BENCHMARKED HARD DISK DRIVE PERFORMANCE CHARACTERIZATION AND OPTIMIZATION BASED ON DESIGN OF EXPERIMENTS TECHNIQUES

A Thesis

presented to

the Faculty of California Polytechnic State University,

San Luis Obispo

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Electrical Engineering

by

Yu-wei Lin

June 2010

COMMITTEE MEMBERSHIP

TITLE: Benchmarked hard disk drive performance characterization and optimization based on Design of Experiments techniques

AUTHOR: Yu-wei Lin

DATE SUBMITTED: June 2010

COMMITTEE CHAIR: John Y. Oliver, Assistant Professor

COMMITTEE MEMBER: Jianbiao Pan, Associate Professor

COMMITTEE MEMBER: James G. Harris, Professor Emeritus

Abstract

Benchmarked hard disk drive performance characterization and optimization based on Design of Experiments techniques

Yu-wei Lin

This paper describes an experimental study offered by Designs of Experiments (DOE) within the defined factor domains to evaluate the factor effects of simultaneous characteristics on the benchmarked hard disk drive performance by proposing well-organized statistical models for optimizations. The numerical relations of the obtained models permit to predict the behaviors of benchmarked disk performances as functions of significant factors to optimize relevant criteria based on the needs.

The experimental data sets were validated to be in satisfying agreement with predicted values by analyzing the response surface plots, contour plots, model equations, and optimization plots. The adequacy of the model equations were verified effectively by a prior generation disk drive within the same model family. The retained solutions for potential industrializations were the concluded response surface models of benchmarked disk performance optimizations.

The comprehensive benchmarked performance modeling procedure for hard disk drives not only saves experimental costs on physical modeling but also leads to hard-to-find quality improvement solutions to manufacturing decisions.

Acknowledgements

First and foremost I would like to thank my thesis advisor, Dr. John Y. Oliver for allowing me to join this project, for his support, for his advice, and for his kindness. I would also like to acknowledge Dr. Jianbiao (John) Pan for his guidance, for his expertise, and especially for his patience during the development in this thesis. My appreciation also goes to Dr. James G. Harris for his feedback and for his participation on my thesis committee.

I thank Western Digital Corporation for providing me with the opportunity to engage in this study, for their support, and for their trust in me. It has been a pleasure to study at the California Polytechnic State University in San Luis Obispo during the years. I thank my fellow students for their friendship, for their company, and for their support. Many individuals have been a part of my graduate education and I am highly grateful to all of them.

Finally, I would like to acknowledge my family for their continuous and unconditional love, support, and guidance. They have been a constant inspiration throughout my life. I would like to thank them for all they are and all they have done for me.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Overview of Benchmarked Hard Disk Drive Performance Modeling

Rapid advancements in high-capacity, low-cost hard disk drives have been one of the key factors enforcing the development of the information technology in modern societies that deeply rely on digital data. The deployments of hard disk drives range from consumer electronics and home computers to enterprise storage arrays and network servers. It is critical to identify and resolve hard disk drive performance issues when dealing with increasing amount of digital data to be accessed. In the real world, it is intense to deliver an operational and optimal solution while facing stringent challenges of diversified market demands, rapid time-to-market, huge capital investment, and cost-sensitive competition.

Hard disk drive performance was one of the most underrated aspects back when the disk storage density was the primary concern of hard drive revolution. Over the past few decades, the price per gigabyte of disk storage has dropped from about $850 [19] in 1994 to about $0.07 [19] and lower in 2009. The researchers' attention has therefore gradually shifted from the subject of disk density to disk performance of the fundamental disk operations [6].

Hard disk drive performance means different things to different people and how the hard disk drive is used in a system. One common way to gauge the performance is to benchmark hard disk drives using high-level benchmark programs specifically developed for the purpose of measuring the performance of a system or components based on logic and analysis. These programs are application-derived benchmarks that simulate the impact of hard disk drive performance on the use of popular real world input/output bound applications such that a lot of reading and writing to the hard disk drive are required. The intention of benchmarking is to express the overall performance of a system or individual components in numeric terms.

