IMPROVING THE PRACTICE OF EXPERIMENTAL DESIGN IN MANUFACTURING ENGINEERING

By

KHALID A. AL-GHAMDI

A thesis submitted to
The University of Birmingham
For the degree of
DOCTOR OF PHILOSOPHY

School of Mechanical Engineering
The University of Birmingham
May 2011

ABSTRACT

Design of Experiments (DOE) is a powerful technique for understanding, characterising and modelling products and processes and improving their performance. Whilst the bulk of its literature revolves around how it should be applied, little attention, if any, is devoted to the manner in which it is being implemented in practice particularly in manufacturing. One objective of this study was to bridge this gap by reviewing practical applications in three manufacturing journals. This revealed not only limited use but also multiple deficiencies. Many of these concerned a lack of familiarity with the concept of aliasing; the use of fractional factorial designs and pooling methods to analyse unreplicated trials; and a misunderstanding of the concepts underpinning the use and interpretation of p-values and factorial effects' importance measures. With respect to aliasing, a novel simple method for generating its pattern is proposed. Besides its ease of application, it can be linked to the three main criteria for measuring the degree of aliasing (maximum resolution, minimum aberration and generalised minimum aberration) in a manner devoid of mathematical complications. Regarding the use of fractional factorial designs and pooling methods, simulation experiments were used to assess the performance of certain experimentation strategies to arrive at the same conclusions had a full factorial trial been performed. In the context of two-level designs, the L_{16} together with the Pooling Up method or the Half Normal Probability plot yielded a satisfactory performance. Similarly, the strategy of using the Best Subset selection procedure in conjunction with the L_{18} design was the best among the examined three-level ones. To attain a robust performance, it was found that the use of small designs such as the L_8 and the L_9 should, as far as possible, be avoided. The concepts concerning the use of the p-values and the effect's importance measures are clarified and to facilitate communication between Engineers, Managers and Statisticians, an importance measure that can be related to three quality engineering techniques is suggested.

DEDICATION

This work is dedicated to my parents for their love and endless support

ACKNOWLEDGEMENT

Firstly, I would like to express my sincerest appreciation to my major supervisor Mrs. Elaine Aspinwall to whom I owe a huge debt of gratitude. I am indebted to Mrs Aspinwall for her patience, invaluable advice, guidance, support, time, encouragement and sharp attention to details. Working under the supervision of Mrs. Aspinwall made a substantial contribution to my academic career and taught me that learning can be fun. Her insightful suggestions and advice will always be remembered and her great kindness will be deeply ingrained in my heart.

I am also very grateful to my co-supervisor Mr. David Aspinwall for his very instructive guidance regarding the planning and execution of the EDM experiments and also for being very helpful and supportive especially during the writing up period. I would like also to thank Mr. R. Fasham for providing the technical support needed for the experimental work.

My special thanks are due to my employer, King Abdul Aziz University for sponsoring this research and providing the necessary funds and financial support, without which this work would not have been completed. I also would like to thank all staff members of the School of Mechanical Engineering at the University of Birmingham who provided me with academic/ administrative assistance.

Finally, I am very thankful to my colleagues in the Machining Research Group and all my friends in the beautiful city of Birmingham for the very nice time we spent together.

TABLE OF CONTENTS

List of Figures

List of Tables

Acronyms and Abbreviations

ANOVA	Analysis of Variance
BE	Backward Elimination
BM	Box Method
BSS	Best Subset Selection
CA	Control Array
CCD	Central Composite Design
CIRP-MT	CIRP Annals - Manufacturing Technology
df	degrees of freedom
DFSS	Design For Six Sigma
DMAIC	Define-Measure-Analyse-Improve-Control
DO	D-optimal
DOE	Design of Experiments
DSS	Decision Support Systems
FE	Factorial Effects
FS	Forward Selection
HNP	Half Normal Probability
HT	Hypothesis Testing
IDOV	Identify-Design-Optimise-Verify
IJMTM	International Journal of Machine Tools and Manufacture
JMPT	Journal of Materials Processing Technology
LM	Lenth Method
MLM	Modified Lenth Method
MSE	Mean Square Error
NA	Noise Array
NPP	Normal Probability Plot
OA	Orthogonal Arrays
PC	Percentage Contribution
PD	Pooling Down
PP	Pareto Plot
PSE	Pseudo Standard Error
PU	Pooling Up
ROT	Rules of Thumb
S/N	signal-to-noise ratio
SES	Standardised Effect Size
SPC	Statistical Process Control SPC
SS	Sum of Squares
SWR	Stepwise Regression
SWR	Stepwise Regression
TS	Test Statistic
UC	Unassigned Columns
UC	Unassigned Columns
VIF	Variance Inflation Factor

CHAPTER 1: INTRODUCTION

1.1 Background

Organisations are facing rapid and continuous technological, economical, social and regulatory changes that impact their ability to remain competitive. With markets globalisation, business competition is very severe and attaining customer satisfaction and delight becomes an essential requirement for an organisation to survive. The breakthrough improvement in information technology has led to increase customer awareness of the varied products, services and levels of quality available. Consequently, the task of meeting the expectations of today's well informed customers and gaining their loyalty requires organisations to steadily improve the quality of their products and services. In addition they need to provide high quality goods and services to the customers who want it, where they want it, and in the quantity and at the price they want it.

Quality here is reflected by what Goh (2000) called the universal set of criteria: on-target performance, least variation and minimum cost. Meeting these criteria entails making vital decisions to improve an organisation's products and processes based on the recognised needs and the identified objectives. The latter include productivity, efficiency, response time and cost targets which should be translated into performance measures quantifying the extent to which the specified objectives are met. The process or product parameters that are known, or at least are believed to affect these measures, should then be identified. Finally, a set of feasible alternative courses of actions in terms of the different combinations of the values of the product/process parameters should be formulated.

As is the case in any decision situation, each course of action can have a significant effect on the performance of the entity under study and there is always some doubt as to which alternative should be selected. One option is to make the decision on the basis of past experiences and trial and error. As this relies on subjective personal judgment and opinions this approach suffers from problems such as cognitive bias and the limited capability of human brains to deal with complexity. Consequently it can result in serious time and financial waste besides threatening an organisation’s survival in today’s highly competitive market.

Alternatively a scientific data-driven approach can be adopted so that decisions regarding product or process improvement can be made on the basis of facts rather than subjective feelings and opinions. In this respect, an effective methodology is the Design of Experiments (DOE). This is a cost effective technique for identifying the key product/process parameters and establishing a functional relationship between them and the appropriate performance measure so that their settings can be adjusted to bring the performance measure on-target and minimise its variation around that target. Utilising DOE, organisations can improve the quality of their products while lowering the cost of production. Furthermore, they can bring products to market quicker and at minimal research and development costs (Antony, 2001).

Although DOE can be utilised in any area, Manufacturing Engineering is the main field around which the scope of its study in this thesis is centred. This is due to the pivotal role this branch of engineering plays in any industry. In fact, Manufacturing Engineering is concerned with providing the know-how needed to make a product. According to Finkin (1987), it has the principal responsibility within a company for attaining meaningful cost reductions in products through changes in the methods, processes and equipment used in their manufacture.

1.2 Motive, Objectives and Methodology

Most studies in the DOE literature tend to focus on how the technique should be applied. Far too little attention has been devoted to investigating the way in which it is actually being employed particularly in Manufacturing Engineering. The initial interest and motivation to carry out this research stem from the need to examine the extent of agreement between what is practiced and what is "preached" regarding the application of DOE. Its aim is to attain a better understanding of the technique by suggesting improvements to its practice in Manufacturing Engineering thereby helping practitioners assimilate how to foster this powerful technique in their product and process improvement endeavours.

The steps towards this end can be represented by the following objectives:

- Highlight the importance of DOE as an effective technique for performance improvement;
- Review DOE methodology and its practice in Manufacturing Engineering;
- Investigate any differences between the way in which it is applied in Manufacturing and that recommended in the DOE literature;
- Identify the aspects that warrant further clarification and assimilation and relate them to potential gaps in the literature ;
- Propose methods and provide suggestions and recommendations to bridge any gap identified and mitigate the impediments to an effective use of DOE in practice;

The methodology adopted to meet these objectives involves conducting a literature review of DOE and its applications in Manufacturing Engineering; performing simulation experiments to assess the performance of certain DOE strategies; examining published experimental data; and carrying out physical trials in an effort to assess the substantiality of the conclusions arrived at using the simulation experiments.

1.3 Structure of the Thesis

The thesis comprises nine Chapters. Following the introduction, Chapter 2 provides an overview of the main theoretical concepts of DOE and the fundamentals underpinning its use. The importance of DOE as a powerful performance improvement technique is also emphasised.

Chapter 3 presents a review of how DOE has been applied in three manufacturing journals. The need for the review is emphasised and its methodology explained together with the aspects investigated. Its findings are also presented and discussed, and the practices of DOE that need improvements are highlighted. Some of these are dealt with in the Chapter whereas others are handled in subsequent ones.

Aliasing which was found to be one of the main misconceived topics in the conducted review is addressed in Chapter 4. Its nature and underlying concepts are discussed together with its impact on the experimental inferences. A simple method is proposed for constructing its pattern that can be used with two- and three-level designs. Also highlighted is the way in which the postulated method can be used to employ the main criteria for measuring the degree of aliasing.

Chapter 5 describes the planning stage of the simulation experiments that were conducted to assess the performance of certain DOE strategies. It points out the main objectives of the performed trials along with the performance measure that was adopted in their execution. The manner in which the examined factors and their levels were specified and the simulation methodology are also presented. The results of the experiments are highlighted and discussed in Chapter 6.

In Chapter 7, the credibility of the inferences drawn from the simulation trials is assessed. For this purpose, data from published and performed experiments are utilised. The conclusions arrived at using the latter are compared with those that were drawn from the simulation experiments and the reasons for any disagreement outlined.

The issues regarding the p-value and the factorial effects' importance measures are the focus of Chapter 8. The misuses of the former and the problem associated with its misinterpretation are indicated and dealt with. The Chapter also investigates the underpinning concepts of two importance measures that are frequently confused in practice. The rational, derivation and bias of each quantity is examined and used to compare their performances. A third measure that is rarely employed is recommended and its usefulness with regard to improving the communications between Engineers, Managers and Statisticians is highlighted.

The main conclusions from this research are outlined in Chapter 9. It comprises an aggregation of this study's chief findings and outcomes, and culminates with a discussion of its limitations and suggestions for future research directions.

Apart from the first and the last Chapters, each one is prefaced with a brief introduction that offers an overview of its major theme and objectives and each concludes with a summary of its main findings.

CHAPTER 2: DESIGN OF EXPERIMENTS

2.1 Introduction

DOE is a powerful technique for improving product or process quality and reducing their performance measures variability around the targeted level in an effective and efficient manner. It is an effective approach for exploring, understanding and establishing the causal relationship between the system parameters and their performance measures. The execution of an experiment involves purposefully changing the parameters that are known or at least believed to affect the performance of a studied process so that knowledge about their performance can be obtained, extended or verified. DOE was initially developed by Sir R. A. Fisher at Rothamsted Agricultural Station in England in the early 1920s (Fisher, 1971). Its emergence marked a significant change from the old "scientific" tradition of varying only one parameter at a time to the simultaneous examination of the studied parameters in order to assess their various interaction effects in a cost effective way. Since its evolution, DOE has undergone a remarkable development involving its design and data analysis aspects. This Chapter will cast some light on the main theoretical concepts and ideas that underline DOE, its bulk being devoted to the importance of the technique which has been accorded scant attention in the literature. DOE is firstly defined together with its associated jargon. The importance of DOE and the role it plays in the product and process improvement endeavour are then introduced. This is followed by a discussion of the types of DOE and its history. The Chapter culminates with a summary of its main conclusions.

2.2 What is Statistical Experimental Design?

An experiment can be defined as "a test under controlled conditions that is made to demonstrate a known truth, examine the validity of a hypothesis, or determine the efficacy of something previously untried” (Elias, 2007). Montgomery (2010) defined it as “a test or series of tests in which purposeful changes are made to the input variables of a process or system so that it may be possible to observe and identify the reasons for changes that may occur in the output response”. The process of planning an experiment so that appropriate data that can be analysed by statistical methods will be collected, resulting in valid and objective conclusions is referred to as DOE (Montgomery, 2010). Antony and Kaye (2000) defined DOE as a powerful technique used for discovering the most influential process’s parameters and determine at what levels they must be set to optimise the associated performance measures. Throughout the remainder of this thesis, the terms product and process are used interchangeably since what is discussed is equally applicable to both.

Clearly, none of the above presented definitions is complete. In fact, each of them focuses on important aspects of DOE. Consequently, in this research they are viewed as complementary to each other.

As is inevitable with any subject, certain terms are used in the literature when discussing DOE. Since these will be used throughout the remainder of this thesis, the most important will now be explained.

1. **Response** is the experimental result corresponding to the setting of the process’s parameters (Anderson and McLean, 1974). It represents a key measurable characteristic of a process performance that reflects the customers’ needs (Barrentine, 1999). In quality engineering studies, the response is referred to as a Critical to Quality characteristic (Goh and Xie, 2004) or a Performance Measure.
2. **Factors** are those process’s parameters that are expected to have some influence on the response. They can be quantitative or measurable (e.g. temperature, pressure, time, etc.) or qualitative or categorical (e.g. different material suppliers,

or presence/absence of some attribute etc.). Factors can be classified as either design /control or nuisance/noise. The former are those that can be easily controlled not only while experimenting but also in day-to-day operations, while the latter are ones which are difficult or impossible to control or those which vary during the use of the product or process such as ambient temperature and humidity. In this work the terms factor and parameter will be used interchangeably.

3. **Levels** are the values that a factor can assume or take; they are sometimes referred to as treatments especially in agricultural, medical and biological experiments (Hicks and Turner, 1999). The simplest designs use only two levels of each factor. In the absence of non-linear responses, this will usually provide the information needed in the most efficient manner. The levels may be fixed i.e. specific levels of interest; which means that the statistical inferences made about the factors are confined to these specific values, or they can be random i.e. chosen at random from a larger population of possible levels. In this case the experimenter can draw conclusions about the entire population of levels, not just those that were used in the experimental design.
4. **Effect** is the change in the average response rendered by a change in the levels of a factor. This is frequently called a main effect because it refers to the primary factors of interest in the experiment.
5. **Interaction** is the failure of one factor to produce the same effect on the response at different levels of another factor i.e. the difference in response between the levels of one factor is not the same at all levels of the other factors.
6. **Experimental unit** is an object, material or unit to which treatments are applied.

2.3 Performance Improvement Approaches

The business world is being transformed worldwide as a result of globalisation. Managers have been complaining of the increasing competition pressures and its associated high cost (Mintzberg, 1994). The substantial improvements that have been made in information technology means that customers are more informed today than ever before and so their demands for better products and more responsive services have greatly increased. Improved products and services constantly raise the level of expectation for the next generation. This has led to rapidly changing and highly competitive business environments in which organisations must steadily improve the quality of their products and services to ensure their profitability and survival. To be competitive in such an environment, organisations need to effectively survey the market so that the customer needs, wants and expectations from an existing or even a potential product can be defined or predicted. Broadly, these needs fall under the categories of timely delivery, competitive pricing and zero-defect quality. Having identified the customer's requirements, organisations need firstly to translate them into process performance metrics such as cycle times, operational costs and defect rates and secondly to determine the target performance level for each of the established measures. In order to attain the required levels with minimum variation, critical decisions have to be made regarding the organisation's processes pertaining to the identification of the most influential parameters and their settings. Broadly, there are two approaches to making such decisions: experiential and data-driven.

2.3.1 Experiential Approach

The decision about how to improve a process using this approach relies on personal judgement utilising experience or using a trial and error method. Being dependent primarily on past experience and previous decision situations that are similar to the one currently confronted, Ackoff (1999) calls this approach clinical. Although its use may involve utilising formal observed quantitative data, they are incorporated into a pool of attitudes and beliefs that is dominated by the qualitative experience outputs. Ackoff et al (2005) noted that most managers and many consultants adopt this attitude in decision making. The outcomes of implementing this approach can be "satisficing"; a term combining the ideas of satisfying and optimising coined by Simon (1976) to characterise solutions that are good enough but not necessarily the best. In this era, where customer expectations for product integrity have never been higher and where constant market competition pressures have led to minimum standards for quality and reliability below which it is impossible for any organisation to survive, this subjective approach can lead to results that are less than attractive. In fact it can be wasteful in terms of time and money as there is no guarantee that the best results will be attained. A further, rather more serious, limitation of this approach pertains to the fact that the accelerating rates of technological, market and social change has rendered the experience no longer the best teacher. Ackoff, (1999) stated that it is not even a good teacher as it is too slow, too ambiguous and too imprecise. Solberg (1988) admonished managers not to use the experiential approach and emphasised the risk of its associate trial and error practice in studying Manufacturing systems. He stated that:

"The ability to apply trial and error learning to tune the performance of Manufacturing systems becomes almost useless in an environment in which changes occur faster than the lessons can be learned. There is now a greater need for formal predictive methodology based on understanding of cause and effect"

Considering the above highlighted limitations, one is entitled to infer that the decision making process can be improved by replacing the experiential, qualitative, judgemental and satisficing approach with a more formal, experimental, quantitative, scientific and "optimising" one.

2.3.2 Data-Driven Approach

Effective decisions concerning process improvements can be made when a rational basis such as collected data is utilised. However, as Hunter (1986) observed, data can sometimes be collected mindlessly resulting in numbers that are difficult to interpret or reflect upon. Thus careful thought must be given to what data to collect and how they are going to be analysed (Hines et al, 2003). Statistics is a science concerned with the collection, analysis, interpretation and presentation of data (Montgomery and Runger, 2010). Consequently it is instructive to use statistical tools to gather and objectively analyse data so that judicious decisions regarding the necessary process changes are made on the basis of facts rather than subjective feelings or opinions.

Broadly, the use of statistical tools for process improvement (particularly in Manufacturing) has progressed in three stages. The first is product inspection, where sampling plans were used to sort good from bad products at the final inspection stage of production. The objective was merely to detect products not conforming to specifications and prevent them from reaching customers. At the next stage, attention was turned upstream to the process (on-line) that generated the products. The main concept that underpins the on-line quality and performance improvement effort is that every process generates information that can be used to improve it (Bisgaard, 1990). Such information may be obtained by informative observational statistical techniques such as Statistical Process Control (SPC) (Box and Bisgaard, 1987). Hunter (1986) called this "listening" to the processes. According to Goh (1993) and Grant and Leavenworth (1996) this approach is effective in preventing the generation of unsatisfactory products but it is passive as no attempt is made to change the process so that its performance is improved. The attainment of the latter requires "conversing" with processes by purposefully changing their parameters to observe and identify the reasons for any associated changes that may occur. DOE is an effective technique for accomplishing this task. Attention in the third stage was moved further upstream to the design stage (off-line) aiming to prevent the occurrence of any problems i.e. doing the right thing first time. DOE is the dominant technique at this stage as its effective use can result in having "built in" product and process quality thereby eliminating the need for final product inspection and reducing the on-line quality endeavour.

2.4 Importance of Experimental Design

A process, as schematically represented in Figure 2-1 (Montgomery, 2010), is a set of connected activities in which inputs are transformed into outputs for a specific purpose (Hoerl and Snee, 2002). Some of the process's parameters (X_1, X_2….X_n) are controllable, whereas others (N_1, N_2,….N_k) are uncontrollable (although they may be controllable during experimentation). There is no direct control over the performance measure (response variable) Y_i; in the classical cause and effect approach, it is the effect. The causes are what dictate the response. To control the response, the causes i.e. the process's variables X_1, X_2….X_n must be controlled. Both manufacturing and service systems may be viewed as processing systems as they process items through a series of activities. In Manufacturing, raw materials (inputs) are transformed into finished products (outputs); whereas in services, customers enter with some service need (inputs) and depart as serviced customers (outputs). The performance measures, Y_i, are derived from the process output. For example in Manufacturing the surface roughness of a finished product may be used as a performance measure while the customer waiting times in, say, a Bank can be used to measure the performance of its service system.

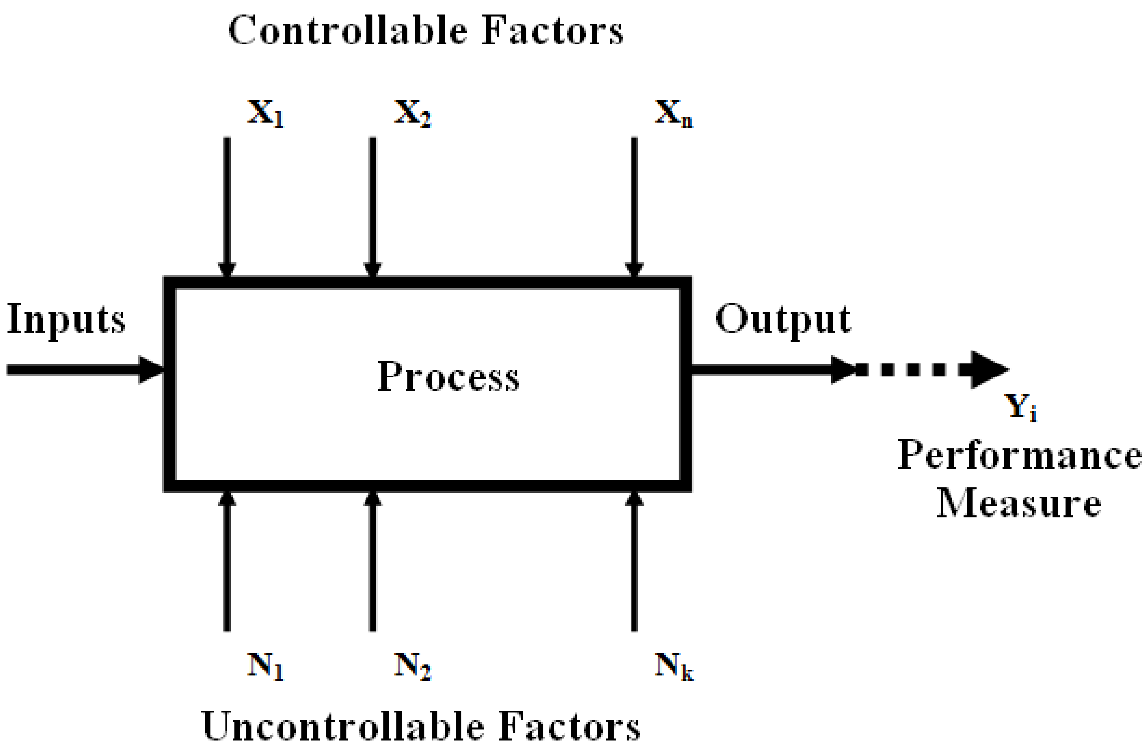


Figure 2-1: General Schematic Representation of a Process (adapted from Montgomery (2010))

DOE can be used to improve the performance of Manufacturing and service processes in many ways as discussed below.

2.4.1 Screening

Processes are often described by many parameters, the role of which may not be understood especially at the early stages of the design and development effort when the relevant scientific and engineering knowledge may be limited. Applying DOE to identify the most influential controllable process's parameters (factors) is critical to its eventual successful improvement. Without a proper screening experiment a considerable amount of time may be wasted in guessing which parameters are important, in lengthening the performance improvement lead time and in possibly missing important deadlines regarding the improvement projects. Therefore, screening experiments are essential for both on-line and off-line improvement endeavours. Goh (1996) suggested that in such experiments all conceivable controllable parameters X_1, X_2….X_n should be included so that the key ones that exert considerable effect on the performance measure of interest Y_i can be detected.

2.4.2 Characterising

DOE can be used to identify how the subset of controllable factors, already identified as important, should be adjusted so that the average value of the associated performance measure could be shifted to its desired level (Goh, 1989). DOE can also be used to identify where to set the influential X's so that the variability of Y_i around its target level is minimised (Kacker, 1985).

2.4.3 Optimising

Another use is to manipulate the most influential process's parameters to levels or settings that result in the "best" obtainable set of operating conditions. If it is the case that the larger the value of Y_i the better the performance, the setting of the influential X's to maximise Y_i must be determined. Alternatively if the best performance is associated with the smallest value of the performance measure, then the settings rendering the minimum Y_i must be specified.

2.4.4 Dealing with Complexity

According to Gershenson and Heylighen (2005) complexity stems from the Latin complexus, meaning entwined or connected together. Although a process may consist of many activities, parameters and resources, it is not the number of these elements that make it complex. It is their level of interdependences and variability. It may be quite easy to understand and predict the effect of each of the process's parameters when taken in isolation from the others. However, the performance of a process does not depend on the independent effect of each parameter but on how each interact with the rest to affect its behaviour. The effect of simple interdependences may be easy to comprehend in the absence of variability in the activities involved. Similarly the variability of the activities that operate independently of each other can be simple to analyse. In fact it is the combination of interdependences and variability that produce a complex unpredictable process performance. Unaided, the human mind is not very good at analysing and understanding complex processes. Simon (1957) referred to this as "the principal of bounded rationality" and stated that "the capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problem whose solution is required for objectively rational behaviour in the real world, or even for a reasonable approximation to such objective rationality" (Simon, 1957). This is in agreement with the observation of the philosopher Alfred North Whitehead that "we think in generalities; we live in details" (Ceder, 2007). As Goh (1999) explained, most of the processes are not only too complex to be understood by the unaided human mind but also by the theoretical principles of Science and Engineering. Dealing with complexity requires both a philosophy and powerful techniques to put it in action. The philosophical requirement is embodied in statistical thinking (Britz, 2000; John et al, 2001; Hoerl and Snee, 2002) which is best and more effectively implemented using DOE. Statistical thinking is a philosophy of learning and action based on the following fundamental principles (ASQ, 1996):

- all work occurs in a system of interconnected processes;
- variation exists in all processes;
- understanding and reducing variation are keys to success.

Snee (1990) defined statistical thinking as " thought processes which recognise that variation is all around us and present in everything we do, all work is a series of interconnected processes, and identifying, characterising, quantifying, controlling and reducing variation provide opportunities for improvement". Contrary to the traditional view of statistics as a set of tools and techniques, statistical thinking views the work as a network of interdependent processes that are joined together to accomplish certain goals. Of vital importance to its effective use is the ability to understand the degree and nature of the relationship and interdependence between the work processes and between the input, activities and output of each process which is an essential aspect of dealing with complexity. To this end, DOE is a powerful technique for gaining knowledge relating to the interdependences (interactions) between a process's parameters X_1, X_2….X_n , their joint effect on and functional linkage with the performance measure of interest Y_i. DOE enables the experimenter to vary all the studied parameters simultaneously hence allowing their interactions, which are forms of interdependence, to be investigated utilising small numbers of experiments.

Another way in which DOE deals with interdependence is through its ability to understand and empirically model the cause-and-effect relationship between the process's parameters and the performance measure in an effective and efficient manner (Montgomery, 1999). This enables the experimenter to judiciously manipulate the influential parameters to attain the target performance level. There are two types of cause-and-effect models: deterministic and probabilistic (Harvey, 2001). In the former, the process's parameters are the necessary and sufficient variables for determining the value of the response or the dependent variable. Although such models are very common in scientific areas such as physics, they are rarely encountered in engineering practice. On the other hand, probabilistic cause-and-effect models in which the process's parameters are the necessary but not sufficient variables for identifying the value of the response are commonly dealt with in practice. In fact most of the DOE empirical models are of this type. Mathematically, they can be represented as follows:

$$Y_i = f(X_1, X_2 \ldots X_n) + e \qquad (2.1)$$

where e is a surrogate for all those variables that are omitted from the model but that collectively affect Y_i . As Box (1993) stated "knowledge about our world is, and always must be, partial knowledge". Consequently, no model is perfect. The term e includes (i) all the noise variables either known or unknown, controllable or uncontrollable; (ii) factors and interactions that have such a small effect that their inclusion in the model is not practical; (iii) the error in measurements of the studied factors and the observed response along with the latter intrinsic unexplainable variability; and (iv) the error resulting from using a wrong functional mathematical model to link the Xs with Y_i such as using a linear equation to represent a quadratic relationship.

In addition to modelling process's parameters interdependences, the development of a formal empirical model provides an explicit and, to a large extent, an unambiguous representation of the process and offers a common language for communicating the findings to those involved in the decision making or evaluation process. Furthermore, it facilitates a what-if analysis which is very useful in devising robust recommendations (Bouyssou et al, 2000). The way in which statistical thinking and DOE deal with the second aspect of complexity (i.e. variability) is discussed next.

2.4.5 Dealing with Variability

DOE is a vital technique for studying and reducing processes' variability which is a major aspect of complexity and an essential component of the statistical thinking philosophy. Variability is a characteristic inherent in any process (Britz at al,1997; Makrymichalos et al, 2005). For example different people can perform the same activity in different ways; different suppliers of supposedly identical inputs can provide variable inputs; different machines that are supposed to perform identically can be inconsistent. A fundamental principal of statistical thinking that is rooted in work of the quality pioneer Deming (1986) is variance understanding and reduction. In fact, Deming's 14 points are all based on understanding variation (Hare et al, 1995). He once said "If I had to reduce my message for management to just a few words, I would say it all had to do with reducing variation" (Stupak and Leitner, 2001). Certainly, manufacturers prefer working with suppliers whose performance is stable; employees value working for managers with predictable behaviours and actions and managers in turn reward employees whose performance level is consistent around their expectations.

Deming (1986) highlighted two approaches for reducing variation: problem solving and process improvement. The former concerns the elimination of special causes of variation i.e. causes that result in an unanticipated and excessive variation causing a process to become unstable. In this context, DOE is very effective in identifying the influential X's so that the variability of the average value of the performance measure Y_i around its target is minimised. The process improvement approach revolves around the reduction of the common cause variation that is inherent in the process. By incorporating the noise factors into the experiment, DOE can be used to determine the settings of the important X's so that the common cause variability, to which Deming attributes more than 85% of the processes' problems (Snee, 1990), is minimised. Figure 2-2 shows schematically how variation reduction can be used to improve product or process quality. Another approach to variability reduction is to anticipate its sources and then design processes that are insensitive or robust to them. This predict-and-prepare approach focuses on designing processes with high "built in" quality and productivity rather than transforming the existing ones to attain such capability i.e. prevention rather than cure. To this end, the DOE approach pioneered by Taguchi (Taguchi and Wu, 1985; Taguchi et al, 1989) is of vital importance. Taguchi advocated using DOE for parameter design, which aims to predict the sources of variation at the design phase and use experimental design to identify the settings of the process's parameters that

- make the process robust to environmental, and other factors that are difficult, expensive or even impossible to control;
- eliminate or reduce the process sensitivity to the variation transmitted from its components;
- attain the target performance level with minimum variation.

The conventional approach to achieving these objectives is to use high quality components with tighter tolerances, advanced technologies or expensive input material. This can result in overdesigned and expensive products. The power of DOE lies in its ability to attain these objectives without additional capital investment.

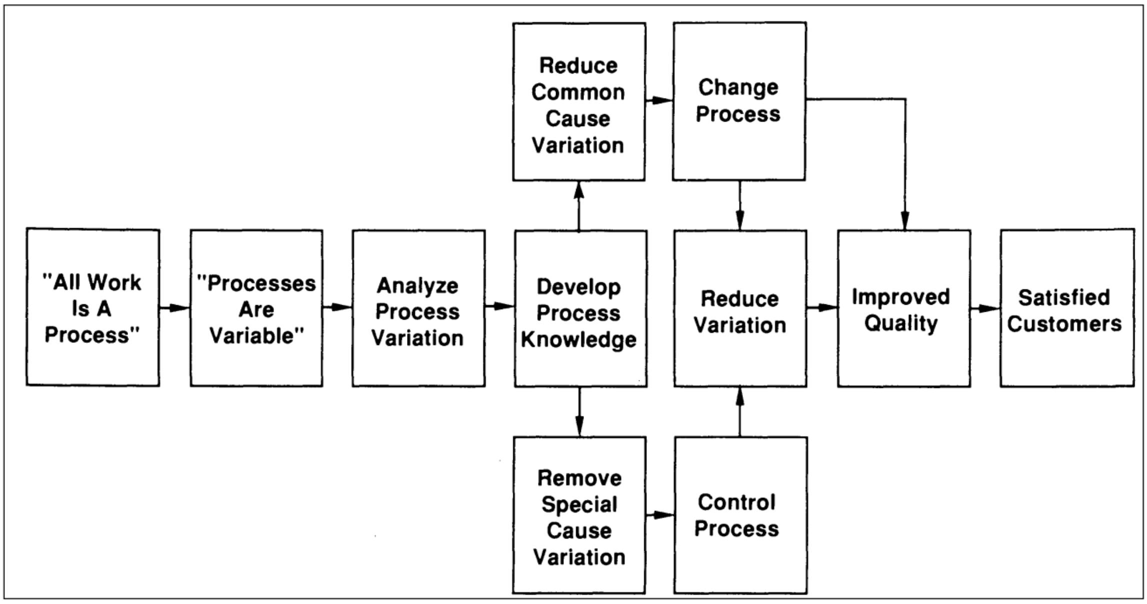


Figure 2-2: Improving Quality through Variation Reduction (Snee, 1990)

2.4.6 Flexibility

One of the distinct features of DOE is its ability to be integrated with other process improvement techniques such as simulation (Law and Kelton, 2000) and mathematical programming (Taha, 2002). The former can be defined as the imitation of the operation of a real-world process or system over time (Banks, 1999) while the latter may be defined as a mathematical representation aimed at programming or planning the best possible allocation of scarce resources (Bradley et al, 1977). Broadly speaking experiments can be conducted on the actual process under study or on a model (Figure 2-3). The latter approach is adopted whenever the experimentation time or the cost using the actual process is prohibitive. A model can, broadly speaking, be classified into two categories: simulation and analytical mathematical (see Figure 2-3). DOE is becoming an integral part of simulation modelling, in fact many simulation books such as those by Banks (1998) and Law and Kelton (2000) devote a separate Chapter to explaining the technique; in addition an entire text (Kleijnen, 2008) and many research papers (Sacks et al, 1989; Welch et al 1990; Kleijnen and Standridge, 1988; Kleijnen 1995, 2005) are dedicated to explaining the use of DOE in

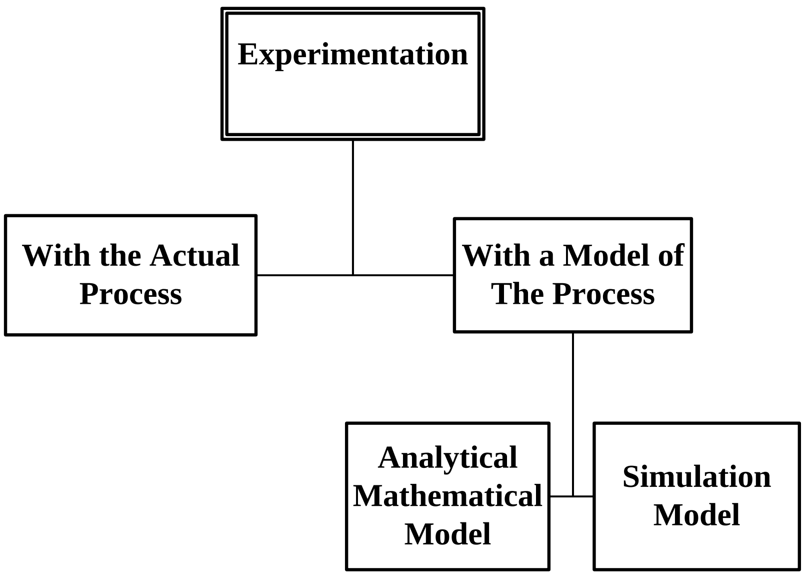


Figure 2-3: Ways of Experimenting

simulation modelling. This is natural as simulation is essentially an experimentation tool in which a computer model of a new or existing process is created for the purpose of conducting what-if analyses.

Regarding mathematical programming which represents an example of analytical mathematical models, DOE can provide a valuable input to its modelling stage which involves developing a mathematical model representing the objective of the study in terms of a measurable quantity such as profit, cost or revenue i.e. the objective function. As a powerful technique for establishing empirical mathematical models linking the process's performance measures with its main parameters, DOE can be very effective in developing the objective functions in studies that utilise mathematical programming as an improvement technique. At the same time, mathematical programming methods are of vital importance to optimisation DOE studies, the objective being to utilise the derived empirical model to identify the "best" settings of the process's parameters.

2.4.7 Improving the Process of Formal Decision Making

According to Goh (1992, 2000) the factors behind the achievement of superior quality can be put into three building blocks of an overall quality framework; namely, a quality management system, quality technology and quality information. DOE is an integral part of the quality information gathering and learning process which should be utilised to inform decisions and actions. According to Sage and Armstrong (2000) a formal approach to decision-making should involve: i) formulating alternative courses of action; ii) analysing their impacts; and iii) interpreting and selecting the appropriate options for implementation. As its studies incorporate all of these, DOE possesses the capability of being an effective formal decision-making technique. For example, screening identifies the most influential parameters that affect process performance. The identification of all or a subset of the combinations of the influential factors' levels parallels the formulation of alternative courses of action as each represents a candidate decision to be taken.

Regarding the analysis of the action impacts, DOE firstly enables the experimenter to obtain (in a physical experiment) or predict (in a simulation experiment) the outcome of each of the candidate courses of action. Furthermore, it encompasses a powerful set of statistical tools such as Analysis of Variance (ANOVA) that can be used to analyse the experimental results and test the statistical significance of the studied factors. It also provides quantitative techniques that can be exploited to isolate and estimate the effect of each of the factors and interactions under study.

With regard to the interpretation stage, DOE utilises very powerful, yet simple, graphical tools such as main effects and contour plots (Antony, 2003; Cornell, 1990) which are very helpful in interpreting the obtained results. Furthermore, mathematical procedures such as steepest ascent (Myers et al, 2009) and linear and non-linear programming methods (Tang and Xu, 2002; Kim and Lin, 2006) which are response surface DOE tools are very powerful in the process of selecting the "best" course of action to choose.

There have been many attempts to classify the types of decisions made in an organisation. Of particular interest are the following four types postulated in Anthony (1965) and Anthony et al (1992):

- *Strategic planning decisions*: those made to choose highest level policies and objectives, and associated resource allocations;
- *Management control decisions*: those made to ensure effectiveness in the acquisition and use of resources;

- *Operational control decisions*: those made to ensure effectiveness in the performance of operations;
- *Operational performance decisions*: day-to-day decisions made while performing operations.

Whenever the cost impact of the decision exceeds that of performing the experiment, DOE is very effective in dealing with the management control and operational performance and control decisions. Generally it is very powerful in applications that involve well defined and repetitive processes that encompass interdependent and variable activities and events. Although intrinsic in Manufacturing processes, these characteristics are not uncommon in all business processes. In fact, Harrington, (1991) observed that 80 percent of all business processes were repetitive. Despite this, the use of DOE is not common in this sector. In a recent article, this was criticised by Ariely (2010) who argued that "companies continue to pay overly confident consultants big money to supply answers rather than gather evidence in experiments to help them make their own decisions". He went on to emphasise that "only an experiment gives you the evidence you need" and that "intuition is a remarkably bad thing to rely on". It is the author's opinion that a combination of experiments and intuition can prove to be very effective in dealing with business decisions. Although DOE may not be as effective a tool in making strategic decisions as it is in operational ones, it is important to remember that the quality, effectiveness, precision, and strategic forethought of decisions made throughout an organisation (whether they be management control or operational), have vital consequences for the overall success of its strategy. Moreover, the flexibility of DOE enables its integration with many helpful tools such as Decision Support Systems (DSS) (Turban and Aronson, 1997) to provide a valuable input to strategic decision-makers. In fact, DOE is of essential importance to the development (Alberti et al, 2009) and the empirical assessments

(King and Rodriguez, 1981; Sharda et al,1988; Daily et al, 1996) of the DSS. It can also be very useful in their application.

A further possible use of DOE in strategic decision-making relates to the use of strategic interactive planning proposed by Ackoff (1979, 1981) who suggested that in the case of failing to generate a consensus among the participants in the planning process, experiments should be designed to resolve the issue causing the disagreement. He further suggested that the participants should be involved in designing these experiments as their agreement on this should facilitate the consensus achievement on the experiments' results and subsequently on the conflict. The main reason for such disputes is the absence of data without which everyone is an expert and as Snee (1986) observed "discussions produce more heat than light". By generating the required data to understand the scope and the root causes of the disagreement, DOE can be an effective technique for generating consensus among the strategic planning participants.

2.4.8 Ability to Express the Improvement Effects in Monetary Terms

Of vital importance to attract the attention of any organisation's management to the value of a technique is the ability to show the impact of its use on the bottom line. Taguchi (1986) proposed the loss function to communicate the monetary gains of using DOE to reduce process variation. He established a quadratic relationship between the loss and the functional specifications. For example, if Y is a performance characteristic with target T and tolerance interval $T \pm \Delta$, then when $Y = T$, no loss is incurred. However, the loss continually increases as Y deviates from the target value and when it exceeds $T - \Delta$ or $T + \Delta$, the quality loss is equal to the cost C of the product disposal or manufacture as indicated by the quadratic loss function in Figure 2-4. Taguchi modelled this loss function, L (Y) (Phadke, 1995), as

$$L(Y) = K [(Y-T)^2] \quad (2.2)$$

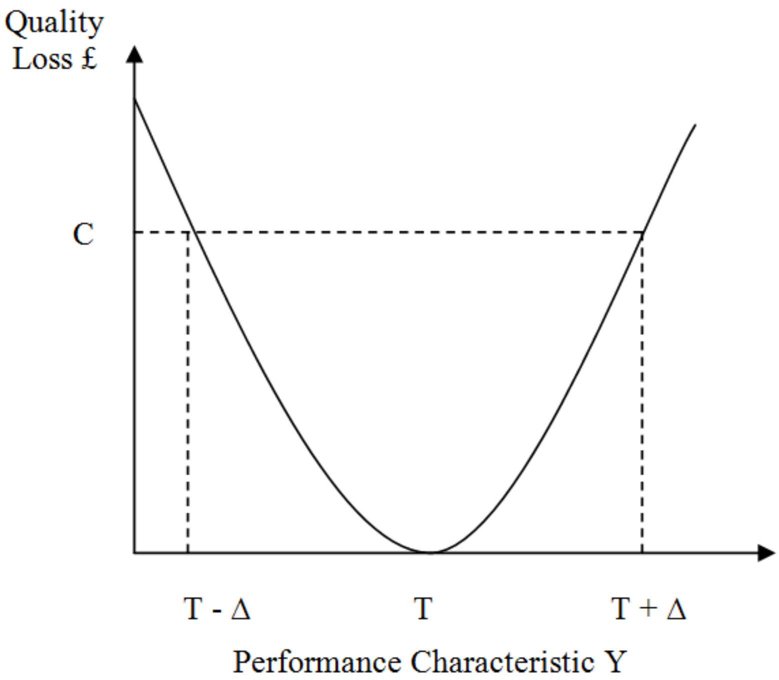


Figure 2-4: Taguchi's Loss Function

where the value of K is a constant depending on the cost at the specification limits and the width. The above equation represents the loss for an individual part. The average loss per part for a group of parts is

$$L(Y) = K [S^2 + (\overline{Y} -T)^2] \quad (2.3)$$

where $\overline{Y}$ and S^2 are respectively the average and variance of the performance measures of the group. Clearly the above functions will penalise even small deviations from the target which is a departure from traditional thinking, where costs are only incurred when Y is outside specification. Depending on the objective of the experiment, Taguchi derived several loss functions as detailed in Ross (1996). It was through the use of these measures, according to Pignatiello and Ramberg (1991), that the DOE technique gained considerable attention and support among managers. It is important to note that the expected bottom line monetary outcomes of any experimental design initiative are dependent upon its objective. One should not expect the outcomes of a screening experiment where the objective is to identify the key parameters to be the same as an optimisation one that aims to identify the best parameter settings (Goh, 2001).

2.4.9 Tool for Scientific Investigation

Crombie (1953) observed that the idea that scientific progress is attained by an iterative process comprising both induction and deduction has been known since the time of Robert Grosseteste, one of the founders of Oxford University, who attributed it to Aristotle (384-322 BC). Generally the investigator starts with some data or facts, from which a possible theory, hypotheses conjecture, or idea is inferred. This can then be represented by a tentative model which leads, using deduction, to consider what should happen if that model was true and what data ought to be collected to compare that with what actually occurred. This can lay the ground for appropriately modifying the model and so on (Box and Youle, 1955). This iterative process is represented diagrammatically in Figure 2-5 and was shown by Box and Liu (1999) and Box (1999) to be equivalent to the plan- do- check- act process (Shewhart

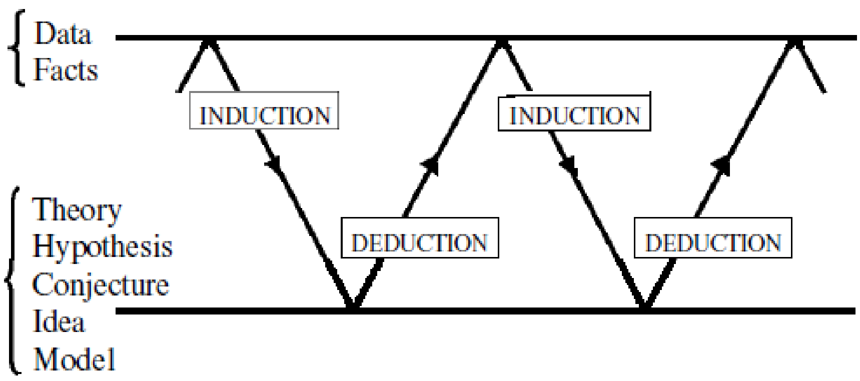


Figure 2-5: Scientific Investigation Process (Box, 2001)

1980; Deming 1986). Box (1994, 2001) ascertained that sequential DOE was vital in this case. He quoted R. A. Fisher: "The best time to design an experiment is after you have done it" (Box, 1992) indicating that experimentation should (whenever feasible) be sequential, with subsequent experiments being designed using the knowledge obtained from prior ones. The reason is that as the initial experimental runs are executed, insight into factor effects is accumulated and used to select the next set of runs. Daniel (1976) suggested using 50–67 % of the available resources on the first experiment, whereas Box et al (1978) recommended at most 25 %.

It is clear from the above discussion that experimental design is, as Goh (2002) emphasised, too important to be left to Statisticians; thus Engineers, particularly in Manufacturing, around which this work is centred, should have a good understanding of the technique and its associated benefits. Publications by Antony (1998, 2002, 2006) and Antony and Antony (2001) are devoted to explaining DOE to Engineers and Managers. Moreover, there are many case studies presenting its successful applications to improve product or process performance within the Manufacturing sector. These include Antony (1999, 2001), Antony et al (1999, 2001, 2004, 2006), Rowlands et al (2000), Antony and Kaye (2000), Antony and Roy (1999). There have also been many examples of the use of DOE to improve non-Manufacturing processes including Donahue et al (1996) in Marketing; Pean et al (1998) and Gratteri et al (1996) in Pharmacy; and Smith et al (1996) and Stolle et al (2002) in Law. Moreover, Condra (1995) explained how Managers should use experimental design as a method for adding value to their organisations' products.

The basic principles of DOE together with its implementation stages are presented in Appendix 1 and a detailed discussion of the statistical analysis associated with the technique is provided in Appendix 2.

2.5 Types of Statistical Design of Experiments

The appropriate DOE to choose depends on such things as the objective of the experiment, the number of factors involved, the number of levels of these factors and the restrictions that need to be considered while experimenting. There are several ways to classify the DOE types. An excerpt of the common criteria that are adopted for this purpose is shown in Figure 2-6. Each will now be discussed apart from Experiment Objectives which has already been described in Section 2.4.

2.5.1 Number of Studied Factors

On the basis of the number of factors under study, the design type can be classified into single factor and factorial. The former involves studying one factor only at two or more levels. The objective is usually to compare these levels and find the best with respect to a certain performance measure. In factorial experiments, two or more factors are varied simultaneously. It is a more efficient alternative to the one-factor-at-a-time experiments as it requires fewer runs and enables the interaction effects to be estimated. Moreover, factorial experiments allow the effects of a factor to be estimated at several levels of the other factors, yielding conclusions that are valid over a range of experimental conditions. Factorial experiments are further categorised into two-level, three-level or four-level experiments according to the number of levels of the studied factors. By the same token, when factors with different numbers of levels are used, the experiment is called a mixed level trial.

2.5.2 Execution Restrictions

Designs can be classified into completely randomised, blocked, split-plot or nested ones depending on the presence (or absence) and the nature of certain restrictions on the way in which the experiment is performed. Some of these relate to how noise factors are dealt with

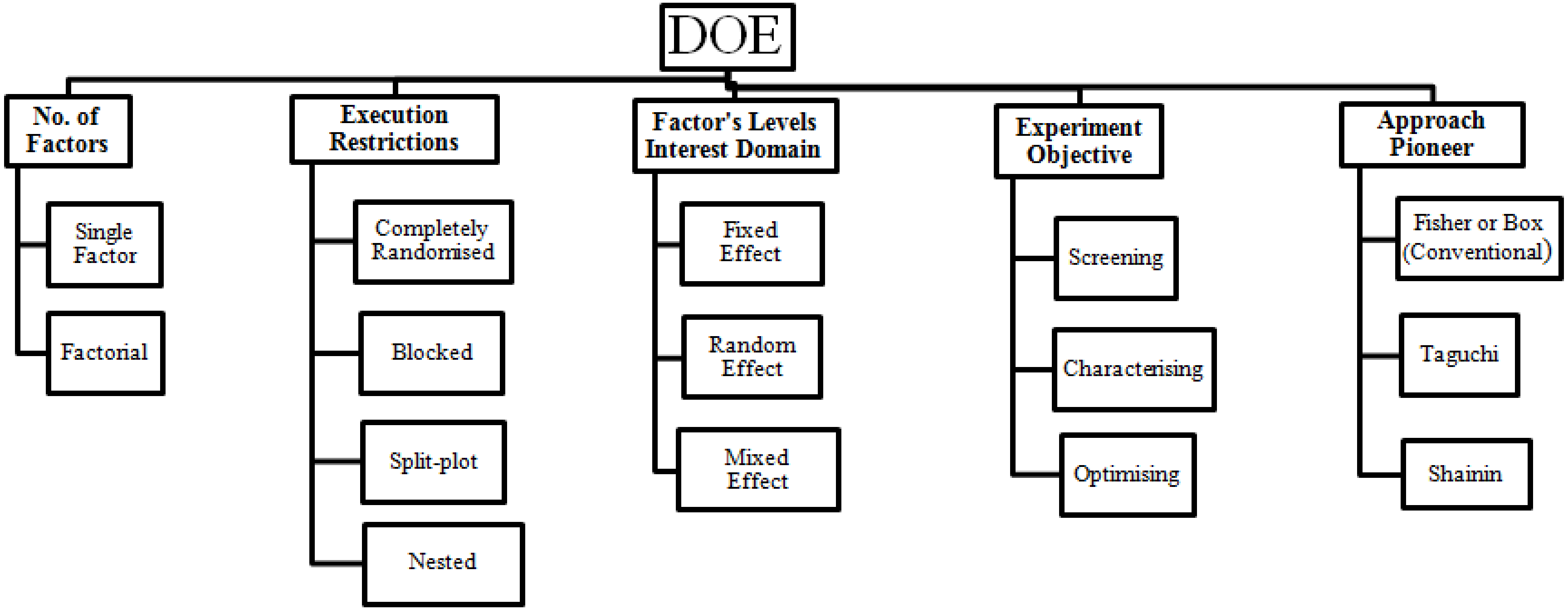


Figure 2-6: Classification of DOE Types

For example, in cases where the noise factors are unknown and uncontrolled, completely randomising the run is the best way as their effects are “averaged out”. When noise factors are known but uncontrollable, their values should at least be observed at each run and accounted for in the analysis using such techniques as Analysis of Covariance (Montgomery, 2010). When noise factors are known and controllable, blocking is a systematic method for eliminating their effects, randomisation being performed within each block. However, there are instances where this is not at all feasible. For example, the factors of interest may be hard-to-vary while the others are easy. The order in which the experimental runs are performed is determined by the ordering of these hard-to-vary factors. So for the same settings, the runs may be performed sequentially (varying only the easy-to-vary factors) without resetting them between the runs. In such experiments, each setting of the hard-to-vary factors is called a plot (a word inherited from agricultural applications). Thinking of the settings of the easy-to-vary factors as a splitting of the plot into a number of parts, one for each setting, leads to a split-plot type of experiment. A nested type of experiment is used when the levels of one or more factors are similar but not identical for different levels of another factor. This differs from the factorial design in the property that the levels of each factor are unique to that factor. An example might be a company that purchases three batches of raw material from three different suppliers and needs to compare their quality. If the batches from each supplier are unique to that particular supplier then the experiment should be performed as a nested trial.

2.5.3 Factors’ Levels Interest Domain

Depending on the interest domain of the levels of the factors under study, DOE may be classified into fixed, random and mixed effect experiments. If interest lies in studying particular levels of each factor, then the experiment is called fixed effect. In this case, the domain of the inferences’ validity regarding each factor is limited to its selected levels i.e. the conclusions concerning each factor cannot be generalised to its possible levels that were not explicitly considered in the experiment. In random effect experiments, the experimenter is interested in examining all the possible levels of each of the studied factors. As the number of possible levels may be very large, it is necessary to choose a sample of levels from the population of all possible levels. In such a case the conclusions regarding the significance of each factor can be generalised to the entire population of its levels, not just those that were considered in the experiment. In some experimental situations, the objective may entail studying some fixed effect factors along with other random effect ones. Such experiments are called mixed effect.

2.5.4 Approach Pioneer

The DOE approach may be identified by the name of its pioneer, examples include Fisher (or Box), Taguchi and Shainin. The first, referred to as "conventional" or "classical" designs, were named after Sir R. A. Fisher and one of his eminent scholars George Box. One of the main characteristics of this approach is that it is centred around studying the average response and its variation around its target along with the causes of this variation i.e. the control factors. The variation of the individual response values around their average is treated as constant and its causes- the noise factors- are handled by either randomisation or blocking. A sequential and adaptive approach is strongly advocated when performing such experiments. In fact the approach encourages starting with a two-level screening experiment to identify the key factors and following this by the use of optimising experiments with factors at more than two levels for studying their non-linear effects and identifying their "best" settings.

The Taguchi approach (see Section 2.4.5) on the other hand, focuses on studying both the variation of the average response around its target and that of the individual responses around their average. It emphasises the importance of including the causes of both types of variation i.e. the control and noise factors in the experiment. In so doing, Taguchi advocates a one-shot large experiment approach where a crossed (product) array design comprising an inner array containing the control factors (dominated by a three-level arrays), and an outer one containing the noise factors is utilised. Taguchi derives a novel variability performance statistic called the signal-to-noise ratio (S/N) that represents the ratio of the average (signal) to the variance (noise) and is directly related to and derived from Taguchi's loss function. Although his philosophy has been endorsed by both researchers and practitioners, his experimentation strategy and data analysis method have generated much controversy among Statisticians and Researchers (Box, 1988; Box et al 1988; Hunter 1985; Montgomery, 1990; Nair et al, 1992).

Shainin DOE was initially introduced in 1957. (Shainin, 1957). Its main objective is to identify the most influential parameters on a process performance. Shainin and Shainin (1988) proposed a five-step methodology for implementing this approach. The first four may be considered as a systematic search for the key causes of a performance measure variation utilising such statistical techniques as confidence intervals and SPC. The fifth step involves conducting a conventional factorial to determine the "best" settings of the key factors. The main advantage of this approach is its simplicity. It is primarily a problem solving approach i.e. in most cases it aims to bring a deviant process performance to its acceptable conditions rather than seeking any breakthrough improvement. Despite this, Ledolter and Swersey (1997) found no good reason to use such an approach as an alternative to the conventional one. Further details on Shainin's design and its performance compared with Taguchi's can be found in Thomas and Antony (2005).

The types of DOE that have been discussed were developed over four eras according to Montgomery (2010). These are presented in the following section.

2.6 A Brief History of Statistical Experimental Design

The first era of DOE is the agricultural which was led by the pioneering work of Fisher in the 1920s and early 1930s, who according to Cox and Reid (2000), was the first to provide a systematic discussion of DOE. While being responsible for statistics and data analysis at the Rothamsted Agricultural Experimental Station (England), Fisher found that carrying out experiments improperly hampered the analysis of the data. Due to the nature of agriculture, its experiments tended to be large, took a long time to complete and were subject to various sources of variation in the field. These characteristics led Fisher, through interacting with Scientists and Researchers in many fields, to introduce the concepts of randomisation, replication, and blocking alongside the principals of factorial designs and ANOVA.

The second (the industrial) era initially started with attempting to apply Fisher's DOE techniques to solve problems in Chemical Engineering. It was pioneered by G.E.P Box and his co-workers at Imperial Chemical Industries who discovered that new techniques had to be developed to deal with the two unique characteristics of industrial experiments: immediacy and sequentiality. Immediacy relates to the fact that the performance measure can usually be observed (nearly) immediately whereas sequentiality relates to the ability to quickly learn crucial information from a small group of runs that can be used to plan the next experiment. In contrast to agricultural experiments where the principal objective is to compare treatments, process modelling and optimisation were the main objectives in industrial experiments. This led to new techniques for DOE, notably response surface and optimal designs, which became widely recognised in the chemical and process industries, mostly in research and development work. However, the application of statistical design at the plant or Manufacturing process level was still not widespread.

As a result of the increasing interest of Western industry in quality improvement that began in the late 1970s, the third era of statistical design began. This was led by the work of Genichi Taguchi who suggested highly fractioned factorial designs and other Orthogonal Arrays (OAs) along with some novel statistical methods including robust parameter and tolerance designs. By this time, designed experiments became more widely used in many industries, including automotive and aerospace Manufacturing, electronics and semiconductors.

One of the positive outcomes of the Taguchi approach was the start of the fourth era of statistical experimental design. The debate regarding the effectiveness of his design and analysis methods resulted in renewing the interest in DOE and led to the development of many new and useful approaches to put Taguchi's philosophy regarding variance reduction and robustness into action in an efficient and effective manner.

Goh (2002) added a fifth era which he associated with the emergence of the Six Sigma quality initiative. As Taguchi simplified DOE and presented it in a way that Managers could understand, Goh attributed to Mikel Harry (the pioneer of the fifth era) the “packaging” of DOE to gain the support of bottom-line oriented CEOs. DOE is considered to be an important technique in Six Sigma due to its power to simultaneously investigate the potential causes of variation. It is prominently utilised in the improve phase of its five projects’ implementation phases: Define-Measure-Analyse-Improve-Control (DMAIC). Moreover, it is a vital technique for the design and optimise phases of the Identify-Design-Optimise-Verify (IDOV) implementation framework of Design For Six Sigma (DFSS) (Goh, 2009). Examples of the use of DOE in the context of Six Sigma can be found in Chan and Spedding (2001) and Conklin (2004). Goh (2002) argues that Harry has taken DOE a step further than Taguchi in several ways. These include extending its applications to the transactional processes alongside physical ones; demonstrating the impact of its outcome on the business bottom line thereby attracting the attention of CEOs and top management; and integrating its training with the "Belts" certification of competence system which comprises a hierarchy of designations such as "Master Black Belts", "Black Belts" and "Green Belts" (Goh, 2002).

2.7 Summary

There are two main approaches to making decisions regarding process improvement, experiential and data-driven. The former is based on subjective feeling, opinion and past experience. It utilises a trial and error approach and can be wasteful in terms of time and money as there is no guarantee that effective decisions will be reached. The data-driven approach, on the other hand, makes use of actual data and objective mathematical principles to produce sound decisions. DOE is one of the powerful techniques of this approach. It has been proven to be among the most effective and reliable techniques employed by twenty-first century globally competitive organisations (Antony et al, 2004). It is very useful for identifying key process's parameters and determining their settings so that the "best" performance is attained. In terms of the cause-and-effect relationship between the key process's parameters and the performance measure, DOE is not only powerful for understanding its nature but also for representing it by means of an empirical model. Alongside the statistical thinking philosophy, DOE comprises an arsenal of field tested and proven approaches for dealing with complexity and its two main constituting elements: interdependence and variability. Due to the appropriateness of its sequential approach for the induction-deduction cycle, DOE is a pivotal technique for scientific research. Although valuable in its own right, DOE can, utilising its flexibility, be effectively integrated with many powerful quantitative techniques such as simulation and mathematical programming.

In general, the attention of top management can be drawn to the importance of DOE through expressing its application's outcomes in monetary and bottom line terms using such techniques as Taguchi's loss functions.

Several experimental designs are available, their classification depending on criteria such as the number of studied factors, execution restrictions, factor levels domain of interest, objectives and approach pioneer. The development of the various types has spanned over five eras; namely agricultural, industrial, Taguchi's variance reduction, alternative robust DOE and Six Sigma.

Considering the importance of DOE and the concepts underpinning its use, it is interesting to examine the way in which it is being employed in practice. This is the topic of the next Chapter.

CHAPTER 3: A REVIEW OF THE PRACTICE OF DOE IN MANUFACTURING

3.1 Introduction

Generally, much of the DOE literature is centred around the question of how the methodology should be applied, while less attention has been devoted to how it is actually being practiced. Alongside other statistical quality improvement tools and techniques, many survey studies examine the extent to which DOE is used within various industries in different countries. For example, Araujo et al (1996) and Antony and Banuelas, (2002) studied the frequency with which DOE has been used compared to other quality improvement techniques in the UK. Similarly, Arvidsson et al (2003), Gremyr et al (2003) and Bergquist and Albing (2006) surveyed the use of several statistical methods including DOE in Swedish industry. In the Basque Country, Tanco et al (2008) examined the knowledge and use of DOE within Manufacturing companies. Ilzarbe et al (2008) provided a collection of published engineering case studies in which DOE was employed during the period 2001-2005. A common thread throughout these studies is their emphasis on how frequently DOE has been applied and the practitioners' perception concerning its usefulness rather than the way in which it has been employed particularly in Manufacturing Engineering.

The objectives of this Chapter are to identify gaps between the way in which DOE should be used (as presented in the literature) and how it is actually employed and to highlight aspects that warrant further clarification and assimilation in order to bridge these gaps. This will enable recommendations to be made for improving Manufacturing Engineering DOE practice. The methodology implemented in this review is firstly described; then its results are presented and discussed. Finally, the upshot of this Chapter is provided.

3.2 Review Methodology

A valuable indicator of any methodological practice is the published literature that has been peer reviewed. Assuming that the way in which DOE is being employed in publications is indicative of the skills and knowledge held by the authors, an examination of how DOE is being applied can be used to assess which of its aspects practitioners are both acquainted and not very well acquainted with. The first step in this review was to select relevant Manufacturing Engineering journals which had recent published papers on DOE applications. Manufacturing covers such a broad area; Groover (2007) for example described it as “the application of physical and chemical processes to alter the geometry, properties and/or the appearance of a given starting material to make parts or products”, while Sharma (2005) stated that it includes such aspects as behaviour and properties of materials and their processes; the design of product, equipment and tooling necessary for their manufacture; management of the Manufacturing enterprises; and the design and operations of Manufacturing systems. Therefore, it was also necessary to define the scope for the review. In so doing the extent to which experimental work was likely to be performed was a key criterion. This led to focusing on journals that deal with the science and technology of Manufacturing processes, materials (including, metal, ceramics, polymers and composites) and the machines and tools applied to their manufacture. Another criterion that was used in selecting the journals was their impact factor. The journals listed under the category of Manufacturing Engineering in the journal citation reports (Thompson Reuters, 2010) were therefore examined to find those that were in line with the scope of the review and had attained a reasonable impact factor. The Journal of Materials Processing Technology (JMPT), the International Journal of Machine Tools and Manufacture (IJMTM) and the CIRP Annals - Manufacturing Technology (CIRP-MT) fitted the bill and so these were selected.

All of the articles published in these journals during the year 2009 were accessed through the ScienceDirect® database. The process of selecting the relevant papers was performed in two stages. Firstly, the abstract of each article was read to identify those that involved experimentation, thus review articles and non-experimental studies were omitted. The methodology section of each selected articles was then read to exclude those in which DOE was not employed. The aspects that were considered for the review were:

1. Design Related and Conducting Aspects

- The number of studied factors and the number of their levels;
- The type of design used, for example orthogonal array and response surface designs;
- Whether a full or fractional factorial design was used;
- Whether conventional or Taguchi DOE was used - the latter was recognised by the use of Taguchi's designs or by the data analysis techniques employed;
- The objective of the experiment i.e. was it conducted to study the response average, variation or both;
- Whether the experiment was replicated or not;
- The examination of aliasing patterns in fractional factorial designs;
- Whether any factor interactions were studied;
- Whether the experimental runs were randomised.

2. Data Analysis Aspects

- The use of pooling methods in analysing unreplicated experiments;
- The use of ANOVA;
- In the case of hypothesis testing, whether the tested hypotheses were stated and the p-values reported;

- Whether regression modelling was employed;
- When used, the kind of effect size or measure employed to identify the relative importance of the studied effects.

3. Results Interpretation and other Aspects

- Whether the p-values were interpreted and the meaning of statistical significance explained;
- The distinction between fixed-effect and random-effect inferences;
- The performance of confirmation trials;
- The use of DOE references including books and articles that explain the DOE methodology;
- When used and stated, the statistical package employed in performing the DOE data analyses.

3.3 Results and Discussion

Having examined all the concerned DOE aspects in the reviewed articles, the collected data were analysed using SPSS v17 (SPSS, 2007). Details of the investigated articles are given in Appendix 3. As shown in Figure 3-1, a total of 765 articles were published in the JMPT, 148 in the IJMTM and 145 in the CIRP-MT during 2009. Of these, the numbers of articles that involved experimental work were, respectively 550 (71.9%), 119 (80.4%) and 74 (51%) with DOE being used in 41 (7.5%), 13 (10.9%) and 4 (5.4%) respectively. Since some of these incorporated more than one experimental phase, a total of 62 applications were noted.

3.3.1 Design Related and Conducting Aspects

As shown in Figure 3-2, the number of factors studied in the experiments ranged between 1 and 10 with the most frequent being 4. In fact, in 85.5% of the cases, six factors or fewer

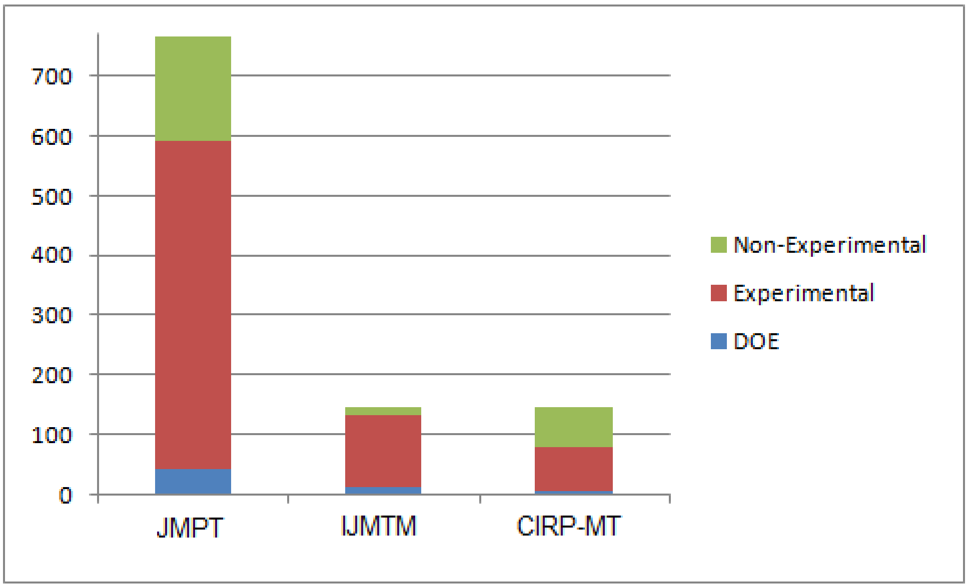


Figure 3-1: Article Types in the Reviewed Journals

were examined. In all but three of the applications, these were control factors; noise factors were used in the others (their numbers ranged from 1 to 3). Generally the number of factors used in Manufacturing experiments is relatively low, which could explain why in some of the studies, a sequential approach was adopted when the numbers of studied factors were 10 (Lauderbaugh, 2009), 8 (Tsai et al, 2009) and even 6 (Ali et al, 2009). Normally initial

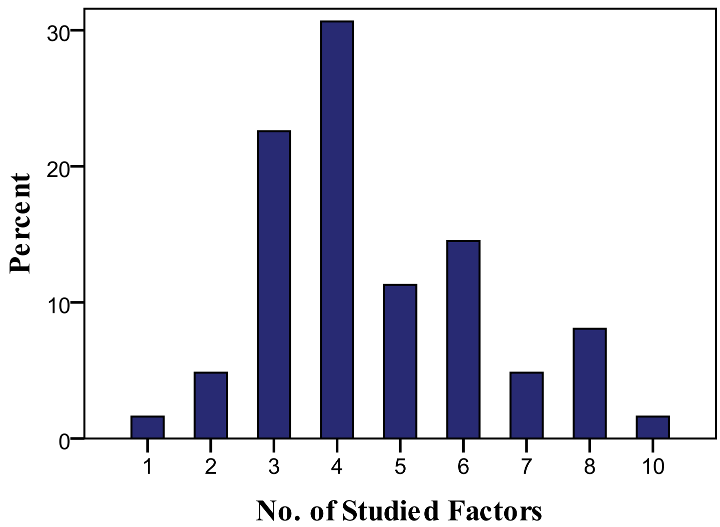


Figure 3-2: Number of Studied Factors in the Surveyed Experiments

screening experiments are conducted when the number of studied factors is regarded as large. This observation is in agreement with that of Ilzarbe et al (2008) who noted that, generally, Engineers utilise their prior knowledge and unpublished initial experiments to minimise the number of factors in their published ones. The reason for omitting initial experiments is their likely poor results (Ilzarbe et al, 2008).

The majority of the applications were three-level experiments (43.5%), the percentages of two- and mixed-level trials were respectively 25.8% and 21%. The most frequently used designs were the L_{18} (22.6%), L_9 (19.4%), L_{16} (11.3%), L_8 (9.7%), L_{27} (9.7%) and L_{12} (6.5%) orthogonal arrays. The Central Composite Design (CCD) was used in 4.8% of the experiments. Approximately, three quarters of the applications were based on fractional designs. Very large full factorial designs were associated with simulation experiments as can be seen in Venkatachalam et al (2009) and Jeang and Li (2009) where an L_{243} (3^5) design was employed. Smaller full factorial designs were also observed in such applications as training a neural network model as was the case in Karayel (2009) who conducted a 4^3(=64) experiment. Another application included a sequential investigation where the screening experiments detected four three-level factors as the most influential and subsequently a 3^4 (=81) experiment was performed (Tsai et al 2009).

While the conventional objective of a DOE study is to improve the average response, Taguchi's philosophy emphasises the reduction of its variation. It was found that the conventional philosophy was more frequently adopted; in fact, the scope of 69.4% of the experiments was limited to studying the average response. A joint study of the response average and variation was the objective of 22.6% of the experiments while an examination of the response variation alone was observed in the remainder. Despite the literature on its pitfalls and inefficiencies (Box et al, 1988; Nair et al, 1992; Montgomery, 1999) Taguchi's orthogonal arrays and data analysis methods were more common than the conventional ones. In addition, Taguchi's signal-to-noise (S/N) ratios were the only performance measures used in all the experiments where the response variation was examined despite their criticism (León et al, 1987; Box, 1988). It is unclear whether the rationale behind their use and what exactly they measure is understood since in one of the studies, the factors that significantly affected the S/N ratio were automatically taken to have a substantial effect on the average response (Tsai et al, 2009). In fact, a S/N ratio was used to screen the important factors which were then examined in a full factorial design to model their effect on the average - the implication being that the S/N ratio and the average response are interchangeable.

The popularity of Taguchi's techniques is not surprising as their simplicity appears to have made them more attractive to practitioners than those of the conventional approach. However, it is interesting to note that they have been used predominantly for merely studying the average response. In their study, Phatak et al (2009) adopted an amalgamated approach in which their experiment was conducted using a Taguchi orthogonal array but the results were analysed using conventional regression analysis. In the study carried out by Yang et al (2009), a conventional 2^3 design was used but it was reported as a Taguchi orthogonal array.

Just over half of the experiments were unreplicated. This may be attributed to time, cost and technological constraints. Of the replicated experiments, the majority were performed to study the response variation. Partial replication, in the form of replicating the runs associated with the central point settings was observed in five of the studies.

Of those experiments based on fractional trials, aliasing patterns were examined in only two of them. Taguchi's linear graphs and the concept of resolution (Chapter 4) were the only techniques employed in this respect. Despite the criticism of the former (Tsui, 1988) and the extensions of the latter (Fries and Hunter, 1980; Chen et al, 1993), none of the contemporary aliasing examination techniques and criteria were applied. As will be detailed in the following Chapter, aliasing can affect both the estimated effects and the experimental error variance. Therefore, failing to examine it can distort the conclusions drawn regarding which factors and interactions are significant. One possible reason for omitting aliasing is that the postulated methods to generate and evaluate its patterns in the DOE literature are not easy to assimilate. A further reason relates to the practitioners possible lack of awareness of the impact of overlooking aliasing consequences. These issues are dealt with in Chapter 4.

A related problem concerns the examination of factor interactions. 69.4% of the studies dealt with the investigation of main effects only. Interactions were ignored even where there were sufficient degrees of freedom to estimate some or all of them. This can, in no small part, be ascribed to the adoption of Taguchi's view regarding interactions. In fact, Taguchi and Wu (1985) stated that "no interactions are calculated even if they exist". They went on to say that "these interactions are treated as errors, so it is advantageous to have the effects of these interactions uniformly distributed in all (design matrix) columns". Thus Taguchi believed that the main effect estimates would not be affected by the presence of factor interactions, since their effects are evenly distributed across all the array columns. This is confirmed in his suggestion that non-regular orthogonal arrays (where main effects are partially aliased with their interactions) such as the L_{18} and L_{12} should be used so that the interaction effects can cancel each other out (Taguchi et al, 2004). This view has been shown to be fallacious (Hamada and Wu, 1992; Box and Meyer, 1993). Yet, it was explicitly adopted in the studies undertaken by Tsai et al (2009) and Marafona and Araujo (2009). The presence of interactions can distort the main effect estimates in terms of both their magnitude and direction thereby impairing the drawing of valid inferences. It must be borne in mind that a major reason for recommending the factorial experimental approach as an alternative to the one-factor-at-a-time is its ability to study factor interactions. Ignoring this feature may indicate a lack of understanding of one of the chief driving forces behind the proposal of factorial designs; namely the potential interdependency between factors in exerting their effect on the studied response.

Another important aspect relates to randomising the experimental runs. Both the allocation of experimental material and the order in which the experiments are performed should be randomly determined. Randomisation provides protection against any unknown factors that may impact the response under study. It reduces the potential bias that could result from the improvement of the experimenter's skill as the runs are performed. The DOE statistical analysis is based on the assumption that the experimental results are independent random variables. Although randomisation validates this assumption it was reported in only 12.9% of the experiments reviewed. It is not clear whether randomisation was performed, but not reported, in the others – this could be because of a lack of appreciation of the importance of this aspect. Of course, constraints such as investigating hard- or expensive-to -vary factors may render complete randomisation infeasible. In such cases, techniques such as blocking may be employed so that the randomisation restrictions can be accounted for in the analysis stage (further details on randomisation can be found in Cox (2006) and Hinkelmann and Kempthorne (2008)).

3.3.2 Data Analysis Aspects

The data analysis aspects comprise testing the statistical significance of the studied factors and interactions and estimating their effect sizes. In the experiments that were reviewed, over half incorporated one or more formal statistical significance testing procedures while the remainder were analysed graphically. As most of the experiments were unreplicated, there were no degrees of freedom available to estimate the experimental error. Thus, in the studies where the effects' significance were formally tested (53.13%), pooling techniques were used to identify those that should be combined to estimate the error variance. These included: Normal Probability Plot (NPP) (Daniel, 1959), Stepwise Regression (SWR) Pareto Plot (PP), the Unassigned (empty) Columns (UC) pooling and the Rules of Thumb (ROT) pooling. The last one refers to the process of selecting what are determined or even predicted to be reasonably small effects and pooling them (Logothetis and Wynn, 1989; Peace, 1993). Figure 3-3 shows the percentage of cases in which each of the pooling methods was used. Clearly, subjective and informal methods such as the UC and ROT were more popular than the

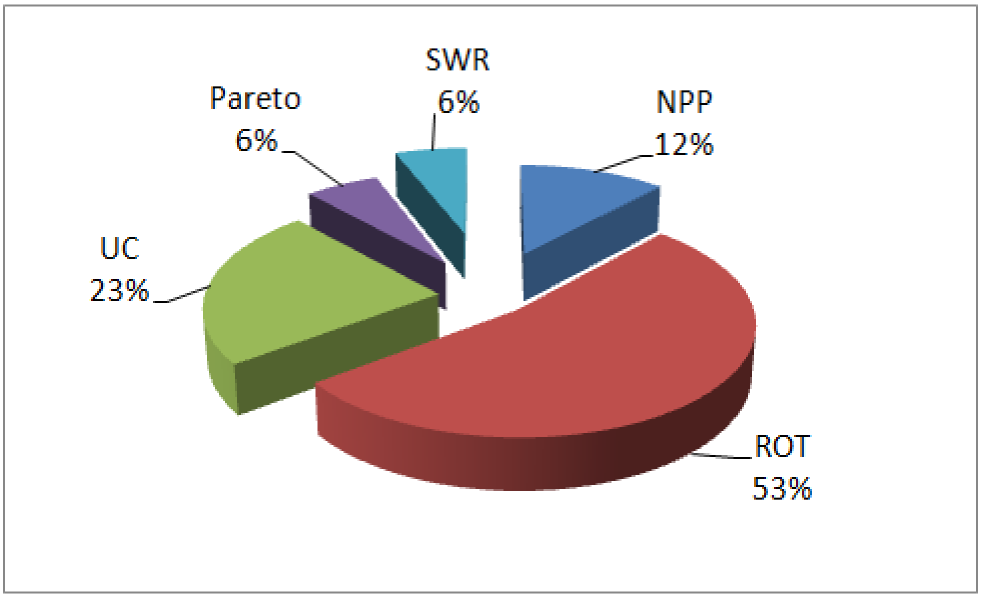


Figure 3-3: Percentage of Pooling Methods Usage in the Studies that Employ them

formal and even the simple graphical ones. This could be attributed to a lack of awareness regarding the usefulness of such a simple method as the NPP, a lack of familiarity with how formal methods can be employed or perhaps a combination of both. With regard to the publications in which NPP and PP were used, the statistical package Minitab was employed to construct both of them so that the effects that should be pooled and those that were deemed significant were automatically selected. In fact, this package uses Lenth's method (Lenth,

1989) to analyse unreplicated experiments and to identify the effects that should be pooled to test the significance of the remaining ones. Therefore, both the NPP and PP were merely used as graphical representations of the results of applying Lenth's method. Yet, the latter was not mentioned in any of the studies that reported their use despite the fact that no control was exercised as regards the selection of the effects that should be pooled or those that were regarded as noticeably large.

