E, F, G, P, U, and V. The regression equation is found interactively through trial and error using forward and backward stepwise selection (to include quadratic terms and three-variable interactions) using the dominant main effect variables with various subsets of the non-dominant main effect variables. The Akaike information criterion and sum of squares are the primary measures used to build the model.

An initial regression analysis is done and results in three quadratic terms, four linear effects, and seven two-variable interactions. In building the model, caution is maintained against deriving an over fitted model, yet balanced with the goal of the model achieving sufficient explanation. The resulting model, shown in (5.1), has an R^2 of .66 and a residual error of .1584 with 114 degrees of freedom. The exchange ratio is

$$\begin{aligned} ER = 1.201 &+ (2.385\text{e-}007)E^2 + (2.654\text{e-}007)P^2 + (2.341\text{e-}008)U^2 \\ &- (0.000221)C + (0.00435)F + (0.00770)G - (0.00325)V + (2.400\text{e-}006)BN \\ &- (6.666\text{e-}006)CF - (4.201\text{e-}006)CG - (0.0000255)EL - (0.0000171)FV \\ &- (0.0000351)GU + (0.0000223)QR\,. \end{aligned} \tag{5.1}$$

Figure 5.5 shows that the predictive ability of the model is susceptible to significant error. An advantage of the model, as shown by Figures 5.6 and 5.7, is that the estimated errors appear patternless and uncorrelated with the fitted values, and the normal distribution is tenable for describing their distribution.

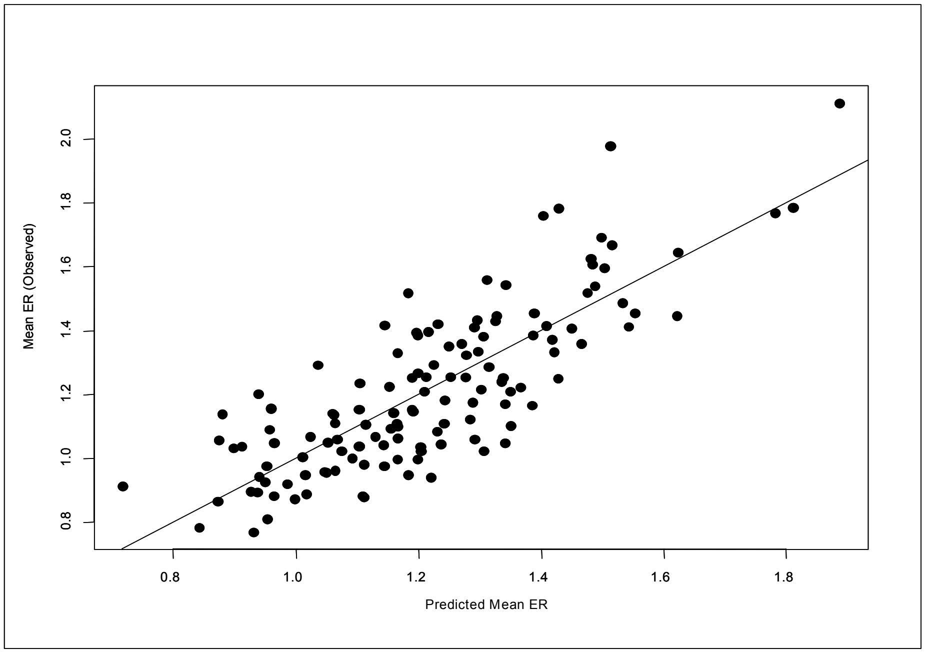

Figure 5.5. First experiment predicted values versus true values for the MANA peace enforcement scenario indicating significant noise in the model.

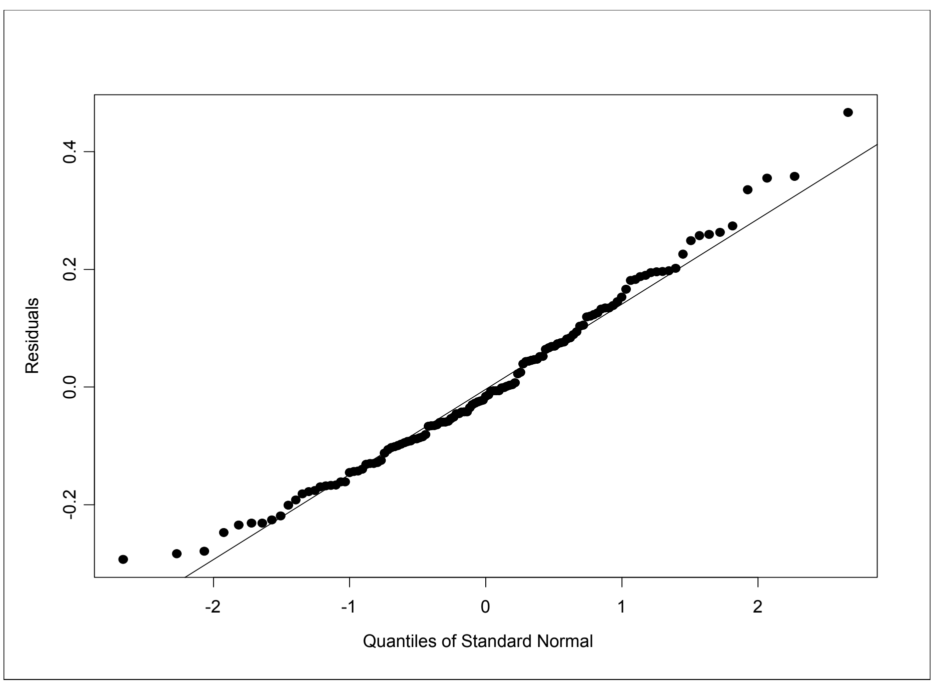

Figure 5.6. Quantile-normal plot of residuals (first experiment) for the MANA peace enforcement scenario indicating relative normality.

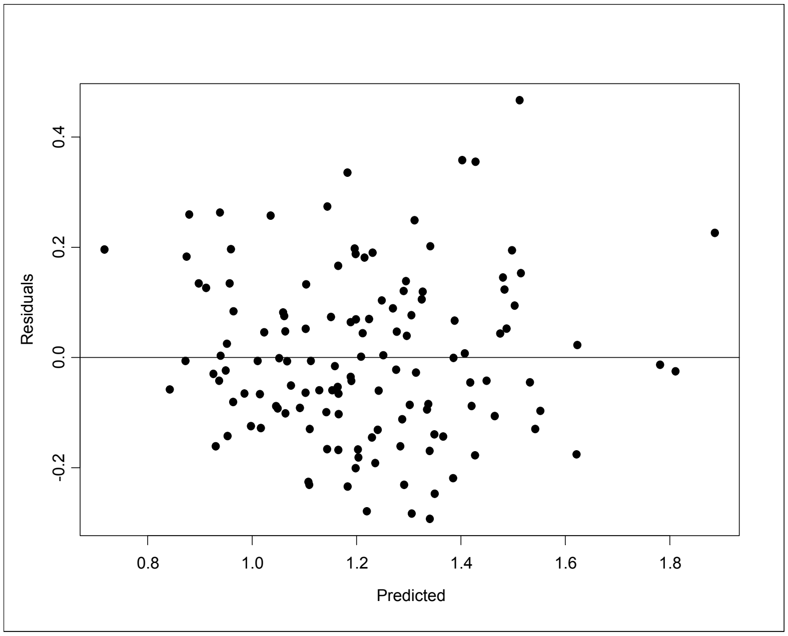

Figure 5.7. Residuals versus predicted value plot (first experiment) for the MANA peace enforcement scenario indicating relative normality.

Recall that for each of the 129 input variable combinations, the response (ER) is the mean of the 100 runs and is denoted as the mean ER. The mean ER's appear to have a gamma shape (see Figure 5.8). Parameters using maximum likelihood estimators are identified. These include a scale parameter of 0.0671 and shape parameter of 18.315. From the Kolmogorov-Smirnov goodness-of-fit test (based on known values for the parameters), it appears that the gamma distribution is a plausible model for the mean ER's (*p*-value=0.586).[26]

[26] Recall that the mean ER's are the mean of the 100 replications of the 129 input combinations.

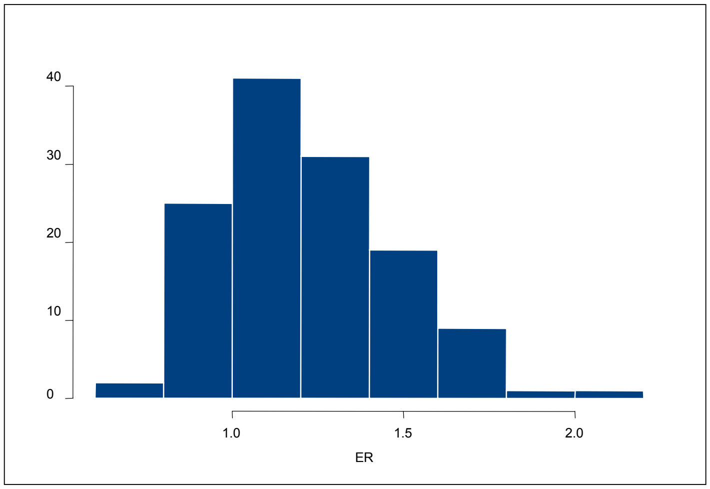

Figure 5.8. Histogram of mean ER's (first experiment). The mean ER's appear to have a gamma distribution.

Although (5.1) does reasonably well in attempting to explain the relationship between the ER and the variables, it may be of limited value for decision-makers due to its poor predictive ability and interpretability. Furthermore, a simulation scenario of this type cannot be replicated exactly in the real world. Finally, due to the chaotic nature of warfare, providing a point forecast for an ER, or even an ER with some predictive interval, could be misleading. Instead, the focus is on gaining significant military insights ("golden nuggets") and identifying regions of good and poor performance.

Although the regression equation can be presented to the decision-maker, the following bullet comments are more representative of the type of information that the author believes should be presented to military decision-makers. Future experimentation can confirm these insights, cast doubt on them, or create new ones. These comments are culled by studying what the regression terms actually mean in terms of the simulation and extensively visualizing the scenario playbacks. Each insight is found by using data analysis, coupled with the author's military education and experience of over 20 years. Each term in (5.1) is investigated to determine the impact upon the ER's.

1. Elements should consider moving towards other friendly elements when in contact with threat elements.
2. An element with injured soldiers should consider reducing the distance between individual soldiers in urban-type terrain.

3. Expedited execution might be critical in peace enforcement operations.
4. The lead squad or unit should have some predictability in their movement in order to provide follow-on units a better picture of where they are on the battlefield.
5. A threat element that is not overtly aggressive might produce more friendly casualties. This problem can be compounded if friendly elements reduce the distance between soldiers against a threat of this type.
6. When a friendly element sustains casualties and is reducing the distance between soldiers, their movement in doing so should not be predictable.
7. When in contact and no casualties have been sustained, elements should consider being less random in their movement.
8. When in contact, elements might consider refraining from reducing the distance between soldiers while simultaneously advancing towards the threat.
9. When the lead element is in contact with the threat, if the element attempts to mass with other elements, the lead element might consider doing so in a measured and deliberate fashion as opposed to an expedited manner.
10. When elements with injured or killed soldiers are in contact with threat elements, continuing the operation instead of ceasing it might be more advantageous.

It is also beneficial to examine the tails of the mean ER distribution to see what insights exist. The best mean ER runs (approximately 10 percent or 13 runs) and worst mean ER runs (approximately 10 percent or 13 runs) were segregated. Since only a subset of the runs are taken, there are now significant correlations between the input variables (i.e., the removal of cases has eliminated the near orthogonality property held by the entire design matrix).

The correlations are computed for each variable and the best and worst mean ER's. The absolute values of the correlations are then rank ordered for each set of segregated runs and these sums added. The significant variables based on an exact binomial test (*p*-values<0.10) are variables B (Blue Squad 1 move precision), K (Blue Squad 2 in contact personality element w2), N (Blue Squad 1 injured personality element w1), Q (Blue Platoon HQ injured personality element w2), and S (Blue Squad 2 injured personality element w2). This indicates they have a significant effect on whether the mean ER was high or low when compared to all of the runs. An analysis of their boxplots, shown in Figures 5.9-5.13, is useful in generating insights.