It is essential for hard disk drive manufacturers to flexibly and efficiently use the existing resources and capability to evaluate the benefits and impact of the proposed solutions for performance improvements as quick responses to market demands and competitiveness over competitors. The fast evolving complexity of hard disk drive technology makes it difficult to quickly identify and resolve performance issues with associated economic and operational costs. The research in the field of hard disk drive performance often involves both analytical or simulation models to compare alternative approaches. From performance optimization's point of view, an analytical model applying design of experiments (DOE) technique is more suitable to characterize the benchmarked disk performance due to the fact that an accurate simulator is not able to be fully developed without knowing how the particular metric of interest, typically known as the "benchmark score", which was defined by the vendor when the benchmarking software was written. Nevertheless, only few extant DOE models regarding overall hard disk drive performance benchmarking have been published.

Thousands of companies have provided documented stories of substantial savings to their business through the application of DOE. Especially during the tough times of recent economic recession, most manufacturers are challenged to find more economical and cost-effective ways to maintain market share without impacting product quality and delivery. DOE is an effectively organized approach designed to have minimal disruption to normal business operations which not only saves industrial experimental costs but also greatly increases the odds of identifying the hard-to-find solution to quality problems, reducing product variation, and optimizing product performance [27].

This empirical study contributes to benchmarked hard disk drive performance characterization by applying the DOE approach to develop statistical models. Such models express the way hard disk drive technologies cooperate to perform various tasks and workloads. Furthermore, the performance of future storage devices within the same model family can be predicted based on the major trend of the current model as a function of the significant factors.

1.2 Previous Work

Hard disk drive performance has been studied by many researchers in the past decade. Grochowski et al. [1] pointed out that the characteristics of future hard disk drives can be estimated by analyzing the specification trends of the disk drive designs with the assumption that no major change in hard disk drive technology will occur. In fact, not all advances in hard disk drive technology are necessarily beneficial to disk drive performance. The considerations of how or where an advance should be applied when weighing technology options for hard disk drive designs have been studied [2]. As a result, the researchers' attentions have focused on increased rotational speeds, faster seek times, and higher data transfer rates as the principal drivers of hard disk drive performance optimization.

Ruemmler et al. [4] emphasized the importance of a high quality and accurate disk performance model by proposing a detailed simulation model in terms of I/O time and describing general techniques for disk drive modeling. The introduction of the demerit figure is used as the metric for hard disk drive performance evaluations. This simulation model has been validated and implemented by Kotz et al. [12] and Triantafillou et al. [6] with detailed analytical characterizations of modern disk behaviors, such as command queuing to minimize rotational delay.

Ng [11] proposed a simple analytic queuing model to improve overall subsystem performance via reducing rotational latency and rotational position sensing (RPS) miss delay.

Automated hard disk drive characterization programs such as Disk Extraction (DIXtrac), presented by Schindler et al. [13], quickly and automatically characterized disk drives via extraction of disk layout, mechanics parameters, cache parameters, and command processing overhead. Shim and Park [10] developed Disk Geometry Analyzer (DIG) which efficiently extracted comprehensive internal information and characterized the performance metric of hard disk drives.

Although there was an awareness for hard disk drive performance characterization, these studies have neither reported on detailed simulation models nor explored the hard disk drives closely enough to provide insights into the interactions between the factors.

The Design of Experiments (DOE) methodology has been commonly used for optimizing mechanics parameter settings of the hard disk drives. Hao et al. [16] proposed a self-tuning robust control scheme based on the Response Surface Methodology (RSM) to optimize the performance of the hard disk drive servo system. The optimization process of actuator dynamics, as one of the most significant design factors, was studied by Oh et al. [17] based on the RSM to insure volume production capability with reasonable amount of tolerances. Li et al. [18] discussed the application of RSM to suppress the cogging torque to an acceptable level as an effective way to optimize the performance of the hard disk drive spindle motor. However, reports on the feasibility of the DOE methodology have been rare in the literature for benchmarked hard disk drive performance.

This study is unique in that the data is collected based on parameter settings typically only available from the hard disk drive manufacturers but provides insight into the depth of detailed benchmarking result that is often studied by end-users. The real-world application of DOE is applied to improve benchmarked performance score in selected Western Digital hard disk drives.

Chapter 2

Experiment Setup

2.1 Characteristics of Parameters

Many parameters are relevant to hard disk drive performance, such as cache size, data density, data rate, drive form factor, interface, overhead, platter diameter, platter count, recording technology, rotation speed, seek time, and etc. This study focused on the impacts of data rate, seek time, and overhead on hard disk drive performance benchmarked using PCMark software. The ranges of these parameter settings were determined based on the technical specifications of each hard disk drive under investigation and knowledge of the experienced engineers participated in this project. All other parameters were held constant throughout the experiments to avoid response variations induced by uncontrollable factors.