In analysing the fully and partially replicated experiments, pooling methods such as NPP, PP and ROT were used in 60% of them. Two different ways of use were distinguished. The first was associated with experiments where the signal-to-noise ratio was calculated for each trial rendering an unreplicated response variable. In this case, the pooling methods were used as already described in the context of unreplicated experiments. The second way was observed in cases where replicated experiments were used merely to study the average response and also in some of the partially replicated experiments. Pooling methods were used along with replication to estimate the error variance. Conventionally, when small effects are pooled in a replicated experiment, their overall significance should be tested utilising the error variance estimated from replication i.e. using a lack of fit test. However, in the experiments examined, the way in which the significance testing was performed varied according to the pooling method used. When ROT pooling was performed, no lack of fit test was conducted, thus there was no objective method to examine the validity of the resultant estimate of the error. Fortunately, when NPP and PP were applied, the replications were appropriately exploited to obtain an unbiased estimate of the error variance. This is due to the use of the Minitab package which employs a t-test (Appendix 2) in such cases and uses the NPP and PP as graphical representations of the test outcome.

The use of the UC pooling method is another reflection of the impact of Taguchi's strategy on the practice of DOE in which the convention is to deemphasise or ignore the effects allocated to the unassigned columns. When using an orthogonal array to conduct a full or fractional factorial experiment each column, be it explicitly assigned an effect or not, accommodates a main or an interaction effect. If any of those that the unassigned columns accommodate is large, the use of the UC method will result in a biased estimate of the experimental error. Consequently, the statistical inferences arrived at using such an estimate are invalid. Rather than deciding to use the unassigned columns to estimate the error, it is better to estimate the effects that these columns accommodate and ensure that they are small enough to be pooled (Daniel, 1959).

The ANOVA technique was used in just over half of the reviewed experiments. Its standard version was employed in 28 studies while the remaining used Regression ANOVA (see Appendix 2). Regression Modelling was reported in 13 of the reviewed experiments. In 4 of these, the developed regression models only were reported with no ANOVA. Their quality was judged by comparing their predicted values with the experimental results.

Hypothesis testing was performed as part of implementing ANOVA, Lenth's method or the t-test in 40 of the examined experiemnts. Of these, none stated the tested hypothesis and the p-value was reported in only 19. An examination of the assumptions that underlie the applied inferential statistical techniques was reported in only 7 studies. It is possible that the assumptions' verification was performed but not reported, however, it is important to emphasise that adopting any statistical test without giving thought to its associated assumptions could produce invalid results. An awareness of the importance of validating the assumptions is a necessary but not sufficient condition for the appropriate use of statistical tests. It is also essential to know how to appositely examine these assumptions. For example, in the study of Tsai et al (2009), it was noted that the Normality assumption was tested by plotting the residuals against the order in which the experimental runs were performed. Although such a plot will detect any unusual patterns in the experimental results, it does not illustrate the statistical distribution of the residuals; simple graphical tools such as NPP and Histograms are more effective.

With regard to measuring the Factorial Effects (FE) importance, the employed measures can be classified as mean-related and variance-related. The former examines the average change in the response variable produced by changing the levels of a FE whereas the latter estimates the amount of variability in the response explained by each FE. The mean-related measures were presented in numerical and graphical forms, the latter involving main effect and interaction plots. The numerical mean-related measures were arrived at using either least squares estimation (in regression models) or the level average method – in this case for each of the FEs under study, the average response associated with each of its levels was computed; the relative impact of each FE can then be determined by computing the range of the relevant average response values.

Four variance-related importance measures were used depending on the analysis technique employed. Where ANOVA was used, two different measures of the variation explained by each FE were used under the same label i.e. the Percentage Contribution (PC). Denoting the sum of squares by SS, these were:

$$PC_1 = SS_{FE}/SS_{total} \qquad (3.1)$$

and

$$PC_2 = \frac{SS_{FE} - (df_{FE})MSE}{SS_{total}} \qquad (3.2)$$

where df_{FE} is the FE's degrees of freedom and the MSE is the mean square error of the experiment. In the context of regression analysis, the two conventional measures of the overall explained variation, R^2 and R^2_{adj}, were used.

The percentages of the reviewed applications using the mean- and variance-related importance measures are shown in Table 3-1. As some studies reported the use of more than one measure, the total percentage exceeds 100%. Clearly, the mean-related measures were the more common with their graphical representations more frequently used than the numerical ones. One major reason for this is that the main objective of the studies was to identify the "best" settings of the factors that exert a significant effect on the response. Hence, it is not the knowledge of the individual numerical value of each FE mean-related measure that matters but rather the sign (direction) and relative value of each to the others. These are best communicated in terms of the main effect and interaction plots.

Table 3-1: Mean and Variance related Importance Measures Usage

FE Importance Measures		**% Use**
Mean-Related	Numerical	31
	Graphical	40
Variance-Related	PC_1	10
	PC_2	8
	R^2	11
	R^2_{Adj}	6

With regard to the use of variance-related measures, the way in which both PC_1 and PC_2 were used causes some concern. One is the extent to which the difference between them is appreciated. They seem to have been used interchangeably as not only did they share the same label but also the same interpretation. Moreover, in three of the experiments, negative values were rendered when the PC_2 measure was used. No explanation was given as regards the meaning and reason for this. Furthermore, the negative values were handled in two different ways. Luo and Chen (2009) and Senthilkumar et al (2009) pooled their effects with the error. However, Shyha et al (2009) replaced them with zeros. The former treatment is more appropriate than the latter - see Chapter 8 for more details. Another concern regarding the use of the PC relates to confusing its estimation with the ANOVA. Although the estimation of PC_1 and PC_2 can be considered as a complementary step in performing the ANOVA, it was taken to be the ANOVA in 5 studies. None of the conventional elements of the ANOVA such as the sum of squares, degrees of freedom and the F-statistic were stated. In fact, only the effects' PCs were reported and used to subjectively judge their significance.

3.3.3 Results Interpretation and Other Aspects

An examination of how the results of applying the DOE techniques were interpreted in the studies revealed several issues most of which were centred around statistical significance and the p-value. Among the prime concerns are:

- Confusing statistical and practical significance - this was common in most of the studies
- The use of the F-statistic as a measure of the FE's importance (Jeang and Li, 2009; Chattopadhyay et al, 2009)
- Misinterpreting the p-value. For example Boronat et al (2009) plotted the p-values in descending order and interpreted each value as a measure of "how far an effect is from behaving randomly". By the same token, a p-value of less than 0.05 was taken to "suggest that the large F ratio is real and not a random error" in Rakwal and Bamberg (2009). Rosa et al (2009) defined the p-value as "the probability value which gives the degree of confidence at which the factor (or interaction) is significant". On the other hand, Lin and Ho, (2009) subtracted the p-value from one and used the results as a measure of the FE's importance.

A detailed discussion of these is given in Chapter 8. A further interpretation issue relates to differentiating between fixed and random effect inferences. In all of the reviewed experiments a subset of the levels of interest for each factor was specified but the conclusions were generalised to all possible levels. While such inferences are valid when the factor levels are randomly selected, this is not the case in fixed effect experiments where the conclusions regarding any factor must be conditioned on its selected levels.

As already mentioned, Taguchi's approach to DOE was adopted in most of the reviewed studies. The approach is characterised by running a one-shot fractional experiment to identify the important factors and their "best" settings. Several confirmation runs are then performed to verify that the experiment's objective is met at the identified settings. In fact, the confirmation experiment is necessary for validating Taguchi's assumptions regarding the absence of interactions, the statistical inference assumptions and the suitability of the S/N ratio. Such runs were only performed in 19 of the surveyed studies. Of these, 5 relied merely on comparing the result of the confirmation trial with either an initial value of the response variable or a value predicted using the developed model. No confidence interval was constructed. This is not a sound method for confirming the experimental results as no information regarding the predicted response variability was considered. Thus repeating the whole experiment may render a predicted response value that can be quite different from the confirmation experiment result. In fact, without a proper confirmation experimental procedure, there is no way to validate the findings of any conducted experiment especially when the limitations highlighted in this Chapter are considered.

The statistical software package used to analyse the experimental results was specified in only 22 of the studies. The reported packages and their usage percentage are shown in Figure 3-4.

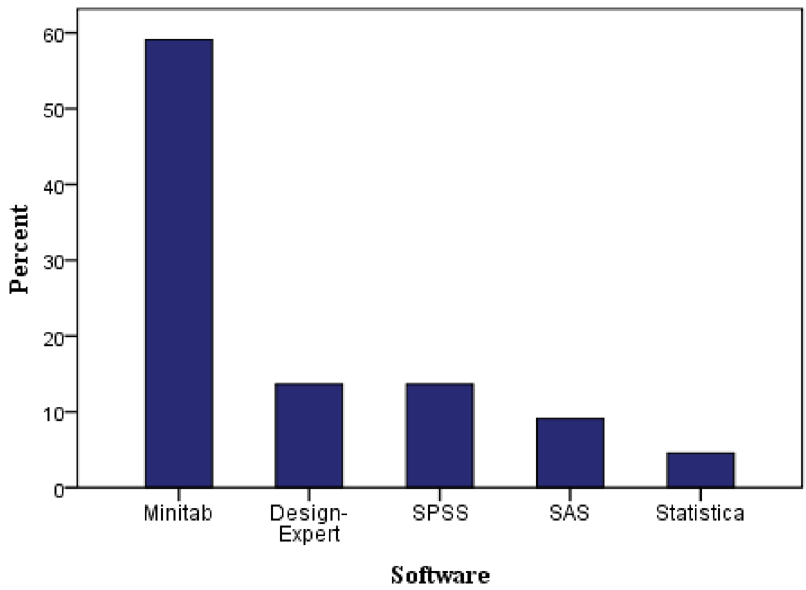


Figure 3-4: Reported Software Packages and their Usage

Only 34 of the reviewed articles included a reference to DOE books and/or articles. The references used in the remaining articles involved other Manufacturing Engineering papers in which DOE was applied. It would appear therefore, that the same methodology was applied perhaps without a rigorous understanding of the underpinning concepts which can have serious consequences.

The tendency to favour Manufacturing Engineering DOE application articles over other DOE references in the reviewed studies may, in no small part, be ascribed to their practicality, similarity to the application at hand and their limited DOE theoretical discussion. This is borne out by Tanco et al (2008) in their survey where they found that one of the major barriers for applying DOE was the lack of "real" applications in DOE references. In the same context, Goh (2001) highlighted that industrial applications could not be based on the standard experimental format presented in textbooks. Moreover, Penzias (1989) and Bisgaard (1991) argued that it is due to emphasising the theoretical aspects of statistics that practitioners are finding it difficult to employ them in practice.

In general, this review has cast some light on the limited use of DOE in the recent publications of the examined journals. Although, they may involve some non-academic participants, the published studies are predominantly conducted by academics. Interestingly, the reported frequency of DOE use in various industries, mainly by non-academic practitioners, exceeded that of the surveyed academics use. For example Antony and Banuelas (2002) found that among the companies that adopted six sigma programmes in the UK about 54% used DOE. This is in line with the findings of Arvidsson et al (2003) and Gremyr et al (2003) who respectively reported that 46% and 52% of the surveyed Swedish Manufacturing companies employ DOE. A lesser rate of use was observed in the Basque industries (Tanco et al, 2008) as only 20% of the companies investigated in that region of Spain were found to apply DOE.

Box (2001) noted that while journals in medical, social, and agricultural sciences feature a wide use of statistics, this is not the case in Engineering. In part, he attributed this to the lack of interest and awareness of the benefits of using statistics. Box's observation seems to be valid as far as the examined journals are concerned.

3.4 Summary

A snapshot of the contemporary practice of DOE in Manufacturing Engineering was provided in this Chapter by reviewing how it was applied in articles of three prominent journals in the field published in 2009. The review showed that DOE was far from being widely used. In the cases where it was applied, it was found that a gap existed between the DOE design and analysis techniques recommended in the statistical research literature and those that were actually used in the reviewed articles. Taguchi's philosophy regarding variance reduction was not widely adopted, but authors relied heavily on his design and analysis techniques to study the response average; rarely were any of the contemporary alternative robust techniques reported. Although highly fractional designs were commonly used, very seldom was the aliasing structure examined despite its potential impact. Moreover, Taguchi's dogmatic view regarding deemphasising interactions was adopted in the majority of the studies examined. Randomisation was not apparent in most of the articles reviewed and the assumptions underlying the used statistical tests were rarely reported.

Although the majority of the reviewed experiments were unreplicated, the practice of using pooling methods was poor. This was due to an overreliance being placed on either unaided subjective methods or unquestioned default options in the statistical packages used to analyse the results. Various misconceptions were identified with respect to the meaning of statistical significance, p-value interpretation, FE importance measures and the difference between fixed and random effect inferences. Despite the predominant use of Taguchi's approach, confirmation experiments were only conducted in a small proportion of the cases. DOE textbooks and papers were not cited in many of the articles reviewed but rather use was made of similar applications.

The results of this study cannot be generalised for all Manufacturing Engineering journals however, the importance and the high quality of the research reported in those reviewed cannot be ignored. It is reflective of an essential and considerable segment within the arena of Manufacturing Engineering research. Thus, even if the findings cannot be generalised, they do provide a platform for understanding and improving the practice of DOE in Manufacturing Engineering. Since this is a major objective of this research project, some of the issues have been dealt with in this Chapter; the others will be discussed in subsequent ones.

CHAPTER 4: ALIASING CONCEPT AND A METHOD FOR CONSTRUCTING ITS PATTERN

4.1 Introduction

Once the factors to be studied are decided upon, the next two interrelated steps are the selection of the OA and the assignment of the factors to its columns. The latter is straightforward when a full factorial OA is selected; however, this is not the case when a fractional factorial is used. The problem associated with selecting the fractional OA columns to accommodate the factors is that a pattern of aliasing exists which dictates what effects are interdependently estimated. In fact, losing the capability to independently estimate all the factorial effects is the price of using smaller OAs in such experiments. As revealed in the previous Chapter, little attention has been drawn to the aliasing problem in Manufacturing Engineering DOE applications. The main objectives of this Chapter are firstly to explain the nature of aliasing and the main methods for developing its pattern and measuring its degree; secondly to propose a simple method for obtaining the aliasing patterns and to demonstrate how it can be used to obtain the main measures of aliasing degree; and thirdly to illustrate how aliasing can impact the conclusions drawn from an experiment. After addressing each of the aforementioned objectives, the main conclusions to be gleaned from this Chapter are summarised.

4.2 The Concept of Aliasing

Aliasing is a form of linear relationship between some or all of the columns of the examined factorial effects. It may be perfect (complete) or partial. In the former case, given the entries of one column the elements of its alias columns can be completely determined whereas in the latter each of the effect's aliases carries only a certain part of its information.

Aliasing is a design related phenomenon in the sense that its presence in one design does not mean that the levels of the actual factors under study are interrelated. Its impact depends a great deal on the validity of the hierarchical ordering principle, according to which three-factor and higher order interaction effects are likely to be negligible. For example, in studying the effects of five two-level factors on a certain response, one option is to conduct a full factorial (2^5 =32 runs) experiment. Of the 31 degrees of freedom associated with this design, 16 are devoted to estimating three-factor and higher interactions. Assuming these are negligible, the remaining 15 degrees of freedom may be estimated by a smaller design such as the L_{16}. If due to resources constraints only 8 trials could be carried out, an L_8 (Table 1 in Appendix 1) comprising 3 main effect and 4 interaction columns could be used. A key question then is in which of the 4 interactions columns should the other 2 new main effects be put. Suppose the five factors under study are A, B, C, D and E and that the AB and AC interaction columns will respectively accommodate factors D and E then the resultant L_8 design is not capable of distinguishing the effect estimates of D from that of AB and similarly E from AC. Although, conventionally the expressions D=AB and E= AC are called *design generators*, they are not the only aliased effect in this case.

Before explaining the classical method of obtaining the complete aliasing pattern it is necessary to explain a multiplication rule that is commonly used for this purpose. When two identical letters (columns) are multiplied the resultant letter is I which denotes the identity column. All of its entries are 1s as it is used to estimate the overall average. Multiplying an effect represented by one or more letters by another yields an effect represented by the combination of the multiplied letters excluding the common ones. For example, multiplying ABC by ADE yields BCDE. The mathematical foundation of this multiplication can be found in Burton and Connor (1957). Applying this multiplication rule, the *generating relation* of the aforementioned L_8 design can be arrived at by multiplying the generators D=AB and E= AC respectively by D and E yielding I= ABD and I=ACE. These along with their generalised interaction (ABD*ACE =BDCE) constitute what is called the *defining relation*:

$$I = ABD=ACE= BDCE$$

Assuming that interactions of order five or higher are negligible, the complete aliasing pattern shown in Table 4-1 can be obtained by multiplying each effect by the elements of the defining relation using the above rule.

Table 4-1: The Complete Aliasing pattern for the 2^{5-2} design with generators D=AB and E =AC

Aliasing Pattern for the L_8 Design with the Defining Relation I = ABD = ACE = BCDE
A = BD =CE
B =AD = CDE =BACE
C =AE = BDE=ABCD
D =AB =BCE=ACED
E =AC =BCD =ABDE
BC = ACD = ABE =DE
BE =ADE= ABC = CD

Generally, a S^{k-p} fractional factorial design is called a $½^p$ fraction of the S^k design. It has k factors each at S levels and S^{k-p} runs. If p = 0, it is a full factorial design whereas if p > 0, the design is fractional i.e. it requires a selection of p independent generators. These along with their generalised interactions form a defining relation of the design. If Y and Z represent the selected generators (i.e. p = 2), then I = Y and I = Z are called the generating relations for the design. The complete defining relation for the design consists of all the columns that are equal to the identity column I. This comprises the $(S^p -1)/(S-1)$ elements, which are Y, Z, and their generalised interaction YZ i.e. the defining relation is I = Y = Z = YZ. Each of these is called a *word* and the number of letters contained within it is known as the *word length*. The aliases of any effect are obtained by multiplying it by each word in the defining relation.

On the basis of their aliasing pattern, OAs can be classified into two categories: regular and non-regular. A regular OA can be determined or constructed by its defining relation and has a simple aliasing structure in that any two effects are either orthogonal or fully aliased. A non-regular OA, on the other hand, cannot be determined by a defining relation and exhibits a complex aliasing structure as some of the effects are neither orthogonal nor fully aliased (partially aliased). The L_8, L_9,and L_{16} are examples of regular OAs whereas the Plackett and Burman (1946) L_{12} and the L_{18} exemplify the non-regular OAs.

When assigning factors to OAs, two situations have to be distinguished. The first is when prior knowledge suggests that the main effects and certain interactions are likely to be important. The general practice in this case is to assume that any unspecified interaction is negligible. To deal with this, Greenfield (1976) suggested using a searching algorithm for rendering the appropriate factor assignment so that none of the important effects is aliased. Franklin and Bailey (1977) and Franklin (1985) proposed some improvements on this but since their implementation requires computer programming their algorithms were rarely used in practice. As an alternative, Taguchi proposed a series of linear graphs which diagrammatically illustrate where each factor and interaction are located in the columns of an OA (Taguchi et al, 2004). Their main drawback is that they do not provide the complete aliasing relationship. Tsui (1988) suggested that aliasing tables should be used to rectify this drawback. Graphical representations of these tables were provided for both two-level (Kacker and Tsui, 1990; Wu and Chen, 1992) and three-level (Sun and Wu, 1994) OAs.

The second situation when assigning factors to columns of an OA arises when there is little or no knowledge about the relative sizes of the factorial effects or when all the factorial effects are equally important. This is the most frequently encountered situation in practice

(Box et al, 2005). In this case it is required to select designs with good overall properties under the hierarchical principal using certain criteria to measure the degree of aliasing.

4.3 Measures of Overall Aliasing

Several criteria have been proposed in the DOE literature, of which the main three are maximum resolution, minimum aberration and minimum generalised aberration. The first two are proposed for measuring the aliasing degree in regular OAs and the last in non-regular two-level and three-level (regular and non-regular) OAs.

4.3.1 Maximum Resolution

A design resolution (R) is defined as the length of the shortest word in the defining relation. Generally, a design is of resolution R if no p-factor effect is aliased with another effect containing less than R – p factors (Box and Hunter, 1961a). Therefore the best fractional design is the one that has the highest possible resolution. For example consider the defining relation of the 2^{5-2} design in Table 4-1. Its shortest two words consist of three letters i.e. it is of resolution III. In such designs the main effects are not aliased with each other but they are aliased with two-factor interactions. By the same token designs of resolution IV do not alias main effects with each other or with any two-factor interaction but they do alias two-factor interactions with each other. Clearly, a lower resolution design involves words with shorter length reflecting the presence of aliasing among lower order effects. As these are more important than the higher order effects, it is best to select a design with maximum resolution.

4.3.2 Minimum Aberration

Fries and Hunter (1980) observed that fractional designs which have maximum resolution are not equally good and suggested an alternative criterion which they called minimum aberration. Let $A_i(d)$ denote the number of words of length i in the defining relation for the design d. The vector

$$W(d) = (A_1(d), A_2(d), A_3(d), A_4(d), \ldots) \qquad (4.1)$$

is called the *world length* pattern (Wu and Zhang, 1993). The resolution of a design d is the smallest i such that $A_i(d) \geq 1$. For any two 2^{k-p} designs d_1 and d_2, let i be the smallest integer such that $A_i(d_1) \neq A_i(d_2)$. Then design d_1 is said to have less aberration than d_2 if $A_i(d_1) < A_i(d_2)$. If there is no design with less aberration than d_1 then it has the minimum aberration.

Consider the 2^{5-2} design in Table 4-1. As its defining relation has no word of length 1 or 2 and as there are 2 words of length 3 and 1 of length 4, the word length pattern of this design is (0, 0, 2, 1). Since the first non-zero entry of the vector is associated with the number of words of length 3, the design is of resolution III. The rationale behind the minimum aberration criterion is that by sequentially minimising the number of words of shortest length, the number of aliased low order effects which are regarded more important than the high order ones is minimised. For illustration, consider the following two possible 2^{7-2} designs (d_1 and d_2) obtained by two different sets of generators. For d_1 let the generators be F=ABD and G=ACE and for d_2 let them be F=ABCD and G=ABCE. Consequently the defining relation of design d_1 is

$$I = ABDF = ACEG = BCDEFG$$

and that of d_2 is

$$I = DEFG = ABCDF = ABCEG$$

The word length patterns for designs d_1 and d_2 are (0, 0, 0, 2, 0, 1) and (0, 0, 0, 1, 2) respectively. Clearly both of the designs are of resolution IV; however since $A_4(d_2) < A_4(d_1)$,

d_2 is considered to be better than d_1 because it has lower aberration i.e. d_1 has a fewer number of aliased low order effects than d_2.

Details on the methods of obtaining minimum aberration designs can be found in Franklin (1984), Chen (1992) and Chen and Wu (1991). A useful catalogue of minimum aberration designs is published in Chen et al (1993).

4.3.3 Minimum Generalised Aberration

For two-level non-regular OAs, Tang and Deng (1999) proposed an extension of the minimum aberration criterion. For an experimental design d with n runs and m columns, let s $=\{c_1, c_2 \ldots c_k\}$ be a subset of k columns of d, and define

$$J_k(s) = \left| \sum_{i=1}^{n} c_{i1} c_{i2} \ldots\ldots c_{ik} \right| \qquad (4.2)$$

where c_{ij} is the i^{th} component of column c_j and $J_k(s)$ is the absolute value of the sum of entry wise products for k of the m columns (Deng and Tang, 1999). Given $J_k(s)$ define

$$B_k(d) = n^{-2} \sum_{|s|=k} [J_k(s)]^2 \qquad (4.3)$$

For two designs d_1 and d_2 let i be the smallest integer such that $B_i(d_1) \neq B_i(d_2)$. if $B_i(d_1) < B_i(d_2)$ then d_1 has lower generalised aberration (lower G_2 aberration) than d_2. If no other design has lower G_2 aberration than d_1 then it has the minimum G_2 aberration (Tang and Deng, 1999).

Xu and Wu (2001) extended the above criterion to multilevel designs and proposed a generalised minimum aberration. For a factorial design of size N, the full ANOVA model for an experiment involving m factors is:

$$Y = \alpha_0 X_0 + \alpha_1 X_1 + \ldots\ldots + \alpha_m X_m + e \qquad (4.4)$$

where Y is the response vector, α_0 the general mean, α_k is the vector of all k-factor effects, X_0 is the vector of 1's and $X_k = [x_{ij}^{(k)}]$ is the matrix of contrast coefficients for α_k. Let

$$A_k(d) = N^{-2} \sum_{s=1}^{nk} \left| \sum_{i=1}^{N} x_{ij}^{(k)} \right|^2 \qquad (4.5)$$

where n_k is the number of all k-factor effects. For the vector of *generalised word length* pattern $W(d) = (A_1(d), A_2(d), \ldots A_m(d))$ the generalised minimum aberration can be obtained by sequentially minimising $A_k(d)$ for $k = 1,2, 3 \ldots m$.

For regular two-level designs both G_2 aberration and the generalised aberration reduce to minimum aberration (Cheng and Tang, 2005). Thus their rationale is the same as that of the minimum aberration criteria. Details of the construction of generalised aberration designs can be found in Fang et al, (2003).

4.4 A Method for Generating the Aliasing Pattern

Before explaining the proposed method for obtaining aliasing patterns, it is necessary to recall some algebraic concepts. A vector is a set of n numbers arranged in a definite order (Pease, 1965). When the average of its elements is subtracted from their individual values, it is referred to as a mean-centred vector, the sum of its elements becoming zero (Gentle, 2003). The length of a vector (or its magnitude) is the square root of the sum of its squared elements. When each of the elements of a vector is divided by its length it is said to be normalised.

Two vectors are said to be orthogonal if they are at right angles (Williams, 2009) and aliasing is a form of departure from orthogonality. Therefore, viewing the columns of any OA as vectors, the degree of aliasing or lack of orthogonality between any two columns may be measured by the extent to which the angle between them differs from 90^0. Rather than finding the angles between the columns, it is easier to obtain their cosines. Generally, the inner product of any two mean-centred normalised vectors yields the cosine of the angle between them (Massart and Vandeginste, 1997). If the angle between two columns is 90^0 (or -90^0) then the value of its cosine is zero i.e. they are orthogonal. However, as the cosine value departs from zero, the degree of aliasing increases. For any non-zero cosine value that lies between -1 and +1 (exclusively), the two columns are said to be partially aliased. However, for cosine values of 1 or -1 the two columns are said to be perfectly aliased in the same or opposite directions respectively.

Using these concepts, the steps of implementing the proposed method for obtaining the aliasing pattern are as follows:

1. For a given design array, construct the corresponding model array **M** using an appropriate coding system. This should incorporate all the effects for which the aliases are to be determined.
2. Let x_{ij} be the i^{th} entry of column j, $\bar{x}_j$ be the average of the entries of column j and denote the number of rows in the model array by n. For each column of **M**, use

$$\frac{x_{ij} - \bar{x}_j}{\sqrt{\sum_{i}^{n}(x_{ij} - \bar{x}_j)^2}} \tag{4.6}$$

   to generate the corresponding column in the mean-centred normalised model array $\mathbf{X_n}$
3. Generate the aliasing array $\mathbf{X_n}^T\mathbf{X_n}$

The entries of this resultant array are the cosines of the angles between all the possible pairs of the model array's columns. Therefore, the degree of aliasing between any two effects can be measured using their corresponding entry in the $\mathbf{X_n}^T\mathbf{X_n}$ array. Three examples will now be used to show the applications of this method. The first deals with the regular two-level L_8 2^{5-2} design presented in section 4.2 whereas the second concerns a non-regular two-level L_{12}

design used to study five factors, and the third example is based on a non-regular L_{18} design used to study three factors at three levels.

4.4.1 Regular Two-Level L_8

As highlighted in Section 4.2 the two generators of the L_8 (2^{5-2}) design are D=AB and E= AC (see Table 4-2). The first step in implementing the suggested model is to construct the model array **M**, shown in Table 4-3. Assuming that interactions of order three or higher are negligible, this should comprise the main effects and all of their two factor interaction columns. Due to the adoption of the (-1,+1) coding system, the interaction columns can be rendered by multiplying the individuals elements of the columns that accommodate the interacting factors. If the (1,2) Taguchi coding system is used then the interaction columns can be formed using a special rule whereby an interaction value of "1" is used if the rows of the two interacting factors contain either "1" and "1" or "2" and "2",whereas an interaction value of "2" is used when these rows contain "1" and "2" or "2" and "1”. The second step in the proposed method is to mean-centre and normalise each of the **M** columns as shown in Table 4-4. Finally the aliasing array $\mathbf{X_n}^T\mathbf{X_n}$ is constructed and is shown in Table 4-5. The entries of the aliasing array are referred to as the aliasing coefficients as they quantify the degree of aliasing between the studied effects. An aliasing coefficient of value 1 or -1 reflects a perfect aliasing whereas any other non-zero value indicates a partial aliasing. Zero aliasing coefficients imply that the corresponding effects are orthogonal. As the L_8 is a regular OA, the entries of its $\mathbf{X_n}^T\mathbf{X_n}$ can assume one of three values; namely -1, 0 and 1. This is because in regular designs the effects are either fully aliased (in the same or opposite direction) or orthogonal. The entries on the main diagonal of the $\mathbf{X_n}^T\mathbf{X_n}$ give the aliasing coefficient of one

Table 4-2: 2^{5-2} L_8 Design Array

No.	A	B	C	D =AB	E = AC
1	-1	-1	-1	1	1
2	1	-1	-1	-1	-1
3	-1	1	-1	-1	1
4	1	1	-1	1	-1
5	-1	-1	1	1	-1
6	1	-1	1	-1	1
7	-1	1	1	-1	-1
8	1	1	1	1	1

Table 4-3: Model Array for the L_8 Design

A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
-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	-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	1	1	1	1	1	1	1	1

Table 4-4: Mean-Centred Normalised Model Array for the L_8 Design

A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
-0.35	-0.35	-0.35	0.35	0.35	0.35	0.35	-0.35	-0.35	0.35	-0.35	-0.35	-0.35	-0.35	0.35
0.35	-0.35	-0.35	-0.35	-0.35	-0.35	-0.35	-0.35	-0.35	0.35	0.35	0.35	0.35	0.35	0.35
-0.35	0.35	-0.35	-0.35	0.35	-0.35	0.35	0.35	-0.35	-0.35	-0.35	0.35	0.35	-0.35	-0.35
0.35	0.35	-0.35	0.35	-0.35	0.35	-0.35	0.35	-0.35	-0.35	0.35	-0.35	-0.35	0.35	-0.35
-0.35	-0.35	0.35	0.35	-0.35	0.35	-0.35	-0.35	0.35	-0.35	-0.35	0.35	0.35	-0.35	-0.35
0.35	-0.35	0.35	-0.35	0.35	-0.35	0.35	-0.35	0.35	-0.35	0.35	-0.35	-0.35	0.35	-0.35
-0.35	0.35	0.35	-0.35	-0.35	-0.35	-0.35	0.35	0.35	0.35	-0.35	-0.35	-0.35	-0.35	0.35
0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35

Table 4-5: Aliasing Array for the 2^{5-2} L_8 Design

	A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
A	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0
B	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
C	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0
D	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
E	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
AB	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
AC	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
AD	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
AE	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0
BC	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
BD	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0
BE	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
CD	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
CE	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0
DE	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1

effect with itself which is always 1 and the entries off the main diagonal are the pair-wise aliasing coefficients between all the considered effects. To obtain the aliases of, say, factor A all the non-zero entries of its row in the $\mathbf{X_n}^T\mathbf{X_n}$ array should be examined. Clearly both BD and CE are fully aliased with A. This is in line with the results of using the conventional method of obtaining the aliasing pattern displayed in Table 4-1. As the three-factor and higher order interactions were deemed negligible, they were not incorporated in the model array of the L_8 design. Consequently, their impact on aliasing was not manifested in the $\mathbf{X_n}^T\mathbf{X_n}$. In the cases where these effects are thought to exert considerable effects on the response variable, they should be included in the model array so that their impact on aliasing can be examined.

4.4.2 Non-regular Two-Level L_{12}

One major drawback of the conventional method of generating aliasing patterns relates to the difficulty associated with its extension to deal with aliasing in non-regular designs such as the L_{12}. This problem can be alleviated by adopting the proposed method of which the primary advantage lies in the capability of dealing with regular and non-regular, two-level and three-level designs in a unified manner. To illustrate this consider the L_{12} design array shown in Table 4-6. The first five columns are used to accommodate the five factors under study. The first step of the proposed method is to construct the model array. Assuming that the

Table 4-6: L_{12} Design Array

	A	B	C	D	E							
No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**Response**
1	1	1	1	1	1	1	1	1	1	1	1	23.6
2	1	1	-1	1	-1	1	-1	-1	-1	1	-1	10.4
3	1	1	-1	-1	1	-1	1	-1	-1	-1	1	10.4
4	-1	1	1	-1	-1	1	1	1	-1	-1	-1	3.6
5	1	-1	-1	1	-1	-1	1	1	1	-1	-1	-5.6
6	-1	1	-1	-1	1	-1	-1	1	1	1	-1	-9.6
7	-1	-1	-1	-1	-1	1	1	-1	1	1	1	2.4
8	1	-1	1	-1	-1	-1	-1	1	-1	1	1	-8.4
9	-1	1	1	1	-1	-1	-1	-1	1	-1	1	3.6
10	-1	-1	1	1	1	-1	1	-1	-1	1	-1	-0.4
11	-1	-1	-1	1	1	1	-1	1	-1	-1	1	2.4
12	1	-1	1	-1	1	1	-1	-1	1	-1	-1	-8.4

main effects and two-factor interactions are the only effects of interest, the design array selected columns can be used to generate all the possible two-factor interaction columns so that the model array **M** is constructed - see Table 4-7. The columns of these should then be mean-centred and normalised to yield $\mathbf{X_n}$ (Table 4-8). The aliasing pattern associated with the examined effects can be generated by constructing the $\mathbf{X_n}^T\mathbf{X_n}$ array which is shown in Table 4-9. Clearly each main effect is partially aliased, with an aliasing coefficient of ±0.333, with every two-factor interaction not involving itself. Furthermore, certain two-factor interactions are partially aliased with each other to the same degree. The proposed model, therefore, provides a straightforward method for obtaining the aliasing pattern in non-regular two-level designs.

Table 4-7: Model Array for the L_{12} Design Array

A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
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	-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	-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	-1	1	-1	1	-1	1	-1	1	-1	-1	1	-1

Table 4-8: Mean-Centred Normalised Array for the L_{12} Design Array

A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29
0.29	0.29	-0.29	0.29	-0.29	0.29	-0.29	0.29	-0.29	-0.29	0.29	-0.29	-0.29	0.29	-0.29
0.29	0.29	-0.29	-0.29	0.29	0.29	-0.29	-0.29	0.29	-0.29	-0.29	0.29	0.29	-0.29	-0.29
-0.29	0.29	0.29	-0.29	-0.29	-0.29	-0.29	0.29	0.29	0.29	-0.29	-0.29	-0.29	-0.29	0.29
0.29	-0.29	-0.29	0.29	-0.29	-0.29	-0.29	0.29	-0.29	0.29	-0.29	0.29	-0.29	0.29	-0.29
-0.29	0.29	-0.29	-0.29	0.29	-0.29	0.29	0.29	-0.29	-0.29	-0.29	0.29	0.29	-0.29	-0.29
-0.29	-0.29	-0.29	-0.29	-0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29
0.29	-0.29	0.29	-0.29	-0.29	-0.29	0.29	-0.29	-0.29	-0.29	0.29	0.29	-0.29	-0.29	0.29
-0.29	0.29	0.29	0.29	-0.29	-0.29	-0.29	-0.29	0.29	0.29	0.29	-0.29	0.29	-0.29	-0.29
-0.29	-0.29	0.29	0.29	0.29	0.29	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29	0.29	0.29	0.29
-0.29	-0.29	-0.29	0.29	0.29	0.29	0.29	-0.29	-0.29	0.29	-0.29	-0.29	-0.29	-0.29	0.29
0.29	-0.29	0.29	-0.29	0.29	-0.29	0.29	-0.29	0.29	-0.29	0.29	-0.29	-0.29	0.29	-0.29

4.4.3 Three-level L_{18} designs

To study three factors at three levels using the L_{18}, three columns need to be selected. As they are associated with the generalised minimum aberration (Xu, 2001), columns 3, 4 and 5 were selected as shown in Table 4-10. To implement the proposed method with this design, the

Table 4-9: Aliasing Array for the L_{12} Design

	A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
A	1	0	0	0	0	0	0	0	0	-0.33	0.33	0.33	-0.33	0.33	-0.33
B	0	1	0	0	0	0	-0.33	0.33	0.33	0	0	0	0.33	-0.33	-0.33
C	0	0	1	0	0	-0.33	0	-0.33	0.33	0	0.33	-0.33	0	0	0.33
D	0	0	0	1	0	0.33	-0.33	0	-0.33	0.33	0	-0.33	0	0.33	0
E	0	0	0	0	1	0.33	0.33	-0.33	0	-0.33	-0.33	0	0.33	0	0
AB	0	0	-0.33	0.33	0.33	1	0	0	0	0	0	0	0.33	0.33	0.33
AC	0	-0.33	0	-0.33	0.33	0	1	0	0	0	0.33	0.33	0	0	0.33
AD	0	0.33	-0.33	0	-0.33	0	0	1	0	0.33	0	0.33	0	0.33	0
AE	0	0.33	0.33	-0.33	0	0	0	0	1	0.33	0.33	0	0.33	0	0
BC	-0.33	0	0	0.33	-0.33	0	0	0.33	0.33	1	0	0	0	0	0.33
BD	0.33	0	0.33	0	-0.33	0	0.33	0	0.33	0	1	0	0	0.33	0
BE	0.33	0	-0.33	-0.33	0	0	0.33	0.33	0	0	0	1	0.33	0	0
CD	-0.33	0.33	0	0	0.33	0.33	0	0	0.33	0	0	0.33	1	0	0
CE	0.33	-0.33	0	0.33	0	0.33	0	0.33	0	0	0.33	0	0	1	0
DE	-0.33	-0.33	0.33	0	0	0.33	0.33	0	0	0.33	0	0	0	0	1

Table 4-10: L_{18} Design Array

			A	B	C			
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	1	1	1	1	1	1	1	1
2	1	1	2	2	2	2	2	2
3	1	1	3	3	3	3	3	3
4	1	2	1	1	2	2	3	3
5	1	2	2	2	3	3	1	1
6	1	2	3	3	1	1	2	2
7	1	3	1	2	1	3	2	3
8	1	3	2	3	2	1	3	1
9	1	3	3	1	3	2	1	2
10	2	1	1	3	3	2	2	1
11	2	1	2	1	1	3	3	2
12	2	1	3	2	2	1	1	3
13	2	2	1	2	3	1	3	2
14	2	2	2	3	1	2	1	3
15	2	2	3	1	2	3	2	1
16	2	3	1	3	2	3	1	2
17	2	3	2	1	3	1	2	3
18	2	3	3	2	1	2	3	1

first step is to construct the model array **M**. In so doing there are various methods for coding the design array as discussed in Appendix 2. As some of them are orthogonal and others are not, each coding system leads to a different aliasing pattern. Due to its orthogonality, the L-Q system is adopted in this study. Consequently it was used to generate the **M** array for the L_{18} design as displayed in Table 4-11. Again the interactions of order three and higher were ignored and the two-factor interaction columns were rendered from the inner product of the columns of the factors that comprise them. The second step is to mean-centre and normalise the model array to generate the $\mathbf{X_n}$ array seen in Table 4-12. To obtain the aliasing pattern, the array $\mathbf{X_n}^T\mathbf{X_n}$ should be constructed - see Table 4-13. Clearly, there is partial aliasing between the main effects and certain two-factor interactions and among the two-factor interactions.

4.5 Obtaining the Aliasing Measures using the Proposed Method

The discussion of the three main aliasing measures revealed that they rely to a great extent on the word length pattern. In fact, once this is obtained for, say, each of two designs, the best in terms of aliasing can be identified by sequentially minimising the number of words of length k ($A_k(d)$) for k = 1,2, 3 .....m, where m is the length of the longest word. A key advantage of the method discussed in the previous section is the ease with which the word length pattern of any design can be obtained from the aliasing array $\mathbf{X_n}^T\mathbf{X_n}$. To illustrate this, assume that three two-level factors with negligible interactions of order three or above were studied and consider the four sub-matrices of the $\mathbf{X_n}^T\mathbf{X_n}$ array shown in Table 4-14. The sub-matrix A is a squared matrix in which the number of rows and columns is equal to the number of main effects under study. Its “off diagonal” entries are measures of the aliasing between main effects. If the two main effects, A and B, are perfectly aliased then the corresponding generator would be A=B. Multiplying both sides by A the resultant word would be I = AB

Table 4-11: Model Array for the L_{18} Design Array

Al	**Aq**	**Bl**	**Bq**	**Cl**	**Cq**	**AlBl**	**AlBq**	**AlCl**	**AlCq**	**AqBl**	**AqBq**	**AqCl**	**AqCq**	**BlCl**	**BlCq**	**BqCl**	**BqCq**
-1	1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1	-1	1
0	-2	0	-2	0	-2	0	0	0	0	0	4	0	4	0	0	0	4
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
-1	1	-1	1	0	-2	1	-1	0	2	-1	1	0	-2	0	2	0	-2
0	-2	0	-2	1	1	0	0	0	0	0	4	-2	-2	0	0	-2	-2
1	1	1	1	-1	1	1	1	-1	1	1	1	-1	1	-1	1	-1	1
-1	1	0	-2	-1	1	0	2	1	-1	0	-2	-1	1	0	0	2	-2
0	-2	1	1	0	-2	0	0	0	0	-2	-2	0	4	0	-2	0	-2
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
0	-2	-1	1	-1	1	0	0	0	0	2	-2	2	-2	1	-1	-1	1
1	1	0	-2	0	-2	0	-2	0	-2	0	-2	0	-2	0	0	0	4
-1	1	0	-2	1	1	0	2	-1	-1	0	-2	1	1	0	0	-2	-2
0	-2	1	1	-1	1	0	0	0	0	-2	-2	2	-2	-1	1	-1	1
1	1	-1	1	0	-2	-1	1	0	-2	-1	1	0	-2	0	2	0	-2
-1	1	1	1	0	-2	-1	-1	0	2	1	1	0	-2	0	-2	0	-2
0	-2	-1	1	1	1	0	0	0	0	2	-2	-2	-2	-1	-1	1	1
1	1	0	-2	-1	1	0	-2	-1	1	0	-2	-1	1	0	0	2	-2

Table 4-12: Mean-Centred Normalised Array for the L_{18} Design

Al	**Aq**	**Bl**	**Bq**	**Cl**	**Cq**	**AlBl**	**AlBq**	**AlCl**	**AlCq**	**AqBl**	**AqBq**	**AqCl**	**AqCq**	**BlCl**	**BlCq**	**BqCl**	**BqCq**
-0.29	0.17	-0.29	0.17	-0.29	0.17	0.35	-0.20	0.35	-0.20	-0.20	0.12	-0.20	0.12	0.35	-0.20	-0.20	0.12
0.00	-0.33	0.00	-0.33	0.00	-0.33	0.00	0.00	0.00	0.00	0.00	0.47	0.00	0.47	0.00	0.00	0.00	0.47
0.29	0.17	0.29	0.17	0.29	0.17	0.35	0.20	0.35	0.20	0.20	0.12	0.20	0.12	0.35	0.20	0.20	0.12
-0.29	0.17	-0.29	0.17	0.00	-0.33	0.35	-0.20	0.00	0.41	-0.20	0.12	0.00	-0.24	0.00	0.41	0.00	-0.24
0.00	-0.33	0.00	-0.33	0.29	0.17	0.00	0.00	0.00	0.00	0.00	0.47	-0.41	-0.24	0.00	0.00	-0.41	-0.24
0.29	0.17	0.29	0.17	-0.29	0.17	0.35	0.20	-0.35	0.20	0.20	0.12	-0.20	0.12	-0.35	0.20	-0.20	0.12
-0.29	0.17	0.00	-0.33	-0.29	0.17	0.00	0.41	0.35	-0.20	0.00	-0.24	-0.20	0.12	0.00	0.00	0.41	-0.24
0.00	-0.33	0.29	0.17	0.00	-0.33	0.00	0.00	0.00	0.00	-0.41	-0.24	0.00	0.47	0.00	-0.41	0.00	-0.24
0.29	0.17	-0.29	0.17	0.29	0.17	-0.35	0.20	0.35	0.20	-0.20	0.12	0.20	0.12	-0.35	-0.20	0.20	0.12
-0.29	0.17	0.29	0.17	0.29	0.17	-0.35	-0.20	-0.35	-0.20	0.20	0.12	0.20	0.12	0.35	0.20	0.20	0.12
0.00	-0.33	-0.29	0.17	-0.29	0.17	0.00	0.00	0.00	0.00	0.41	-0.24	0.41	-0.24	0.35	-0.20	-0.20	0.12
0.29	0.17	0.00	-0.33	0.00	-0.33	0.00	-0.41	0.00	-0.41	0.00	-0.24	0.00	-0.24	0.00	0.00	0.00	0.47
-0.29	0.17	0.00	-0.33	0.29	0.17	0.00	0.41	-0.35	-0.20	0.00	-0.24	0.20	0.12	0.00	0.00	-0.41	-0.24
0.00	-0.33	0.29	0.17	-0.29	0.17	0.00	0.00	0.00	0.00	-0.41	-0.24	0.41	-0.24	-0.35	0.20	-0.20	0.12
0.29	0.17	-0.29	0.17	0.00	-0.33	-0.35	0.20	0.00	-0.41	-0.20	0.12	0.00	-0.24	0.00	0.41	0.00	-0.24
-0.29	0.17	0.29	0.17	0.00	-0.33	-0.35	-0.20	0.00	0.41	0.20	0.12	0.00	-0.24	0.00	-0.41	0.00	-0.24
0.00	-0.33	-0.29	0.17	0.29	0.17	0.00	0.00	0.00	0.00	0.41	-0.24	-0.41	-0.24	-0.35	-0.20	0.20	0.12
0.29	0.17	0.00	-0.33	-0.29	0.17	0.00	-0.41	-0.35	0.20	0.00	-0.24	-0.20	0.12	0.00	0.00	0.41	-0.24

Table 4-13: Aliasing Array for the L_{18} Design

	Al	**Aq**	**Bl**	**Bq**	**Cl**	**Cq**	**AlBl**	**AlBq**	**AlCl**	**AlCq**	**AqBl**	**AqBq**	**AqCl**	**AqCq**	**BlCl**	**BlCq**	**BqCl**	**BqCq**
Al	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.31	0.18	0.18	0.31
Aq	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.31	0.31	-0.18
Bl	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	-0.31	0.18	0.00	0.00	0.18	0.31	0.00	0.00	0.00	0.00
Bq	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.18	0.31	0.00	0.00	0.31	-0.18	0.00	0.00	0.00	0.00
Cl	0.00	0.00	0.00	0.00	1.00	0.00	-0.31	0.18	0.00	0.00	0.18	0.31	0.00	0.00	0.00	0.00	0.00	0.00
Cq	0.00	0.00	0.00	0.00	0.00	1.00	0.18	0.31	0.00	0.00	0.31	-0.18	0.00	0.00	0.00	0.00	0.00	0.00
AlBl	0.00	0.00	0.00	0.00	-0.31	0.18	1.00	0.00	0.13	0.22	0.00	0.00	-0.22	0.13	0.13	0.22	-0.22	0.13
AlBq	0.00	0.00	0.00	0.00	0.18	0.31	0.00	1.00	0.22	-0.13	0.00	0.00	0.13	0.22	-0.22	0.13	-0.13	-0.22
AlCl	0.00	0.00	-0.31	0.18	0.00	0.00	0.13	0.22	1.00	0.00	-0.22	0.13	0.00	0.00	0.13	-0.22	0.22	0.13
AlCq	0.00	0.00	0.18	0.31	0.00	0.00	0.22	-0.13	0.00	1.00	0.13	0.22	0.00	0.00	-0.22	-0.13	0.13	-0.22
AqBl	0.00	0.00	0.00	0.00	0.18	0.31	0.00	0.00	-0.22	0.13	1.00	0.00	-0.13	-0.22	0.22	-0.13	0.13	0.22
AqBq	0.00	0.00	0.00	0.00	0.31	-0.18	0.00	0.00	0.13	0.22	0.00	1.00	-0.22	0.13	0.13	0.22	-0.22	0.13
AqCl	0.00	0.00	0.18	0.31	0.00	0.00	-0.22	0.13	0.00	0.00	-0.13	-0.22	1.00	0.00	0.22	0.13	-0.13	0.22
AqCq	0.00	0.00	0.31	-0.18	0.00	0.00	0.13	0.22	0.00	0.00	-0.22	0.13	0.00	1.00	0.13	-0.22	0.22	0.13
BlCl	-0.31	0.18	0.00	0.00	0.00	0.00	0.13	-0.22	0.13	-0.22	0.22	0.13	0.22	0.13	1.00	0.00	0.00	0.00
BlCq	0.18	0.31	0.00	0.00	0.00	0.00	0.22	0.13	-0.22	-0.13	-0.13	0.22	0.13	-0.22	0.00	1.00	0.00	0.00
BqCl	0.18	0.31	0.00	0.00	0.00	0.00	-0.22	-0.13	0.22	0.13	0.13	-0.22	-0.13	0.22	0.00	0.00	1.00	0.00
BqCq	0.31	-0.18	0.00	0.00	0.00	0.00	0.13	-0.22	0.13	-0.22	0.22	0.13	0.22	0.13	0.00	0.00	0.00	1.00

Table 4-14 Sub-Matrices of the $X_n^TX_n$ Aliasing Array

		Main Effects			Two-factor interaction		
		A	B	C	AB	AC	BC
Main Effects	A	Sub-matrix A			Sub-matrix B		
	B						
	C						
Two-factor Interactions	AB	Sub-matrix C			Sub-matrix D		
	AC						
	BC						

which is of length 2. As the sub-matrix A is symmetrical, each of its aliasing coefficients appears twice (above and below the diagonal elements). Each entry of say 1 (or -1) represents a word of length two. Thus the number of words of length 2 can be obtained by computing the sum of squares of the “off diagonal” entries of the sub-matrix A divided by 2. With regard to sub-matrix B, it is clear that the number of rows is equal to the number of studied main effects whereas the number of its columns is equal to that of the two-factor interactions. As its entries measure the aliasing between main effects and two-factor interactions, each entry of 1 or -1 represents a word of length 3. However, in this case each of the sub-matrix entries appears three times. For example the entry corresponding to the main effect A and the two factor interaction BC appears in two other forms, one of which is represented by the coefficient of aliasing between the main effect B and the two-factor interaction AC and the other corresponds to the aliasing coefficient of the main effect C and the two-factor interaction AB. Therefore, the number of words of length 3 can be obtained by summing the squares of the entries of the sub-matrix B divided by 3. Using a similar argument the number of words of length 4 can be obtained by calculating the sum of the “off diagonal” entries of the sub matrix D squared divided by 6. Hence, by examining the number of times each word appears in the appropriate sub-matrix, the number of words of any length can be obtained from the aliasing array $\mathbf{X_n}^T\mathbf{X_n}$. This is equivalent to obtaining the word length pattern using the defining relation or the generalised minimum aberration measures (equations 4.3 and 4.5). To illustrate this consider the aliasing array of the 2^{5-2} L_8 shown in Table 4-5. Clearly all the “off diagonal” entries of the sub-matrix A are zeros indicating the absence of words of length 2. For the sub-matrix B, the sum of its squared entries is 6. So the number of words of length 3 is 6/3=2. For words of length 4, the sum of the “off diagonal” entries of the sub-matrix D squared is 6. Therefore, the number of words of length 4 is 6/6=1. Consequently, the word length pattern of this 2^{5-2} design is (0, 0, 2, 1). This can be verified from the defining relation of this design (Table 4-1) which comprises two words of length 3 and one of length 4.

Applying the same method described above to the aliasing arrays of the L_{12} and L_{18} the resultant word lengths are (0, 0, 1.11, 0.56) and (0, 0, 0.5) respectively. The latter can be verified by examining Xu’s (2001) study in which the same word lengths pattern was arrived at when investigating the L_{18} model array shown in Table 4-11. An aspect of interest here is that despite the presence of non-zero entries in the sub-matrix D of the L_{18} aliasing array (Table 4-13), no words of length 4 were observed in its word lengths pattern. This is due to the fact that all the latter entries correspond to two-factor interactions that share at least one letter. For example the interactions AlBq and BlCq are partially aliased with a coefficient of 0.13. Nevertheless, their ability to form a word of length 4 is impaired by the presence of B in each of them.

The upshot of the preceding discussion is that the proposed method can be used to

- construct the aliasing array of any fractional design,
- obtain its word length pattern as a means of measuring the degree of its associated aliasing,
- obtain the best subset of fractional designs’ columns that render a word length pattern with minimum aberration.

4.6 Impact of Aliasing on the Statistical Analyses of Experimental Data

The presence of aliasing between the studied effects can seriously affect the quality of the conclusions drawn from an experiment, since it impacts both estimation and hypothesis testing. To explain this, the t-statistic, presented in equation 13 of Appendix 2, will be used to demonstrate how aliasing can affect the regression coefficient estimate and its standard error.

Denoting an aliasing coefficient by ω, the aliasing between factors A and B may be mathematically represented by the following equation

$$X_A = \omega X_B + c \quad (4.7)$$

where X_A and X_B are the factor columns and c is a stochastic error term. If A is perfectly aliased with B then ω =1 and c=0 indicating that column A is fully determined by column B. In the case of partial aliasing ω can assume any non-zero value between -1 and +1 exclusively and c is a non-zero value reflecting the overall impact of all the other partially aliased effects that completely determine X_A. Assume that the true relation between the response variable and factors A and B can be approximated by the following model:

$$Y_i = \hat{\beta}_o + \hat{\beta}_1 X_A + \hat{\beta}_2 X_B \quad (4.8)$$

When A and B are perfectly aliased, substituting the value of X_A from equation 4.7 in equation 4.8 yields

$$Y_i = \hat{\beta}_o + (\omega\hat{\beta}_1 + \hat{\beta}_2) X_B \quad (4.9)$$

Clearly the effect of factor B is indistinguishable from A and may be either overestimated or underestimated, depending on the signs of $\hat{\beta}_1$, $\hat{\beta}_2$ and ω. Generally, the impact of aliasing on the effect estimate $\hat{\beta}_1$ can be quantified using the following formula:

$$E(\hat{\beta}_1) = \beta_1 + \omega_2\beta_2 + ...\omega_k\beta_k \quad (4.10)$$

This arises not only when A and B are perfectly aliased but also when the non-zero partially aliased effects are not included in the fitted model. In fact if all the non-zero effects that are partially aliased with A were included then the estimate of all the regression coefficients would be unbiased. Consider the L_{12} simulated experimental results displayed in Table 4-6 for which the true model is

$$Y_i = 2 + 3X_A + 5X_B + 2.6X_C + 7X_{AB} + 4X_{BC} \quad (4.11)$$

The aliasing array of the L_{12} (Table 4-9) shows that factors A and C were partially aliased with interactions BC and AB respectively with an aliasing coefficient of -0.333. As all of the aliased effects are included in equation 4.11, they are all correctly estimated.

A common, but misconceived view of the L_{12} is that it cannot be used to examine factor interactions. Suppose it was decided to only include the main effects in the fitted model, then it would be

$$Y_i = 2 + 1.67X_A + 5X_B + 0.27X_C \quad (4.12)$$

Evidently, the effect estimates of both A and C were biased by -1.33 (i.e -0.33*4) and -2.33 (i.e. -0.33*7) respectively i.e. the result of multiplying the aliasing coefficent by the size of the excluded non-zero effect. Because none of its aliases was included in the true model, the estimate of the factor B effect was not impacted. Thus the estimated effiects can be considerably distorted by aliasing.

The effect of aliasing on a regression coefficient's standard error $(MSE*C_{jj})^{1/2}$ can be demonstrated by assessing how it affects its design related element i.e. C_{jj}. According to Brownlee (1965):

$$C_{jj} = \frac{1}{\sum_{i}^{n}(x_{ij} - \bar{x}_j)^2 * (1 - R_j^2)} \qquad (4.13)$$

where x_{ij} is the i[th] entry of column j, $\bar{x}_j$ is the average of the x values in column j and R_j^2 is the coefficient of determination that results from regressing the X_j column on the remaining ones included in the model and measures the extent to which the effect X_j is aliased with the other effects in the model and can take any value between 0 and 1 inclusive. In orthogonal arrays, its value is zero; and the value of C_{jj} becomes $1/\Sigma(x_{ij}- \bar{x}_j)$. The larger the value of R_j^2 the larger the value of C_{jj} and consequently the larger the standard error. It was for this reason that Marquardt (1970) called the quantity $1/(1- R_j^2)$ the Variance Inflation Factor (VIF). Increasing the standard error, decreases the value of the t-statistic and can lead to pronouncing significant effects as inert. It must be borne in mind that by influencing the regression coefficients and their corresponding C_{jj}, aliasing affects the sum of squares of each regression coefficient. This is because the latter is the result of dividing the squared regression coefficient by its C_{jj} as shown in equation 16 of Appendix 2.

To shed further light on how aliasing affects C_{jj}, again consider the L_{12} simulated experimental results (Table 4-6). Under the true model (equation 4.11) the only effect that is orthogonal with the other terms in the model is B. Therefore, its C_{jj} = 1/12 = 0.0833. As the remaining effects are partially aliased with each other, their individual R_j^2 is 0.111. Consequently the VIF of each of these is 1.125. Therefore, the partial aliasing of each effect inflated the orthogonal C_{jj} value by 1.125 resulting in a C_{jj} value of 0.094. Interestingly, since the main effect model (equation 4.12) comprises orthogonal effects, their individual C_{jj} is 0.083. Particularly noticeable here is the inverse effect of aliasing on the regression coefficients and the C_{jj} values. When the true model is fitted, the estimated regression coefficients are unbiased but their C_{jj} are inflated whereas under the reduced model the regression coefficients are distorted but their C_{jj} are improved. In the case of over fitting the true model by including non-significant effects the regression coefficients will not be affected; however regarding the C_{jj} there are two outcomes. The first occurs if the added terms are orthogonal to those included in the model and thus exert no effect on their C_{jj} values. The second is when the added effects are partially aliased with some or all the terms of the true model. In this case, the appropriate R_j^2 will increase as the number of added terms increases thereby inflating the standard error of the corresponding regression coefficients. Such situations may only arise in the case of partial aliasing as it is not possible to fit the true model if any of its terms are perfectly aliased.

4.7 Summary

The interdependence between some or all the studied factorial effects is called aliasing. When small fractional designs are employed, two situations are commonly encountered. The first arises when subject knowledge suggests that while certain effects are important the others are inert. Tools such as linear graphs and aliasing tables are apposite here. The second situation occurs when all the studied effects are regarded as equally important and the aim is to select a fractional design so that the overall aliasing is minimised. To attain this, criteria such as maximum resolution, minimum aberration and generalised minimum aberration are postulated in the conventional DOE literature. A simple unified method that can be implemented using an Excel spreadsheet has been presented for constructing the aliasing pattern of regular and non-regular OAs. A discussion has also been included to show how, using the proposed method, the word length pattern can be obtained and used to select a design that satisfies any of the main aliasing measuring criteria. Furthermore, an explanation has been given about how aliasing can affect both the estimated factorial effects and their standard errors and in turn the t-statistics used to test each effect's statistical significance.

CHAPTER 5: ASSESSING THE PERFORMANCE OF CERTAIN DOE STRATIGIES USING SIMULATION

5.1 Introduction

The results of the review presented in Chapter 3 revealed a predominant use of unreplicated fractional factorial designs. This can be ascribed to time and cost limitations since full factorial experiments are likely to entail a prohibitive number of runs as the number of studied factors increases and replication, if not expensive, can be time consuming or even infeasible. Unreplicated responses can also be encountered in replicated experiments, when summary statistics such as a signal-to-noise ratio or a sample variance is calculated for each trial. In such cases an estimate of the experimental error requires certain factorial effects to be pooled together. Although there are many methods that can used to appropriately identify these effects (see Appendix 2) the results of the conducted review showed that (i) unaided subjective pooling methods were extensively relied upon (ii) practitioners did not appear to appreciate the need for and the importance of using formal pooling methods and (iii) they did not appear to be very well acquainted with which method to use under certain circumstances. In general, the use of unreplicated fractional factorial designs stimulates the investigation of several interesting aspects including

- their ability to identify the same important effects that would have been detected had a full factorial experiment been conducted,
- the performance of pooling methods i.e. whether they perform equally well in terms of identifying the "right" effects to be pooled,
- the extent to which the above two aspects are robust to such uncontrollable variables as the size and number of significant effects.

In the DOE literature, some of these aspects have been assessed in the context of two-level experiments. For example, Miller et al (1993) examined the first using data from a published full factorial experiment, while Hamada and Balakrishnan (1998) studied the performance of certain pooling methods using simulated data from an L_{16} OA. Only descriptive statistics were presented in both studies. The author could find no examples related to the third aspect. Moreover, to date no examples were found related to any of the aspects being investigated in the context of three-level experiments which have frequently been used in practice. Simulated multi-phased full factorial experiments were used in this research to study the three aspects simultaneously in the context of both two- and three-level trials using DOE.

As illustrated in Chapter 2, DOE is a powerful technique for decision making. However, its very implementation involves making decisions regarding, say, the selection of a design array and a pooling method and hence it can be exploited in making such decisions. From this perspective, design arrays such as the L_{16} and the L_8 can be viewed as levels of a controllable factor called the employed OAs. Similarly, methods such as the UC and Lenth can be regarded as levels of another controllable factor labelled pooling methods. The noise factors would be those over which the experimenter exerts no control such as the size and number of active effects. Besides enabling any possible interactions between these factors to be formally assessed, this approach allows the concept of robustness postulated by Taguchi and Wu (1985) to be utilised in identifying the settings of the controllable factors that make the detection of the “true” significant effects insensitive to the impact of the noise factors. The planning stage of the conducted experiments is firstly discussed, followed by a description of how they were performed. The Chapter culminates with a summary of the main findings.

5.2 Planning Stage

The planning stage involved stating the objectives, selecting the performance measure, the factors and their levels and choosing the experimental design.

The three main objectives of the simulation experiments were to:

- Study how the selected fractional design, the pooling method, and the number and size of the significant effects influenced the ability to detect the "true" active effects identified in the full factorial experiments.
- Identify the appropriate combination of fractional design and pooling method that would lead to maximising the detection of the "true" active effects whilst reducing the sensitivity to the noise factors i.e. the number and size of significant effects.

Given the objectives of the experiments, a continuous performance measure that would capture the quantities of interest was defined. This was the percentage of effects correctly detected as in the full factorial experiments for each combination of fractional design, pooling method, and number and size of active effects under study i.e. an estimate of the Power corresponding to each configuration of the studied factors.

5.3 Selection of Factors and Their Levels

The term "factor" is henceforth used to denote the controllable and the noise variables examined in the simulation experiments. Those associated with the levels of two of these, namely the number and sizes of active effects are referred to as "parameters". The simulation trials were conducted in four main phases. Each involved a sub-phase in which only the controllable factors were varied according to a full factorial array whilst random settings of the noise factors were used. At each phase, large full factorial experiments were simulated with a certain number of studied parameters of which some of specific sizes were designed to be active. The number and sizes of the specified active effects constituted the levels of two noise factors. Having generated the full factorial experiment, the appropriate runs for the examined fractional designs were extracted and analysed using the appropriate pooling methods (see Appendix 2). As already noted, the assessed fractional designs and pooling methods represent the levels of two controllable factors.

5.3.1 Phase-1

The scope of this phase focused on investigating the performance of two-level fractional designs and six pooling methods in cases where only main effects were significant. The number of studied parameters in the full factorial design was kept constant at a value specified utilising the results of the review discussed in Chapter 3. It was found that the number of studied factors was 4 or less in 60% of the cases and 6 or less in 85.5%. These results were found to be in line with that of Ilzarbe et al (2008) who reported that the number of studied factors was 5 or less in 71% of the general engineering DOE applications they reviewed. The number of studied factors in this phase was set at 7. The reason for not using 4 or 5 was that the size of their associated full factorial experiments ($2^4 = 16$ and $2^5 = 32$ runs) were not felt to be sufficiently large to demonstrate the impact of using fractional designs. In fact, the larger the difference between the full and the fractional factorial designs the clearer is the effect of using smaller designs. The two noise factors investigated in this phase were the number of active parameters out of the studied 7 and their sizes. In order to determine their levels, it was necessary to find an approximate estimate of the percentage of parameters that are likely to be declared significant in practice. To this end, Box and Meyer (1986) found, using the results from 10 published experiments, that the percentage was 0.2. This was estimated by calculating the average of the results of dividing the number of significant effects by the total number of estimable effects in the used OA for each of the examined experiments. One problem with this approach is that for the same number of studied and significant effects, different percentages may be observed depending on the size of the employed OA. As the scope of this phase precluded studying the interaction effects, it was felt that one way to deal with this problem was to estimate the percentage as the number of significant main effects over the number of studied parameters. This was calculated for each of the reviewed experiments and the results are shown in Figure 5-1. In 81% of the cases, the percentage of significant main effects was 75% or less and the percentage lying between 26% and 50% was the largest. On this basis, the selected numbers of significant effects were 3, 4 and 5 representing 42.9%, 57% and 71% respectively thereby accounting for a reasonable segment of the practical situations. The specification of the active parameters out of the studied ones was made randomly. With respect to the sizes of the significant effect, one approach could be to select their values subjectively as appeared to be the case in Hamada and Balakrishnan (1998). One problem with this is that there is no way to examine the extent to which the selected sizes reflect those that may be encountered in practice. Another approach

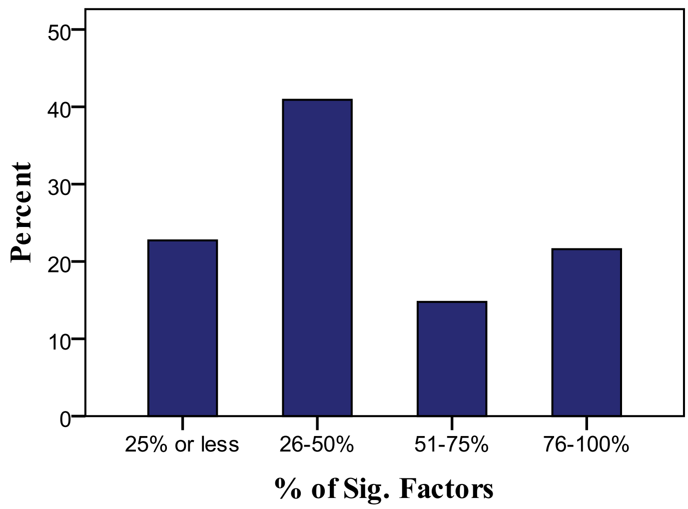


Figure 5-1: Fractions of the Significant Factors Percentage in the Surveyed Experiments

could be to estimate the effect sizes from the reviewed experiments (Chapter 3) but in this case a serious limitation was recognised. This related to the fact that the majority of trials were fractional and their aliasing pattern and its consequences regarding the estimated effect sizes were rarely examined. Generally, the reliability of the effect estimates depends a great deal on the degree of aliasing among the studied effects. Goh (2001) recommended that when reliable estimates of the factorial effects are needed, full factorial or the highest possible resolution fractional design should be used. Consequently, it was necessary to limit the effect size estimation to the full factorial experiments only. Unfortunately, it was found that besides being limited in their number, only a few of them incorporated the raw data or any effect estimates. Consequently, another review was conducted to find studies where full factorial experiments were employed and their data reported. The advanced search properties of the ScienceDirect® database were used to look for articles in which "full factorial" experiments were used during the period 1999-2009 (inclusive) in the same three journals as before, i.e. the JMPT, IJMTM, and CIRP. Of those found only 37 met the requirement that one or more full factorial data sets were reported. In fact, these incorporated 83 full factorial experiments, the results of which were used to estimate the effects of the studied factors and interactions as detailed in Appendix 4. As the trials involved two-and three-level experiments, all the studied effects in the latter case, be they factors or interactions, were decomposed into single degree of freedom effects. For example, factors at three levels were split up into linear and quadratic components using the Linear-Quadratic (L-Q) system (see Appendix 2). Rather than using a mean- or a variance-related effect measure, it was decided to use a standardised mean-related measure. This was arrived at by dividing the effect's regression coefficient by the square root of the mean square error, i.e. the mean-related effect was presented in multiples of σ, the error standard deviation. The advantages of this Standardised Effect Size (SES) are discussed in Chapter 8. After standardising the single degree of freedom effects it was realised that there was no observable difference between the sizes of the main effects and those associated with interactions. Consequently, these were grouped together to provide a practical insight into the sizes of the active factorial effects be they factors or interactions not only in this phase but also in subsequent ones and are shown in Figure 5-2. 83% of the effect sizes were between 0.51 and 3.5. It was infeasible to examine all the possible values within this interval since they are continuous in nature. Thus certain discrete values had to be selected. The first two were the mid-point and the maximum of the modal interval: 0.51-1.5. Two further values were the mid-points of the next two most frequently observed intervals, i.e. 2.51-3.5 and 1.51-2.5. The selected levels of the effect size were therefore 1, 1.5, 2, and 3 which should reflect a plausible range of effects that might be encountered in practice. The signs of these were randomly assigned.

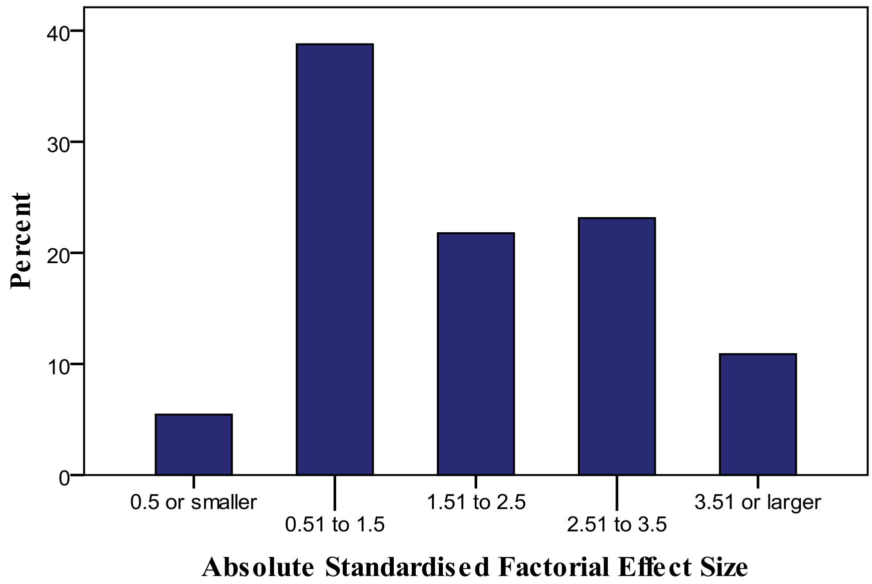


Figure 5-2: Percentages of Observed Absolute Standardised Factorial Effect Sizes

In terms of the fractional designs, the three most frequently used two-level OAs: the L_{16}, L_{12} and the L_8 were chosen as the levels of this factor. 3 and 4 generators were needed in the L_{16} and L_8 respectively to study 7 factors. In specifying these, the minimum aberration criterion (Chapter 4) was employed. There are many ways to choose both the columns and their signs to generate the required design. For example, the following two sets of generators:

$$E = \pm ABC,\ F = \pm BCD,\ G = \pm ACD$$

$$E = \pm ABC,\ F = \pm ABD,\ G = \pm ACD$$

render 16 equivalent L_{16} (2^{7-3}) minimum aberration designs. Examining all the possibilities can result in a prohibitively large experiment, so, the default choice of two of the frequently used statistical packages in practice, Minitab and Design-Expert, was adopted with the generators

$$E = +ABC,\ F = +BCD,\ G = +ACD$$

The selected generator resembled what Montgomery (2010) called the principal fraction i.e. the one with positive generators. The same concept was applied in choosing the L_8 (2^{7-4}) design and the generators:

$$D = +AB,\ E = +AC,\ F = +BC,\ G = +ABC$$

were used. For the L_{12} design, Sun (1996) showed that all the possible ways of choosing 7 out of the design's 11 columns lead to equivalent designs in terms of estimation capability, so again the default selection of the Minitab package was used.

Six of the pooling methods discussed in Appendix 2 were selected for examination. The first was the Half Normal Probability (HNP) plot which was seldom employed in practice despite its simplicity. Its use involves an element of subjectivity in deciding which effect is large enough to be pronounced significant. Thus its use was limited to specifying the effects that should be pooled to form an estimate of the error variance. This was then utilised in the ANOVA to formally test the significance of the other effects. A default pooling method in the Minitab package - Lenth Method (LM) - was used in some of the reviewed experiments.

Therefore, it was important to examine its performance along with that of its Modified version (MLM) proposed by Ye and Hamada (2000). In both cases the individual test's margin of error (see Appendix 2) was adopted, the reason being that in terms of Power, its use has better performance than that of the simultaneous margin of error. The review also showed that Taguchi analysis techniques were commonly employed. Among the popular pooling methods applied in this approach are the Pooling Up (PU), Pooling Down (PD) and Unassigned Columns (UC) methods. Ross (1996) and Roy (2001) presented a limited theoretical discussion about the performance of the first two; however the last one receives neither theoretical nor empirical assessment. Consequently, all these were examined in this study. Of course when 7 factors are examined using the L_8 it is infeasible to use the UC method, all the columns being accommodated. This was dealt with by regarding the columns associated with the factors that were designated to be inert as unassigned. Regarding the inflated residuals' estimates encountered in the initial stages of the PD method, an α value of 0.1 was used to alleviate this problem. The same α level was adopted in implementing the PU method to handle the t-statistic robustness problem associated with its first step (see Chapter 6). The factors examined in this first phase along with their levels are shown in Table 5-1.

Table 5-1: Phase-1 Factors and their Levels

Factor	**Description**	**Type**	**Levels**					
			1	**2**	**3**	**4**	**5**	**6**
A	Fractional OA	Controllable	L_{16}	L_{12}	L_8			
B	Pooling Methods	Controllable	HNP	LM	MLM	PU	PD	UC
C	No. of Active Effects	Noise	3	4	5			
D	Size of Active Effects	Noise	1	1.5	2	3		

One problem in robust designs is that some of the configurations of noise factors may not exactly mirror those encountered in practice. For example, despite the selection of magnitudes derived from actual applications, choosing active factors of the same sizes in each of the noise factor settings may represent rare situations. However, this is unavoidable here as choosing effects of different magnitudes would impair the ability to investigate the impact of the size of active effects which is a major objective of this phase. To circumvent this problem, a further sub-phase of experimentation was added in order to examine the impact of using different numbers and sizes of active effects on the way controllable factors affected the studied response. The numbers and sizes of active effects in six of the reviewed full factorial experiments (Appendix 4) where no interaction effect was detected, were firstly identified and are shown in Table 5-2. These were then used to simulate 6 full factorial experiments each with 7 factors. The appropriate runs for the examined OAs were extracted from each trial and analysed using the six pooling methods. Consequently, against each of the possible configurations of the controllable factors, 6 replications of the studied response were arrived at. The average and sample variance of each set of replications were subsequently calculated and analysed.