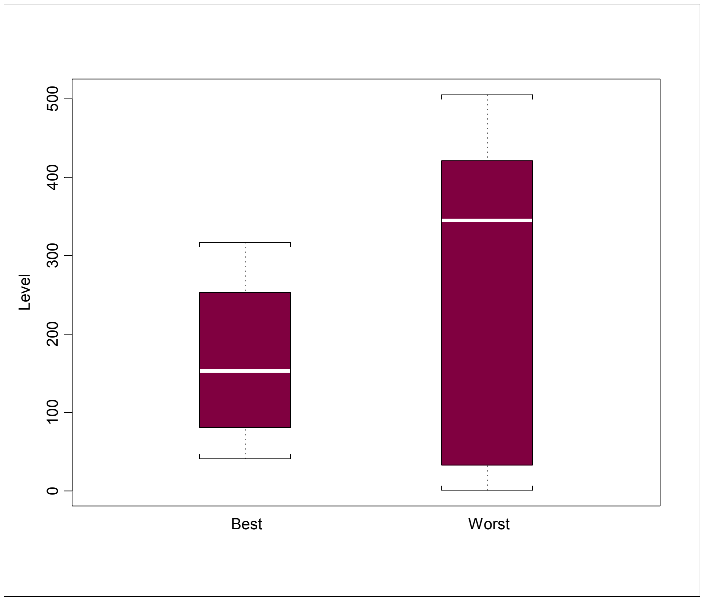

Figure 5.9. Boxplots of levels of variable B (Blue Squad 1 move precision) for best and worst mean ER's (first experiment).

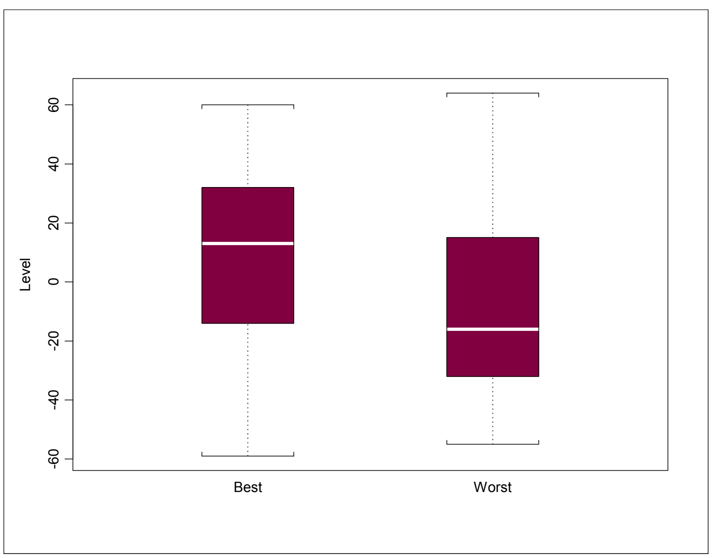

Figure 5.10. Boxplots of levels of variable K (Blue Squad 2 in contact personality element w2) for best and worst mean ER's (first experiment).

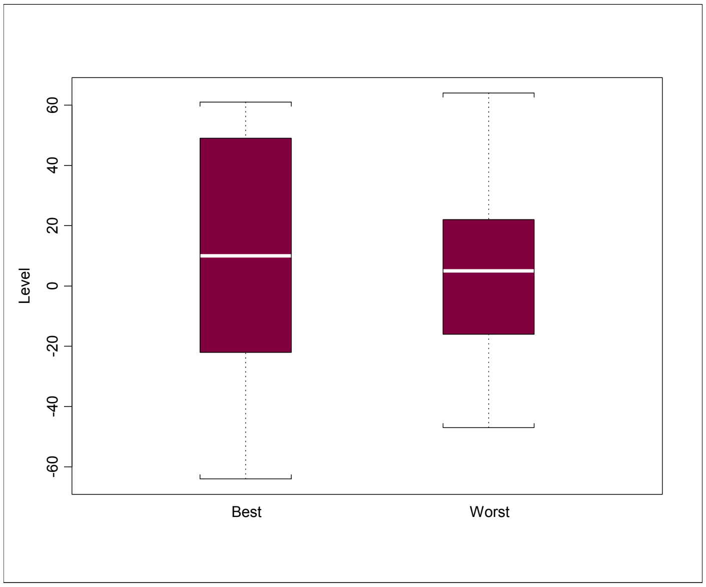

Figure 5.11. Boxplots of levels of variable N (Blue Squad 1 injured personality element w1) for best and worst mean ER's (first experiment).

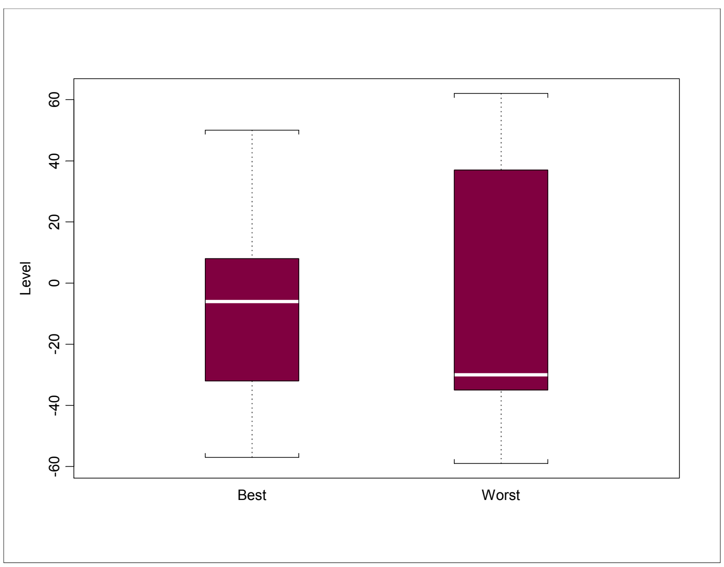

Figure 5.12. Boxplots of levels of variable Q (Blue Platoon HQ injured personality element w2) for best and worst mean ER's (first experiment).

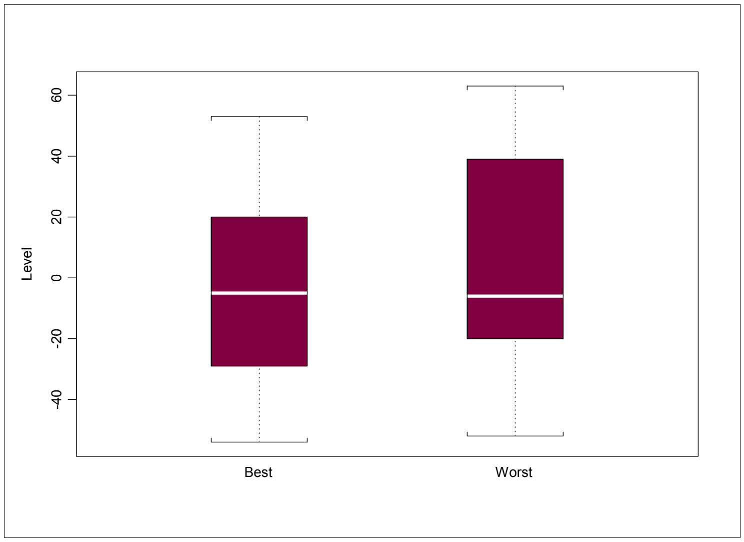

Figure 5.13. Boxplots of levels of variable S (Blue Squad 2 injured personality element w2) for best and worst mean ER's (first experiment).

An analysis of Figures 5.9-5.13 and visualizing the simulation runs from MANA, in conjunction with military judgment, provides some additional possible insights.

11. An element encountering an undetermined element (not identified as friendly or threat) should consider moving in an orderly and systematic manner.
12. When in contact with the threat, with no casualties sustained, the lead element should consider maintaining contact. If casualties are sustained, there should be consideration given to continuing the engagement with a different lead element.
13. When an element has casualties and is engaged with a once non-hostile threat that has become hostile, reducing the distance between soldiers might be beneficial.
14. When the headquarters element has injured or killed soldiers, the element should be cautious in seeking engagement with the threat, although it still provides command and control to its subordinate elements.

Fewer casualties are preferred. However, the mission of securing the areas must be completed. Therefore, an additional proposed measure is considered. The measure is a categorical variable of whether or not each of the AO's is occupied by Blue entities by time step 1,000. This measure requires an analysis of the playback since the output file cannot provide this information; this is a limitation of MANA. For each of the 129 input variable combinations, a subset of 10 runs from the 100 replications is manually selected. If each of the 10 runs in the subset achieves the goal of occupying the AO, the corresponding input variable combination is segregated from those input variable combinations not achieving the goal.

One of the most interesting findings is discovered from this analysis. The most important variable that affects whether the mission is completed on time or not is variable U (Blue movement range) when its levels range from 101 to 114. At these levels, the Blue entities do not advance substantially from their initial starting positions. Yet, at levels below 101 and above 114, the Blue entities do move as specified by the parameter (i.e., at level 90, the Blue entities move slower than at level 120). By using the $(N_O)_{22}^{129}$ design, this model problem, not yet resolved, is identified.

If the experimentation is terminated at this stage, the military decision-maker may have sufficient insight and analysis to make a decision. Since time permits, the next step is to identify the follow-on set of experiments, predict its results, and determine if the initial analysis is substantiated by this subsequent experimentation. Although the next design also covers the entire experimental region, based on the first experiment, the ranges of certain variables could be reduced to focus on regions of particular interest.

D. DATA ANALYSIS FOR SUBSEQUENT EXPERIMENTS

This section describes the analysis associated with permuting and appending the columns of the $(N_O)_{22}^{129}$ design, as specified in Chapter III, and then conducting the computer experimentation. Using (5.1) and the permuted $(N_O)_{22}^{128}$ design, ER's were predicted for this new design. Again, each of the 128 runs (the center point is not repeated) was replicated 100 times and the mean of the number of Red killed divided by the Blue killed (the ER) is the measure.

The mean ER's have a gamma distribution similar to that of the first experiment's mean ER's. The shape parameter is 17.799 (compared to 18.315) and scale parameter of 0.0670 (compared to 0.0671). The Kolmogorov-Smirnov goodness-of-fit test (using the estimated parameters) has a p-value of 0.678.

After the experiment is conducted, the predicted mean ER's are compared with the actual mean ER's. Figure 5.14 illustrates this relationship with both a least squares and a weighted least squares fitted line. There is not much difference between the two fitted lines. A correlation of approximately 0.628 exists between the predicted values and actual values. Although this cross-validation does not achieve as high of agreement as one would desire, considering the complexity of the model, the correlation is certainly reasonable and indicates our initial proposed model seems reasonable.

A separate regression equation is done for the second experiment (to identify additional insights). Although initial insights from the first experiment may not be confirmed by the second experiment, the insights should not be dismissed. Even though 129 runs are done on the first experiment and are not significantly clustered, these design points are still quite sparse in 22 dimensions. The second experiment of 128 points may confirm the initial experiment's findings or generate additional insights since additional areas of the experimental region are explored. The resulting regression equation, built by the author as before, contains one quadratic term, six main effects, and four two-way interactions.

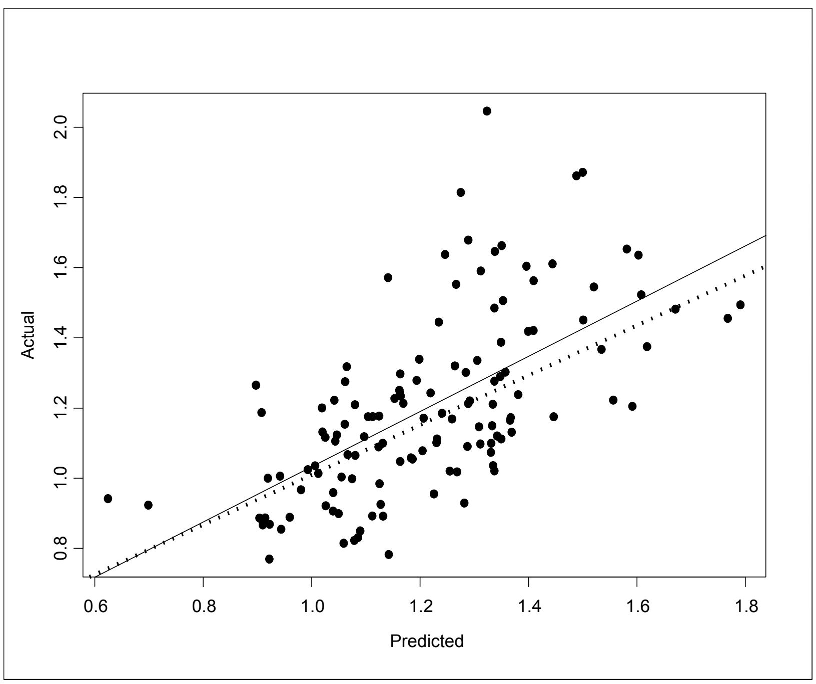

Figure 5.14. Predicted values versus actual values (second experiment) with least squares fitted line (solid) and weighted least squares line (dotted) for the mean ER's.