2.1.1 Data Rate

Modern hard drives employ a Zoned Bit Recording (ZBR) technique that allows different read speeds depending on where the data is located. With this technique, tracks are grouped into zones based on their distances from the center of the disk, and each zone is assigned a number of sectors per track to scale the tremendous amount of data in bytes stored on each track as shown in Figure 2.1.

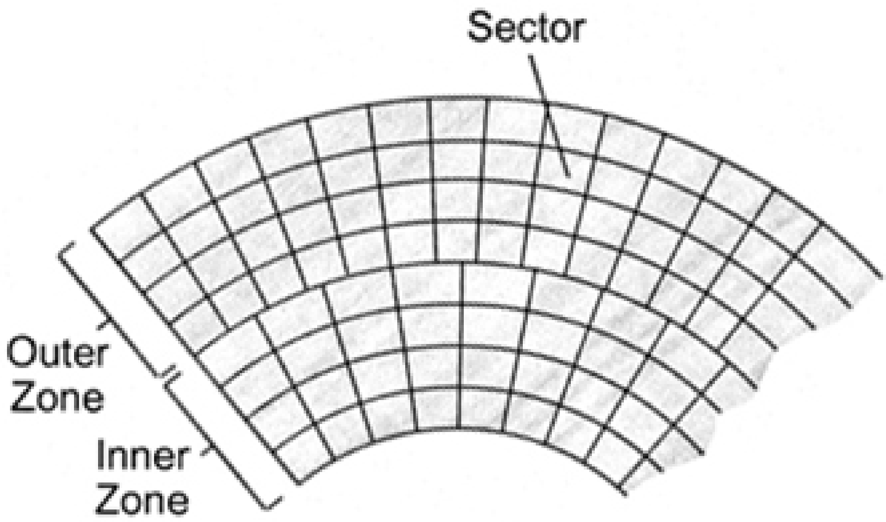

Figure 2.1: A graphical illustration of Zoned Bit Recording (ZBR) [25].

The number of sectors per track is determined by the linear bit density limitations throughout the whole disk where the tracks are concentric circles. As moving from the outer zones through inner zones, each zone contains fewer sectors per track than the one before. ZBR technique allows for more efficient use of the space in outer tracks that were generally underutilized by non-ZBR techniques such that every track had the same number of sectors as the innermost zone.

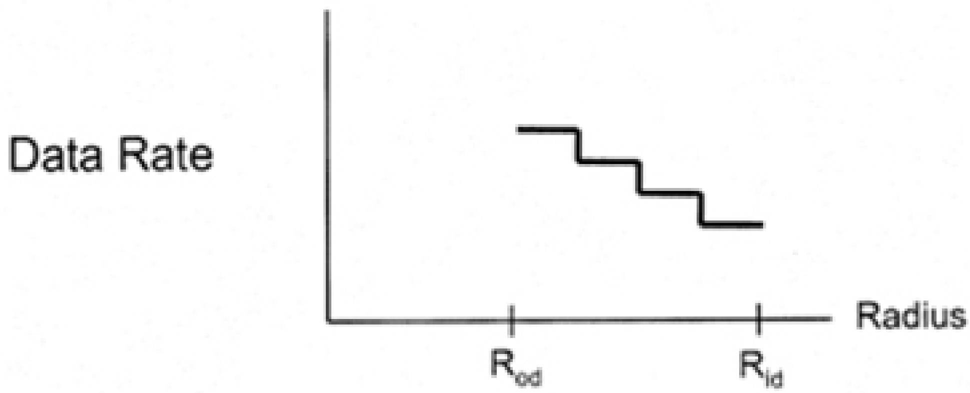

Figure 2.2: Data rate versus cylinder radius due to ZBR [25].

Due to the constant angular velocity throughout the platters, the data rate is the fastest when reading the outermost cylinders which the zone contains the most data. Figure 2.2 shows that the data rate decreases in gradual steps as moving towards inner cylinders. The gradual steps across the entire platter surface are the consequences of ZBR technique as the tracks within the same zone have the same number of sectors and therefore the same data rate.

The selected Western Digital hard disk drives in this study have a few thousand tracks which have been divided into eight zones. Each zone has different number of tracks and different number of sectors. The size (length) of a sector remains fairly constant over the entire surface of the disk.