Table 5-2: Number and Sizes of Active Effects at Sub-Phase-1

Study	**No. of Active Main Effects**	**Effect Sizes**			
		1	**2**	**3**	**4**
Gunaraj and Murugan (1999)	3	1.7	2.4	-3.1	
Sun et al (2004)	3	-1.9	-0.9	0.6	
Dutta and Pratihar (2007)	4	1.3	0.6	-0.5	-1.9
Darwish (2000)	3	5.4	0.9	-1.2	
Raghukandan and Senthilvelan (2004)	3	2.6	4.1	0.7	
Kannan & Murugan (2006)	2	2.5	-3		

5.3.2 Phase-2

The controllable factors' impact on the studied response was assessed in this phase where both main and interaction effects were active in two-level experiments.

Again the number of studied factors was 7 and the same levels of the noise factors examined in phase-1 were adopted in the current one. From the results of the review detailed in Chapter 3, it was clear that the majority of the experiments ignored factor interactions. In the few that did examine them, only two-factor interactions were considered. This was also the case in the full factorial experiments' review (Appendix 4). It was also noted in both that a weak version of the heredity principal was valid, i.e. for each of the active two-factor interactions at least one of the main effects that comprised it was active. As already mentioned in phase-1, since the reviewed experiments involved both two- and three-level experiments all the observed interactions were decomposed into single degree of freedom effects. Irrespective of the experimental type, the observed numbers of active two-factor interactions were 1, 2 or 3 in 95.8% of the cases - see Figure 5-3. Consequently, these were the number of active two-factor interactions selected. After specifying the active main effects,

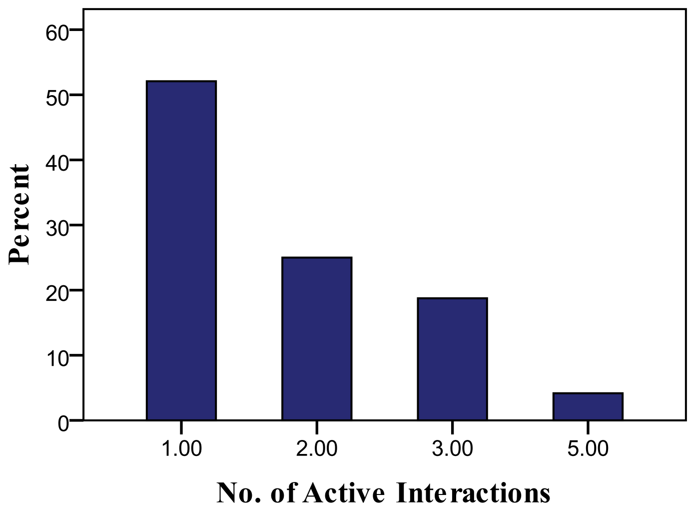


Figure 5-3: Numbers of Active Two-Factor Interactions in the Surveyed Experiments

these were determined randomly from all the possible two-factor interactions that satisfied the weak heredity principal, their signs were randomly assigned.

Owing to the substantial interdependence in choosing their levels, the consideration of the interactions in this phase impacted both the examined fractional OAs and the pooling methods. In fact, selecting any of the former dictates the feasible options of the latter. For example, the examination of two-factor interactions while maintaining orthogonality among the studied effects is feasible in the L_{16}. This is an essential condition for using pooling techniques such as HNP, LM and MLM which require the estimated effects to be uncorrelated. However, the situation is different for the L_{12} as every main effect is partially aliased with every two-factor interaction not involving it. In fact, regression analysis is needed to study interactions in this design where variable selection techniques such as Backward Elimination (BE) and Forward Selection (FS) act as pooling methods. One problem with the BE is that the full model incorporating all the main effects and two-factor interactions has to be fitted. For 7 factors, there are 21 ($^7C_2$) possible two-factor interactions, so 28 degrees of freedom are required. With only 11 degrees of freedom available in such a design the use of BE is infeasible. A better alternative to either the BE or the FS is a procedure that combines the use of both, i.e. Stepwise Regression (SWR). Another applicable technique is the Best Subset Selection (BSS). A description of these variable selection techniques is provided in Appendix 2.

A problem associated with the L_8 is that it has only 7 degrees of freedom. In one of the examined conditions, 5 active main effects and 3 two-factor interactions were simulated, thereby exceeding this number. It was therefore necessary to augment the L_8 designs with further runs to enable the estimation of these active effects. This is advantageous as it allows some of the augmentation strategies to be examined. The first decision regarding the augmentation related to the number of additional runs. This was set at 4 in order to make both the L_8 and L_{12} comparable in terms of the number of runs. Two augmenting methods were applied, the D-optimal (DO) and Box's Method (BM). The D in the former stands for determinate as it aims to augment the design so that the determinant $|\mathbf{X}^T\mathbf{X}|$ is maximised where $\mathbf{X}$ is the resultant model array. By so doing the variance of the least square estimates is minimised (Wu and Hamada, 2000). The DO augmentation process requires the true, or at least a suspected, model to be specified. In order to examine its performance, it was assumed that the true model was known and was used to generate the augmenting runs. Box's method of augmentation (Box et al, 2005) relies upon the principle of effect sparsity. For example, the first step for enabling the 7 two-level factors to be studied using the L_8 design was to generate the full factorial array and regard each of its entries as a candidate for being selected to augment the L_8. Then, utilising the original L_8 data, all the possible models comprising the studied 7 or all possible subsets of them and their 2-factor interactions that required no more than 7 degrees of freedom were fitted. Against each of the candidate runs (2^7=128) each of the derived models was used to predict a response value given the entries of the corresponding run. Consequently, a number of predicted responses equivalent to that of the derived models were generated and their standard deviation calculated. Of the obtained 128 standard deviation values the top largest four were selected as being powerful in distinguishing between the fitted models. Thus they were expected to be the most influential in identifying the true model. The DO and BM were implemented using the Design-Expert and the R-statistical software packages respectively. Once augmented and the two-factor interactions incorporated, the L_8 was no longer orthogonal as it involved partial aliasing. Therefore, variable selection techniques were again needed for analysing its data.

From the above discussion, it is clear that each OA requires a different pooling technique. Therefore, a single controllable factor combining the designs and their appropriate pooling method was used. The HNP method was employed with the L_{16} while SWR and BSS were used to analyse the L_{12} and the augmented L_8. In the case of SWR, α_{in} was taken to be 0.1 and α_{out} 0.1 to account for the inflated variance in the initial steps. For the BSS, the best fitting subset models of each possible size were produced. Choosing between models of different sizes can be difficult, so only the top fitted model (highest R^2_{adj}) of the same size as the "true" model was examined to assess whether it consisted of the same effects detected in the full factorial design. The factors investigated at this phase and their levels are shown in Table 5-3.

Table 5-3: Phase-2 Factors and their Levels

Factor	**Description**	**Type**	**Levels**						
			1	**2**	**3**	**4**	**5**	**6**	**7**
A	OAs & Pooling Methods	Controllable	L_{16}-HNP	L_{12}-BSS	L_{12}-SWR	L_8-BM-BSS	L_8-BM-SWR	L_8-DO-BSS	L_8-DO-SWR
B	No. of Active Main Effects	Noise	3	4	5				
C	No. of Active Two-Factor Interactions	Noise	1	2	3				
D	Size of Active Effects	Noise	1	1.5	2	3			

The issue of using effects of identical sizes was handled in the same manner as in sub-phase-1. Six two-level trials involving active main effects and two-factor interactions were selected from the reviewed full factorial experiments. Their number and size of active effects (Table 5-4) were used to examine the controllable factor performance under a variety of noise factor settings. Consequently, a single factor experiment with six replications was conducted in this sub-phase.

Table 5-4 Number and Sizes of Active Effects at Sub-Phase-2

No.	Study	Effects Number and Sizes								
		Main Effects						Two-Factor Interactions		
		1	2	3	4	5		1	2	3
1	Ganjigatti et al (2007)	-2.9	-2.3	0.6	0.4	-1.2		1	-0.4	0.5
2	Aggarwal et al (2008)	8.4	3.2	7.4	2.5			1.9		
3	Gunaraj and Murugan (1999)	2.7	3.4	-1.3	-1.1			-1.2	1.5	
4	Darwin et al (2008)	0.7	-3.2	-1.4	-0.5			0.4	-0.6	
5	Pei et al (2003)	-2.5	4.9	1.9				-2.3	-0.8	1.7
6	Kannan and Murugan (2006)	0.7	2.1	-1	1.6			-0.6	1.2	

5.3.3 Phase-3

Under the premise that only main effects were significant, the performance of the fractional three-level OAs and pooling methods was investigated in this phase. The analysis of three-level experiments is more time consuming than that for two-level. One reason is that steps such as the array coding and the effects pooling for two-level experiments can be implemented using any of the widely used menu-driven statistical packages. This is not the case for three-level trials where performing these steps entails writing routines in Excel or the R-software. As such it was necessary to keep the number of performed experiments in this and the following phases to a minimum. To accomplish this, one option was to reduce the number of studied factors. This was deemed inappropriate as it would have seriously affected the objectives of this study. Alternatively, though less effective, it was decided to keep the number of factor levels as small as possible bearing in mind the study's main purposes.

As was the case in the previous phases, the number of studied parameters was kept constant. Its value was specified as 4 -the number studied in the majority of the reviewed experiments discussed in Chapter 3. This rendered a large full factorial three-level trial of 84 runs. The number of active main effects was chosen to be 2 and 3 respectively representing

50% and 75% of the studied parameters (as shown in Figure 5-1, they belong to two of the percentage of active effects' intervals that collectively encompass 58% of observed values). Each three-level parameter was decomposed into two single degree of freedom components using the L-Q system, thus it was necessary to specify how many linear and quadratic components should be simulated to be significant when there are 2 and 3 active main effects. According to the results of Chapter 3 and the full factorial reviews, it was rarely the case that all the components comprising the active parameters were significant. In fact, almost 100% of the linear components of the active effects were found significant, whereas only 48% of the quadratic ones were declared active. Therefore, the presence of 2 active parameters was translated to mean 2 active linear components and 1 active quadratic. Similarly, for the 3 active parameters, 3 linear and 2 quadratic components were simulated to be significant. With regard to the sizes of the active effects, the same levels used in the previous phases were adopted as they were determined from the standardised single degree of freedom effects.

The fractional OA, the ANOVA type and the pooling methods are the three controllable factors that were investigated in this phase. With regard to the first, the most frequently used three-level OAs: the L_9 and L_{18} were examined. In the case of the L_9, all the columns were used to accommodate the 4 factors under study. For the L_{18}, the minimum generalised aberration criterion was used to select the four columns based on the L-Q coding system. The review in Chapter 3 revealed that two types of ANOVA, standard and regression, were employed in practice. Although, these are equivalent in the case of two-level experiments, this is not the case for three-levels. Both types of ANOVA were examined in this phase to assess their impact on the studied response. In the case of using the standard ANOVA, the separated OAs' columns were re-united so that each main effect was associated with two degrees of freedom. Since the PU and PD methods were suitable for dealing with both types of ANOVA, they were examined in this phase. In the context of regression ANOVA, these are equivalent to the BE and FS techniques. In implementing them, the same values of α adopted in phase 1 were used. The factors assessed in this phase and their selected levels are shown in Table 5-5.

Table 5-5: Phase-3 Factors and their Levels

Factor	**Description**	**Type**	**Levels**			
			1	**2**	**3**	**4**
A	Fractional OAs	Controllable	L_9	L_{18}		
B	ANOVA Type	Controllable	Standard ANOVA	Regression ANOVA		
C	Pooling Methods	Controllable	PU	PD		
D	No. of Active Main Effects	Noise	2	3		
E	Size of Active Effects	Noise	1	1.5	2	3

A further experiment was conducted to assess the impact of this phase's controllable factors on the studied response under six cases where the settings of the noise factors were specified using actual published data. This allowed for the consequence of using active effects of different numbers and sizes to be investigated. The six combinations are shown in Table 5-6.

Table 5-6: Number and Sizes of Active Effects at Sub-Phase-3

Study	**Active Main Effect Sizes**					
	Linear				**Quadratic**	
	1	**2**	**3**		**1**	**2**
Correia and Ferraresi (2007)	-3.4				-1.8	
Chattopadhyay et al (2009)	-11.2	-1.4	1.2			
Gaitonde et al (2008)	-0.8	1.4	1.4		-0.7	
Dhar et al (2007)	4.3	0.8			-1.9	
Dhar et al (2007)	1.5	1.5	2.7		-1.4	
Davim (2003)	1.2	1.4	1.1		1.1	

5.3.4 Phase-4

This phase involved studying the effect of the three-level fractional OAs and pooling methods on the studied response in the presence of active main effects as well as two-factor interactions. The number of studied parameters and of active main effects selected in the previous phase were also used in this one, as were the values of the sizes of the active effects. Again, the L_9 and L_{18} were examined. In order to study two-factor interactions, partial aliasing was involved and hence effects' pooling was performed using variable selection methods. Techniques such as BE were not applicable in the case of the L_9 since the required degrees of freedom for fitting the full model exceeded those available. Hence, SWR and BSS were employed.

Regarding the number of active two-factor interactions, the two most frequently observed values of 1 and 2 were used. As shown in Figure 5-3, these account for 79% of the encountered number of active interactions in the conducted reviews. Using the L-Q coding system, there were four components associated with each three-level two-factor interaction; namely the LxL, LxQ, QxL and the QxQ. Of the interactions investigated in the reviewed experiments involving three-level factors, these represented 71.5%, 17.8%, 10.7% and 0% respectively. The LxL was the most encountered whereas the QxQ was almost absent. That the LxL is more likely to happen than the others in practice follows from the fact that it is the only interaction component considered in second order response surface models (Box and Draper, 2007). In light of this, it was decided to only use the LxL component in cases where only one interaction was active; however, where two were significant, one was selected to be LxL, and the other was randomly specified as either LxQ or QxL. The factors examined in this phase along with their levels are shown in Table 5-7.

Table 5-7: Phase-4 Factors and their Levels

Factor	Description	Type	Levels			
			1	2	3	4
A	OAs	Controllable	L_9	L_{18}		
B	Pooling Methods	Controllable	SWR	BSS		
C	No. of Active Main Effects	Noise	2	3		
D	No. of Active Interactions	Noise	1	2		
E	Size of Active Effects	Noise	1	1.5	2	3

As in the previous cases, the impact of using different combinations of noise factor levels was assessed in sub-phase 4. Six of the reviewed full factorial three-level experiments involving active main effects and two-factor interactions were used to select the settings of the number and sizes of active effects as shown in Table 5-8.

Table 5-8: Number and Sizes of Active Effects at Sub-Phase-4

Study	Main Effects						Two-Factor Interactions				
	Linear			Quadratic			LxL		LxQ		QxL
	1	2	3	1	2		1	2	1	2	1
Correia and Ferraresi (2007)	-1.8	3.7		1.6			0.9				
Dhar et al (2007)	-1.2	6.3	0.7	-0.8	11		-0.7				
Ghani et al (2004)	0.8	2.4	-1	0.9			-1.5		2.6		-3.6
Davim et al (2008)	-2.9	-0.6		1.4					2.6		
Davim (2000)	0.7	6	1.9	0.7	4.7		0.6	1.9			
Chattopadhyay et al (2009)	12.6	-2.1	-2.6				-1				

5.3.5 Selection of Experimental Design

Having specified the factors under study and their levels, the next step was to select the experimental designs, i.e. the arrays according to which the simulation experiments were performed. Conventionally, this should be based on the required effects to be estimated and their respective degrees of freedom. However, in the absence of prior knowledge regarding the likelihood of the significance of certain effects, all the possible ones, be they factors or interactions, were deemed equally important. Consequently, it was decided to conduct full factorial simulation experiments for the factors stated in the previous sections. Broadly, robust experiments have two execution arrangements, each entailing a certain analysis approach. One is the cross arrays proposed by Taguchi (1986) wherein two separate arrays are generated for controllable and noise factors, referred to as the Control Array (CA) and Noise Array (NA). If there are n_1 and n_2 runs associated respectively with the CA and NA, then the run size of the cross array is n_1 x n_2. Denoting the response for the combination of the i^{th} control setting and the j^{th} noise setting by y_{ij}, for each control setting i, there are n_2 responses across the NA. These can be regarded as the noise replicates and their sample mean and variance can each be modeled as a function of the controllable factors. The objective is to determine the controllable factor settings so that the targeted mean value is attained whilst minimising the variance. The second approach is to use a single array to accommodate both the controllable and noise factors (Welch et al, 1990; Shoemaker et al, 1991), the idea being to model the response y as a function of both the controllable and noise factors and to explicitly investigate the existence of interaction between them. These can be exploited to achieve robustness. For example, consider the two-factor interaction between the controllable factor x and the noise factor z in Figure 5-4. Clearly, when there is no interaction between x and z (Figure 5-4 (a)), the effect of z and hence the variability it causes in the response variable Y is the same at the two levels of x. Thus selecting either level will not affect the variability in Y. On the other hand, when there is a significant interaction between x and z as shown in Figure 5-4 (b) the effect of the noise factor z is very small (as measured by the slope of the line) when x is at the high level (x =1). Consequently, the variability transmitted to the response by that in z is reduced by setting x at the high level. Generally, unless there is at least one considerable

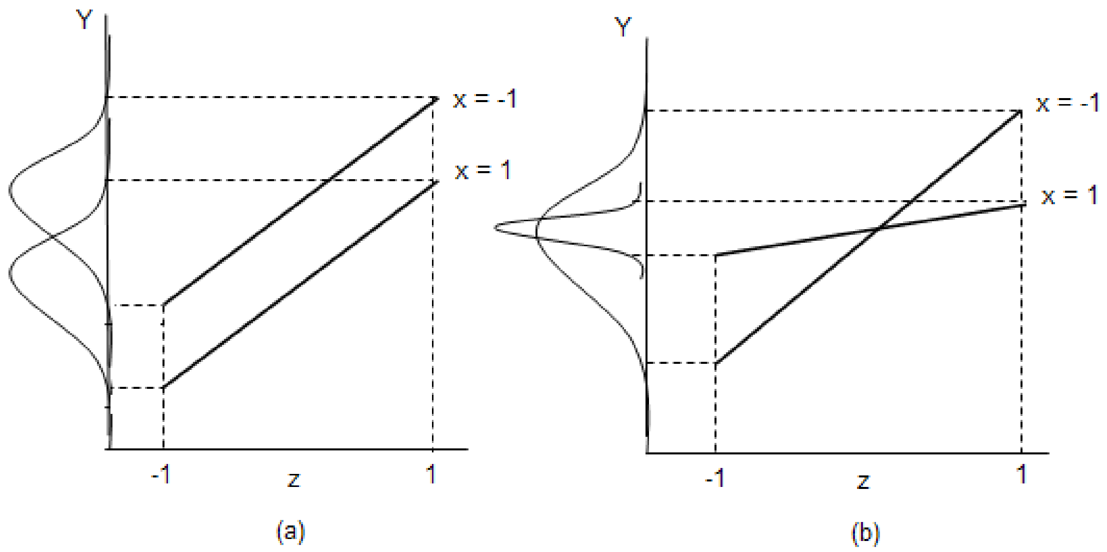


Figure 5-4 The Role of Controllable-Noise Interaction in Robust Design (adapted from Montgomery (2010))

interaction between control and noise factors, there is no robust design problem. Though implicit, this is also the case in the cross array approach. Both the single and cross array arrangements were used in the simulation experiments. In the cases where the number and sizes of active effects were treated as factors at specific levels, the single array design and analysis approach was adopted. This is because a better process understanding is gained from the explicit examination of the interaction between the controllable and noise factors in this approach. The cross array approach was employed in the experiments where six different sets of number and sizes of active effects were used with each of the possible configurations of the controllable factors under study. In these, the selected sets of the number and sizes of active effects were regarded as six levels of a noise factor studied using a NA, whereas all the possible combinations of the studied controllable factors resembled the CA. Consequently, the CA's row sample means and variances were studied.

Considering the large number of trials required and the available time, the maximum feasible number of replications was set at 50. Therefore, each of the recorded responses of the Power represents the average of 50 replicated trials.

5.4 Simulation Experiments

The simulation experiments were conducted utilising a Monte Carlo approach. This is a form of static simulation where the state of the simulated process is independent of time. It relies on generating observations from a certain statistical probability distribution and using them in such a way that the random process under study is directly simulated and its observed behaviour is used to infer the desired solution (Winston, 2003; Kalos and Whitlock, 2008). The simulation procedures can be summarised as follows:

1. Fifty random samples were generated from a Normal distribution with mean zero and variance 4 i.e. N (0, 4) using the Minitab package. Their values were plotted on Normal probability graphs using the same package to ensure their compliance with the requirements. In the case of the two-level experiments the size of each sample was 128 whereas it was 81 in the three-level trials. Each sample was used as the stochastic error component of a simulated response and its individual values were denoted by e_i.
2. The full factorial orthogonal arrays were constructed in Excel. For two-level trials, the size of the array was 128 (2^7) and its entries were coded using the -1 , +1 system. In the case of the three-level experiments, the array size was 81 (3^4) and its entries were coded using the L-Q system (Appendix 2).
3. A regression model was then used to generate the deterministic components of each of the simulated responses where the dependent variable was the deterministic response Y_d and the independent variables were the selected active effects (factors and interactions). Their possible values were the entries of their columns in the full factorial orthogonal array. The coefficients of the independent variables were chosen to attain the required sizes of the effects in terms of the number of error standard deviations. For example, in order to simulate the Y_d in the case where two active two-level factors A and B were both

of size, say, 2σ, the effects should both be 4. This is because the error σ is 2 and 4/2 equals 2. Thus, the initial regression model should be

$$Y_{di} = 2X_{Ai}+2X_{Bi} \qquad (4.1)$$

where i =1, 2, 3.... 128. The simulated response can be arrived at by adding the error random components (step1) to the deterministic components generated using equation 4.1 i.e. $Y_i = Y_{di} + e_i$. The three-level effects can be simulated in the same manner bearing in mind the effects' interpretation difference between the two- and three-level effects (see Appendix 2).

4. For each simulated response in the full factorial designs, the subset of runs that corresponded to the two- and three-level fractional designs under study were extracted. The appropriate pooling methods as explained in each experimental phase were then used to identify the effects that could be flagged as significant.

Against each of the possible configurations of the studied controllable and noise factors, fifty simulated responses were generated. For each, the percentage of the effects correctly detected (the Power) as in the appropriate full factorial design was firstly computed. Then the average Power over the fifty replications was estimated.

5.5 Summary

Of principal importance when conducting an unreplicated fractional factorial experiment is the aspect relating to its capability of detecting the same active effects that would have been identified had a full factorial experiment been conducted. Also important is the extent to which this is impacted by the choice of pooling method. In this Chapter, the planning stage was presented for robust simulation experiments that were conducted to investigate these aspects. In these trials, the selected fractional OAs along with the pooling methods were regarded as controllable factors whereas the number of active effects and their sizes were handled as noise factors. The studied response was the Power and the robust DOE approach was adopted. The major reason for this was to allow the possibility of finding certain settings of the controllable factors that attain ruggedness of the Power against the examined noise factors to be investigated; an aspect that, to date, has not been addressed in the literature. In specifying the levels of the studied controllable and noise factors the results of the review detailed in Chapter 3 were utilised. A further review concerning published full factorial experimental data was conducted in this Chapter. The objective of this was to provide reliable estimates for the active effect sizes which were then used to determine the levels of the corresponding noise factor. The simulated experiments were full factorial types with both single and cross array arrangements for performing robust experiments. Cross arrays were adopted in the cases where the number and sizes of active effects were specified on the basis of published experiments that matched the required properties in terms of the type of studied and active effects.

Having explained the planning and execution procedures for the simulated experiments, their results are presented and discussed in the next Chapter.

CHAPTER 6: RESULTS AND DISCUSSION OF THE SIMULATION EXPERIMENTS

6.1 Introduction

In this Chapter, the results of the simulation experiments, detailed in the previous Chapter, are presented, analysed and discussed. The method adopted in analysing the results is firstly outlined. Then, the results of each experimental phase are addressed and dealt with separately. The Chapter culminates with a summary of the main findings.

6.2 Analysing Method

DOE techniques were generally used to analyse the data, however aliasing was not examined, since the experiments used full factorial designs. For each of the studied factors, the hypothesis of no effect was tested using the ANOVA. As the experiments were unreplicated, the interactions of order three or higher were used to estimate the experimental error so that the statistical significance of all the studied main effects and two-factor interactions could be investigated. Of the latter, the ones that incorporated controllable and noise factors were of particular interest since they could possibly be used to attain robustness. Another reason for using high order interactions to estimate the experimental error was that they rarely exhibit effects that are distinguishable from noise (Montgomery, 2010; Box et al, 2005).

Having performed the ANOVA significance tests, the inert main effects and two-factor interactions were pooled with the initial estimate of the error variance. To assess the importance of the studied effects, their variance PCs were estimated. For reasons which will be explained in Chapter 8, the PC measure of equation 3.2 was adopted. Besides quantifying the effects' importance, this measure provides an objective means for examining both the aptness of selected effects for pooling and also the adequacy of the active effects for representing the process under study. These are judged by examining the PC associated with the experimental error. If this turns to be small (15% or less), the selected active effects may be regarded as adequate. Otherwise, further investigation is necessary on the studied effects, the measurement error and the effect sizes (see Chapter 8).

The assumptions of the ANOVA were examined by plotting the residuals against both the estimated response values and the order of performing the experiments, and by plotting them on a Normal probability plot. Whenever a violation was suspected, the Box-Cox plot (Box and Cox, 1964) was used to assess the suitability of the power family of transformation for rendering satisfactory plots of the residuals. The significant main effects and two-factor interactions were plotted in order to visually assess how they affected the responses. The Minitab and Design-Expert statistical packages were used for analysing the data. The results of the performed simulation experiments are presented in Appendix 5.

6.3 Phase-1

The ANOVA and the PC of the effects (listed in Table 5-1) that had statistically significant impacts on the Power of detecting the correct active factors at an α level of 0.05 are presented in Table 6-1. On the basis of the estimated PCs, the most influential factor was the pooling methods (B) followed by the size of active effects (D) and the employed OA (A). The smallest significant effects were the number of active effects (C) and its interaction with the employed OA. Despite its small explained variance (1.17%) the latter is very instructive in terms of casting some light on how robustness can be attained. This controllable-noise factors' interaction (Figure 6-1) shows that the L_8 was not only associated with the smallest Power but also with the highest sensitivity to the number of active effects - evident from the fact that its response line has a steeper slope than those associated with the L_{12} and the L_{16}. Thus the variance transmitted to the Power due to the variability in the number of active

Table 6-1: Phase-1 ANOVA for Power

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value	PC (%)
A	1.54	2	0.77037	7358.85	< 0.0001	18.13
B	3.56	5	0.71295	6810.35	< 0.0001	41.92
C	0.17	2	0.08474	809.44	< 0.0001	2
D	3.10	3	1.03248	9862.71	< 0.0001	36.51
AC	0.10	4	0.02443	233.37	< 0.0001	1.17
Residual	0.020832	199	0.00010			
Total	8.490943	215				

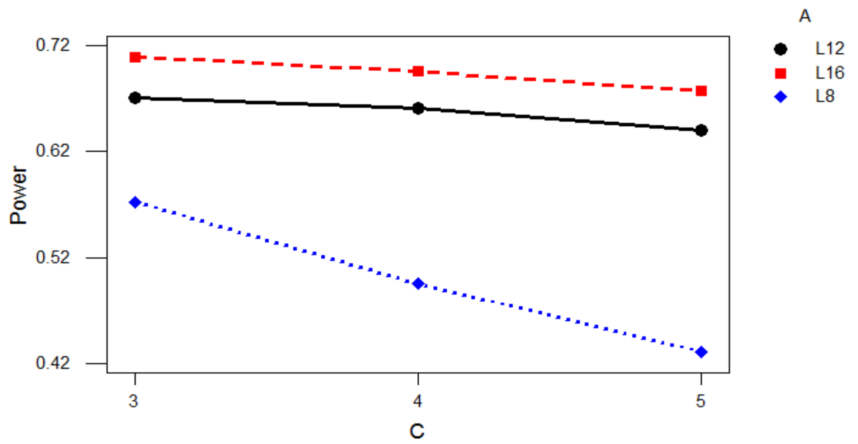


Figure 6-1: Phase-1 AC interaction Plot for Power

effects, over which no control can be exercised in practice, may be minimised by using the L_{16} or the L_{12}. The error PC of 0.27 % indicates that the considered effects provide a very good representation of the response. The main effects plot of the statistically significant factors (shown in Figure 6-2) shows that the highest Power was attained when the L_{16} was analysed using either the PU or the HNP plot in the cases where the number of active effects was 3 and the size of each was 3σ. It also shows that the larger the size of the employed OA

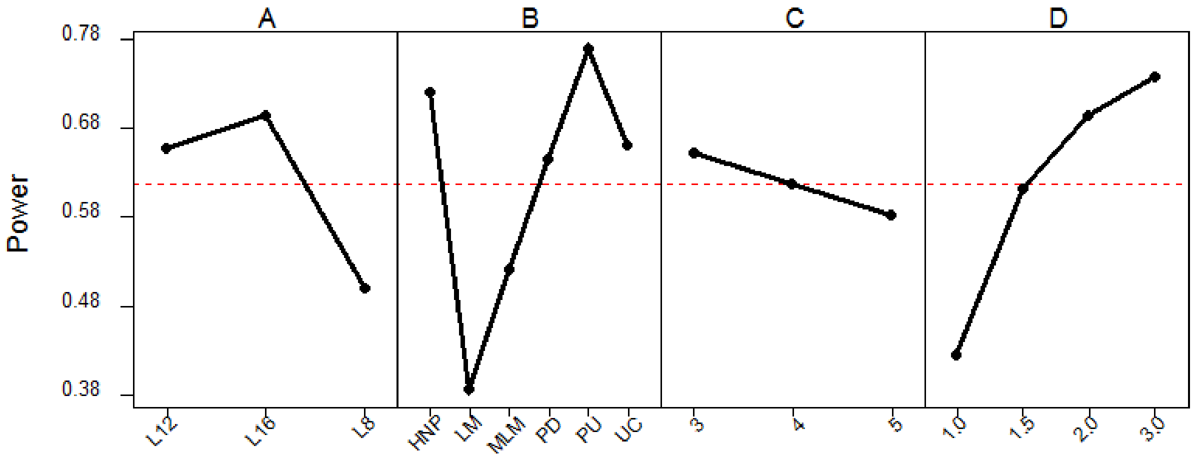


Figure 6-2: Phase-1 Main Effects Plot for Power

the higher the Power which decreases as the number of active effects increases while it increases as their sizes increase.

To provide a theoretical justification for these results, consider the t-distribution used in testing the statistical significance of the simulated parameters. Under the null hypothesis of no effect, the distribution is central and its degrees of freedom are specified based on those available to estimate the error. However, when an effect is significant with a certain SES, the reference distribution is a non-central t-distribution which has, in addition to degrees of freedom, another characterising parameter called the non-centrality parameter (Lynch, 1993; Oehlert and Whitcomb, 2001), defined by

$$\delta = \beta_i / \sqrt{MSE \times C_{ij}} \qquad (6.1)$$

where β_i is the estimated regression coefficient as defined in Appendix 2, the MSE is the error mean square and C_{ij} is as defined in equation (4.13). In the absence of aliasing, as was the case in this phase, the C_{ij} of any of the studied two-level OA is reduced to (1/n). Consequently, δ can be rewritten as SES x $\sqrt{n}$. Clearly, δ increases as either or both of the SES and $\sqrt{n}$ increase. Generally, the larger the δ, the higher the Power. Figure 6-3 shows a

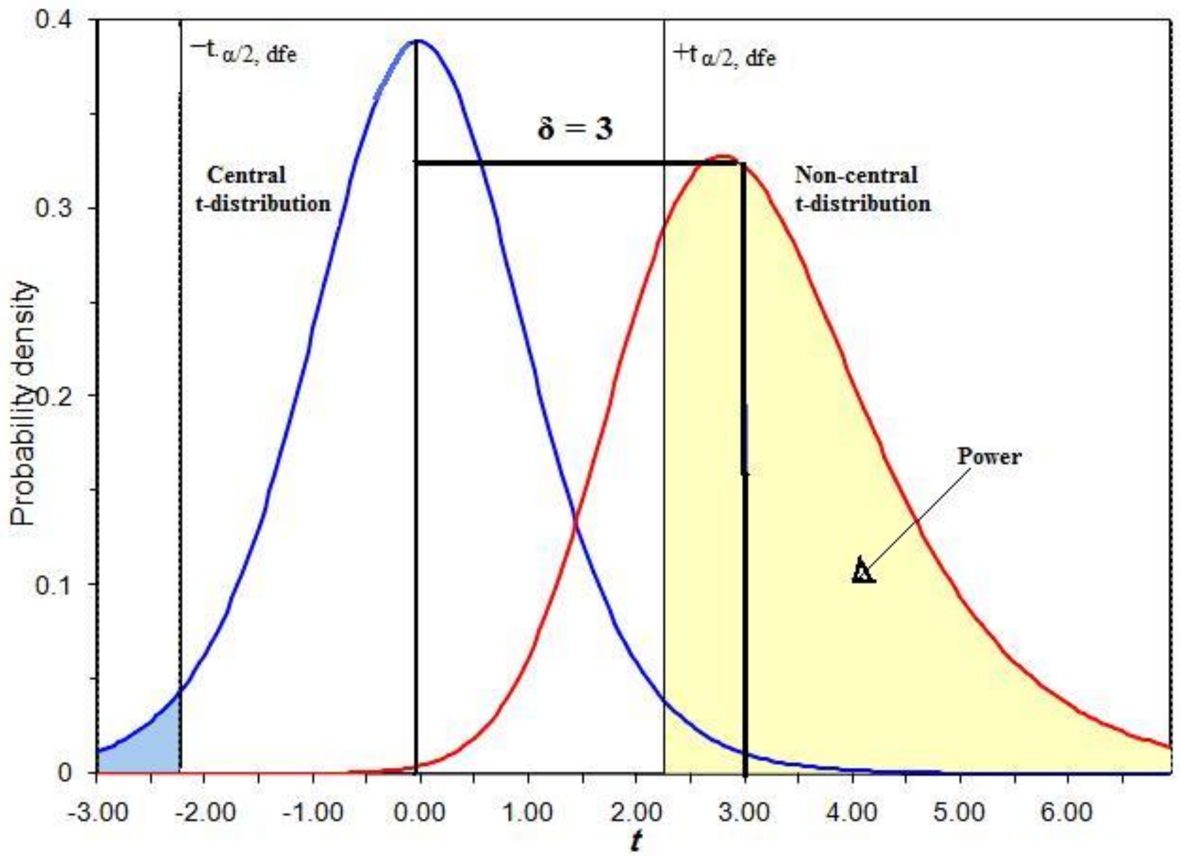


Figure 6-3: Graphical Representation of Power under a Non-central t-distribution with δ =3, df_e =10

graphical representation of the Power for a non-central t distribution with δ = 3 and the error degrees of freedom (df_e) =10. For the same SES and tabulated $+t_{\alpha/2,\ dfe}$ value, the larger the size of the OA the larger the δ and consequently the larger the Power. For a given value of δ, the larger the tabulated $t_{\alpha/2,dfe}$ value the smaller the Power. On the other hand, for a fixed α level the values of $t_{\alpha/2,\ dfe}$ relate inversely to the df_e . For small values of the latter, the rate of change in the $t_{\alpha/2,dfe}$ value is considerably high. For example consider the relationship between the tabulated t-values and the df_e when α =0.05 portrayed graphically in Figure 6-4. As can be seen, the largest t-value (6.314) is associated with a df_e of 1 but this value reduces to 2.92 and 2.353 when the df_e increases to 2 and 3 respectively. Clearly, by increasing the df_e from 1 to 2, the reduction in the t-values markedly exceeded that which was observed when the df_e

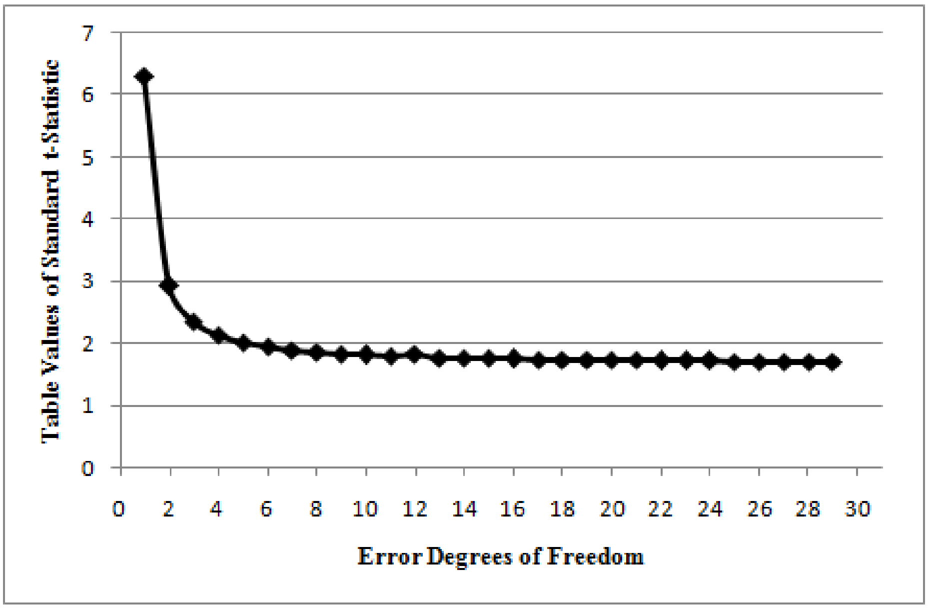


Figure 6-4: Standard t-statistic Values vs. Error Degrees of Freedom at α =0.05

was increased from 2 to 3. Moreover, the values of the t-statistic associated with a df_e of 6 or more exhibit a much smaller rate of change than those associated with smaller df_es. From this, one can infer that for the t-test to be robust, a df_e of 6 or more is required. Furthermore, the t-test reliability is seriously low when the df_e is 1. The value of $t_{\alpha/2,dfe}$ depends on the number of active effects; in fact, the larger the latter the fewer the available df_e and subsequently the larger its value. Consequently, the Power is reduced as the number of active effects increases. Note that due to their equivalence (Appendix 2), the discussion related to the Power of the $t_{\alpha/2,dfe}$ statistic is also applicable to the $F_{\alpha,1,dfe}$ statistic. Generally, any results related to Power can be explained in terms of the components that form δ as well as the level of α and the df_e. However, the complexity arises from the interdependence between these terms. For example, consider the situation where 5 active factors of size 2σ were simulated. Although, the generated full factorial data matched the simulated conditions, this was not the case for the extracted fractional design data. The rendered β_i and MSE values and their consequent SESs were, more often than not, noticeably different from the intended ones. From this it can be inferred that, using a small fractional design, it may be difficult if not impossible to obtain reliable estimates of the factorial effects under study. This can be used to explain Goh's (2001) recommendation regarding the necessity of using a full factorial experiment whenever credible estimates of the factorial effects are needed. In fact, it is not only aliasing that can impact the reliability of the estimated SES but also the reduction in the number of runs along with the reliance on a single replicate of the experiment. It is due to the joint effect of these, that a distorted effects' estimate may be observed in small experiments despite the absence of aliasing. The difference between the full factorial SES estimates and those of the fractional OA should not be surprising in the light of the expected difference in the MSE estimates rendered by the two designs. Due to the larger number of degrees of freedom available to estimate the MSE, its full factorial design estimate is more reliable than that of the fractional one. For the latter, the matter becomes more complicated when considering that its small df_e may not be correctly specified depending on the employed pooling method. It was with regard to such interrelations that the use of simulation experiments was deemed necessary for empirically estimating the Power associated with the investigated DOE strategies. The impact of reducing the experiment's size becomes clearer when considering the pooling methods' effect. As the MSE is estimated by pooling the effects that are intended to be inert in the simulated full factorial data, the UC method should theoretically be associated with the highest Power. However, as can be seen from Figure 6-2, both the HNP plot and the PU method outperformed it. This is due to the fact that as a result of reducing the size of the simulated experiments, some of the full factorial inert effects became large and inflicted a bias on the error estimate that was formed by pooling them using the UC method. Consequently, the inflated MSE reduced the test statistics associated with the active effects and led to declaring some of them as inert. This problem was avoided in the cases where the HNP plot was employed. The reason being that all the effects including those associated with the unassigned columns were plotted and the error estimate was arrived at by pooling the small contrasts. Thus any overestimated effect, as a result of using fractional designs, was not pooled with the error. Similarly, the PU method determined the effects to be pooled based on their observed sizes irrespective of whether their columns were assigned an effect. Although, this may result in increasing the Type I error, this was considered to be far less serious than decreasing the Power.

As can be seen from Figure 6-2, the HNP plot and PU method attained the highest Power with the latter being marginally more powerful than the former. This should not be surprising since a high α value of 0.1 was used with the PU method. Apparently, the use of a value greater than 0.05 with this method (Hines et al, 2003) relates to its first step where the smallest effect is used to estimate the error rendering a df_e of 1. As already explained, the t-statistic is not robust in this case, so if the first tested effect happens to be active but its size is not large enough to portray this problem, it would be pooled with the inert effects yielding a biased estimate of the error. The impact of this problem can be reduced by increasing the α level. For example, while the value of $t_{0.025,1}$ is 12.701 when α =0.05, this is reduced to 6.314 approximately half the size for α =0.1. In practice, however, it is very unlikely that all the effects except the smallest would be active. Thus, adopting such a strategy could result in an increased Type I error.

The UC method exhibited a better performance than that of the PD which in turn attained higher Power than the two versions of Lenth's method (see Figure 6-2). The main problem associated with the use of the PD method relates to its initial steps. As discussed in Appendix 2, this method starts with pooling all the studied effects, except the largest, to yield the initial MSE. Clearly, if there are, say, 4 active effects, three will be pooled with the MSE rendering an overestimated error variance. The degree of bias associated with the obtained MSE increases as both the number and size of the active effects increase. For the largest effect to be declared significant in the first step of the PD method, it needs to be large enough so that it is not camouflaged by the experimental error and the combined impact of the pooled active effects. An aspect that should be borne in mind here is that in this phase the PD method is equivalent to both the FS procedure and the SWR method. With regard to the latter, the absence of aliasing among the simulated effects in this phase prevents the selected effect at each step from being eliminated from the model in the subsequent ones. Consequently both the FS and SWR would be expected to exhibit the same performance in this phase as that of the PD method.

As shown in Figure 6-2, the lowest Power was observed with the two versions of Lenth's method. Perhaps the prime reason for this relates to the way in which the effects' "pseudo" standard error (PSE) is estimated. As explained in Appendix 2, the first step in implementing this method is to find the median of the estimated effects. This is then multiplied by 1.5 to yield S_o which should in turn be multiplied by 2.5. Those effects smaller than the resultant number are determined and their median is computed and multiplied by 1.5 to render an estimate of the effects' PSE. Unless the number of inert effects is more than half the studied ones, a large PSE is likely to be observed rendering small non-significant effects' t-statistics. This is best illustrated by analysing the extracted L_8 Response 1 data displayed in Table 6-2 (a). These were taken from a full factorial simulated experiment where the parameters A, B, C, D and E were all significant and of equal size. Arranged in ascending order, the L_8 estimated main effects of the parameters F, G, E, C, B, D and A are respectively 0.59, 0.75, 1.085, 5.664, 6.412, 7.724, 8.85. Clearly their median and S_0 are 5.664 and 8.496 (1.5x5.664) respectively. All the estimated effects are smaller than $2.5xS_0$. Consequently, the PSE = S_0 = 8.496 is marginally smaller than the largest estimated effect. The t-statistic for each effect is

Table 6-2: Sample Simulated Data for the (a) L_8 and (b) L_{12}

No.	A	B	C	D	E	F	G	Resp. 1	Resp. 2
1	-1	-1	-1	1	1	1	-1	-6.097	-4.59
2	1	-1	-1	-1	-1	1	1	-5.313	-0.813
3	-1	1	-1	-1	1	-1	1	-7.252	0.248
4	1	1	-1	1	-1	-1	-1	7.489	5.989
5	-1	-1	1	1	-1	-1	1	-1.361	-5.861
6	1	-1	1	-1	1	-1	-1	0.102	-1.398
7	-1	1	1	-1	-1	1	-1	-2.830	-1.33
8	1	1	1	1	1	1	1	15.571	13.1

(a)

No.	A	B	C	D	E	F	G	Resp. 1	Resp. 2
1	1	-1	1	-1	-1	-1	1	1.129	-5.076
2	1	1	-1	1	-1	-1	-1	-0.726	-0.900
3	-1	1	1	-1	1	-1	-1	4.800	-1.100
4	1	-1	1	1	-1	1	-1	2.966	6.076
5	1	1	-1	1	1	-1	1	6.008	-0.900
6	1	1	1	-1	1	1	-1	5.855	4.864
7	-1	1	1	1	-1	1	1	1.870	10.273
8	-1	-1	1	1	1	-1	1	4.976	0.100
9	-1	-1	-1	1	1	1	-1	-1.323	-4.918
10	1	-1	-1	-1	1	1	1	-2.622	1.900
11	-1	1	-1	-1	-1	1	1	-3.655	-6.111
12	-1	-1	-1	-1	-1	-1	-1	-11.162	-4.072

(b)

obtained by dividing its estimate by the PSE. Thus the largest possible t-statistic is 1.04 (8.85/8.498). The smallest tabulated t value for $\alpha/2 = 0.025$ is 1.96 rendering all the examined effects non-significant. Nothing was gained by adopting the t-Lenth critical values proposed by Ye and Hamada (2000) in the modified version of LM (see page 245). In fact, for $\alpha = 0.05$ and 7 contrasts, the proposed t-Lenth is 2.297 which is larger than 1.04 implying that all the studied effects are inert. In fact the only reason why the modified version of Lenth method outperformed the standard one is that for any given value of α and a number of contrasts, the critical t-values associated with the former are smaller than those associated with the latter. The implementation of LM showed that its success in detecting the active effects depends a great deal on the value of the median effect. If this is large, then both the S_0 and PSE are likely to be large. Of course the larger the PSE the smaller the chance of declaring any effect as active. As the median value separates the upper half of the estimated effects from the lower one, it follows that to observe a small median the percentage of inert effects needs to exceed 50% of the examined effects. In practice, it is difficult to have prior knowledge regarding the proportion of inert effects. More importantly, such a percentage can result in a remarkably different number of inert effects depending on the size of the OA that is used. For example, assuming that the percentage of inert effects is 50%, if an L_8 is used, then 3 or 4 effects are expected to be inactive. However, this implies that 7 to 8 effects are inert in the case of using the L_{16}.

In their simulation study, Hamada and Balakrishnan (1998) reported a better performance than that observed in this investigation. One reason for this is that the L_{16} was the only OA investigated in their study. More importantly, the numbers of simulated inert effects were 14, 13, 11 and 9. Consequently, their percentages of the available 15 degrees of freedom were respectively 93.33, 86.7, 73.33 and 60%. As already discussed, under such circumstances Lenth's method is likely to show an acceptable performance. However, in practice it is better to use a method that is robust to the unknown and uncontrollable percentage of inert effects. Therefore, unless used in conjunction with other pooling techniques, the use of LM is not recommended. In this context, it is noteworthy to recall that as revealed from the results of the review discussed in Chapter 3, practitioners apply LM unintentionally. This is because of the use of the Minitab statistical package in which the Normal Probability (NP) plot is used as a graphical means for conveying the results of employing LM. In fact, the users of this package have no control over which of the NP plot points to pick as potentially large or pool as possibly inert, these being decided upon automatically using LM. Therefore, it is likely that the poor performance of this method will mistakenly be attributed to such a powerful tool as the NP plot. On the other hand, users of the Design-Expert package are not only able to pick the points representing the potentially active effects in both the NP and HNP plots but also to test their significance using the ANOVA. For illustration, consider the HNP plot for the L_8 data of Table 6-2 portrayed in Figure 6-5 (a). Clearly, the main effects of the parameters A, D,

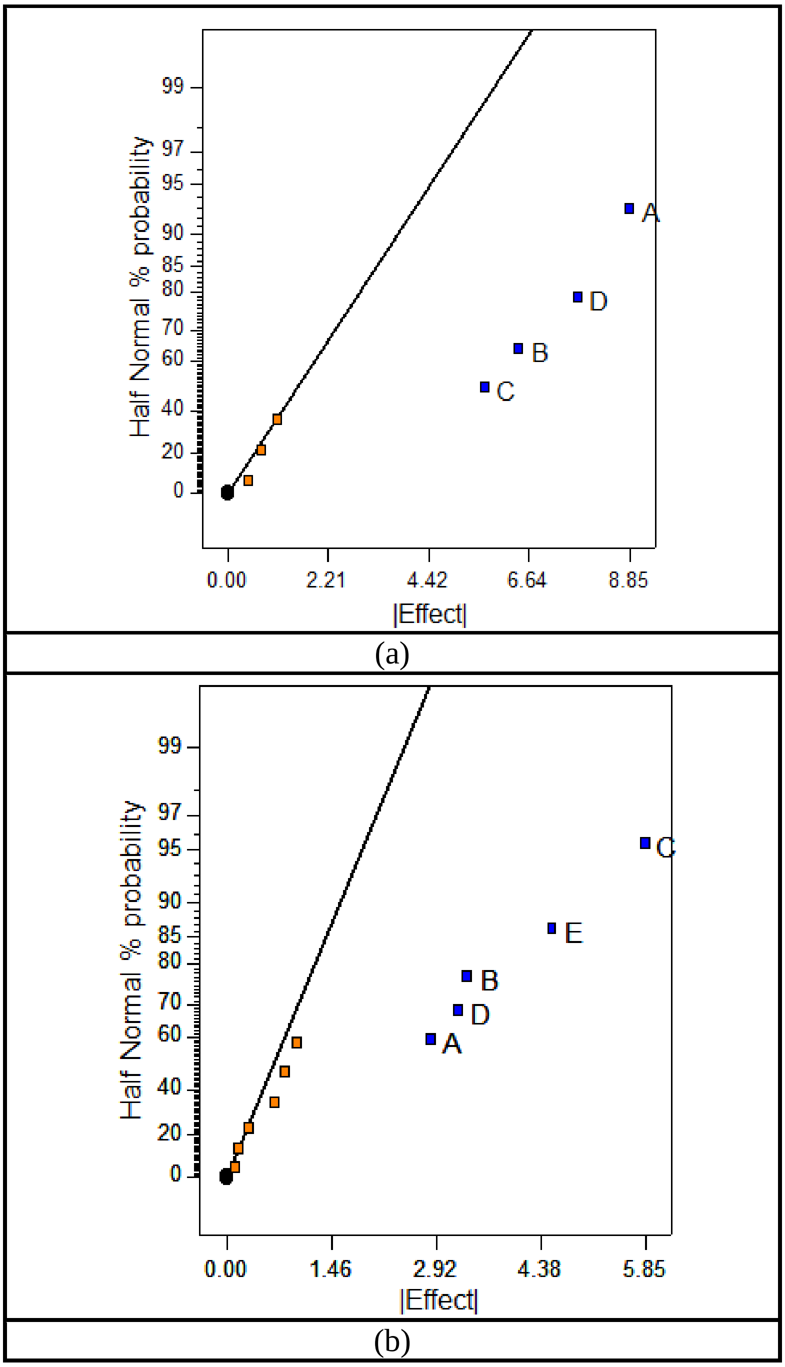


Figure 6-5: HNP Plots for the (a) L_8 and (b) L_{12} Data

B and C are too large to be explained by noise. Thus, they were picked as potentially active and using the ANOVA, they were all found to be significant at the 0.05 level. The same effects were declared significant when both the PU and UC methods were employed. However, as was the case with LM, the PD procedure failed to detect any of the active effects. This is due to its first step's inflated MSE, formed by pooling the active effects D, B, and C with the inert ones. This should not be misread as suggesting that the performance of LM and PD are equivalent. In fact, the PD method attained a better performance than did the LM as can be seen from the analysis of the L_{12} response 1 data shown in Table 6-2 (b). Using the HNP plot (Figure 6-5 (b)), the main effects C, E, B, D and A were flagged as important. The ANOVA showed that they were all significant at α =0.05. As the pooled effects were associated with the unassigned columns the same results were yielded using the UC method. Moreover, the PU and PD methods declared the same effects significant despite the MSE bias associated with the initial steps of the latter. Although 54.5% of the studied effects were inert, none of the active ones were detected using both versions of LM, the reason being that the median effect was not small enough to render a reliable PSE estimate.

A final noteworthy point in the context of the pooling methods' performance pertains to a distinct feature of the HNP plot. Despite the fact that the simulated full factorial data did not need transformation, this was required to validate the data analysis of some of the extracted fractional OAs. In contrast to all other pooling methods, the HNP plot allows any potential violation of the ANOVA assumption to be visually detected. In fact, when the inert effects do not line up with the origin or when there are gaps between them, data transformation is likely to be needed. For example, in the HNP plot (see Appendix 6) of the L_8 Response 2 data displayed in Table 6-2 (a) the inert effects do not fall along the line emanating from the origin. An examination of the residuals' plot revealed the existence of an outlier. Using the

Box-Cox plot (Box and Cox, 1964), a square root transformation was suggested. Having performed this, the inert effects were found to fall approximately along a line stemming from the origin of the HNP plot. Moreover, the plot of the residual outliers showed no point falling beyond its limits. The aforementioned plots are all presented in Appendix 6. Evidently, the HNP plot is more informative than all the other pooling methods investigated in this phase, yet, it has rarely been employed in practice. This might be attributed to the fact that the Taguchi approach which permeates the DOE applications in Manufacturing does not emphasise the use of graphical tools (Antony, 2006). Consequently, despite its simplicity and powerfulness, the HNP plot is not presented in any of the popular Taguchi DOE books such as Peace (1993), Ross (1996) and Roy (2001). It is difficult to understand why analysing unreplicated two-level Taguchi type experiments should not be undertaken using the HNP. The importance of using it may become clearer by calling attention to some of the limitations of what appears to be the most commonly used pooling strategy in practice. As the review in Chapter 3 review showed, most practitioners when analysing unreplicated experiments adopted a rule of thumb approach whereby the effects are pooled until the error combined degrees of freedom is nearly half the available ones (Logothetis and Wynn, 1989; Roy, 2001). Apparently, this is being practiced on the basis that “when a number of factors are included in an experiment, the laws of nature make it probable that half of them would be more influential than the rest” (Roy, 2001). In this regard, it is not clear whether the “laws of nature” would account for the impact of the employed OA in implementing such a rule. As already discussed, the assumption that half of the effects are inert can be translated in a considerably different number of pooled effects depending on the employed OA. It is not clear why prior knowledge regarding the percentage of active or inert effects is needed at all. Surely such knowledge is not necessary when the HNP plot is used. What is needed is to visually inspect the relative magnitudes of the plotted effects and locate a logical breaking point distancing the strong effects from those that are marginal or small. Although this may invoke an element of subjectivity in deciding which effects “fall off” the line, it can be circumvented by accompanying the use of the HNP plot with the ANOVA. The latter can be used to formally test the significance of the selected effects. If the experimenter is not sure whether certain points should be pooled or not, it is advisable to assess the impact of their inclusion and preclusion on the ANOVA result. In parallel, the subject knowledge should be utilised to examine the practical plausibility of pronouncing such effects as significant.

6.3.1 Sub-Phase-1

The ANOVA for the average and the ln S^2 of the Power are shown in Table 6-3. With regard to the average, both the fractional OA and the pooling method factors were significant at the 5% level. As was the case in Phase-1, the latter was more influential than the former, its PC being 64.15%. The error PC was only 3.26%. The main effects plot shown in Figure 6-6(a) exhibits similar patterns (relative values) of Power to those that were observed and discussed in phase-1 albeit with different magnitudes.

With respect to Power robustness, the ANOVA of its ln S^2 (Table 6-3 (b)) identified the fractional OA as significant at $\alpha = 0.05$. The p-value of the pooling method factor was very close to 0.05 indicating a potential significance. Collectively, the two factors explained 65.13% of the Power’s ln S^2 variability rendering an error PC of 34.87 %. This implies that while controlling the fractional OAs and pooling methods can reduce the average Power variability by 96.74%, it can only reduce the individual Power values variation around their average by 65.13%. One possible reason for this is that when using certain OAs and pooling methods, the variation of the individual values of Power around their average is inherently high. Another plausible reason is the existence of other, perhaps unknown or uncontrollable, sources of variation that were not systematically examined in this experiment.

Table 6-3: Sub-Phase-1 ANOVA for Power's (a) Average and (b) ln S^2

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value	PC (%)
A	0.200761	2	0.100381	86.02669	< 0.0001	32.59
B	0.396368	5	0.079274	67.93773	< 0.0001	64.15
Residual	0.011669	10	0.001167			
Total	0.608798	17				
(a)						
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**	**PC (%)**
A	7.758595	2	3.879297	11.28774	0.0027	42.21
B	5.557597	5	1.111519	3.23423	0.0539	22.92
Residual	3.436736	10	0.343674			
Total	16.75293	17				
(b)						

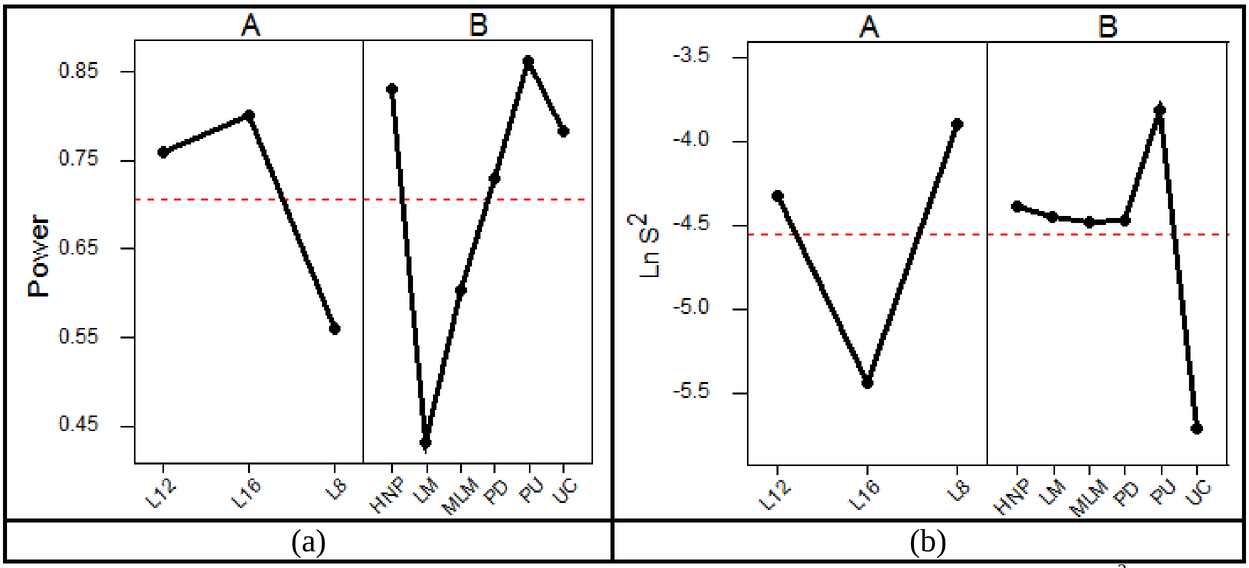


Figure 6-6: Sub-Phase-1 Main Effects Plots for Power's (a) Average and (b) ln S^2

As is evident from Figure 6-6 (b), the highest variability in Power was associated with the L_8 followed by the L_{12}. This confirms the results of Phase-1 regarding the necessity of avoiding the use of the L_8 to attain Power robustness (see Figure 6-1). Among the examined pooling methods, the PU procedure was associated with the highest variance whereas the UC method was the most robust. The agreement between the results of Phase-1 and its sub-phase can be explained by reverting to the L_8 and L_{12} HNP plots presented in Figure 6-5. Although extracted from full factorial designs with equal active effect sizes, the plots show that the fractional designs' active effects were of different sizes. It was to alleviate the impracticality issue pertaining to the use of active effects of equal sizes in phase-1 that the sub-phase-1 experiment was conducted. The extraction process in phase-1 circumvented this problem, to a large extent, by yielding active effects of different sizes. Consequently, the results of the sub-phase-1 experiments were in line with the phase-1 findings.

6.4 Phase-2

The factors declared significant at the 5% level using the ANOVA are shown in Table 6-4. The number of studied two-factor interactions (C) was the most influential factor as it accounted for 55% of the Power variability. The joint effect of the fractional OAs and the pooling methods (A) had a PC of 20.43% whereas the PCs of the number of active effects (B) and their sizes (D) were 2.79% and 4.5% respectively. The controllable-noise factors' interaction (AC) explained 9.76% of the Power variability. As shown in Figure 6-7, there was

Table 6-4: Phase-2 ANOVA for Power

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value	PC (%)
A	2.10699	6	0.351165	116.7914	< 0.0001	20.43
B	0.291654	2	0.145827	48.49958	< 0.0001	2.79
C	5.642464	2	2.821232	938.293	< 0.0001	55.13
D	0.469542	3	0.156514	52.05381	< 0.0001	4.50
AC	1.03355	12	0.086129	28.64507	< 0.0001	9.76
Residual	0.67953	226	0.003007			
Total	10.22373	251				

an opportunity to attain Power robustness against the variability transmitted from the number of active two-factor interactions noise factor by avoiding the use of the L_8 OA. In fact, the top three lines of the plot, associated with the uses of the L_{12} and L_{16}, exhibit less fluctuation than the remaining lines that represent different applications of the L_8. Thus, the use of the latter increases the variability that emanates from the changes in the number of active two-factor interactions.

The Power response was well represented by the significant effects as they collectively explained 92.62% of its variability. The main effects plot of the significant factors is depicted in Figure 6-8. Again increasing the number of active effects decreases the Power while increasing their sizes increases it. Moreover, the larger the number of active two-factor interactions the lower the Power. In fact, increasing the number of active two-factor interactions from 2 to 3 was accompanied by a substantial decrease in the Power. In no small

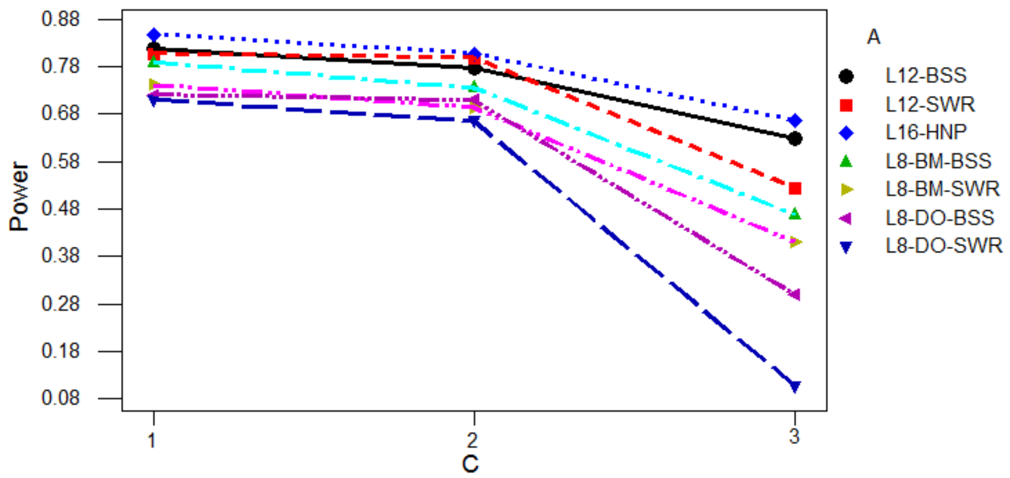


Figure 6-7: Phase-2 AC interaction Plot for Power

part, this is due to the influence of the partial aliasing associated with most of the cases examined in this phase. In fact, except for the case where the L_{16} was used in conjunction with the HNP plot, all the investigated levels of factor A involved partial aliasing. The larger the number of active two-factor interactions, the larger the influence of partial aliasing. As illustrated in Chapter 4, partial aliasing can seriously impact both the estimated parameter effects and their variances. Consequently, while the effects of the true active parameters may be camouflaged, many spurious effects may be pronounced significant.

In terms of the joint effect of the fractional OAs and the pooling methods, the use of the L_{16} analysed using the HNP plot attained the highest Power. Bearing in mind the results of phase-1, one can infer that the L_{16} always outperforms both the L_{12} and the L_8 especially when analysed using the HNP plot. This is not only valid in the cases where only main effects are active but also when some two-factor interactions are significant. Thus, the results of this and the previous phase provide empirical evidence to support Snee's (1985) claim that the L_{16} is the most useful fractional design for studying 5, 6, 7 and 8 factors. With regard to the use of

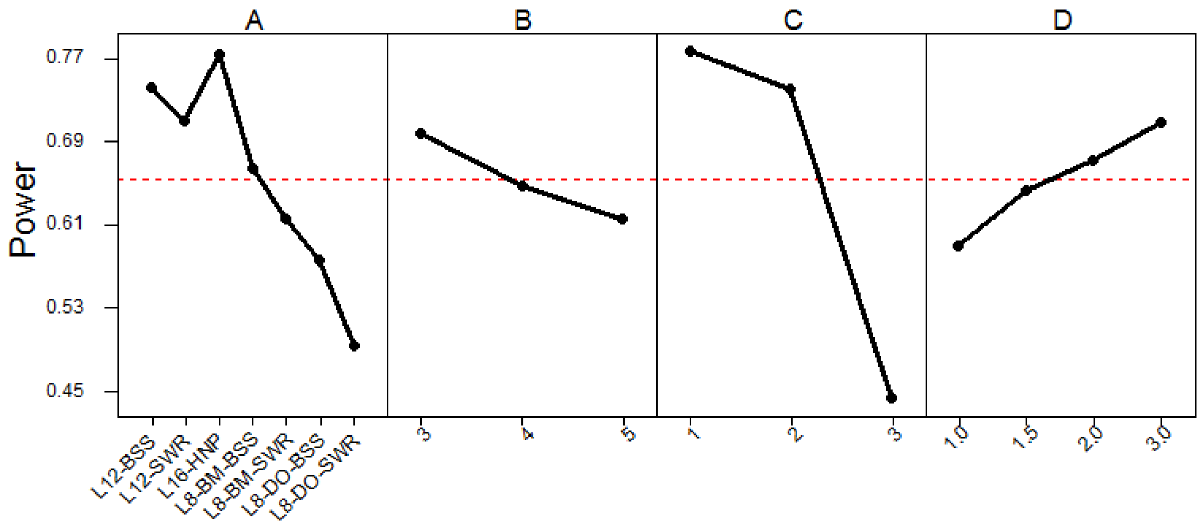


Figure 6-8: Phase-2 Main Effects Plot for Power

the L_{12} in the presence of active two-factor interactions, there has been some controversy in the DOE literature. For example, Daniel (1976) suggested that it should only be used when no interaction is expected to be active. Montgomery et al (1997) criticised the design as having a complex aliasing structure while Snee (1985) offered an opposing view. Based on his practical applications of the L_{12} and other Placket and Burman (1946) designs, he stated that "these designs do a good job of identifying those variables that have important effects." (Snee, 1985). Considering the studied experimental circumstances, the results of this phase are in agreement with Snee's (1985) view. In fact, when analysed using the BSS procedure, the L_{12} attained a Power level close to that achieved with the L_{16} when analysed using the HNP plot. Its performance is, by and large, dependent upon the number of active two-factor interactions, their sizes and the employed pooling method. The results of the review carried out in Chapter 5 revealed that in 77% of the reviewed experiments with active interactions, the number of two-factor interactions was 2 or less. Moreover, their sizes were rarely larger than the main effects. Under these conditions, the L_{12} showed a satisfactory performance especially when analysed using the BSS method. In fact, the latter outperformed the SWR method not only when analysing the L_{12} data but also the L_8. One reason for this, is the inflated MSE associated with the first step in SWR. The other reason relates to the impact of aliasing discussed in Chapter 4. To illustrate this, consider the L_{12} Response 2 data presented in Table 6-2 (b). These were extracted from a full factorial design where the effects A, B, C, D, BC and CD were active. Using the BSS procedure the true effects were detected and the fitted model was

$$Y = 0.011+2.99A+2.02B+2.51C+2.62D+3.02BC+3.02CD \qquad (6.2)$$

Note that although the effects were intended to be of equal sizes, their L_{12} estimates differ. Analysing the data using SWR, the inert interaction DE was the only detected effect. This was due to the aliasing impact on the estimated effect sizes. Consider the first step of the SWR method where the variable that has the largest effect (coefficient) is selected for a model with one independent variable. As candidates for selection, Table 6-5 shows the coefficient of each of the true active effects (equation 6.2) along with the DE interaction when selected for such a model. Clearly, the effects were seriously distorted by the impact of aliasing. For example, as shown in the L_{12} aliasing array presented in Table 4-9, variable A is aliased with the BC and

Table 6-5: SWR First Step Estimates of the L_{12} Data True Active Effects and the DE interaction

Variable	A	B	C	D	BC	CD	DE
Coefficient	0.98	1.01	2.51	1.61	1.15	1.35	-3.51

CD interactions with an aliasing coefficient of -0.333. Thus, as discussed in Chapter 4, excluding these from any model that contains A, should reduce the latter's coefficient by an amount equivalent to the result of multiplying the aliasing coefficient by their true coefficients i.e. ((-0.333x3.02)+(-0.333x3.02))= -2.01. Consequently, the coefficient of A was reduced to 0.98 (2.99-2.01). The same argument can be used to justify the distorted coefficients of the remaining variables. As C was not aliased with any of the true active effects (see Table 4-9), its coefficient was not changed. However, as a consequence of its aliasing with the active effects A, B, C and BC, the coefficient of the inert interaction DE was overestimated. Being the largest effect in the first step of SWR, it was mistakenly selected as significant. This prevented the selection of the true active effects in the subsequent steps. The above problem is also applicable to the FS procedure which was proposed by Hamada and Wu (1992) for analysing the L_{12} when parameter interactions are suspected of being active. The BSS method dealt with the problem of aliasing by collectively selecting the six true active effects to constitute the best fitted model of size 6 with an R^2_{adj} of 0.99. As illustrated in Chapter 4, unless excluded from the fitted model, the active aliased effects impart no bias on the variables included in the model. By examining all the possible models of various sizes, the BSS is likely to find a model comprising all the true active effects. Despite the impact of aliasing, such a model would not only render reliable estimates of the effects but also a high R^2_{adj}. Of course, models of smaller sizes than the true one are expected to be impacted by aliasing. For instance, in the above example, the models of size 1 will suffer from the same problem encountered in the SWR first step. However, instead of yielding one model as was the case in SWR, BSS provides the experimenter with a large number of candidate models.

Consequently, the challenge is to pick the right one from those fitted. This was dealt with, in this study, by utilizing the knowledge of the size of the “true” model simulated in the full factorial data. However, in practice, subject knowledge can be utilised to assess the plausibility of the fitted models. Moreover, a model should not be simply preferred over another for merely having a larger R^2_{adj}. To be selected, the model must make practical sense and its increase in the R^2_{adj} value is bound to be appreciable. A further noteworthy aspect of the BSS is that it is computationally intensive. This should not be a problem as many software packages are capable of handling 30 or more independent variables which is enough to cover a wide range of engineering applications.

The L_8 performance exhibited a substantial reliance on the employed augmenting methods. In fact, when the design was augmented using the BM, it attained a higher Power than the level achieved with the DO method. In the latter, the augmenting runs were selected so as to maximise the determinant $|\mathbf{X}^T\mathbf{X}|$, where **X** is the model array. Thus, the technique aimed to minimise the effects' standard deviation $((MSE*C_{jj})^{1/2})$ by reducing the size of its design related component C_{jj}. No account was taken of the extent, or even the existence of dependence between the response (Y) and the parameters that resemble **X**. On the other hand, the BM relied on the relationship between the studied response and the independent variables in deciding the augmenting runs. As discussed in Chapter 5, when used with the L_8 the BM firstly fitted all the possible models of size 7 or less. Then, all the possible augmenting runs were identified. The entries of each were then substituted into each of the developed models and the standard deviation of the resultant predicted responses was obtained. Consequently, a standard deviation value was recorded for each of the candidate runs. The larger this value, the more influential the run in discriminating between the potential models. Therefore, the candidate runs with the highest standard deviations were selected for augmentation. Since the impact of aliasing was manifested in estimating the coefficients of the fitted models, the BM accounted for both the studied parameters' aliasing and their relationship with the response in identifying the augmenting runs. By so doing it outperformed the DO method. This should not be misread as suggesting that the BM handled aliasing as efficiently as the DO method. Instead, the intention is to highlight that dealing with the two aspects less perfectly is better than confining the emphases to the optimising of a single aspect. As the results of this phase showed, the disadvantage associated with the BM augmentation rendering a smaller $|\mathbf{X}^T\mathbf{X}|$ value than that observed with the DO method was not only offset but also outweighed by the benefits gained from the utilisation of the relationship between the response and the studied parameters in selecting the augmenting runs. The BM has a further advantage over the DO method. Unless the true model is known, which is seldom the case in practice, the latter is likely to demand a large number of augmenting runs. For example, if 7 factors were studied using an L_8 and all the possible two factor interactions were needed to be estimable, then ${}^7C_2$ or 21 additional runs would be necessary. On the other hand, using the BM, any affordable number of augmenting runs may be selected by choosing from the candidate runs associated with the largest standard deviations.

6.4.1 Sub-Phase-2

In this sub-phase the settings of three factors of phase-2, namely, the number of active main effects (B) and two-factor interactions (C) and their sizes (D), were specified based on six published experiments. Hence only one factor was examined in this sub-phase. This was the joint effect of the OAs and pooling method for which the ANOVA is shown in Table 6-6. Clearly, the factor is significant at $\alpha = 0.05$ explaining 47.41% of the Power variability. Consequently, the error PC was 52.59% which is rather high. This should not be surprising in

Table 6-6: Sub-Phase-2 ANOVA for Power

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value	PC (%)
A	1.213865	6	0.202311	7.159139	< 0.0001	47.41
Pure Error	0.989069	35	0.028259			
Total	2.202934	41				

light of the phase-2 results, which demonstrated that Power cannot be well represented by solely considering factor A as it accounted for no more than 20.42% of its variability. In fact, the factors that exhibited a remarkable effect on the Power such as the number of active two-factor interactions were treated as noise sources of variation. Accordingly, their impact was predominantly reflected in the variation of the individual values of Power around their averages, which contributed largely to the error variance estimate in such a single factor experiment. This does not only explain why the error PC was high but also indicates the absence of degrees of freedoms to study the ln S^2. In fact, the latter is confounded with the average Power error variance. Viewed differently, for each level of factor A, ln S^2 consolidates the six replications into one value rendering an unreplicated response that cannot be examined using ANOVA in a single factor experiment. However, as was the case with the average Power, the impact of Factor A on ln S^2 can be investigated graphically. The plots of factor A main effects on the average Power and its ln S^2 are shown in Figure 6-9 (a) and (b) respectively. Clearly factor A exhibited a similar pattern of average Power to the one observed and discussed in phase-2. Regarding the Power ln S^2, the plot demonstrates that the highest level of Power variability was associated with the use of the L_8. In fact, less variability was observed when the latter was augmented using the BM. Moreover, the highest level of robustness was attained when the L_{16} and L_{12} were respectively used in conjunction with the

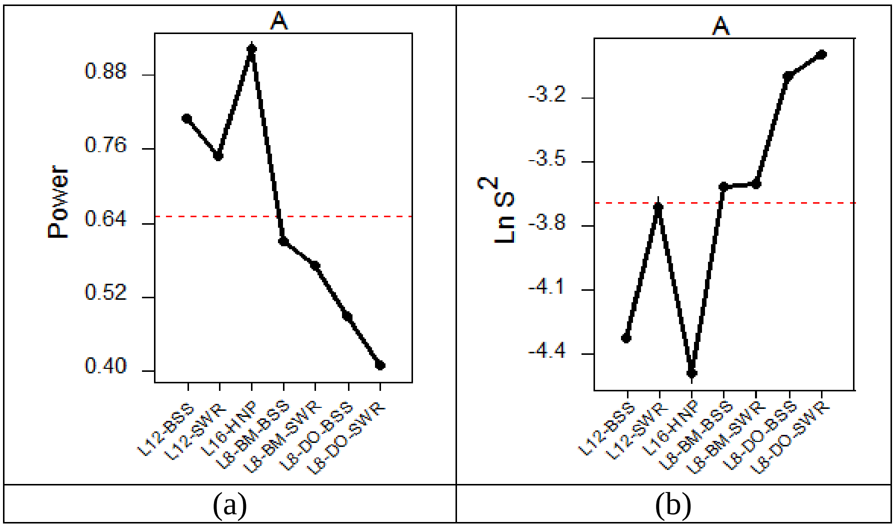


Figure 6-9: Sub-Phase-2 Main Effect Plots for Power's (a) Average and (b) ln S^2

HNP plot and the BSS method. To a large extent, this is in line with the implications of the controllable-noise interaction plot displayed in Figure 6-7, which shows that the variability transmitted from the changes in the number of active two-factor interactions can be minimised by using the L_{16} and the L_{12}. The results of this sub-phase are in agreement with the findings of phase-2. Again, this is largely due to the fact that after reducing the size of the simulated experiments, the effects that were intended to be of equal sizes differed, rendering similar circumstances to those simulated in this sub-phase.