The resulting model, shown in (5.2), has an R^2 of 0.67 and a residual error of 0.1553 with 115 degrees of freedom. These measures are similar to the measures of (5.2), but the model terms are different. Significant two-variable interactions, where each element of that interaction is not necessarily significant as a main effect, are found in (5.2).

$$ER = 1.678 + (4.035\text{e-}007)U^3 - (.000319)B + (.000782)E + (.00213)F + (.00430)G + (.000976)P + (2.082\text{e-}005)FG - (2.184\text{e-}005)GU + (1.977\text{e-}005)KR + 6.757\text{e-}006)RU \quad (5.2)$$

The analysis of the residuals does not indicate a departure from the normality assumption and is omitted. Although the first experiment results in a more complex equation and (5.1) and (5.2) do not have all of the same terms, there is similarity between the experiments when military analysis and judgment are applied.

- The addition of the *E* and *FG* terms reinforces insight 1.
- The addition of the *P* term reinforces insight 2.
- The addition of the *B* term reinforces insight 11.
- The addition of the *KR* term expands upon insights 9 and 10 by incorporating the insight that supporting elements of the lead element must continue to provide support even if the supporting element has sustained casualties.
- The addition of the *RU* term expands insight 13 by incorporating the insight that if the element, with or without casualties, decides to engage a hostile threat that was once non-hostile, they should do so expeditiously.

This detailed analysis indicates that the two experiments generate complementary insights that can be useful to decision-makers. Furthermore, there is considerable noise in the simulation (as would be expected in a true peace enforcement operation), so solely using these regression equations to predict, optimize, or calibrate may be misleading. Instead, applying data analysis and military knowledge leads to potentially useful results from the simulation.

As was done in the first experiment, the next step is to identify the top and bottom 10 percent of the mean ER's. After rank ordering the correlations and applying the exact binomial test (p-values<0.10) to the rank sums, the significant variables are E (Blue Platoon HQ in contact personality element w1), I (Blue Platoon HQ in contact personality element w2), K (Blue Squad 2 in contact personality element w2), Q (Blue Platoon HQ injured personality element w2), and S (Blue Squad 2 injured personality element w2). Variables K, Q, and S share similar boxplots as those in Figures 5.10, 5.12, and 5.13 and support insights 12 and 14. Figures 5.15 and 5.16 show the boxplots of levels of variables E and I for the best and worst mean ER’s.

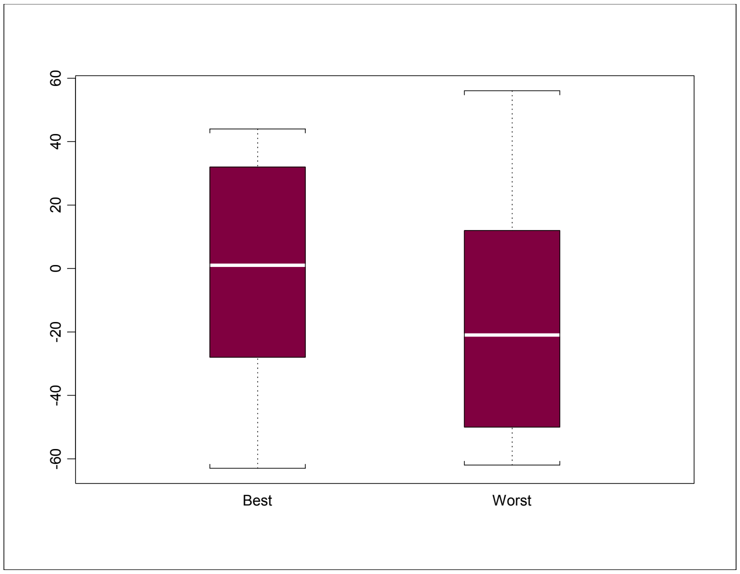

Figure 5.15. Boxplots of levels of variable E (Blue Platoon HQ in contact personality element w1) for best and worst ER’s (second experiment).

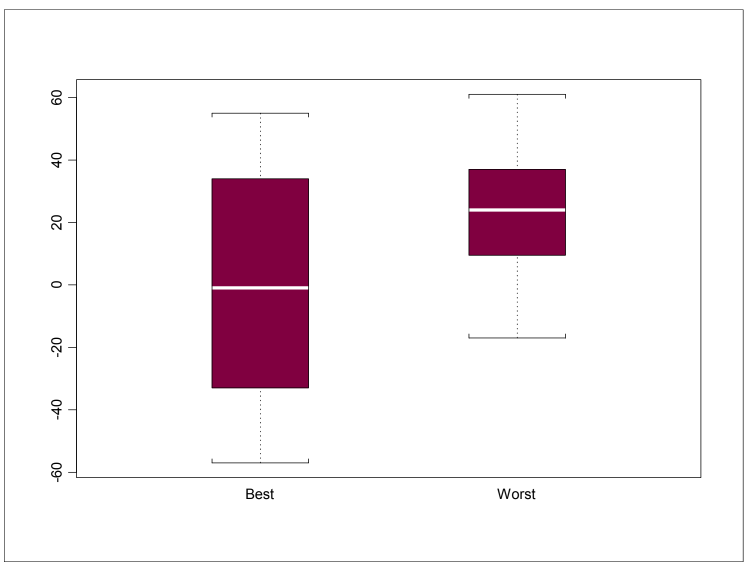

Figure 5.16. Boxplots of levels of variable I (Blue Platoon HQ in contact personality element w2) for best and worst ER’s (second experiment).

An examination of the correlations for variables E and I from the first experiment’s best and worst mean ER’s does not show as strong of a correlation as in the second experiment. Analyzing Figures 5.15 and 5.16 generates one additional insight and confirms a previous insight.

- When the headquarters element is in contact with the threat, it should consider moving towards other friendly elements.
- The I variable reinforces insight 14.

Finally, there are similar problems with Blue movement when variable U had levels of 101 to 114. This problem in MANA has been forwarded to the model developers.

The first and second experiments are now combined and a regression analysis is executed on the 257 input variable combinations. The resulting model, shown in (5.3), has an R^2 of 0.67 and a residual standard error of 0.1505 with 243 degrees of freedom. The fitted exchange ratio is

$$\begin{aligned} ER = 1.890 + (1.928\text{ee-}007)U^2 + (.000457)B + (.000736)E + \\ + (.00237)F + (.00568)G + (.000826)P - (.00898)U - (.00327)V \\ - (4.866\text{E-}006)BU - (3.021\text{e-}005)GU - (2.688\text{e-}005)FV + (1.378\text{e-}005)IJ \\ + (2.225\text{e-}006)BN. \end{aligned} \tag{5.3}$$

An analysis of the quantile-normal plot of the residuals in Figure 5.17 indicates a heavy- tailed right-hand side. This most likely occurs due to the skewed mean ER measures.

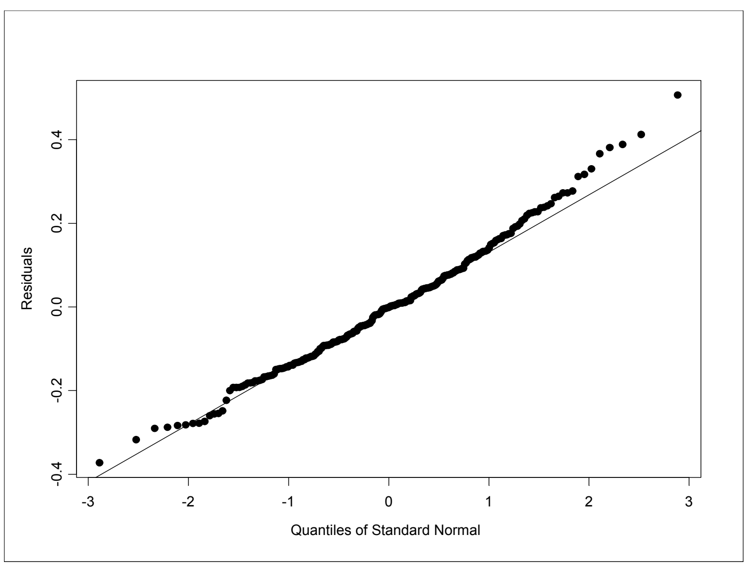

Figure 5.17. Quantile-normal plot of residuals (combined experiment) indicating a heavy tailed right-hand side.

A third experiment is conducted by permuting the columns of the combined experiment. Though the third experiment was not entirely necessary, it is done to illustrate how additional design points are generated. Recall that these permuted columns are hybrid columns of the two original columns; that is, each permuted column consists of 257 values, with 128 values each showing up twice. Table 5.1 shows the composition of each of the columns, where the number represents the variable from Appendix D. For example, column 1 is composed of columns 1 (first experiment) and 3 (second experiment). This hybrid column is then appended with columns 18 and 17.

	Experiment			
Column	First	Second	Third	Fourth
1	1	3	18	17
2	2	16	1	3
3	3	20	21	18
4	4	11	22	7
5	5	9	16	21
6	6	19	8	14
7	7	4	2	16
8	8	14	17	2
9	9	12	14	5
10	10	15	12	8
11	11	22	5	9
12	12	8	4	11
13	13	1	11	22
14	14	5	6	19
15	15	6	20	10
16	16	21	15	6
17	17	2	10	15
18	18	17	13	1
19	19	13	3	20
20	20	10	19	13
21	21	18	7	4
22	22	7	9	12

Table 5.1. Column composition for variables in the four MANA experiments.

The hybrid columns that have significantly better space-filling are hybrid columns 1, 3, 5, and 9. The poorest hybrid columns are hybrid columns 2, 11, 19, and 22. The complete design matrix has an Mm distance of 1.9078, ML_2 discrepancy of 10.1202, maximum pairwise correlation of 0.008, and condition number of 1.037. Since only one of the columns from the third and fourth experiments (see Table 5.1) is required for the third iteration, a comparison using the third or fourth experiment appended to the first two experiments is done. Using the third column from Table 5.1 yields an Mm distance of 1.9422 and ML_2 discrepancy of 13.2759, whereas the fourth column yields an Mm distance of 1.9078 and ML_2 discrepancy of 13.1352. Neither dominates the other, and the third column is chosen for the third experiment.

The predicted mean ER's using (5.3) and the observed mean ER's from the third experiment have a correlation of 0.80 indicating a strong predictive capability. Figure 5.18 illustrates this relationship.

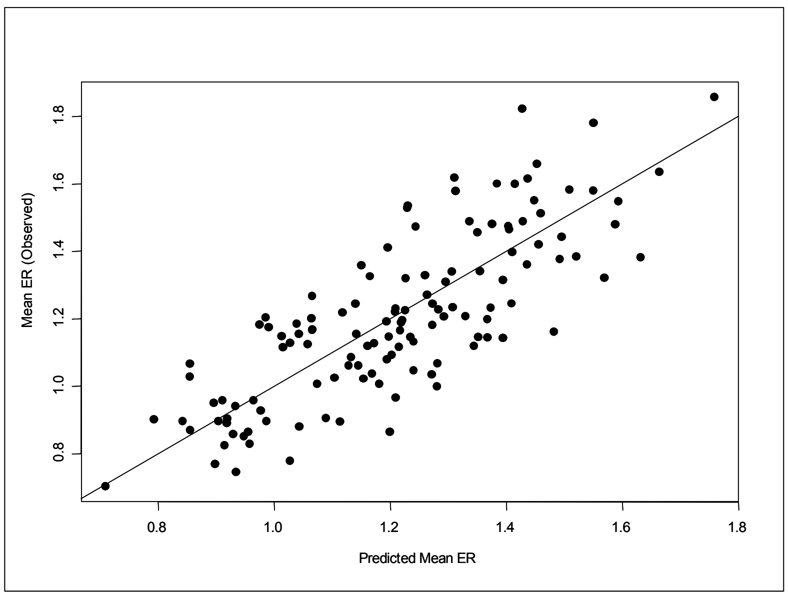

Figure 5.18. Predicted ER versus actual ER for MANA's third peace enforcement scenario experiment resulting in a 0.80 correlation.