Since the data rate by its name is obviously the amount of data transferred per unit of time. The sequential transfer rate can be derived from the amount of data transferred per cylinder and the time per cylinder transfer. These representations are measured on a per cylinder basis because a sequential transfer rate covers an entire cylinder on a disk.

2.1.2 Seek Time

Seek time is the time measured for the read and write head movements between cylinders to position over data blocks before accessing on a specific track. The default seek profile is a lookup table that provides the expected seek time value for a given seek distance in cylinders [7] and often referred as the worst case seek time. The performance of a hard disk drive can often be improved by adjusting the seek profile to be closer to the average of the actual seek time [8].

The seek distance up to where a square root relationship exists between the seek time and seek distance in cylinder is considered to be short seeks and the seek distance after this boundary is so called long seeks where a linear relationship exists between the seek time and seek distance.

$$T_s(d) = \begin{cases} T_{s0} + k\sqrt{d} & d < d' \\ T'_{s0} + k'd & d \geq d' \end{cases} \tag{2.1}$$

Equation 2.1 describes the relationship between seek time and seek distance, where k and k' are the coefficients, d is the seek distance, d' is the boundary between short seeks and long seeks, T_{s0} and T'_{s0} are the settle time for short seeks and long seeks, respectively. Short seeks spend almost all of their time in the constant-acceleration phase, and their time is proportional to the square root of the seek distance plus the settle time. Long seeks spend most of their time moving at a constant speed, taking time that is proportional to distance plus the settle time [3].

The seek profile can be broken into a small number of distance groups that have similar variance value between the start and destination cylinders. On a Western Digital hard disk drive, the distance between the start and destination cylinders is typically divided into eight zones with the seven boundaries: 10, 41, 154, 400, 1000, 1500, and 4500 in cylinders. Each zone has its own data collection, calculation, and adjustment. Read and write operations are separated and each has its own eight zones. The variance of each distance group is the difference between adjusted seek profile and average of the actual seek time, where the adjusted seek profile is an attempt for adjusting seek time measurements and is manually created during servo characterization.

$$\begin{aligned}\text{Adjusted seek profile} &= \text{Default seek profile} + \text{Seek profile variable offset} \\ &\quad + \text{Seek profile fixed offset}\end{aligned} \tag{2.2}$$

Adjusted seek profile can be described as in Equation 2.2, where seek profile variable offset is the average of the difference between the actual seek time and the default seek profile. Normally, this is a negative value as the actual seek time should be less than the worst case seek time.

The seek time parameter discussed in this study is referred to the seek profile fixed offset, which is an explicitly conservative value included to account for the variability. It is normally a positive value, and the smaller this value is the closer the adjusted seek profile is to the actual seek time. It is stored as byte number 64 to 79 (total 16 bytes) in the reserved file (id 0x4002). The first three bytes, byte number 64 to 66, are referred to short seek reads follow by the next five bytes, byte number 67 to 71, as long seek reads. The next eight bytes of the seek profile fixed offset are for the write operation; where byte number 72 to 74 are referred to short seek writes, and byte number 75 to 79 are long seek writes. Each byte refers to a value measured in wedges as number of sectors on disk, and corresponds to each zone of the variance groups. The seek profile fixed offset ranges from 0 to 127 in wedges for each byte as only positive delay offset can be added. The seek profile fixed offset in wedges can be converted to units in time domain by multiplying a wedge-to-wedge time. The wedge-to-wedge time is different from drive to drive because it is affected by the rotational speed and number of wedges per track. The wedge-to-wedge time for each hard disk drive under investigation in this study was provided by Western Digital Corporation.

After powering up, seek profile variable offset and seek profile fixed offset of each zone and the initialized bit will be loaded from the reserved file to the static buffer. The Adaptive seek feature will be enabled if bit 0 of byte 80, the initialized bit is set. This feature can be disabled by resetting this bit to 0.

2.1.3 Overhead

Overhead is the time it takes for the controller to process an interrupt service routine (ISR) and handle servo hook code requests. The interrupt signal is generated every wedge to signify the determination of the correction factor in response to the servo hook codes loaded from the event detector; hence, there is an interaction with the wedge-to wedge time.