6.5 Phase-3

Using ANOVA to analyse the results of this phase, the significant effects at an α level of 0.05 are shown in Table 6-7. The fractional OAs (A) and the pooling methods (C) were the two most influential factors. The number of active main effects (D) and their sizes (E) respectively accounted for 18.08% and 17.74% of the Power variability. Collectively, the effects explained 97.76% of the variability implying a strong representation of the response. Although, associated with the smallest PC (2.72%), the controllable-noise factors' interaction AE provided an insight with regard to attaining Power robustness. As shown in Figure 6-10,

Table 6-7: Phase-3 ANOVA for Power

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value	PC (%)
A	0.267936	1	0.267936	736.6949	< 0.0001	26.11
B	0.082872	1	0.082872	227.8584	< 0.0001	8.05
C	0.257176	1	0.257176	707.1104	< 0.0001	25.06
D	0.185653	1	0.185653	510.4577	< 0.0001	18.08
E	0.182867	3	0.060956	167.5986	< 0.0001	17.74
AE	0.028937	3	0.009646	26.52136	< 0.0001	2.72
Residual	0.019276	53	0.000364			
Total	1.024717	63				

the response curve associated with the L_{18} was flatter than that representing the L_9; hence the variance transmitted to the Power as a result of the changes in the sizes of the active effects can be minimised by using the L_{18}.

The main effects plot of the factors studied in this phase is shown in Figure 6-11. Clearly, the best level of Power was attained when the L_{18} was analysed using the PU method with regression ANOVA in the cases where the number of active main effects was 2 and the size of each was 3σ. Regarding factors A, D and E, the same argument presented in phase-1 regarding the relationship between the Power and the size of the used OA, the number of active factors and their size is applicable here. Moreover, the reasons why the PU method outperformed the PD procedure have already been addressed.

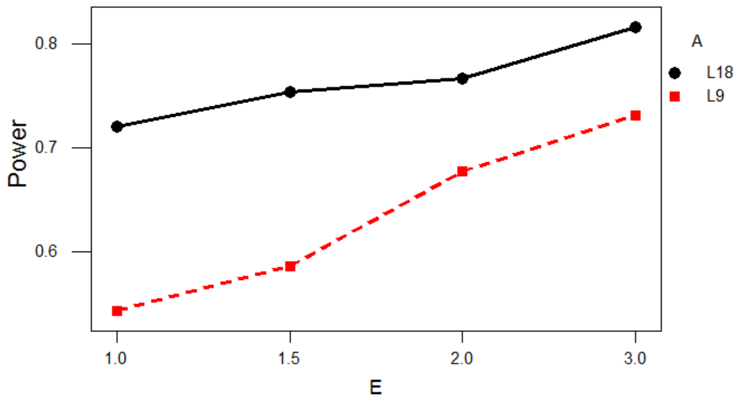


Figure 6-10: Phase-3 AD interaction Plot for Power

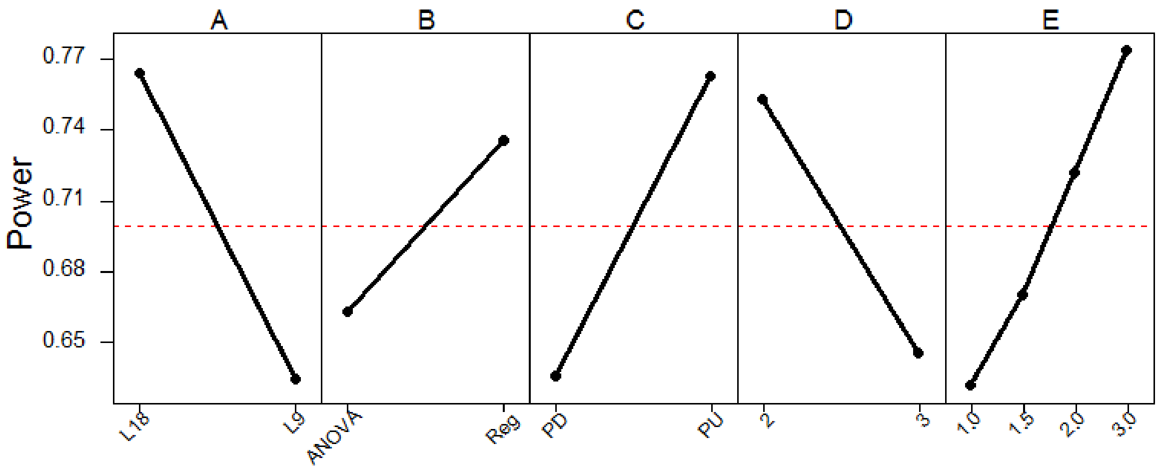


Figure 6-11: Phase-3 Main Effects Plot for Power

When the columns of the L_{18} and the L_9 were decomposed into single degree of freedom components using the L-Q system and analysed using the regression ANOVA, the attained Power was larger than that associated with the use of the standard ANOVA. To explain this, consider the L_9 data presented in Table 6-8. These were simulated so that the 3 main effects A, B and D were active. More specifically, the linear components of the three main effects along with the quadratic components of the parameters B and D were simulated to be significant. The conventional L_9 data (Table 6-8 (a)) were firstly analysed using the PU method in conjunction with the standard ANOVA. The results of the first and second steps of this analysis are shown in Table 6-9. These were performed by firstly pooling the smallest effect C, and testing the significance of D, the next larger parameter (Table 6-9 (a)). As the latter was not detected at the 1% significance level, it was pooled with parameter C to form the error estimate in the second step (Table 6-9 (b)). Clearly, both A and B were identified as significant at α =0.1. However, factor D was missed. Using α =0.05, none of the parameters was found significant in the second step of the PU method. This was also the case in its last step where the next larger effect B was pooled with D and C to estimate the error.

Consequently, using this approach, there were two possible outcomes depending on the α

Table 6-8: Simulated Data for the L_9 (a) in Its Conventional Form and (b) After Decomposition

A	B	C	D	Resp.
1	1	1	1	4.514673
2	1	2	2	5.643341
3	1	3	3	7.900677
1	2	2	3	6.772009
2	2	3	1	10.15801
3	2	1	2	9.029345
1	3	3	2	6.772009
2	3	1	3	7.900677
3	3	2	1	11.28668

(a)

A_l	A_q	B_l	B_q	C_l	C_q	D_l	D_q	Resp.
-1	1	-1	1	-1	1	-1	1	4.514673
0	-2	-1	1	0	-2	0	-2	5.643341
1	1	-1	1	1	1	1	1	7.900677
-1	1	0	-2	0	-2	1	1	6.772009
0	-2	0	-2	1	1	-1	1	10.15801
1	1	0	-2	-1	1	0	-2	9.029345
-1	1	1	1	1	1	0	-2	6.772009
0	-2	1	1	-1	1	1	1	7.900677
1	1	1	1	0	-2	-1	1	11.28668

(b)

level used in the PU method. The first was to overlook one of the active parameters which was the case when α =0.1 was used whereas the second was to declare none of the true parameters as active which was associated with using α =0.05. On the other hand, the true active parameters were all declared significant at α =0.05 in the second step of the PU procedure when the regression ANOVA was used with the decomposed L_9 data (Table 6-8 (b)) as shown in Table 6-10. Clearly, this came at the expense of falsely identifying the linear component of parameter C as important. However, this is, by far, less serious than overlooking all or even one of the true active parameters as was the case when the standard

Table 6-9: The ANOVA of the Two First Steps of PU Method Applied to the Conventional L_9 Data

Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	17.26831	2	8.634156	8.714286	0.1029
B	13.87127	2	6.935633	7	0.1250
D	3.680132	2	1.840066	1.857143	0.3500
Residual	1.98161	2	0.990805		
Total	36.80132	8			
(a)					
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	17.26831	2	8.634156	6.1	0.0610
B	13.87127	2	6.935633	4.9	0.0840
Residual	5.661742	4	1.415435		
Total	36.80132	8			
(b)					

ANOVA was employed. One reason why the regression ANOVA outperformed the latter relates to its ability to deal with the components of each of the studied effects separately. In fact, the standard ANOVA is not flexible in the sense that the two degrees of freedom associated with each three-level parameter are treated as one unit. They are either collectively pooled with the error or included with the tested effects. When the linear component is, say, the only active element of a three-level parameter, it is not wise to include the quadratic component with the tested effects. One reason for this is that by doing so the df_e will be reduced thereby increasing the MSE. In fact, it is clear that the MSE obtained using the regression ANOVA in Table 6-10 was smaller than its two corresponding values that were rendered in the standard ANOVA (Table 6-9). Of course, the larger the MSE the smaller the effects' calculated F-statistics and the smaller the Power. A further negative consequence of failing to incorporate the inert component of a three-level parameter with the error is that it increases the parameter degrees of freedom. Consequently, the numerator degrees of freedom (υ_1) associated with the F-statistic is increased. Generally, given the value of α and the df_e (υ_2), the larger the υ_1 the larger the tabulated F value and the smaller the Power.

Table 6-10: The ANOVA of the Second Step of PU Method Applied to the Decomposed L_9 Data

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value
A_l	17.19754	1	17.19754	243	0.0041
B_l	10.40345	1	10.40345	147	0.0067
B_q	3.467817	1	3.467817	49	0.0198
C_l	1.910838	1	1.910838	27	0.0351
D_l	1.910838	1	1.910838	27	0.0351
D_q	1.769294	1	1.769294	25	0.0377
Residual	0.141544	2	0.070772		
Total	36.80132	8			

The review in Chapter 5 revealed that the linear element was the only significant component in 52% of the cases that investigated three-level factors. If the standard ANOVA is used in their analysis, then for a linear effect to be detected it must be sufficiently large to not be masked by the combined impact of the inflated MSE and the increased υ_1. Otherwise, misleading conclusions regarding it are likely to be arrived at. One implication of this is that when analysing three-level experiments, sole reliance on the standard ANOVA should be avoided. In fact, the analysis can always be enhanced by accompanying the standard ANOVA with a regression one so that each component of the studied parameters can be addressed separately. Unfortunately, such an aspect is rarely, if ever, mentioned in the DOE literature. In fact, while most of the texts emphasise the use of the response surface method, Wu and Hamada's (2000) DOE book seems to be the only one, in which the use of regression ANOVA along with the L-Q decomposition system is mentioned as an alternative or at least a complement to the standard ANOVA in the context of analysing three-level experiments. However, no discussion or empirical evidence regarding which approach outperforms the other, especially in the absence of active parameters' interactions, was provided. One plausible reason for this is that, compared with two-level experiments, the analysis of three-level trials is accorded considerably less attention in the DOE literature.

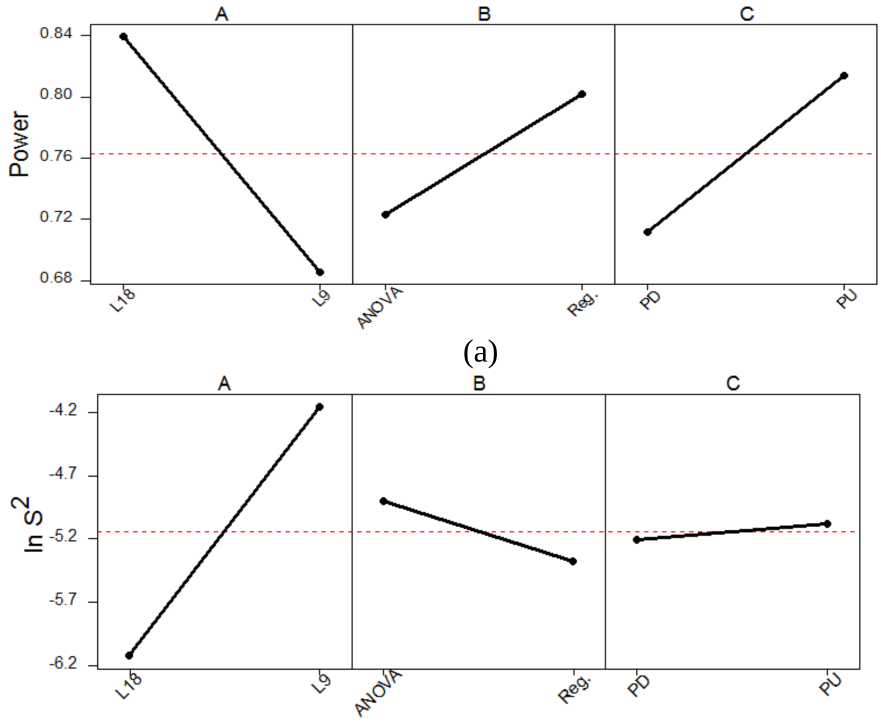
6.5.1 Sub-phase-3

The average Power ANOVA (Table 6-11 (a)) shows that all the studied factors were significant at $\alpha = 0.05$. As was the case in phase-3, the fractional OAs (A) were more influential than the pooling methods factor (C) which, in turn, was more influential than the type of ANOVA used (B). Taken together these factors explained 95.08% of the average Power. Their main effects plot is depicted in Figure 6-12 (a). The factors exhibited the same pattern of effects to that observed in phase-3, however, the fractional OAs was the only factor that significantly impacted the Power's ln S^2 (Table 6-11 (b)). The error, comprising all the

Table 6-11: Sub-Phase-3 ANOVA for Power's (a) Average and (b) ln S^2

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value	PC (%)
A	0.047586	1	0.047586	80.94599	0.0008	56.21
B	0.012561	1	0.012561	21.367	0.0099	14.32
C	0.021115	1	0.021115	35.91771	0.0039	24.55
Residual	0.002352	4	0.000588			
Total	0.083614	7				
(a)						
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**	**PC (%)**
A	7.800855	1	7.800855	44.67154	0.0005	86.19
Residual	1.047762	6	0.174627			
Total	8.848617	7				
(b)						

excluded factors, accounted for only 13.81% of the ln S^2 variability. As portrayed in Figure 6-12 (b), the use of the L_9 was associated with the highest variability of the individual values of the Power around their averages. To a certain extent, this is in line with the outcome of the controllable-noise interaction (Figure 6-10) which indicates that the Power variability can be reduced by using the L_{18} design. As was the case with the previous sub-phases, the results of this one are in agreement with those arrived at in phase-3.



(b)

Figure 6-12: Sub-Phase-3 Main Effects Plots for Power's (a) Average and (b) ln S^2

6.6 Phase-4

All the studied main effects along with two controllable-noise factor interactions were significant at the 5% level (see Table 6-12). Collectively, they explained 89.90% of the Power variability; so the error PC was only 10.10%. The most influential factor was the fractional OAs (A) followed by the size of the active effects (E). The number of active two-factor interactions (D) explained 14.47% of the Power variability. This was very close to the PC associated with the number of active main effects (C).

The controllable-noise factor interactions AD and AE are plotted in Figure 6-13. Clearly, the variation transmitted to the Power as a result of the changes in the number of active two-factor interactions as well as the size of active effects can be reduced by using the L_{18}. This is because the line representing it in both interaction plots was flatter than that associated with the L_9. The main effects plot of the factors studied in this phase is shown in Figure 6-14. For the same reasons presented in the previous phases, increasing the number of active main effects and two-factor interactions decreases the Power. However, increasing the size of the active effects decreases it.

Table 6-12: Phase-4 ANOVA for Power

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value	PC (%)
A	0.364212	1	0.364212	168.4479	< 0.0001	26.85
B	0.08673	1	0.08673	40.11267	< 0.0001	6.27
C	0.181476	1	0.181476	83.93251	< 0.0001	13.30
D	0.197358	1	0.197358	91.27795	< 0.0001	14.47
E	0.242435	3	0.080812	37.37533	< 0.0001	17.50
AD	0.096877	1	0.096877	44.80533	< 0.0001	7.02
AE	0.067005	3	0.022335	10.32992	< 0.0001	4.49
Residual	0.112433	52	0.002162			
Total	1.348526	63				

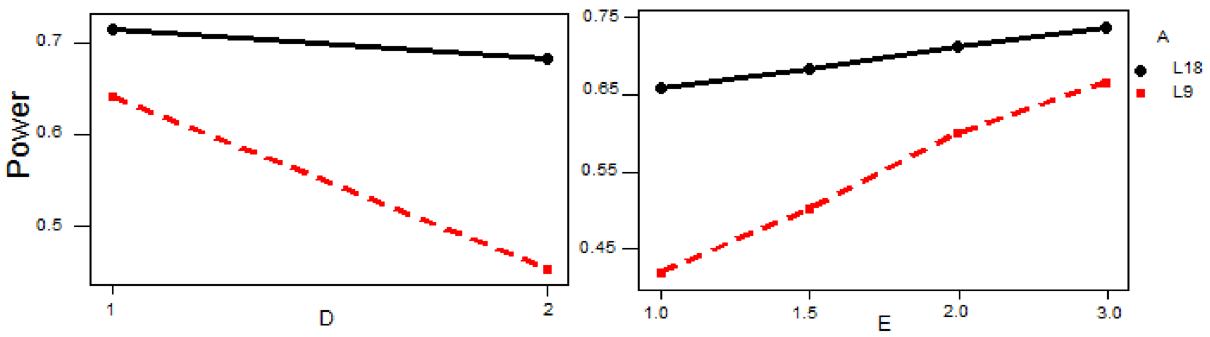


Figure 6-13: Phase-4 AD and AE interactions Plots for Power

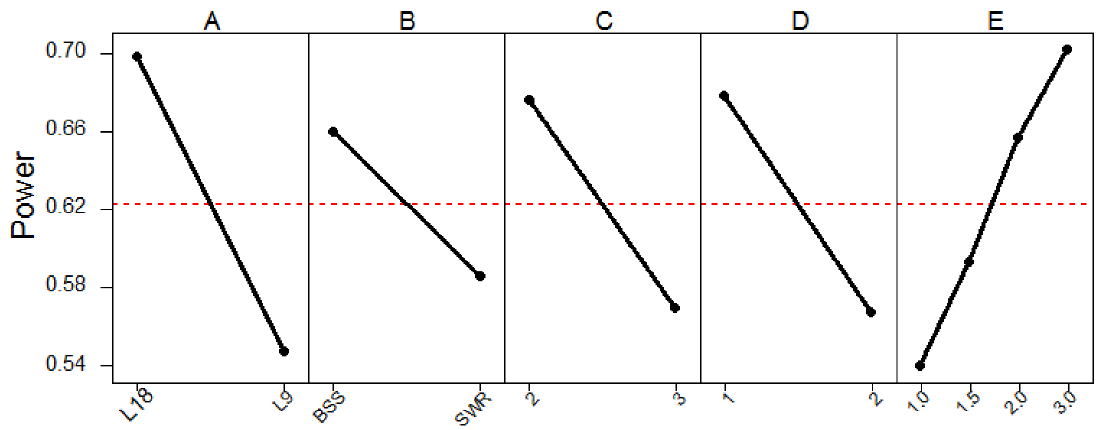


Figure 6-14: Phase-4 Main Effects Plot for Power

The L_{18} was associated with a higher Power level than that of the L_9. Moreover, for the same reasons as given in phase-2, the BSS method outperformed the SWR procedure. Despite being frequently used in practice, the L_{18} and L_9 were rarely decomposed into single degree of freedom elements and analysed using either of these. This is not only due to the lack of coverage of these approaches in the DOE literature but also to the ill-treatment that the analysis of the three-level trials received. The inconsistency of dealing with the latter analysis in some of the Taguchi DOE books is a case in point. For example, whilst Roy (2001) used only the standard ANOVA for analysing three-level experiments, Peace (1993) ignored its use altogether. Ross (1996) barely considered three-level trials, two-level ones being predominantly discussed. With regard to conventional DOE, again Wu and Hamada's (2000) book appears to be the only one that theoretically highlights the importance of using the L-Q system to decompose three-level OAs in the presence of active two-factor interactions. For example, conventionally it is not possible to examine interactions when four three-level parameters are studied using the L_9. However, Wu and Hamada (2000) showed that by decomposing the array using the L-Q system, the degrees of freedom associated with the inert components of the studied parameters can be exploited to examine some two-factor interactions. They also mentioned that both SWR and BSS could be used to detect the active effects in such cases; however, they only employed the former. Moreover, there was no assessment of how well these pooling methods performed or which one was more powerful than the other. In fact, no research has been found that empirically investigated this data analysis approach. To this end, the results of this phase showed empirically that the success of this approach depends substantially on the fractional OA employed, the number of active main effects and two-factor interactions as well as their sizes. As per, the pooling methods, it was revealed that BSS is more powerful than SWR.

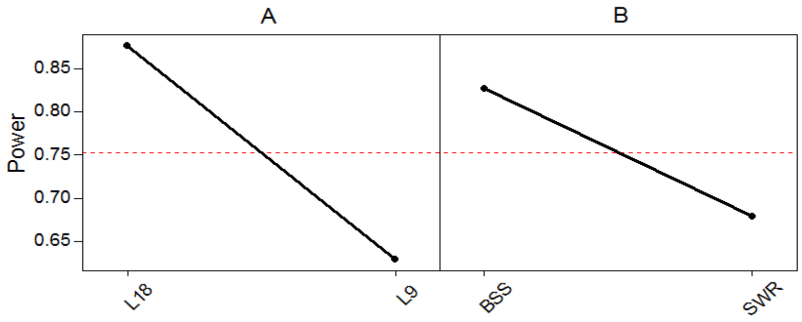
6.6.1 Sub-Phase-4

None of the studied factors were found to be significant in the average Power ANOVA at the 5% level (see Table 6-13 (a)) even though the factors collectively explained 98.63% of the variability. The reason for not declaring them significant is that the df_e is only one. Since the tabulated value of $F_{0.05,1,1} = t^2_{0.025,1}$, it follows that the former suffers from the same robustness problem as the latter when there is only one degree of freedom associated with the error. Consequently, care should be exercised when using the ANOVA in such cases.

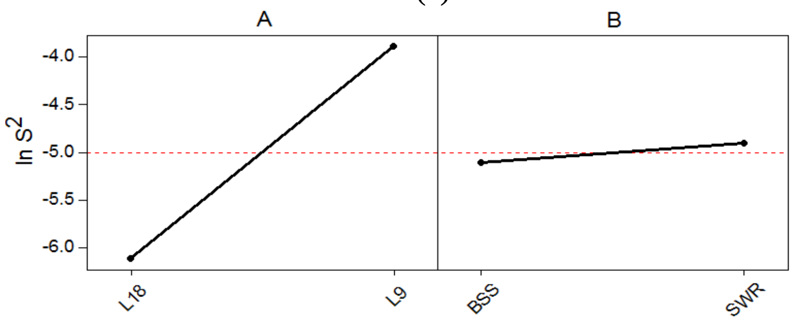
The main effects plot of the fractional OAs and the pooling methods factors is presented in Figure 6-15 (a). As was the case in phase-4, the best performance was attained when the L_{18} was analysed using the BSS method. With regard to the variability of the individual values of the Power around their averages the ANOVA in Table 6-13 (b) implies that the employed fractional OA was the only significant factor at α=0.05. Moreover, as shown in Figure 6-15(b), the highest level of ln S^2 was observed with the L_9. This confirms the inferences drawn from the controllable-noise interactions plots displayed in Figure 6-13.

Table 6-13: Sub-Phase-4 ANOVA for Power's (a) Average and (b) ln S^2

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value	PC (%)
A	0.061413	1	0.061413	160.6821	0.0501	72.82
B	0.022018	1	0.022018	57.60719	0.0834	25.81
Residual	0.000382	1	0.000382			
Total	0.083813	3				
(a)						
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**	**PC (%)**
A	4.95911	1	4.95911	29.51829	0.0322	90.48
Residual	0.336003	2	0.168001			
Total	5.295112	3				
(b)						



(a)



(b)

Figure 6-15: Sub-Phase-3 Main Effects Plots for Power's (a) Average and (b) ln S^2

6.7 Types I and II Errors and Their Consequences

The importance of using the above recommended fractional OAs and pooling methods can be illustrated by examining the consequences of failing to detect significant effects and falsely identifying inert ones as significant. As these are dependent upon the objectives of the experiment and the approach adopted in its design and analysis, it is important to distinguish between the conventional and Taguchi's approach to experimental design. As discussed in Chapter 2, the former is sequential in that the initial stage aims to identify the most influential factors so that their effects are studied in detail in the subsequent characterisation and optimisation experiments. On the other hand, Taguchi's approach utilises a one-shot experiment followed by confirmation runs. Falsely detecting some effects as significant should be of little concern when the conventional approach is adopted as these are likely to be discovered and eliminated at subsequent experimentation stages. However, missing an active factor at the screening stage can have serious consequences. In fact, unless the experimentation team decides to augment the initial experiment with further runs so that the missed effects are detected, the conclusions drawn from the subsequent experiments can be seriously distorted.

On the other hand, falsely detecting effects as significant when using Taguchi's approach has several consequences. Firstly, faulty conclusions relating to the cause and effect relationship between the process parameters and the response may be drawn. Consequently, the opportunity of gaining a better understanding of the process is lost. As previously pointed out, this is one of the main purposes of using DOE. Secondly, funds may be appropriated to control factors that have marginal effect on the response thereby unnecessarily incurring an additional cost. Moreover, as only some of the marginal effects would be pooled, the estimate of the error variance which represents the variation of the response around its average would be inaccurate. Since variance reduction is a main target in any process improvement endeavour, failure to obtain a reliable estimate of it can seriously compromise the attainment of the intended performance level. A further issue is that, unless the "best" prescribed combination of factor levels is identical to one of those performed in the experiment, its corresponding response value must be predicted. The prediction equation should only involve the detected significant effects. Incorporating falsely flagged ones would result in an overestimation of the predicted results which may lead to a disappointing confirmation run when actually the results would have been validated had the predicted results not been biased by the inclusion of marginal effects. The same problem may be encountered as a result of using an inaccurate estimate of the error variance in constructing the confidence interval around the predicted response at the "best" settings.

A more crucial problem than falsely detecting effects as significant, especially in one-shot experiments, is the failure to detect active factors or interactions. This again could lead to a failure in confirming the experimental results thereby necessitating further experimentation. Otherwise the attainment of the intended process improvement will be impaired.

To reduce the impact of committing Types I and II errors, the recommended fractional OAs and pooling methods should be used and whenever possible, sequential experimentation should be adopted so that the conclusions drawn at one stage can gradually improve in the following stage as new data become available.

6.8 Summary

Unreplicated fractional factorial designs provide a formal cost effective way of improving products and processes. Their ability to identify the "real" active effects as would be obtained from performing full factorial trials depends on several variables. Some of these are controllable such as the fractional OAs that are used and the pooling methods applied. In contrast, however, others such as the number of active main effects and interactions and their sizes are uncontrollable (noise). The simulations performed here allowed the impact of these variables on the Power of the fractional designs to detect the true active effects to be examined using DOE techniques. This not only allowed the statistical significance testing of each variable but also the quantification of its explained variance. More importantly, it enabled the controllable-noise interactions to be exploited to desensitise the impact of the noise variable on the Power. To this end, it was found that the use of small designs such as the L_8 and the L_9 should, as far as possible, be avoided to attain Power robustness. It is well established in statistics that small designs are associated with higher variance than that of large ones. However, this relates to the size of the estimated MSE which affects the average Power. It should not be confused with the variation of the individual Power values around their averages, which can only be reduced by exploiting the interactions between the noise and the controllable factors. In fact, to date, DOE research has tended to focus on the average Power not its variability. In this study, this gap was bridged by using the DOE technique.

In comparing the performance of fractional OAs with full factorial ones, the classical view is only to emphasise the impact of aliasing. In fact, this merely represents one way in which the impact of the fractional OAs factor may be manifested. Of course the other way relates to the size of the OA. As the results of this study revealed, the success of certain fractional OAs in detecting the true active effects involves an element of complexity. This is due to the fact that reducing the size of an experiment may result not only in distorting the actual effects' sizes but also in requiring data transformation. The consequences of the former can be very serious, especially in cases where the true active effects are underestimated whilst the inert ones stand out as potentially active. While the impact of this cannot be eliminated it can be reasonably reduced by appropriately choosing the fractional OAs as well as the pooling methods. For example, in the context of two-level experiments, the use of the L_{16} with the PU method and the HNP plot is recommended. In fact, the latter is more informative than all the other pooling methods as it enables the relative magnitude of the studied effects to be investigated visually. Moreover, it has the capability of revealing the possible need for data transformation. If the available time and resources limit the number of trials to less than 16, the use of the L_{12} is suggested. In this instance, the HNP plot and the PU method should only be used if no two-factor interaction is suspected. If this is not the case, BSS should be employed. When using the L_8, it must be borne in mind that lower and inconsistent levels of Power are expected. Moreover, if augmenting is necessary, it is advisable to use BM as it outperformed the DO method.

With regard to three-level designs, the L_{18} was associated with a high and robust level of Power and hence should, whenever possible, be used. Irrespective of which OA is chosen, the decomposition of each parameter's column into single degree of freedom components using the L-Q system is recommended. Of course, when qualitative factors are studied, it is necessary to examine the extent to which the interpretation of each component as a comparison between certain level-averages (see Appendix 2) makes practical sense. The decomposed designs should be analysed using the BSS procedure. It is instructive to accompany the use of the latter with the SWR analysis and examine the reasons behind any results' disagreements. In fact, Tukey (1969) wisely advised "a body of data can and usually should be analyzed in more than one way". Consequently, it is always advisable to analyse the experimental data using more than one pooling method. When several methods pronounce the same parameters as active, especially at α=0.05, the experimenter can have more confidence in the drawn conclusions. However, when certain effects are declared active only by some of the pooling methods, subject knowledge can be utilised to assess the practical plausibility of their significance. Otherwise, one-factor-at-a-time confirmation trials may be performed varying only the settings of the suspected parameters.

In order to assess the validity of this Chapter's findings, data from practical experiments were investigated as will be discussed in the next Chapter.

CHAPTER 7: VALIDATING THE RESULTS OF THE SIMULATION EXPERIMENTS

7.1 Introduction

The effort in this Chapter is concerned with validating the findings of the previous Chapter simulation experiments. In the context of simulation modeling, validation can be defined as "substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model" (Schlesinger et al. 1979). The exercise detailed in the previous Chapter evaluated certain DOE strategies. Judging its possession of "a satisfactory range of accuracy" that is in line with its purpose, is not straightforward. One reason for this relates to the concept of validating any simulation results. As Kleijnen (1995) noted there is "no standard theory" on validation, nor is there "a standard 'box of tools' from which tools are taken in a natural order". Moreover, Davis (1992) regarded the validation of simulation results as "a complex subject that has troubled model developers and users for many years". Thus, there is generally an intrinsic complexity in the validation task.

In addition there are also sources of difficulties that relate particularly to the simulation experiments performed in this study. One such example concerned obtaining actual experimental data conforming to that simulated. Large full factorial experiments are rarely conducted in practice, and even when they are, the observed data are seldom reported in the literature. Indeed, in cases when such experimental data are found, there is no guarantee that their analysis will render results that match the simulated conditions, especially with regard to the settings of the noise factors such as the number of active effects and their sizes. By the same token, no control can be exercised over the results of such observed data with regard to pronouncing as significant only main effects or even some two-factor interactions. One alternative to relying on published experimental data is to generate them by conducting some "physical" experiments. However, no experimenter can exercise any control over either the rendered number of active effects or their sizes.

Despite the infeasibility of finding actual experimental data sets that exactly match the simulated conditions, it was possible to locate certain case studies that bore some resemblance to them. Moreover, it was necessary to conduct a large three-level experiment similar to the simulated ones in an attempt to yield circumstances comparable with those modelled. Having obtained appropriate "real" full factorial experimental data sets for each of the simulation phases, the empirical evaluation approach (Davis, 1992) to validation was adopted. This entailed comparing the conclusions rendered from the simulation experiments with those drawn from the "real" experimental data. Rather than judging whether the simulation results were valid or not, the intention was to assess the "degree" of agreement between the compared inferences. In so doing, Sargent (2000) distinguished three approaches of comparison: graphical, confidence intervals and hypothesis testing. The first was adopted for two reasons. First, there is a considerable difference in the sample sizes between the results upon which the simulation experiment conclusions were drawn and the ones that underlie the validation experiments' inferences. Second, the cornerstone underpinning statistical methods is randomisation. For the conclusions drawn on the basis of hypothesis testing or confidence intervals to be valid, the compared samples should be randomly selected. The validation case studies were not selected randomly, so their calculated statistics were not random variables and therefore should not be analysed in this way. Furthermore, as the simulated conditions and those associated with the validation case studies were not exactly the same, it is the pattern of the results and their relative rather than their absolute magnitudes that should be compared. To this end, graphical techniques are very informative. The validation of each of the simulation phases will now be discussed followed by a summary of the main conclusions.

7.2 Phase-1

In order to assess the credibility of the conclusions drawn from this phase, it was necessary to find published examples of large two-level full factorial experimental data where main effects were the only significant contrasts. Three data sets that met these requirements were found in Box et al (2005). These related to the manufacturing process of a certain dyestuff and incorporated three responses; its strength, hue, and brightness. Six parameters were identified as having a potential effect on the responses: polysulfide index (A), flux ratio (B), moles of polysulfide (C), reaction time (D), amount of solvent (E) and reaction temperature (F). Two levels were chosen for each parameter, and a full factorial (2^6=64 runs) experiment conducted. The data for the three responses are shown in Appendix 7. For each response, the runs that corresponded to the L_{16}, L_{12} and L_8 were extracted. The results showing the ability of these fractional OAs to declare as active the same effects pronounced significant in the full factorial trials when analysed using the six pooling methods that were investigated in this phase, are given in Appendix 8. A detailed analysis of the dyestuff hue response is presented as an example in Appendix 9. The average Power associated with each of the fractional OAs and the pooling methods for the dyestuff data were computed and are plotted in Figure 7-1. Generally, the patterns of the “curves” are similar to those observed in Figure 6-2. As was the case in the simulation experiments, the L_{16} outperformed the L_{12} which in turn attained a higher average Power than the L_8. However, while the results related to the performance of the HNP plot, LM, MLM and the PU methods were in line with their corresponding simulation findings, this was not the case with the PD and the UC methods.

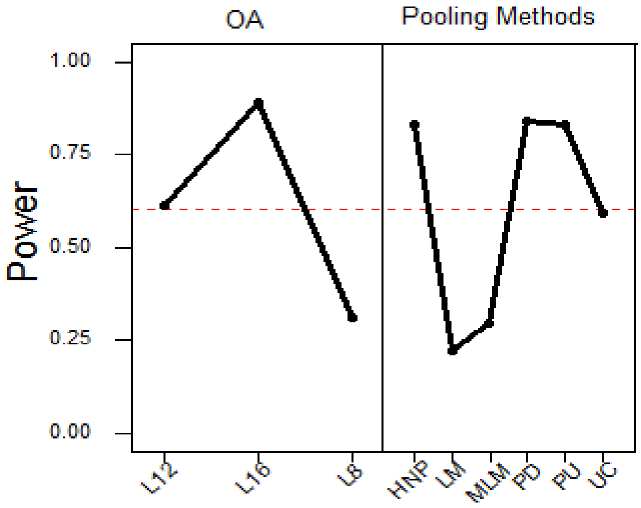


Figure 7-1: Average Power of each Fractional OA and Pooling Method in Phase-1 Validation

In the case of the simulation results, the latter outperformed the former whereas in the dyestuff experiment the opposite was true. In fact the PD method attained the same average Power as that observed with the HNP plot and the PU method. Before explaining the reasons for this, it is important to remember that the average Power was calculated over 50 identical replications against each combined setting of the fractional OAs, pooling methods and number and sizes of active effects in the simulation experiments. Then the average of these Power values associated with each of the examined fractional OAs and pooling methods was computed and plotted in Figure 6-2. In the case of analysing the dyestuff experiment, however, each of the three responses was associated with different values of the number and sizes of active effects. Thus, for each response the Power (associated with each of the OAs and pooling methods) was estimated on the basis of a single replicate. Furthermore, the average Power shown in Figure 7-1 was computed over the three responses. Considering these differences between the simulation and dyestuff experiments, it should not be surprising that the observed results were not the same. In addition, it was due to the use of $\alpha = 0.1$ in the PD technique (to overcome the impact of the inflation in the MSE associated with the initial steps) that it outperformed the UC method in the dyestuff experiment. In fact, had the same α

level been used with the UC method, it would have attained the same level of Power as that observed with the PD procedure.

The results of the simulation experiments revealed that the superiority of the UC method over the PD technique was very marginal (Figure 6-2). The exceptional performance of the PD method attained in the analysis of the dyestuff experiment must be interpreted reflectively. This is because not only did it outperform the UC method but also it achieved an equivalent level of performance to both the HNP plot and the PU method. Ross (1996) argued that the latter is more powerful than the PD method. Moreover, if the problem of the MSE inflation in its initial steps is considered, there is no reason to believe that the PD is as powerful as the HNP plot which is free from such a pitfall. The chief point to be gleaned from the above discussion is that, in practice, pooling methods may exhibit an unexpected performance owing, in no small part, to the variability in the observed active effect sizes. As this is a noise factor over which there is no control, it is important to guard against its effect by adopting a robust DOE strategy. To this end, the importance of avoiding the use of small OAs such as the L_8 cannot be overemphasised. With regard to the pooling methods, as suggested in the previous Chapter, it is better to use more than one technique to analyse the experimental data and investigate the reasons for any disagreement in their inferences. Adopting such an approach should not only increase the Power but also improve its robustness to the noise factors.

7.3 Phase-2

This phase of the simulation experiments entailed two-level full factorial designs where only main effects and two-factor interactions were active. Three experiments satisfying these conditions were found in the literature, two involving the study of 6 parameters (2^6=64 runs) and the other 5 parameters (2^5=32 runs). The first experiment was reported in Taguchi (1987) and related to the wool washing and carding processes of wool spinning. Its main objective was to investigate the key parameters that influenced the number of defects (number of neps, pinholes and troubles) in the output. The studied parameters were the number of revolutions of worker (A), worker gauge (B), feed roller gauge (C), card feed quantity (D), water content (E) and residual fat rate (F). The second experiment related to a Metal Inert Gas (MIG) welding process and was presented in Ganjigatti et al (2007). The six two-level parameters were the welding speed (A), the arc voltage (B), the wire feed rate (C), the gas flow rate (D), the nozzle to plate distance (E) and the torch angle (F). Examining the impact of these on the bead width response was one of the main purposes of the experiment. The third experiment concerned the voice-coil actuated radial contour turning process and was discussed in Reddy et al (2001). The main objective was to understand the impact of the five two-level parameters on the turning process performance. The parameters were the contour's order of angular symmetry (A), the maximum acceleration requirements on the contour (B), the depth of cut (C), the feed rate (D) and the type of controller (E). The maximum tracking error (peak error within one revolution) was used as the performance measure of the process.

The data for all these experiments are given in Appendix 7. For each data set, the L_{16}, L_{12} and L_8 fractional factorial runs were extracted and analysed using the pooling methods investigated at this phase. The effects declared significant using these fractional factorial designs and analysis strategies along with those that were pronounced significant in the full factorial analysis are shown in Appendix 8. As an example of how these results were implemented, a detailed analysis of the MIG welding experiment is provided in Appendix 9. The average Power attained for each fractional OA/pooling method strategy was computed and is shown in Figure 7-2. The observed pattern is very similar to that generated from the simulation results and depicted in Figure 6-8. The maximum Power in both cases was observed with the L_{16} when analysed using the HNP plot. The BSS method outperformed SWR when used to analyse both the L_{12} and the L_8. When augmented using the BM method, the latter attained a higher Power level than that associated with its DO augmentation.

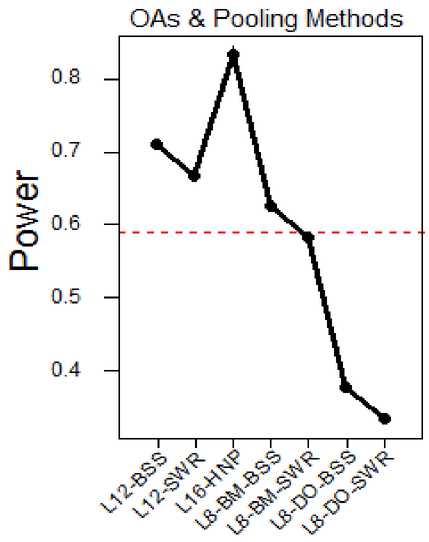


Figure 7-2: Average Power of each Fractional OA-Pooling Method strategy in Phase-2 Validation

Hence a high degree of credibility can be attached to the conclusions arrived at using this phase of simulation experiments as they were in line with the inferences drawn from the actual full factorial experiments.

7.4 Phase-3

In assessing the viability of the phase-3 simulation results, it was necessary to at least find a large published three-level full factorial experiment with the main effects being the only active contrasts. Unfortunately, no experiment that met these requirements was found so, it was necessary to conduct one. The performed experiment concerned an Electric Discharge Machining (EDM) process whereby the material removal takes place as a result of the discharge of energy between a tool and a workpiece, which are separated by a small gap filled with a dielectric fluid. The process involves discrete sparks initiated using a DC pulse generator for a certain duration followed by a similar period during which deionisation of the dielectric occurs and the gap is flushed of debris (McGeough, 1988).

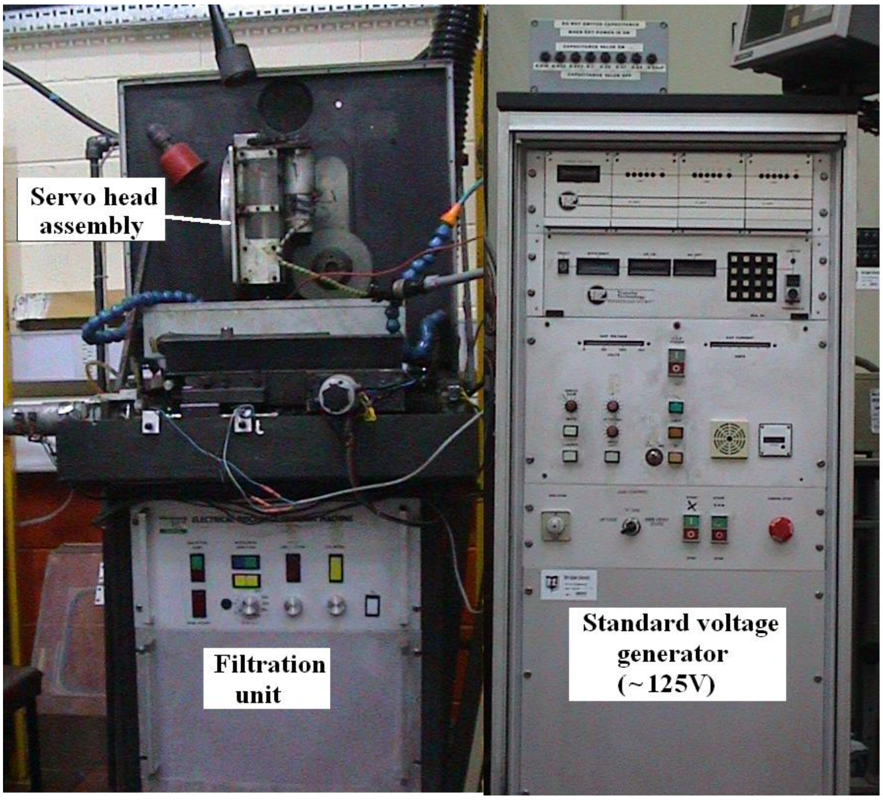
7.4.1 Equipment, Workpiece and Tool Electrode Materials

The experiment was carried out using an EDM machine built in the School of Mechanical Engineering at the University of Birmingham which was fitted with a miniature DC vertically mounted servo unit (330 mm long with an 80 mm stroke). The machine was connected to a Spark Tec international generator which employed an open circuit voltage of ~125 V with programmable peak current and on/off time selectable in steps of 1 amp and 1 μs respectively. A hydrocarbon oil dielectric BP 180 was used. This was supplied externally to the electrode/workpiece interface using two adjustable nozzles at a flow rate of 2 litres/min. and re-circulated through a Filtermist Superfine dielectric filtration unit (filtration down to 1 μm particle size). Figure 7-3 shows the generator, the EDM unit, the servo head and the machining setup. The workpiece material was Titanium (Ti-6Al-4V) supplied in bars which were cut into 15 mm long cylindrical discs each with a diameter of 20 mm. Copper electrodes of 50 mm length and 12 mm diameter were used.

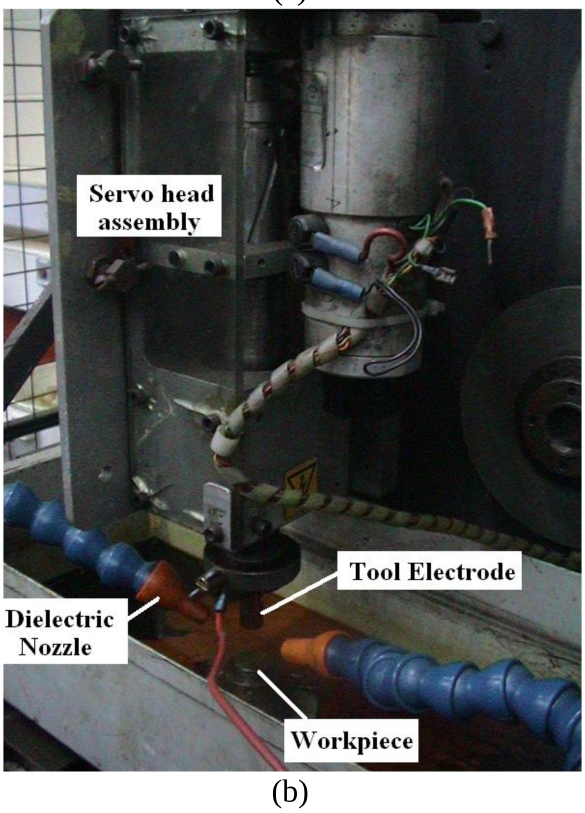
7.4.2 Experimental Parameters and Procedures

Four parameters were chosen: peak current, pulse on-time, capacitance and pulse off-time, each at three levels (see Table 7-1). Fixed parameters included the tool polarity which was positive and the machining time which was 10 minutes.

A full factorial three-level experiment comprising 3^4= 81 runs was conducted. The responses of interest were the Material Removal Rates (MRR) of both the workpiece and the tool. The weights of both were recorded before and after the EDM operation using a micro-level balance, the volume of the material removed was then obtained by dividing the weight



(a)



(b)

Figure 7-3: (a) EDM Unit with the Power generator; (b) Servo Head and Machining Setup

Table 7-1: Parameters of the EDM Experiment and Their Levels

Parameter	Description	Levels		
		1	2	3
A	Peak Current (A)	10	15	20
B	Pulse On-Time (μs)	50	100	200
C	Capacitance (μF)	0	0.1	0.22
D	Pulse Off-Time (μs)	15	20	25

differences by the appropriate material density. The MRR (per minute) was then calculated by dividing the volume of the removed material by the machining time.

7.4.3 Results and Validation Process

The full factorial data of the EDM experiment are listed in Appendix 7. Their analyses (detailed in Appendix 9) revealed that only one of the responses namely the workpiece MRR was comparable to the simulation results with main effects only being active contrasts. Consequently, the full factorial workpiece MRR response values associated with the settings that corresponded to the L_{18} and L_9 were extracted and analysed using the ANOVA types and pooling methods investigated at this phase (see Appendix 9). The effects detected as active using these strategies along with those that were declared significant in the full factorial analysis are shown in Appendix 8. The Power values associated with each fractional OA, ANOVA type and pooling method for the workpiece MRR are presented in Figure 7-4. Clearly, the patterns match their counterparts from the simulation results seen in Figure 6-11. The L_{18} attained a higher Power level than that of the L_9, regression was associated with a higher Power level than that of the standard ANOVA and the PU outperformed the PD

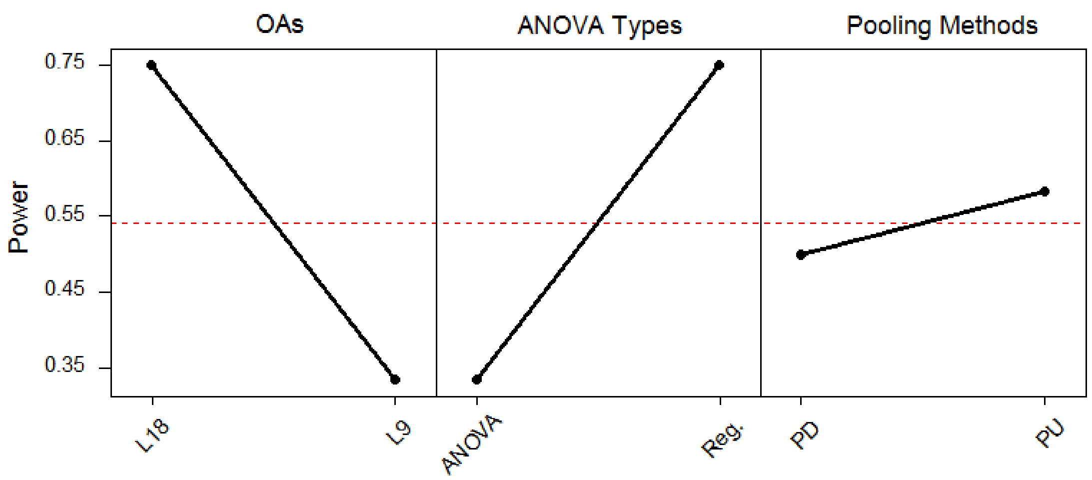


Figure 7-4: Power of each Fractional OA, ANOVA Type and Pooling Method in Phase-3 Validation

method. Although the computations were based on analysing a single full factorial experiment, the results depicted in Figure 7-4 provide evidence supporting the conclusions drawn from the simulation experiments conducted in Phase-3.

7.5 Phase-4

Two three-level full factorial experiments of 81 (3^4) runs were used to investigate the extent to which the findings of the phase-4 simulation experiments were valid. Main effects and two-factor interactions were the only active contrasts. The first involved the tool MRR data yielded in the EDM experiment and the second related to analysing the surface roughness data in a micro-end-milling experiment described by Tansel et al (2006). In the latter, four parameters, namely the cutting speed (A), feed rate (B), radial depth of cut (C) and the tolerance (D) were investigated. The objective was to assess their impact on the surface roughness of Aluminum 6061 blocks used to manufacture a critical part of a mould. The data resulting from this experiment are given in Appendix 7. After extracting the appropriate runs for the L_{18} and the L_9 OA both the BSS and the SWR methods were used to analyse the data. The effects pronounced significant using these strategies in conjunction with those that were declared active in the full factorial analyses are provided in Appendix 8. The full analysis of the EDM tool MRR data are presented in Appendix 9 . Using the results of the two experiments considered here, the average Power associated with each of the employed fractional OA and pooling method was computed and plotted in Figure 7-5. As was observed

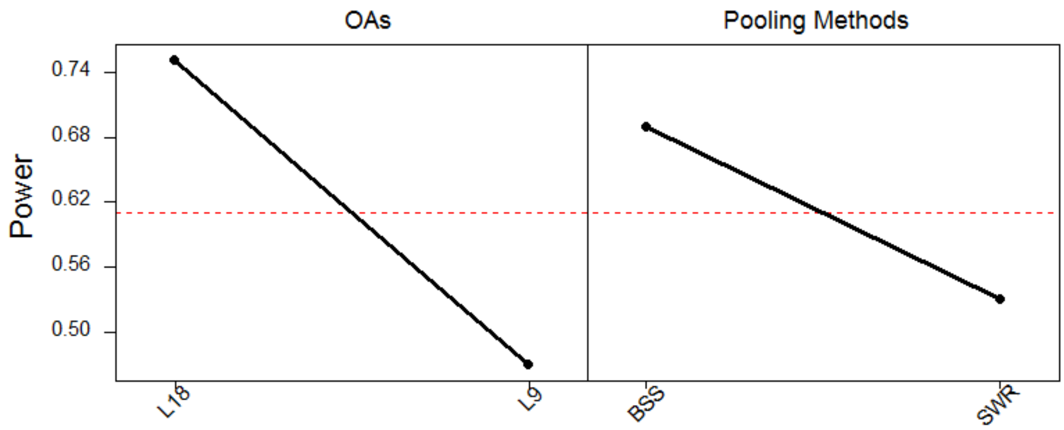


Figure 7-5: Average Power of each Fractional OA and Pooling Method in Phase-4 Validation

in this phase of the simulation experiments, the L_{18} attained a higher Power than the L_9 and the BSS procedure outperformed the SWR method. Consequently, the results from the experimental data in this phase are compliant with those that were obtained using the simulated data.

7.6 Summary

In this Chapter data from "real" large full factorial experiments were used to examine the credibility of the conclusions drawn from the simulation experiments conducted in this study. Published experimental data that reasonably matched the simulated circumstances were used togather with data generated from a large EDM experiment conducted in the laboratory . The conclusions of the "physical" experiments were graphically compared with those from the simulations. In general, the findings were found to be in reasonably good agreement. In fact, the Power patterns generated on the basis of the results of the "physical" experiments examined in phases 2, 3 and 4 conformed well to those produced using the simulation data. Some disagreements were experienced in the phase-1 analysis and their reasons discussed. As indicated by Balci (1998) the outcomes of validating a simulation study should not be considered as a binary variable where they are regarded as either absolutely correct or incorrect. Instead, the degree of their credibility should be assessed. From this perspective, the results of the performed simulation experiments can be regarded as having a reasonable degree of credibility. An interesting view of the validation decision that is apppropriate here is the one expressed by Neelamkavil (1987) when he stated that “true validation is a philosophical impossibility and all we can do is either to invalidate or ‘fail to invalidate’ ”. Drawing on the results of the analyses presented in this Chapter, there is not enough evidence to invalidate the simulation experiments’ inferences.

Several issues relating to the practice of DOE have been addressed so far in this thesis including aliasing, the use of unreplicated fractional designs and pooling methods. The next Chapter will deal with two further issues namely the use of p-values and effect sizes.

CHAPTER 8: P-VALUES AND EFFECTS' IMPORTANCE MEASURES

8.1 Introduction

Among the prime issues concerning the implementation of DOE in practice are those related to the use and interpretation of the p-value and the FE measures. As was revealed in the review undertaken in Chapter 3, the former was more often than not misunderstood and consequently misinterpreted. Also brought to light was the problem related to the use of the variance-related measures of the effects' importance. To this end, it was observed that two measures PC_1 (equation 3.1) and PC_2 (equation 3.2) were not only used interchangeably but also interpreted as if they were equivalent; the implication being that they perform equally well in estimating the FEs' explained variance.

These aspects cast some doubt on the extent to which the authors of the reviewed articles are acquainted with both the appropriate interpretation of the p-values and the variance-related FE measures and the concepts underpinning their use. The objective of this Chapter is to argue the purpose and the appropriate application of statistical significance tests and effects' importance measures. The p-value will be addressed first, then a discussion of the FE measures is presented. The main conclusions are discussed in the Chapter's closing section.

8.2 P-values Uses and Interpretation

As highlighted in Chapter 3, the misconceptions regarding the use and interpretation of the p-value can be classified into three categories: (i) the belief that its complement (1-p) reflects the degree of confidence associated with H_1 being true; (ii) its use as a measure of the likelihood that the results are due to random error or chance; and (iii) its use as a measure of the effect's importance.

8.2.1 P-value complement as a Measure of Confidence Level

As pointed out in Appendix 2, the p-value is the probability of rendering a Test Statistics' (TS) value as extreme as or more extreme than the observed one given that H_o is true. Symbolically, this can be represented as P (TS/ H_o true). Its use is based on Fisher's philosophy that "every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis" (Fisher, 1971). He also stated that "the null hypothesis is never proved or established, but is possibly disproved in the course of experimentation" (Fisher, 1971). It must be borne in mind that "disproving" per se is not possible in Hypothesis Testing (HT). In fact, when the resultant p-value is smaller than the designated significance level, two conclusions are possible. One is that H_o is true and a rare event has occurred and the other is that it is false and a non-zero effect exists. It is due to the former that a certain level of Type I error is accepted when an HT is performed.

One possible reason for using the complement of the p-value as an indicator of the degree of confidence that the alternative H_1 is true in practice is the belief that "disproving" H_o entails "proving" the alternative, i.e. if H_o is highly unlikely then H_1 is very likely. This follows from the fact that H_o and H_1 are mutually exclusive and exhaustive. However, such an interpretation would only be valid if the probability of H_o was known which is not the case in HT the only rendered probability being the p-value i.e. P (TS/H_o true). The latter is frequently confused with the P (H_o true/TS) and consequently incorrect conclusions could be drawn.

Another plausible explanation for misinterpreting the complement of the p-value stems from relating it to the confidence level that is conventionally used in the interval estimation of population parameters i.e. (1-α). This implies some confusion between the α-level and the p-value. The former is a prior conditional probability, P (Reject H_o/H_o true) that is specified before experimentation. Unlike the p-value, it relates to a nominal event (reject or fail to reject) rather than to a particular value of a TS. Thus, irrespective of the experimental results, its value remains fixed. It is due to this characteristic that the level of confidence (1-α) makes sense. In fact, the probability statement attached to this level should be understood in the long-run sense. More specifically, if an experiment is repeated a large number of times, then it is expected that the Type I error will be committed in α of the performed trials. Consequently, the confidence with respect to replicating the appropriate decision regarding H_o is 1- α. However, if a single experiment is performed and the p-value associated with its results is obtained, the Type I error is either committed or not. Thus its probability is either 0 or 1. As the p-value is a posterior probability estimated based on the results of a single experiment, its complement cannot be used to estimate the confidence level. This is because repeating any experiment would result in a different p-value. The truth or otherwise of H_o and the level of confidence regarding the decision made about it cannot change with replication.

8.2.2 P-value as a Probability that the Results are Due to Chance

A common misinterpretation of the p-value concerns using it as a measure of the odds that the observed effect is due to chance or random sampling. This is a fallacy because the p-value is calculated under the assumption that the sampling error is what caused the test statistic to depart from zero. Consequently, the likelihood of sampling error is already assumed to be 1 when the HT is performed. Put differently, when the MSE is used to compute, say, the t-statistic, it is taken as an estimate of the variance due to common or chance causes. Thus, interpreting the p-value as the odds that this caused the observed TS, implies a violation of an essential assumption that underpins the calculation of the TS and consequently the p-value.

8.2.3 P-value as a Measure of FE Importance

Another fallacy with regard to the interpretation of the p-value relates to its use as a numerical index of the FE individual or relative importance. As already noted its value is determined on the basis of the calculated TS. As an example of the latter, consider the t-statistic which is a function of both the SES and C_{ij} (see equation 4.13) i.e. $t = SES/(C_{ij})^{0.5}$. While its value gets larger as the SES increases, it is an inverse function of C_{ij} which increases as the degree of aliasing increases but decreases as the experiment size (n) increases. Consequently, for a certain value of the SES, several p-values may be observed depending on the aliasing level and n. Put simply, the p-value is a confounded measure of the FE importance. In fact, the use of the p-value should be confined to assessing the statistical significance of the observed TS, since the experiment size can seriously impact its level. To shed light on this, consider the simulated L_8 and L_{16} full factorial data given in Table 8-1. In both cases, the true significant effects were A, B, C, AB and BC. From the ANOVA shown in Table 8-2, it can be seen that the SESs of the true active effects in both analyses are identical. Their PCs are almost the same, however, while all the true active effects were declared significant at the 5% level in the L_{16} analysis, none were in the L_8. It is interesting to note that although they were all identified as inert in the latter case, their collective effect accounted for 89.1% of the response variability. Since the SESs were identical in both analyses, the only reason for the difference in their rendered p-values relates to the value of C_{ij} which reduces to 1/n in the absence of aliasing as was the case in the investigated two-level experiments. Thus, it was due to the difference in the sizes of the experiments that unequal p-values were observed. In general, as n gets larger, the t-statistic increases thereby reducing the p-value. Consequently, the impact of the size of the experiment must be considered when interpreting the p-value.

8.3 FE Importance Measures

When the null hypothesis concerning a certain FE is rejected, the implication is that an effect value of zero is very unlikely. However, even if a non-zero effect occurs, its size may

Table 8-1: Simulated Full factorial (a) L_8 and (b) L_{16} Data

No.	A	B	C	Res
1	-1	-1	-1	-5.03415
2	1	-1	-1	-17.1673
3	-1	1	-1	4.340672
4	1	1	-1	37.01398
5	-1	-1	1	17.66631
6	1	-1	1	22.81531
7	-1	1	1	-0.43875
8	1	1	1	29.74567

(a)

No.	A	B	C	D	Res
1	-1	-1	-1	-1	-12.0224
2	1	-1	-1	-1	-20.7173
3	-1	1	-1	-1	0.42144
4	1	1	-1	-1	34.16814
5	-1	-1	1	-1	17.24011
6	1	-1	1	-1	17.55249
7	-1	1	1	-1	-4.16578
8	1	1	1	-1	28.13705
9	-1	-1	-1	1	-9.95224
10	1	-1	-1	1	-11.4222
11	-1	1	-1	1	-1.36355
12	1	1	-1	1	22.27604
13	-1	-1	1	1	10.69849
14	1	-1	1	1	8.505269
15	-1	1	1	1	-3.98535
16	1	1	1	1	14.73723

(b)

be small enough to render it of no practical importance. Therefore, it is important to quantify the extent to which the effect diverges from zero. In fact, mere reliance on the HT cannot help to attain any scientific progress. As Yates stated the HT "has caused scientific research workers to pay undue attention to the results of the test of significance that they perform on

Table 8-2: ANOVA for the Simulated (a) L_8 and (b) L_{16} Data

Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**	**PC (%)**	**SES**
A	390.2315	1	390.2315	10.24	0.0853	14.44	1.13
B	342.9768	1	342.9768	9	0.0955	12.51	1.06
C	320.4928	1	320.4928	8.41	0.1012	11.58	1.03
AB	609.7365	1	609.7365	16	0.0572	23.45	1.41
BC	698.0872	1	698.0872	18.3184	0.0505	27.07	1.51
Residual	76.21706	2	38.10853				
Total	2437.742	7					
(a)							

Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**	**PC (%)**	**SES**
A	580.3999	1	580.3999	20.48	0.0011	14.54	1.13
B	510.117	1	510.117	18	0.0017	12.69	1.06
C	476.676	1	476.676	16.82	0.0021	11.81	1.03
AB	906.8746	1	906.8746	32	0.0002	23.15	1.41
BC	1038.281	1	1038.281	36.6368	0.0001	26.61	1.51
Residual	283.3983	10	28.33983				
Total	3795.747	15					
(b)							

their data and too little attention to the estimate of the magnitude of the effects they are investigating" (Yates, 1951). He warned that sole emphasis on the HT has led to the "unfortunate consequences" of regarding the performance of the significance test as the "ultimate objective" (Yates, 1951). Consequently, the execution of a significance test should be accompanied with estimates of the studied FE importance measures. Broadly, this can be accomplished by means of two types of measures: mean-related and variance-related. Regarding the latter, the review discussed in Chapter 3 concluded that there was some confusion regarding the difference between the PC_1 and PC_2 measures.

8.3.1 PC_1 and PC_2 Importance Measures

Originally, PC_1 (equation 3.1) was proposed by Fisher (1925) to supplement the significance test in the ANOVA under the label of the correlation ratio (η). On the other hand, Ross (1996) appeared to have coined the term Percentage Contribution when referring to the PC_2 (equation 4.2) measure. In practice, Ross's term is used to refer to both measures as they have been used interchangeably for the purpose of estimating the variance explained by each of the studied FEs.

To understand the major difference between PC_1 and PC_2, it is essential to distinguish between the population variability explained by the FEs and its corresponding sample estimator. Assuming that there are k FEs that relate to the response variable (Y) according to the following linear model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + e \tag{8.1}$$

the objective is to estimate the extent to which their consideration in the fitted model reduces uncertainty about Y; a prime measure of which is its variance σ^2. When the FEs are not considered, the variance of Y is σ^2_Y. Fitting a model relating the k FEs to the response Y, the conditional variance of Y is σ^2_e (assuming the response variance is homogeneous across the values of the FEs). The reduction in uncertainty resulting from relating FEs to the response Y can be represented by σ^2_Y - σ^2_e i.e. the difference between the overall and the conditional variance of Y. This can be expressed as a relative reduction by dividing it by σ^2_Y :

$$\rho = \frac{\sigma^2_Y - \sigma^2_e}{\sigma^2_Y} \tag{8.2}$$

Thus ρ represents a population measure of the Y explained variability due to the k FEs in equation 8.1. Clearly if the consideration of the k FEs eliminates the uncertainty in Y, σ^2_e will be zero and ρ will be 1.

One possible way of estimating ρ from a sample is to replace the population parameters in equation 8.2 with their corresponding sample statistics. For PC_1, this was performed using the general unbiased estimator of the population variance of a variable which can be obtained by dividing the Sum of Squares (SS) by (n-1). Substituting this for the σ^2_Y and σ^2_e in equation 8.2 yields

$$PC_1 = \frac{SST/(n-1) - SSE/(n-1)}{SST/(n-1)} = \frac{SST - SSE}{SST} \qquad (8.3)$$

where SST and SSE are the total and error sums of squares respectively and n is the size of the experiment. Since SST-SSE = $\sum_{i}^{k} SS_{FEi}$ = SS_{Model}, PC_1 can be obtained collectively for all the FEs under study i.e. for the fitted model. In the context of regression analysis, this is equivalent to the Coefficient of Determination, R^2. Being independently estimated in regular orthogonal designs, a separate PC_1 can be obtained for each FE by dividing its sum of squares by the SST so that

$$PC_{1\,Model} = PC_{1\,FE1} + PC_{1\,FE2} + \ldots.\ PC_{1\,FEk} \qquad (8.4)$$

Consequently, each PC_{1FEi} measures the FE_i independent contribution to the response variance and the amount by which the latter can be reduced by controlling that factorial effect. Whereas PC_{1FEi} is conventionally employed as a measure of the partial explained variance due to each FE tested in the ANOVA, R^2 is often used as an overall measure of the explained variance of the FEs that comprise the fitted regression model. Generally, PC_1 is a biased estimator of ρ. Barten (1962) showed that its bias is

$$\left[\frac{1-\rho}{n}\right]\left[(k+1)-(1-\rho)(1+2*\rho)\right] \qquad (8.5)$$

which is a function of the true size of ρ, the number (k) of FEs included in the model and the size of the experiment (n). The impact of ρ on the bias can simply be explained by considering the case where its value is 1. In such a situation, irrespective of k and n, all the sampled overall PC_1s (R^2) will be 1 (this can be verified by substituting 1 in equation 8.5).

With regard to k and n, it follows from equation 8.5 that increasing the former increases the bias while increasing the latter decreases it. Except when the relationship is perfect, PC_1 tends systematically to overestimate ρ. This can be illustrated by examining the case when ρ =0 which renders a bias of k/n; hence the bias is always positive since for any fitted model, k is greater than or at least equal to one.

Kelley (1935) proposed an alternative estimate of ρ which he believed to be unbiased. He suggested that unbiased estimators should be substituted for σ^2_Y and σ^2_e in equation 8.2. Although SST/(n-1) is an unbiased estimator of σ^2_Y, this is not the case for SSE/(n-1) which is a biased estimator of σ^2_e . The unbiased estimator of the latter is the MSE which, when substituted for σ^2_e in equation 8.2, yields what Kelley called epsilon squared (ε^2):

$$\varepsilon^2 = \frac{SST/(n-1) - SSE/(n-\nu-1)}{SST/(n-1)} \tag{8.6}$$

where ν is the model degrees of freedom i.e. $\nu = \sum_{i=1}^{k} df_{FEi}$ .This measure and Ross's PC_2 are numerically equivalent (see Appendix 10 for details). Moreover, Kelley's ε^2 is numerically equivalent to the Adjusted- R^2 that was proposed by Wherry (1931) to reduce the bias associated with using R^2 in regression analysis. Thus, both Kelley (1935) and Wherry (1931) independently proposed ε^2 and Adjusted- R^2 respectively as unbiased estimators of ρ. Although, these two estimators are the same, the former is conventionally used in the ANOVA to estimate the partial explained variance due to each FE under study whereas the latter is normally employed in the context of regression analysis to estimate the overall explained variance by all the FEs encompassed in the fitted regression model.

The equivalence between ε^2 and the PC_2 was not addressed in Ross (1996). In fact, he provided no reference to the discussion of the PC_2 derivation. Hence, it is not clear whether Ross (1996) independently arrived at PC_2. Indeed, he should be given the credit for popularising the use of this measure within the engineering community. He also seems to be the first author to suggest the use of a rule of thumb in interpreting it whereby the adequacy of the experimental endeavour can be judged. According to Ross (1996), if the value of PC_2 due to error ($1-\Sigma PC_{FEi}$) is "low" (15% or less) then the experiment is assumed adequate. However, if it is "High" (50% or more) then some important factors were omitted, the measurement error was excessive or the control of the experimental conditions was not sufficiently precise. No guidelines are provided to interpret the PC_2 when the error contribution lies between 15 and 50%. It is not clear why an error PC_2 of, say, 40% would be different from one of 50%; thus it must be borne in mind that these guidelines have no theoretical basis and hence should not be strictly adhered to.

Despite the use of unbiased estimators for the terms of equation 8.2, PC_2 is also a biased estimator of ρ under the standard least square assumptions. Barten (1962) estimated this bias to be

$$\frac{-\rho(1-\rho)(1-2\rho)}{n} \tag{8.7}$$

Clearly it is independent of the number of FEs under study and is an inverse function of the size of the experiment (n). It is also dependent upon the true value of ρ in that when it is equal to 0, 0.5 or 1 the bias is zero. In addition, the bias is positive when $\rho > 0.5$ and negative when $\rho < 0.5$. Its maximum is approximately (0.1/n) which occurs when ρ is approximately 0.2 and 0.8.

Although ρ cannot be negative, PC_2 can assume negative values; when this is the case the estimate of ρ is taken to be zero. As its value is quite small, the bias associated with PC_2 should be of little concern. In fact, Olkin and Pratt (1958) argued that any unbiased estimator of ρ has characteristics that make it inferior to certain biased estimators. For an estimator to be unbiased, its mean over an infinite number of estimates from independent random samples must be equal to the parameter estimated. Thus when $\rho = 0$, its unbiased estimator must be one which can be negative in some samples. This means that in order for the estimator to be unbiased it must assume values which the parameter cannot assume as ρ must always be greater than or equal to zero. When negative estimates are observed and taken to mean that ρ is zero, the estimation procedure no longer provides an unbiased estimate of ρ. Thus, in this case the unbiased estimator of ρ is not the best one.

Although, negative values of PC_2 are taken to be zero, the review conducted in Chapter 3 showed that they were dealt with in two different ways. In some studies, they were replaced with zeros whereas in others their associated FEs were pooled with the error estimate. The latter approach is more appropriate than the former because by pooling, the MSE is decreased thereby reducing the error percentage contribution. In fact, keeping the FE with a negative $PC_{FE(i)}$ in the developed model and replacing its estimate with zero leads to inflating the MSE. It is because of this that the value of the $PC_{FE(i)}$ should be used to aid in selecting the FEs that should be pooled to estimate the error. Its use for this purpose has two advantages. The first is that it alleviates the problem of reporting a negative $PC_{FE(i)}$, and the second is that it increases the degrees of freedom associated with the residuals and consequently increases the Power of the performed statistical test.

The chief point to be gleaned from the discussion presented in this section is that due to the smaller bias associated with it, PC_2 is more effective than PC_1 in estimating the FE explained variance. Consequently, PC_1 should be avoided. The use of PC_2 however is not without limitations. One common problem of variance-related measures is the fact that no information is provided regarding the amount by which the average response has changed. Another issue concerns the question of how large is a large effect. For example, if a FE explained 10% or 20% of the variability would it be regarded as a large, medium or small effect?. A related problem is the one concerning the absence of any theoretical or empirical justifications for such guidelines as those provided by Ross (1996) for judging the adequacy of the experimental inferences. Such problems may be alleviated by using a measure that incorporates elements of both the mean- and the variance-related importance measures such as the SES.

8.3.2 SES as a Measure of FE Importance

Before discussing the SES as a measure of the FE importance, the mean-related measures and their use in practice will be recalled. As highlighted in Chapter 3, these were used both numerically and graphically. An example of the former was the least squares estimate of a FE regression coefficient i.e. $\hat{\beta}_{FE(i)}$ (see Appendix 2). Its main advantage is that it is unbiased and has the smallest variance among the class of unbiased estimators. Besides allowing an examination of the FEs relative importance, it enables an evaluation to be made of both the plausibility of the experimental results and whether they make engineering sense. This can be done by assessing the consistency of the average response change per unit change of each FE with that expected using (whenever available) the theoretical subject knowledge. Furthermore, the signs of the estimated β_{FEi} allow the comparison of the direction of the average change resulting from each FE with its prior expectation to judge its validity.

The conducted review (Chapter 3) highlighted the fact that the numerical mean-related measures were rarely if ever interpreted in the studies that reported them. Typically, they were presented in a regression equation or in a summary table without commenting on their individual values or signs. One possible reason for this is the difficulty associated with justifying or interpreting the magnitude of the numerical mean-related measures in terms of the units of the response. In fact, it is rarely the case that there is a prior theoretical background regarding the extent to which the average response is expected to vary as a result of changing a FE by one unit. For example in an EDM experiment, it is possible to predict that increasing the current increases, say, the removal rate of a certain material and hence the sign of its coefficient (in the fitted model) should be positive. However, it is very difficult to have a prior knowledge regarding the average level by which it will change with an ampere change in the current. It is due to this that most practitioners prefer the use of the mean-related measures in their graphical form particularly in two-level experiments. By doing so, it is possible to efficiently obtain information regarding the FEs' relative importance and, to an extent, the feasibility of the experimental results judged respectively by the gradient and the direction of the lines in the main effects plot. However, when factors with more than two levels are studied this becomes more difficult and hence a great deal of subjectivity is needed.

As with the PC_2 limitations, the problem concerning the interpretation of the numerical mean-related importance measure can be overcome by using the SES (β_{FEi} /$(MSE)^{0.5}$). In fact, this measure is unique in the sense that it represents the ratio of a mean-related measure to an estimate that quantifies the common causes' variability i.e. the standard deviation of the experimental error. Consequently, the SES conveys information about both the average change and the variability of the response. Generally, assuming that the response variable is Normally distributed, it may be easier to express the average change in response per unit change of each FE in terms of numbers of standard errors. This enables one to relate the relative importance of a FE to the shift it can cause in the process average (upward or downward) and its consequent out-of-control state.

The SES has the advantage of being metric free so it does not depend on the response variable unit of measurement. Furthermore, its interpretation can be facilitated by relating it to the use of statistical quality control techniques such as control charts and process capability. For example in a variables mean-monitoring control chart the control limits are set at $\pm 3\sigma_{\bar{x}}$ from target where $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ and n is the sample size. Assuming that an experiment is conducted to study the impact of changing the settings of certain parameters (FEs) on the average of a specific performance measure (response), the $\sqrt{MSE}$ can be used as an estimate of the process standard deviation. Consequently, assuming that a control chart is being used to monitor the behaviour of the process under study using a subgroup size of, say, 5, and since $\sigma = \sigma_{\bar{x}} * \sqrt{n}$, a $SES_{FE(i)}$ of, say, 2 gives a value of approximately $4.5\sigma_{\bar{x}}$ which would lie beyond the control limit. Consequently the investigated FE may be regarded as having a large effect if its corresponding control chart value is larger than $3\sigma_{\bar{x}}$. This helps to answer the question of how large is a large effect. However, it results in classifying the observed SES into two categories: large or not large. Thus, all SES values that are translated to a control chart value greater (smaller) than $3\sigma_{\bar{x}}$ ($-3\sigma_{\bar{x}}$) are deemed equivalent.

Rather than relying on such a nominal conclusion, it is better to use a continuous scale in assessing the importance of the FEs. This can be accomplished by interpreting the SES in terms of a process capability index such as C_{pk}, the proportion of the response values expected to fall beyond the process's specification limits or the Taguchi loss function. In fact, given a target performance value and a specification/tolerance interval for a response variable, defined by the customers' needs and expectations, it is straightforward to assess the impact of a $SES_{FE(i)}$ of a certain value on these measures. As an illustration consider the following example. In a process for drilling steel plates, suppose that the hole diameter was the response under study with a target of 6.35 mm and a tolerance of ± 0.20 mm and that its measurements are Normally distributed. Assume further that

(i) the process can be modelled using equation 8.1 and that an experiment was conducted to identify the most influential FEs on the hole diameter size. In this case the $\sqrt{MSE}$ of the conducted experiment can be used as an estimate of the process standard deviation.

(ii) the FEs of feed rate, cutting speed, tool material and tool geometry were the important main effects and that the estimated $\sqrt{MSE}$ was 0.05.

(iii) the process under study was centred on its target as shown in Figure 8-1(a).