Applying regression and data analysis to the third experiment does not yield any new terms that were not already identified in (5.1), (5.2), or (5.3). Furthermore, segregating the best and poorest ER's also does not generate any further insights. As noted previously, the main purpose for executing the third experiment was to demonstrate how to identify additional design points from the design matrix containing both the first and second experiment's design points.

E. SUMMARY

This section summarizes the application of the $(N_O)_{22}^{129}$ design matrix and its permuted designs to a peace enforcement scenario using the agent-based simulation MANA in order to obtain insights suitable for a military decision-maker. The methodology achieved the intended objectives of capturing significant insights from a complex model in an efficient manner. A recap of the notable accomplishments follows.

- The peace enforcement scenario used was assessed as doctrinally correct and plausible by the U.S. Army Infantry Simulation Center at Fort Benning, Georgia.
- Twenty-two variables were incorporated into the analysis, where each variable was sampled uniformly across the applicable ranges. In most agent-based simulation studies, five or fewer variables are used. The $(N_O)_{22}^{129}$ design had design points sufficiently dispersed throughout the entire experimental region.
- The nearly orthogonal designs facilitated regression analysis, and models were built using the output and the author's military experience.
- Applying military expertise and judgment to these results generated significant insights for military decision-makers and illustrated the methodology's strength. This type of analysis is more applicable to military operations than optimizing, predicting, or calibrating.
- The permuting and appending of columns of the design matrix successfully generated additional design points that improved space-filling and strengthened the analysis.
- The design showed an excellent capability for identifying model problems or flaws.

Although the design was used in an agent-based simulation to analyze a military problem, the applicability of these designs to any problem or simulation is evident. The peace enforcement example in this chapter serves as just one illustration.

VI. SUMMARY OF CONCLUSIONS AND FUTURE RESEARCH

This chapter summarizes the contents of the previous chapters and presents a coherent overview of the contributions to the body of knowledge and potential areas of further research. Chapter I provides the motivation for why experimental designs are necessary and important in military simulations. It discusses the trade-off between required resources for conducting experiments and the quantity and quality of information obtainable. The main goal in any experiment is to collect as much quality information as possible while expending minimal resources. The need in military analyses for generating insights or "golden nuggets," instead of strictly predicting, optimizing, or calibrating is articulated.

Chapter II outlines the characteristics desired in an experimental design. The development of orthogonal Latin hypercubes and the importance of space-filling is given. A comprehensive discussion of the measures used to assess concepts of near orthogonality and space-filling is presented. The proposed designs blend these two important properties and offer advantages over other competing designs.

Chapter III is the crux of the dissertation. In it, Ye's [1998] OLHC algorithm is extended to include far more variables (e.g., an 83 percent increase when 129 runs are taken). If some orthogonality is sacrificed, a substantial gain in space-filling can be achieved. An argument follows for examining both the maximum pairwise correlation and the condition number in order to assess the quality of a proposed design matrix. The concept of space-filling is emphasized. Drawing on uniform design theory that previously ignored the issue of orthogonality, we implement the ML_2 discrepancy in conjunction with the Mm distance. All of this was done in order to enhance the ability to discriminate between candidate designs. The proposed designs are listed in the appendices. The merits of the proposed designs are illustrated by comparison to existing designs. Modifications of the proposed designs to incorporate fewer variables are shown. An extensive justification on how additional design points are generated to improve both near orthogonality and space-filling concludes the chapter.

Chapters IV and V illustrate the use of the proposed designs. Chapter IV uses a $(N_O)_{11}^{33}$ design on a known response surface. The advantages of this design over some competing designs is depicted. Chapter V uses a $(N_O)_{22}^{129}$ design for a peace enforcement scenario in an agent-based simulation (MANA). Numerous insights, as well as an extensive data analysis, including regression equations, are generated.

The dissertation extends the field of experimental design by melding near orthogonality and space-filling. Furthermore, the appendices contain ready-to-use designs. The designs are being considered for use by two major Army analytical agencies, CAA and TRAC. Furthermore, two Naval Postgraduate School Operations Research students are using these designs in their master's theses. The major contributions to the existing body of knowledge include:

- Extending the orthogonal Latin hypercube design construction to significantly increase the number of variables examined, while retaining orthogonality or near orthogonality.
- Combining the theory of Latin hypercubes and uniform designs to create design matrices with excellent orthogonality <u>and</u> space-filling properties.
- Constructing an algorithm and using associated measures to assess and then improve the orthogonality and space-filling of design matrices, and increase the likelihood of choosing a best possible design matrix for experimentation.
- Developing an approach that generates additional design points and gracefully handles certain classes of premature experiment termination.
- Illustrating the methodology's applicability and potential by implementing a design with 22 variables in an agent-based simulation.

The major disadvantage of the methodology is that, except for the $(O)_7^{17}$ design, there is no guarantee that the proposed designs are globally optimal. Although better nearly orthogonal and space-filling designs may exist, the listed designs in the appendices are excellent. Their usefulness was demonstrated in Chapters IV and V.

Possible future research in this area is both extensive and exciting. There are two major areas that are particularly worthy of exploration. The first area concerns design matrices that contain both continuous quantitative and qualitative variables. Currently, when a variable contains fewer levels than runs, the levels are used more than once. This method works reasonably well when the number of levels is relatively close to the number of runs. A thorough examination when certain variables have only two or three levels is necessary. This line of inquiry arose from a discussion with the U.S. Army Center for Army Analysis and their value-added analysis for determining which weapon systems will be acquired. In the past, they used Plackett-Burman designs (Loerch et al. [1996]). Recently, they have been using highly fractionated two-level resolution IV designs.[27] We presented the $(O)_7^{17}$ design for their consideration. Unfortunately, they required two variables having only two levels and one variable having three levels. A preliminary methodology was able to achieve a design matrix having a condition number of 1.34 with good space-filling properties. Another major analytical agency (TRAC-White Sands Missile Range) has also expressed similar interest in our designs in their simulation studies of the U.S. Army's Future Combat System. Further research into the effect of having qualitative variables and how to improve the design's near orthogonality and space-filling properties is needed.

The second area concerns sequencing, combining, and crossing the proposed designs with full-factorial, fractional factorial, or group screening designs. One possible approach is to use a nearly orthogonal design for the perceived important variables and a full-factorial, fractional factorial, or group screening design for the perceived non-important variables (or vice versa) to conduct analysis. An investigation of this methodology's ability to find chaotic regions and determine if the *a priori* knowledge of important and non-important variables is correct or incorrect would be beneficial. A further study of how to combine different experimental designs and under what circumstances would be useful. For example, a group screening design, followed by a fractional factorial design, followed by a nearly orthogonal design might be an excellent course of action for a complex model with fewer than 10 variables. Conversely, if there are more than 10 variables, perhaps a nearly orthogonal design followed by a fractional factorial design might be the best approach. This area of research could yield important

[27] From Box et al. [1978], "a design of resolution *R* is one in which no *p*-variable effect is confounded with any other effect containing fewer than *R* - *p* variables."

insights into how experiments should be conducted to gain the most information while expending the minimal resources.

The nearly orthogonal and space-filling experimental designs constructed in this dissertation have demonstrated their usefulness in high-dimensional complex models. The blending of Latin hypercubes and uniform designs, while jointly considering multiple orthogonality and space-filling measures, is an important contribution to the field of experimental design. The actual use of these designs in the MANA scenario shows their value. Presently, two other students are using these designs and the peace enforcement scenario in their research, and two U.S. Army analytical agencies are using or considering the use of these designs in major studies involving billions of dollars. It is the author's hope that these designs continue to merit serious consideration in future military and business analyses.

APPENDIX A. EXAMPLE OF FLORIAN'S [1992] METHOD

We will use the example from Florian [1992]. Assume a design matrix exists with five variables where each variable has 10 levels.

Let the Spearman rank matrix **W** = **T** =

1	3	4	1	5
8	6	10	2	4
5	5	9	3	7
9	4	1	10	3
6	10	7	8	1
10	2	2	6	6
2	1	5	9	10
4	7	6	4	8
7	8	8	7	9
3	9	3	5	2

Rank correlation matrix of **W** = **C** =

1	0.0303	-0.0424	0.309	-0.2
0.0303	1	0.37	-0.0303	-0.467
-0.0424	0.37	1	-0.406	0.224
0.309	-0.0303	-0.406	1	0.00606
-0.2	-0.467	0.224	0.00606	1

In order for **C** = $\mathbf{Q}*\mathbf{Q}^T$, then **Q =**

1	0	0	0	0
0.0303	0.9995	0	0	0
-0.0424	0.3715	0.9275	0	0
0.309	-0.0397	-0.4077	0.8583	0
-0.2	-0.4612	0.4171	0.2559	0.7127

Then, **D** = $\mathbf{Q}^{-1}$ =

1	0	0	0	0
-0.0303	1.0005	0	0	0
0.05786	-0.4007	1.07817	0	0
-0.3339	-0.1441	0.51214	1.16509	0
0.34705	0.9337	-0.8149	-0.4183	1.40311

Then, $\mathbf{D}^T$ =

1	-0.0303	0.05786	-0.3339	0.34705
0	1.0005	-0.4007	-0.1441	0.9337
0	0	1.07817	0.51214	-0.8149
0	0	0	1.16509	-0.4183
0	0	0	0	1.40311

$\mathbf{W}_B = \mathbf{W}*\mathbf{D}^T$ =

1	2.971	3.169	2.448	6.482
8	5.76	8.842	3.916	5.001
5	4.851	7.99	5.715	7.632
9	3.729	-0.0024	8.581	6.064
6	9.823	3.89	9.463	3.763
10	1.698	1.934	4.387	9.613
2	0.9398	5.106	12.23	7.818
4	6.882	3.898	5.39	12.58
7	7.791	5.827	8.763	13.07
3	8.913	-0.195	5.065	7.704

Rearranging the columns of $\mathbf{W}_B$ to correspond to the ordering of $\mathbf{W}$ yields:

1	3	4	1	4
8	6	10	2	2
5	5	9	6	5
9	4	2	7	3
6	10	5	9	1
10	2	3	3	8
2	1	7	10	7
4	7	6	5	9
7	8	8	8	10
3	9	1	4	6

The corresponding correlation matrix of the above matrix is:

1	0.0303	0.01818	-0.006061	-0.07879
0.0303	1	0.006061	0.1394	-0.1394
0.01818	0.006061	1	0.1394	0.0303
-0.006061	0.1394	0.1394	1	0.103
-0.07879	-0.1394	0.0303	0.103	1

Thus, the correlations are reduced. The above procedures may be repeated until there is no further improvement (decrease) in the maximum pairwise correlation and condition number.