While the read and write heads are moving, they periodically read the servo hook codes written in special data areas on the disk which provide information about the locations of the heads. Servo hook codes are embedded during manufacturing either completely on one side of each platter or among the data in non-writable and engineered positions, not able to be modified. Each track loads different codes to the closed-loop control logic to dynamically guide the actuator to the correct track. The positions of the read and write heads are adjusted to compensate for any changes in the platter or head dimensions due to thermal expansion or physical stress [20]. The ISR mentioned in this chapter results from the actuator arm movement to a specific position. As the cycles prior to the execution become shorter, the process by which the read and write heads positioning becomes much faster. Therefore, lower overhead yields higher disk drive performance.

As mentioned earlier, the overhead is associated with wedge-to-wedge time since the ISR is processed every wedge to handle servo hook code requests. From the perspective of CPU utilization, overhead can be represented as how much system resource, such as CPU time, is required. The higher the percentage of the CPU time is used for overhead, the less resource can be devoted to other tasks and thus can cause slowdowns. Available CPU resources remaining after overhead can be represented as Equation 2.3, where the CPU clock is in megahertz.

$$\text{CPU clock} \times \frac{(\text{wedge-to-wedge time} - \text{overhead})}{\text{wedge-to-wedge time}} \times 100\% \tag{2.3}$$

Another key issue is that each drive may have different CPU clock and wedge-to-wedge time. It is totally invalid to compare the CPU utilization of different drives without normalizing this consideration.

2.2 Hardware Setup

The experimental data acquisition setup was a workstation with a RAID controller custom-built by Western Digital Corporation specifically for this study. Four different Western Digital hard disk drives were selected for investigation.

2.2.1 Workstation

The custom-built workstation was a Dell Precision Firmware Workstation T3400 that is powered by Intel Core 2 Quad Q6600 (2.40GHz/1066MHz/2X4MB L2)

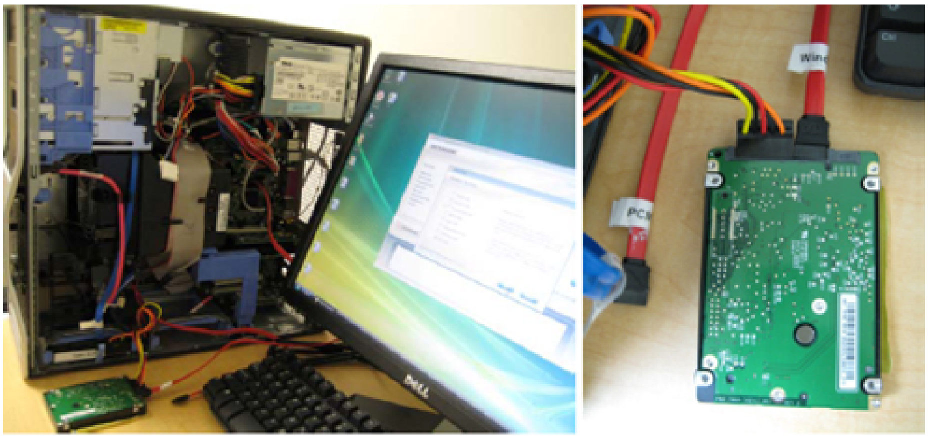

Figure 2.3: Customized workstation with RAID controller connected.

along with 2GB of 667MHz DDR2 SDRAM Memory and configured dual-boot with Windows Vista Business and Windows XP Business.

2.2.2 RAID Controller

The RAID controller was a Promise SuperTrak EX STEX8658 RAID Controller with 8 external ports that supports both SAS and SATA 3Gb/s drives and maintains configuration optimization for performance. The card incorporates a single chip solution for optimized reliability with 512MB data cache onboard, and an x8 PCI-Express interface to the host.

2.2.3 Hard Disk Drives

The hard disk drives under investigation in this study were known as Mars-RE, Viking, and Mercury. Atlantis-RE disk drive was used to confirm the Mars-RE models. The official product names and technical specifications of these hard disk drives are summarized in the following subsections.

Mars-RE

Mars-RE, model: WD1002FBYS has an official product name known as WD terabyte RAID Edition 3 (RE3) which is a third-generation RE disk drive. The RE3 is a 3.5-inch enterprise-class hard drive which uses the three 334GB platters backed by 32MB of cache, spins them at 7,200RPM, and shares a 3 Gb/s Serial ATA interface.

Atlantis-RE

Atlantis-RE, Model: WD5002ABYS is officially known as WD 500GB RAID Edition 3 (RE3) which is in the same product family of Western Digital's RAID Edition 3 as Mars-RE. The 500 GB version is also a 3.5-inch enterprise-class, 7,200 RPM drive, and shares a 3 Gb/s Serial ATA interface. However, it only comes with 16 MB of cache memory.