C_{pk} can be used to assess the process performance in terms of the deviation of its average from its target value and is given by

$$C_{pk} = \frac{Min[(USL - \bar{X}) \text{ or } (\bar{X} - LSL)]}{3\sigma} \quad (8.8)$$

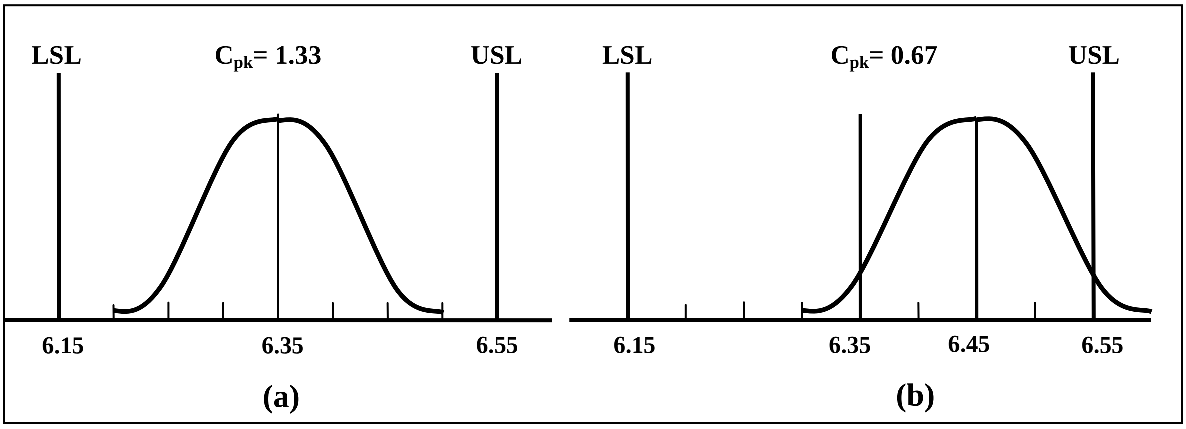


Figure 8-1: Capability Index for the Hole Drilling Process

where USL and LSL are the process upper and the lower specification limits respectively. For the process in its initial state (Figure 8-1 (a)) C_{pk} is 1.33 i.e. (0.2/(3*0.05)) which is conventionally deemed satisfactory. Suppose that the $SES_{FE(i)}$ of the cutting speed was 2. This means that this FE has the potential to shift the process average by 2σ towards the USL yielding a C_{pk} value of 0.67. Consequently, increasing the cutting speed by one m/min. has the impact of changing the process C_{pk} from 1.33 to 0.67 as illustrated in Figure 8-1 (b).

The new C_{pk} associated with a specific absolute value of $SES_{FE(i)}$ (see Appendix 10) can be computed from the conventional one using the following equation:

$$\text{New } C_{pk} = C_{pk} - (SES_{FE(i)}/3) \qquad (8.9)$$

Alternatively the cutting speed $SES_{FE(i)}$ can be interpreted in terms of the potential percentage of holes that will be drilled with a diameter larger than the USL (6.55 mm) i.e. the percentage of nonconforming holes. This can be illustrated by considering the process at its initial state, where its average is centred at its target value and is 4σ away from each of the specification limits. In this case the percentage of nonconforming holes is approximately 0.006% ($P(z > 4$ & $z < 4)$ where z is the score of a Standard Normal distribution). As its $SES_{FE(i)}$ is 2, increasing the cutting speed by one m/min increases the percentage of nonconforming holes to 2.275% ($P(z > 2)$). Of course, this can be translated into monetary terms by multiplying the total number of produced units by the percentage of nonconforming holes to yield the expected number of defects and then multiplying this by the unit cost, rendering the total loss associated with the estimated $SES_{FE(i)}$. This loss can also be arrived at using Taguchi’s loss function (equation 2.3). In fact, assuming that the process average is centred at its target value the loss function associated with certain $SES_{FE(i)}$ can be rewritten as follows:

$$L(Y) = K[MSE + ((SES_{FE(i)})^2 * MSE)] \qquad (8.10)$$

The expression of the FE importance in monetary terms enables ones to precisely evaluate the value and the possible gain associated with controlling and changing the concerned FE. It facilitates the interpretation of the FE in the sense of making the process of attaching such attributes as “large” or “small” to a FE more objective and presentable. In short, the use of SES as a measure of a FE’s importance has the advantage of making its relative importance easier to understand and communicate.

A noteworthy aspect regarding the use of the $SES_{FE(i)}$ relates to the bias associated with estimating the process standard deviation. Assuming that the process can be modelled using equation 8.1, the experiment MSE is an unbiased estimator of its variance. However, its square root is a biased estimator of the process standard deviation. Therefore, despite the fact that the numerator of the $SES_{FE(i)}$ is an unbiased estimator of the mean-related effect, its denominator is biased rendering it a biased estimator of its population parameter. However, as presented in Montgomery (2009), a bias correction factor, conventionally referred to as c_4, can be used to obtain an unbiased estimator of the population standard deviation i.e. $E(S/c_4)= \sigma$ where σ is the population standard deviation and S is its sample estimator.

The correction factor c_4 is tabulated against various values of n (Montgomery, 2009). Before being able to use it to correct for the bias associated with the MSE, its formula needs to be rearranged so that it is of the same form as the conventional standard deviation i.e. $(SS/n\text{-}1)^{0.5}$. Generally

$$\sqrt{MSE} = \sqrt{\frac{SSE}{n-\nu-1}} \qquad (8.11)$$

The number of degrees of freedom associated with the MSE is smaller than that associated with the general sample variance as the estimation of both the average and the FEs render less degrees of freedom for estimating the former. Substituting n_c for n - ν in equation 8.11 yields a MSE that is equivalent in form to $(SS/n\text{-}1)^{0.5}$. n_c should now be used to determine the appropriate value of c_4. Substituting the corrected estimate of the process standard deviation in the $SES_{FE(i)}$ equation yields an adjusted version of it i.e.

$$\text{Adj-SES}_{FE(i)} = \frac{\hat{\beta}_{FE(i)}}{\dfrac{\sqrt{MSE}}{c_4}}. \qquad (8.12)$$

In fact, even when both the numerator and the denominator of the $SES_{FE(i)}$ are unbiased, their ratio is biased i.e. E(X÷Y) ≠ E(X) ÷ E(Y). According to Chiang (1966), E(X÷Y) = $E(X)*E(\frac{1}{Y}) > E(X) \div E(Y)$. Consequently, the estimated $SES_{FE(i)}$ (both the original and the adjusted) tends to overestimate the true value. This is not only due to the aforementioned mathematical relationship but also to the fact that the expected value of the experimental MSE may well be smaller than the process variance. This is due to the difficulty of rigidly controlling all the process's sources of variation in practice as can be the case in a laboratory experiment. Although this should not affect the relative importance of the FEs under study as each $\hat{\beta}_{FE(i)}$ is divided by the $\sqrt{MSE}$, it is important to bear in mind that the true $SES_{FE(i)}$s are likely to be smaller than the observed ones when interpreting them in terms of falling beyond the limits of a control chart, C_{pk} or Taguchi loss function.

8.4 Summary

This Chapter aimed at presenting the concepts underlying the use and interpretation of the p-values and the FE importance measures. The major misconception concerning the interpretation of the p-values in practice has been pinpointed and the fact that it has been given more meaning than is warranted revealed. In this context, the appropriate interpretation of the p-value was emphasized and the necessity of using the FE importance measures as instructive adjuncts to the HT procedures was ascertained.

The examination of the use of the effect's measures in practice draws attention to a lack of familiarity regarding the difference between the PC_1 and PC_2 estimates. This is evident from their interchangeable use and interpretation. Though both are biased estimators of the explained variance, the PC_2 bias is far less serious than that of the PC_1. Consequently, the latter use should be avoided.

Although popularised by Ross (1996) PC_2 was originally derived by Wherry (1931) under the title of Adjusted- R^2 and Kelley (1935) who called it ε^2. Apparently, many practitioners attribute this measure to Ross (1996) as is evident from the frequent use of his guidelines in judging the adequacy of the manufacturing experimental results. In fact, Ross seems to be unique in emphasising the importance of this variance-related measure and providing guidelines relating to its use in the assessment of the performed experiment adequacy in the engineering arena. Other well known engineering experimental design books such as Wu and Hamada (2000), Box et al (2005) and Montgomery (2010) pay little attention to such a measure but prefer the absolute mean-related important measure, especially in the context of two-level experiments. R^2 and Adjusted- R^2 are only discussed as an overall measure of all the FEs in the context of regression analysis. This may be one of the main reasons for the inconsistent use of the FEs' importance measures in the Manufacturing Engineering literature as different references emphasise different importance measures.

In using the variance-related importance measure, it is difficult to answer the question how large is a large effect. Put differently, the translation of a certain value of explained variance in monetary terms or a language understandable by practitioners is not straightforward. To this end, the use of the SES measure is recommended. Though common in simulation studies, the SES is rarely, if ever, employed in practice. This should not be surprising in light of the absence of its use and discussion in both conventional and Taguchi's DOE textbooks. The chief advantage of the SES stems from the ability to link its interpretation to such quality engineering techniques as control charts, process capability and Taguchi's loss function. This facilitates the expression of the FE importance in shop floor terminology thereby improving the communication between Engineers, Managers and Statisticians.

Of note regarding the use of the SES and the variance-related importance measures is the fact that each is a biased estimate. Generally, the problems relating to bias present no difficulty as long as its presence is recognised and an appropriate correction factor is known. The bias associated with the variance related measures is not only recognised but also quantified in the statistics literature. The dependence of the bias on the true population value of these measures renders the estimation of their correction factor infeasible. On the other hand, despite the recognition of the bias associated with the SES, there is still a need for more studies to quantify its bias and investigate the possibility of reducing or even eliminating its impact using an appropriate correction factor.

Being point estimates, there may be no knowledge of the sampling error of PC_1, PC_2, or the SES. Thus, in contrast to the mean-related measure $\hat{\beta}_{FE(i)}$ for which the sampling distribution is known, no confidence interval can be constructed to gain an idea about the sampling variability associated with such measures. Consequently, when interpreting them it must be borne in mind that their sampling error may be high, especially when the number of experimental runs is small, rendering their reliability dubious. Thus it is recommended, whenever possible, to replicate the experiment and compare the FEs importance measures estimated from each replicate with each other to gain some idea about these measures' sampling variability. It is also recommended that further research be undertaken to investigate and establish the sampling distribution of PC_1, PC_2, and the SES.

CHAPTER 9: CONCLUSIONS AND FUTURE WORK

9.1 Conclusions

The industrial world has been experiencing unprecedented social and technological changes, and it seems that more profound changes lie ahead. Recently, it has been established that to have an edge over competitors, more effective decision making processes must be adopted. Davenport (2006) argued that owing to advances in information technology, communication and reverse engineering, companies' vulnerability to product imitations has considerably increased. He observed that, in many industries, companies are providing similar products and employ comparable technologies. Moreover, what used to be regarded as a distinct competitive advantage can now be easily learned and copied. Drawing on this, Davenport and Harris (2007) argued that, in this era, to effectively differentiate itself, a company needs to compete on analytics, i.e. "extensive use of data, statistical and quantitative analysis, explanatory and predictive models and fact-based management to drive decisions and actions". To this end, DOE is among the most effective statistical techniques for understanding the cause-and-effect relationships between process and product parameters and their performance measures thereby generating the required knowledge to inform any performance improvement decision in a timely and cost effective manner.

Owing to the vital role of Manufacturing Engineering in any industry, the aim of this research was to improve the practice of DOE in this field. While the literature on how DOE should be applied abound, the manner in which it is actually being employed in practice is rarely, if ever, addressed. Though alluded to in research by Antony et al (1998) and Makrymichalos et al (2005) for example, no study that explicitly discusses the topic using empirical evidence was found. In this work, an attempt was made to fill this void by reviewing its practice in a sample of three manufacturing journals. The intention was to examine the extent to which the methods suggested in the published DOE literature are being applied by practitioners.

One problem that was encountered in the conducted review was the fact that the use of DOE was limited, thereby highlighting the need for improving the extent of awareness of the importance and capabilities of this powerful technique. The bulk of the second Chapter in this thesis was therefore devoted to this end. Several aspects of the DOE power were explored including

- its effectiveness in characterising and optimising product and process parameters,
- its ability to deal with the main two dimensions of complexity i.e. interdependence and variability,
- the ease with which it can be integrated with other process improvement techniques,
- its potential for improving the formal decision making process as well as the scientific investigation endeavour.

Merely being aware of the importance of DOE is not enough; it takes desire and ability to successfully employ any improvement technique. Consequently, it is essential to understand the appropriate way of implementing DOE and mitigating the impediments to its effective use. The main problems associated with DOE practice identified in the review were as follows:

(i) a lack of familiarity with the concept of aliasing, the method of generating its pattern and the consequences of overlooking its impact;

(ii) the use of fractional factorial designs and pooling methods to analyse unreplicated experiments;

(iii) the misconceptions and misunderstanding of the fundamental foundations that underpin the use and interpretation of the p-value and the FE importance measures.

Regarding aliasing, the ramifications of failing to recognise its impact were discussed, including its influence on the estimated factorial effects and their standard error in addition to the effect it exerts on the employed test statistic and the statistical significance outcome. Also addressed were the main criteria for measuring the degree of aliasing, including maximum resolution, minimum aberration and generalised minimum aberration. It was highlighted that these were seldom used in practice due, in no small part, to the complexity associated with their presentation in the DOE literature. To deal with this, a simple method for generating the aliasing pattern was proposed. Its main advantage lies in the ease with which it can be implemented. In fact, it requires nothing more than applying the basic matrix operations using Microsoft Excel functions. Another advantage of the suggested method stems from the fact that it can be linked to the aforementioned criteria for measuring aliasing in a manner devoid of mathematical complications. Moreover, in addition to its suitability for two- and three-level designs, the proposed method can be used to handle both regular and non-regular OAs. By drawing attention to the concepts of aliasing and the implications of ignoring its consequences and providing a simple method for dealing with it, it is hoped that the aliasing related issues in the practice of DOE will be alleviated.

With respect to the strategies of experimentation, the main question that was posed in this research related to the extent to which the conclusions arrived at using a certain fractional factorial OA and a pooling method in an unreplicated trial match those that would have been drawn had a full factorial experiment been performed. The question was addressed by designing and analysing a set of simulation experiments using DOE techniques. The percentage of effects that were correctly detected by the employed fractional OA and pooling method as identified in the full factorial trial i.e. the Power, was used as the response variable. The employed OAs and the applied pooling methods were treated as controllable factors whereas the number of active effects and their sizes were dealt with as noise factors. This allowed not only the average Power to be examined but also its variability around its average. In other words, the objective was not only to maximise the average Power but also to seek the settings of the controllable factors that desensitised the impact of the noise factors on it.

The scope of the conducted experiments incorporated both two- and three-level OAs. In the DOE literature, the latter receive nothing like the attention given to the former. Moreover, whenever the topic of Power is discussed, its average was the main concern. In fact, the Power variability has been accorded scant attention. Thus, one of the purposes of this study was to redress the balance by considering both types of OAs and examining the Power variability (robustness) as well as its average.

The results of the conducted simulation experiments revealed that in the context of two-level trials, the highest Power was attained when the L_{16} was used in conjunction with the PU method and the HNP plot. In terms of pooling methods, the latter is more informative than all the others for it allows the relative magnitude of the studied effects to be investigated visually. Moreover, it has the advantage of highlighting the need for data transformation. When the time and resource limitations entail performing less than 16 trials, the use of the L_{12} is suggested. In this case the appropriate pooling method should be specified on the basis of process knowledge. If two-factor interactions are suspected, then BSS should be employed, otherwise the use of the HNP plot and the PU method is recommended. Whenever possible, the use of the L_8 should be avoided. This is due to its low and inconsistent level of Power.

Moreover, if augmenting is necessary, it is advisable to use BM as it outperformed the DO method.

In the case of three-level experiments, irrespective of which OA is used, it is recommended that each of the OA's columns is decomposed into single degree of freedom components using the L-Q system and the data analysed using regression. This allows for better utilisation of the available degrees of freedom thereby increasing the Power to detect the true active effects. In terms of the OAs, the L_{18} attained a high level of Power especially when analysed using BSS which outperformed SWR. Not surprisingly, the L_9 suffered from the same pitfalls as the L_8.

To attain Power robustness, different approaches should be adopted with the two investigated controllable factors. For the OAs, the use of the L_8 and the L_9 should, as far as possible, be avoided. With regard to the pooling methods, it is always recommended that more than one technique is used to analyse the experimental data and the reasons for any disagreements investigated. In fact, if several methods detect the same effects as active, the confidence regarding their significance is increased. On the other hand, when certain effects are pronounced active only by some of the pooling methods, the subject knowledge should be used to examine the practical plausibility of their significance. Otherwise, one-factor-at-a-time confirmation trials may be performed varying only the settings of the suspected parameters. Adopting these approaches should not only increase the average Power but also reduce its variability.

Among the prime issues that were revealed in the conducted review were those related to the misconception and the misuse of p-values and the FE importance measures. Regarding the former, it was observed that it was used to convey more meaning than warranted; this included the treatment of its complement (1-p) as a quantity reflecting the degree of confidence that H_1 was true, its use as a measure of the likelihood that the result was due to chance, and regarding its value as reflective of the effect importance. The pitfalls associated with each of these interpretations were discussed and the appropriate use of the p-value as a probability of obtaining a statistic value as extreme as or more extreme than the observed one was emphasised.

With respect to FE importance measures, it was noted that the two variance related measures PC_1 and PC_2 were not only used interchangeably in the manufacturing literature but also were interpreted as equivalent. After examining their roots and the rationale behind their derivation, it was noted that both are biased estimators of the FE explained variance. Moreover, the bias associated with PC_2 was demonstrably smaller than that with PC_1. Consequently, the use of the former is suggested. Being reflective of the change caused by the FE in both the response average and its variability, it was recommended that the SES be used to measure the effect's importance. The main advantage of this measure lies in the possibility of relating its interpretation to such quality engineering techniques as control charts, process capability and the Taguchi loss function. By so doing, communication between Engineers, Managers and Statisticians can be greatly improved.

Promoting the proficient practice of DOE requires a multifaceted approach incorporating textbook authors, software packages' developers and journal editors. Drawing on the results of this study, one is entitled to infer that the DOE literature has lacked a comprehensive textbook on understanding the impediments of the proper application of this technique and developing practices for successfully mitigating them. In fact, one of the major reasons for the inappropriate practices of the DOE is the textbooks' unbalanced coverage of topics such as graphical and formal data analysis techniques, pooling methods, two-and three-level OAs and their data analysis, FE importance measures and the proper interpretation of significance testing inferences. To gain a rigorous understanding of these, it is necessary to refer to more than one textbook. The reason being that while some of them are emphasised in certain texts, others may be completely ignored. Moreover, some textbooks adopt Taguchi's approach to DOE and overlook all the powerful techniques presented in the conventional literature on the subject. To deal with these issues, there is a need for a comprehensive text that integrates the strengths of the conventional approach with those of Taguchi's without losing sight of the problematic issues highlighted in this study and their mitigation.

The developers of software packages have a responsibility to assist practitioners in implementing best practices in DOE. Most of the widely used packages do not provide, in menu-driven interfaces, many of the powerful DOE statistical analysis techniques such as the BSS method and the OAs' L-Q and other coding procedures. Moreover, the FE importance measures such as PC_2 and SES are not provided as a part of the data analysis options in such packages as Minitab, Design-Expert or SPSS which were found to be widely used in practice. The inclusions of the above mentioned procedures and options will not only improve the practice of DOE but also encourage its adoption.

As the gatekeepers for what appears in publications, journal editors have an essential role to play in improving the practice of DOE. They should be aware of its importance in nourishing scientific research, and be knowledgeable about its appropriate practices and ensure that the authors adhere to them. A pertinent aspect worth mentioning here is what was brought to light in the conducted review regarding the limited citation of DOE references in the examined articles. This is reflective of the fact that in learning how to apply DOE, practitioners appear to rely on articles that deal with similar case studies to the one under consideration. As these are normally published in the Manufacturing Engineering journals, they are preferred over the theoretical statistically-centered papers presented in the DOE literature. To a large extent, this can be attributed to the practitioners prevailing tendency to adopt cookbook approaches in which the statistical and philosophical framework out of which the implemented methodology emerge is ignored. However, it also calls attention to a lack of communication between practitioners and researchers in the field of DOE. In fact, it seems that most of the DOE publications are addressed to other researchers in this field, not to practitioners. This is evident from the excessive use of jargon and mathematical complication in the DOE literature. It should be conceivable that Manufacturing Engineers may not fully understand the DOE concepts that they rarely read about in their literature. To attain an effective communication, DOE researchers should publish their findings in applied journals such as Manufacturing Engineering ones. The manner in which their material is presented should be readily understandable. As DOE is not a static field, this can help increase the awareness of the new DOE developments thereby enhancing the way in which this technique is applied.

9.2 Limitations and Future Work

As being the case in any research, the results of this work should be considered without losing sight of its limitations. The topics of DOE that warranted further illustration and assimilation were pinpointed in this study based on reviewing one year's publications in three journals listed under the category of Manufacturing Engineering in the citation reports issued by Thompson Reuters. This was due to a time constraint that rendered reviewing all the journals listed under the appropriate category a prohibitive task. This issue however, should not have a tremendous bearing on the main objectives of this research, the principal reason being that no attempt has been made to generalise the findings of the conducted review to all Manufacturing Engineering journals. Rather, the intention was to use it to represent an important segment of the Manufacturing Engineering literature as a platform for issues that are crucial to enhance the understanding and practice of DOE in this field. Moreover, although based on reviewing only three journals, it did highlight several gaps in the DOE literature. These included (i) the need for a simple and unified method for generating the aliasing pattern for both the two- and three-level OAs and (ii) the absence of empirical performance evaluation studies that, besides aiming to identify the fractional DOE strategy that maximises Power, address the problem of robustness against such noise factors as the number of active effects and their sizes. Further research on the practice of DOE in all Manufacturing Engineering journals is needed over a longer period. Such investigations are likely to highlight other gaps in the DOE literature.

With respect to the simulation experiments, it must be borne in mind that they were fixed-effect trials and consequently their conclusions cannot be generalised. This points to the possibility of expanding this study in the future by exploring further two- and three-level OAs, pooling methods and number and sizes of active effects. The response variable of interest in the performed trials was Power but it would be interesting to undertake similar simulation experiments to study the impact of the investigated factors on the Type I error.

In the context of the FE importance measures, there is a need for a thorough study of the bias associated with the SES measure and the possibility of eliminating or even reducing its impact using an appropriate correction factor. Equally important is the need for studying the sampling distribution of not only the SES but also the PC_2. In this regard, it must be noted that in estimating the sampling error and constructing the confidence interval of any statistic, the bulk of the statistics literature focuses on the central sampling distributions which assume a zero effect. The distributions of the SES and the PC_2 are, however, non-central in the sense that besides specifying the degrees of freedom, their construction entails identifying the value of the effect. Therefore, for each effect value there is a specific distribution. Research on developing algorithms for constructing such distributions and examining the possibility of incorporating them into the standard DOE software packages is called for. By casting light on the sampling variability of the SES and the PC_2 and reducing the sole reliance on their point estimates, the outcome of such studies is likely to contribute significantly to the enhancement of DOE practice.

Finally, it is hoped that this research will help practitioners gain a better insight into how to employ this powerful technique and pave the way for further research aimed at improving its practice.

REFERENCES

Goh, T. (2000), "Operating Frameworks For Statistical Quality Engineering", *International Journal of Quality and Reliability Management*, Vol. 17, No**.** 2, pp. 180-188.

Antony, J. (2001), "Improving The Manufacturing Process Quality Using Design Of Experiments: A Case Study", *International Journal of Operations and Production Management*, Vol. 21, No**.** 5/6, pp. 812-822.

Finkin, E. F. (1987), *Successful Corporate Turnarounds: A Guide For Board Members, Financial Managers, Financial Institutions, And Other Creditors,* Greenwood Publishing Group.

Fisher, R. (1971), *The Design Of Experiment*, Hafner Press.

Elias, J. (2007), *Science Terms Made Easy: A Lexicon Of Scientific Words And Their Root Language Origins,* Greenwood Publishing Group, Inc.

Montgomery, D. (2010), *Design And Analysis Of Experiments,* Wiley

Antony, J. and Kaye, M. (2000), *Experimental Quality: A Strategic Approach To Achieve And Improve Quality,* Kluwer Academic Publishers.

Anderson, V. and McLean, R. (1974), *Design Of Experiments: A Realistic Approach,* CRC.

Barrentine, L. (1999), *An Introduction To Design Of Experiments: A Simplified Approach,* American Society for Qualit.

Goh, T. and Xie, M. (2004), "Improving On The Six Sigma Paradigm", *The TQM Magazine*, Vol. 16, No**.** 4, pp. 235-240.

Hicks, C. and Turner, K. (1999), *Fundamental Concepts In The Design Of Experiments,* OUP USA.

Mintzberg, H. (1994), "That's Not "Turbulence," Chicken Little, It's Really Opportunity", *Strategy & Leadership*, Vol. 22, No**.** 6, pp. 7-9.

Ackoff, R. (1999), *Re-Creating The Corporation: A Design Of Organizations For The 21st Century,* Oxford University Press.

Ackoff, R., Ackoff, R., Emery, F. and Ruben, B. (2005), *On Purposeful Systems: An Interdisciplinary Analysis Of Individual And Social Behavior As A System Of Purposeful Events,* Aldine Transaction.

Simon, H. A. (1976), *Administrative Behavior,* Macmillan.

Solberg, J. (1988), "Integrated Manufacturing Systems: An Overview", in Compton, W. D. (Ed.), *Design and Analysis of Integrated Manufacturing Systems*, National Academy Press, Washington, DC, pp. 1-6.

Hunter, W. G. (1986), "Managing Our Way to Economic Success: Two Untapped Resources", *Center for Quality and Productivity Improvement*.

Hines, W. W., Montgomery, D. C., Goldsman, D. M. and Borror, C. M. (2003), *Probability And Statistics In Engineering And Management Science*, Wiley.

Montgomery, D. C. and Runger, G. C. (2010), *Applied Statistics and Probability for Engineers* Wiley.

Bisgaard, S. (1990), "Quality Engineering and Taguchi Methods: A Perspective", *Center for Quality and Productivity Improvement*.

Box, G. and Bisgaard, S. (1987), "The Scientific Context of Quality Improvement", *Center for Quality and Productivity Improvement*.

Goh, T. (1993), "Taguchi Methods: Some Technical, Cultural And Pedagogical Perspectives", *Quality and Reliability Engineering International*, Vol. 9, No**.** 3, pp. 185-202.

Grant, E. and Leavenworth, R. (1996), *Statistical Quality Control*, McGraw-Hill.

Hoerl, R. and Snee, R. (2002), *Statistical Thinking: Improving Business Performance*, Duxbury-Thomson Learning, Pacific Grove, CA.

Goh, T. (1996), "Economical Experimentation via ‘Lean Design’", *Quality and Reliability Engineering International*, Vol. 12, No**.** 5, pp. 383-388.

Goh, T. (1989), "An Efficient Empirical Approach To Process Improvement", *International Journal of Quality & Reliability Management*, Vol. 6, No**.** 1, pp. 7-18.

Kacker, R. (1985), "Off-Line Quality, Parameter Design, And The Taguchi Method (with discussions)", *Journal of Quality Technology*, Vol. 17, No**.** 4, pp. 176–209.

Gershenson, C. and Heylighen, F. (2005), "How Can We Think Complex?", in Richardson, K. (Ed.), *Managing Organizational Complexity: Philosophy, Theory And Application*, Information Age Pub Inc, pp. 47-62.

Simon, H. (1957), *Models Of Man: Social And Rational: Mathematical Essays On Rational Human Behavior In A Social Setting*, Wiley

Ceder, A. (2007), *Public Transit Planning And Operation: Theory, Modelling And Practice*, Butterworth-Heinemann.

Goh, T. (1999), "Perspectives On Statistical Quality Engineering", *The TQM Magazine*, Vol. 11, No**.** 6, pp. 461-466.

Britz, G. (2000), *Improving Performance Through Statistical Thinking,* American Society for Quality.

John, J., Whitaker, D. and Johnson, D. (2001), *Statistical Thinking For Managers,* CRC Press.

ASQ, S. D. (1996), *Statistical Thinking,* American Society for Quality.

Snee, R. (1990), "Statistical Thinking And Its Contribution To Total Quality", *American Statistician*, Vol. 44, No**.** 2, pp. 116-121.

Montgomery, D. (1999), "Experimental Design For Product And Process Design And Development", *Journal of the Royal Statistical Society: Series D (The Statistician)*, Vol. 48, No**.** 2, pp. 159-177.

Harvey, N. (2001), "Studying Judgement: Models And Methods", *Thinking & Reasoning*, Vol. 7, No**.** 1, pp. 1-3.

Box, G. (1993), "Quality Improvement: The New Industrial Revolution", *International Statistical Review*, Vol. 61, No**.** 1, pp. 3-19.

Bouyssou, D., Marchant, T., Pirlot, M., Perny, P., Tsoukias, A. and Vinche, P. (2000), *Evaluation And Decision Models: A Critical Perspective,* Springer.

Britz, G., Emerling, D., Hare, L., Hoerl, R. and Shade, J. (1997), "How To Teach Others To Apply Statistical Thinking", *Quality Progress,* Vol. 30, No**.** 6, pp. 67-80.

Makrymichalos, M., Antony, J., Antony, F. and Kumar, M. (2005), "Statistical Thinking And Its Role For Industrial Engineers And Managers In The 21st Century", *Managerial Auditing Journal*, Vol. 20, No**.** 4, pp. 354-363.

Deming, W. (1986), *Out Of the Crisis,* MIT Center for Advanced Engineering Study.

Hare, L., Hoerl, R., Hromi, J. and Snee, R. (1995), "The Role Of Statistical Thinking In Management", *Quality Progress*, Vol. 28, No**.** 2, pp. 53-62.

Stupak, R. and Leitner, P. (2001), *Handbook Of Public Quality Management,* CRC Press.

Taguchi, G. and Wu, Y. (1985), *Introduction To Off-Line Quality Control,* Central Japan Quality Control Association.

Taguchi, G., Elsayed, E. and Hsiang, T. (1989), *Quality Engineering In Production Systems,* McGraw-Hill College.

Law, A. and Kelton, W. (2000), *Simulation Modeling And Analysis,* McGraw-Hill

Taha, H. (2002), *An Introduction to Operations Research,* Prentice Hall.

Banks, J. (1999), "Introduction To Simulation", in Farrington, P. A., Nembhard, H. B., Sturrock, D. T. and Evans, G. W. (Ed.), *Proceedings of the 31th Winter Simulation Conference*

Bradley, S., Hax, A. and Magnanti, T. (1977), *Applied Mathematical Programming,* Addison-Wesley.

Banks, J. (1998), *Handbook Of Simulation,* Wiley, New York.

Kleijnen, J. (2008), *Design And Analysis Of Simulation Experiments,* Springer Verlag.

Sacks, J., Welch, W., Mitchell, T. and Wynn, H. (1989), "Design And Analysis Of Computer Experiments", *Statistical science*, Vol. 4, No**.** 4, pp. 409-423.

Welch, W., Yu, T., Kang, S. and Sacks, J. (1990), "Computer Experiments For Quality Control By Parameter Design", *Journal of Quality Technology*, Vol. 22, No**.** 1, pp. 15-22.

Kleijnen, J. and Standridge, C. (1988), "Experimental Design And Regression Analysis In Simulation: An FMS Case Study", *European Journal of Operational Research*, Vol. 33, No**.** 3, pp. 257-261.

Kleijnen, J. (1995), "Sensitivity Analysis And Optimization Of System Dynamics Models: Regression Analysis And Statistical Design Of Experiments", *System Dynamics Review,* Vol. 11, No**.** 4, pp. 275-288.

Kleijnen, J. (2005), "An Overview Of The Design And Analysis Of Simulation Experiments For Sensitivity Analysis", *European Journal of Operational Research*, Vol. 164, No**.** 2, pp. 287-300.

Goh, T. (1992), "An Organizational Approach To Product Quality Via Statistical Experiment Design", *International Journal of Production Economics*, Vol. 27, No**.** 2, pp. 167-173.

Sage, A. and Armstrong, J. (2000), *Introduction To Systems Engineering,* Wiley, New York.

Antony, J. (2003), *Design Of Experiments For Engineers And Scientists,* Butterworth-Heinemann.

Cornell, J. (1990), *How To Apply Response Surface Methodology,* American Society for Quality.

Myers, R., Montgomery, D. and Anderson-Cook, C. (2009), *Response Surface Methodology: Process And Product Optimization Using Designed Experiments,* Wiley

Tang, L. and Xu, K. (2002), "A Unified Approach For Dual Response Surface Optimization", *Journal of Quality Technology*, Vol. 34, No**.** 4, pp. 437-447.

Kim, K. and Lin, D. (2006), "Optimization Of Multiple Responses Considering Both Location And Dispersion Effects", *European Journal of Operational Research*, Vol. 169, No. 1, pp. 133-145.

Anthony, R. (1965), *Planning And Control Systems: A Framework For Analysis*, Harvard Business School Press.

Anthony, R. N., Dearden, N. J. and Govindarajan, V. (1992), *Management Control Systems*, Irwin, Homewood, IL.

Harrington, H. (1991), *Business Process Improvement: The Breakthrough Strategy For Total Quality, Productivity, And Competitiveness*, McGraw-Hill Professional.

Ariely, D. (2010), "Why Businesses Don't Experiment", *Harvard business review*, Vol. 88, No. 4, pp. 34-34.

Turban, E. and Aronson, J. (1997), *Decision Support Systems And Intelligent Systems*, Prentice Hall, NJ, USA.

Alberti, M., Ciurana, J., Rodríguez, C. and Özel, T. (2009), "Design Of A Decision Support System For Machine Tool Selection Based On Machine Characteristics And Performance Tests", *Journal of Intelligent Manufacturing*, Vol. 20, No. 4, pp. 1-15.

King, W. and Rodriguez, J. (1981), "Participative Design Of Strategic Decision Support Systems: An Empirical Assessment", *Management science*, Vol. 27, No. 6, pp. 717-726.

Sharda, R., Barr, S. and MCDonnell, J. (1988), "Decision Support System Effectiveness: A Review And An Empirical Test", *Management science*, Vol. 34, No. 2, pp. 139-159.

Daily, B., Whatley, A., Ash, S. and Steiner, R. (1996), "The Effects Of A Group Decision Support System On Culturally Diverse And Culturally Homogeneous Group Decision Making", *Information & Management*, Vol. 30, No. 6, pp. 281-289.

Ackoff, R. (1979), "Resurrecting The Future Of Operational Research", *Journal of the Operational Research Society*, Vol. 30, No. 3, pp. 189-199.

Ackoff, R. (1981), *Creating The Corporate Future: Plan Or Be Planned For*, John Wiley & Sons Inc.

Snee, R. (1986), "In Pursuit Of Total Quality", *Quality Progress*, Vol. 19, No. 8, pp. 153-159.

Taguchi, G. (1986), *Introduction To Quality Engineering: Designing Quality Into Products And Processes*, Asian Productivity Organization Tokyo.

Phadke, M. (1995), *Quality Engineering Using Robust Design*, Prentice Hall

Ross, P. J. (1996), *Taguchi Techniques for Quality Engineering*, McGraw-Hill.

Pignatiello, J. and Ramberg, J. (1991), "Top Ten Triumphs And Tragedies Of Genichi Taguchi", *Quality Engineering*, Vol. 4, No. 2, pp. 211-225.

Goh, T. (2001), "A Pragmatic Approach To Experimental Design In Industry", *Journal of Applied Statistics*, Vol. 28, No. 3, pp. 391-398.

Crombie, A. (1953), *Robert Grosseteste And The Origins Of Experimental Science, 1100-1700,* Clarendon Press.

Box, G. and Youle, P. (1955), "The Exploration And Exploitation Of Response Surfaces: An Example Of The Link Between The Fitted Surface And The Basic Mechanism Of The System", *Biometrics*, Vol. 11, No. 3, pp. 287-323.

Box, G. and Liu, P. (1999), "Statistics As A Catalyst To Learning By Scientific Method Part I-An Example", *Journal of Quality Technology*, Vol. 31, No. 1, pp. 1-15.

Box, G. (1999), "Statistics As A Catalyst To Learning By Scientific Method Part II-A Discussion", *Journal of Quality Technology*, Vol. 31, No. 1, pp. 16-29.

Shewhart, W. (1980), *Economic Control Of Quality Of Manufactured Product/50th Anniversary Commemorative Issue.* , American Society for Quality.

Box, G. (1994), "Statistics and Quality Improvement", *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, Vol. 157, No. 2, pp. 209-229.

Box, G. (2001), "Statistics For Discovery", *Journal of Applied Statistics*, Vol. 28, No. 3, pp. 285-299.

Box, G. (1992), "George's Column", *Quality Engineering*, Vol. 5 No. 2, pp. 321 - 330.

Daniel, C. (1976), *Applications Of Statistics To Industrial Experimentation*, Wiley-Interscience.

Box, G., Hunter, S. and Hunter, W. (1978), *Statistics For Experimenters: Design, Innovation, And Discovery,* Wiley.

Goh, T. (2002), "The Role Of Statistical Design Of Experiments In Six Sigma: Perspectives Of A Practitioner", *Quality Engineering*, Vol. 14, No. 4, pp. 659-671.

Antony, J. (1998), "Some Key Things Industrial Engineers Should Know About Experimental Design", *Logistics Information Management*, Vol. 11, No. 6, pp. 386-392.

Antony, J. (2002), "Training For Design Of Experiments Using A Catapult", *Quality and Reliability Engineering International*, Vol. 18, No. 1, pp. 29-35.

Antony, J. (2006), "Taguchi Or Classical Design Of Experiments: A Perspective From A Practitioner", *Sensor Review*, Vol. 26, No. 3, pp. 227-230.

Antony, J. and Antony, F. (2001), "Teaching The Taguchi Method To Industrial Engineers", *Work study*, Vol. 50, No. 4/5, pp. 141-149.

Antony, J. (1999), "Improving The Wire Bonding Process Quality Using Statistically Designed Experiments", *Microelectronics Journal*, Vol. 30, No. 2, pp. 161-168.

Antony, J., Hughes, M. and Kaye, M. (1999), "Reducing Manufacturing Process Variability Using Experimental Design Technique: A Case Study", *Integrated manufacturing systems*, Vol. 10, No. 3, pp. 162-169.

Antony, J., Warwood, S., Fernandes, K. and Rowlands, H. (2001), "Process Optimisation Using Taguchi Methods Of Experimental Design", *Work study*, Vol. 50, No. 2/3, pp. 51-57.

Antony, J., Somasundarum, V., Fergusson, C. and Blecharz, P. (2004), "Applications Of Taguchi Approach To Statistical Design Of Experiments In Czech Republican Industries", *International Journal of Productivity and Performance Management*, Vol. 53, No. 5, pp. 447-457.

Antony, J., Perry, D., Wang, C. and Kumar, M. (2006), "An Application Of Taguchi Method Of Experimental Design For New Product Design And Development Process", *Assembly Automation*, Vol. 26, No. 1, pp. 18-24.

Rowlands, H., Antony, J. and Knowles, G. (2000), "An Application Of Experimental Design For Process Optimisation", *The TQM Magazine*, Vol. 12, No. 2, pp. 78-83.

Antony, J. and Roy, R. (1999), "Improving The Process Quality Using Statistical Design Of Experiments: A Case Study", *Quality Assurance*, Vol. 6, No. 2, pp. 87-95.

Donahue, D., Sowell, R. and Bengtson, N. (1996), "Simulation Of Alternative Agricultural Marketing Systems", *Agricultural Systems*, Vol. 51, No. 4, pp. 395-406.

Pean, J., Venier-Julienne, M., Filmon, R., Sergent, M., Phan-Tan-Luu, R. and Benoit, J. (1998), "Optimization Of HSA And NGF Encapsulation Yields In PLGA Microparticles", *International Journal of Pharmaceutics*, Vol. 166, No. 1, pp. 105-115.

Gratteri, P., Furlanetto, S., La Porta, E., Pinzauti, S. and Leardi, R. (1996), "Development And Set-up Of Drug Electroanalysis By Experimental Design: A Survey", *Il Farmaco*, Vol. 51, No. 4, pp. 231-246.

Smith, B., Penrod, S., Otto, A. and Park, R. (1996), "Jurors' Use Of Probabilistic Evidence", *Law and Human Behavior*, Vol. 20, No. 1, pp. 49-82.

Stolle, D., Robbennolt, J., Patry, M. and Penrod, S. (2002), "Fractional Factorial Designs For Legal Psychology", *Behavioral Sciences & the Law*, Vol. 20, No. 1-2, pp. 5-17.

Condra, L. (1995), *Value-Added Management With Design Of Experiments*, Springer.

Box, G. (1988), "Signal-to-Noise Ratios, Performance Criteria, And Transformations", *Technometrics*, Vol. 30, No. 1, pp. 1-17.

Box, G., Bisgaard, S. and Fung, C. (1988), "An Explanation And Critique Of Taguchi's Contributions To Quality Engineering", *Quality and Reliability Engineering International*, Vol. 4, No. 2, pp. 123-131.

Hunter, J. (1985), "Statistical Design Applied To Product Design", *Journal of Quality Technology*, Vol. 17, No. 4, pp. 210-221.

Montgomery, D. (1990), "Using Fractional Factorial Designs For Robust Process Development", *Quality Engineering*, Vol. 3, No. 2, pp. 193-205.

Nair, V., Abraham, B., MacKay, J., Nelder, J., Box, G., Phadke, M., Kacker, R., Sacks, J., Welch, W. and Lorenzen, T. (1992), "Taguchi's Parameter Design: A Panel Discussion", *Technometrics*, Vol. 34, No. 2, pp. 127-161.

Shainin, D. (1957), "The Statistically Designed Experiment", *Harvard business review*, Vol. 35, No. 4, pp. 67-73.

Shainin, D. and Shainin, P. (1988), "Better Than Taguchi Orthogonal Tables", *Quality and Reliability Engineering International*, Vol. 4, No. 2, pp. 143-149.

Ledolter, J. and Swersey, A. (1997), "Dorian Shainin's Variables Search Procedure: A Critical Assessment", *Journal of Quality Technology*, Vol. 29, No. 3, pp. 237-247.

Thomas, A. and Antony, J. (2005), "A Comparative Analysis Of The Taguchi And Shainin DOE Techniques In An Aerospace Environment", *International Journal of Productivity and Performance Management*, Vol. 54, No. 8, pp. 658.

Cox, D. and Reid, N. (2000), *The Theory Of The Design Of Experiments,* CRC Press.

Goh, T. (2009), "Guest Editorial: Statistical Thinking And Experimental Design As Dual Drivers Of DFSS", *International Journal of Six Sigma and Competitive Advantage*, Vol. 5, No. 1, pp. 2-9.

Chan, K. and Spedding, T. (2001), "On-line Optimization Of Quality In A Manufacturing System", *International Journal of Production Research*, Vol. 39, No. 6, pp. 1127-1145.

Conklin, J. (2004), "DOE and Six Sigma", *Quality Progress*, Vol. 37, No. 3, pp. 66-69.

Araujo, C., Benedetto-Neto, H., Campello, A., Segre, F. and Wright, I. (1996), "The Utilization Of Product Development Methods: A Survey Of UK Industry", *Journal of Engineering Design*, Vol. 7, No. 3, pp. 265-277.

Antony, J. and Banuelas, R. (2002), "Key Ingredients For The Effective Implementation Of Six Sigma Program", *Measuring Business Excellence*, Vol. 6, No. 4, pp. 20-27.

Arvidsson, M., Gremyr, I. and Johansson, P. (2003), "Use And Knowledge Of Robust Design Methodology: A Survey Of Swedish Industry", *Journal of Engineering Design*, Vol. 14, No. 2, pp. 129-143.

Gremyr, I., Arvidsson, M. and Johansson, P. (2003), "Robust Design Methodology: Status In The Swedish Manufacturing Industry", *Quality and Reliability Engineering International*, Vol. 19, No. 4, pp. 285-293.

Bergquist, B. and Albing, M. (2006), "Statistical Methods–Does Anyone Really Use Them?", *Total Quality Management & Business Excellence*, Vol. 17, No. 8, pp. 961-972.

Tanco, M., Viles, E., Ilzarbe, L. and Alvarez, M. (2008), "Is Design Of Experiments Really Used? A Survey Of Basque Industries", *Journal of Engineering Design*, Vol. 19, No. 5, pp. 447-460.

Ilzarbe, L., Álvarez, M., Viles, E. and Tanco, M. (2008), "Practical Applications Of Design Of Experiments In The Field Of Engineering: A Bibliographical Review", *Quality and Reliability Engineering International*, Vol. 24, No. 4, pp. 417-428.

Groover, M. (2007), *Fundamentals Of Modern Manufacturing: Materials Processes, And Systems*, Wiley-India.

Sharma, S. (2005), *Dictionary of Engineering*, Gyan Books.

Thomson-Reuters (2010), "Thomson Citation Reports", Retrieved 5/10/2010, from http://thomsonreuters.com/products_services/science/science_products/a-z/journal_citation_reports.

SPSS (2007), "SPSS Statistics Base 17.0 User's Guide", Retrieved 5/4/2010, from http://support.spss.com/ProductsExt/SPSS/ESD/17/Download/User%20Manuals/English/SPSS%20Statistcs%20Base%20User's%20Guide%2017.0.pdf.

Lauderbaugh, L. (2009), "Analysis Of The Effects Of Process Parameters On Exit Burrs In Drilling Using A Combined Simulation And Experimental Approach", *Journal of Materials Processing Technology*, Vol. 209, No. 4, pp. 1909-1919.

Tsai, K., Hsieh, C. and Lo, W. (2009), "A Study Of The Effects Of Process Parameters For Injection Molding On Surface Quality Of Optical Lenses", *Journal of Materials Processing Technology*, Vol. 209, No. 7, pp. 3469-3477.

Ali, S., Hinduja, S., Atkinson, J. and Pandya, M. (2009), "Shaped Tube Electrochemical Drilling Of Good Quality Holes", *CIRP Annals-Manufacturing Technology*, Vol. 58, No. 1, pp. 185-188.

Venkatachalam, S., Li, X. and Liang, S. (2009), "Predictive Modeling Of Transition Undeformed Chip Thickness In Ductile-Regime Micro-Machining Of Single Crystal Brittle Materials", *Journal of Materials Processing Technology*, Vol. 209, No. 7, pp. 3306-3319.

Jeang, A. and Li, H. (2009), "An Empirical Study On The Robust Design Of A Semiconductor-Bumping Process", *Journal of Materials Processing Technology*, Vol. 209, No. 6, pp. 2986-2993.

Karayel, D. (2009), "Prediction And Control Of Surface Roughness In CNC Lathe Using Artificial Neural Network", *Journal of Materials Processing Technology*, Vol. 209, No. 7, pp. 3125-3137.

León, R., Shoemaker, A. and Kacker, R. (1987), "Performance Measures Independent Of Adjustment: An Explanation And Extension Of Taguchi's Signal-To-Noise Ratios", *Technometrics*, Vol. 29, No. 3, pp. 253-265.

Phatak, U., Bukkapatnam, S., Kong, Z. and Komanduri, R. (2009), "Sensor-Based Modeling Of Slurry Chemistry Effects On The Material Removal Rate (MRR) In Copper-CMP Process", *International Journal of Machine Tools and Manufacture*, Vol. 49, No. 2, pp. 171-181.

Yang, Y., Chuang, M. and Lin, S. (2009), "Optimization Of Dry Machining Parameters For High-Purity Graphite In End Milling Process Via Design Of Experiments Methods", *Journal of Materials Processing Technology*, Vol. 209, No. 9, pp. 4395-4400.

Tsui, K. (1988), "Strategies For Planning Experiments Using Orthogonal Arrays And Confounding Tables", *Quality and Reliability Engineering International*, Vol. 4, No. 2, pp. 113-122.

Fries, A. and Hunter, W. (1980), "Minimum Aberration 2^{k-p} Designs", *Technometrics*, Vol. 22, No. 4, pp. 601-608.

Chen, J., Sun, D. and Wu, C. (1993), "A Catalogue Of Two-Level And Three-Level Fractional Factorial Designs With Small Runs", *International Statistical Review/Revue Internationale de Statistique*, Vol. 61, No. 1, pp. 131-145.

Taguchi, G., Subir. Chowdhury, Y. and Wu, Y. (2004), *Taguchi's Quality Engineering Handbook,* John Wiley & Sons.

Hamada, M. and Wu, C. (1992), "Analysis Of Designed Experiments With Complex Aliasing", *Journal of Quality Technology*, Vol. 24, No. 3, pp. 130-137.

Box, G. and Meyer, R. (1993), "Finding The Active Factors In Fractionated Screening Experiments", *Journal of Quality Technology*, Vol. 25, No. 2, pp. 94-94.

Marafona, J. and Araújo, A. (2009), "Influence Of Workpiece Hardness On EDM Performance", *International Journal of Machine Tools and Manufacture,* Vol. 49, No. 9, pp. 744-748.

Cox, D. (2006), *Principles Of Statistical Inference*, Cambridge Univ Press.

Hinkelmann, K. and Kempthorne, O. (2008), *Design And Analysis Of Experiments: Introduction To Experimental Design,* Wiley-Interscience.

Daniel, C. (1959), "Use Of Half-Normal Plots In Interpreting Factorial Two-Level Experiments", *Technometrics*, Vol. 1, No. 4, pp. 311-341.

Logothetis, N. and Wynn, H. (1989), *Quality Through Design: Experimental Design, Off-Line Quality Control, And Taguchi's Contributions,* Clarendon Press Oxford.

Peace, G. (1993), *Taguchi Methods: A Hands-On Approach,* Addison Wesley Publishing Company.

Lenth, R. (1989), "Quick And Easy Analysis Of Unreplicated Factorials", *Technometrics,* Vol. 31 No. 4, pp. 469-473.

Luo, S. and Chen, K. (2009), "An Experimental Study Of Flat Fixed Abrasive Grinding Of Silicon Wafers Using Resin-Bonded Diamond Pellets", *Journal of Materials Processing Technology*, Vol. 209, No. 2, pp. 686-694.

Senthilkumar, B., Ponnambalam, S. and Jawahar, N. (2009), "Process Factor Optimization For Controlling Pull-Down Defects In Iron Castings", *Journal of Materials Processing Technology*, Vol. 209, No. 1, pp. 554-560.

Shyha, I., Aspinwall, D., Soo, S. and Bradley, S. (2009), "Drill Geometry And Operating Effects When Cutting Small Diameter Holes In CFRP", *International Journal of Machine Tools and Manufacture*, Vol. 49, No. 12-13, pp. 1008-1014.

Chattopadhyay, K., Verma, S., Satsangi, P. and Sharma, P. (2009), "Development Of Empirical Model For Different Process Parameters During Rotary Electrical Discharge Machining Of Copper-Steel (EN-8) system", *Journal of Materials Processing Technology*, Vol. 209, No. 3, pp. 1454-1465.

Boronat, T., Segui, V., Peydro, M. and Reig, M. (2009), "Influence Of Temperature And Shear Rate On The Rheology And Processability Of Reprocessed ABS In Injection Molding Process", *Journal of Materials Processing Technology,* Vol. 209, No. 5, pp. 2735-2745.

Rakwal, D. and Bamberg, E. (2009), "Slicing, Cleaning And Kerf Analysis Of Germanium Wafers Machined By Wire Electrical Discharge Machining", *Journal of Materials Processing Technology*, Vol. 209, No. 8, pp. 3740-3751.

Rosa, J., Robin, A., Silva, M., Baldan, C. and Peres, M. (2009), "Electrodeposition Of Copper On Titanium Wires: Taguchi Experimental Design Approach", *Journal of Materials Processing Technology,* Vol. 209, No. 3, pp. 1181-1188.

Lin, Z. and Ho, C. (2009), "Performance Of Coated Tungsten Carbide Tools On Milling Printed Circuit Board", *Journal of Materials Processing Technology*, Vol. 209, No. 1, pp. 303-309.

Penzias, A. (1989), "Teaching Statistics To Engineers", *Science*, Vol. 244, No. 4908, pp. 1025-1025.

Bisgaard, S. (1991), "Teaching Statistics To Engineers", *American Statistician*, Vol. 45, No. 4, pp. 274-283.

Burton, R. and Connor, W. (1957), "On the Identity Relationship for Fractional Replicates of the 2^n Series", *The Annals of Mathematical Statistics*, Vol. 28 No. 3, pp. 762-767.

Plackett, R. L. and Burman, J. P. (1946), "The Design of Optimum Multifactorial Experiments", *Biometrika*, Vol. 33, No. 4, pp. 305-325.

Greenfield, A. (1976), "Selection Of Defining Contrasts In Two-Level Experiments", *Applied Statistics*, Vol. 25, No. 1, pp. 64-67.

Franklin, M. and Bailey, R. (1977), "Selection Of Defining Contrasts And Confounded Effects In Two-Level Experiments", *Applied Statistics*, Vol. 26, No. 3, pp. 321-326.

Franklin, M. (1985), "Selecting Defining Contrasts And Confounded Effects In p^{n-m} Factorial Experiments", *Technometrics*, Vol. 27, No. 2, pp. 165-172.

Kacker, R. and Tsui, K. (1990), "Interaction Graphs: Graphical Aids For Planning Experiments", *Journal of Quality Technology*, Vol. 22, No. 1, pp. 1-14.

Wu, C. and Chen, Y. (1992), "A Graph-aided Method For Planning Two-Level Experiments When Certain Interactions Are Important", *Technometrics*, Vol. 34, No. 2, pp. 162-175.

Sun, D. and Wu, C. (1994), "Interaction Graphs For Three-Level Fractional Factorial Designs", *Journal of Quality Technology*, Vol. 26, No. 4, pp. 297-307.

Box, G. E. P. and Hunter, J. S. (1961 a), "The 2^{k-p} Fractional Factorial Designs Part I", *Technometrics*, Vol. 3, No. 3, pp. 311-351.

Wu, C. and Zhang, R. (1993), "Minimum Aberration Designs With Two-Level And Four-Level Factors", *Biometrika*, Vol. 80, No. 1, pp. 203-209.

Franklin, M. (1984), "Constructing Tables Of Minimum Aberration p^{n-m} Designs", *Technometrics*, Vol. 26, No. 3, pp. 225-232.

Chen, J. (1992), "Some Results on 2^{n-k} Fractional Factorial Designs and Search for Minimum Aberration Designs", *The Annals of Statistics*, Vol. 20, No. 4, pp. 2124-2141.

Chen, J. and Wu, C. (1991), "Some Results on S^{n-k} fractional Factorial Designs With Minimum Aberration or Optimal Moments", *The Annals of Statistics*, Vol. 19, No. 2, pp. 1028-1041.

Tang, B. and Deng, L. (1999), "Minimum G_2-aberration For Nonregular Fractional Factorial Designs", *Annals of Statistics*, Vol. 27, No. 6, pp. 1914-1926.

Deng, L. and Tang, B. (1999), "Generalized Resolution And Minimum Aberration Criteria For Plackett-Burman And Other Nonregular Factorial Designs", *Statistica Sinica*, Vol. 9, No**.** 4, pp. 1071-1082.

Xu, H. and Wu, C. (2001), "Generalized Minimum Aberration For Asymmetrical Fractional Factorial Designs", *Annals of Statistics*, Vol. 29, No**.** 2, pp. 549-560.

Cheng, C. and Tang, B. (2005), "A General Theory Of Minimum Aberration And Its Applications", *Annals of Statistics*, Vol. 33, No**.** 2, pp. 944-958.

Fang, K., Ge, G., Liu, M. and Qin, H. (2003), "Construction Of Minimum Generalized Aberration Designs", *Metrika*, Vol. 57, No**.** 1, pp. 37-50.

Pease, M. (1965), *Methods Of Matrix Algebra*, Academic Press Inc.

Gentle, J. (2003), *Random Number Generation And Monte Carlo Methods*, Springer Verlag.

Williams, G. (2009), *Linear Algebra With Applications*, Jones & Bartlett Pub.

Massart, D. and Vandeginste, B. (1997), *Handbook Of Chemometrics And Qualimetrics*, Elsevier Science.

Xu, H. (2001), *Optimal Factor Assignment for Asymmetrical Fractional Factorial Designs: Theory and Applications*, Ph.D. thesis, University of Michigan, Ann Arbor.

Brownlee, K. (1965), *Statistical Methodology In Science And Engineering*, John Wiley

Marquardt, D. (1970), "Generalized Inverses, Ridge Regression, Biased Linear Estimation, And Nonlinear Estimation", *Technometrics*, Vol. 12, No**.** 3, pp. 591-612.

Miller, A., Sitter, R., Wu, C. F. J. and Long, D. (1993), "Are Large Taguchi-style Experiments Necessary? A Reanalysis of Gear and Pinion Data", *Quality Engineering*, Vol. 6, No**.** 1, pp. 21-37.

Hamada, M. and Balakrishnan, N. (1998), "Analyzing Unreplicated Factorial Experiments: A Review With Some New Proposals", *Statistica Sinica*, Vol. 8, No**.** 1, pp. 1-28.

Box, G. and Meyer, R. (1986), "An Analysis For Unreplicated Fractional Factorials", *Technometrics*, Vol. 28, No**.** 1, pp. 11-18.

Sun, D. X. (1996), *Estimation Capacity and Related Topics in Experimental Design*, Ph.D. Thesis, University of Waterloo.

Ye, K. and Hamada, M. (2000), "Critical Values Of The Lenth Method For Unreplicated Factorial Design", *Journal of Quality Technology* Vol. 32, No**.** 1, pp. 57 - 66.

Roy, R. (2001), *Design Of Experiments Using The Taguchi Approach: 16 Steps To Product And Process Improvement,* Wiley-Interscience.

Gunaraj, V. and Murugan, N. (1999), "Prediction And Comparison Of The Area Of The Heat-Affected Zone For The Bead-On-Plate And Bead-On-Joint In Submerged Arc Welding Of Pipes", *Journal of Materials Processing Technology,* Vol. 95, No. 1-3, pp. 246-261.

Sun, X., Pei, Z., Xin, X. and Fouts, M. (2004), "Waviness Removal In Grinding Of Wire-Sawn Silicon Wafers: 3D Finite Element Analysis With Designed Experiments", *International Journal of Machine Tools and Manufacture,* Vol. 44, No. 1, pp. 11-19.

Dutta, P. and Pratihar, D. (2007), "Modeling Of TIG Welding Process Using Conventional Regression Analysis And Neural Network-Based Approaches", *Journal of Materials Processing Technology,* Vol. 184, No. 1-3, pp. 56-68.

Darwish, S. (2000), "The Impact Of The Tool Material And The Cutting Parameters On Surface Roughness Of Supermet 718 Nickel Superalloy", *Journal of Materials Processing Technology*, Vol. 97, No. 1-3, pp. 10-18.

Raghukandan, K. and Senthilvelan, T. (2004), "Analysis of P/M Hollow Extrusion Using Design Of Experiments", *Journal of Materials Processing Technology*, Vol. 153, No. 1, pp. 416-419.

Kannan, T. and Murugan, N. (2006), "Effect Of Flux Cored Arc Welding Process Parameters On Duplex Stainless Steel Clad Quality", *Journal of Materials Processing Technology,* Vol. 176, No. 1-3, pp. 230-239.

Wu, C. and Hamada, M. (2000), *Experiments: Planning, Analysis, And Parameter Design Optimization,* Wiley New York.

Box, G., Hunter, S. and Hunter, W. (2005), *Statistics For Experimenters: Design, Innovation, And Discovery,* Wiley.

Ganjigatti, J., Pratihar, D. and Roy Choudhury, A. (2007), "Global Versus Cluster-Wise Regression Analyses For Prediction Of Bead Geometry In MIG Welding Process", *Journal of Materials Processing Technology*, Vol. 189, No. 1-3, pp. 352-366.

Aggarwal, A., Singh, H., Kumar, P. and Singh, M. (2008), "Optimizing Power Consumption For CNC Turned Parts Using Response Surface Methodology And Taguchi's Technique--A Comparative Analysis", *Journal of Materials Processing Technology*, Vol. 200, No. 1-3, pp. 373-384.

Darwin, J., Mohan Lal, D. and Nagarajan, G. (2008), "Optimization Of Cryogenic Treatment To Maximize The Wear Resistance Of 18% Cr Martensitic Stainless Steel By Taguchi Method", *Journal of Materials Processing Technology*, Vol. 195, No. 1-3, pp. 241-247.

Pei, Z., Xin, X. and Liu, W. (2003), "Finite Element Analysis For Grinding Of Wire-Sawn Silicon Wafers: A Designed Experiment", *International Journal of Machine Tools and Manufacture*, Vol. 43, No. 1, pp. 7-16.

Correia, D. and Ferraresi, V. (2007), "Welding Process Selection Through A Double Criteria: Operational Costs And Non-Quality Costs", *Journal of Materials Processing Technology*, Vol. 184, No. 1-3, pp. 47-55.

Gaitonde, V., Karnik, S., Rubio, J., Correia, A., Abrão, A. and Davim, J. (2008), "Analysis Of Parametric Influence On Delamination In High-Speed Drilling Of Carbon Fiber Reinforced Plastic Composites", *Journal of Materials Processing Technology*, Vol. 203, No. 1-3, pp. 431-438.

Dhar, S., Purohit, R., Saini, N., Sharma, A. and Kumar, G. (2007), "Mathematical Modeling Of Electric Discharge Machining Of Cast Al-4Cu-6Si alloy-10 wt.% SiCP Composites", *Journal of Materials Processing Technology*, Vol. 194, No. 1-3, pp. 24-29.

Davim, J. (2003), "Design Of Optimisation Of Cutting Parameters For Turning Metal Matrix Composites Based On The Orthogonal Arrays", *Journal of Materials Processing Technology*, Vol. 132, No. 1-3, pp. 340-344.

Box, G. and Draper, N. (2007), *Response Surfaces, Mixtures, And Ridge Analyses*, Wiley-Interscience.

Ghani, J., Choudhury, I. and Hassan, H. (2004), "Application Of Taguchi Method In The Optimization Of End Milling Parameters", *Journal of Materials Processing Technology*, Vol. 145, No. 1, pp. 84-92.

Davim, J., Gaitonde, V. and Karnik, S. (2008), "Investigations Into The Effect Of Cutting Conditions On Surface Roughness In Turning Of Free Machining Steel by ANN Models", *Journal of Materials Processing Technology*, Vol. 205, No. 1-3, pp. 16-23.

Davim, J. (2000), "An Experimental Study Of The Tribological Behaviour Of The Brass/Steel Pair", *Journal of Materials Processing Technology*, Vol. 100, No. 1-3, pp. 273-277.

Shoemaker, A., Tsui, K. and Wu, C. (1991), "Economical Experimentation Methods For Robust Design", *Technometrics*, Vol. 33, No. 4, pp. 415-427.

Winston, W. (2003), *Operations Research: Applications And Algorithms*, Duxbury press Belmont, CA.

Kalos, M. and Whitlock, P. (2008), *Monte Carlo Methods*, Wiley-VCH.

Box, G. and Cox, D. (1964), "An Analysis Of Transformations", *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 26, No. 2, pp. 211-252.

Lynch, R. (1993), "Minimum Detectable Effects For 2^{k-p} Experimental Plans", *Journal of Quality Technology*, Vol. 25, No. 1, pp. 12-12.

Oehlert, G. and Whitcomb, P. (2001), "Sizing Fixed Effects For Computing Power In Experimental Designs", *Quality and Reliability Engineering International*, Vol. 17, No**.** 4, pp. 291-306.

Snee, R. (1985), "Experimenting With A Large Number Of Variables", in Snee, R., Hare, L. and Trout, J. (Ed.), *Experiments in Industry: Design, Analysis, and Interpretation of Results*, American Society for Quality Control, Milwaukee, WI, pp.

Montgomery, D., Borror, C. and Stanley, J. (1997), "Some Cautions In The Use Of Plackett-Burman Designs", *Quality Engineering*, Vol. 10, No**.** 2, pp. 371-381.

Tukey, J. (1969), "Analyzing Data: Sanctification On Detective Work", *American Psychologist*, Vol. 24, No**.** 2, pp. 83-91.

Schlesinger, S., Crosbie, R., Gagne, R., Innis, G., Lalwani, C., Loch, J., Sylvester, R., Wright, R., Kheir, N. and Bartos, D. (1979), "Terminology For Model Credibility", *Simulation*, Vol. 32, No**.** 3, pp. 103-104.

Kleijnen, J. (1995), "Verification And Validation Of Simulation Models", *European Journal of Operational Research*, Vol. 82, No**.** 1, pp. 145-162.

Davis, P. (1992), Generalizing Concepts And Methods Of Verification, Validation, And Accreditation (vv&a) For Military Simulations, RAND Publications R-4249-ACQ.

Sargent, R. (2000), *Verification And Validation Of Simulation Models,* Proceedings of the 2000 Winter Simulation Conference.

Taguchi, G. (1987), *System of Experimental Design,* White Plains, New York: UNIPUB/Krauss International.

Reddy, R., DeVor, R., Kapoor, S. and Sun, Z. (2001), "A Mechanistic Model-Based Force-Feedback Scheme For Voice-Coil Actuated Radial Contour Turning", *International Journal of Machine Tools and Manufacture*, Vol. 41, No**.** 8, pp. 1131-1147.

McGeough, J. (1988), *Advanced Methods Of Machining*, Springer.

Tansel, I., Ozcelik, B., Bao, W., Chen, P., Rincon, D., Yang, S. and Yenilmez, A. (2006), "Selection Of Optimal Cutting Conditions By Using GONNS", *International Journal of Machine Tools and Manufacture*, Vol. 46, No**.** 1, pp. 26-35.

Balci, O. (1998), "Verification, Validation, And Testing", in Banks, J. (Ed.), *Handbook of Simulation: Principles, Methodology, Advances, Applications, And Practice*, Wiley-Interscience, pp. 335-393.

Neelamkavil, F. (1987), *Computer Simulation And Modelling,* John Wiley & Sons, Inc. New York, NY, USA.

Yates, F. (1951), "The Influence Of Statistical Methods For Research Workers On The Development Of The Science Of Statistics", *Journal of the American Statistical Association*, Vol. 46, No**.** 253, pp. 19-34.

Fisher, R. A. (1925), *Statistical Methods For Research Workers, ,* Oliver and Boyd, Edinburgh, UK

Barten, A. P. (1962), "Note On Unbiased Estimation Of The Squared Multiple Correlation Coefficient", *Statistica Neerlandica*, Vol. 16, No**.** 2, pp. 151-164.

Kelley, T. L. (1935), "An Unbiased Correlation Ratio Measure", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 21, No**.** 9, pp. 554-559.

Wherry, R. (1931), "A New Formula For Predicting The Shrinkage Of The Coefficient Of Multiple Correlation", *The Annals of Mathematical Statistics*, Vol. 2, No**.** 4, pp. 440-457.

Olkin, I. and Pratt, J. W. (1958), "Unbiased Estimation Of Certain Correlation Coefficients", *The Annals of Mathematical Statistics*, Vol. 29, No**.** 1, pp. 201-211.

Montgomery, D. C. (2009), *Introduction To Statistical Quality Control,* Wiley.

Chiang, C. (1966), "On The Expectation Of The Reciprocal Of A Random Variable", *American Statistician*, Vol. 20, No**.** 4, pp. 28-28.

Davenport, T. H. (2006), "Competing on Analytics", *Harvard Business Review*, Vol. 84, No**.** 1, pp. 99-107.

Davenport, T. H. and Harris, J. G. (2007), *Competing on Analytics: the New Science Of Winning*, Harvard Business Press.

Antony, J., Kaye, J. A. M. and Frangou, A. (1998), "A Strategic Methodology To The Use Of Advanced Statistical Quality Improvement Techniques", *The TQM Magazine*, Vol. 10, No**.** 3, pp. 169-176.

APPENDIX 1:
DOE BASIC PRINCIPALS AND IMPLEMENTATION STAGES

1. DOE Basic Principles

A major aspect of DOE is the concept of analysing the sources of variation. In fact every response in an experiment demonstrates variation, of which there are three sources: those due to changing the controllable variables, those associated with the noise variables and those due to chance and the measurement process (inherent). A good experimental design will enable the variability associated with each source to be estimated. The conventional purpose of an experiment is to study and estimate the variation due to controllable factors (or the variability of the average response around its target). The noise variables' variation includes the variability that results from using a different batch of materials or carrying out the experiment under different environmental conditions (e.g. ambient temperature, humidity etc.) or by different operators etc. This is the variability of the individual response values around their average but also includes errors in the measurement process.

In any experiment, if two measurements are taken under the same conditions the results will almost never be identical. Although this variation is inherent and unavoidable it has two properties (Cobb, 1997). Firstly, some chance errors will be positive, while others will be negative, but on balance they will tend to cancel each other out, at least partially, when an average is computed. The more measurements that go into an average, the smaller the chance variation will tend to be, and the more accurate the average itself will be. Secondly, if the experiment is well-planned, it will be possible to estimate the size of the chance variation, and this will make it possible to determine the precision of the obtained average.

In order to account for the above mentioned sources of variation, the experiment should be planned and designed very carefully. According to Montgomery (2010) there are three basic principles for experimental design which should be considered, replication, randomisation, and blocking.

Replication means repeating the experiment under the same conditions not performing the experiment once and taking several measurements. This allows the experimenter to obtain an estimate of the experimental error, which becomes a basic unit of measurement for determining whether observed differences in the data are really statistically different. Also, if the sample mean is used to estimate the effect of a factor in the experiment, replication allows the experimenter to obtain a more precise estimate of this effect because the variance of the sample mean is less than that of the individual observations. On the other hand the variability that results from repeated measurements is a direct reflection of the inherent variability in the measurement system or gauge.

Randomisation means that both the allocation of the experimental material and the order in which the trials are conducted are determined randomly. This is important for the following three reasons: (i) statistical methods require observations (or errors) to be independently distributed random variables, randomisation usually validates this assumption; (ii) it assists in "averaging out" the effects of extraneous factors that may be present; and (iii) where the operation is repetitious the order in which the experiments are conducted may be important, either because a learning process is involved which tends to make later runs better than earlier ones, or because of fatigue which has the opposite effect. This systematic bias may be eradicated by randomisation.

Although randomisation helps to "average out" the effects of extraneous factors, it does not eliminate their resultant variability. Blocking is another design technique that reduces or eliminates the variability arising from 'nuisance' factors i.e. those that may influence the experimental response but in which there is no direct interest such as different batches of raw material. Using this technique ensures that any variability that results from these factors, will be isolated from the experimental error, which in turn becomes smaller enabling more precise conclusions to be drawn from the experiment. Generally, a block is a set of relatively homogeneous experimental conditions. Each level of the nuisance factor constitutes a block. The experimenter divides the observations from the statistical design into groups that are run in each block.

Alongside the above three principles, there are another three concerning the relative importance and the relationships between the studied main effects and their interactions (Box et al 2005). These are particularly important for justifying the development and use of fractional factorial designs (addressed in section 2). The first is the *hierarchical ordering* principle which entails considering lower order effects as being more likely to be important than higher order effects and regarding those of the same order as equally likely to be important. Thus it suggests that when resources are scarce, priority should be given to lower order effects. So the estimation of, say, main effects is more important than that of two-factor interactions. The second principle is *effect sparsity* (Box and Meyer, 1986) which states that the number of relatively important effects in an experiment is small. It parallels the Pareto principle (Juran, 1993) with respect to focusing on the "vital few" not the "trivial many". *Effect heredity*, which is the third principle, demands that at least one of the parent factors of an interaction is significant in order for the interaction to be pronounced significant. It must be borne in mind that, as Wu and Hamada, (2000) highlighted, these three principles are empirical, their validity having been confirmed in many real experiments. However, it is not uncommon to encounter experimental situations where one or more of these principles is violated.

2. Stages of Design of Experiment Study

As Antony (2003) illustrated, the successful application of DOE to improve the performance of engineering processes requires planning, statistical, teamwork and engineering skills. Generally a DOE study is performed in three stages: planning, conducting and analysis and interpretation, of which the first is the most important by far.

2.1 The Planning Stage

This consists of seven steps; namely recognising the problem or the improvement opportunity, stating the objectives, selecting the performance measure(s) and the measurement system(s), selecting the factors that may influence the chosen performance measure(s), selecting levels for the factors, selecting interactions that may be important, selecting the experimental design and assigning the factors and interactions to the selected design.

1. *Recognising the problem or the improvement opportunity*: the need for conducting a DOE study to deal with an existing problem or to attain better performance should be understood. This can be achieved by developing a clear and succinct description of the issue or the difference between the current level of performance and the target value. Failure Mode and Effect Analysis (FMEA), Process Flow Diagram, Scrap and Pareto Analysis are among the tools that can be utilised in this step. It is important to involve all the parties concerned with the process;. in fact a team should be formed involving all those who may be able to contribute key information about it. Those whom the experiment and its ramifications may affect should also be involved. As Peace (1993) suggested the team may include a DOE specialist, a Process Engineer, a Management Representative, a Customer, and Operating Personnel. A detailed discussion of the importance of teamwork in designing and carrying out experiments can be found in Van Matre and Diamond (1996).
2. *Stating the objectives*: the objectives should be clear, specific, measurable and of practical value. They should incorporate the target performance level as specified utilising customers' input and competitive benchmark information (Ross, 1996). To be of practical value, there should be a novelty aspect to the experiment, such as proposing new operating conditions for the process or suggesting the use of a new material. All

interested parties should agree that the proper objectives have been set, such tools as Brainstorming, Quality Function Deployment (QFD), and Pareto Analysis are very useful here. In order to secure management approval of the objectives, it is better to express the expected outcomes in monetary terms utilising Taguchi's loss function. This is also important for justifying that the benefits to be gained from the experiment will exceed the expenses incurred. A Gantt chart showing the steps to be followed to attain the objectives along with their associated dates is also advantageous.

3. *Selecting the performance measure(s) and the measurement system(s)*: This step includes determining the performance measure (response variable), the appropriate measurement system and the personnel who will perform the measurements. Process Flow Diagrams and reproducibility studies are useful aids for this step (Ross, 1996). Coleman and Montgomery (1993) suggested that response variables should

   i), whenever possible, be continuous. This is because binary and ordinal data carry much less information and continuous data measured on a well-defined numerical scale are typically easier to analyse.

   ii) capture a quantity or quality of interest for the experimental unit

   iii) be in appropriate units, for example absolute, such as kilograms, or relative, such as percentage of concentration by weight or by volume or proportional deviation from a standard.

   iv) be associated with a target or desirable condition (which motivates the experiment). Continuous and ordinal responses can be classified according to the experiment's objectives into three categories: Nominal-the-Best (the objective being to achieve a target level of performance), Larger-the-Better (to maximise the value of the performance measure) and Smaller-the-Better (to minimise the value of the performance measure).

v) preferably be obtained by non-destructive and non-damaging methods so that repeated measures can be made and measurement error can be quantified

A thorough understanding of the process is essential to obtain a response variable that satisfies most of these criteria. This can be accomplished utilising

- information from previous experiments (if any),
- observational data that may have been collected routinely by process operating personnel,
- field quality or reliability data,
- knowledge based on physical laws or theories, and expert opinion.

This process understanding will be helpful not only for selecting an appropriate response variable but also for quantifying what new knowledge could be gained from the experiment and for motivating discussion by all team members. In order to understand the variation due to the measurements that are to be taken, it is essential to define its system. This incorporates the identification of what to measure, the units, and where and how to measure (Antony and Preece, 2002). It is also important to have a well-established system of ensuring both accuracy and precision of the measurement methods so that the amount of error introduced by the equipment used, for example, can be appreciated. If this is large relative to a change that is important to detect in the response variable, then a measurement system's capability study should be performed to improve the system.

4. *Selecting the factors that may influence the chosen performance measure(s)*: in this step a list of factors to be evaluated for their effect on the selected response variable(s) should be chosen. Brainstorming with product and process technical experts, Process Flow Diagrams, Statistical Process Control Charts, Product Design Specifications and Process Control Plans and Cause and Effect Diagrams are useful aids here. Various

characteristics of the studied factors need to be considered as they can impact the choice of the experimental design. For example, the factor may be hard to change, e.g. after changing its settings, it may take some time to stabilise at the new settings as is the case with furnace temperature. Furthermore, the factor may be hard to set accurately i.e. its actual levels used in the experiment may be different from the intended ones. The experiment's objectives and the information regarding the characteristics of the factors are valuable inputs to enable informed decisions to be made, such as whether the factors should be treated as design or noise and which should be varied and which held constant.

5. *Selecting levels for the factors*: in this step the number and values for all the levels of the selected factors are determined. A thorough knowledge of the product or process is vital to ensure that appropriate level values are selected. The process specification or operating limits are also useful aids (Ross, 1996). To evaluate the effect of each factor, a minimum of two levels is required. If the objective of the experiment is to screen a few factors out of the many possible that actually have a significant effect on the response variable then it is recommended that the experiments use only two levels where possible to minimise its size (Montgomery, 2010). For quantitative factors, the levels should be far enough apart to allow the effect to be detected. However, as Montgomery (2010) noted, it must be borne in mind that by choosing levels that are too far apart a factor can be made to look significant. Conversely, a factor can be made to look insignificant by choosing levels that are too close together. If curvature is expected, three or more levels are required to be examined. In such cases, it is recommended that equal intervals between levels should be used. The flexibility in choosing the levels of qualitative factors is limited. For example if one of the experiment's objectives is to compare three types of machine tooling then the

experimenter has no choice but to include the factor at three levels. Otherwise, the experiment's objective should be amended.

6. *Selecting interactions that may be important*: Product and process expertise should be utilised to identify all the interactions that are suspected to affect the response. Failure to recognise the presence of a significant interaction can result in arriving at misleading conclusions.
7. *Selecting the experimental design*: This step involves choosing an appropriate array to represent the design layout. The columns of the array accommodate the factors and interactions and the rows contain the combination levels. Arrays can be classified as either orthogonal or non-orthogonal the former being most common in practice. An Orthogonal Array (OA) is a matrix whose columns have the property that in every pair of columns, all the possible combinations of levels occur an equal number of times. Any array that does not possess this characteristic is called non-orthogonal. OAs were initially introduced by Jacques Hadamard in 1897 (Ross, 1996). Denoting the number of factors under study by k, the size of the experiments (number of rows) by n and the number of levels by S, a symmetrical OA has associated with it S^k-1 degrees of freedom. These can be split into (S^k -1)/(S -1) mutually orthogonal sets of (S -1) degrees of freedom (Bose, 1947; Bose and Bush, 1952). The term degrees of freedom refers to the number of independent units of information in a sample (in this case the experimental results) relevant to the estimation of a parameter or the calculation of a statistic (Everitt, 2006).