Figure A.1 contains S-Plus program code that would enable the reader to implement Florian's [1992] procedure.

```
function(mat, facnum, subnum)
{
	#
	# This function takes a nearly orthogonal Latin hypercube and improves
	# its condition number and maximum pairwise correlation by decreasing
	# both measures.
	#
	#	mat - the incoming matrix
	#	facnum - the number of variables or columns
	#	subnum - the number of levels or runs
	#
	# The returning argument (bettermatrix) is the improved design matrix.
	#
	newmatrix <- matrix(data = NA, nrow = facnum, ncol = facnum)
	for(i in 1:facnum) {
		for(j in 1:facnum) {
			newmatrix[i, j] <- cor(rank(mat[, i]), rank(mat[, j]))
		}
	}
	bettermatrix <- mat %*% t(ginverse(t(chol(newmatrix))))
	for(i in 1:facnum) {
		bettermatrix[, i] <- rank(bettermatrix[, i])
	}
	return(bettermatrix)
```

A.1. S-Plus program code to implement Florian's [1992] procedure that may decrease the maximum pairwise correlation and condition number of the original design matrix.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. A $(N_O)_{11}^{33}$ DESIGN WITH ORDINAL LEVELS FOR THE VARIABLES

33	4	15	16	7	29	23	21	33	20	23
30	33	5	11	13	16	15	7	30	28	25
29	15	30	2	6	2	32	20	11	13	24
19	29	33	3	14	31	4	6	15	8	27
31	2	16	19	8	23	5	24	3	23	14
32	31	11	29	10	15	18	8	2	25	6
23	16	32	30	9	1	1	22	29	10	13
18	23	31	33	12	30	31	9	18	7	8
22	9	8	7	18	24	20	11	20	3	1
25	22	10	13	23	8	7	18	28	2	4
24	8	25	6	32	12	21	3	13	22	5
26	24	22	14	31	25	6	32	8	19	16
20	6	7	26	19	20	10	5	12	1	32
28	20	13	24	29	6	26	19	9	5	31
21	7	28	25	30	13	12	1	24	30	22
27	21	20	22	33	27	25	30	27	18	19
17	17	17	17	17	17	17	17	17	17	17
1	30	19	18	27	5	11	13	1	14	11
4	1	29	23	21	18	19	27	4	6	9
5	19	4	32	28	32	2	14	23	21	10
15	5	1	31	20	3	30	28	19	26	7
3	32	18	15	26	11	29	10	31	11	20
2	3	23	5	24	19	16	26	32	9	28
11	18	2	4	25	33	33	12	5	24	21
16	11	3	1	22	4	3	25	16	27	26
12	25	26	27	16	10	14	23	14	31	33
9	12	24	21	11	26	27	16	6	32	30
10	26	9	28	2	22	13	31	21	12	29
8	10	12	20	3	9	28	2	26	15	18
14	28	27	8	15	14	24	29	22	33	2
6	14	21	10	5	28	8	15	25	29	3
13	27	6	9	4	21	22	33	10	4	12
7	13	14	12	1	7	9	4	7	16	15

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. A $(N_O)_{16}^{65}$ DESIGN WITH ORDINAL LEVELS FOR THE VARIABLES

47	4	24	13	22	9	53	50	41	45	52	32	63	47	36	61
62	47	8	16	28	23	23	17	23	11	36	49	47	60	50	32
58	24	62	31	15	20	18	56	34	58	11	30	27	38	52	57
42	58	47	6	30	5	56	28	14	22	18	18	9	64	59	41
60	31	13	42	2	7	46	11	39	52	17	40	40	32	5	55
35	60	16	58	32	11	12	43	5	4	10	56	58	8	27	51
50	13	35	62	3	17	7	22	48	36	58	28	11	29	23	59
53	50	60	47	21	25	36	60	3	13	46	2	17	21	2	42
45	3	2	22	53	27	34	40	16	65	26	7	52	20	55	26
63	45	32	28	50	1	2	19	55	31	60	25	44	1	44	3
34	2	63	15	35	14	16	29	4	54	21	45	2	16	47	13
64	34	45	30	60	12	60	5	38	15	25	53	21	25	35	17
36	15	22	64	42	26	41	32	12	48	13	19	62	62	10	4
51	36	28	34	58	18	26	63	44	26	4	1	34	35	17	20
38	22	51	63	62	29	38	15	20	56	31	61	10	54	1	29
44	38	36	45	47	10	55	62	64	29	65	37	30	53	24	8
56	29	26	27	9	44	44	65	42	25	51	63	38	39	21	19
37	56	18	1	23	38	24	30	6	46	54	52	24	48	8	1
48	26	37	14	20	51	3	58	57	34	3	27	35	52	32	6
40	48	56	12	5	36	31	12	17	47	29	12	53	57	4	28
54	14	27	40	7	64	65	25	53	27	2	46	20	7	28	12
52	54	1	48	11	34	29	46	10	64	32	58	12	23	60	21
65	27	52	37	17	63	14	9	36	17	47	9	37	22	37	18
39	65	54	56	25	45	61	35	15	60	38	15	65	5	46	14
41	17	7	9	39	53	58	48	1	24	61	11	7	11	18	43
49	41	11	23	65	50	4	13	37	61	42	24	25	30	3	56
55	7	49	20	52	35	27	59	8	9	22	62	50	17	15	39
59	55	41	5	54	60	45	21	47	38	9	43	48	24	25	64
61	20	9	59	40	42	47	14	21	7	27	6	15	56	57	36
46	61	23	55	48	58	17	64	40	43	23	35	5	51	40	44
43	9	46	49	37	62	9	27	7	16	50	50	60	63	53	35
57	43	61	41	56	47	51	42	35	63	59	44	43	40	54	50
33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33
19	62	42	53	44	57	13	16	25	21	14	34	3	19	30	5
4	19	58	50	38	43	43	49	43	55	30	17	19	6	16	34

8	42	4	35	51	46	48	10	32	8	55	36	39	28	14	9
24	8	19	60	36	61	10	38	52	44	48	48	57	2	7	25
6	35	53	24	64	59	20	55	27	14	49	26	26	34	61	11
31	6	50	8	34	55	54	23	61	62	56	10	8	58	39	15
16	53	31	4	63	49	59	44	18	30	8	38	55	37	43	7
13	16	6	19	45	41	30	6	63	53	20	64	49	45	64	24
21	63	64	44	13	39	32	26	50	1	40	59	14	46	11	40
3	21	34	38	16	65	64	47	11	35	6	41	22	65	22	63
32	64	3	51	31	52	50	37	62	12	45	21	64	50	19	53
2	32	21	36	6	54	6	61	28	51	41	13	45	41	31	49
30	51	44	2	24	40	25	34	54	18	53	47	4	4	56	62
15	30	38	32	8	48	40	3	22	40	62	65	32	31	49	46
28	44	15	3	4	37	28	51	46	10	35	5	56	12	65	37
22	28	30	21	19	56	11	4	2	37	1	29	36	13	42	58
10	37	40	39	57	22	22	1	24	41	15	3	28	27	45	47
29	10	48	65	43	28	42	36	60	20	12	14	42	18	58	65
18	40	29	52	46	15	63	8	9	32	63	39	31	14	34	60
26	18	10	54	61	30	35	54	49	19	37	54	13	9	62	38
12	52	39	26	59	2	1	41	13	39	64	20	46	59	38	54
14	12	65	18	55	32	37	20	56	2	34	8	54	43	6	45
1	39	14	29	49	3	52	57	30	49	19	57	29	44	29	48
27	1	12	10	41	21	5	31	51	6	28	51	1	61	20	52
25	49	59	57	27	13	8	18	65	42	5	55	59	55	48	23
17	25	55	43	1	16	62	53	29	5	24	42	41	36	63	10
11	59	17	46	14	31	39	7	58	57	44	4	16	49	51	27
7	11	25	61	12	6	21	45	19	28	57	23	18	42	41	2
5	46	57	7	26	24	19	52	45	59	39	60	51	10	9	30
20	5	43	11	18	8	49	2	26	23	43	31	61	15	26	22
23	57	20	17	29	4	57	39	59	50	16	16	6	3	13	31
9	23	5	25	10	19	15	24	31	3	7	22	23	26	12	16

APPENDIX D. A $(N_O)_{22}^{129}$ DESIGN WITH ORDINAL LEVELS FOR THE VARIABLES

115	32	58	51	34	59	44	89	73	98	72	120	100	98	78	70	129	120	80	124	116	109
98	115	40	56	29	60	13	59	55	27	62	50	119	77	80	75	122	94	104	94	79	117
90	58	98	1	62	36	54	21	97	84	79	74	61	21	63	20	111	128	82	85	108	72
72	90	115	39	33	48	57	98	10	53	35	60	54	49	44	47	127	87	125	100	76	79
91	1	51	72	7	31	14	69	47	120	129	82	15	128	110	87	35	58	57	84	94	113
129	91	56	90	11	2	52	43	76	6	33	16	24	129	81	113	63	41	45	95	98	119
74	51	129	98	4	38	21	30	32	121	124	94	91	14	1	45	15	61	41	88	121	118
79	74	91	115	3	9	45	119	112	3	40	28	64	12	22	13	37	44	56	75	125	61
127	4	7	34	79	27	26	126	94	56	94	110	96	36	77	126	34	122	103	4	43	101
126	127	11	29	74	35	16	14	27	71	26	19	63	18	90	90	16	118	59	48	29	94
119	7	126	62	129	37	41	10	100	29	80	107	22	70	36	28	30	97	121	1	39	93
123	119	127	33	91	50	19	129	66	117	37	41	32	87	33	26	57	109	70	9	26	76
97	62	34	123	72	24	53	93	7	47	85	115	2	28	117	107	94	42	15	62	60	99
68	97	29	119	90	46	12	31	118	70	27	51	3	42	109	63	121	47	28	32	64	87
101	34	68	126	98	30	61	28	5	19	127	104	109	97	31	15	68	39	10	60	19	81
96	101	97	127	115	5	23	67	126	94	32	55	124	80	43	49	76	34	12	37	74	127
125	30	24	27	59	96	6	90	89	64	28	77	81	78	27	116	128	6	101	10	107	10
100	125	46	35	60	101	55	26	20	62	96	6	77	93	57	100	80	11	88	39	123	62
84	24	100	37	36	68	28	5	88	69	31	119	41	63	92	54	124	9	87	41	50	8
106	84	125	50	48	97	17	103	45	21	70	32	8	61	85	7	79	3	78	31	67	47
80	37	27	106	31	123	64	86	40	129	14	83	23	126	3	86	59	106	11	50	128	15
93	80	35	84	2	119	49	17	111	35	125	73	7	75	2	80	17	71	49	8	113	23
95	27	93	100	38	126	43	45	8	122	7	113	104	6	116	33	52	103	69	51	105	22
103	95	80	125	9	127	25	81	105	37	128	52	82	17	120	58	38	85	32	28	83	4
121	38	31	59	103	79	42	128	61	34	57	106	71	56	35	118	14	7	107	119	18	44
92	121	2	60	95	74	20	38	50	105	81	38	85	44	4	121	26	49	117	128	49	20
128	31	92	36	93	129	8	16	78	40	39	108	60	91	66	18	23	50	75	112	30	64
99	128	121	48	80	91	22	97	37	76	126	8	38	105	79	41	5	66	92	107	57	28
82	36	59	99	106	72	47	70	1	49	41	61	5	10	5	114	87	99	17	109	11	57
94	82	60	128	84	90	63	55	96	118	74	46	40	20	48	108	91	129	6	71	41	46
70	59	94	92	100	98	10	51	6	2	17	112	121	104	70	62	84	73	8	101	68	24
71	70	82	121	125	115	18	95	86	99	101	4	95	90	88	39	89	95	34	76	6	52
112	10	47	42	6	44	71	64	117	73	117	67	116	96	74	52	110	84	14	115	1	6
120	112	63	20	55	13	70	52	49	5	43	68	117	123	62	42	85	107	25	78	48	33
67	47	120	8	28	54	94	20	121	106	69	54	31	45	19	71	82	76	19	127	33	41
83	67	112	22	17	57	82	123	15	17	52	103	36	16	25	77	77	92	44	81	15	55
108	8	42	83	64	14	99	108	62	86	76	2	29	100	96	10	31	60	68	106	55	40
122	108	20	67	49	52	128	39	77	79	11	101	19	95	113	31	29	48	84	111	42	2
110	42	122	120	43	21	92	8	29	110	88	12	78	51	55	125	60	20	76	118	34	59
88	110	108	112	25	45	121	111	128	48	21	100	87	68	18	92	27	18	79	97	46	30
105	43	64	6	88	26	103	88	72	14	110	31	112	9	121	73	28	105	2	47	120	9

87	105	49	55	110	16	95	68	3	92	24	93	114	27	93	37	11	67	21	34	110	14
81	64	87	28	122	41	93	36	116	15	107	45	47	101	24	111	69	125	36	58	127	18
66	81	105	17	108	19	80	96	31	119	38	86	12	107	32	106	64	77	1	25	102	32
113	28	6	66	83	53	106	73	38	72	118	43	28	5	115	6	108	17	77	61	69	26
102	113	55	81	67	12	84	3	67	87	55	109	46	22	72	25	117	37	90	63	99	50
75	6	102	87	120	61	100	18	23	22	108	3	80	59	11	94	123	26	97	5	59	34
124	75	113	105	112	23	125	101	91	126	59	95	79	83	76	101	109	19	99	40	86	39
107	61	53	26	44	124	96	124	113	123	53	33	126	57	30	27	90	14	7	14	8	125
69	107	12	16	13	75	101	53	21	52	122	105	74	76	7	64	106	29	4	38	35	95
118	53	69	41	54	102	68	24	114	67	15	39	27	3	83	95	75	55	20	73	24	82
77	118	107	19	57	113	97	72	46	50	100	71	10	41	124	98	112	56	22	64	40	111
111	41	26	77	14	66	123	107	11	89	48	13	44	84	23	2	72	108	100	16	45	74
89	111	16	118	52	81	119	13	70	33	67	125	58	119	39	1	42	98	64	13	13	85
114	26	89	69	21	87	126	9	28	91	25	9	129	48	84	109	25	79	106	20	53	105
104	114	111	107	45	105	127	80	74	28	86	129	68	13	59	122	49	102	114	43	52	92
85	21	14	44	104	88	79	84	108	18	16	49	75	8	28	46	33	13	18	113	126	123
109	85	52	13	114	110	74	25	35	104	84	72	73	38	8	3	47	16	35	104	118	103
78	14	109	54	89	122	129	47	71	23	18	42	25	106	114	74	10	68	37	86	93	129
116	78	85	57	111	108	91	118	51	88	121	123	33	99	89	79	56	40	58	74	114	114
73	54	44	116	77	83	72	115	48	16	66	40	37	19	12	11	118	100	91	103	103	63
76	73	13	78	118	67	90	48	106	85	83	64	17	58	26	34	86	126	127	108	72	88
117	44	76	109	69	120	98	54	43	30	19	14	88	115	69	82	126	115	67	77	109	77
86	117	73	85	107	112	115	74	104	75	120	96	110	66	101	69	98	78	83	123	92	70
65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65	65
15	98	72	79	96	71	86	41	57	32	58	10	30	32	52	60	1	10	50	6	14	21
32	15	90	74	101	70	117	71	75	103	68	80	11	53	50	55	8	36	26	36	51	13
40	72	32	129	68	94	76	109	33	46	51	56	69	109	67	110	19	2	48	45	22	58
58	40	15	91	97	82	73	32	120	77	95	70	76	81	86	83	3	43	5	30	54	51
39	129	79	58	123	99	116	61	83	10	1	48	115	2	20	43	95	72	73	46	36	17
1	39	74	40	119	128	78	87	54	124	97	114	106	1	49	17	67	89	85	35	32	11
56	79	1	32	126	92	109	100	98	9	6	36	39	116	129	85	115	69	89	42	9	12
51	56	39	15	127	121	85	11	18	127	90	102	66	118	108	117	93	86	74	55	5	69
3	126	123	96	51	103	104	4	36	74	36	20	34	94	53	4	96	8	27	126	87	29
4	3	119	101	56	95	114	116	103	59	104	111	67	112	40	40	114	12	71	82	101	36
11	123	4	68	1	93	89	120	30	101	50	23	108	60	94	102	100	33	9	129	91	37
7	11	3	97	39	80	111	1	64	13	93	89	98	43	97	104	73	21	60	121	104	54
33	68	96	7	58	106	77	37	123	83	45	15	128	102	13	23	36	88	115	68	70	31
62	33	101	11	40	84	118	99	12	60	103	79	127	88	21	67	9	83	102	98	66	43
29	96	62	4	32	100	69	102	125	111	3	26	21	33	99	115	62	91	120	70	111	49
34	29	33	3	15	125	107	63	4	36	98	75	6	50	87	81	54	96	118	93	56	3
5	100	106	103	71	34	124	40	41	66	102	53	49	52	103	14	2	124	29	120	23	120
30	5	84	95	70	29	75	104	110	68	34	124	53	37	73	30	50	119	42	91	7	68
46	106	30	93	94	62	102	125	42	61	99	11	89	67	38	76	6	121	43	89	80	122
24	46	5	80	82	33	113	27	85	109	60	98	122	69	45	123	51	127	52	99	63	83
50	93	103	24	99	7	66	44	90	1	116	47	107	4	127	44	71	24	119	80	2	115

37	50	95	46	128	11	81	113	19	95	5	57	123	55	128	50	113	59	81	122	17	107
35	103	37	30	92	4	87	85	122	8	123	17	26	124	14	97	78	27	61	79	25	108
27	35	50	5	121	3	105	49	25	93	2	78	48	113	10	72	92	45	98	102	47	126
9	92	99	71	27	51	88	2	69	96	73	24	59	74	95	12	116	123	23	11	112	86
38	9	128	70	35	56	110	92	80	25	49	92	45	86	126	9	104	81	13	2	81	110
2	99	38	94	37	1	122	114	52	90	91	22	70	39	64	112	107	80	55	18	100	66
31	2	9	82	50	39	108	33	93	54	4	122	92	25	51	89	125	64	38	23	73	102
48	94	71	31	24	58	83	60	129	81	89	69	125	120	125	16	43	31	113	21	119	73
36	48	70	2	46	40	67	75	34	12	56	84	90	110	82	22	39	1	124	59	89	84
60	71	36	38	30	32	120	79	124	128	113	18	9	26	60	68	46	57	122	29	62	106
59	60	48	9	5	15	112	35	44	31	29	126	35	40	42	91	41	35	96	54	124	78
18	120	83	88	124	86	59	66	13	57	13	63	14	34	56	78	20	46	116	15	129	124
10	18	67	110	75	117	60	78	81	125	87	62	13	7	68	88	45	23	105	52	82	97
63	83	10	122	102	76	36	110	9	24	61	76	99	85	111	59	48	54	111	3	97	89
47	63	18	108	113	73	48	7	115	113	78	27	94	114	105	53	53	38	86	49	115	75
22	122	88	47	66	116	31	22	68	44	54	128	101	30	34	120	99	70	62	24	75	90
8	22	110	63	81	78	2	91	53	51	119	29	111	35	17	99	101	82	46	19	88	128
20	88	8	10	87	109	38	122	101	20	42	118	52	79	75	5	70	110	54	12	96	71
42	20	22	18	105	85	9	19	2	82	109	30	43	62	112	38	103	112	51	33	84	100
25	87	66	124	42	104	27	42	58	116	20	99	18	121	9	57	102	25	128	83	10	121
43	25	81	75	20	114	35	62	127	38	106	37	16	103	37	93	119	63	109	96	20	116
49	66	43	102	8	89	37	94	14	115	23	85	83	29	106	19	61	5	94	72	3	112
64	49	25	113	22	111	50	34	99	11	92	44	118	23	98	24	66	53	129	105	28	98
17	102	124	64	47	77	24	57	92	58	12	87	102	125	15	124	22	113	53	69	61	104
28	17	75	49	63	118	46	127	63	43	75	21	84	108	58	105	13	93	40	67	31	80
55	124	28	43	10	69	30	112	107	108	22	127	50	71	119	36	7	104	33	125	71	96
6	55	17	25	18	107	5	29	39	4	71	35	51	47	54	29	21	111	31	90	44	91
23	69	77	104	86	6	34	6	17	7	77	97	4	73	100	103	40	116	123	116	122	5
61	23	118	114	117	55	29	77	109	78	8	25	56	54	123	66	24	101	126	92	95	35
12	77	61	89	76	28	62	106	16	63	115	91	103	127	47	35	55	75	110	57	106	48
53	12	23	111	73	17	33	58	84	80	30	59	120	89	6	32	18	74	108	66	90	19
19	89	104	53	116	64	7	23	119	41	82	117	86	46	107	128	58	22	30	114	85	56
41	19	114	12	78	49	11	117	60	97	63	5	72	11	91	129	88	32	66	117	117	45
16	104	41	61	109	43	4	121	102	39	105	121	1	82	46	21	105	51	24	110	77	25
26	16	19	23	85	25	3	50	56	102	44	1	62	117	71	8	81	28	16	87	78	38
45	109	116	86	26	42	51	46	22	112	114	81	55	122	102	84	97	117	112	17	4	7
21	45	78	117	16	20	56	105	95	26	46	58	57	92	122	127	83	114	95	26	12	27
52	116	21	76	41	8	1	83	59	107	112	88	105	24	16	56	120	62	93	44	37	1
14	52	45	73	19	22	39	12	79	42	9	7	97	31	41	51	74	90	72	56	16	16
57	76	86	14	53	47	58	15	82	114	64	90	93	111	118	119	12	30	39	27	27	67
54	57	117	52	12	63	40	82	24	45	47	66	113	72	104	96	44	4	3	22	58	42
13	86	54	21	61	10	32	76	87	100	111	116	42	15	61	48	4	15	63	53	21	53
44	13	57	45	23	18	15	56	26	55	10	34	20	64	29	61	32	52	47	7	38	60

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Appleget, J.A., "The Combat Simulation of Desert Storm with Applications for Contingency Operations," *Naval Research Logistics*, Vol. 42, No. 4, pp. 691-713, 1995.

Barnes, P. and Steffey, J., "Combat Service Support (CSS) Vector-in-Commander (VIC) Analysis in Support of Force XXI Analyses Division Design Analysis -- Phase III CSS Analysis of VIC Dynamic Gaming Strike Division Interim Design (North East Asia 2.0)," TRAC-TR-1095, Fort Lee, VA, 1997.

Bettonvil, B., "Variable Screening by Sequential Bifurcation," *Communications in Statistics: Simulation*, Vol. 24, No. 1, pp. 165-185, 1995.

Box, G.E.P., Hunter, W.G., and Hunter, J.S., *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*, John Wiley and Sons, New York, 1978.

Chaloner, K. and Verdinelli, I., "Bayesian Experimental Design: A Review," Technical Report #599, Department of Statistics, Carnegie Mellon University, 1994.

Chatterjee, K., Deng, L., and Lin, D., "Resolution V.2 Search Designs," *Communications in Statistics: Theory and Method*s," Vol. 29, Nos. 5 and 6, pp. 1143-1154, 2000.

Clyde, M., Muller, P., and Parmigiani, G., "Optimal Design for Heart Defibrillators," In *Case Studies in Bayesian Statistics II*, C. Gatasonis et al., eds., Springer-Verlag, New York, pp. 278-292, 1996.

Conover, W.J., *Practical Nonparametric Statistics*, John Wiley and Sons, Inc., New York, 1999.

Currin, C., Mitchell, T., Morris, M., and Ylvisaker, D., "A Bayesian Approach to the Design and Analysis of Computer Experiments," ORNL Technical Report 6498, National Technical Information Service, Springfield, Virginia, 1988.

Demmel, J.W., *Applied Numerical Linear Algebra*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1997.

Department of the Army, *Field Manual 100-23 Peace Operations*, Washington, DC, 1994.

Dorfman, R., "The Detection of Defective Numbers of Large Populations," *The Annals of Mathematical Statistics*, Vol. 14, pp. 436-440, 1943.

Fang, K.T. and Wang, Y., *Number-Theoretic Methods in Statistics*, Chapman and Hall, London, 1994.

Fang, K.T., Ma, C., and Winker, P., "Centered L_2-Discrepancy of Random Sampling and Latin Hypercube Design, and Construction of Uniform Designs," *Mathematics of Computation*, Vol. 71, No. 237, pp. 275-296, 2000.

Fang, K.T., Lin, D.K.J., Winker, P., and Zhang, Y., "Uniform Design: Theory and Application," *Technometrics*, Vol. 42, No. 3, pp. 237-248, 2000.

Flourney, N., "A Clinical Experiment in Bone Marrow Transplantation: Estimating a Percentage Point of a Quantal Response Curve, " In *Case Studies in Bayesian Statistics*, C. Gatsonis et al., eds., Springer-Verlag, New York, pp. 324-336, 1993.

Golub, G.H. and Van Loan, C.F., *Matrix Computations*, The John Hopkins University Press, Maryland, 1983.

Hickernell, F.J., "A Generalized Discrepancy and Quadrature Error Bound," *Mathematics of Computation*, Vol. 67, No. 221, pp. 299-322, 1998.

Hickernell, F.J., "Goodness-of-Fit Statistics, Discrepancies, and Robust Designs," *Statistics and Probability Letters*, Vol. 44, No. 1, pp. 73-78, 1999.