   In an asymmetrical (mixed level) OA($S^{k1}P^{k2}$) each S-level column has associated with it S-1 degrees of freedom and each P-level column has associated with it P-1 degrees of freedom. The total number of degrees of freedom is K_1x (S -1) + K_2 x (P -1) where K_1 and K_2 are the number of the S-level and P-level factors respectively. In conventional

DOE, OAs are denoted by S^k (e.g. 2^3 and 3^2) whereas the denotation L_n (e.g. L_8 and L_9) is used in Taguchi DOE. L_{18} ($2^1 3^7$) is an example of a mixed level OA. Further details on OAs can be found in Addelman and Kempthorne (1961) and Addelman (1962a, 1962b). In practice, it is commonplace to classify OAs into two categories, conventional and Taguchi, however they are equivalent. To shed light on this, consider the conventional 2^3 design and Taguchi's L_8 OAs shown respectively in Tables 1 and 2. As is the case in all two-level experiments, the low and high levels of each factor in the conventional design are represented by the numbers -1 and +1 respectively. Its interaction columns are obtained by multiplying the corresponding columns of the main effects. In the Taguchi L_8 design, 1 and 2 are used to respectively denote low and high levels of each factor. As can be seen the main factor columns of the conventional design are in the opposite order to that in the Taguchi one. In fact the Taguchi L_8 can be obtained from the conventional one by reversing the signs of the 1s in the interaction columns; rearranging the columns as 4,2,6,1,5,7, and 3; and re-labelling -1 as 1 and +1 as 2. In general, two orthogonal arrays are defined to be equivalent if one can be

Table 1: Conventional L8

Col. No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Runs	**A**	**B**	**AB**	**C**	**AC**	**BC**	**ABC**
1	-1	-1	1	-1	1	1	-1
2	1	-1	-1	-1	-1	1	1
3	-1	1	-1	-1	1	-1	1
4	1	1	1	-1	-1	-1	-1
5	-1	-1	1	1	-1	-1	1
6	1	-1	-1	1	1	-1	-1
7	-1	1	-1	1	-1	1	-1
8	1	1	1	1	1	1	1

Table 2: Taguchi's L8

Col. No.	1	2	3	4	5	6	7
Runs	**A**	**B**	**AB**	**C**	**AC**	**BC**	**ABC**
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

obtained from the other through permuting the rows, the columns, or the levels within a column (Bose and Bush, 1952). Kacker et al (1991) explained how Taguchi's OAs were constructed and showed that they are equivalent to conventional ones. The selection of the appropriate OA is dependent upon the number of factors to be studied, the number of important interactions, the size of the factorial effect that needs to be detected, the available time and the resource limitations. One option is to conduct a full factorial trial where all the possible combinations of factor levels are tested. However, this can lead to a prohibitively large experiment as the number of factors increases. To alleviate this, fractional factorial designs in which only a subset of the full factorial runs is performed are commonly employed in practice. Useful though these can be, their adoption gives rise to another type of problem regarding the estimability of the studied effects (see Chapter 4).

8. *Assigning the factors and interactions to the selected design*: having selected the appropriate OA, the next step is to assign factors to it. With full factorial experiments, this is straightforward; however it is not as clear cut in the case of fractional factorial experiments. The reason is that each way of assigning factors to the columns of an OA generates a certain pattern of aliasing which can be defined as the failure to obtain an independent estimate of some or all the factors under study. The topics of aliasing and

of how to assign the factors to an OA so as to reduce its impact are dealt with in Chapter 4.

A collection of graphical methods that can be utilised in the planning phase is presented in Barton (1997). A more detailed discussion of each step can be found in Coleman and Montgomery (1993) where a very useful set of guidance sheets is provided.

2.2 Conducting Stage

At this stage, the experiment is carried out as planned in the previous stage. Before embarking on performing the experiment, it is important to ensure that the necessary materials, machines, operators and other resources are available. It is also important to define the roles and responsibilities of those who are involved in performing the experiment. Coleman and Montgomery (1993) suggested conducting a few trial runs or pilot runs prior to starting the experiment. These provide information about the consistency of the experimental material, a check on the measurement system, a rough idea of experimental error, and a chance to practice the overall experimental technique. It may also lead to revisiting the decisions made in the previous steps.

2.3 Analyses and Interpretation Stage

The experimental results can be analysed using informal graphical methods, formal statistical techniques or a combination of both. Graphical methods involve plotting the main effects and the factors' interactions. In the main effect plot, the averages of all the observations at each level of the factor are plotted and connected by a line. In the case of two-level factors, the vertical height of the line is the difference between the two averages which is the main effect. For the interaction plot, the levels of one factor are displayed on the horizontal axis of the graph. For each of these, the average associated with each level of the second factor is plotted as a point and the points that represent each level of this are connected by a line. The main objective of plotting the main effects and interactions is to identify the strongest effects and determine the combination of factor levels that can produce the most desirable results. Therefore, it provides a valuable input to the process of analysing and interpreting the experimental results. The formal statistical techniques involve the t-test, ANOVA and regression modelling. Their main purposes are to test the statistical significance of each effect and derive an empirical model from the experimental data expressing the relationship between the response variable and the important factors and interactions. A detailed discussion of these techniques is presented in Appendix 2. Once the experimental data have been analysed, a practical interpretation should be provided utilising the process technical expertise on the basis of which conclusions about the results can be drawn leading to recommending a certain course of action. Confirmation runs should then be performed to validate the conclusions.

REFERENCES

Cobb, G. (1997), *Introduction To Design And Analysis Of Experiments,* Springer New York.

Montgomery, D. (2010), *Design And Analysis Of Experiments,* Wiley

Box, G., Hunter, S. and Hunter, W. (2005), *Statistics For Experimenters: Design, Innovation, And Discovery,* Wiley.

Box, G. and Meyer, R. (1986), "An Analysis For Unreplicated Fractional Factorials", *Technometrics*, Vol. 28, No**.** 1, pp. 11-18.

Juran, J. (1993), *Quality Planning And Analysis: From Product Development Through Use,* McGraw-Hill

Wu, C. F. J. and Hamada, M. (2000), *Experiments: Planning, Analysis and Parameter Design Optimization,* John Wiley.

Antony, J. (2003), *Design Of Experiments For Engineers And Scientists,* Butterworth-Heinemann.

Peace, G. (1993), *Taguchi Methods: A Hands-On Approach,* Addison Wesley Publishing Company.

Van Matre, J. and Diamond, N. (1996), "Team Work and Design of Experiments", *Quality Engineering*, Vol. 9, No**.** 2, pp. 343–348.

Ross, P. J. (1996), *Taguchi Techniques for Quality Engineering*, McGraw-Hill.

Coleman, D. and Montgomery, D. (1993), "A Systematic Approach To Planning For A Designed Industrial Experiment", *Technometrics*, Vol. 35, No**.** 1, pp. 1-12.

Antony, J. and Preece, D. (2002), *Understanding, Managing, And Implementing Quality: Frameworks, Techniques, And Cases*, Routledge.

Bose, R. (1947), "Mathematical Theory Of The Symmetrical Factorial Design", *Sankhy : The Indian Journal of Statistics*, Vol. 8, No**.** 2, pp. 107-166.

Bose, R. and Bush, K. (1952), "Orthogonal Arrays Of Strength Two And Three", *The Annals of Mathematical Statistics*, Vol. 23, No**.** 4, pp. 508-524.

Everitt, B. (2006), *The Cambridge Dictionary Of Statistics*, Cambridge University Press Cambridge.

Addelman, S. and Kempthorne, O. (1961), "Orthogonal Main-Effect Plans", *Aeronautical Research Laboratory Technical Report*, Vol. 79, No**.** 2, pp. 3.

Addelman, S. (1962 a), "Symmetrical and Asymmetrical Fractional Factorial Plans", *Technometrics*, Vol. 4, No**.** 1, pp. 47-58.

Addelman, S. (1962 b), "Orthogonal Main-Effect Plans For Asymmetrical Factorial Experiments", *Technometrics*, Vol. 4, No**.** 1, pp. 21-46.

Kacker, R., Lagergren, E. and Filliben, J. (1991), "Taguchi's Fixed-Element Arrays Are Fractional Factorials", *Journal of Quality Technology*, Vol. 23, No**.** 2, pp. 107-107.

Barton, R. (1997), "Pre-experiment Planning For Designed Experiments: Graphical Methods", *Journal of Quality Technology*, Vol. 29, No**.** 3, pp. 307-316.

APPENDIX 2:
STATISTICAL ANALYSIS OF EXPERIMENTAL DATA

1. Introduction

Broadly, the statistical analysis of experimental data comprises two stages: estimation and hypothesis testing. In the context of DOE, the former relates to estimating the effect size of the studied factors and interactions and the latter pertains to testing their statistical significance. The procedure for performing both is dependent upon the number of factor levels, whether the trial is replicated and whether regression modelling is performed. The aim of this Appendix is to present an overview of the statistical analysis of two- and three-level experimental data. A brief discussion of the estimation and hypothesis testing concepts is firstly presented. Then, the most commonly used statistical significance tests i.e. t-test and ANOVA are explained together with the way in which regression technique can be employed. Some limitations are highlighted regarding the literature treatment of analysing three-level experiments and their level coding systems. The way in which the regression coefficients should be interpreted under each coding system is discussed. Attention then centres on the analysis of unreplicated experiments including the most commonly used pooling strategies since in such cases no degrees of freedom are available to estimate the error variance. Finally the main conclusions are summarised.

2. Estimation and Hypothesis Testing

Estimation is the process of providing a numerical value for a population parameter on the basis of information collected from a sample (Walpole et al, 2006). Point estimation is the process whereby a single figure (a statistic) is calculated for the unknown parameter, while interval estimation is a procedure for setting bounds within which the parameter is likely to lie. Assessing whether sample data is consistent or otherwise with assumptions made about the population is known as hypothesis testing. It involves stating a null hypothesis (H_0) regarding the studied population along with an alternative (H_1). H_0 is a supposition or a statement expressing a certain expectation regarding the population. It usually concerns a parameter and is represented in such forms as “no effect”, “no difference” or “no association” (Kanji, 2006). To test the null hypothesis on the basis of sample data, it is necessary to establish an inferential linkage between the sample and the population. This is accomplished by means of a sampling distribution which is the probability distribution of a statistic calculated from repeated random samples of a particular size collected from a population about which the null hypothesis is true. The procedure of testing a hypothesis consists of firstly stating the null hypothesis. Then, on the basis of its relevant parameter, the appropriate statistic should be specified and its probability distribution under the null hypothesis should be defined. The next step is to collect a random sample and calculate the concerned statistic. Utilising its sampling distribution, the probability of obtaining a statistic as extreme as or more extreme than the observed one is then calculated (Devore, 2008). This is referred to as the p-value and is used as a quantitative measure of the plausibility of the null hypothesis. Whenever the p-value is less than a particular threshold the result is said to be significant meaning that the null hypothesis should be rejected. This threshold relates to one of the two errors that are possible whenever hypothesis testing is performed i.e. Type I and Type II errors. The former occurs when a true null hypothesis is rejected while the latter arises in the case of failing to reject a false null hypothesis (Black, 2009).

Before testing a certain hypothesis, the experimenter should specify the probability of a Type I error which is conventionally called the significance level and represented by α. It is against this probability that the p-value is examined and the decision regarding the rejection of the null hypothesis or otherwise is made. The value of α is determined by subjective judgment about the acceptable level of such error given the context of the study. A common rule of thumb is to use $\alpha = 0.05$. Given the value of α, the standardised effect size, the sample size and the error degrees of freedom, the probability of a Type II error commonly denoted by β can be determined. In general it is an inverse function of the three aforementioned parameters. The complement of the Type II error is a measure of the extent to which the hypothesis testing process is successful i.e. the Power. It can be defined as the probability of correctly rejecting the false null hypothesis which is equivalent to 1- β (Downing and Clark, 2003).

In the fixed effect DOE analysis, which is the theme of this Appendix, the null hypothesis may take two general forms. The first is associated with single factor experiments where the hypothesis of no difference between the factor levels is the main concern. The second involves two or more factors where the null hypothesis of no individual or collective effect of the studied factors and interactions is commonly tested. Figure 1 summarises the conventional procedures of testing the latter form of null hypothesis. Clearly the student (t) and Fisher (F) statistical tests along with regression analysis, pooling and variable selection methods are key techniques in hypothesis testing.

3. t-Statistical Test

Before discussing the t-test, it is instructive to present a theory upon which most of the statistical methods rely. This is the Central Limit (CL) theory which, in one of its forms, states that if a large enough sample is drawn from a population with mean μ and variance σ^2 then the distribution of the sample average $\overline{X}$ is approximately Normal with mean μ but with variance σ^2/n no matter from what population it was drawn (Ross, 2004). As any linear combination of Normally distributed variables is itself Normally distributed, the Central Limit theory is applicable to any linear combination of $\overline{X}$ such as $\overline{X}_1 - \overline{X}_2$. The necessary sample size to satisfy these conditions is dependent on the distribution shape of the underlying population. A symmetrical distribution such as the Normal requires a fairly small sample size whereas a large one may be required for a heavily skewed population (Cramer, 1999; Stuart and Ord, 2009). Box et al (2005) however, noted that irrespective of the

.

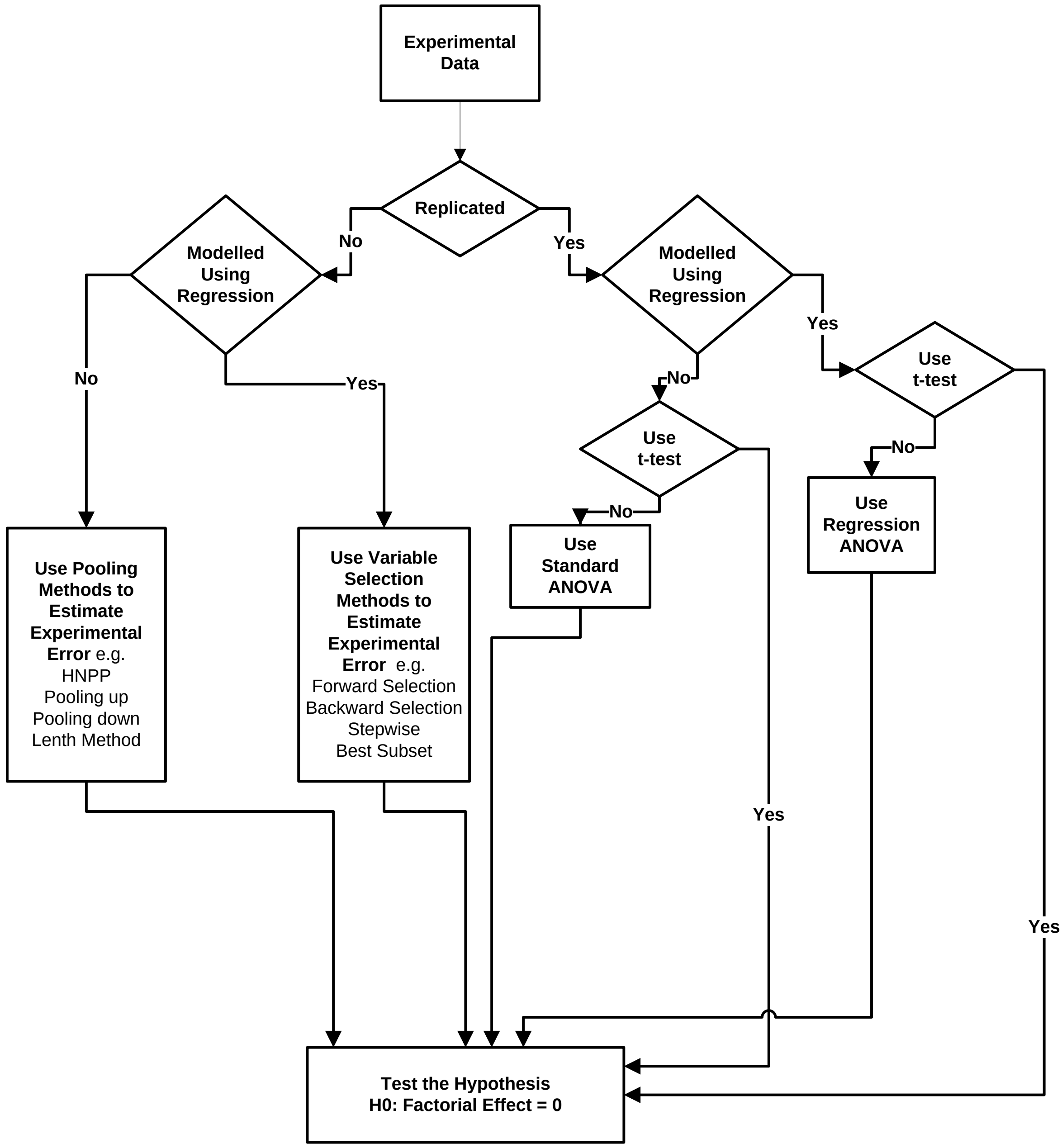


Figure 1: Conventional Approaches for Testing H_0: Factorial Effect =0

distribution of the individual observations, the statistical methods that depend on the distribution of their averages tend to be insensitive to non-Normality. They also observed that in practice, σ^2 is almost always unknown. Consequently, its estimate S^2 is used. If this is calculated from a fairly large sample size (30 or more as a general rule of thumb) its value will be very close to σ^2 and the distribution of $\overline{X}$ will be approximately Normal. However, for small sample sizes, S^2 may not be close to σ^2 and $\overline{X}$ conforms to the Student-t distribution with mean μ and variance S^2/n. Since it depends on how "reliable" the value of S^2 is, its shape is dependent upon the number of degrees of freedom (υ) used to calculate S^2. In DOE, the t-test is important for assessing the significance of the studied effects when expressed as differences between the sample averages of the response at each of the factor levels.

3.1 Two-Level Experiments

In the case of single factor experiments at two levels, the t-test can be used to test whether the difference between the averages associated with the two levels is statistically significant. For replicated factorial experiments at two levels, there are two forms in which the t-test may be used to test the following hypothesis:

H_0: Factorial Effect = 0

H_1: Factorial Effect ≠ 0

The first is associated with replicated experiments that are analysed using regression (see section 5.2.1) and the second to cases where the experiment is replicated and no regression analysis is employed. The procedure for using the t-test in such cases is as follows (Barrentine, 1999) :

- Estimate the effect of each factor and interaction under study. This is performed by subtracting the average of the experimental results associated with the low level from those that are associated with the high level for each effect.
- Calculate the sample variance S_i^2 for each of the replicated runs using the following equation:

$$S_i^2 = \frac{\sum(y_i - \bar{y}_i)}{r - 1} \tag{1}$$

where r is the number of replications and $\bar{y}_i$ is the average of the replicated responses at the i^{th} run

- Calculate an estimate of the experimental error variance S_e^2 by averaging the sample variances
- Calculate the variance of the effects:

$$S_{eff}^2 = S_e^2 \frac{4}{nxr} \quad (2)$$

- Calculate the degrees of freedom associated with the error variance:

$$df = n \text{ x } (r-1) \quad (3)$$

- Select a value of α and obtain the $t_{\alpha/2,df}$ from tables of the "t" distribution
- Using the null hypothesis value (0) and the $t_{\alpha/2,df}$ along with the S_{eff}^2,, calculate the decision limits also called the Confidence Limits (CL):

$$CL = 0 \pm t_{\alpha/2,df} \text{ x } S_{eff}^2 \quad (4)$$

  The calculated effects that fall within the above computed limits are deemed inert whereas those that fall outside it are pronounced significant. The interval in equation 4 is called the Confidence Interval (CI) and its level is obtained by subtracting the α from 1 and multiplying the result by 100. If α = 0.05, then the confidence level is (1-0.05)x100 =95% and the CI is interpreted as follows: if the same experiment is repeated 100 times, it is expected that the truly inert effects will fall within the interval 95 times.

In the case of unreplicated two-level experiments, specific versions of the t-test can be employed as part of the Lenth pooling method (see Section 6.2).

3.2 Three-Level Experiments

For single factor three-level experiments, all possible pairs of the level-means can be compared using the t-test. However, such multiple test procedures can lead to inflating the Type I error as discussed in Section 5.2.2. Thus the ANOVA technique should be used instead. With regard to replicated and unreplicated factorial three-level experiments, the use of the t-test requires each factor and interaction to be decomposed into single degree of freedom elements using an appropriate coding system. As this step is performed as a part of regression analysis, its discussion is presented in Section 5.1.2.

4. ANOVA

ANOVA is a method of decomposing the observed total variance into components due to different sources of variation. It utilises the statistical theory that if two random samples of size n_1 and n_2 are taken from two different populations then the ratio of their sample variances (S^2_1/S^2_2) follows an F distribution with n_1-1 and n_2-1 degrees of freedom. In DOE, the use of ANOVA depends on whether the experiment is a fixed or random effect. As the scope of this study precludes random effect experiments, this Section is concerned with fixed effect ANOVA.

In replicated single factor experiments, ANOVA is referred to as one-way and is applied when the number of levels is three or more. In such cases, two variances are estimated: one for the variance in the means of the levels and the other for the variance within them combined for all the levels. Under the null hypothesis of no difference between the levels, these two estimates should be very similar rendering an F ratio close to 1. However, if the level means differ substantially, then their variance should be greater than the combined within sample variances yielding a large F ratio.

Table 1 shows the ANOVA table for a two-factor replicated factorial experiments, (Montgomery 2010). The first column shows how the total variation is broken down into its constituent elements. The second column shows the Sums of Squares (SS) for each factor and interaction. Regarding the factors, this can be obtained by subtracting the average of the experimental results associated with each of its levels from the overall average of the experimental results and computing the sum of the squares of the resultant difference values.

Table 1: ANOVA Table the Fixed Effect Two-Factor Experiment (Montgomery, 2010)

Sources of Variation	**Sum of Squares**	**Degree of Freedom**	**Mean Squares**	**F-value**
Factor A	SS_A	a-1	$MS_A=SS_A/a-1$	MS_A/MS_E
Factor B	SS_B	b-1	$MS_B=SS_B/b-1$	MS_A/MS_E
Interaction AB	SS_{AB}	(a-1)(b-1)	$MS_{AB}=SS_{AB}/(a-1)(b-1)$	MS_A/MS_E
Error	SS_E	ab (r-1)	$MS_E=SS_E/ab\ (r-1)$	
Total	SS_T	abr-1		
a: no of factor A levels;		b: no of factor B levels;		r: no of replications

The interaction sum of squares can be calculated in two stages. Firstly the average of the runs associated with each of the level combinations of the interacting factors should be subtracted from the overall average and the squares of the resultant differences summed. Secondly, the SS of the factors that comprise the interaction should be subtracted from the value arrived at in the first stage yielding the interaction SS. The error SS can also be calculated in two stages. In the first the average of the replicated runs at each level combination should be computed and then subtracted from the values of the individual runs at its level combination. The resultant differences should then be squared and summed up over all the level combinations to arrive at the error SS. The total SS can be obtained by adding together the SSs due to the factors and interactions under study and the error SS.

Each is divided by its degrees of freedom to yield the variance or mean square (MS) as shown in the fourth column of Table 1. For fixed effect models and under the null hypotheses of no factorial effect, each of the mean squares in Table 1 is an estimate of the true experimental error σ^2. Thus when any of them is divided by the MS_E as illustrated in the fifth column of Table 1 the resultant F ratio should be close to one. However, if the effect of, say, factor A is large then a larger F ratio will result. This should conventionally be compared with the tabulated $F_{\alpha,a-1,\ ab\ (r-1)}$ value and if it exceeds it, factor A is deemed statistically significant. Alternatively, the p-value corresponding to the calculated F can be estimated and if it is found to be smaller than α then the same result applies.

Although there are simpler formulae for computing the SS in two-level experiments, the aforementioned procedure is applicable to both two- and three-level trials. In the case of unreplicted experiments, there are no degrees of freedom available for estimating the error variance. To deal with this, one approach is to assume that the high order interactions are inert and use their degrees of freedom to estimate it. This can result in misleading conclusions as the effect of some high order interactions may be large. Alternatively, pooling methods discussed in Section 6 may be used.

4.1 Assumptions of ANOVA

For the conclusions drawn from the ANOVA to be valid, certain assumptions must be satisfied. These are related to the residuals which are estimates of the experimental error obtained by subtracting the observed responses from the predicted ones. The latter are calculated from the derived model (as discussed in Section 5) after all the coefficients have been estimated from the experimental data. For the conclusions of an experiment to be valid, the residuals must be (approximately) Normally and independently distributed with a zero mean and constant variance. In general the error is the failure to obtain exactly the same response value when replicating the experiment at exactly the same settings of its conditions. It is a function of a number of component errors such as those due to measurement, environmental conditions and other factors not explicitly included in the experiment. As the overall error can be viewed as the summation of these components, it should, according to the CL theory, be Normally distributed. In fact, the CL theory states that for a large enough random sample the sum of the sample observations is Normally distributed irrespective of the individual distribution of the observations. The reason for the zero mean is to ensure that the positive error component values cancel out the negative ones thereby eliminating any systematic effect on the average response value. The constant variance requirement ensures that all the response values corresponding to each of the factor settings are equally reliable -

reliability being judged by how variable the response values are around their predicted values. The assumption of Normality can be assessed using histograms or more formally using Normal probability plots. Plotting the residuals in time order of data collection is helpful in detecting any correlation between the residuals and thereby verifying the independence assumption. The plot should not demonstrate any pattern or trend. The constant variance assumption can be assessed by plotting the residuals against their predicted values.

The presence of any unusual pattern violates this assumption. Transformation techniques can then be used to overcome the problem. Their need can be recognised using a Cox-Box plot (Box and Cox, 1964), which is constructed using the power family of transformations $y^* = y^\lambda$, where λ is the power to be detected. The procedure consists of performing ANOVA on y^λ for various values of λ to obtain its maximum likelihood estimate for which the error sum of squares (SS_E) is minimised. Its outcome can be presented graphically by plotting the SS_E values against those of λ and locating the minimum. A confidence interval around the identified λ value should then be constructed. If this contains 1 then no transformation is needed. Otherwise the appropriate transformation can be determined on the basis of the value of λ as shown in Table 2.

Table 2: Transformations Based on λ Values

λ	**Transformation**
0.5	Square Root
0	Log
-0.5	Reciprocal square root
-1	Reciprocal

5. Regression Analysis

In fixed effect DOE, Regression analysis is a powerful tool for studying the dependence of the response variable on one or more factors and interactions the objective being to estimate their effect size, test their statistical significance and predict the average response

(Kutner et al 2004). In general the relationship between the response variable (y) and the factors (X_i) and interactions (X_iX_j) under study may be represented by the following model:

$$Y=\beta_0+\sum_{i}^{k}\beta_i X_i + \sum_{i}^{k}\beta_{ii}X_i^2+\sum_{ij}^{k}\beta_{ij}X_iX_j+e \qquad (5)$$

The term e represents the residuals which, for reasons already explained, are Normally and independently distributed with a zero mean and constant variance. The above equation describes a plane, the dimension of which is equal to the number of factors included in the equation (Rawlings et al 1998). Although non-linear in terms of the variable, it is a linear model in the parameters which are conventionally estimated using least squares procedures. The objective is to find estimate values for the β_is so that the error e is minimised. The resultant estimated model is as follows

$$\hat{Y}=\hat{\beta}_0+\sum_{i}^{k}\hat{\beta}_i X_i + \sum_{i}^{k}\hat{\beta}_{ii}X_i^2+\sum_{ij}^{k}\hat{\beta}_{ij}X_iX_j \qquad (6)$$

where $\hat{\beta}_0$ defines the estimated intercept of the modelled plane. The coefficients ($\hat{\beta}_i$s, $\hat{\beta}_{ii}$s, $\hat{\beta}_{ij}$s) provide estimates for the studied factorial effects.

5.1 Estimating the Regression Coefficients

Before discussing the estimation of the coefficients of a regression model, it is insightful to distinguish between: the design and the model arrays. For an experiment with m factors, the design array consists of at least m columns corresponding to the m factors irrespective of the model to be fitted and its number of parameters. Furthermore, it conventionally lists the factor level combinations in terms of their actual values or numbers or symbols representing them. The model array, on the other hand, is a coded matrix for the design that has a column for each parameter in the intended model. Thus if an m x n design array is used to fit a model with r parameters then its model array will be an r x n array whose columns represent the effects of the model parameters.

The least squares estimators of the β_i in equation 6 can be obtained using the following formula (Montgomery et al 2006):

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathbf{T}}\mathbf{Y} \quad (7)$$

where $\hat{\boldsymbol{\beta}}$ is the coefficient vector, $\mathbf{X}$ is the model array, $\mathbf{X}^{\mathbf{T}}$ is its transpose and $\mathbf{Y}$ is the responses vector. Clearly the values of $\hat{\beta}_i$ are reliant upon the model array and its coded entries. Depending on whether the $(\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}$ matrix is diagonal, the model arrays may be classified into orthogonally and non-orthogonally coded arrays. The former have a diagonal $(\mathbf{X}^{\mathbf{T}}\mathbf{X})^{-1}$matrix whereas the latter have a non-diagonal one. The departure from diagonallity leads to $\hat{\beta}_i$ estimates that are aliased and associated with high variance. Consequently, it is recommended that the model matrix should be coded orthogonally so that the estimated $\hat{\beta}_i$s are orthogonal.

5.1.1 Two-Level Arrays Coding

There are several ways in which two-level arrays may be coded. Examples include using 0 and 1 or 1 and 2 to represent the two levels of the studied factors - the latter being the case in Taguchi two-level arrays. One problem with these coding methods is that although the design arrays are orthogonal they render non-orthogonally coded model arrays thereby generating non-orthogonal high variance estimates of the $\hat{\beta}_i$s. An alternative coding scheme is to use -1 and +1 for the two levels of each factor. The principal advantage of this scheme is the orthogonality of its resultant model array and the minimum variance associated with its estimated $\hat{\beta}_i$s. A further advantage relates to the interpretation of the $\hat{\beta}_i$s as the amount by which the average response variable changes per unit change in X_i. The common method of estimating the effects of a two-level factorial effect is to subtract the average of the response values associated with the -1 level from that of those associated with +1 i.e. over two units (-1 to 0 and 0 to 1). Therefore, each of the $\hat{\beta}_i$s estimated using the (-1,+1) coding system is one-half the conventionally estimated factorial effect. This gives the (-1,+1) coding system the advantage of being simple in terms of its estimated coefficients' interpretation.

5.1.2 Three-Level Arrays Coding

The two-degrees of freedom associated with a three-level factor can be split up into two components, each carrying one degree of freedom, in various coding ways. As was the case in two-level coding, each scheme renders a different interpretation of the estimated $\hat{\beta}_i$s. The choice of an appropriate coding system depends on such factors as the effect to be estimated, the comparison to be made and whether the studied factor is qualitative or quantitative. Table 3 lists the four most common coding alternatives for the rendition of the two degrees of freedom associated with a three-level factor. For the purpose of referencing, they are called

Table 3: Four Systems for Coding Three-Level factors

	Coding ID	**Dummy Variable**			**Level-Square**			**ANOVA-Model**			**Linear-Quadratic**	
	Column No	1	2		3	4		5	6		7	8
		X_1	X_2		X_1	X_2		X_1	X_2		X_1	X_2
Levels	**1**	1	0		-1	1		1	0		-1	1
	2	0	1		0	0		0	1		0	-2
	3	0	0		1	1		-1	-1		1	1

the Dummy Variable (DV), Level-Square (LSQ), ANOVA-Model (ANOVA-M) and Linear-Quadratic (L-Q) systems. There are several limitations associated with the presentation of these coding systems in the DOE literature. Firstly, they are rarely if ever discussed; in fact, different references use different coding systems depending on the circumstances. For example, Montgomery (2010) used the ANOVA-M system to handle analyses incorporating three-level qualitative factors whereas those involving quantitative ones were dealt with using the LSQ system. No explanation was provided regarding why a particular coding system was used nor the possible alternatives. Furthermore, Ross (1996) implicitly utilised ANOVA-M in deriving the mean response models. His discussion regarding the L-Q system was limited to how the data could be analysed using ANOVA. He limited the use of such a system to cases where the studied factors were quantitative with evenly spaced levels. No discussion about either the estimation of the mean response model using such a system was provided or the interpretation of the coefficient of such a model. The L-Q system was also discussed in Wu and Hamada (2000) as being an apposite system for coding quantitative factors while Draper and Smith (1998) showed how it could be used to analyse qualitative factors. Such treatment results in practitioners being confused about which coding system should be used and when. A further problem is that different statistical packages employ different coding systems in their analyses; hence for a certain model derived from particular experimental data, different coefficients may be estimated. One of the main reasons for this state of affairs is that the implications of using different coding systems have more profound consequences in experiments that encompass factors at three or more levels, the analyses of which have been accorded substantially lesser emphasis than that placed on two-level experiments. To deal with this an explanation of the four coding systems presented in Table 3 follows in conjunction with the interpretation of the regression coefficients estimated using each.

1. **The DV system**

This system, also called the indicator variable system, involves a successive dichotomising in three-level designs so that each of the coded two levels is distinguished from the remainder as representing one aspect of the studied factor. For example column 1 of Table 3, shows that the results associated with level 1 of the appropriate three-level factor are assigned 1 in the model array whereas the rest are assigned 0. Consequently this column carries only some of the information (one degree of freedom) of the studied factors. Level 2 is coded in the same manner. Hence, columns 1 and 2 of Table 3 exhaust the information associated with the coded three-level factor. In fact the third level is represented by zero in columns 1 and 2. This combination can be interpreted as not level 1 and not level 2 leading to level 3. Although predominantly used with qualitative factors, there is no reason why the DV system could not be used with quantitative factors as long as the implications of its use are understood. The interaction column of any two or more three-level factors can be obtained by multiplying the individual elements of the columns corresponding to the interacting factors. If a model array is constructed using this system and equation 7 used to estimate the coefficients of a fitted regression model, each should be interpreted as follows. For main effects, each $\hat{\beta}_i$ represents the net contribution or consequence of level 1 or 2 relative to the excluded level i.e. 3. For example, if $\hat{\beta}_1$ is, say, 2, this means that using level 1 results in increasing the average response by 2 units over its value corresponding to level 3. With single factor experiments and when the main effects are the principal concern in multi-factors experiments, this coding system has the advantage of simplicity. It lends itself to experiments involving a control setting as this can be represented by the levels excluded from the model. In this way, the impact of the other level combinations relative to the control one can simply be assessed. As the interpretation of factor interactions is not straightforward, this system might not be the best for studying three-level factors, be they qualitative or quantitative, in factorial experiments when interactions are important and expected.

2. The LSQ system

This system is suitable for coding quantitative factors. It decomposes the three-level factor into linear and quadratic components (columns 3 and 4 in Table 3-3). The former is obtained by associating the highest and lowest levels of all factors under study with 1 and -1 respectively. Any value in between can be computed so that its difference from 1 and -1 is proportional to the difference between the corresponding actual value and the actual highest and lowest values. The quadratic main effect can be obtained by squaring the linear main effect levels. The interactions are each obtained from the entry-wise multiplication of the columns corresponding to the interacting factors. Using this system, the least squares regression coefficient corresponding to each main effect is interpreted as the average change in the response variable per unit change in the linear main effect. As Box et al (2005) noted the coefficient of a two factor interaction, say, $X_A X_B$ is a quantity that measures how the effect of factor A changes as factor B is changed and vice versa. By the same token, the quadratic effect, say, A^2 can be viewed as an $X_A X_A$ interaction. Thus its coefficient is a measure of how factor A changes as its values are changed (Box et al, 2005). This system is the best option for coding quantitative three-level factors involving non-equally spaced numerical levels.

3. The ANOVA-M system

Again two columns are used for each three-level factor in this system. The first is formed by associating 1, 0 and -1 with the levels 1, 2 and 3 respectively while the second assigns 0, 1 and -1 respectively to the levels 1, 2 and 3. The interaction columns are generated using the inner product of their comprising factor columns. To interpret the least squares estimate of the coefficients of a model derived on the basis of this coding system, let $\overline{Y}_{A1}$, $\overline{Y}_{A2}$,and $\overline{Y}_{A3}$ be the average responses associated with levels 1, 2 and 3 of factor A and let $\overline{Y}_{00}$ be the overall average. For each three-level factor, two main effect coefficients $\hat{\beta}_{A1}$ and $\hat{\beta}_{A2}$ are estimated. The former is equal to $\overline{Y}_{A1}$- $\overline{Y}_{00}$ whereas the latter is equal to $\overline{Y}_{A2}$- $\overline{Y}_{00}$. Although not included in the model the effect associated with the third level ($\hat{\beta}_{A3}$) can be obtained using the following equation:

$$\hat{\beta}_{A3} = -(\hat{\beta}_{A1} + \hat{\beta}_{A2}) \quad (8)$$

For interactions, a coefficient is estimated for each of the possible level combinations of the interacting factors. The interaction coefficient of, say, level x of factor A with level z of factor B obtained using equation 7 is equivalent to the following:

$$\hat{\beta}_{xz} = \overline{Y}_{xz} - \overline{Y}_{Ax} - \overline{Y}_{Bz} + \overline{Y}_{00} \quad (9)$$

where $\overline{Y}_{xz}$ is the average of the response variable at the level combination xz. This system has the advantage of being simple in terms of its coefficients' interpretation as each represents the specific difference between the average responses at a certain level or a level combination and the overall average. It is called ANOVA model because it yields a regression model equivalent to the average response model that is conventionally obtained when ANOVA is conducted. While it is commonly used with qualitative factors, it can also be used with quantitative factors when the objective is limited to assessing their effect at the discrete values considered in the experiment.

4. The L-Q system

A common limitation of all the coding systems discussed is that they render a non-orthogonally coded model array. The primary advantage of the L-Q system lies in its ability to yield an orthogonally coded one. Therefore, when used with quantitative factors, it is referred to as the orthogonal polynomial coding system. It splits each three-level factor into linear and quadratic components. The former is generated in the same manner as its correspondent component discussed in the LSQ system. The quadratic column is obtained by associating 1 with levels 1 and 3 and assigning -2 with level 2. One-half the difference between the response averages associated with levels 1 and 3 is estimated by the linear effect regression coefficient. When used to code qualitative factors, the term "linear" has no practical meaning. Therefore, its coefficient should be interpreted as a measure of the difference between levels 1 and 3.

Before interpreting the $\widehat{\beta}$ associated with the quadratic effects, it is important to demonstrate that in the absence of a quadratic effect of, say, factor A the differences $(\overline{Y}_{A1} - \overline{Y}_{A2})$ and $(\overline{Y}_{A2} - \overline{Y}_{A3})$ are approximately the same i.e.

$$(\overline{Y}_{A1} - \overline{Y}_{A2}) - (\overline{Y}_{A2} - \overline{Y}_{A3}) = 0 \quad (10)$$

so

$$(\overline{Y}_{A1} - 2\overline{Y}_{A2} + \overline{Y}_{A3}) = 0 \quad (11)$$

In this case, the quantity in equation 11 is equivalent to the difference between the overall average and the average of the response values associated with A_2 level i.e. $\overline{Y}_{00} - \overline{Y}_{A2}$. The quadratic regression coefficient estimated using the L-Q system is one-half the latter contrast. Again in the case of qualitative factors, the term "quadratic" should not be used and its effect should be interpreted as a measure of the extent to which the average response at level 2 (the middle level) is different from the average of the responses associated with the other two levels. The interaction columns are formed by multiplying the entries of their comprising factors. Their coefficients can be interpreted in two ways depending on whether the interacting factors are qualitative or quantitative. In the latter case the regression coefficient of, say, $X_AX_B^2$ interaction is interpreted as the rate by which factor A varies linearly as factor B changes quadratically or equivalently the rate by which factor B varies quadratically as factor A changes linearly. For qualitative factors, the interpretation of the interaction coefficient entails the use of the conditional main effect concept (Bohrer et al, 1981; Winer and Brown, 1991; Bonett and Woodward, 1993) This demands that the interaction between two factors be expressed in terms of the main effects of one of the interacting factors at each of the levels of the other factor. For example assume that the fitted model is:

$$Y = 2 + 2X_{A1} + 3X_{B2} + 4\,X_{A1}X_{B2} \qquad (12)$$

where X_{A1} and X_{B2} are the first and second coded columns (columns 7 and 8 in Table 3) of the qualitative factors A and B respectively. One way of interpreting the $X_{A1}X_{B2}$ interaction is to substitute the three possible values of X_{A1} (-1, 0, 1) into equation 12 to yield three main effect only equations in X_{B2}. The main effect terms of each of these is then interpreted conditionally on the value of X_{A1} which was used to obtain it. In so doing, the above discussed way of interpreting the qualitative main effects when the L-Q coding system is used is adopted. Alternatively, the possible values of X_{B2} (1,-2) can be used to obtain three main effect only equations in X_{A1} (two of which are identical). Similarly their terms can be interpreted as already described. This method should also be applied in the case of interpreting the qualitative-quantitative factors' interactions. While the L-Q system can be used to study qualitative and quantitative factors, it must be borne in mind that in the latter case the numerical levels must be equally spaced. Otherwise they should either be treated as qualitative factors or coded using the LSQ system.

5.2 Regression Hypothesis Testing

In testing the statistical significance of the regression model coefficients there are generally two approaches; the t-test and the ANOVA.

5.2.1 Regression t-test

The t-statistic can be used to test the hypothesis of zero effect where each effect size is represented by a regression coefficient $\hat{\beta}_i$. From equation 5, it is clear that, when the X values are assumed to be fixed (fixed model), the only random element is the error (e). The response Y along with the β_is are linear functions of e. Consequently, as the latter is assumed to be Normally distributed, the response Y and the β_is are also Normally distributed. Since the error variance is almost always unknown, its estimate MSE (S^2) can be used. Thus, if the null hypothesis of no effect (H_0: $\hat{\beta}_i = 0$) is true then the t-statistic

$$t_0 = \frac{\hat{\beta}_i - 0}{\sqrt{MSE * C_{ii}}} \tag{13}$$

where C_{ii} is the diagonal element of $(\mathbf{X}^T\mathbf{X})^{-1}$matrix, is appropriate for testing its statistical significance. The denominator of equation 13 is called the standard error. Whenever $|t_0| > t_{\alpha/2,dfe}$ or equivalently if the p-value is less than α the null hypothesis is rejected.

5.2.2 Regression ANOVA

When regression analysis is employed in DOE, the ANOVA can be used to perform two types of significance tests; namely, that of the individual factorial effects and that of the overall effect. Before describing the former, recall the following statistical relationship between the F and t test statistics (Allen, 2004):

$$F_{\alpha,1,\, dfe} = t^2_{\alpha/2,dfe} \tag{14}$$

This equality is applicable in regression analysis due to the fact that all the effects that are associated with more than one degree of freedom are decomposed into single degree of freedom components and consequently their F statistic’s first degree of freedom is one. Another way of viewing the equality in equation 14 is to express the F value in terms of a ratio between the SS of each component (SSc) and the MSE and substitute the t value in equation 13 for that in 14. This gives:

$$\frac{SSc}{MSE} = \frac{\hat{\beta}_c^{\,2}}{MSE * C_{ii}} \tag{15}$$

Multiplying both sides of equation 15 by MSE gives

$$SSc = \frac{\hat{\beta}_c^{\,2}}{C_{ii}} \tag{16}$$

Therefore, the SS of the coded components can be obtained using equation 16.

Given the SS due to each term in the regression model and their degrees of freedom, the same procedure of performing the ANOVA presented in Section 4 can be implemented to estimate the MSE and the statistical significance of the individual factorial effects under study.

Alternatively, ANOVA can be used to test the joint significance of all the studied factorial effects i.e.

$$H_0: \hat{\beta}_1 = \hat{\beta}_2 = ... = \hat{\beta}_k = 0$$
$$H_1: \hat{\beta}_i \neq 0 \text{ for at least one i}$$

To perform this, the SS due to all the factorial effects are aggregated to form the regression SS. By dividing this by the total degrees of freedom associated with the amalgamated effects, the regression MS can be obtained. This should then be divided by the MSE and F-tested as discussed in Section 4.

Of particular importance when choosing between individual and simultaneous significance testing is the distinction between the individual error rate and the experiment wise one. The latter is the α level when one statistical test is performed to test the overall significance of all the studied effects. In such a case the probability of not rejecting a true H_0 is $(1-\alpha)$. However, this is not the case when multiple individual tests are performed. If for example, there are U individual tests to be conducted then the probability of not rejecting a true H_0 is $(1-\alpha)$ for each of the conducted tests. For all the individual tests taken together, the probability of not rejecting a true H_0 for any one of the tests is $(1-\alpha)^U$ (Lee and Comrey, 2009). Consequently the true α level is $1-(1-\alpha)^U$ which exceeds α for $U > 1$ and increases as U increases.

6. Analysis of Unreplicated Experiments

The discussion so far has centred around replicated experiments so that an independent estimate of the experimental error variance can be obtained, however due to economical, technical or time related reasons, it might not be possible to replicate the conducted experiment. Consequently, it becomes necessary to pool the degrees of freedom associated with the smallest studied effects to form an estimate of the experimental error variance. The need for pooling might also arise in replicated experiments where summary statistics such as signal-to-noise ratio or a sample variance are used. These statistics consolidate the replicated runs at each experimental setting into one value that measures their variability yielding a single replicate response. In the subsequent Sections, the most widely used methods for identifying which effects should be pooled are discussed.

6.1 Normal and Half Normal Probability Plots

Daniel (1959) proposed a simple and effective graphical tool for identifying the effects that can be pooled together to provide an estimate of the error variance in unreplicated experiments. The idea being that in most of the experiments especially two-level ones the effects are calculated by subtracting the averages of the responses associated with certain levels. When these are uncorrelated and have the same variance, they tend, according to the CL theory, to be approximately Normally distributed. Utilising this, Daniel (1959) proposed the Normal Probability (NP) plot as a tool for distinguishing active effects from those that are inert. To construct it, the calculated effects should be ranked in ascending order and then plotted against their observed cumulative Normal probabilities:

$$100*[i-0.5]/I \tag{17}$$

where i is the effect rank and I is the number of effects under study. The next step is to plot a straight line that passes through the middle group of points. Any effect whose corresponding points fall far off this line is declared significant. A primary limitation of this plot is its sensitivity to the effect signs. In fact, there can be many different Normal plots of the estimated effects for the same experimental results when different model arrays are used; the reason being the possibility of using model arrays with opposite sign columns. Furthermore, the arbitrariness in specifying the "high" and "low" factor level label especially in qualitative factors makes clear the need for a method that is insensitive to the effect signs.

To alleviate this problem Daniel (1959) proposed the use of a Half Normal Probability (HNP) plot which can be constructed by ranking the estimated effects in ascending order of absolute magnitude and plotting them on the upper half of a Normal probability scaled sheet. The inert effects are expected to be Normally distributed with a zero mean so they will tend to fall approximately along a straight line passing through the origin whereas the significant effects will have nonzero means and hence will constitute the outliers in the plot.

6.2 Lenth Method

Lenth Method (LM) (Lenth, 1989) is a formal means for estimating the effects' standard error when they are uncorrelated and of common variance. If C_1, C_2,C_3....C_m represent the calculated effects, Lenth (1989) suggested that a robust estimator of the effects' standard error which he termed pseudo standard error when there are few significant effects could be obtained as follows:

$$PSE = 1.5 * \text{median} \left(|C_j| : |C_j| < 2.5 S_0 \right) \tag{18}$$

where the median is computed over the $|C_j| < 2.5S_0$ and

$$S_0 = 1.5 * \text{median} \left(|C_j| \right) \tag{19}$$

A t-like statistic can then be derived by dividing the effects C_i by the PSE

$$t\text{-Lenth} = C_i / PSE \tag{20}$$

Lenth (1989) proposed that this statistic could be approximated by a t-distribution with m/3 degrees of freedom. Thus an individual effect C_i is declared significant at an α level if it exceeds the Margin of Error (ME):

$$ME = t_{\alpha/2,\,(m/3)} * PSE \tag{21}$$

As already discussed, using equation 20 to individually test the significance of each of the effects under study results in inflating α. To account for this, Lenth suggested using the Simultaneous Margin of Error (SME):

$$SME = t_{\gamma,\,(m/3)} * PSE \tag{22}$$

where $\gamma = 1\text{-}(1+0.95^{1/m})/2$. However, despite the use of equation 22, the problem of controlling α was not alleviated as described in the following Section.

6.3 Modified Lenth Method

Haaland and O'Connell, 1995 and Hamada and Balakrishnan, 1998 studied the performance of Lenth's method and found that it performed acceptably in terms of power and that it was one of the simplest methods to implement. However they also highlighted that it failed to control the significance level α at its nominal value. Loughin and Noble (1997) and Loughin (1998) ascribed this to the use of the t-distribution as they found that it was not a good approximation for the reference distribution of the t-Lenth statistic. Thus even multiple comparison methods cannot account for the failure to control α as they need a good approximation of the t-Lenth distribution. Ye and Hamada (2000) proposed a Modified version of Lenth's Method (MLM) where simulation was used to generate samples of the null distribution of t-Lenth. Consequently, extensive calibrated individual and simultaneous critical values for the Lenth method (for many values of m that arise with two-level experiments) so that α is maintained at its nominal value were presented. On the basis of these an effect is declared significant when its t-Lenth statistic exceeds any of them.

6.4 Pooling Up Method

For analysing unreplicated experiments, Taguchi suggested the pooling up (PU) method (Ross, 1996). This entails using the smallest effect as an estimate of the experimental error and utilising the ANOVA to test the significance of the next larger one at a specified significance level. If the test yields an insignificant result, the two effects are pooled to test the next larger effect and the procedure continues in the same manner until a significant value is observed.

6.5 Pooling Down Method

In the pooling down (PD) method (Ross, 1996) all the effects but the largest are pooled to form an initial estimate of the experimental error which is then used in the ANOVA to test the significance of the largest effect. If this is found to be insignificant the procedure terminates and no effect is pronounced significant. Otherwise, the next largest effect is excluded from the error and the remaining effects form a new pooled estimate of the error. This is then used to test the significance of the two largest effects. If either of these is found to be insignificant it will be pooled with the error terms and the selection procedure terminates. Otherwise the process continues in the same manner by excluding the next largest effect from those pooled.

6.6 Unassigned Columns Method

When columns are not assigned factors or used to estimate particular interactions, they are deemed inactive thereby freeing up degrees of freedom so that the error variance can be estimated. In this method, therefore, the effects associated with the Unassigned Columns (UC) are pooled together to provide an error estimate.

6.7 Variable Selection Methods

When regression is used to analyse unreplicated experiments, variable selection techniques (Miller, 2002) can be used to determine the factorial effects that should be included in the fitted model. Consequently, the remaining effects should be pooled to form an estimate of the experimental error variance. The most commonly employed variable selection are discussed in the subsequent Sections.

6.7.1 Stepwise Regression

In Stepwise Regression (SWR) several regression models are iteratively constructed by adding or removing factorial effects at each step (Bajpai, 2009), the idea being to compare the current model with a new one obtained by adding or deleting a factorial effect from it. Using equation 23, partial F statistics should be calculated to decide whether an effect should be added or deleted:

$$F_{partial} = [(SS_{E(Model\ I)} - SS_{E(Model\ II)})/(k-q)] / [SS_{E(Model\ II)}/(N-k-1)] \quad (23)$$

where k is the number of effects in Model II (the one with the larger number of effects), q is the number of effects in Model I and N is the number of experimental runs. For each model, the SS_E is obtained by pooling the SS associated with the effects not included in that model. The p-value of the calculated partial F should be compared as appropriate with one of the two preselected threshold α-values for adding (α_{in}) or deleting (α_{out}) variables. The effect is added (or deleted) if the p-value of its partial F is smaller (or larger) than the corresponding threshold α-value. The stepwise procedure starts with two forward selections where in the first the factorial effect with the smallest p-value (or largest effect) is selected provided that it is smaller than α_{in}. In the second, the remaining effects are examined one at a time as candidates for the second effect in the model. The one with the smallest p-value is added provided again that its p-value is less than α_{in}. In order to examine whether the addition of the second effect has increased the p-value of the first one a backward elimination step is performed where the first effect is dropped if its new p-value exceeds α_{out}. The procedure continues by alternating between one step of forward selection and one step of backward elimination. It terminates when no effect meets the criteria for being added or eliminated from the model. The non-added effects are then used to form an estimate of the experimental error.

6.7.2 Backward Elimination

Backward Elimination (BE) parallels that of the pooling up strategy. It starts with all the studied effects included in the model. Then the smallest effect associated with which the smallest partial F is observed is dropped from the model and used as an estimate of the error given that its p-value is larger than that of α_{out}. Then the next effect for potential elimination is examined and pooled if its p-value is larger than α_{out}. The procedure terminates when no further effects can be pooled. The effects that the final model precludes are pooled to estimate the error variance.

6.7.3 Forward Selection

Again Forward Selection (FS) is equivalent to the pooling down method given that the same α_{in} is used. The first effect to enter the model is the one that has the smallest p-value associated with its partial F statistic given that this is smaller than the predetermined α_{in} i.e. the largest effect. In the next step, all the non-selected effects are examined and the largest one is selected given that its p-value is smaller than the α_{in}. The procedure continues in the same manner until no more effects can be selected. The effects not included in the last model are the one to be used to estimate the experimental error.

6.7.4 Best Subset Selection

In Best Subset Selection (BSS) all the possible models that can be fitted to the experimental results are constructed and the subset of variables that attain the best value of a certain criterion is selected. The selection criterion should reward good model fitting and penalise model complexity. R^2 obtained by dividing the amalgamated SS due to all the effects selected in the constructed model by the SS_T is not a suitable criterion as it increases as the number of variables in the model increases. An alternative criterion that can circumvent this problem is

$$\text{Adj. } R^2 = 1 - MSE/(SS_T/n-1) \tag{24}$$

Its main advantage is that it accounts for the model degrees of freedom and increases only when the added variable(s) results in reducing the error variance.

7. Summary

Statistical analyses of experimental data comprise the factorial effects estimation and their significance testing. The latter, is a procedure by which sample results are used to verify the plausibility of a null hypotheses. Whenever a true null hypothesis is rejected a Type I error is committed whereas a Type II error is committed when failing to reject a false null hypothesis. The probability of correctly rejecting a false hypothesis is called the Power. A common null hypothesis in DOE is that of no effect. This can be tested in many ways depending on whether the experiment is replicated; whether regression analysis is employed; and the number of studied factors' levels. In the case of replicated experiments, the experimental error variance can be estimated and the effects' significance can be tested using a t-test or ANOVA either independently or as part of regression analysis. This is true in the case of two-level experiments. However, for three-level ones, a t-test can only be used after decomposing each of the studied effects into single degree of freedom components. One way to accomplish this is to use one of the four coding systems: DV, LSQ, ANOVA-M or L-Q system. In response to the confusing treatment of these in the DOE literature, the conditions under which each system should be applied along with the implications of its adoption have been detailed in this Appendix. In the case of unreplicated experiments, no degrees of freedom are available to estimate the error variance. Consequently, certain effects should be pooled together to form an estimate of the latter. These are specified using one of the pooling techniques presented in this Appendix i.e. the NP and HNP plots, Lenth, its modified version, the PU, PD and UC methods. Alternatively, if regression modelling is employed variable selection methods including stepwise regression, backward elimination, forward selection and best subset methods should be used.

REFERENCES

Walpole, R., Myers, R., Myers, S. and Ye, K. (2006), *Probability And Statistics For Engineers And Scientists*, Pearson Education.

Kanji, G. (2006), *100 Statistical Tests,* SAGE publications Ltd.

Devore, J. (2008), *Probability and Statistics for Engineering and the Sciences,* Cengage Learning.

Black, K. (2009), *Business Statistics: Contemporary Decision Making,* Wiley.

Downing, D. and Clark, J. (2003), *Business Statistics*, Barron's Educational Series.

Ross, S. (2004), *Introduction to Probability and Statistics for Engineers and Scientists,* Academic Pr.

Cramer, H. (1999), *Mathematical Methods Of Statistics,* Princeton Univ Pr.

Stuart, A. and Ord, J. (2009), *Kendall's Advanced Theory Of Statistics. Vol. 1: Distribution theory,* Wiley.

Box, G., Hunter, S. and Hunter, W. (2005), *Statistics For Experimenters: Design, Innovation, And Discovery,* Wiley.

Barrentine, L. (1999), *An Introduction To Design Of Experiments: A Simplified Approach,* American Society for Qualit.

Montgomery, D. (2010), *Design And Analysis Of Experiments,* Wiley

Box, G. E. P. and Cox, D. R. (1964), "An Analysis of Transformations", *Journal of the Royal Statistical Society. Series B (Methodological),* Vol. 26, No**.** 2, pp. 211-252.

Kutner, M., Nachtsheim, C., Neter, J. and Wasserman, W. (2004), *Applied Linear Statistical Models: Regression, Analysis Of Variance, And Experimental Designs,* McGraw-Hill/Irwin.

Rawlings, J., Pantula, S. and Dickey, D. (1998), *Applied Regression Analysis: A Research Tool,* springer New York.

Montgomery, D., Peck, E. and Vining, G. (2006), "Introduction To Linear Regression Analysis", Wiley-New York.

Ross, P. J. (1996), *Taguchi Techniques for Quality Engineering,* McGraw-Hill.

Wu, C., Hamada, M. (2000), *Experiments: Planning, Analysis, And Parameter Design Optimization,* Wiley New York.

Draper, N. and Smith, H. (1998), *Applied Regression Analysis,* Wiley New York.

Bohrer, R., Chow, W., Faith, R., Joshi, V. and Wu, C. (1981), "Multiple Three-Decision Rules for Factorial Simple Effects: Bonferroni Wins Again!", *Journal of the American Statistical Association*, Vol. 76, No**.** 373, pp. 119-124.

Winer, B. and Brown, D. (1991), *Statistical Principles In Experimental Design,* Mc.Graw-Hill

Bonett, D. and Woodward, J. (1993), "Analysis of simple main effects in fractional factorial experimental designs of resolution v", *Communications in Statistics-Theory and Methods*, Vol. 22, No**.** 6, pp. 1585-1593.

Allen, M. (2004), *Understanding regression analysis*, Kluwer Academic Pub.

Lee, H. and Comrey, A. (2009), *Elementary Statistics: A Problem Solving Approach 4th Edition*, Lulu. com.

Daniel, C. (1959), "Use Of Half-Normal Plots In Interpreting Factorial Two-Level Experiments", *Technometrics*, Vol. 1, No**.** 4, pp. 311-341.

Lenth, R. (1989), "Quick And Easy Analysis Of Unreplicated Factorials", *Technometrics*, Vol. 31, No**.** 4, pp. 469-473.

Haaland, P. and O'Connell, M. (1995), "Inference For Effect-Saturated Fractional Factorials", *Technometrics*, Vol. 37, No**.** 1, pp. 82-93.

Hamada, M. and Balakrishnan, N. (1998), "Analyzing Unreplicated Factorial Experiments: A Review With Some New Proposals", *Statistica Sinica*, Vol. 8, No**.** 1, pp. 1-28.

Loughin, T. and Noble, W. (1997), "A Permutation Test For Effects In An Unreplicated Factorial Design", *Technometrics*, Vol. 39, No**.** 2, pp. 180-190.

Loughin, T. (1998), "Calibration Of The Lenth Test For Unreplicated Factorial Designs", *Journal of Quality Technology*, Vol. 30, No**.** 2, pp. 171-175.

Ye, K. and Hamada, M. (2000), "Critical Values Of The Lenth Method For Unreplicated Factorial Design", *Journal of Quality Technology* Vol. 32, No**.** 1, pp. 57 - 66.

Miller, A. (2002), *Subset Selection In Regression,* CRC Press.

Bajpai, N. (2009), *Business Statistics,* Pearson Education

APPENDIX 3: SUMMARY OF RESULTS OF THE DOE PRACTICE REVIEW

1. Abbreviations

NR: Not Reported

NA: Not Applicable

ML: Mixed Level

CCD: Central Composite Design

CP: Central Points

NPP: Normal Probability Plot

ROT: Role of Thumb

UC: Unassigned Columns

SWR: Stepwise Regression

BE: Backward Elimination

2-Design Related and Conducting Aspects

No	Author(s)	No. Factors	No. Levels	Design	Study Avg.	Study Variation	Replication	Aliasing Examined	Interactions Examined	Randomisation
1	Karayel	3	4	L64	Yes	No	Yes	NA	No	No
2	Gunasegaram et al	5	2	L16	Yes	No	CP	Yes	Yes	Yes
3	Shanmugam & Masood	4	4	L64	Yes	No	No	No	No	No
4	Cakir et al	3	3	L27	Yes	No	No	NA	No	No
5	Fang et al	4	3	L9	Yes	No	No	No	No	No
6	Yang et al	3	2	L8	Yes	No	No	NA	Yes	No
7	Venkatachalam et al	5	3	L243	Yes	No	No	NA	No	No
8	Lin et al	6	ML	L18	Yes	Yes	Yes	No	No	No
9	Jeang & Lin	5	3	L243	Yes	No	No	NA	No	No
10	Lauderbaugh	10 6	2 2	L64 L12	Yes Yes	No No	No No	No No	Yes Yes	No No
11	Xavior & Adithan	4	3	L27	Yes	No	No	No	No	No
12	Boronat et al	4	2	L16	Yes	No	No	NA	Yes	Yes
13	Wang et al	5	ML	L18	No	Yes	Yes	No	No	No

No	Author(s)	No. Factors	No. Levels	Design	Study Avg.	Study Variation	Replication	Aliasing Examined	Interactions Examined	Randomisation
14	Majewski et al	2	ML	NR	Yes	No	No	No	Yes	No
15	Bi & Jiang	2	3	L9	No	Yes	Yes	No	No	No
16	Senthilkumar et al	3	3	L18	No	Yes	Yes	No	Yes	No
17	Mi & Lackey	3	5	CCD	Yes	No	CP	No	Yes	No
18	Mata et al	2	ML	L12	Yes	No	No	NA	Yes	Yes
19	Chattopadhyay et al	3	3	L27	Yes	Yes	Yes	NA	No	No
20	Huang et al	7	3	L18	Yes	Yes	Yes	No	No	No
21	Palanikumar & Davim	4	2	L16	Yes	No	Yes	NA	Yes	Yes
22	Azmir & Ahsan	6	ML	L18	Yes	No	No	No	No	No
23	Lin & Ho	4	3	L9	Yes	No	Yes	No	No	No
24	Ajaal & Smith	4 7	3 2	L9 L8	Yes Yes	Yes Yes	Yes Yes	No No	No No	No No
25	Dey et al	3	3	CCD	Yes	No	CP	No	Yes	No
26	Sahin	3	3	L9	Yes	Yes	Yes	No	No	No
27	Dabade & Joshi	5	3	L27	Yes	No	No	No	No	No

No	Author(s)	No. Factors	No. Levels	Design	Study Avg.	Study Variation	Replication	Aliasing Examined	Interactions Examined	Randomisation
28	Oudjene et al	8	ML	L18	Yes	No	No	No	No	No
29	Tsai et al	8 4	3 3	L18 L81	No Yes	Yes No	Yes No	No NA	No Yes	No No
30	Chang & Choi	6	ML	L18	Yes	Yes	Yes	No	No	No
31	Rakwal & Bamberg	3	ML	L36	Yes	No	No	No	Yes	No
32	Xueping et al	3	3	L9	Yes	No	No	No	No	No
33	Sharma & Rout	4	3	L9	Yes	No	No	No	No	No
34	Lu et al	5	ML	L18	Yes	Yes	Yes	No	No	No
35	Jeang et al	8	3	L18	Yes	Yes	Yes	No	No	No
36	Luo & Chen	4	3	L9	Yes	Yes	Yes	No	No	No
37	Sorrentino & Carrino	1	6	L6	Yes	No	Yes	NA	No	Yes
38	Muthukrishnan & Davim	3	3	L27	Yes	Yes	Yes	NA	No	No
39	Rosa et al	6	2	L16	Yes	Yes	Yes	Yes	Yes	Yes
40	Tzeng et al	4	3	L9	Yes	No	No	No	No	No
41	Zhang et al	4	3	L9	Yes	No	No	No	No	No

No	Author(s)	No. Factors	No. Levels	Design	Study Avg.	Study Variation	Replication	Aliasing Examined	Interactions Examined	Randomisation
42	Tsai et al	8	ML	L18	No	Yes	Yes	No	No	No
43	Phatak et al	4	2	L12	Yes	No	No	No	Yes	No
44	Prihandana et al	4	ML	L18	Yes	Yes	Yes	No	No	No
45	Sadasivam et al	4	3	L9	Yes	No	No	No	No	No
46	Saha & Choudhury	6	5	CCD	Yes	No	CP	NA	Yes	Yes
47	Courbon et al	5	3	L27	Yes	No	No	No	Yes	No
48	Shyha et al	6	2	L12	Yes	No	No	No	No	No
49	Marafona & Araujo	8	3	L18	Yes	No	No	No	No	No
50	Zhang & Guo	4	4	L16	Yes	No	No	No	No	No
51	Arai et al	6	2	L16	Yes	No	No	No	Yes	No
52	Nandy et al	4	2	L8	Yes	No	CP	No	No	Yes
53	Kwak	4	3	L9	Yes	No	No	No	No	No
54	Totis	3	ML	NR	Yes	No	Yes	No	Yes	No
55	Galantucci et al	3	2	L8	Yes	No	Yes	NA	No	No

No	Author(s)	No. Factors	No. Levels	Design	Study Avg.	Study Variation	Replication	Aliasing Examined	Interactions Examined	Randomisation
56	Martins et al	3	2	L8	Yes	No	No	No	No	No
57	Ali et al	6 4	2 2	L8 L16	Yes Yes	No No	No No	No NA	No Yes	No No
58	Mori et al	7	ML	L18	Yes	Yes	Yes	No	No	No

3- Data Analysis Aspects

No	Author(s)	Pooling Method	Standard ANOVA	Regression ANOVA	Hypothesis Stated	Reporting P-value	Examining Assumptions	FE's Importance Measures					
								Mean-Related		Variance-Related			
								Graphical	Absolute	PC_1	PC_2	R^2	R^2-Adj
1	Karayel	No	No	No	NA	NA	NA	Yes	No	No	No	No	No
2	Gunasegaram et al	NPP	No	No	No	No	No	Yes	No	No	No	No	No
3	Shanmugam & Masood	No	No	No	NA	NA	NA	Yes	No	No	No	No	No
4	Cakir et al	No	No	No	NA	NA	NA	No	Yes	No	No	Yes	Yes
5	Fang et al	No	No	No	NA	NA	NA	Yes	No	No	No	No	No
6	Yang et al	ROT	Yes	No	No	Yes	Yes	Yes	No	Yes	No	No	No
7	Venkatachalam et al	No	No	No	NA	NA	NA	Yes	No	No	No	No	No
8	Lin et al	UC	Yes	No	No	No	No	Yes	No	No	No	No	No
9	Jeang & Lin	UC	Yes	No	No	Yes	No	Yes	Yes	No	No	No	No
10	Lauderbaugh	NPP NPP	No No	Yes Yes	No No	Yes Yes	Yes No	No No	Yes Yes	No No	No No	No No	No No
11	Xavior & Adithan	UC	Yes	No	No	No	No	Yes	Yes	Yes	No	No	No
12	Boronat et al	Pareto	No	No	No	No	No	No	No	No	No	No	No
13	Wang et al	UC	Yes	No	No	No	No	Yes	No	No	No	Yes	Yes

No	Author(s)	Pooling Method	Standard ANOVA	Regression ANOVA	Hypothesis Stated	Reporting P-value	Examining Assumptions	FE's Importance Measures					
								Mean-Related		Variance-Related			
								Graphical	Absolute	PC_1	PC_2	R^2	R^2-Adj
14	Majewski et al	No	Yes	No	No	No	No	No	No	No	No	No	No
15	Bi & Jiang	No	No	No	NA	NA	NA	No	Yes	No	No	No	No
16	Senthilkumar et al	ROT	Yes	No	No	No	No	No	No	No	Yes	No	No
17	Mi & Lackey	Pareto	No	No	No	No	No	Yes	No	No	No	No	No
18	Mata et al	ROT	No	Yes	No	No	Yes	Yes	Yes	No	No	Yes	No
19	Chattopadhyay et al	UC	Yes	No	No	No	No	Yes	Yes	No	No	Yes	No
20	Huang et al	UC	Yes	No	No	Yes	Yes	Yes	No	No	Yes	No	No
21	Palanikumar & Davim	ROT	Yes	No	No	Yes	No	Yes	No	No	No	No	No
22	Azmir & Ahsan	UC	Yes	No	No	No	No	Yes	No	No	Yes	No	No
23	Lin & Ho	No	Yes	No	No	No	No	No	No	No	No	No	No
24	Ajaal & Smith	No No	No No	No No	No No	NA NA	NA NA	Yes Yes	Yes Yes	No No	No No	No No	No No
25	Dey et al	No	No	No	NA	NA	NA	No	Yes	No	No	No	No
26	Sahin	UC	Yes	No	No	No	No	Yes	Yes	Yes	No	No	No

No	Author(s)	Pooling Method	Standard ANOVA	Regression ANOVA	Hypothesis Stated	Reporting P-value	Examining Assumptions	FE's Importance Measures					
								Mean-Related		Variance-Related			
								Graphical	Absolute	PC_1	PC_2	R^2	R^2-Adj
27	Dabade & Joshi	ROT	Yes	No	No	Yes	No	Yes	No	No	No	No	No
28	Oudjene et al	No	No	No	NA	NA	NA	No	No	No	No	No	No
29	Tsai et al	UC	Yes	No	No	Yes	No	Yes	No	No	No	No	No
		No	No	No	No	Yes	Yes	No	Yes	No	No	Yes	No
30	Chang & Choi	No	No	No	NA	NA	NA	Yes	Yes	No	No	No	No
31	Rakwal & Bamberg	ROT	No	Yes	No	Yes	No	No	Yes	No	No	Yes	No
32	Xueping et al	No	No	No	NA	NA	NA	Yes	Yes	No	No	No	No
33	Sharma & Rout	ROT	Yes	No	No	No	No	Yes	Yes	No	Yes	No	No
34	Lu et al	UC	Yes	No	No	No	No	No	Yes	Yes	No	No	No
35	Jeang et al	ROT	Yes	No	No	No	No	Yes	Yes	No	Yes	No	No
36	Luo & Chen	No	Yes	No	No	Yes	No	Yes	Yes	No	Yes	No	No
37	Sorrentino & Carrino	No	Yes	No	No	Yes	Yes	Yes	No	No	No	No	No
38	Muthukrishnan & Davim	UC	Yes	No	No	No	No	Yes	Yes	Yes	No	No	No
39	Rosa et al	ROT	Yes	No	No	Yes	No	Yes	No	No	No	No	No

No	Author(s)	Pooling Method	Standard ANOVA	Regression ANOVA	Hypothesis Stated	Reporting P-value	Examining Assumptions	FE's Importance Measures					
								Mean-Related		Variance-Related			
								Graphical	Absolute	PC_1	PC_2	R^2	R^2-Adj
40	Tzeng et al	No	Yes	No	No	No	No	No	Yes	Yes	No	No	No
41	Xiaoyun et al	No	No	No	NA	NA	NA	Yes	No	No	No	No	No
42	Tsai et al	UC	Yes	No	NA	NA	NA	No	Yes	Yes	No	No	No
43	Phatak et al	ROT	No	Yes	No	Yes	Yes	No	No	No	No	Yes	Yes
44	Prihandana et al	No	No	No	NA	NA	NA	Yes	No	Yes	No	No	No
45	Sadasivam et al	No	No	No	NA	NA	NA	No	Yes	Yes	No	No	No
46	Saha & Choudhury	BE	No	Yes	No	Yes	No	Yes	Yes	No	No	Yes	Yes
47	Courbon et al	SWR	No	Yes	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes
48	Shyha et al	UC	Yes	No	No	Yes	No	Yes	No	No	Yes	No	No
49	Marafona & Araujo	ROT	Yes	Yes	No	Yes	No	No	Yes	Yes	No	Yes	Yes
50	Zhang & Guo	No	No	No	NA	NA	NA	No	No	No	No	No	No
51	Arai et al	ROT	Yes	No	NA	NA	NA	No	No	No	Yes	No	No
52	Nandy et al	ROT	No	Yes	No	No	No	Yes	Yes	No	No	No	No
53	Kwak	ROT	Yes	No	No	No	No	Yes	No	No	No	No	No

No	Author(s)	Pooling Method	Standard ANOVA	Regression ANOVA	Hypothesis Stated	Reporting P-value	Examining Assumptions	FE's Importance Measures					
								Mean-Related		Variance-Related			
								Graphical	Absolute	PC_1	PC_2	R^2	R^2-Adj
54	Totis	SWR	No	No	No	No	No	No	Yes	No	No	Yes	No
55	Galantucci et al	No	No	No	NA	NA	NA	Yes	No	No	No	No	No
56	Martins et al	No	No	No	NA	NA	NA	Yes	No	No	No	No	No
57	Ali et al	No	No	No	NA	NA	NA	No	Yes	No	No	No	No
		No	No	No	NA	NA	NA	Yes	Yes	No	No	No	No
58	Mori et al	No	No	No	NA	NA	NA	Yes	No	No	No	No	No

4-Results Interpretation and Other Aspects

No.	Author(s)	Confirmation runs	DOE References	Software
1	Karayel	No	No	No
2	Gunasegaram et al	No	Yes	Minitab
3	Shanmugam & Masood	No	No	No
4	Cakir et al	No	No	No
5	Fang et al	No	No	No
6	Yang et al	Yes	Yes	Minitab
7	Venkatachalam et al	No	No	No
8	Lin et al	Yes	No	No
9	Jeang & Lin	No	Yes	SAS
10	Lauderbaugh	No No	No	Minitab
11	Xavior & Adithan	No	No	Minitab
12	Boronat et al	No	Yes	Design-Expert
13	Wang et al	Yes	Yes	No
14	Majewski et al	No	No	No
15	Bi & Jiang	No	No	No
16	Senthilkumar et al	Yes	Yes	No
17	Mi & Lackey	No	No	No
18	Mata et al	No	Yes	No
19	Chattopadhyay et al	No	Yes	SPSS
20	Huang et al	Yes	Yes	No

No.	Author(s)	Confirmation runs	DOE References	Software
21	Palanikumar & Davim	Yes	Yes	Minitab
22	Azmir & Ahsan	No	Yes	No
23	Lin & Ho	No	Yes	No
24	Ajaal & Smith	Yes Yes	Yes	No
25	Dey et al	No	Yes	Minitab
26	Sahin	Yes	Yes	Minitab
27	Dabade & Joshi	No	Yes	No
28	Oudjene et al	No	No	No
29	Tsai et al	Yes Yes	Yes	SPSS
30	Chang & Choi	No	Yes	No
31	Rakwal & Bamberg	Yes	Yes	No
32	Xueping et al	No	No	No
33	Sharma & Rout	Yes	Yes	No
34	Lu et al	Yes	Yes	No
35	Jeang et al	Yes	Yes	No
36	Luo & Chen	No	Yes	No
37	Sorrentino & Carrino	No	No	Minitab
38	Muthukrishnan & Davim	Yes	No	No
39	Rosa et al	Yes	Yes	STATISTICA
40	Tzeng et al	Yes	No	No

No.	Author(s)	Confirmation runs	DOE References	Software
41	Xiaoyun et al	No	No	Minitab
42	Tsai et al	No	Yes	No
43	Phatak et al	No	Yes	No
44	Prihandana et al	No	No	No
45	Sadasivam et al	No	Yes	SAS
46	Saha & Choudhury	residual	Yes	Design-Expert
47	Courbon et al	No	Yes	Design-Expert
48	Shyha et al	No	Yes	Minitab
49	Marafona & Arau jo	Yes	No	No
50	Zhang & Guo	No	No	No
51	Arai et al	No	No	No
52	Nandy et al	No	No	No
53	Kwak	No	No	No
54	Totis	No	Yes	No
55	Galantucci et al	No	No	Minitab
56	Martins et al	No	No	No
57	Ali et al	No No	No	No
58	Mori et al	No	Yes	No

REFERENCES

Karayel, D. (2009), "Prediction And Control Of Surface Roughness In CNC Lathe Using Artificial Neural Network", *Journal of Materials Processing Technology*, Vol. 209, No. 7, pp. 3125-3137.

Gunasegaram, D., Farnsworth, D. and Nguyen, T. (2009), "Identification Of Critical Factors Affecting Shrinkage Porosity In Permanent Mold Casting Using Numerical Simulations Based On Design Of Experiments", *Journal of Materials Processing Technology*, Vol. 209, No. 3, pp. 1209-1219.

Shanmugam, D. and Masood, S. (2009), "An Investigation On Kerf Characteristics In Abrasive Waterjet Cutting Of Layered Composites", *Journal of Materials Processing Technology*, Vol. 209, No. 8, pp. 3887-3893.

Cakir, M., Ensarioglu, C. and Demirayak, I. (2009), "Mathematical Modeling Of Surface Roughness For Evaluating The Effects Of Cutting Parameters And Coating Material", *Journal of Materials Processing Technology*, Vol. 209, No. 1, pp. 102-109.