Hicks, C.R., *Fundamental Concepts in the Design of Experiments*, Saunders College Publishing, New York, 1993.

Hoeber, F.P., *Military Applications of Modeling: Selected Case Studies*, Gordon and Breach Science Publishers, New York, 1981.

Horne, G. and Leonardi, M., editors, *Maneuver Warfare Science 2001*, Marine Corps Combat Development Command, Quantico, Virginia, 2001.

Hua, L.K. and Wang,Y., *Applications of Number theory to Numerical Analysis*, Springer-Verlag and Scientific Press, Berlin, 1981.

Iman, R.L. and Conover, W.J., "Small Sample Sensitivity Analysis Techniques For Computer Models, With An Application To Risk Assessment," *Communications in Statistics – Theory and Methods*, Vol. A9, No. 17, pp. 1749-1845, 1980.

Ipekci, A.I., "How Agent Based Models Can Be Utilized To Explore And Exploit Non-Linearity And Intangibles Inherent In Guerrilla Warfare," Master's Thesis, United States Naval Postgraduate School, Monterey, CA, June 2002.

Jacoby, J.E. and Harrison, S., "Multi-variable Experimentation and Simulation Models," *Naval Research Logistics Quarterly*, Vol. 9, pp. 121-136, 1962.

Johnson, M., Moore, L., and Ylvisaker, D., "Minimax and Maximin Distance Designs," *Journal of Statistical Planning and Inference*, Vol. 26, pp. 131-148, 1990.

Lauren, M.K. and Stephen, R.T., Map Aware Non-Uniform Automata Version 1.0 Users Manual, Land Operations Division, Defence Science and Technology Organisation, Australia, 2001.

Leon, S.J., *Linear Algebra with Applications*, Prentice Hall, Inc., Upper Saddle River, New Jersey, 1998.

Loerch, A.G., Pudwill, R.A., and LaBarbera, L., *Army Program Value Added Analysis 98-03 (VAA 98-03)*, Volume I, Report CAA-SR-96-1, U.S. Army Concepts Analysis Agency, Bethesda, Maryland, 1996.

Matousek, J., "On the L-2 Discrepancy For Anchored Boxes," *Journal of Complexity*, Vol. 14, No. 4, pp. 527-556, 1998.

McKay, M.D., Beckman, R.J., and Conover, W.J., "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, Vol. 21, No. 2, pp. 239-245, May 1979.

Morrice, D.J., "A comparison of frequency domain methodology and conventional variable screening methods," *Operations Research Letters*, Vol. 17, pp. 165-174, 1995.

Morris, M.D. and Mitchell, T.J., *Exploratory Designs for Computational Experiments*, ORNL, National Technical Information Service, Springfield, VA, 1992.

Morris, M.D. and Mitchell, T.J., "Exploratory Designs for Computational Experiments," *Journal of Statistical Planning and Inference*, Vol. 43, No. 3, pp. 381-402, 1995.

Okten, C., "High Dimensional Simulation," *Mathematics and Computers in Simulations*, Vol. 55, No. 3, pp. 215-222, 2001.

Owen, A.B., "Controlling Correlations in Latin Hypercube Samples," *Journal of the American Statistical Association: Theory and Methods*, Vol. 89, No. 428, pp. 1517-1522, December 1994.

Owen, A.B., "Latin Supercube Sampling for Very High-Dimensional Simulations," *ACM Transactions on Modeling and Computer Simulation*, Vol. 8, No. 1, pp. 71-102, 1998.

Patel, M.S., "Group-Screening With More Than Two Stages," *Technometrics*, Vol. 4, No. 2, pp. 209-217, 1962.

Posadas, S., "Simulation of a Commander's Decision Cycle (SSIM Code)," Master's Thesis, United States Naval Postgraduate School, Monterey, CA, June 2001.

Pukelsheim, F., *Optimal Design of Experiments*, John Wiley and Sons, Inc., New York, 1993.

Rodgers, D. and Prueitt, I., *10K Force Analysis*, TRAC-TR-0793, Fort Leavenworth, KS, 1993.

Sacks, J., Welch, W.J., Mitchell, T.J., and Wynn, H.P., "Design and Analysis of Computer Experiments," *Statistical Science*, Vol. 4, No. 4, pp. 409-422, 1989.

Satterthwaite, F.E., "Random Balance Experimentation," *Technometrics*, Vol. 1, No. 2, pp. 111-137, May 1959.

Schmidt, M., "Friendly Fire and Combat Identification," AMSAA-TR-534, Aberdeen Proving Ground, Maryland, 1992.

Schruben, L.W., "Simulation Optimization Using Frequency Domain Methods," *Proceedings of the 1986 Winter Simulation Conference*, pp. 369-396, 1986.

Shupenas, J. and Armstrong, J., "Optimal Mix of Army Aviation Assets," Technical Report, United States Military Academy, West Point, New York, 1998.

S-Plus User's Manual and S-PLUS Reference Manual, Statistical Sciences, Incorporated, Seattle, Washington, 1991.

Srivastava, J.N., "Design for Searching Non-Negligible Effects," *A Survey of Statistical Design and Linear Models*, Editor: J.N. Srivastava, North-Holland, Amsterdam, pp. 507-519, 1975.

Srivastava, J.N., "On The Inadequacy of the Customary Arrays in Quality Control and General Scientific Experimentation and The Need of Probing Designs of Higher Revealing Power," *Communications in Statistics – Theory and Methods*, Vol. 16, No. 10, pp. 2901-2941, 1987.

Taguchi, G., "*Introduction to Quality Engineering: Designing Quality into Products and Processes*," Asian Productivity Organization, Japan, 1988.

Tang, B., "Orthogonal Array-Based Latin Hypercubes," *Journal of the American Statistical Association: Theory and Methods*, Vol. 88, No. 424, pp. 1392-1397, December 1993.

Tang, B., "Selecting Latin Hypercubes Using Correlation Criteria," *Statistic Sinica*, Vol. 8, pp. 965-977, 1998.

Telephone conversation between LTC Michael McGuire, United States Army, Chief Analytical Simulations, Fort Benning, Georgia and the author, 16 November 2001.

Wild, R.H. and Pignatiello, J.J., "An Experimental Design Strategy for Designing Robust Systems Using Discrete-Event Simulation," *Simulation*, Vol. 57, No. 6, pp. 358-368, 1991.

Wilmer, A., "Javelin Analysis Using Mathematical Modeling," Master's Thesis, United States Naval Postgraduate School, Monterey, CA, June 1994.

Ye, K.Q., "Orthogonal Column Latin Hypercubes and Their Application in Computer Experiments," *Journal of the American Statistical Association – Theory and Methods*, Vol. 93, No. 444, pp. 1430-1439, December 1998.

Youden, W.J., Kempthorne, O., Tukey, J.W., Box, G.E.P., and Hunter, J.S., "Discussion of the Papers of Messrs. Satterthwaite and Budne," *Technometrics*, Vol. 1, No. 2, pp. 157-193, May 1959.

THIS PAGE INTENTIONALLY LEFT BLANK

BIBLIOGRAPHY

Anscombe, F.J., “Quick Analysis Methods for Random Balance Screening Experiments,” *Technometrics*, Vol. 1, No. 2, pp. 195-209, May 1959.

Atwood, C.L., Easterling, R.G., McKay, M.D, Beckman, R.J., Iman, R.L., and Conover, W.J., “Comments and Rejoinder to Comments on Small Sample Sensitivity Analysis Techniques For Computer Models, With An Application To Risk Assessment,” *Communications in Statistics – Theory and Methods*, Vol. A9, No. 17, pp. 1847-1874, 1980.

Bates, R.A., Buck, R.J., Riccomagno, E., and Wynn, H.P., “Experimental Design and Observation for Large Systems,” *Journal of the Royal Statistical Society 58*, No.1, pp. 77-94, 1996.

Bettonvil, B. and Kleijnen, J.P.C., “Searching for Important Variables in Simulation Models with Many Variables: Sequential Bifurcation*,” European Journal of Operations Research*, Vol. 96, pp. 180-194, 1996.

Biles, W.E., “Experimental Design in Computer Simulation,” *Proceedings of the 1979 Winter Simulation Conference*, pp. 3-9, 1979.

Biles, W.E. and Hatfield, I.T., “Simultaneous Variable Screening and Region Reduction in Computer Simulation Experiments,” *Computers and Industrial Engineering*, Vol. 21, No. 1-4, pp. 617-621, 1991.

Booth, K.H.V. and Cox, D.R., “Some Systematic Supersaturated Designs,” *Technometrics*, Vol. 4, No. 4, pp. 489-495, November 1962.

Box, G.E.P. and Wilson, K.B., “On the Experimental Attainment of Optimum Conditions,” *Journal of the Royal Statistical Society B*, Vol. 13, pp. 1-45, 1951.

Box, G.E.P. and Meyer, R.D., “An Analysis for Unreplicated Fractional Factorials,” *Technometrics*, Vol. 28, No. 1, pp. 11-18, February 1986.

Budne, T.A., “The Application of Random Balance Designs,” *Technometrics*, Vol. 1, No. 2, pp. 139-155, May 1959.

Buss, A.H., “Some Extensions and Limitations of Frequency Domain Experiments,” *Proceedings of the 1988 Winter Simulation Conference*, pp. 549-557, 1988.

Butler, N.A., “Optimal and Orthogonal Latin Hypercube Designs for Computer Experiments,” *Biometrika*, Vol. 88, No. 3, pp. 847-857, 2001.

Carlyle, W.M., Montgomery, D.C., and Runger, G.C., "Optimization Problems and Methods in Quality Control and Improvement," *Journal of Quality Technology*, Vol. 32, No. 1, pp. 1-17, January 2000.

Cheng, R.C., "Searching For Important Variables: Sequential Bifurcation Under Uncertainty," *Proceedings of the 1997 Winter Simulation Conference*, pp. 275-280, 1997.

Cochran, J.K. and Chang, J., "Optimization of Multivariate Simulation Output Models Using a Group Screening Method," *Computers and Industrial Engineering*, Vol. 18, No. 1, pp. 95-103, 1990.

Coleman, D.E. and Montgomery, D.C., "A Systematic Approach to Planning for a Designed Industrial Experiment," *Technometrics*, Vol. 35, No. 1, pp. 1-12, February 1993.

Dalal, S.R. and Mallows, C.L., "Variable-Covering Designs for Testing Software," *Technometrics*, Vol. 40, No. 3, pp. 234-243, 1998.

Daniels, C., "Use of Half-Normal Plots in Interpreting Factorial Two-Level Experiments," *Technometrics*, Vol. 1, pp. 311-341, 1959.

Derringer, G. and Suich, R, "Simultaneous Optimization of Several Response Variables," *Journal of Quality Technology*, Vol. 12, No. 4, pp. 214-219, October 1980.

DeWit, M.S., "Identification of the Important Parameters in Thermal Building Simulation Models," *Journal of Statistical Computation: Simulation*, Vol. 57, pp. 305-320, 1997.

Duckworth II, W.M., "Some Binary Maximin Distance Designs," *Journal of Statistical Planning and Inference*, Vol. 88, pp. 149-170, 2000.

DuMouchel, W. and Jones, B., "A Simple Bayesian Modification of D-Optimal Designs to Reduce Dependence on an Assumed Model," *Technometrics*, Vol. 36, No. 1, pp. 37-47, February 1994.

Federer, W.T., "Fractional Replication in Simulation Studies," *Communications in Statistics – Simulation*, Vol. 16, No. 1, pp. 233-237, 1987.

Federer, W.T., "On Screening Variables in the Laboratory and Variables in Factorial Investigations," *Communications in Statistics – Theory and Methods*, Vol. 16, No. 10, pp. 3033-3049, 1987.

Fisher, R.A., "The Arrangement of Field Experiments," *Journal of Ministry of Agriculture*, Vol. 33, pp. 503-512, 1926.

Ghosh, S. and Avila, D., “Some New Variable Screening Designs Using The Search Linear Model,” *Journal of Statistical Planning and Inference*, Vol. 11, pp. 259-266, 1985.

Hamada, M., Martx, H.F., Reese, C.S., and Wilson, A.G., “Finding Near-Optimal Bayesian Experimental Designs via Genetic Algorithms,” *The American Statistician*, Vol. 55, No. 3, pp. 175-181, August 2001.

Herzberg, A.M. and Cox, D.R., “Recent Work on the Design of Experiments: A Bibliography and a Review,” *Journal of the Royal Statistical Society*, Series A., No. 132, pp. 29-39, 1969.

Hunter, J.S. and Naylor, T.H., “Experimental Designs for Computer Simulation Experiments,” *Management Science*, Vol. 16, No. 7, pp. 422-434, 1970.

Ignall, E.J., “On Experimental Designs For Computer Simulation Experiments,” *Management Science*, Vol. 18, No. 7, pp. 384-388, 1972.

John, P.W.M., Johnson, M.E., Moore, L.M., and Ylvisaker, D., “Minimax designs in two-level factorial experiments,” *Journal of Statistical Planning and Inference*, Vol. 44, pp. 249-263, 1995.

Kjell, G., “The Level-Based Stratified Sampling Plan,” *Journal of the American Statistical Association – Theory and Methods*, Vol. 95, No. 452, pp. 1185-1191, December 2000.

Kleijnen, J.P.C., “Screening Designs for Poly-Variable Experimentation,” *Technometrics*, Vol. 17, pp. 487-493, 1975.

Kleijnen, J.P.C., “Design and Analysis of Simulations: Practical Statistical Techniques,” *Simulation*, Vol. 28, pp. 81-90, 1977.

Kleijnen, J.P.C., *Statistical Tools for Simulation Practitioners*, Marcel Dekker, New York, 1987.

Kleijnen, J.P.C., “Simulation with Too Many Variables: Review of Random and Group Screening Designs,” *European Journal of Operational Research*, Vol. 31, pp. 31-36, 1987.

Kleijnen, J.P.C., “Review of Random and Group-Screening Designs,” *Communications in Statistics – Theory and Methods*, Vol. 16, No. 10, pp. 2885-2900, 1987.

Kleijnen, J.P.C., “Sensitivity Analysis and Related Analyses: A Review of Some Statistical Techniques,” *Journal of Statistical Computer Simulation*, Vol. 57, pp. 111-142, 1997.

Kleijnen, J.P.C. and Rubinstein, R.Y., “Optimization and Sensitivity Analysis of Computer Simulation Models by the Score Function Method,” *European Journal of Operational Research*, Vol. 88, pp. 413-427, 1996.

Kuei, C. and Madu, C.N., “Sequential Experimentation and Multilevel Taguchi Designs in Simulation Metamodeling of a Maintenance Float Problem,” *Microelectronics and Reliability*, Vol. 34, No. 5, pp. 831-843, 1994.

Lahkin, M.B., Garcia, L.A., and Nuckols, J.R., “Stochastic Modeling of Exposure and Risk in a Contaminated Heterogeneous Aquifer. 2: Application of Latin Hypercube Sampling,” *Environmental Engineering Science*, Vol. 16, No. 5, pp. 329-343, 1999.

Lam, R.L.H., Welch, W.J., and Young, S.S., “Uniform Coverage Designs for Molecule Selection,” submitted to and revised for *Technometrics*, 2001.

Law, A.M. and Kelton, W.D., *Simulation Modeling and Analysis*, McGraw-Hill Companies, Inc., New York, 2000.

Li, C.H., “A Sequential Method for Screening Experimental Variables,” *Journal of the American Statistical Association* 57, pp. 455-477, 1962.

Lin, D.K.J., “A New Class of Supersaturated Designs,” *Technometrics*, Vol. 35, No. 1, pp. 28-31, February 1993.

Loh, W., “On Latin Hypercube Sampling,” *The Annals of Statistics*, Vol. 24, No. 5, pp. 2058-2080, 1996.

Madu, C.N. and Kuei, C., “Group Screening and Taguchi Design in the Optimization of Multi-Echelon Maintenance Float Simulation Metamodels,” *Computers and Operations Research*, Vol. 19, No. 2, pp. 95-105, February 1992.

Mauro, C.A., “On the Performance of Two-Stage Group Screening Experiments,” *Technometrics*, Vol. 26, No. 3, pp. 255-264, August 1984.

Mauro, C.A., “Efficient Identification of Important Variables in Large Scale Simulations,” *Proceedings of the 1986 Winter Simulation Conference*, pp. 296-305, 1986.

Mauro, C.A. and Burns, K.C., “A Comparison of Random Balance and Two-Stage Group Screening Designs – Part II,” Technical Report 113-14, Desmatics, Inc., State College, Pennsylvania, 1984.

Mauro, C.A. and Smith, D.E., “Variable Screening In Simulation: Evaluation Of Two Strategies Based on Random Balance Sampling,” *Management Science*, Vol. 30, No. 2, pp. 209-221, February 1984.

Mitchell, T.J. and Scott, D.S., "A Computer Program for the Design of Group Testing Experiments," *Communications in Statistics – Theory and Methods*, Vol. 16, No. 10, pp. 2943-2955, 1987.

Montgomery, D.C., "Methods for Variable Screening in Computer Simulation Experiments," Final Report, Office of Naval Research, Contract N0014-78-C-0312, Georgia Institute of Technology, Atlanta, GA, 1979.

Morrice, D.J. and Schruben, L.W., "Simulation variable screening using cross-spectral methods," *Operations Research Letters*, Vol. 13, pp. 247-257, 1993.

Morris, M.D., "Two-Stage Variable Screening Procedures Using Multiple Group Assignments," *Communications in Statistics – Theory and Methods*, Vol. 16, No. 10, pp. 3051-3067, 1987.

Morris, M.D., "Factorial Sampling Plans for Preliminary Computational Experiments," *Technometrics*, Vol. 33, No. 2, pp. 161-174, May 1991.

Morris, M.D., "A Class of Three-Level Experimental Designs for Response Surface Modeling," *Technometrics*, Vol. 42, No. 2, pp. 111-121, May 2000.

Morris, M.D., Easterling, R.G., Johnson, M.E., Ylvisaker, D., Owen, A., Koehler, J., Sharifzadeh, S., O'Hagen, A., and Stein, M.L., "Comments on Design and Analysis of Computer Experiments," *Statistical Science*, Vol. 4, No. 4, pp. 423-435, 1989.

Muller, M., "Supersaturated Designs for One or Two Effective Variables," *Journal of Statistical Planning and Inference*, Vol. 37, pp. 237-244, 1993.

Nychka, D., Yang, Q., and Royle, J.A., "Constructing Spatial Designs for Monitoring Air Pollution using Regression Subset Selection," *Statistics for The Environment 3, Pollution Assessment and Control*, John Wiley and Sons, England, 1997.

Odhiambo, J.W. and Patel, M.S., "On Multistage Group Screening Designs," *Communications in Statistics – Theory and Methods*, Vol. 15, No. 5, pp. 1627-1645, 1986.

Odhiambo, J.W. and Patel, M.S., "Three-Stage Group Screening With Unequal Group Sizes and With Errors in Observations," *Communications in Statistics – Theory and Methods*, Vol. 16, No. 10, pp. 2957-2979, 1987.

Ott, E.R. and Wehrfritz, F.W., "A Special Screening Program for Many Treatments," *Statistica Neerlandica*, Vol. 26, pp. 165-170, 1972.

Owen, A.B., "Orthogonal Arrays for Computer Experiments, Integration and Visualization," *Statistica Sinica*, Vol. 2, pp. 439-452, 1992.

Ozdemirel, N.E., Yurttas, G.Y., and Koksal, G., "Computer-Aided Planning and Design of Manufacturing Simulation Experiments," *Simulation*, Vol. 63, No. 3, pp. 171-191, 1996.

Park, J.S., "Optimal Latin-Hypercube Designs for Computer Experiments," *Journal of Statistical Planning and Inference*, Vol. 39, pp. 95-111, 1994.

Patel, M.S. and Ottieno, J.A.M., "Optimum Two Stage Group-Screening With Unequal Group Sizes and With Errors in Decisions," *Communications in Statistics – Theory and Methods*, Vol. 16, No. 10, pp. 2799-2820, 1987.

Peak, D. and Frame, M., *Chaos Under Control: The Art and Science of Complexity*, W.H. Freeman, Publishers, New York, 1994.

Pignatiello, K.J., "An Overview of the Strategy and Tactics of Taguchi," *IIE Transactions*, Vol. 20, No. 3, pp. 247-254, September 1988.

Plackett, R.L. and Burman, J.P., "Design of Optimal Multifactorial Experiments," *Biometrika*, Vol. 23, pp. 305-325, 1946.

Sacks, J., Schiller, S.B., and Welch, W.J., "Designs for Computer Experiments," *Technometrics*, Vol. 31, No. 1, pp. 41-47, February 1989.

Sanchez, P.J. and Buss, A.H., "A Model For Frequency Domain Experiments," *Proceedings of the 1987 Winter Simulation Conference*, pp. 424-427, 1987.

Sanchez, S.M., Sanchez, P.J., and Ramberg, J.S., "A Framework for Robust System and Process Design Through Simulation," *Concurrent Products, Manufacturing Processes, and Systems*, 1998.

Sanchez, S.M., Moeeni, F., and Sanchez, P., "So Many Variables, So Little Time…A Frequency Domain Approach for Variable Screening," Unpublished Paper, 2001.

Smith, D.E., "Examination of a Parsimonious Variable Screening Technique: Screening K=2N-1-1 Variables in N Runs," *Communications in Statistics – Theory and Methods*, Vol. 13, No. 3, pp. 383-393, 1984.

Smith, D.E. and Mauro, C.A., "Variable Screening in Computer Simulation," *Simulation* 38, pp. 49-54, 1982.

Sobel, M. and Groll, P.A., "Binomial Group-Testing With an Unknown Proportion of Defectives," *Technometrics*, Vol. 8, No. 4, pp. 631-656, November 1966.

Tang, B., "A Theorem for Selecting OA-Based Latin Hypercubes Using a Distance Criterion," *Communications in Statistics—Theory and Methods*, Vol. 23, pp. 2047-2058, 1994.

Toman, B., "Bayes Optimal Designs for Two- and Three-Level Factorial Experiments," *Journal of the American Statistical Association: Theory and Methods*, Vol. 89, No. 427, pp. 937-946, September 1994.

Warnock, T.T., "Computational Investigations of Low Discrepancy Point Sets," *Applications of Number Theory to Numerical Analysis*, ed. S.K. Zaremba, Academic Press, New York, 1972.

Watson, G.S., "Reflections on Group Screening," *Communications in Statistics – Theory and Methods*, Vol. 16, No. 10, pp. 2795-2798, 1987.

Welch, W.J, Buck, R.J., Sacks, J., Wynn, H.P., Mitchell, T.J., and Morris, M.D., "Screening, Predicting, and Computer Experiments," *Technometrics*, Vol. 34, No. 1, pp. 15-25, 1992.

Wu, C.F.J., "Construction of Supersaturated Designs Through Partially Aliased Interactions," *Biometrika*, Vol. 80, No. 3, pp. 661-669, 1993.

Yamada, S. and Lin, D.K.J., "Three-Level Supersaturated Designs," *Statistics and Probability Letters* 45, pp. 31-39, 1999.

Ye, K.Q., Li, W., and Sudjianto, A., "Algorithmic Construction of Optimal Symmetric Latin Hypercube Designs," *Journal of Statistical Planning and Inference*, Vol. 90, No. 1, pp. 145-159, 2000.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center…
   Fort Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Dr. Gary Horne
   Marine Corps Combat Development Command
   Quantico, Virginia

4. Director, TRADOC Analysis Center
   Fort Leavenworth, Kansas

5. Director, Center for Army Analysis
   Fort Belvoir, Virginia

6. Professor Gordon Bradley
   Department of Operations Research
   Naval Postgraduate School
   Monterey, California

7. Professor Thomas Lucas
   Department of Operations Research
   Naval Postgraduate School
   Monterey, California

8. LTC Eugene Paulo
   Department of Operations Research
   Naval Postgraduate School
   Monterey, California

9. Professor Craig Rasmussen
   Department of Mathematics
   Naval Postgraduate School
   Monterey, California

10. Professor Robert Read
    Department of Operations Research
    Naval Postgraduate School
    Monterey, California

11. Professor Susan Sanchez
Department of Operations Research
Naval Postgraduate School
Monterey, California

12. LTC Thomas M. Cioppa
TRAC-Monterey
Naval Postgraduate School
Monterey, California