Fang, W., Cho, T., Yoon, J., Song, K., Hur, S., Youn, S. and Chun, H. (2009), "Processing Optimization, Surface Properties And Wear Behavior Of HVOF Spraying WC-Crc-Ni Coating", *Journal of Materials Processing Technology*, Vol. 209, No. 7, pp. 3561-3567.

Yang, Y., Chuang, M. and Lin, S. (2009), "Optimization Of Dry Machining Parameters For High-Purity Graphite In End Milling Process Via Design Of Experiments Methods", *Journal of Materials Processing Technology*, Vol. 209, No. 9, pp. 4395-4400.

Venkatachalam, S., Li, X. and Liang, S. (2009), "Predictive Modeling Of Transition Undeformed Chip Thickness In Ductile-Regime Micro-Machining Of Single Crystal Brittle Materials", *Journal of Materials Processing Technology*, Vol. 209, No. 7, pp. 3306-3319.

Lin, Y., Chen, Y., Wang, D. and Lee, H. (2009), "Optimization Of Machining Parameters In Magnetic Force Assisted EDM Based On Taguchi Method", *Journal of Materials Processing Technology*, Vol. 209, No. 7, pp. 3374-3383.

Jeang, A. and Li, H. (2009), "An Empirical Study On The Robust Design Of A Semiconductor-Bumping Process", *Journal of Materials Processing Technology*, Vol. 209, No. 6, pp. 2986-2993.

Lauderbaugh, L. (2009), "Analysis Of The Effects Of Process Parameters On Exit Burrs In Drilling Using A Combined Simulation And Experimental Approach", *Journal of Materials Processing Technology*, Vol. 209, No. 4, pp. 1909-1919.

Xavior, M. and Adithan, M. (2009), "Determining The Influence Of Cutting Fluids On Tool Wear And Surface Roughness During Turning Of AISI 304 Austenitic Stainless Steel", *Journal of Materials Processing Technology*, Vol. 209, No. 2, pp. 900-909.

Boronat, T., Segui, V., Peydro, M. and Reig, M. (2009), "Influence Of Temperature And Shear Rate On The Rheology And Processability Of Reprocessed ABS In Injection Molding Process", *Journal of Materials Processing Technology*, Vol. 209, No. 5, pp. 2735-2745.

Wang, C., Chow, H., Yang, L. and Lu, C. (2009), "Recast Layer Removal After Electrical Discharge Machining Via Taguchi Analysis: A Feasibility Study", *Journal of Materials Processing Technology*, Vol. 209, No. 8, pp. 4134-4140.

Majewski, C., Zarringhalam, H., Toon, D., Ajoku, U., Hopkinson, N. and Caine, M. (2009), "The Use Of Off-Line Part Production To Predict The Tensile Properties Of Parts Produced By Selective Laser Sintering", *Journal of Materials Processing Technology*, Vol. 209, No. 6, pp. 2855-2863.

Bi, C. and Jiang, B. (2009), "Study Of Residence Time Distribution In A Reciprocating Single-Screw Pin-Barrel Extruder", *Journal of Materials Processing Technology*, Vol. 209, No. 8, pp. 4147-4153.

Senthilkumar, B., Ponnambalam, S. and Jawahar, N. (2009), "Process Factor Optimization For Controlling Pull-Down Defects In Iron Castings", *Journal of Materials Processing Technology*, Vol. 209, No. 1, pp. 554-560.

Mi, J. and Lackey, W. (2009), "SiC Line Deposition Using Laser CVD", *Journal of Materials Processing Technology*, Vol. 209, No. 8, pp. 3818-3829.

Mata, F., Gaitonde, V., Karnik, S. and Davim, J. (2009), "Influence Of Cutting Conditions On Machinability Aspects Of PEEK, PEEK CF 30 And PEEK GF 30 Composites Using PCD Tools", *Journal of Materials Processing Technology*, Vol. 209, No. 4, pp. 1980-1987.

Chattopadhyay, K., Verma, S., Satsangi, P. and Sharma, P. (2009), "Development Of Empirical Model For Different Process Parameters During Rotary Electrical Discharge Machining Of Copper-Steel (EN-8) System", *Journal of Materials Processing Technology*, Vol. 209, No. 3, pp. 1454-1465.

Huang, M., Li, C., Yu, J., Huang, Y. and Hsieh, L. (2009), "Robust Parameter Design Of Micro-Injection Molded Gears Using A LIGA-Like Fabricated Mold Insert", *Journal of Materials Processing Technology*, Vol. 209, No. 15-16, pp. 5690-5701.

Palanikumar, K. and Davim, J. (2009), "Assessment Of Some Factors Influencing Tool Wear On The Machining Of Glass Fibre-Reinforced Plastics By Coated Cemented Carbide Tools", *Journal of Materials Processing Technology*, Vol. 209, No. 1, pp. 511-519.

Azmir, M. and Ahsan, A. (2009), "A Study Of Abrasive Water Jet Machining Process On Glass/Epoxy Composite Laminate", *Journal of Materials Processing Technology*, Vol. 209, No. 20, pp. 6168-6173.

Lin, Z. and Ho, C. (2009), "Performance Of Coated Tungsten Carbide Tools On Milling Printed Circuit Board", *Journal of Materials Processing Technology*, Vol. 209, No. 1, pp. 303-309.

Ajaal, T. and Smith, R. (2009), "Employing The Taguchi Method In Optimizing The Scaffold Production Process For Artificial Bone Grafts", *Journal of Materials Processing Technology*, Vol. 209, No. 3, pp. 1521-1532.

Dey, V., Pratihar, D., Datta, G., Jha, M., Saha, T. and Bapat, A. (2009), "Optimization Of Bead Geometry In Electron Beam Welding Using A Genetic Algorithm", *Journal of Materials Processing Technology*, Vol. 209, No. 3, pp. 1151-1157.

Sahin, Y. (2009), "Comparison Of Tool Life Between Ceramic And Cubic Boron Nitride (CBN) Cutting Tools When Machining Hardened Steels", *Journal of Materials Processing Technology*, Vol. 209, No. 7, pp. 3478-3489.

Dabade, U. and Joshi, S. (2009), "Analysis Of Chip Formation Mechanism In Machining Of Al/Sicp Metal Matrix Composites", *Journal of Materials Processing Technology*, Vol. 209, No. 10, pp. 4704-4710.

Oudjene, M., Ben-Ayed, L., Delamézière, A. and Batoz, J. (2009), "Shape Optimization Of Clinching Tools Using The Response Surface Methodology With Moving Least-Square Approximation", *Journal of Materials Processing Technology*, Vol. 209, No. 1, pp. 289-296.

Tsai, K., Hsieh, C. and Lo, W. (2009), "A Study Of The Effects Of Process Parameters For Injection Molding On Surface Quality Of Optical Lenses", *Journal of Materials Processing Technology*, Vol. 209, No. 7, pp. 3469-3477.

Chang, D. and Choi, J. (2009), "Geometric Parameter Design Of A Cantilever Probing Needle Used In Epoxy Ring Probe Card", *Journal of Materials Processing Technology*, Vol. 209, No. 1, pp. 38-50.

Rakwal, D. and Bamberg, E. (2009), "Slicing, Cleaning And Kerf Analysis Of Germanium Wafers Machined By Wire Electrical Discharge Machining", *Journal of Materials Processing Technology*, Vol. 209, No. 8, pp. 3740-3751.

Xueping, Z., Erwei, G. and Richard Liu, C. (2009), "Optimization Of Process Parameter Of Residual Stresses For Hard Turned Surfaces", *Journal of Materials Processing Technology*, Vol. 209, No. 9, pp. 4286-4291.

Sharma, A. and Rout, D. (2009), "Finite Element Analysis Of Sheet Hydromechanical Forming Of Circular Cup", *Journal of Materials Processing Technology*, Vol. 209, No. 3, pp. 1445-1453.

Lu, H., Chang, C., Hwang, N. and Chung, C. (2009), "Grey Relational Analysis Coupled With Principal Component Analysis For Optimization Design Of The Cutting Parameters In High-Speed End Milling", *Journal of Materials Processing Technology*, Vol. 209, No. 8, pp. 3808-3817.

Jeang, A., Chung, C., Chen, C. and Li, H. (2009), "Optimizing Process Parameters Of Hot-Bar Soldering Process Through Quality Function Deployment And Taguchi Method", *Journal of Materials Processing Technology*, Vol. 209, No. 6, pp. 2967-2977.

Luo, S. and Chen, K. (2009), "An Experimental Study Of Flat Fixed Abrasive Grinding Of Silicon Wafers Using Resin-Bonded Diamond Pellets", *Journal of Materials Processing Technology*, Vol. 209, No. 2, pp. 686-694.

Sorrentino, L. and Carrino, L. (2009), "2024 Aluminium Alloy Wettability And Superficial Cleaning Improvement By Air Cold Plasma Treatment", *Journal of Materials Processing Technology*, Vol. 209, No. 3, pp. 1400-1409.

Muthukrishnan, N. and Davim, J. (2009), "Optimization Of Machining Parameters Of Al/Sic-MMC With ANOVA And ANN Analysis", *Journal of Materials Processing Technology*, Vol. 209, No. 1, pp. 225-232.

Rosa, J., Robin, A., Silva, M., Baldan, C. and Peres, M. (2009), "Electrodeposition Of Copper On Titanium Wires: Taguchi Experimental Design Approach", *Journal of Materials Processing Technology*, Vol. 209, No. 3, pp. 1181-1188.

Tzeng, C., Lin, Y., Yang, Y. and Jeng, M. (2009), "Optimization Of Turning Operations With Multiple Performance Characteristics Using The Taguchi Method And Grey Relational Analysis", *Journal of Materials Processing Technology*, Vol. 209, No. 6, pp. 2753-2759.

Zhang, X., Chen, G., Zhang, Y. and Lai, X. (2009), "Improvement Of Resistance Spot Weldability For Dual-Phase (DP600) Steels Using Servo Gun", *Journal of Materials Processing Technology*, Vol. 209, No. 5, pp. 2671-2675.

Tsai, M., Huang, J. and Kao, W. (2009), "Robotic Polishing Of Precision Molds With Uniform Material Removal Control", *International Journal of Machine Tools and Manufacture*, Vol. 49, No. 11, pp. 885-895.

Phatak, U., Bukkapatnam, S., Kong, Z. and Komanduri, R. (2009), "Sensor-Based Modeling Of Slurry Chemistry Effects On The Material Removal Rate (MRR) In Copper-CMP Process", *International Journal of Machine Tools and Manufacture*, Vol. 49, No. 2, pp. 171-181.

Prihandana, G., Mahardika, M., Hamdi, M., Wong, Y. and Mitsui, K. (2009), "Effect Of Micro-Powder Suspension And Ultrasonic Vibration Of Dielectric Fluid In Micro-EDM Processes--Taguchi Approach", *International Journal of Machine Tools and Manufacture*, Vol. 49, No. 12-13, pp. 1035-1041.

Sadasivam, B., Hizal, A., Park, S. and Arola, D. (2009), "An Evaluation Of Abrasive Waterjet Peening With Elastic Prestress", *Journal of Manufacturing Science and Engineering*, Vol. 131, No. 1, pp. 011010.

Saha, S. and Choudhury, S. (2009), "Experimental Investigation And Empirical Modeling Of The Dry Electric Discharge Machining Process", *International Journal of Machine Tools and Manufacture*, Vol. 49, No. 3-4, pp. 297-308.

Courbon, C., Kramar, D., Krajnik, P., Pusavec, F., Rech, J. and Kopac, J. (2009), "Investigation Of Machining Performance In High-Pressure Jet Assisted Turning Of Inconel 718: An Experimental Study", *International Journal of Machine Tools and Manufacture*, Vol. 49, No. 14, pp. 1114-1125.

Shyha, I., Aspinwall, D., Soo, S. and Bradley, S. (2009), "Drill Geometry And Operating Effects When Cutting Small Diameter Holes In CFRP", *International Journal of Machine Tools and Manufacture*, Vol. 49, No. 12-13, pp. 1008-1014.

Marafona, J. and Araújo, A. (2009), "Influence Of Workpiece Hardness On EDM Performance", *International Journal of Machine Tools and Manufacture*, Vol. 49, No. 9, pp. 744-748.

Zhang, S. and Guo, Y. (2009), "An Experimental And Analytical Analysis On Chip Morphology, Phase Transformation, Oxidation, And Their Relationships In Finish Hard Milling", *International Journal of Machine Tools and Manufacture*, Vol. 49, No**.** 11, pp. 805-813.

Arai, S., Wilson, S., Corbett, J. and Whatmore, R. (2009), "Ultra-Precision Grinding Of PZT Ceramics--Surface Integrity Control And Tooling Design", *International Journal of Machine Tools and Manufacture*, Vol. 49, No**.** 12-13, pp. 998-1007.

Nandy, A., Gowrishankar, M. and Paul, S. (2009), "Some Studies On High-Pressure Cooling In Turning Of Ti-6Al-4V", *International Journal of Machine Tools and Manufacture*, Vol. 49, No**.** 2, pp. 182-198.

Kwak, J. (2009), "Enhanced Magnetic Abrasive Polishing Of Non-Ferrous Metals Utilizing A Permanent Magnet", *International Journal of Machine Tools and Manufacture*, Vol. 49, No**.** 7-8, pp. 613-618.

Totis, G. (2009), "RCPM--A New Method For Robust Chatter Prediction In Milling", *International Journal of Machine Tools and Manufacture*, Vol. 49, No**.** 3-4, pp. 273-284.

Galantucci, L., Lavecchia, F. and Percoco, G. (2009), "Experimental Study Aiming To Enhance The Surface Finish Of Fused Deposition Modeled Parts", *CIRP Annals-Manufacturing Technology*, Vol. 58, No**.** 1, pp. 189-192.

Martins, P., Kwiatkowski, L., Franzen, V., Tekkaya, A. and Kleiner, M. (2009), "Single Point Incremental Forming Of Polymers", *CIRP Annals-Manufacturing Technology*, Vol. 58, No**.** 1, pp. 229-232.

Ali, S., Hinduja, S., Atkinson, J. and Pandya, M. (2009), "Shaped Tube Electrochemical Drilling Of Good Quality Holes", *CIRP Annals-Manufacturing Technology*, Vol. 58, No**.** 1, pp. 185-188.

Mori, M., Mizuguchi, H., Fujishima, M., Ido, Y., Mingkai, N. and Konishi, K. (2009), "Design Optimization And Development Of CNC Lathe Headstock To Minimize Thermal Deformation", *CIRP Annals-Manufacturing Technology*, Vol. 58, No**.** 1, pp. 331-334.

APPENDIX 4:
RESULTS OF THE EFFECT SIZES' REVIEW

No	Author(s)	Process	No. Factors	No of Levels	OA	No. of Sig. Effects		Standardised Effects	
						Main	2fi*	Main	2fi
1	Ganjigatti et al 2007	Welding	6	2	L64	5	3	-2.94 -2.27 0.64 0.38 -1.17	1.00 -0.42 0.48
			6	2	L64	3	1	-1.67 1.95 0.50	-0.98
			6	2	L64	4	3	0.87 -1.11 -0.66 1.11	-0.54 0.48 0.38
2	Gunaraj & Murugan 1999	Welding	4	2	L16	3		1.70 2.37 -3.10	
			4	2	L16	4	2	2.72 3.40 -1.26 -1.12	-1.16 1.48
			4	2	L16	3	1	1.28 1.53 -0.69	-0.63
3	Darwish 2000	Turning	4	2	L16	2		4.13 1.08	
			4	2	L16	3		5.40 0.87 -1.23	
4	Aggarwal et al 2008	Turning	4	2	L16	4	2	5.68 1.92 4.50 0.89	0.82 0.75
			4	2	L16	4	2	5.95 1.08 3.79 -0.72	0.72 2.35
			4	2	L16	4	1	3.41 3.15 7.36 2.48	1.91

*** Two-Factor Interactions**

No	Author(s)	Process	No. Factors	No of Levels	OA	No. of Sig. Effects		Standardised Effects	
						Main	2fi	Main	2fi
5	Mahadevan et al 2008	Hardening	3	2	L8	3	3	0.88 5.95 1.55	7.20 2.95 1.76
6	Francis & El-Midany 2008	Carbothemic Reduction	3	2	L8	1		4.75	
			3	2	L8	1		3.20	
7	Darwin et al 2008	Deep Cryogenic Treatment	4	2	L16	4	2	0.68 -3.23 -1.44 -0.48	0.39 -0.64
8	Senthilvelan et al 2003	Powder Metallurgy	3	2	L8	3	2	8.62 6.44 3.28	2.99 -2.39
			3	2	L8	3	3	-2.35 1.25 2.30	-1.69 -1.69 1.47
9	Puertas & Luis 2003	Electric Discharge Machining	3	2	L8	1	1	-3.47	1.96
			3	2	L8	1	1	-3.51	1.98
10	Onwubolu & Kumar 2006	Drilling	3	2	L8	2	0	1.21 1.74	
			3	2	L8	0	0		
11	Chatterjee et al 2007	Powder Metallurgy	3	2	L8	2	1	-1.59 3.96	2.34
12	Mahadevan et al 2006	Heat Treatment	3	2	L8	3	3	9.61 1.31 9.46	-9.28 -4.41 -2.02

No	Author(s)	Process	No. Factors	No of Levels	OA	No. of Sig. Effects		Standardised Effects	
						Main	2fi	Main	2fi
13	Raghukandan & Senthilvelan 2004	Extrusion	3	3	L8	3	0	2.62 4.13 0.67	
14	Yang et al 2009	Milling	3	2	L8	1	0	3.72	
			3	2	L8	1	0	1.32	
15	Noordin et al 2004	Turning	3	2	L8	1	0	1.97	
			3	2	L8	1	0	1.22	
16	Kannan & Murugan 2006	Arc Welding	4	2	L16	2	0	2.47 -3.01	
			4	2	L16	4	1	1.43 1.46 0.94 2.51	1.21
			4	2	L16	1	0	-3.82	
			4	2	L16	4	2	0.72 2.13 -1.00 1.63	-0.62 1.22
17	Dutta & Pratihar 2007	Welding	5	2	L32	4	0	1.33 0.62 -0.49 -1.88	
			5	2	L32	2	1	-2.32 2.47	-0.93
			5	2	L32	3	0	-1.37 0.48 0.90	
			5	2	L32	3	0	-3.49 0.98 4.03	
18	Krajnik &Kopac 2004	Grinding	3	2	L8	1	0	4.73	
19	Mohammadi et al 2007	Grit Blasting	3	2	L8	2	0	1.06 1.29	

No	Author(s)	Process	No. Factors	No of Levels	OA	No. of Sig. Effects		Standardised Effects	
						Main	2fi	Main	2fi
20	Hung et al 2002	Ion Beam Machining	4	2	L16	2	1	-2.48 2.77	-3.01
			4	2	L16	1	0	1.64	
			4	2	L16	1	0	1.15	
			4	2	L16	2	1	-1.61 1.25	-1.07
21	Huang & Lin 2008	Injection Molding	3	2	L8	3	1	-8.66 -3.98 2.20	0.96
			3	2	L8	3	1	9.90 4.55 -2.51	1.22
22	Pei & Strasbaugh 2002	Fine Grinding	3	2	L8	2	1	0.74 1.82	0.64
			3	2	L8	2	1	-0.52 -3.57	-0.66
			3	2	L8	1	2	2.07	0.55 0.62
23	Pei at al 2003	Grinding	4	2	L16	3	3	-2.51 4.87 1.87	-2.30 -0.83 1.69
24	Pei 2002	Grinding	3	2	L8	3	2	3.65 0.91 2.70	0.86 -2.78
			3	2	L8	2	2	-9.21 7.46	0.59 -1.84
			3	2	L8	2	1	1.78 1.20	-1.36
			3	2	L8	0	0	0.00	0.00
			3	2	L8	3	2	-1.15 -0.65 1.35	0.75 -0.65

No	Author(s)	Process	No. Factors	No of Levels	OA	No. of Sig. Effects		Standardised Effects	
						Main	2fi	Main	2fi
25	Sun et al 2004	Grinding	4	2	L16	3	1	1.43 -1.68 6.57	1.09
			4	2	L16	3	0	-1.94 -0.92 0.60	
26	Li et al 2005	Ultrasonic Machining	3	2	L8	1	0	0.65	
			3	2	L8	3	2	0.93 -0.77 1.55	-1.75 -0.75
			3	2	L8	1	2	1.23	0.92 0.98
			3	2	L8	2	1	0.70 -0.78	0.85
27	Gorana et al 2004	Abrasive Flow Machining	3	2	L8	3	1	1.68 2.17 4.25	2.75
			3	2	L8	1	0	2.07	
			3	2	L8	1	0	0.97	
			3	2	L8	0	0	0.00	0.00
28	Pei et al 1999	Grinding	3	2	L8	1	0	-0.51	
29	Reddy et al 2001	Turning	5	2	L32	5	3	1.44 0.81 -0.79 0.94 -4.77	1.15 0.93 -1.04

No	Author(s)	Process	No. Factors	No of Levels	OA	No. of Sig. Effects		Standardised Effects			
						Main	2fi	Main	Type	2fi	Type
30	Correia & Ferraresi 2007	Welding	2	3	L9	3	1	-1.77	L	0.92	LxL
								3.75	L		
								1.63	Q		
			2	3	L9	1	0	-1.93	Q		
			2	3	L9	2		-3.40	L		
								-1.77	Q		
31	Chattopadhyay et al 2009	Electric Discharge Machining	3	3	L27	3	1	2.80	L	-0.93	LxL
								-1.98	L		
								-2.44	L		
			3	3	L27	4	1	-3.20	L	0.63	LxL
								0.56	L		
								2.30	L		
								1.95	Q		
			3	3	L27	3		0.98	L		
								-1.45	L		
								1.18	L		
32	Gaitonde et al 2008	Drilling	3	3	L27	4		-0.84	L		
								1.43	L		
								1.45	L		
								-0.74	Q		
33	Dhar et al 2007	Electric Discharge Machining	3	3	L27	4	1	-1.24	L	-0.71	LxL
								1.49	L		
								0.74	L		
								-0.77	Q		
								0.97	Q		
			3	3	L27	3		4.30	L		
								0.84	L		
								-1.90	Q		
			3	3	L27	4		1.49	L		
								1.55	L		
								2.72	L		
								-1.40	Q		
34	Ghani et al 2004	End Milling	3	3	L27	4	3	0.80	L	-1.51	LxL
								2.44	L	2.60	LxQ
								-1.00	L	-3.60	QxL
								0.94	Q		
			3	3	L27	2	0	1.58	L		
								1.38	L		

No	Author(s)	Process	No. Factors	No of Levels	OA	No. of Sig. Effects		Standardised Effects			
						Main	2fi	Main	Type	2fi	Type
35	Davim et al 2008	Turning	3	3	L27	3	1	-2.89 -0.64 1.39	L L Q	2.61	LxQ
36	Davim 2003	Turning	3	3	L27	6	3	1.33 -1.72 4.12 1.67 0.89 -1.46	L L L Q Q Q	1.49 -2.58 -1.09	LxL LxL LxL
			3	3	L27	5	5	1.20 5.63 1.43 0.86 -3.25	L L L Q Q	-2.17 -1.08 1.38 -1.90 1.62	LxL LxL QxL LxQ QxL
			3	3	L27	4	0	1.20 1.35 1.06 1.06	L L L Q		
37	Davim 2000	Drilling	3	3	L27	1	3	1.02	L	-0.89 -0.67 0.62	LxL LxL LxL
			3	3	L27	5	2	0.67 0.50 2.60 0.68 2.60	L L L Q Q	0.58 1.93	LxL LxL

REFERENCES

Ganjigatti, J., Pratihar, D. and Roy Choudhury, A. (2007), "Global Versus Cluster-Wise Regression Analyses For Prediction Of Bead Geometry In MIG Welding Process", *Journal of Materials Processing Technology*, Vol. 189, No. 1-3, pp. 352-366.

Gunaraj, V. and Murugan, N. (1999), "Prediction And Comparison Of The Area Of The Heat-Affected Zone For The Bead-On-Plate And Bead-On-Joint In Submerged Arc Welding Of Pipes", *Journal of Materials Processing Technology*, Vol. 95, No. 1-3, pp. 246-261.

Darwish, S. (2000), "The Impact Of The Tool Material And The Cutting Parameters On Surface Roughness Of Supermet 718 Nickel Superalloy", *Journal of Materials Processing Technology*, Vol. 97, No. 1-3, pp. 10-18.

Aggarwal, A., Singh, H., Kumar, P. and Singh, M. (2008), "Optimizing Power Consumption For CNC Turned Parts Using Response Surface Methodology And Taguchi's Technique--A Comparative Analysis", *Journal of Materials Processing Technology*, Vol. 200, No. 1-3, pp. 373-384.

Mahadevan, K., Raghukandan, K., Pai, B. and Pillai, U. (2008), "Influence Of Precipitation Hardening Parameters On The Fatigue Strength Of AA 6061-SiCp Composite", *Journal of Materials Processing Technology*, Vol. 198, No. 1-3, pp. 241-247.

Francis, A. and El-Midany, A. (2008), "An Assessment Of The Carbothermic Reduction Of Ilmenite Ore By Statistical Design", *Journal of Materials Processing Technology*, Vol. 199, No. 1-3, pp. 279-286.

Darwin, J., Mohan Lal, D. and Nagarajan, G. (2008), "Optimization Of Cryogenic Treatment To Maximize The Wear Resistance Of 18% Cr Martensitic Stainless Steel By Taguchi Method", *Journal of Materials Processing Technology*, Vol. 195, No. 1-3, pp. 241-247.

Senthilvelan, T., Raghukandan, K. and Venkatraman, A. (2003), "Modelling Of Process Parameters On The Working Of P/M Copper Preforms", *Journal of Materials Processing Technology*, Vol. 142, No. 3, pp. 767-772.

Puertas, I. and Luis, C. (2003), "A Study On The Machining Parameters Optimisation Of Electrical Discharge Machining", *Journal of Materials Processing Technology*, Vol. 143, No. 1, pp. 521-526.

Onwubolu, G. and Kumar, S. (2006), "Response Surface Methodology-Based Approach To CNC Drilling Operations", *Journal of Materials Processing Technology*, Vol. 171, No. 1, pp. 41-47.

Chatterjee, D., Oraon, B., Sutradhar, G. and Bose, P. (2007), "Prediction Of Hardness For Sintered HSS Components Using Response Surface Method", *Journal of Materials Processing Technology*, Vol. 190, No. 1-3, pp. 123-129.

Mahadevan, K., Raghukandan, K., Senthilvelan, T., Pai, B. and Pillai, U. (2006), "Investigation On The Influence Of Heat-Treatment Parameters On The Hardness Of AA6061-SiCp Composite", *Journal of Materials Processing Technology*, Vol. 171, No. 2, pp. 314-318.

Raghukandan, K. and Senthilvelan, T. (2004), "Analysis of P/M Hollow Extrusion Using Design Of Experiments", *Journal of Materials Processing Technology*, Vol. 153, No. 1, pp. 416-419.

Yang, Y., Chuang, M. and Lin, S. (2009), "Optimization of Dry Machining Parameters For High-Purity Graphite In End Milling Process Via Design of Experiments Methods", *Journal of Materials Processing Technology*, Vol. 209, No. 9, pp. 4395-4400.

Noordin, M., Venkatesh, V., Sharif, S., Elting, S. and Abdullah, A. (2004), "Application of Response Surface Methodology In Describing The Performance Of Coated Carbide Tools When Turning AISI 1045 Steel", *Journal of Materials Processing Technology*, Vol. 145, No. 1, pp. 46-58.

Kannan, T. and Murugan, N. (2006), "Effect Of Flux Cored Arc Welding Process Parameters On Duplex Stainless Steel Clad Quality", *Journal of Materials Processing Technology*, Vol. 176, No. 1-3, pp. 230-239.

Dutta, P. and Pratihar, D. (2007), "Modeling of TIG Welding Process Using Conventional Regression Analysis and Neural Network-Based Approaches", *Journal of Materials Processing Technology*, Vol. 184, No. 1-3, pp. 56-68.

Krajnik, P. and Kopac, J. (2004), "Adequacy of Matrix Experiment in Grinding", *Journal of Materials Processing Technology*, Vol. 157, No. 1, pp. 566-572.

Mohammadi, Z., Ziaei-Moayyed, A. and Mesgar, A. (2007), "Grit Blasting of Ti-6Al-4V Alloy: Optimization and Its Effect on Adhesion Strength of Plasma-Sprayed Hydroxyapatite Coatings", *Journal of Materials Processing Technology*, Vol. 194, No. 1-3, pp. 15-23.

Hung, N., Fu, Y. and Ali, M. (2002), "Focused Ion Beam Machining Of Silicon", *Journal of Materials Processing Technology*, Vol. 127, No. 2, pp. 256-260.

Huang, M. and Lin, T. (2008), "An Innovative Regression Model-Based Searching Method For Setting The Robust Injection Molding Parameters", *Journal of Materials Processing Technology*, Vol. 198, No. 1-3, pp. 436-444.

Pei, Z. and Strasbaugh, A. (2002), "Fine Grinding Of Silicon Wafers: Designed Experiments", *International Journal of Machine Tools and Manufacture*, Vol. 42, No. 3, pp. 395-404.

Pei, Z., Xin, X. and Liu, W. (2003), "Finite Element Analysis For Grinding Of Wire-Sawn Silicon Wafers: A Designed Experiment", *International Journal of Machine Tools and Manufacture*, Vol. 43, No. 1, pp. 7-16.

Pei, Z. (2002), "A Study On Surface Grinding Of 300 Mm Silicon Wafers", *International Journal of Machine Tools and Manufacture*, Vol. 42, No. 3, pp. 385-393.

Sun, X., Pei, Z., Xin, X. and Fouts, M. (2004), "Waviness Removal In Grinding Of Wire-Sawn Silicon Wafers: 3D Finite Element Analysis With Designed Experiments", *International Journal of Machine Tools and Manufacture*, Vol. 44, No. 1, pp. 11-19.

Li, Z., Jiao, Y., Deines, T., Pei, Z. and Treadwell, C. (2005), "Rotary Ultrasonic Machining Of Ceramic Matrix Composites: Feasibility Study And Designed Experiments", *International Journal of Machine Tools and Manufacture*, Vol. 45, No. 12-13, pp. 1402-1411.

Gorana, V., Jain, V. and Lal, G. (2004), "Experimental Investigation Into Cutting Forces And Active Grain Density During Abrasive Flow Machining", *International Journal of Machine Tools and Manufacture*, Vol. 44, No. 2-3, pp. 201-211.

Pei, Z., Billingsley, S. and Miura, S. (1999), "Grinding Induced Subsurface Cracks In Silicon Wafers", *International Journal of Machine Tools and Manufacture*, Vol. 39, No. 7, pp. 1103-1116.

Reddy, R., DeVor, R., Kapoor, S. and Sun, Z. (2001), "A Mechanistic Model-Based Force-Feedback Scheme For Voice-Coil Actuated Radial Contour Turning", *International Journal of Machine Tools and Manufacture*, Vol. 41, No. 8, pp. 1131-1147.

Correia, D. and Ferraresi, V. (2007), "Welding Process Selection Through A Double Criteria: Operational Costs And Non-Quality Costs", *Journal of Materials Processing Technology*, Vol. 184, No. 1-3, pp. 47-55.

Chattopadhyay, K., Verma, S., Satsangi, P. and Sharma, P. (2009), "Development Of Empirical Model For Different Process Parameters During Rotary Electrical Discharge Machining Of Copper-Steel (EN-8) System", *Journal of Materials Processing Technology*, Vol. 209, No. 3, pp. 1454-1465.

Gaitonde, V., Karnik, S., Rubio, J., Correia, A., Abrão, A. and Davim, J. (2008), "Analysis of Parametric Influence on Delamination in High-Speed Drilling of Carbon Fiber Reinforced Plastic Composites", *Journal of Materials Processing Technology*, Vol. 203, No. 1-3, pp. 431-438.

Dhar, S., Purohit, R., Saini, N., Sharma, A. and Kumar, G. (2007), "Mathematical Modeling Of Electric Discharge Machining Of Cast Al-4Cu-6Si Alloy-10 wt.% SiCP Composites", *Journal of Materials Processing Technology*, Vol. 194, No. 1-3, pp. 24-29.

Ghani, J., Choudhury, I. and Hassan, H. (2004), "Application Of Taguchi Method In The Optimization Of End Milling Parameters", *Journal of Materials Processing Technology*, Vol. 145, No. 1, pp. 84-92.

Davim, J., Gaitonde, V. and Karnik, S. (2008), "Investigations Into The Effect Of Cutting Conditions On Surface Roughness In Turning Of Free Machining Steel By ANN Models", *Journal of Materials Processing Technology*, Vol. 205, No. 1-3, pp. 16-23.

Davim, J. (2003), "Design Of Optimisation Of Cutting Parameters For Turning Metal Matrix Composites Based On The Orthogonal Arrays", *Journal of Materials Processing Technology*, Vol. 132, No. 1-3, pp. 340-344.

Davim, J. (2000), "An Experimental Study Of The Tribological Behaviour Of The Brass/Steel Pair", *Journal of Materials Processing Technology*, Vol. 100, No. 1-3, pp. 273-277.

APPENDIX 5:
RESULTS OF THE SIMULATION EXPERIMENTS

Phase-1 Results

No.	Fractional OA (A)	Pooling Methods (B)	No. of Active Factors (C)	Size of Active Factors (D)	Power
1	L12	MLM	4	3	0.685
2	L12	PU	4	3	0.921
3	L12	PD	5	2	0.754
4	L16	PD	4	1	0.527
5	L12	HNP	5	1	0.547
6	L16	PU	3	3	0.98
7	L12	HNP	3	3	0.911
8	L8	PD	5	2	0.537
9	L16	UC	5	1.5	0.695
10	L16	UC	5	1	0.517
11	L12	UC	5	1.5	0.665
12	L16	UC	4	3	0.862
13	L16	LM	5	1.5	0.429
14	L12	LM	5	1.5	0.389
15	L12	PU	5	1	0.596
16	L16	MLM	5	2	0.655
17	L12	MLM	3	3	0.685
18	L16	PU	4	3	0.99
19	L8	MLM	3	1.5	0.488
20	L8	UC	3	2	0.685
21	L8	MLM	4	1	0.202
22	L12	HNP	4	1	0.566
23	L12	PD	5	3	0.773
24	L16	HNP	4	3	0.911
25	L8	UC	5	1	0.291
26	L16	PU	3	1	0.645
27	L12	UC	3	1	0.527
28	L8	PU	4	1.5	0.645
29	L12	LM	3	1	0.261
30	L12	MLM	4	1.5	0.566
31	L16	LM	3	1.5	0.468
32	L12	UC	3	3	0.832
33	L16	LM	4	1	0.261
34	L8	PU	3	1	0.547
35	L8	HNP	3	1.5	0.665
36	L8	LM	5	3	0.33

37	L8	HNP	5	3	0.645
38	L16	MLM	3	2	0.685
39	L8	PU	3	2	0.813
40	L12	PD	5	1	0.468
41	L12	PU	5	2	0.852
42	L12	PD	4	1.5	0.665
43	L8	PD	4	3	0.635
44	L8	MLM	5	1	0.143
45	L8	UC	3	1.5	0.606
46	L16	LM	3	2	0.557
47	L16	PD	4	1.5	0.724
48	L16	UC	3	1.5	0.744
49	L8	LM	3	1.5	0.35
50	L12	LM	4	1.5	0.419
51	L12	PU	5	1.5	0.793
52	L12	MLM	5	3	0.665
53	L12	PU	3	3	0.951
54	L8	UC	4	1.5	0.527
55	L8	PD	3	3	0.724
56	L16	PU	4	1.5	0.852
57	L8	UC	3	3	0.734
58	L16	LM	5	1	0.261
59	L8	UC	5	2	0.557
60	L16	PU	5	1	0.645
61	L8	PU	4	3	0.763
62	L8	HNP	3	3	0.793
63	L8	MLM	5	2	0.409
64	L16	HNP	5	2	0.852
65	L16	HNP	5	1.5	0.783
66	L12	LM	3	3	0.557
67	L8	UC	4	2	0.635
68	L8	LM	4	1.5	0.232
69	L8	HNP	4	1.5	0.606
70	L16	PD	5	2	0.783
71	L8	UC	5	1.5	0.468
72	L16	HNP	4	1.5	0.793
73	L12	UC	3	2	0.803
74	L8	HNP	3	1	0.488
75	L16	HNP	3	1	0.645
76	L8	MLM	5	3	0.448
77	L8	HNP	3	2	0.734
78	L8	LM	5	1	0.005

79	L12	LM	4	2	0.498
80	L12	HNP	3	2	0.832
81	L16	UC	3	2	0.842
82	L8	MLM	4	3	0.517
83	L12	LM	5	3	0.527
84	L16	PD	5	3	0.832
85	L12	HNP	4	1.5	0.763
86	L8	PD	4	1	0.33
87	L8	HNP	4	1	0.419
88	L8	MLM	5	1.5	0.35
89	L12	HNP	5	2	0.842
90	L8	LM	3	3	0.468
91	L16	PD	3	1.5	0.744
92	L16	MLM	4	1.5	0.596
93	L12	PU	4	2	0.882
94	L16	MLM	3	1	0.429
95	L16	LM	4	2	0.547
96	L16	PD	5	1	0.517
97	L12	HNP	4	3	0.892
98	L8	PU	5	1	0.389
99	L12	PU	4	1.5	0.813
100	L16	LM	5	2	0.527
101	L12	LM	4	3	0.566
102	L8	PD	4	1.5	0.527
103	L16	MLM	4	2	0.675
104	L12	HNP	5	1.5	0.754
105	L8	PU	4	1	0.438
106	L16	UC	4	1.5	0.724
107	L8	LM	4	3	0.379
108	L12	MLM	5	1.5	0.537
109	L12	UC	5	3	0.813
110	L8	PU	5	1.5	0.586
111	L16	UC	4	1	0.547
112	L12	PD	3	2	0.754
113	L16	PU	4	2	0.911
114	L12	PD	4	1	0.488
115	L12	LM	5	1	0.232
116	L16	PU	5	2	0.911
117	L12	MLM	3	2	0.655
118	L8	MLM	3	2	0.547
119	L12	MLM	5	2	0.616
120	L12	LM	3	1.5	0.448

121	L16	HNP	3	2	0.882
122	L16	PU	5	3	0.941
123	L16	HNP	4	1	0.596
124	L8	LM	3	2	0.409
125	L12	UC	4	3	0.832
126	L16	LM	3	3	0.616
127	L16	MLM	5	1.5	0.586
128	L16	PD	3	1	0.537
129	L8	HNP	4	2	0.665
130	L8	LM	4	2	0.35
131	L8	PD	3	1	0.409
132	L16	PU	4	1	0.655
133	L12	MLM	3	1.5	0.557
134	L8	UC	4	3	0.665
135	L12	PU	3	2	0.911
136	L16	PD	5	1.5	0.714
137	L8	LM	3	1	0.153
138	L16	PD	3	2	0.803
139	L12	PU	3	1	0.635
140	L12	PU	5	3	0.921
141	L16	MLM	5	1	0.389
142	L16	PD	4	2	0.803
143	L8	MLM	3	3	0.596
144	L8	PU	4	2	0.734
145	L12	PD	3	3	0.813
146	L8	PD	3	1.5	0.596
147	L8	MLM	4	2	0.488
148	L16	LM	3	1	0.281
149	L16	UC	3	1	0.566
150	L8	PD	4	2	0.606
151	L12	MLM	3	1	0.389
152	L16	UC	5	2	0.793
153	L8	PU	5	3	0.704
154	L8	HNP	5	1	0.35
155	L12	HNP	3	1.5	0.773
156	L16	HNP	5	3	0.892
157	L12	UC	4	1	0.517
158	L8	PD	5	3	0.586
159	L12	PU	3	1.5	0.823
160	L16	PU	3	1.5	0.852
161	L12	PD	3	1	0.517
162	L8	PD	3	2	0.685

163	L8	HNP	4	3	0.714
164	L12	UC	4	2	0.783
165	L12	LM	5	2	0.488
166	L12	PU	4	1	0.616
167	L12	PD	3	1.5	0.695
168	L8	PD	5	1	0.241
169	L12	PD	5	1.5	0.665
170	L16	HNP	3	3	0.931
171	L16	HNP	5	1	0.606
172	L16	MLM	5	3	0.704
173	L8	PU	3	3	0.823
174	L16	PD	4	3	0.842
175	L12	PD	4	3	0.813
176	L8	MLM	3	1	0.301
177	L16	PU	3	2	0.941
178	L12	HNP	4	2	0.842
179	L16	MLM	3	3	0.724
180	L12	UC	5	2	0.754
181	L12	PD	4	2	0.793
182	L8	PU	3	1.5	0.724
183	L8	LM	4	1	0.064
184	L8	LM	5	1.5	0.202
185	L12	HNP	3	1	0.576
186	L12	MLM	4	1	0.369
187	L12	LM	4	1	0.222
188	L16	LM	4	1.5	0.468
189	L16	MLM	3	1.5	0.606
190	L16	PD	3	3	0.872
191	L16	LM	4	3	0.586
192	L12	LM	3	2	0.507
193	L8	LM	5	2	0.281
194	L16	UC	5	3	0.832
195	L16	MLM	4	3	0.724
196	L16	HNP	4	2	0.872
197	L16	PU	5	1.5	0.832
198	L16	UC	4	2	0.823
199	L16	HNP	3	1.5	0.813
200	L16	LM	5	3	0.557
201	L12	HNP	5	3	0.862
202	L12	UC	3	1.5	0.704
203	L16	MLM	4	1	0.409
204	L8	UC	5	3	0.596

205	L16	UC	3	3	0.872
206	L8	MLM	4	1.5	0.389
207	L8	UC	3	1	0.409
208	L8	HNP	5	1.5	0.507
209	L12	UC	5	1	0.498
210	L8	UC	4	1	0.35
211	L12	MLM	5	1	0.35
212	L12	MLM	4	2	0.645
213	L8	HNP	5	2	0.616
214	L8	PD	5	1.5	0.448
215	L12	UC	4	1.5	0.714
216	L8	PU	5	2	0.655

Sub-Phase-1 Results

No.	Fractional OA (A)	Pooling Methods (B)	Replicates						Power		
			1	2	3	4	5	6	Avg.	S^2	ln S^2
1	L16	PU	1	1	0.64	1	0.928	1	0.928	0.021	-3.86323
2	L12	PU	1	0.96	0.752	1	1	0.73	0.907	0.017	-4.07454
3	L8	PU	0.918	0.897	0.778	0.835	0.536	0.536	0.75	0.03	-3.50656
4	L16	HN	0.956	1	1	1	0.872	0.872	0.95	0.004	-5.52146
5	L12	HN	1	1	0.67	0.954	0.899	0.997	0.92	0.017	-4.07454
6	L8	HN	0.878	0.743	0.449	0.601	0.449	0.601	0.62	0.028	-3.57555
7	L16	EC	0.836	0.938	0.924	0.93	0.826	0.826	0.88	0.003	-5.80914
8	L12	EC	0.782	0.896	0.878	0.887	0.768	0.768	0.83	0.004	-5.52146
9	L8	EC	0.598	0.7	0.682	0.69	0.585	0.585	0.64	0.003	-5.80914
10	L16	PD	0.756	0.856	0.841	0.848	0.75	0.75	0.8	0.003	-5.80914
11	L12	PD	0.898	0.916	0.867	0.89	0.645	0.645	0.81	0.017	-4.07454
12	L8	PD	0.769	0.732	0.573	0.653	0.377	0.377	0.58	0.029	-3.54046
13	L16	LM2	0.644	0.768	0.755	0.762	0.666	0.666	0.71	0.003	-5.80914
14	L12	LM2	0.656	0.794	0.677	0.737	0.488	0.488	0.64	0.016	-4.13517
15	L8	LM2	0.583	0.687	0.419	0.552	0.268	0.268	0.463	0.03	-3.50656
16	L16	LM1	0.504	0.594	0.579	0.587	0.488	0.488	0.54	0.003	-5.80914
17	L12	LM1	0.455	0.632	0.46	0.546	0.304	0.304	0.45	0.017	-4.07454
18	L8	LM1	0.401	0.56	0.177	0.401	0.139	0.139	0.303	0.031	-3.47377

Phase-2 Results

No	OAs & Pooling Methods (A)	No. of Active Factors (B)	No. of Active Two-Factor Interactions (C)	Size of Active Effects (D)	Power
1	L8-DO-SWR	3	3	1	0.222
2	L8-BM-SWR	4	2	3	0.765
3	L12-BSS	4	1	1.5	0.783
4	L8-BM-SWR	4	2	1.5	0.67
5	L16-HNP	5	2	3	0.827
6	L12-SWR	3	3	2	0.579
7	L8-BM-BSS	3	2	1	0.745
8	L8-DO-BSS	4	3	2	0.34
9	L12-SWR	3	2	2	0.958
10	L8-DO-BSS	3	3	3	0.25
11	L12-BSS	3	1	1	0.731
12	L8-DO-SWR	5	3	1.5	0.093
13	L12-SWR	3	3	1.5	0.709
14	L8-BM-BSS	4	3	1.5	0.388
15	L8-DO-BSS	3	3	1	0.287
16	L12-SWR	5	3	1	0.392
17	L8-DO-SWR	5	1	3	0.76
18	L8-DO-SWR	5	2	1	0.519
19	L8-DO-BSS	4	2	2	0.625
20	L12-BSS	4	3	1	0.626
21	L16-HNP	5	2	1.5	0.776
22	L8-BM-SWR	5	2	1	0.56
23	L16-HNP	5	3	2	0.717
24	L8-DO-SWR	3	2	2	0.735
25	L8-DO-BSS	5	2	1	0.554
26	L12-SWR	3	3	3	0.625
27	L8-BM-BSS	4	1	1.5	0.889
28	L8-BM-SWR	3	1	1	0.778
29	L8-DO-BSS	3	1	3	0.73
30	L12-BSS	5	2	3	0.81
31	L12-SWR	5	2	1	0.678
32	L8-BM-SWR	5	3	2	0.418
33	L8-DO-SWR	4	3	1.5	0.068
34	L8-DO-BSS	4	3	1.5	0.298
35	L8-BM-BSS	4	1	1	0.597

36	L12-SWR	4	3	1.5	0.417
37	L8-DO-SWR	3	2	1	0.675
38	L8-BM-BSS	4	3	1	0.45
39	L12-BSS	3	1	2	0.917
40	L8-BM-BSS	5	1	1	0.684
41	L8-DO-BSS	3	1	1.5	0.789
42	L8-BM-SWR	3	3	1.5	0.374
43	L16-HNP	5	3	3	0.725
44	L8-DO-BSS	3	3	2	0.393
45	L8-DO-SWR	5	3	2	0.02
46	L8-DO-BSS	3	2	1.5	0.655
47	L8-BM-BSS	4	2	2	0.805
48	L12-BSS	3	2	3	0.85
49	L12-SWR	5	3	2	0.485
50	L12-BSS	3	3	2	0.642
51	L12-BSS	3	3	1	0.519
52	L8-DO-BSS	3	2	2	0.734
53	L8-BM-BSS	5	2	2	0.684
54	L16-HNP	4	3	1	0.511
55	L8-BM-BSS	5	1	1.5	0.725
56	L16-HNP	4	2	2	0.761
57	L8-BM-SWR	4	1	3	0.776
58	L16-HNP	3	1	1.5	0.953
59	L8-BM-SWR	4	1	2	0.791
60	L12-BSS	4	1	3	0.888
61	L8-BM-BSS	3	2	1.5	0.804
62	L12-SWR	4	2	1	0.773
63	L12-BSS	3	2	2	0.889
64	L12-BSS	5	1	2	0.8
65	L16-HNP	5	3	1	0.492
66	L8-DO-BSS	3	1	2	0.845
67	L8-DO-BSS	5	2	3	0.728
68	L8-DO-BSS	4	2	1	0.635
69	L8-BM-SWR	4	3	2	0.417
70	L12-SWR	5	1	1.5	0.765
71	L16-HNP	4	1	3	0.871
72	L8-BM-SWR	5	2	3	0.786
73	L12-SWR	4	1	1.5	0.821
74	L8-BM-SWR	5	1	2	0.654
75	L8-BM-SWR	3	1	2	0.768
76	L16-HNP	4	1	2	0.891
77	L12-SWR	3	1	1	0.792

78	L8-DO-SWR	4	1	1	0.648
79	L8-DO-BSS	5	1	1	0.664
80	L16-HNP	3	1	1	0.788
81	L16-HNP	3	3	1.5	0.721
82	L8-DO-BSS	5	1	2	0.672
83	L16-HNP	4	1	1.5	0.864
84	L12-SWR	4	2	3	0.841
85	L8-BM-BSS	3	2	3	0.811
86	L12-BSS	4	2	2	0.75
87	L12-BSS	5	1	1	0.824
88	L12-SWR	5	2	3	0.769
89	L12-BSS	3	1	1.5	0.956
90	L16-HNP	5	1	3	0.906
91	L12-BSS	4	1	2	0.856
92	L8-DO-SWR	5	1	2	0.687
93	L16-HNP	3	2	2	0.851
94	L12-BSS	3	1	3	0.907
95	L8-BM-BSS	5	3	1.5	0.292
96	L12-SWR	3	2	1	0.673
97	L8-DO-BSS	3	1	1	0.714
98	L8-BM-SWR	3	2	2	0.647
99	L8-DO-BSS	4	1	2	0.851
100	L16-HNP	5	2	2	0.83
101	L8-BM-SWR	5	1	3	0.821
102	L8-DO-SWR	3	1	1	0.704
103	L12-BSS	5	2	2	0.817
104	L8-DO-SWR	4	1	3	0.798
105	L8-DO-SWR	5	3	3	0.069
106	L16-HNP	3	1	3	0.858
107	L12-BSS	4	3	2	0.621
108	L8-DO-BSS	4	1	1.5	0.688
109	L8-DO-SWR	3	3	3	0.261
110	L12-BSS	3	2	1.5	0.772
111	L8-DO-SWR	3	3	2	0.115
112	L8-DO-SWR	5	1	1	0.695
113	L8-BM-SWR	3	3	1	0.415
114	L8-DO-SWR	4	1	1.5	0.662
115	L12-BSS	4	2	1	0.705
116	L8-BM-BSS	4	2	1.5	0.744
117	L12-SWR	5	2	2	0.768
118	L12-BSS	3	3	1.5	0.635
119	L12-SWR	3	1	3	0.833

120	L8-DO-SWR	4	1	2	0.746
121	L8-BM-SWR	4	2	1	0.682
122	L16-HNP	4	3	1.5	0.679
123	L8-DO-SWR	4	3	2	0.124
124	L8-DO-SWR	3	2	3	0.861
125	L8-DO-BSS	5	3	1.5	0.3
126	L8-BM-SWR	4	1	1.5	0.71
127	L8-BM-SWR	3	2	1	0.659
128	L16-HNP	4	2	1.5	0.783
129	L8-BM-SWR	4	3	1	0.33
130	L12-BSS	4	3	3	0.781
131	L8-DO-BSS	3	3	1.5	0.34
132	L16-HNP	5	1	1	0.813
133	L8-BM-BSS	4	2	1	0.64
134	L16-HNP	3	3	2	0.836
135	L12-SWR	5	1	3	0.881
136	L12-SWR	4	3	1	0.422
137	L12-BSS	4	1	1	0.666
138	L12-SWR	4	2	2	0.8
139	L8-DO-BSS	5	2	1.5	0.687
140	L8-BM-BSS	5	2	3	0.662
141	L8-DO-SWR	5	2	1.5	0.568
142	L8-BM-SWR	5	3	1	0.251
143	L12-BSS	5	1	1.5	0.736
144	L8-DO-SWR	5	2	3	0.7
145	L8-BM-SWR	3	3	3	0.528
146	L8-DO-BSS	5	2	2	0.684
147	L12-BSS	3	3	3	0.776
148	L16-HNP	4	3	2	0.724
149	L8-DO-BSS	5	3	3	0.239
150	L8-BM-BSS	3	3	3	0.565
151	L8-DO-BSS	4	3	3	0.371
152	L8-BM-BSS	4	3	3	0.498
153	L8-BM-BSS	3	2	2	0.791
154	L12-BSS	3	2	1	0.751
155	L8-DO-SWR	3	1	2	0.787
156	L8-BM-BSS	3	3	2	0.526
157	L8-DO-SWR	4	2	3	0.667
158	L8-DO-BSS	3	2	1	0.782
159	L8-BM-SWR	3	1	1.5	0.723
160	L12-SWR	5	3	1.5	0.482
161	L12-SWR	5	2	1.5	0.797

162	L16-HNP	3	2	3	0.88
163	L8-DO-BSS	4	2	3	0.816
164	L12-SWR	4	3	2	0.461
165	L16-HNP	5	1	1.5	0.745
166	L8-DO-BSS	3	2	3	0.861
167	L16-HNP	3	1	2	0.9
168	L8-BM-BSS	5	1	3	0.828
169	L16-HNP	4	1	1	0.737
170	L16-HNP	5	2	1	0.798
171	L8-BM-SWR	3	1	3	0.899
172	L12-SWR	5	1	2	0.877
173	L12-SWR	3	1	2	0.871
174	L8-BM-SWR	3	3	2	0.485
175	L8-BM-BSS	3	1	1.5	0.96
176	L8-DO-BSS	5	1	3	0.729
177	L16-HNP	3	2	1.5	0.841
178	L16-HNP	4	2	3	0.863
179	L8-BM-SWR	5	2	1.5	0.623
180	L8-BM-SWR	5	2	2	0.726
181	L8-DO-SWR	4	2	1.5	0.624
182	L16-HNP	3	3	1	0.643
183	L8-DO-SWR	4	3	1	0
184	L8-BM-BSS	3	1	2	0.792
185	L8-BM-BSS	4	1	3	0.857
186	L8-BM-BSS	5	3	1	0.417
187	L8-DO-SWR	3	1	1.5	0.68
188	L8-BM-BSS	5	2	1.5	0.746
189	L8-BM-BSS	3	3	1	0.488
190	L8-BM-SWR	5	3	1.5	0.394
191	L8-BM-BSS	5	1	2	0.726
192	L8-BM-BSS	4	1	2	0.773
193	L12-BSS	5	3	2	0.57
194	L8-DO-SWR	5	1	1.5	0.625
195	L12-SWR	3	2	1.5	0.833
196	L8-BM-SWR	5	1	1	0.635
197	L8-DO-BSS	5	1	1.5	0.687
198	L12-SWR	4	1	3	0.881
199	L12-BSS	5	3	3	0.659
200	L8-BM-SWR	3	2	1.5	0.794
201	L8-DO-BSS	5	3	1	0.285
202	L8-DO-SWR	4	3	3	0.147
203	L16-HNP	3	2	1	0.738

204	L16-HNP	4	2	1	0.75
205	L12-SWR	3	3	1	0.551
206	L12-SWR	3	2	3	0.943
207	L8-DO-SWR	3	1	3	0.744
208	L8-BM-BSS	5	3	2	0.417
209	L8-BM-BSS	5	3	3	0.564
210	L12-BSS	5	2	1	0.657
211	L8-BM-BSS	3	1	3	0.86
212	L8-BM-SWR	4	2	2	0.684
213	L12-BSS	5	2	1.5	0.66
214	L8-DO-SWR	5	2	2	0.61
215	L16-HNP	5	1	2	0.863
216	L8-DO-SWR	4	2	1	0.664
217	L12-SWR	4	1	1	0.719
218	L12-BSS	5	3	1.5	0.547
219	L8-BM-BSS	3	3	1.5	0.524
220	L8-DO-SWR	3	3	1.5	0.143
221	L12-BSS	4	2	1.5	0.842
222	L12-SWR	4	1	2	0.793
223	L8-DO-BSS	4	1	1	0.559
224	L8-DO-SWR	4	2	2	0.687
225	L16-HNP	4	3	3	0.696
226	L12-BSS	5	1	3	0.766
227	L12-SWR	5	1	1	0.686
228	L8-BM-BSS	4	2	3	0.784
229	L8-DO-SWR	5	3	1	0.017
230	L12-SWR	5	3	3	0.58
231	L8-BM-SWR	4	1	1	0.715
232	L8-DO-SWR	3	2	1.5	0.676
233	L12-SWR	4	2	1.5	0.762
234	L8-DO-BSS	4	2	1.5	0.75
235	L12-SWR	4	3	3	0.593
236	L8-BM-SWR	5	1	1.5	0.636
237	L12-BSS	4	3	1.5	0.627
238	L8-BM-BSS	5	2	1	0.617
239	L8-DO-BSS	5	3	2	0.315
240	L8-DO-BSS	4	1	3	0.72
241	L8-BM-SWR	3	2	3	0.729
242	L8-BM-BSS	3	1	1	0.789
243	L16-HNP	5	3	1.5	0.511
244	L8-BM-SWR	4	3	1.5	0.42
245	L12-BSS	5	3	1	0.539

246	L16-HNP	3	3	3	0.744
247	L8-DO-BSS	4	3	1	0.169
248	L8-BM-SWR	4	3	3	0.464
249	L8-BM-SWR	5	3	3	0.427
250	L12-BSS	4	2	3	0.82
251	L12-SWR	3	1	1.5	0.778
252	L8-BM-BSS	4	3	2	0.488

Sub-Phase-2 Results

No.	OAs & Pooling Methods (A)	Replications						Power		
		1	2	3	4	5	6	Avg.	S^2	ln S^2
1	L16-HNP	1	1	0.874	0.984	0.73	0.951	0.923	0.011201	-4.49175
2	L12-BSS	1	0.765	0.734	0.819	0.674	0.869	0.810	0.0132	-4.3275
3	L12-SWR	0.997	0.554	0.689	0.862	0.675	0.724	0.750	0.024405	-3.71297
4	L8-BM-BSS	0.3	0.736	0.587	0.702	0.724	0.612	0.610	0.026805	-3.61916
5	L8-BM-SWR	0.241	0.602	0.684	0.631	0.667	0.596	0.570	0.027203	-3.60443
6	L8-DO-BSS	0.349	0.154	0.576	0.471	0.675	0.714	0.490	0.045004	-3.101
7	L8-DO-SWR	0.041	0.362	0.429	0.69	0.58	0.347	0.408	0.049995	-2.99583

Phase-3 Results

No.	Fractional OAs (A)	ANOVA Type (B)	Pooling Methods (C)	No. of Active Main Effects (D)	Size of Active Effects (E)	Power
1	L18	ANOVA	PU	2	3	0.881
2	L18	ANOVA	PD	3	1.5	0.601
3	L18	ANOVA	PU	2	2	0.832
4	L18	Reg.	PD	2	1	0.762
5	L18	ANOVA	PD	3	2	0.622
6	L9	Reg.	PU	3	1	0.562
7	L18	Reg.	PD	3	3	0.712
8	L18	Reg.	PD	2	3	0.862
9	L18	Reg.	PD	3	2	0.712
10	L18	ANOVA	PU	3	3	0.791
11	L9	ANOVA	PD	2	1	0.511
12	L18	Reg.	PD	3	1.5	0.681
13	L18	Reg.	PU	3	1	0.791
14	L9	ANOVA	PU	3	3	0.711
15	L9	ANOVA	PU	2	3	0.791
16	L9	Reg.	PU	3	1.5	0.641
17	L9	ANOVA	PD	3	2	0.521
18	L18	Reg.	PD	2	2	0.801
19	L18	Reg.	PU	2	1	0.871
20	L9	Reg.	PU	3	2	0.731
21	L18	Reg.	PU	3	3	0.851
22	L9	ANOVA	PU	2	1.5	0.682
23	L18	ANOVA	PU	3	1.5	0.732
24	L18	Reg.	PU	2	3	1.002
25	L9	ANOVA	PU	2	1	0.612
26	L9	Reg.	PU	2	2	0.812
27	L9	Reg.	PD	3	3	0.642
28	L18	ANOVA	PU	2	1.5	0.842
29	L18	Reg.	PD	2	1.5	0.792
30	L18	ANOVA	PD	2	1	0.682
31	L9	Reg.	PD	2	1	0.562
32	L18	ANOVA	PD	2	3	0.782
33	L9	Reg.	PD	2	2	0.702
34	L9	ANOVA	PU	3	1.5	0.572
35	L18	Reg.	PU	3	2	0.791

36	L9	Reg.	PD	2	3	0.771
37	L18	ANOVA	PU	3	2	0.761
38	L9	ANOVA	PU	3	1	0.551
39	L9	ANOVA	PD	2	2	0.641
40	L18	ANOVA	PD	2	2	0.701
41	L9	ANOVA	PU	2	2	0.781
42	L9	ANOVA	PD	3	3	0.571
43	L9	Reg.	PU	3	3	0.801
44	L9	Reg.	PD	3	1.5	0.501
45	L18	ANOVA	PU	2	1	0.771
46	L18	Reg.	PD	3	1	0.631
47	L9	ANOVA	PD	3	1.5	0.401
48	L9	ANOVA	PU	3	2	0.672
49	L9	ANOVA	PD	2	1.5	0.502
50	L18	Reg.	PU	3	1.5	0.792
51	L9	ANOVA	PD	2	3	0.712
52	L18	ANOVA	PU	3	1	0.702
53	L18	ANOVA	PD	3	3	0.662
54	L9	Reg.	PD	3	2	0.572
55	L9	Reg.	PD	3	1	0.462
56	L9	Reg.	PU	2	1	0.702
57	L18	ANOVA	PD	3	1	0.562
58	L9	Reg.	PD	2	1.5	0.642
59	L18	Reg.	PU	2	2	0.922
60	L18	Reg.	PU	2	1.5	0.882
61	L18	ANOVA	PD	2	1.5	0.722
62	L9	Reg.	PU	2	1.5	0.762
63	L9	ANOVA	PD	3	1	0.402
64	L9	Reg.	PU	2	3	0.862

Sub-Phase-3 Results

No.	Fractional OAs (A)	ANOVA Type (B)	Pooling Methods (C)	Replications						Power		
				1	2	3	4	5	6	Avg.	S^2	Ln S^2
1	L9	ANOVA	PD	0.440	0.513	0.466	0.669	0.687	0.440	0.571	0.0122	-4.4056
2	L18	ANOVA	PD	0.742	0.740	0.705	0.774	0.726	0.742	0.757	0.0029	-5.8283
3	L9	Reg.	PD	0.552	0.508	0.586	0.769	0.771	0.552	0.668	0.0179	-4.0207
4	L18	Reg.	PD	0.807	0.851	0.826	0.909	0.876	0.807	0.849	0.0014	-6.5588
5	L9	ANOVA	PU	0.533	0.557	0.644	0.892	0.877	0.533	0.726	0.0278	-3.5829
6	L18	ANOVA	PU	0.866	0.750	0.850	0.854	0.910	0.866	0.838	0.0031	-5.7808
7	L9	Reg.	PU	0.704	0.621	0.770	0.839	0.835	0.704	0.777	0.01	-4.6002
8	L18	Reg.	PU	0.953	0.876	0.853	0.918	0.937	0.953	0.915	0.0018	-6.3413

Phase-4 Results

No.	Fractional OAs (A)	Pooling Methods (B)	No. of Active Main Effects (C)	No. of Active Two-Factor Interactions (D)	Size of Active Effects (E)	Power
1	L9	BSS	3	1	3	0.721
2	L18	SWR	2	1	1	0.68
3	L9	SWR	2	1	3	0.792
4	L18	BSS	2	1	1	0.721
5	L18	SWR	3	1	2	0.62
6	L18	BSS	2	2	2	0.77
7	L9	BSS	2	1	2	0.85
8	L18	BSS	2	2	1	0.761
9	L18	BSS	2	2	1.5	0.76
10	L18	BSS	3	1	2	0.722
11	L18	BSS	2	1	2	0.831
12	L18	SWR	2	2	1	0.63
13	L9	SWR	3	1	1.5	0.48
14	L9	SWR	2	2	3	0.59
15	L9	BSS	2	2	1	0.511
16	L9	SWR	2	2	1	0.31
17	L18	BSS	3	1	1	0.702
18	L9	SWR	3	2	3	0.461
19	L18	BSS	3	2	3	0.72
20	L18	SWR	2	1	2	0.78
21	L18	BSS	3	2	1.5	0.69
22	L18	SWR	3	2	3	0.631
23	L9	SWR	3	2	1.5	0.34
24	L9	BSS	3	2	1	0.262
25	L18	SWR	3	2	1.5	0.631
26	L18	SWR	3	2	2	0.67
27	L9	SWR	2	1	1	0.54
28	L9	SWR	3	1	2	0.57
29	L9	SWR	3	1	3	0.591
30	L18	SWR	2	2	1.5	0.67
31	L9	SWR	2	1	1.5	0.632
32	L9	BSS	2	2	1.5	0.501
33	L9	BSS	3	2	1.5	0.39
34	L9	BSS	3	2	3	0.56
35	L18	SWR	2	2	2	0.7
36	L18	BSS	3	2	1	0.641

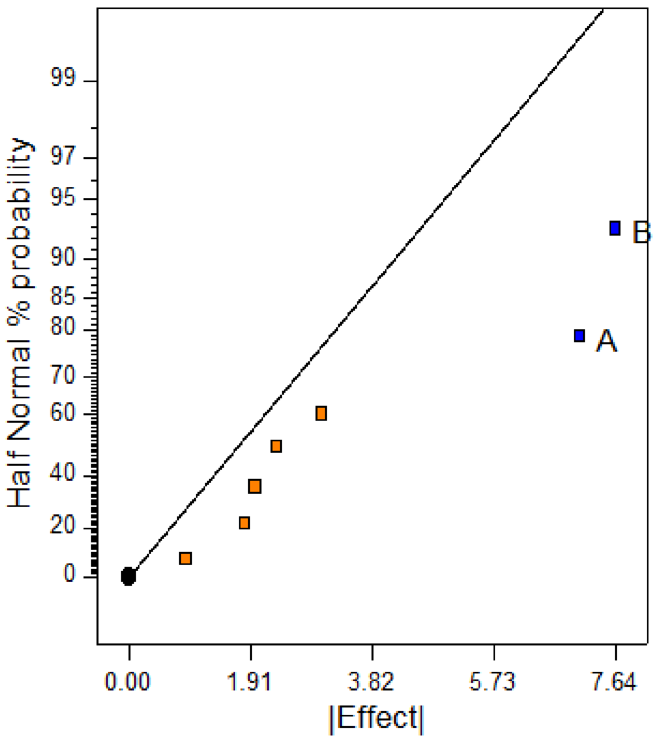
37	L9	BSS	2	1	1	0.64
38	L9	BSS	3	1	1.5	0.612
39	L9	BSS	3	1	1	0.451
40	L18	BSS	3	1	1.5	0.64
41	L9	BSS	2	2	2	0.48
42	L9	BSS	3	1	2	0.68
43	L18	BSS	3	1	3	0.701
44	L9	BSS	3	2	2	0.58
45	L18	SWR	2	1	3	0.782
46	L9	SWR	3	2	1	0.231
47	L18	BSS	2	2	3	0.77
48	L18	SWR	3	1	3	0.77
49	L9	BSS	2	1	3	0.95
50	L18	SWR	3	2	1	0.521
51	L9	SWR	2	2	1.5	0.44
52	L9	BSS	2	1	1.5	0.632
53	L18	SWR	2	1	1.5	0.681
54	L18	BSS	2	1	3	0.79
55	L9	SWR	2	2	2	0.52
56	L18	SWR	3	1	1.5	0.59
57	L18	BSS	3	2	2	0.611
58	L9	SWR	3	1	1	0.42
59	L18	SWR	2	2	3	0.742
60	L9	SWR	3	2	2	0.411
61	L18	BSS	2	1	1.5	0.81
62	L18	SWR	3	1	1	0.63
63	L9	SWR	2	1	2	0.72
64	L9	BSS	2	2	3	0.672

Sub-Phase-4 Results

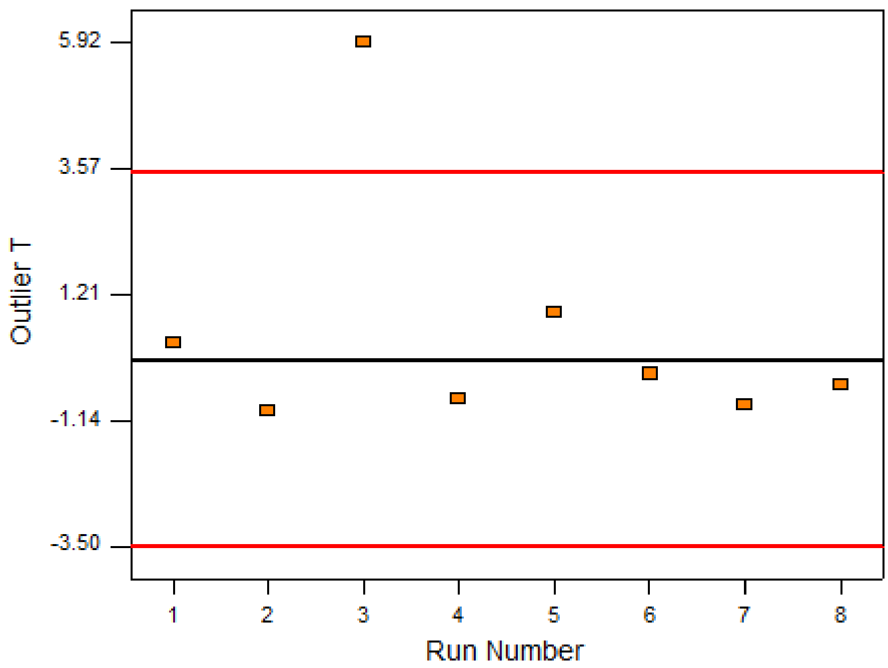
No.	Fractional OAs (A)	Pooling Methods (B)	Replications						Power		
			1	2	3	4	5	6	Avg.	S^2	Ln S^2
1	L9	SW	0.476	0.678	0.658	0.430	0.449	0.697	0.565	0.0173	-4.0567
2	L18	SW	0.793	0.890	0.782	0.781	0.743	0.769	0.793	0.0032	-5.7409
3	L9	BS	0.600	0.859	0.858	0.529	0.573	0.744	0.694	0.0243	-3.7177
4	L18	BS	0.936	0.952	1.000	0.987	0.987	0.904	0.961	0.0015	-6.4873

APPENDIX 6:
ANALYSIS OF THE L_8 RESPONSE 2 DATA USING HNP PLOT

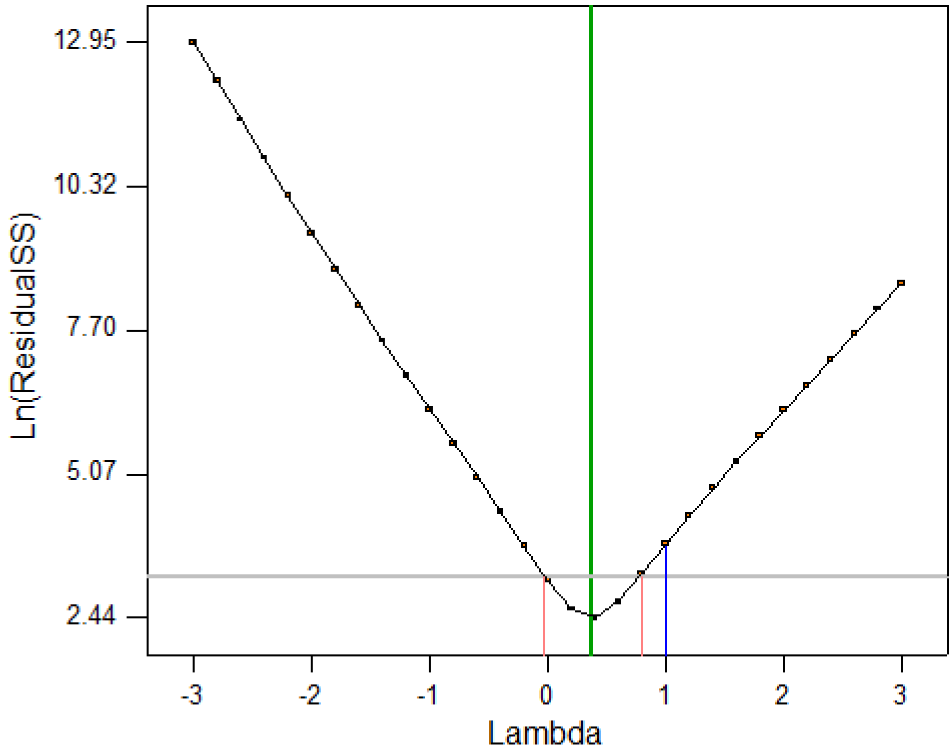
1. The HNP plot of the original data:



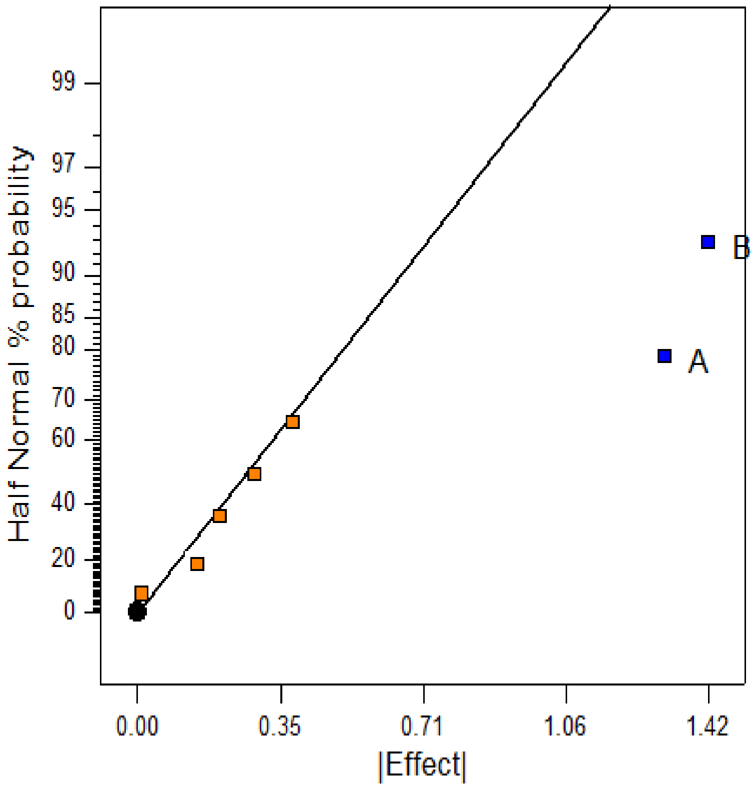
2. The residuals outliers plot of the original data



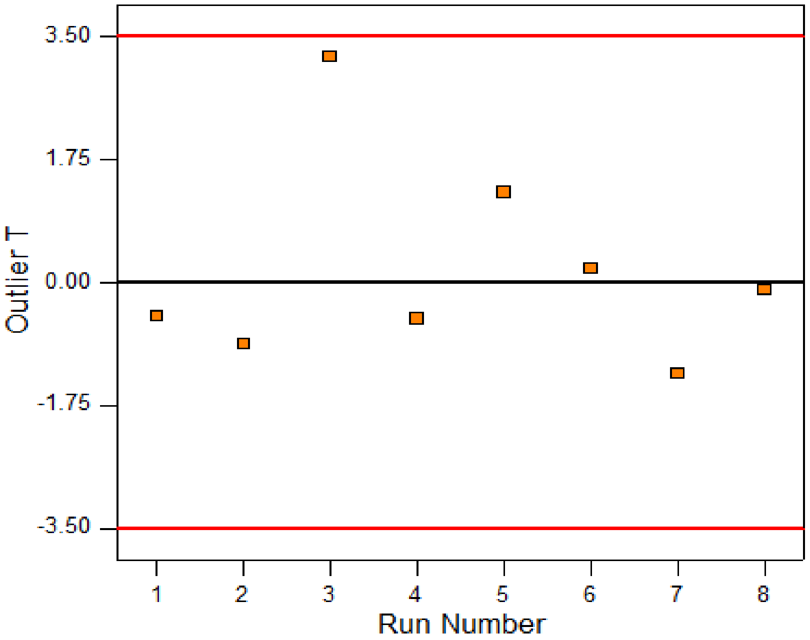
3. Box-Cox plot:



4. The HNP plot of the transformed data:



5. The residuals outliers plot of the transformed data:



APPENDIX 7:
FULL FACTORIAL DATA FOR THE PUBLISHED AND CONDUCTED EXPERIMENTS

Phase-1: Dyestuff Manufacturing Experiment

No.	A	B	C	D	E	F	Strength	Hue	Brightness
1	-1	-1	-1	-1	-1	-1	3.4	15	36
2	1	-1	-1	-1	-1	-1	9.7	5	35
3	-1	1	-1	-1	-1	-1	7.4	23	37
4	1	1	-1	-1	-1	-1	10.6	8	34
5	-1	-1	1	-1	-1	-1	6.5	20	30
6	1	-1	1	-1	-1	-1	7.9	9	32
7	-1	1	1	-1	-1	-1	10.3	13	28
8	1	1	1	-1	-1	-1	9.5	5	38
9	-1	-1	-1	1	-1	-1	14.3	23	40
10	1	-1	-1	1	-1	-1	10.5	1	32
11	-1	1	-1	1	-1	-1	7.8	11	32
12	1	1	-1	1	-1	-1	17.2	5	28
13	-1	-1	1	1	-1	-1	9.4	15	34
14	1	-1	1	1	-1	-1	12.1	8	26
15	-1	1	1	1	-1	-1	9.5	15	30
16	1	1	1	1	-1	-1	15.8	1	28
17	-1	-1	-1	-1	1	-1	8.3	22	40
18	1	-1	-1	-1	1	-1	8	8	30
19	-1	1	-1	-1	1	-1	7.9	16	35
20	1	1	-1	-1	1	-1	10.7	7	35
21	-1	-1	1	-1	1	-1	7.2	25	32
22	1	-1	1	-1	1	-1	7.2	5	35
23	-1	1	1	-1	1	-1	7.9	17	36
24	1	1	1	-1	1	-1	10.2	8	32
25	-1	-1	-1	1	1	-1	10.3	10	20
26	1	-1	-1	1	1	-1	9.9	3	35
27	-1	1	-1	1	1	-1	7.4	22	35
28	1	1	-1	1	1	-1	10.5	6	28
29	-1	-1	1	1	1	-1	9.6	24	27
30	1	-1	1	1	1	-1	15.1	4	36
31	-1	1	1	1	1	-1	8.7	10	36
32	1	1	1	1	1	-1	12.1	5	35
33	-1	-1	-1	-1	-1	1	12.6	32	32
34	1	-1	-1	-1	-1	1	10.5	10	34
35	-1	1	-1	-1	-1	1	11.3	28	30
36	1	1	-1	-1	-1	1	10.6	18	24
37	-1	-1	1	-1	-1	1	8.1	22	30
38	1	-1	1	-1	-1	1	12.5	31	20

39	-1	1	1	-1	-1	1	11.1	17	32
40	1	1	1	-1	-1	1	12.9	16	25
41	-1	-1	-1	1	-1	1	14.6	38	20
42	1	-1	-1	1	-1	1	12.7	12	20
43	-1	1	-1	1	-1	1	10.8	34	22
44	1	1	-1	1	-1	1	17.1	19	35
45	-1	-1	1	1	-1	1	13.6	12	26
46	1	-1	1	1	-1	1	14.6	14	15
47	-1	1	1	1	-1	1	13.3	25	19
48	1	1	1	1	-1	1	14.4	16	24
49	-1	-1	-1	-1	1	1	11	31	22
50	1	-1	-1	-1	1	1	12.5	14	23
51	-1	1	-1	-1	1	1	8.9	23	22
52	1	1	-1	-1	1	1	13.1	23	18
53	-1	-1	1	-1	1	1	7.6	28	20
54	1	-1	1	-1	1	1	8.6	20	20
55	-1	1	1	-1	1	1	11.8	18	20
56	1	1	1	-1	1	1	12.4	11	36
57	-1	-1	-1	1	1	1	13.4	39	20
58	1	-1	-1	1	1	1	14.6	30	11
59	-1	1	-1	1	1	1	14.9	31	20
60	1	1	-1	1	1	1	11.8	6	35
61	-1	-1	1	1	1	1	15.6	33	16
62	1	-1	1	1	1	1	12.8	23	32
63	-1	1	1	1	1	1	13.5	31	20
64	1	1	1	1	1	1	15.8	11	20

Phase-2: Wool Washing and Carding Experiment

No.	A	B	C	D	E	F	No. of Defects
1	-1	-1	-1	-1	-1	-1	18
2	1	-1	-1	-1	-1	-1	19
3	-1	1	-1	-1	-1	-1	33
4	1	1	-1	-1	-1	-1	22
5	-1	-1	1	-1	-1	-1	16
6	1	-1	1	-1	-1	-1	17
7	-1	1	1	-1	-1	-1	28
8	1	1	1	-1	-1	-1	21
9	-1	-1	-1	1	-1	-1	17
10	1	-1	-1	1	-1	-1	21
11	-1	1	-1	1	-1	-1	34
12	1	1	-1	1	-1	-1	21
13	-1	-1	1	1	-1	-1	16
14	1	-1	1	1	-1	-1	25
15	-1	1	1	1	-1	-1	26
16	1	1	1	1	-1	-1	27
17	-1	-1	-1	-1	1	-1	14
18	1	-1	-1	-1	1	-1	16
19	-1	1	-1	-1	1	-1	24
20	1	1	-1	-1	1	-1	19
21	-1	-1	1	-1	1	-1	13
22	1	-1	1	-1	1	-1	14
23	-1	1	1	-1	1	-1	28
24	1	1	1	-1	1	-1	17
25	-1	-1	-1	1	1	-1	14
26	1	-1	-1	1	1	-1	18
27	-1	1	-1	1	1	-1	28
28	1	1	-1	1	1	-1	18
29	-1	-1	1	1	1	-1	13
30	1	-1	1	1	1	-1	21
31	-1	1	1	1	1	-1	27
32	1	1	1	1	1	-1	19
33	-1	-1	-1	-1	-1	1	18
34	1	-1	-1	-1	-1	1	18
35	-1	1	-1	-1	-1	1	26
36	1	1	-1	-1	-1	1	24
37	-1	-1	1	-1	-1	1	17
38	1	-1	1	-1	-1	1	17
39	-1	1	1	-1	-1	1	25
40	1	1	1	-1	-1	1	21
41	-1	-1	-1	1	-1	1	16
42	1	-1	-1	1	-1	1	21
43	-1	1	-1	1	-1	1	25

44	1	1	-1	1	-1	1	19
45	-1	-1	1	1	-1	1	18
46	1	-1	1	1	-1	1	22
47	-1	1	1	1	-1	1	29
48	1	1	1	1	-1	1	26
49	-1	-1	-1	-1	1	1	16
50	1	-1	-1	-1	1	1	21
51	-1	1	-1	-1	1	1	32
52	1	1	-1	-1	1	1	21
53	-1	-1	1	-1	1	1	17
54	1	-1	1	-1	1	1	19
55	-1	1	1	-1	1	1	26
56	1	1	1	-1	1	1	20
57	-1	-1	-1	1	1	1	17
58	1	-1	-1	1	1	1	18
59	-1	1	-1	1	1	1	35
60	1	1	-1	1	1	1	21
61	-1	-1	1	1	1	1	20
62	1	-1	1	1	1	1	22
63	-1	1	1	1	1	1	24
64	1	1	1	1	1	1	26

Phase-2: Metal Inert Gas Welding Experiment

No.	A	B	C	D	E	F	Bead Width (mm)
1	-1	-1	-1	-1	-1	-1	8.601
2	1	-1	-1	-1	-1	-1	8.361
3	-1	1	-1	-1	-1	-1	11.447
8	1	1	-1	-1	-1	-1	7.946
4	-1	-1	1	-1	-1	-1	8.935
9	1	-1	1	-1	-1	-1	7.917
13	-1	1	1	-1	-1	-1	11.828
23	1	1	1	-1	-1	-1	9.873
5	-1	-1	-1	1	-1	-1	8.551
10	1	-1	-1	1	-1	-1	7.775
14	-1	1	-1	1	-1	-1	11.759
24	1	1	-1	1	-1	-1	7.518
17	-1	-1	1	1	-1	-1	9.458
27	1	-1	1	1	-1	-1	8.948
33	-1	1	1	1	-1	-1	13.402
43	1	1	1	1	-1	-1	10.059
6	-1	-1	-1	-1	1	-1	9.057
11	1	-1	-1	-1	1	-1	8.029
15	-1	1	-1	-1	1	-1	13.136
25	1	1	-1	-1	1	-1	8.002
18	-1	-1	1	-1	1	-1	9.498
28	1	-1	1	-1	1	-1	8.298
34	-1	1	1	-1	1	-1	12.787
44	1	1	1	-1	1	-1	9.561
20	-1	-1	-1	1	1	-1	8.566
30	1	-1	-1	1	1	-1	7.559
36	-1	1	-1	1	1	-1	11.565
46	1	1	-1	1	1	-1	7.41
39	-1	-1	1	1	1	-1	8.885
49	1	-1	1	1	1	-1	8.022
53	-1	1	1	1	1	-1	12.457
58	1	1	1	1	1	-1	9.34
7	-1	-1	-1	-1	-1	1	9.268
12	1	-1	-1	-1	-1	1	7.784
16	-1	1	-1	-1	-1	1	12.024
26	1	1	-1	-1	-1	1	10.297

19	-1	-1	1	-1	-1	1	9.287
29	1	-1	1	-1	-1	1	8.727
35	-1	1	1	-1	-1	1	12.582
45	1	1	1	-1	-1	1	10.083
21	-1	-1	-1	1	-1	1	8.798
31	1	-1	-1	1	-1	1	8.235
37	-1	1	-1	1	-1	1	12.276
47	1	1	-1	1	-1	1	9.591
40	-1	-1	1	1	-1	1	8.904
50	1	-1	1	1	-1	1	8.238
54	-1	1	1	1	-1	1	12.325
59	1	1	1	1	-1	1	9.849
22	-1	-1	-1	-1	1	1	7.821
32	1	-1	-1	-1	1	1	7.678
38	-1	1	-1	-1	1	1	11.512
48	1	1	-1	-1	1	1	7.803
41	-1	-1	1	-1	1	1	8.726
51	1	-1	1	-1	1	1	7.435
55	-1	1	1	-1	1	1	11.822
60	1	1	1	-1	1	1	8.965
42	-1	-1	-1	1	1	1	8.258
52	1	-1	-1	1	1	1	8.216
56	-1	1	-1	1	1	1	11.789
61	1	1	-1	1	1	1	10.21
57	-1	-1	1	1	1	1	8.885
62	1	-1	1	1	1	1	7.67
63	-1	1	1	1	1	1	12.216
64	1	1	1	1	1	1	10.225

Phase-2: Radial Contour Turning Experiment

No.	A	B	C	D	E	Tracking Error (Micron)
1	-1	-1	-1	-1	-1	24.7721
2	1	-1	-1	-1	-1	24.2379
3	-1	1	-1	-1	-1	52.7276
4	1	1	-1	-1	-1	60.8756
5	-1	-1	1	-1	-1	22.253
6	1	-1	1	-1	-1	20.2671
7	-1	1	1	-1	-1	48.9596
8	1	1	1	-1	-1	57.7356
9	-1	-1	-1	1	-1	33.5641
10	1	-1	-1	1	-1	33.0299
11	-1	1	-1	1	-1	58.5898
12	1	1	-1	1	-1	67.7836
13	-1	-1	1	1	-1	26.649
14	1	-1	1	1	-1	24.2379
15	-1	1	1	1	-1	52.7276
16	1	1	1	1	-1	62.1316
17	-1	-1	-1	-1	1	11.1688
18	1	-1	-1	-1	1	14.4899
19	-1	1	-1	-1	1	39.1894
20	1	1	-1	-1	1	45.8036
21	-1	-1	1	-1	1	11.7123
22	1	-1	1	-1	1	15.1179
23	-1	1	1	-1	1	39.8174
24	1	1	1	-1	1	46.4316
25	-1	-1	-1	1	1	11.2389
26	1	-1	-1	1	1	13.6651
27	-1	1	-1	1	1	39.5396
28	1	1	-1	1	1	45.8036
29	-1	-1	1	1	1	11.2389
30	1	-1	1	1	1	14.4899
31	-1	1	1	1	1	39.5396
32	1	1	1	1	1	45.8036

Phase-3: EDM Experiment

No.	A	B	C	D	Work Peace MRR (mm^3/min)
1	1	1	1	1	9.029
2	2	1	1	1	13.544
3	3	1	1	1	18.059
4	1	2	1	1	13.544
5	2	2	1	1	15.801
6	3	2	1	1	22.573
7	1	3	1	1	15.801
8	2	3	1	1	18.059
9	3	3	1	1	27.088
10	1	1	2	1	11.287
11	2	1	2	1	11.287
12	3	1	2	1	15.801
13	1	2	2	1	13.544
14	2	2	2	1	20.316
15	3	2	2	1	20.316
16	1	3	2	1	15.801
17	2	3	2	1	22.573
18	3	3	2	1	22.573
19	1	1	3	1	13.544
20	2	1	3	1	11.287
21	3	1	3	1	11.287
22	1	2	3	1	13.544
23	2	2	3	1	20.316
24	3	2	3	1	18.059
25	1	3	3	1	15.801
26	2	3	3	1	15.801
27	3	3	3	1	22.573
28	1	1	1	2	13.544
29	2	1	1	2	13.544
30	3	1	1	2	15.801
31	1	2	1	2	11.287
32	2	2	1	2	13.544
33	3	2	1	2	18.059
34	1	3	1	2	13.544
35	2	3	1	2	15.801
36	3	3	1	2	22.573

37	1	1	2	2	11.287
38	2	1	2	2	11.287
39	3	1	2	2	15.801
40	1	2	2	2	13.544
41	2	2	2	2	15.801
42	3	2	2	2	18.059
43	1	3	2	2	13.544
44	2	3	2	2	15.801
45	3	3	2	2	22.573
46	1	1	3	2	11.287
47	2	1	3	2	11.287
48	3	1	3	2	15.801
49	1	2	3	2	11.287
50	2	2	3	2	13.544
51	3	2	3	2	18.059
52	1	3	3	2	13.544
53	2	3	3	2	15.801
54	3	3	3	2	22.573
55	1	1	1	3	11.287
56	2	1	1	3	11.287
57	3	1	1	3	15.801
58	1	2	1	3	13.544
59	2	2	1	3	15.801
60	3	2	1	3	18.059
61	1	3	1	3	13.544
62	2	3	1	3	15.801
63	3	3	1	3	22.573
64	1	1	2	3	11.287
65	2	1	2	3	11.287
66	3	1	2	3	31.603
67	1	2	2	3	13.544
68	2	2	2	3	15.801
69	3	2	2	3	15.801
70	1	3	2	3	13.544
71	2	3	2	3	15.801
72	3	3	2	3	22.573
73	1	1	3	3	11.287
74	2	1	3	3	11.287
75	3	1	3	3	15.801
76	1	2	3	3	11.287

77	2	2	3	3	13.544
78	3	2	3	3	20.316
79	1	3	3	3	13.544
80	2	3	3	3	20.316
81	3	3	3	3	22.573

Phase-4: EDM Experiment

No.	A	B	C	D	Tool MRR (mm^3/min)
1	1	1	1	1	7.813
2	2	1	1	1	7.813
3	3	1	1	1	4.464
4	1	2	1	1	5.580
5	2	2	1	1	6.696
6	3	2	1	1	7.813
7	1	3	1	1	6.696
8	2	3	1	1	8.929
9	3	3	1	1	11.161
10	1	1	2	1	7.813
11	2	1	2	1	6.696
12	3	1	2	1	6.696
13	1	2	2	1	5.580
14	2	2	2	1	6.696
15	3	2	2	1	7.813
16	1	3	2	1	6.696
17	2	3	2	1	8.929
18	3	3	2	1	12.277
19	1	1	3	1	1.116
20	2	1	3	1	6.696
21	3	1	3	1	6.696
22	1	2	3	1	5.580
23	2	2	3	1	6.696
24	3	2	3	1	7.813
25	1	3	3	1	6.696
26	2	3	3	1	10.045
27	3	3	3	1	12.277
28	1	1	1	2	4.464
29	2	1	1	2	7.813
30	3	1	1	2	5.580
31	1	2	1	2	5.580
32	2	2	1	2	6.696
33	3	2	1	2	7.813
34	1	3	1	2	5.580
35	2	3	1	2	10.045
36	3	3	1	2	12.277

37	1	1	2	2	7.813
38	2	1	2	2	6.696
39	3	1	2	2	6.696
40	1	2	2	2	5.580
41	2	2	2	2	6.696
42	3	2	2	2	7.813
43	1	3	2	2	6.696
44	2	3	2	2	10.045
45	3	3	2	2	12.277
46	1	1	3	2	12.277
47	2	1	3	2	6.696
48	3	1	3	2	6.696
49	1	2	3	2	4.464
50	2	2	3	2	5.580
51	3	2	3	2	7.813
52	1	3	3	2	3.348
53	2	3	3	2	8.929
54	3	3	3	2	12.277
55	1	1	1	3	7.813
56	2	1	1	3	7.813
57	3	1	1	3	7.813
58	1	2	1	3	5.580
59	2	2	1	3	6.696
60	3	2	1	3	7.813
61	1	3	1	3	3.348
62	2	3	1	3	8.929
63	3	3	1	3	11.161
64	1	1	2	3	7.813
65	2	1	2	3	6.696
66	3	1	2	3	6.696
67	1	2	2	3	5.580
68	2	2	2	3	6.696
69	3	2	2	3	7.813
70	1	3	2	3	6.696
71	2	3	2	3	8.929
72	3	3	2	3	12.277
73	1	1	3	3	12.277
74	2	1	3	3	6.696
75	3	1	3	3	6.696
76	1	2	3	3	5.580

77	2	2	3	3	6.696
78	3	2	3	3	7.813
79	1	3	3	3	8.929
80	2	3	3	3	10.045
81	3	3	3	3	12.277

Phase-4: Micro-End-Milling Experiment

No.	A	B	C	D	Surface Roughness (µm)
1	1	1	1	1	0.26
2	2	1	1	1	0.31
3	3	1	1	1	0.27
4	1	2	1	1	0.32
5	2	2	1	1	0.36
6	3	2	1	1	0.85
7	1	3	1	1	0.48
8	2	3	1	1	0.37
9	3	3	1	1	1.58
10	1	1	2	1	0.36
11	2	1	2	1	0.56
12	3	1	2	1	0.52
13	1	2	2	1	0.51
14	2	2	2	1	0.53
15	3	2	2	1	0.81
16	1	3	2	1	0.53
17	2	3	2	1	0.47
18	3	3	2	1	0.93
19	1	1	3	1	0.49
20	2	1	3	1	0.5
21	3	1	3	1	1.22
22	1	2	3	1	0.42
23	2	2	3	1	0.58
24	3	2	3	1	1.31
25	1	3	3	1	0.67
26	2	3	3	1	0.47
27	3	3	3	1	0.98
28	1	1	1	2	0.37
29	2	1	1	2	0.3
30	3	1	1	2	0.37
31	1	2	1	2	0.53
32	2	2	1	2	0.47
33	3	2	1	2	0.64
34	1	3	1	2	0.65
35	2	3	1	2	0.52
36	3	3	1	2	1.15

37	1	1	2	2	0.49
38	2	1	2	2	0.89
39	3	1	2	2	0.71
40	1	2	2	2	0.8
41	2	2	2	2	0.66
42	3	2	2	2	0.76
43	1	3	2	2	0.58
44	2	3	2	2	0.32
45	3	3	2	2	0.77
46	1	1	3	2	0.57
47	2	1	3	2	1.22
48	3	1	3	2	0.83
49	1	2	3	2	0.69
50	2	2	3	2	0.91
51	3	2	3	2	0.9
52	1	3	3	2	0.66
53	2	3	3	2	0.73
54	3	3	3	2	0.82
55	1	1	1	3	0.42
56	2	1	1	3	0.2
57	3	1	1	3	0.57
58	1	2	1	3	0.47
59	2	2	1	3	0.4
60	3	2	1	3	0.51
61	1	3	1	3	0.48
62	2	3	1	3	0.59
63	3	3	1	3	0.47
64	1	1	2	3	0.66
65	2	1	2	3	0.78
66	3	1	2	3	1.27
67	1	2	2	3	0.5
68	2	2	2	3	0.58
69	3	2	2	3	0.66
70	1	3	2	3	0.62
71	2	3	2	3	0.6
72	3	3	2	3	0.62
73	1	1	3	3	0.58
74	2	1	3	3	0.7
75	3	1	3	3	0.71
76	1	2	3	3	0.68

77	2	2	3	3	0.72
78	3	2	3	3	1.05
79	1	3	3	3	0.85
80	2	3	3	3	0.61
81	3	3	3	3	0.62

APPENDIX 8:
RESULTS OF VALIDATION EXPERIMENTS ANALYSIS

Phase-1: Dyestuff Experiment – Strength Response Analysis Results

	Full Factorial	L16						L12						L8					
		HNP	LM	MLM	PU	PD	UC	HNP	LM	MLM	PU	PD	UC	HNP	LM	MLM	PU	PD	UC
A	*	*			*	*													
B					*	*													
C					*														
D	*	*		*	*	*	*	*			*	*	*	*			*	*	
E					*														
F	*	*			*	*	*	*		*	*	*	*						
AB					*														
AC					*														
AD					*														
AF					*	*													
BD					*														
BF					*														
ABD		*			*	*													
ABF					*														
Power		**1**	**0**	**0.33**	**1**	**1**	**0.67**	**0.67**	**0**	**0.33**	**0.67**	**0.67**	**0.67**	**0.33**	**0**	**0**	**0.33**	**0.33**	**0**

Phase-1: Dyestuff Experiment – Hue Response Analysis Results

	Full Factorial	L16						L12						L8					
		HNP	**LM**	**MLM**	**PU**	**PD**	**UC**	**HNP**	**LM**	**MLM**	**PU**	**PD**	**UC**	**HNP**	**LM**	**MLM**	**PU**	**PD**	**UC**
A	*	*	*	*	*	*	*	*			*	*	*	*			*	*	
B		*			*	*	*				*	*							
C					*														
D					*														
E																			
F	*	*	*	*	*	*	*	*			*	*	*						
AB					*														
AC					*														
AD		*			*	*													
AE					*														
AF					*														
BD					*														
BF		*			*	*													
ABF					*														
Power		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**0**	**0**	**1**	**1**	**1**	**0.5**	**0**	**0**	**0.5**	**0.5**	**0**

Phase-1: Dyestuff Experiment – Brightness Response Analysis Results

	Full Factorial	L16						L12						L8					
		HNP	**LM**	**MLM**	**PU**	**PD**	**UC**	**HNP**	**LM**	**MLM**	**PU**	**PD**	**UC**	**HNP**	**LM**	**MLM**	**PU**	**PD**	**UC**
A																			
B								*			*	*							
C								*			*	*	*						
D											*	*		*			*	*	
E								*			*	*	*	*			*	*	
F	*	*	*	*	*	*	*	*			*	*	*	*			*	*	
Power		**1**	**1**	**1**	**1**	**1**	**1**	**1**	**0**	**0**	**1**	**1**	**1**	**1**	**0**	**0**	**1**	**1**	**0**

Phase-2: Wool Washing and Carding Experiment – Number of Defects Analysis Results

	Full Factorial	L16	L12		L8			
					DO		BM	
		HNP	BSS	SWR	BSS	SWR	BSS	SWR
A	*						*	*
B	*	*	*	*	*	*	*	*
C								
D								
E								
F						*		
AB	*	*	*	*				
AC			*	*				
AD					*			
BC			*	*	*	*		
BE						*		*
CD					*		*	
CE				*		*		*
DE							*	*
EF	*							
Power		**0.5**	**0.5**	**0.5**	**0.25**	**0.25**	**0.5**	**0.5**

Phase-2: Metal Inert Gas Welding Experiment – Bead Width Analysis Results

	Full Factorial	L16	L12		L8			
					DO		BM	
		HNP	BSS	SWR	BSS	SWR	BSS	SWR
A	*	*	*	*		*		
B	*	*	*	*			*	*
C	*	*						
D		*				*		
E		*				*		
F						*		
AB	*	*	*	*	*		*	*
AD		*			*			
AE		*				*		
AF		*						
BE					*	*		
BD			*		*			
CE						*	*	*
CD							*	
CF						*		
ABD		*						
ABF		*						
Power		**1**	**0.75**	**0.75**	**0.25**	**0.25**	**0.5**	**0.5**

Phase-2: Radial Contour Turning Experiment – Tracking Errors Analysis Results

	Full Factorial	L16	L12		L8			
					DO		BM	
		HNP	BSS	SWR	BSS	SWR	BSS	SWR
A	*	*	*	*	*	*	*	*
B	*	*	*	*	*	*	*	*
C	*	*	*	*			*	
D	*	*	*				*	*
E	*	*	*	*	*	*	*	*
F								
AB	*	*	*	*	*	*	*	*
AC				*				
AD								
AE								
AF								
BC				*	*			
BD			*					
BE				*			*	
CE	*	*		*	*			
DE	*	*	*				*	*
Power		**1**	**0.88**	**0.75**	**0.63**	**0.5**	**0.88**	**0.75**

Phase-3: EDM Experiment – Workpiece MRR Analysis Results

No.	Full Factorial	L18				L9			
		Standard ANOVA		Regression ANOVA		Standard ANOVA		Regression ANOVA	
		PU	PD	PU	PD	PU	PD	PU	PD
Al	*	*	*	*	*			*	*
Bl	*			*	*			*	*
Cl								*	
Dl								*	
Aq	*	*	*	*					
Bq								*	
Cq				*	*				
Dq				*				*	
Eq				*	*				
Fq				*	*				
Power		**0.667**	**0.667**	**1**	**0.667**	**0**	**0**	**0.667**	**0.667**

Phase-4: EDM Experiment – Tool MRR Analysis Results

No.	Full Factorial	L18		L9	
		BSS	SWR	BSS	SWR
Al	*	*	*		
Bl	*	*	*		
Cl					
Dl					
Aq					
Bq	*	*	*	*	*
Cq					
Dq					
AlBl	*	*	*	*	*
BlCq					*
CqDl				*	*
BlCl				*	*
Power		**1**	**1**	**0.5**	**0.5**

Phase-4: Micro-End-Milling Experiment – Surface Roughness Analysis Results

No.	Full Factorial	L18		L9	
		BSS	SWR	BSS	SWR
Al	*	*	*	*	
Bl				*	
Cl	*	*		*	
Dl					*
Aq	*			*	
Bq			*		
AqBl	*				
AlDl	*	*	*		
BlCq	*	*		*	*
BlCl	*			*	
BlDl	*	*	*		*
BlDq					*
AlDq			*	*	*
CqDl		*	*		*
AqCq					*
AlCq					*
AlBq			*		
ClDl		*	*		
CqDq			*		
Power		**0.63**	**0.38**	**0.63**	**0.25**

APPENDIX 9:
DETAILED ANALYSIS OF SELECTED VALIDATION EXPERIMENTS

1. Phase-1: Analysis of the Dyestuff Hue Data

1-1 Analysing the Full Factorial Dyestuff Hue Data

The ANOVA for the full factorial dyestuff hue data is shown in Table 1-1. Factors A and F were the only significant effects at $\alpha = 0.05$ since their p-values were less than this.

Table 1-1: ANOVA of the Full Factorial Dyestuff Hue Data

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value
A	2036.266	1	2036.266	61.02323	< 0.0001
B	118.2656	1	118.2656	3.544209	0.0649
C	58.14063	1	58.14063	1.74237	0.1921
D	1.890625	1	1.890625	0.056659	0.8127
E	28.89062	1	28.89062	0.8658	0.3560
F	1881.391	1	1881.391	56.38191	< 0.0001
Residual	1902.016	57	33.3687		
Total	6026.859	63			

1-2 Extracting the Fractional Factorial Dyestuff Hue Data

The extracted dyestuff hue full factorial data that corresponded to the L_{16}, L_{12} and L_8 OAs are presented in Table 1-2.

1-3 HNP Plot Analysis

The HNP plots for the dyestuff hue L_{16}, L_{12} and L_8 data are shown in Figure 1-1. From the L_{16} plot the active effects appeared to be F, A, B and the BF and AD interactions. In the case of the L_{12}, only A, F and B were distinguishable from noise while for the L_8 factors A and F were flagged as potentially active. The statistical significance of the selected effects was tested at the 5% level using the ANOVA as shown in Table 1-3. In the case of the L_{16} the ANOVA confirmed the significance of all the selected effects at the 5% level meaning that factor B and the BF and AD interactions were falsely identified as significant. The L_{12} ANOVA correctly detected A and F as significant but not B as its p-value was larger than

Table 1-2: Extracted (a) L_{16}, (b) L_{12} and (a) L_8 Dyestuff Hue Data

Runs	A	B	C	D	E	F	Response
1	-1	-1	-1	-1	-1	-1	15
2	1	-1	-1	-1	1	-1	8
3	-1	1	-1	-1	1	1	23
4	1	1	-1	-1	-1	1	18
5	-1	-1	1	-1	1	1	28
6	1	-1	1	-1	-1	1	31
7	-1	1	1	-1	-1	-1	13
8	1	1	1	-1	1	-1	8
9	-1	-1	-1	1	-1	1	38
10	1	-1	-1	1	1	1	30
11	-1	1	-1	1	1	-1	22
12	1	1	-1	1	-1	-1	5
13	-1	-1	1	1	1	-1	24
14	1	-1	1	1	-1	-1	8
15	-1	1	1	1	-1	1	25
16	1	1	1	1	1	1	11

(a)

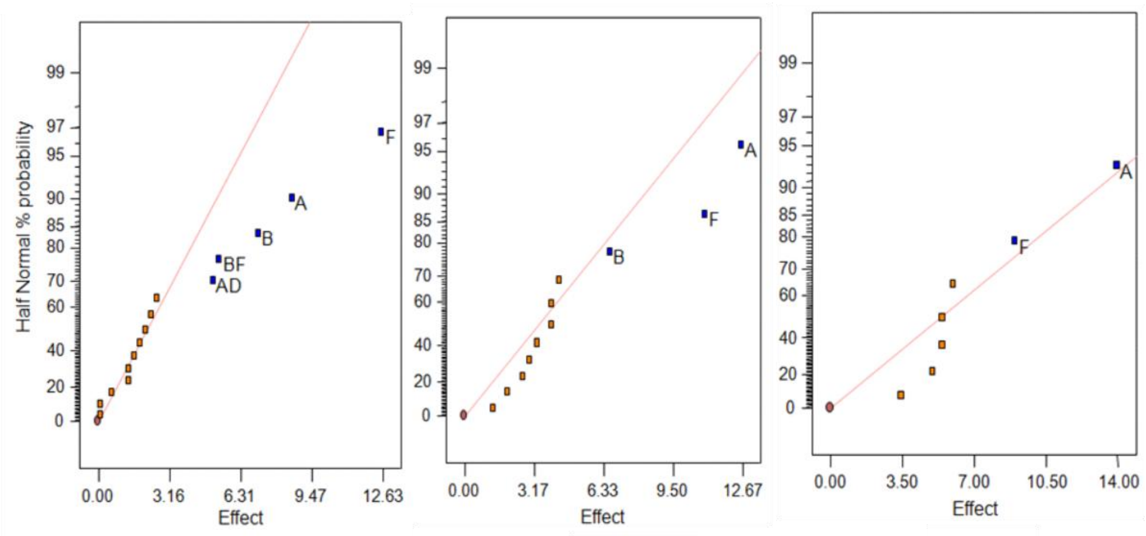
Runs	A	B	C	D	E	F	Response
1	1	1	1	1	1	1	11
2	1	1	-1	1	-1	1	19
3	1	1	-1	-1	1	-1	7
4	-1	1	1	-1	-1	1	17
5	1	-1	-1	1	-1	-1	1
6	-1	1	-1	-1	1	-1	16
7	-1	-1	-1	-1	-1	1	32
8	1	-1	1	-1	-1	-1	9
9	-1	1	1	1	-1	-1	15
10	-1	-1	1	1	1	-1	24
11	-1	-1	-1	1	1	1	39
12	1	-1	1	-1	1	1	20

(b)

Runs	A	B	C	D	E	F	Response
1	-1	-1	-1	1	1	1	39
2	1	-1	-1	-1	-1	1	10
3	-1	1	-1	-1	1	-1	16
4	1	1	-1	1	-1	-1	5
5	-1	-1	1	1	-1	-1	15
6	1	-1	1	-1	1	-1	5
7	-1	1	1	-1	-1	1	17
8	1	1	1	1	1	1	11

(c)

0.05. The ANOVA of the L_8 only declared A as significant its p-value being smaller than 0.05.



(a) (b) (c)

Figure 1-1: HNPs of the Dyestuff Hue (a) L_{16}, (b) L_{12} and (c) L_8 Data

Table 1-3: ANOVA for the HNP analysis of (a) L_{16}, (b) L_{12} and (c) L_8 Dyestuff Hue Data

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value
A	297.6	1.0	297.6	27.1	0.0004
B	203.1	1.0	203.1	18.5	0.0016
F	637.6	1.0	637.6	58.2	< 0.0001
AD	105.1	1.0	105.1	9.6	0.0113
BF	115.6	1.0	115.6	10.5	0.0088
Residual	109.6	10.0	11.0		
Total	1468.4	15.0			
(a)					

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value
A	481.3	1.0	481.3	15.3	0.0045
B	133.3	1.0	133.3	4.2	0.0733
F	363.0	1.0	363.0	11.6	0.0094
Residual	251.3	8.0	31.4		
Total	1229.0	11.0			
(b)					

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value
A	392	1	392	7.33	0.0424
F	162	1	162	3.03	0.1423
Residual	267.5	5	53.5		
Total	821.5	7			
(c)					

1-4 LM and MLM Analysis

The results of implementing LM and the MLM are presented by means of Pareto charts of the examined effects' t-lenth statistics. For each of the extracted OA data, the absolute values of the effect t-lenths were ranked and displayed in descending order as shown in Figure 1-2. In each chart, two reference lines were drawn representing the critical t-lenth of both LM and MLM at α =0.05. Any effect's t-lenth that extended past any of the lines was declared significant by the appropriate method. Implementing this rule on the Pareto charts displayed in Figure 1-2, it is clear that while the main effect B was falsely pronounced significant by the MLM method in the L_{16} data analysis, A and F were correctly declared as active by both LM and MLM. As none of the effect t-lenth values extended beyond any of the critical t-lenths, no effects were found to be distinguishable from noise when the L_{12} and L_8 data were analysed using both methods.

1-5 PU Method Analysis

The results of implementing the PU procedure in analysing the dyestuff hue L_{16} data are presented in Table 1-4. The upper part of the table shows the pooled effects in the sequence of their selection along with the p-values at which each effect was pooled. Since the main effect E was the smallest, it was the first to be pooled to estimate the error. The next smallest effect was the ABD interaction, its p-value when E was used as an error estimate was 0.5. As this was greater than α (=0.1), ABD was pooled with E to estimate the error variance. The next smallest effect was the AC interaction. When its significance was tested using this new error estimate, it was pronounced significant as its p-value (0.0377) was smaller than 0.1. Consequently, the AC interaction and all the remaining effects were declared significant at α =0.1 as shown in the lower part of Table 1-4. Clearly, while the main effects A and F were correctly judged to be significant, many spurious effects were falsely deemed important.

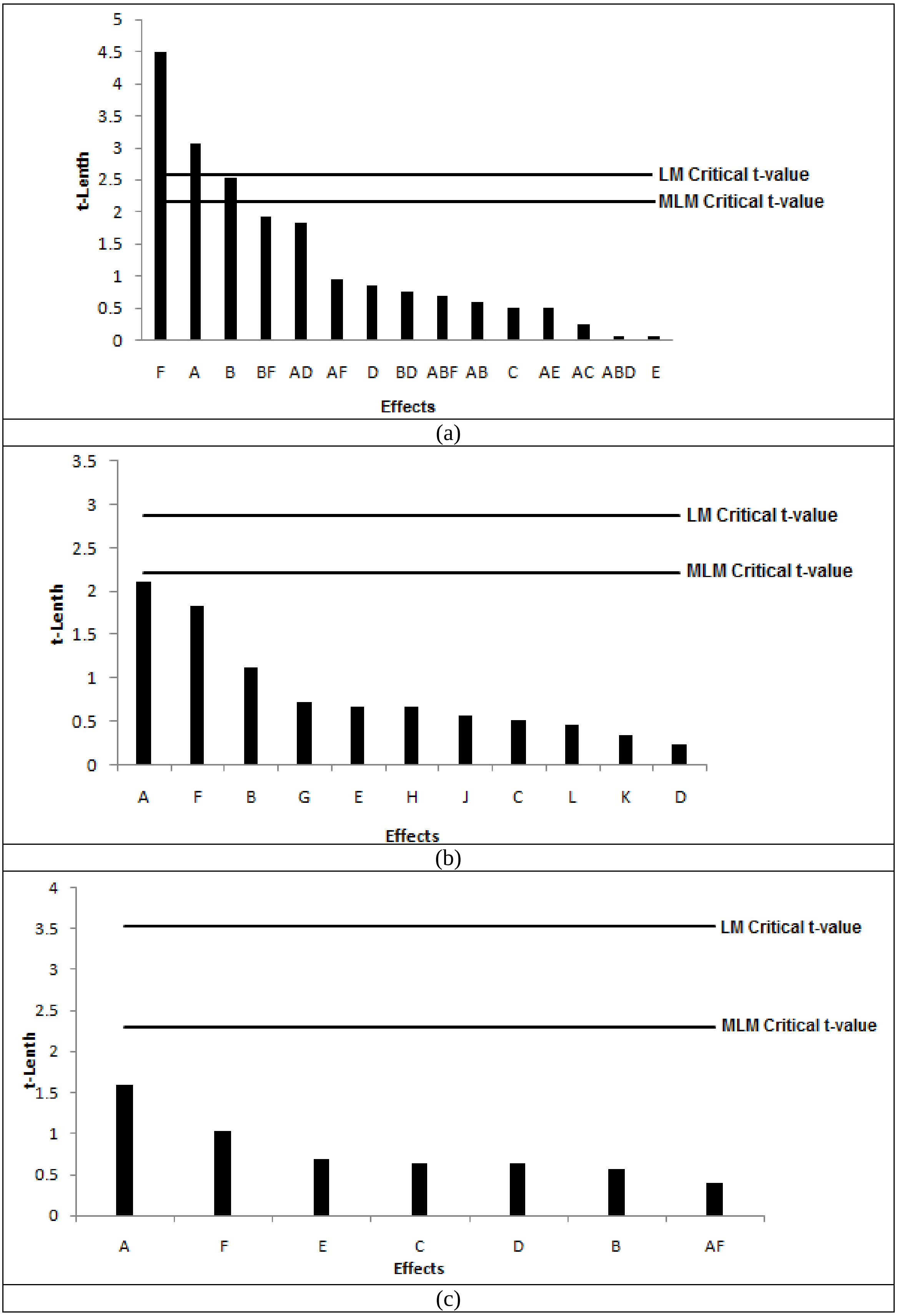


Figure 1-2: Pareto Chart of the Effect t-Lenths for the Dyestuff Hue (a) L_{16}, (b) L_{12} and (c) L_8 Data

Table 1-4: The PU Analysis of the Dyestuff Hue L_{16} Data

	Steps	Pooled Effects	p-value		
	1	E			
	2	ABD	0.5		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	297.5625	1	297.5625	4761	0.0002
B	203.0625	1	203.0625	3249	0.0003
C	7.5625	1	7.5625	121	0.0082
D	22.5625	1	22.5625	361	0.0028
F	637.5625	1	637.5625	10201	< 0.0001
AB	10.5625	1	10.5625	169	0.0059
AC	1.5625	1	1.5625	25	0.0377
AD	105.0625	1	105.0625	1681	0.0006
AE	7.5625	1	7.5625	121	0.0082
AF	27.5625	1	27.5625	441	0.0023
BD	18.0625	1	18.0625	289	0.0034
BF	115.5625	1	115.5625	1849	0.0005
ABF	14.0625	1	14.0625	225	0.0044
Residual	0.125	2	0.0625		
Total	1468.438	15			

The results of implementing the PU method in analysing the L_{12} and L_8 data are given in Tables 1-5 and 1-6. In the case of the L_{12} the true active effects A and F were declared significant at $\alpha = 0.1$ along with the inert effect B. However, only A was identified as active at the designated level of significance when the L_8 data were analysed using this method.

1-6 PD Method Analysis

Table 1-7 shows the results of applying the PD method to analyse the dyestuff hue L_{16} data. The upper part of the table illustrates the picked-up effects in the order of their selection. In the first step, factor F was selected as the largest effect and its significance was tested using an error estimate obtained by pooling all the remaining effects including the active ones. As its p-value (0.0055) was smaller than 0.1, factor F was declared significant. Factor A was then picked-up as the next largest effect. An estimate of the error was then rendered by pooling all

Table 1-5: The PU Analysis of the Dyestuff Hue L_{12} Data

	Steps	**Pooled Effects**	**p-value**		
	1	D			
	2	K	0.3743		
	3	L	0.2572		
	4	C	0.2436		
	5	J	0.2273		
	6	E	0.1802		
	7	H	0.2111		
	8	G	0.198		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	481.3333	1	481.3333	15.32095	0.0045
B	133.3333	1	133.3333	4.244032	0.0733
F	363	1	363	11.55438	0.0094
Residual	251.3333	8	31.41667		
Total	1229	11			

Table 1-6: The PU Analysis of the Dyestuff Hue L_8 Data

	Steps	**Pooled Effects**	**p-value**		
	1	AF			
	2	B	0.3888		
	3	C	0.3306		
	4	D	0.3302		
	5	E	0.2916		
	6	F	0.1423		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	392	1	392	5.476135	0.0578
Residual	429.5	6	71.58333		
Total	821.5	7			

the studied effects except F and A. This was used to test the significance of A which was found to be active with a p-value (0.0184) smaller than 0.1. By the same token, the effects B, BF and AD were selected in the third, fourth and the fifth steps of the PD method. Among the remaining effects, the largest was AF. In the sixth step, its significance was tested using an error estimate formed by pooling all the effects smaller than AF. As the resultant p-value was 0.1161 which is greater than 0.1, AF was not selected and the PD procedure was terminated at the fifth step pronouncing the effects A, B, F, AD and BF as the only significant ones. Consequently, the PD method successfully detected the true active effects along with three spurious ones.

Table 1-7: The PD Analysis of the Dyestuff Hue L_{16} Data

	Steps	**Selected Effects**	**p-value**		
	1	F	0.0055		
	2	A	0.0184		
	3	B	0.0187		
	4	BF	0.0332		
	5	AD	0.0113		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	297.5625	1	297.5625	27.14367	0.0004
B	203.0625	1	203.0625	18.52338	0.0016
F	637.5625	1	637.5625	58.15849	< 0.0001
AD	105.0625	1	105.0625	9.583808	0.0113
BF	115.5625	1	115.5625	10.54162	0.0088
Residual	109.625	10	10.9625		
Total	1468.438	15			

When applied to analyse the L_{12} data, the PD method correctly identified factors A and F as significant as shown in Table 1-8. Also detected as active was factor B its p-value being smaller than 0.1. With regard to the L_8 data, only factor A was highlighted as active at $\alpha = 0.1$ by the PD method (see Table 1-9).

Table 1-8: The PD Analysis of the Dyestuff Hue L_{12} Data

	Steps	Selected Effects	p-value		
	1	A	0.0295		
	2	F	0.0172		
	3	B	0.0733		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	481.3333	1	481.3333	15.32095	0.0045
B	133.3333	1	133.3333	4.244032	0.0733
F	363	1	363	11.55438	0.0094
Residual	251.3333	8	31.41667		
Total	1229	11			

Table 1-9: The PD Analysis of the Dyestuff Hue L_8 Data

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value
A	392	1	392	5.476135	0.0578
Residual	429.5	6	71.58333		
Total	821.5	7			

1-7 UC method Analysis

For implementing this method, all the columns that were not assigned to any of the six factors in the dyestuff experiments were pooled to estimate the error variance. The ANOVAs using this strategy to analyse the L_{16} L_{12} and the L_8 are shown in Table 1-10. Factors A, F and B were found significant at the 5% level in the L_{16} case. Thus, despite the detection of the true active effects, one inert effect was falsely declared significant. In the case of the L_{12}, only the true active effects A and F were pronounced significant. No effect was found significant in the L_8 analysis.

Table 1-10: ANOVA for the UC analysis of (a) L16, (b) L12 and (c) L8 Dyestuff Hue Data

Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	297.6	1	297.6	8.9	0.0153
B	203.1	1	203.1	6.1	0.0357
C	7.6	1	7.6	0.2	0.6452
D	22.6	1	22.6	0.7	0.4320
E	0.1	1	0.1	0.0	0.9664
F	637.6	1	637.6	19.1	0.0018
Residual	300.1	9	33.3		
Total	1468.4	15			

(a)

Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	481.3	1	481.3	14.1	0.0133
B	133.3	1	133.3	3.9	0.1053
C	27.0	1	27.0	0.8	0.4150
D	5.3	1	5.3	0.2	0.7092
E	48.0	1	48.0	1.4	0.2894
F	363.0	1	363.0	10.6	0.0225
Residual	171.0	5	34.2		
Total	1229.0	11			

(b)

Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	392	1	392	16	0.1560
B	50	1	50	2.0	0.3888
C	60.5	1	60.5	2.5	0.3608
D	60.5	1	60.5	2.5	0.3608
E	72	1	72	2.9	0.3362
F	162	1	162	6.6	0.2361
Residual	24.5	1	24.5		
Total	821.5	7			

(c)

2. Phase-2: Analysis of the Metal Inert Gas Welding Experiment

2-1 Analysing the Full Factorial Bead Width Data

The full factorial bead width data was analysed using the ANOVA as shown in Table 2-1. At $\alpha = 0.05$, the only active effects were A, B, C and AB.

Table 2-1: ANOVA of the Full Factorial Bead Width Data

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value
A	57.7619	1	57.7619	183.4729	< 0.0001
B	79.3636	1	79.3636	252.0877	< 0.0001
C	5.269894	1	5.269894	16.73911	0.0001
D	0.054581	1	0.054581	0.173368	0.6788
E	1.334891	1	1.334891	4.240102	0.0442
F	0.135884	1	0.135884	0.431618	0.5139
AB	19.79026	1	19.79026	62.8611	< 0.0001
Residual	17.31539	55	0.314825		
Total	181.0264	62			

2-2 Extracting the Fractional Factorial Bead Width Data

The extracted L_{16} and L_{12} data are displayed in Table 2-2. For the L_8 design augmented using the DO method and the BM, the extracted bead width data are given in Table 2-3.

2-3 Analysis of the L_{16} Data Using HNP Plot

The HNP plot of the extracted L_{16} bead width data is depicted in Figure 2-1. Clearly the effects A, B, C, D, E, AB, AD, AE, AF, ABD and ABF are distinguishable from noise. Their statistical significance at $\alpha = 0.05$ was confirmed using the ANOVA as shown in Table 2-4. All the true active effects were therefore correctly detected as active. However, the reduction in the experiment's size resulted in overestimating many of the inert effects. Some of these

Table 2-2: Extracted (a) L_{16}, and (b) L_{12} Bead Width Data

Runs	A	B	C	D	E	F	Response
1	-1	-1	-1	-1	-1	-1	8.601
2	1	-1	-1	-1	1	-1	8.029
3	-1	1	-1	-1	1	1	7.803
4	1	1	-1	-1	-1	1	10.297
5	-1	-1	1	-1	1	1	8.726
6	1	-1	1	-1	-1	1	8.727
7	-1	1	1	-1	-1	-1	11.828
8	1	1	1	-1	1	-1	9.561
9	-1	-1	-1	1	-1	1	8.798
10	1	-1	-1	1	1	1	8.216
11	-1	1	-1	1	1	-1	11.565
12	1	1	-1	1	-1	-1	7.518
13	-1	-1	1	1	1	-1	8.885
14	1	-1	1	1	-1	-1	8.948
15	-1	1	1	1	-1	1	12.325
16	1	1	1	1	1	1	10.225

(a)

Runs	A	B	C	D	E	F	Response
1	1	1	1	1	1	1	10.225
2	1	1	-1	1	-1	1	9.591
3	1	1	-1	-1	1	-1	8.002
4	-1	1	1	-1	-1	1	12.582
5	1	-1	-1	1	-1	-1	7.775
6	-1	1	-1	-1	1	-1	13.136
7	-1	-1	-1	-1	-1	1	9.268
8	1	-1	1	-1	-1	-1	7.917
9	-1	1	1	1	-1	-1	13.402
10	-1	-1	1	1	1	-1	8.885
11	-1	-1	-1	1	1	1	8.258
12	1	-1	1	-1	1	1	7.435

(b)

were found to be larger than the true active effects. For example, while the ABF, AF, ABD and AD interaction effects were smaller than the main effect A in the full factorial analysis, their sizes, as can be seen in Figure 2-1, surpassed the latter in the extracted L_{16} data. Moreover, contrary to what was observed in the full factorial analysis, these effects along with factor F and the AE interaction became larger than the AB interaction effect in the extracted L_{16} analysis. Despite this, the use of the HNP plot enabled these biased effects to be isolated from the error estimate thereby allowing the true active effects to be detected. Of course this came at the expense of falsely declaring 7 effects as active. However, as explained in Chapter 6, this is far less serious than overlooking the true active effects.

Table 2-3: Extracted (a) DO and (b) BM Augmented L_8 Bead Width Data

Runs	**A**	**B**	**C**	**D**	**E**	**F**	**Response**
1	-1	-1	-1	1	1	1	8.258
2	1	-1	-1	-1	-1	1	7.784
3	-1	1	-1	-1	1	-1	13.136
4	1	1	-1	1	-1	-1	7.518
5	-1	-1	1	1	-1	-1	9.458
6	1	-1	1	-1	1	-1	8.298
7	-1	1	1	-1	-1	1	12.582
8	1	1	1	1	1	1	10.225
9	1	1	-1	1	1	1	10.21
10	-1	-1	1	1	1	1	8.885
11	1	1	1	-1	1	-1	9.561
12	-1	-1	-1	-1	-1	1	9.268

(a)

Runs	**A**	**B**	**C**	**D**	**E**	**F**	**Response**
1	-1	-1	-1	1	1	1	8.258
2	1	-1	-1	-1	-1	1	7.784
3	-1	1	-1	-1	1	-1	13.136
4	1	1	-1	1	-1	-1	7.518
5	-1	-1	1	1	-1	-1	9.458
6	1	-1	1	-1	1	-1	8.298
7	-1	1	1	-1	-1	1	12.582
8	1	1	1	1	1	1	10.225
9	1	1	1	-1	1	1	8.965
10	-1	-1	-1	1	1	1	8.258
11	1	1	-1	1	1	1	10.21
12	1	1	1	-1	-1	1	10.083

(b)

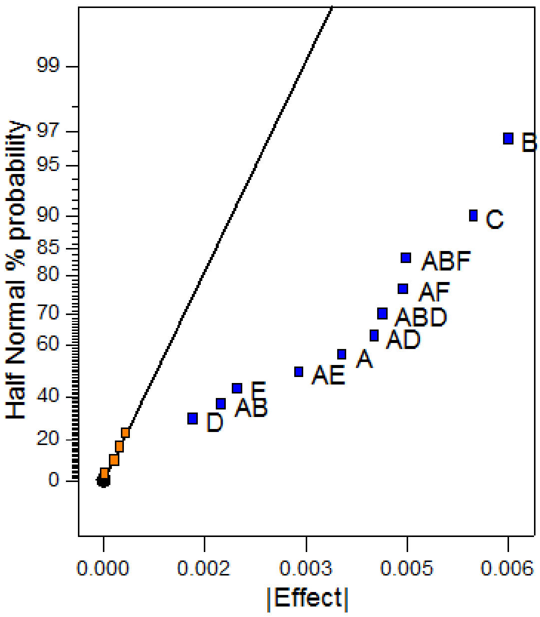


Figure 2-1: HNP of the Bead Width L_{16} Data

Table 2-4: ANOVA for the HNP analysis of the L_{16} Bead Width Data

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value
A	0.00005	1	0.00005	264.8643	< 0.0001
B	0.000144	1	0.000144	764.2592	< 0.0001
C	0.00012	1	0.00012	638.1302	< 0.0001
D	0.00000709	1	0.00000709	37.57729	0.0036
E	0.0000157	1	0.0000157	83.2852	0.0008
AB	0.0000122	1	0.0000122	64.63058	0.0013
AD	0.0000645	1	0.0000645	341.7604	< 0.0001
AE	0.0000335	1	0.0000335	177.5793	0.0002
AF	0.0000791	1	0.0000791	419.2732	< 0.0001
ABD	0.0000686	1	0.0000686	363.4727	< 0.0001
ABF	0.0000808	1	0.0000808	428.2752	< 0.0001
Residual	0.000000754	4	0.000000189		
Total	0.000677	15			

2-4 Analysis the L_{12} Data Using SWR and BSS methods

The results of performing the SWR analysis on the extracted L_{12} data are shown in Table 2-5. Displayed in the upper part of the table are the effects declared significant by the SWR method in the order of their selection. Also shown is the p-value at which each effect was picked-up. As evident from the ANOVA table all the selected effects were significant at $\alpha = 0.1$. Consequently, the true active effect C was falsely declared inert using SWR. On the other hand, the interaction BD was incorrectly identified as active.

Table 2-5: SWR Analysis of the Bead Width L_{12} Data

	Steps	**Selected Effects**	**p-value**		
	1	B	0.0118		
	2	A	0.0022		
	3	AB	0.0088		
	4	BD	0.056		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	19.22281	1	19.22281	65.04549	< 0.0001
B	25.23	1	25.23	85.37242	< 0.0001
AB	5.35736	1	5.35736	18.12805	0.0038
BD	1.54686	1	1.54686	5.234214	0.0560
Residual	2.068701	7	0.295529		
Total	51.9322	11			

In applying the BSS method, all the possible models comprising the main effects and two factor interactions were fitted using the extracted L_{12} data. To save space, only the selected model is presented here. As explained in Chapter 5, the selection was made utilising the knowledge regarding the true number of active effects which was 4 in this case. All the fitted models comprising 4 terms were examined and the one associated with the highest R^2_{adj} was selected. This was

$$Y = 9.71 - 1.34\,A + 1.45B - 0.67AB + 0.38BD$$

Its R^2_{adj} was 0.94. As they declared the same effects as significant, the ANOVA of the results of the BSS method is equivalent to the one associated with SWR which is given in Table 2-5.

2-5 Analysis of the DO Augmented L_8 Data Using SWR and BSS Methods

The effects declared significant by SWR when applied to analyse the DO augmented L_8 are displayed in Table 2-6. Again the upper part of the table shows the effects selected at each step of the SWR procedure. As can be seen in the ANOVA part of Table 2-6, all the selected effects were significant at α =0.1.

Table 2-6: SWR Analysis of the Bead Width DO Augmented L_8 Data

	Steps	**Selected Effects**	**p-value**		
	1	CE	0.0651		
	2	CF	0.0655		
	3	D	0.0124		
	4	BE	0.0513		
	5	AE	0.0307		
	6	E	0.0113		
	7	A	0.0363		
	8	F	0.0332		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Block	0.082955	1	0.082955		
A	0.124677	1	0.124677	119.6454	0.0083
D	6.141672	1	6.141672	5893.798	0.0002
E	1.002329	1	1.002329	961.8752	0.0010
F	0.029791	1	0.029791	28.58837	0.0332
AE	2.063565	1	2.063565	1980.28	0.0005
BE	2.150788	1	2.150788	2063.983	0.0005
CE	6.895274	1	6.895274	6616.984	0.0002
CF	13.28494	1	13.28494	12748.77	< 0.0001
Residual	0.002084	2	0.001042		
Total	33.9302	11			

As should be the case when analysing any augmented design, the extracted runs were analysed as comprising two blocks. One was associated with the original L_8 data whereas, the other represented the augmenting runs. Consequently, the variation due to the block effect was isolated from the error variance as shown in Table 2-6. Given the effects that were pronounced significant by the SWR method, it is clear that only one of the true active effects (factor A) was detected. Moreover, 7 inert effects were falsely declared significant.

With regard to the BSS method, the best model of size 4 was

$$Y= 9.85-0.79AB+1.21AD-1.20BD+0.46BE$$

Its R^2_{adj} was 0.842. Considering the block impact, the ANOVA of the BSS selected effects is given in Table 2-7. Clearly, the interactions AB, AD, BD and BE were significant at α =0.05 whereas the BE interaction was significant at α =0.1. Consequently, of the true active effects only the AB interaction was correctly detected.

Table 2-7: ANOVA of the BSS Selected Effects of the Bead Width DO Augmented L_8 Data

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value
Block	0.082955	1	0.082955		
AB	5.026035	1	5.026035	9.408506	0.0220
AD	15.44073	1	15.44073	28.90434	0.0017
BD	15.02789	1	15.02789	28.13152	0.0018
BE	2.262696	1	2.262696	4.235663	0.0853
Residual	3.205207	6	0.534201		
Total	33.9302	11			

2-6 Analysis of the BM Augmented L_8 Data Using SWR and BSS Methods

Performing the SWR analysis on the BM augmented L_8 data rendered three selected effects as shown in Table 2-8. Also displayed in this table is the ANOVA of the selected effects which was performed taking the block impact into consideration. Clearly all the selected effects were significant at α =0.1. As a result only factor B and the AB interaction were correctly detected using this strategy.

Table 2-8: SWR Analysis of the Bead Width BM Augmented L_8 Data

	Steps	Selected Effects	p-value		
	1	B	0.0499		
	2	CE	0.0159		
	3	AB	0.0386		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Block	0.206647	1	0.206647		
B	15.14523	1	15.14523	19.4563	0.0031
AB	5.026035	1	5.026035	6.45669	0.0386
CE	12.16358	1	12.16358	15.62593	0.0055
Residual	5.44896	7	0.778423		
Total	35.72968	11			

As with the BSS method the model with four terms that attained the highest R^2_{adj} (0.836) was

$$Y = 9.60 + 1.18B - 0.79\,AB + 0.49CD - 1.06CE$$

Its associated ANOVA is presented in Table 2-9. Clearly only B, AB and CE effects were significant at α =0.05. Thus, besides its false selection of the interactions CD and CE, this strategy failed to identify two of the true active effects, namely factors A and C.

Table 2-9: ANOVA of the BSS Selected Effects of the Bead Width BM Augmented L_8 Data

Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Block	0.206647	1	0.206647		
B	15.14523	1	15.14523	25.95132	0.0022
AB	5.026035	1	5.026035	8.612102	0.0261
CD	1.947351	1	1.947351	3.336782	0.1175
CE	12.16358	1	12.16358	20.84228	0.0038
Residual	3.501609	6	0.583601		
Total	35.72968	11			

3. Phase-3: Analysis of the EDM Experiment

3-1 Analysing the Full Factorial Workpiece MRR Data

The standard ANOVA of the workpiece MRR full factorial data is presented in Table 3-1 (a). The only active effects with p-values smaller than α =0.05 were factors A and B. The same effects were declared significant using the regression ANOVA. However, only certain components of these were distinguishable from noise as displayed in Table 3-1 (b) which revealed that at α =0.05, A_l, B_l and A_q were the only significant components.

Table 3-1: ANOVA of the Full Factorial Workpiece MRR Data in its (a) Standard and (b) Regression Forms

Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	662.8012	2	331.4006	51.98684	< 0.0001
B	295.9204	2	147.9602	23.21053	< 0.0001
C	13.58818	2	6.79409	1.065789	0.3498
D	30.5734	2	15.2867	2.398026	0.0981
Residual	458.9785	72	6.374702		
Total	1461.862	80			
(a)					
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Al	634.4925	1	634.4925	95.84096	< 0.0001
Bl	295.9204	1	295.9204	44.69917	< 0.0001
Aq	28.30871	1	28.30871	4.276069	0.0421
Bq	0	1	0	0	1.0000
Residual	503.1401	76	6.620264		
Total	1461.862	80			
(b)					

3-2 Extracting the Fractional Factorial Workpiece MRR Data

The appropriated runs for the L_{18} and L_9 OAs were taken from the full factorial L_{81} workpiece MRR data as shown in Table 3-2. Since only four factors were studied, dummy factors were used to denote the unassigned three level columns of the L_{18} OA. As the scope of this phase was confined to three-level effects, the L_{18} two-level column was incorporated into the error estimate.

Table 3-2: Extracted (a) L_{18}, and (b) L_9 Workpiece MRR Data

e	A	B	C	D	E	F	G	Response
1	1	1	1	1	1	1	1	9.029
1	2	2	2	2	1	2	2	15.801
1	3	3	3	3	1	3	3	22.573
1	1	1	2	2	2	3	3	11.287
1	2	2	3	3	2	1	1	13.544
1	3	3	1	1	2	2	2	27.088
1	1	2	1	3	3	2	3	13.544
1	2	3	2	1	3	3	1	22.573
1	3	1	3	2	3	1	2	15.801
2	1	3	3	2	1	2	1	13.544
2	2	1	1	3	1	3	2	11.287
2	3	2	2	1	1	1	3	20.316
2	1	2	3	1	2	3	2	13.544
2	2	3	1	2	2	1	3	15.801
2	3	1	2	3	2	2	1	31.603
2	1	3	2	3	3	1	2	13.544
2	2	1	3	1	3	2	3	11.287
2	3	2	1	2	3	3	1	18.059

(a)

A	B	C	D	Response
1	1	1	1	9.029
2	1	2	2	11.287
3	1	3	3	15.801
1	2	2	3	13.544
2	2	3	1	20.316
3	2	1	2	18.059
1	3	3	2	13.544
2	3	1	3	15.801
3	3	2	1	22.573

(b)

3-3 PU Analysis of the L_{18} Data

The PU method was applied to analyse the L_{18} workpiece data in two forms. One was implemented with the standard ANOVA where each of the examined effects had two degrees of freedom whereas the other was associated with the regression ANOVA where each effect was decomposed into two single degree of freedom effects namely linear and quadratic. Table

3-3 shows the results of using the PU method with the standard ANOVA. The smallest three-level effect G was pooled at the first step. Owing to the use of the two-level column degree of freedom to estimate the error, the significance of G was tested and found to be insignificant at α =0.1. Thus it was pooled with the error estimate. This was also the case with all the remaining effects except A which was flagged as significant at α =0.1 in the seventh step of the PU analysis. Consequently, using this analysis strategy factor B was falsely pronounced inert.

Table 3-3: PU with the Standard ANOVA of the L_{18} Workpiece MRR Data

	Steps	**Pooled Effects**	**p-value**		
	1	G	0.4822		
	2	D	0.3430		
	3	E	0.1968		
	4	F	0.1804		
	5	C	0.1856		
	6	B	0.2184		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	3.334766	2	1.667383	8.951368	0.0028
Residual	2.794069	15	0.186271		
Total	6.128835	17			

The results of applying the PU method with the regression ANOVA to analyse the L_{18} MRR data are shown in Table 3-4. The upper part of this table lists the seven pooled effects along with their p-values which were in every case larger than 0.1. The selected effects were analysed using ANOVA and found significant at α =0.1. Evidently, alongside four inert effects this procedure detected all the true active effects as identified in the full factorial analysis.

Table 3-4: PU with the Regression ANOVA of the L_{18} Workpiece MRR Data

	Steps	Pooled Effects	p-value		
	1	Dl	0.845		
	2	El	0.819		
	3	Cl	0.611		
	4	Gq	0.53		
	5	Bq	0.263		
	6	Fl	0.182		
	7	Gl	0.133		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Al	3.095557	1	3.095557	47.13362	< 0.0001
Aq	0.239209	1	0.239209	3.642241	0.0854
Bl	0.513803	1	0.513803	7.823276	0.0189
Cq	0.566174	1	0.566174	8.62069	0.0149
Dq	0.239209	1	0.239209	3.642241	0.0854
Eq	0.409061	1	0.409061	6.228448	0.0317
Fq	0.409061	1	0.409061	6.228448	0.0317
Residual	0.656762	10	0.065676		
Total	6.128835	17			

3-4 PD Analysis of the L_{18} Data

As was the case with the PU method the PD procedure was implemented with both the standard and the regression ANOVA. In the former case, the results were equivalent to those yielded using the PU method (Table 3-3). In fact, the first selected effect was factor A which was found significant at α =0.1 with a p-value of 0.0028. In the second step, factor B was selected and its significance test yielded a p-value of 0.2184 as shown in Table 3-3. As this value was larger than 0.1, factor B was falsely deemed inert.

The regression ANOVA results of the PD analysis are presented in Table 3-5. All the selected effects were significant at α =0.1. As A_q was not among the selected effects, only two of the true active effects were detected by this analysis strategy.

Table 3-5: PD with the Regression ANOVA of the L_{18} Workpiece MRR Data

	Steps	Selected Effects	p-value		
	1	Al	0.001		
	2	Cq	0.083		
	3	Bl	0.076		
	4	Eq	0.086		
	5	Fq	0.06		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Al	3.095557	1	3.095557	32.72319	< 0.0001
Bl	0.513803	1	0.513803	5.431421	0.0380
Cq	0.566174	1	0.566174	5.985037	0.0308
Eq	0.409061	1	0.409061	4.32419	0.0597
Fq	0.409061	1	0.409061	4.32419	0.0597
Residual	1.135179	12	0.094598		
Total	6.128835	17			

3-5 PU Analysis of the L_9 Data

The results of using the PU method in conjunction with the standard ANOVA to analyse the L_9 MRR data are shown in Table 3-6. None of the four studied effects was found significant at α =0.1 as their p-values were larger than or equal to α. evidently, the use of this analysis strategy resulted in overlooking all the true active effects.

Table 3-6: PU with the Standard ANOVA of the L_9 Workpiece MRR Data

	Steps	Pooled Effects	p-value		
	1	C			
	2	D	0.3500		
	3	B	0.1		
	4	A	0.1495		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	0.690732	2	0.345366	2.652174	0.1495
Residual	0.78132	6	0.13022		
Total	1.472053	8			

As demonstrated in Table 3-7, when used with the regression ANOVA the PU method successfully detected two of the true active effects namely A_l and B_l. The only missing effect was A_q which was the smallest and thus was the first to be used to estimate the error. Furthermore, four spurious effects were declared significant by this analysis strategy.

Table 3-7: PU with the Regression ANOVA of the L_9 Workpiece MRR Data

	Steps	**Pooled Effects**	**p-value**		
	1	Aq			
	2	Cq	0.5		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Al	0.687902	1	0.687902	243	0.0041
Bl	0.416138	1	0.416138	147	0.0067
Bq	0.138713	1	0.138713	49	0.0198
Cl	0.076434	1	0.076434	27	0.0351
Dl	0.076434	1	0.076434	27	0.0351
Dq	0.070772	1	0.070772	25	0.0377
Residual	0.005662	2	0.002831		
Total	1.472053	8			

3-6 PD Analysis of the L_9 Data

Analysing the L_9 MRR data using the PD along with the standard ANOVA yielded the same conclusions as those drawn using the PU - see Table 3-6. In fact, factor A was the first effect to be selected and tested using an error estimate that was formed by pooling the B, C and D effects. As shown in Table 3-6, the resultant p-value was 0.1495 which exceeded the value of α (0.1). Therefore, this strategy failed to detect any of the true active effects.

The results of performing the PD method along with the regression ANOVA are displayed in Table 3-8. Attaining p-values smaller than 0.1, the two effects A_l and B_l were correctly detected as identified in the full factorial analysis. A_l was the only overlooked effect using this analysis strategy.

Table 3-8: PD with the Regression ANOVA of the L_9 Workpiece MRR Data

	Steps	**Selected Effects**	**p-value**		
	1	Al	0.0423		
	2	Bl	0.0404		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Al	0.687902	1	0.687902	11.21538	0.0154
Bl	0.416138	1	0.416138	6.784615	0.0404
Residual	0.368013	6	0.061336		
Total	1.472053	8			

4. Phase-4: Analysis of the EDM Experiment

4-1 Analysing the Full Factorial Tool MRR Data

Table 4-1 (a) presents the standard ANOVA of the tool MRR data of the EDM experiment. The only effects with p-values smaller than 0.05 were factors A and B and their interaction. The regression ANOVA confirmed the significance of certain components of these effects at $\alpha = 0.05$ as shown in Table 4-1 (b) since the only components with a p-value smaller than 0.05 were A_l, B_l, B_q and A_lB_l.

Table 4-1: ANOVA of the Full Factorial Tool MRR Data in its (a) Standard and (b) Regression Forms

Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
A	75.13675	2	37.56838	19.3708	< 0.0001
B	100.9717	2	50.48587	26.03125	< 0.0001
C	1.96838	2	0.98419	0.507463	0.6043
D	3.444665	2	1.722332	0.88806	0.4162
AB	117.0571	4	29.26427	15.08909	< 0.0001
Residual	131.8815	68	1.939433		
Total	430.4601	80			
(a)					

Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Al	74.94453	1	74.94453	39.30242	< 0.0001
Bl	59.99714	1	59.99714	31.46371	< 0.0001
Aq	0.192225	1	0.192225	0.100806	0.7518
Bq	40.9746	1	40.9746	21.4879	< 0.0001
AlBl	116.3958	1	116.3958	61.04032	< 0.0001
AlBq	0.00001	1	0.00001	0.00000	1.0000
AqBl	0.415205	1	0.415205	0.217742	0.6422
AqBq	0.246047	1	0.246047	0.129032	0.7205
Residual	137.2945	72	1.906868		
Total	430.4601	80			
(b)					

4-2 Extracting the Fractional Factorial Tool MRR Data

The full factorial tool MRR data that corresponded to the rows of the L_{18} and the L_9 fractional OAs were extracted as displayed in Table 4-2.

Table 4-2: Extracted (a) L_{18}, and (b) L_9 Tool MRR Data

e	A	B	C	D	E	F	G	Response
1	1	1	1	1	1	1	1	7.81
1	2	2	2	2	1	2	2	6.7
1	3	3	3	3	1	3	3	12.28
1	1	1	2	2	2	3	3	7.81
1	2	2	3	3	2	1	1	6.7
1	3	3	1	1	2	2	2	11.16
1	1	2	1	3	3	2	3	5.58
1	2	3	2	1	3	3	1	8.93
1	3	1	3	2	3	1	2	6.7
2	1	3	3	2	1	2	1	3.35
2	2	1	1	3	1	3	2	7.81
2	3	2	2	1	1	1	3	7.81
2	1	2	3	1	2	3	2	5.58
2	2	3	1	2	2	1	3	10.04
2	3	1	2	3	2	2	1	6.7
2	1	3	2	3	3	1	2	6.7
2	2	1	3	1	3	2	3	6.7
2	3	2	1	2	3	3	1	7.81

(a)

A	B	C	D	Response
1	1	1	1	7.81
2	1	2	2	6.7
3	1	3	3	6.7
1	2	2	3	5.58
2	2	3	1	6.7
3	2	1	2	7.81
1	3	3	2	3.35
2	3	1	3	8.93
3	3	2	1	12.28

(b)

4-3 Analysis of the L_{18} Tool MRR Data Using SWR and BSS Methods

The effects that were picked-up by the SWR procedure when applied to analyse the L_{18} MRR data are displayed in their selection order in the upper part of Table 4-3. The interaction A_lB_l was the first to be selected followed by the A_l and the B_q. The component B_l was the last

Table 4-3: SWR Analysis of the L_{18} Tool MRR Data

	Steps	**Selected Effects**	**p-value**		
	1	AlBl	0.0004		
	2	Al	0.0023		
	3	Bq	0.0287		
	4	Bl	0.0094		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Al	20.35808	1	20.35808	28.36076	0.0001
Bl	6.645408	1	6.645408	9.257696	0.0094
Bq	6.786025	1	6.786025	9.453588	0.0089
AlBl	30.45901	1	30.45901	42.43235	< 0.0001
Residual	9.331729	13	0.717825		
Total	73.58025	17			

selected effect. The ANOVA of the chosen effects is given in the lower part of Table 4-3. As their p-values were smaller than 0.1, they were all declared significant. Consequently, using this strategy all the true significant effects were correctly detected.

As the true number of active effects was four, the BSS method was used to identify the best model of size four that could be fitted to the L_{18} MRR data. This comprised the same terms identified by the SWR method i.e.

$$Y = 7.56 + 1.30\ A_l + 0.744\ B_l + 0.434\ B_q + 1.95\ A_lB_l$$

Thus its associated ANOVA was equivalent to the one presented in Table 4-3 for the SWR results. The R^2_{adj} for this model was 0.834.

The main implication of the SWR and the BSS analyses of the L_{18} data is that the same conclusions drawn from the full factorial data would have been inferred had a smaller design such as the L_{18} been performed instead.

4-4 Analysis of the L_9 Tool MRR Data Using SWR and BSS Methods

The results of performing the SWR analysis on the L_9 data are presented in Table 4-4. As can be seen from the ANOVA results, the phenomenon of overfitting was experienced in this analysis, the MSE being equal to zero. The chief reason for this is the complexity of the aliasing pattern associated with the selected effects. Despite the perfect fit of the yielded model, only two of the four true active effects namely B_q and A_lB_l were detected.

Table 4-4: SWR Analysis of the L_9Tool MRR Data

	Steps	**Selected Effects**	**p-value**		
	1	AlBl	0.0262		
	2	CqDl	0.0035		
	3	BlCl	0.0093		
	4	Bq	0.0290		
	5	BlCq	< 0.0001		
Source	**Sum of Squares**	**DF**	**Mean Square**	**F-Value**	**p-Value**
Bq	0.553607	1	0.553607	63660000	< 0.0001
AlBl	20.34505	1	20.34505	63660000	< 0.0001
BlCl	2.18529	1	2.18529	63660000	< 0.0001
BlCq	0.293086	1	0.293086	63660000	< 0.0001
CqDl	5.605269	1	5.605269	63660000	< 0.0001
Residual	0	3	0		
Total	47.61019	8			

The best model of size four that the BSS analysis of the L_9 data rendered was

$$Y = 7.32 + 2.74\, A_lB_l - 0.897\, B_lC_l - 0.292\, B_q + 1.20\, C_qD_l$$

its associate ANOVA is given in Table 4-5 and its R^2_{adj} was 0.988. Obviously, the selected terms were the same as the first four effects that were picked-up by the SWR procedure. As their p-values were smaller than 0.05, they were all declared significant. Again using this analysis strategy, only two of the true active effects were detected.

Table 4-5: ANOVA of the BSS Selected Effects of the L_9Tool MRR Data

Source	Sum of Squares	DF	Mean Square	F-Value	p-Value
Bq	0.814128	1	0.814128	11.11111	0.0290
AlBl	27.25874	1	27.25874	372.0238	< 0.0001
BlCl	2.052823	1	2.052823	28.01667	0.0061
CqDl	7.388209	1	7.388209	100.8333	0.0006
Residual	0.293086	4	0.073271		
Total	47.61019	8			

APPENDIX 10:
CHAPTER 8 DERIVATIONS

1. Relating the PC_2 to the ε^2

To show that PC_2 and ε^2 are numerically equivalent, both the numerator and denominator of equation 8.6 should be multiplied by (n-1) and the MSE should be substituted for SSE/(n-ν-1) so that

$$\varepsilon^2 = \frac{SST - (n-1)MSE}{SST}$$

Replacing the SST in the numerator with $\sum_i^k SS_{FEi} + SSE$ yields

$$\varepsilon^2 = \frac{\sum_i^k SS_{FEi} + SSE - (n-1)MSE}{SST}$$

Substituting (n-ν-1)MSE for SSE gives

$$\varepsilon^2 = \frac{\sum_i^k SS_{FEi} - (\nu)\,MSE}{SST}$$

This equals the model (encompassing k FEs) ε^2 or Adjusted- R^2. Although, it provides an overall measure of the FEs collective importance in terms of the overall explained variance, it does not identify the relative importance of each FE. In orthogonal designs, the model sum of squares can be decomposed into k independent sums of squares corresponding to the k FEs that comprise it. Moreover, $\nu = \sum_{i=1}^k df_{FEi}$. Thus a partial measure of explained variance can be obtain for each FE as follows

$$\varepsilon^2_{FEi} = \frac{SS_{FEi} - (df_{FEi})MSE}{SS_{total}}$$

which is the PC_2 formula presented in Ross (1996) i.e. equation 3.2.

2. Impact of SES on C_{pk}

Considering the shift in the process average that results from a specific $SES_{FE(i)}$, a new C_{pk} value can be obtained from the conventional one as follows:

$$\text{New } C_{pk} = \frac{\text{Min}[(USL - \overline{X} - SES_{FE(i)}) \text{ or } (\overline{X} - LSL - SES_{FE(i)})]}{3\sigma}$$

$$\text{New } C_{pk} = \text{Min}\left(\frac{USL - \overline{X}}{3\sigma} - \frac{SES_{FE(i)}}{3\sigma} \text{ or } \frac{\overline{X} - LSL}{3\sigma} - \frac{SES_{FE(i)}}{3\sigma}\right)$$

$$\text{New } C_{pk} = \left(\frac{\text{Min}[(USL - \overline{X}) \text{ or } (\overline{X} - LSL)]}{3\sigma}\right) - \frac{SES_{FE(i)}}{3\sigma}$$

As the $SES_{FE(i)}$ expresses the factorial effect in numbers of standard deviations the above equation can be rewritten as:

$$\text{New } C_{pk} = \left(\frac{\text{Min}[(USL - \overline{X}) \text{ or } (\overline{X} - LSL)]}{3\sigma}\right) - \frac{SES_{FE(i)}}{3}$$

Consequently

$$\text{New } C_{pk} = C_{pk} - (SES_{FE(i)}/3)$$

APPENDIX 11:
LIST OF JOURNAL ARTICLES

Journal Articles in Preparation

- Al-Ghamdi K., Aspinwall E., "The Importance of Experimental Design", to appear in the Journal of Manufacturing Technology Management.
- Al-Ghamdi K., Aspinwall E., "A Review of the Practice of Experimental Design in Manufacturing", to appear in the International Journal of Quality and Reliability Management.
- Al-Ghamdi K., Aspinwall E., "A Simple Method for Generating Aliasing Patterns"
- Al-Ghamdi K., Aspinwall E., "Performance Evaluation of Certain Two-Level Experimental Design Strategies"
- Al-Ghamdi K., Aspinwall E., "Performance Evaluation of Certain Three-Level Experimental Design Strategies"
- Al-Ghamdi K., Aspinwall E., " An Investigation of the Use of p-values and Factorial Effects' Importance Measures in Manufacturing"
- Al-Ghamdi K., Ho S.K., Aspinwall E, Aspinwall D.K., "A Comparative Study of Taguchi Main Effects Analysis"