Mercury

Mercury has the model number: WD5000BEVT, which is known as WD Scorpio Blue for its official product name. This 2.5-inch, 500 GB Scorpio Blue that spins at 5400 RPM utilizes a SATA 3 Gb/s interface and comes with 8 MB of cache memory. Its low power consumption and cool operation in addition to its high performance making them ideal for notebooks and other portable devices.

Viking

Viking, model number: WD3000BLFS is officially known as WD VelociRaptor for its product name. Western Digital's standout 10,000 RPM and SATA 3 Gb/s interface Raptor family offers maximum speed, low power consumption, and cool operation. This WD VelociRaptor is available in a 2.5-inch form-factor, 300 GB and comes with 16 MB of cache memory for use in enterprise applications.

2.3 Software Setup

2.3.1 Windex

Windows Drive Exerciser (Windex) is a Windows-based, Western Digital proprietary test interface that enables the user to adjust the settings of performance parameters such as seek time and overhead by editing the header format file of a hard disk drive.

2.3.2 HD Tach

HD Tach is a Windows-based low-level benchmark tool used for testing the sequential read speed, the random access speed, interface burst speed, and CPU utilization of the random access read/write storage device attached. The data rate parameter investigated in this study is the sequential read speed measured at various points on the hard disk drive. The HD Tach sequential read test is a little bit different from other benchmarks. HD Tach reads from areas all over the hard drive and reports an average speed while most benchmarks create a file on the hard drive and test within that file.

2.3.3 PCMark

From the perspective of benchmarking software, the most influential applications regarding hard disk performance are I/O bound tasks which intensive reading from the hard disk drive and writing to it are involved. Multimedia editing applications which deal with large audio and video files are greatly influenced by the speed of the storage devices. On top of these, the starting up of an operating system and loading applications are also intensive I/O processes.

The performance characteristics measured during the experiment were standardized benchmarking software scores. Measurements were obtained using PCMark05 in Windows XP and PCMark Vantage in Windows Vista emulating various tasks, such as virus scanning, application loading, importing files, and media editing.

PCMark05 and PCMark Vantage Hard Disk Drive (HDD) Suite measures the hard disk drive performance based on the results of various tests with different workloads within the suite. The individual test scores are combined using a geometric mean. The geometric mean provides a fair mechanism to combine a large number of test results as compared to assigning arbitrary weights to individual scores. It is scaled using a multiplier based on performance results from reference hard disk drives to produce the appropriate range of scores. The general formula for the geometric mean is described in Equation 2.4.

$$\text{Geometric Mean} = (\text{Test1} \times \text{Test2} \times \text{Test3} \times \ldots)^{\frac{1}{\text{Number of Tests}}} \tag{2.4}$$

The formulas of PCMark05 and PCMark Vantage HDD Scores are described in Equation 2.5 and Equation 2.6 respectively.

$$\begin{aligned} \text{PCMark05 Score} &= 300 \times [\text{Geometric Mean of HDD Suite Test Results}] \\ &= 300 \times (\text{Windows XP Startup} \times \text{Application Load} \\ &\quad \times \text{General Usage} \times \text{Virus Scan} \times \text{File Write})^{\frac{1}{5}} \end{aligned} \tag{2.5}$$

$$\begin{aligned} \text{PCMark Vantage Score} &= 214.65 \times [\text{Geometric Mean of HDD Suite Test Results}] \\ &= 214.65 \times (\text{Windows Vista Startup} \times \text{Windows Defender} \\ &\quad \times \text{Windows Media Center} \times \text{Gaming} \times \text{Video Editing} \\ &\quad \times \text{Adding Music} \times \text{Importing Pictures} \\ &\quad \times \text{Application Loading})^{\frac{1}{8}} \end{aligned} \tag{2.6}$$

2.3.4 Minitab

Minitab is a commercially available statistical package that allows the user to generate a randomized experimental procedure for conducting the experiments and statistical manipulation. The software sets up a variety of multi-level designed experiments using the data set obtained from experiments and offers several analytical and graphing tools for analysis. Minitab helps us understand the results that can lead to potential improvement.

Chapter 3

Design of Experiments (DOE)

A three phase experimental study was performed in order to illustrate various aspects and benefits of the DOE methodology in the framework of benchmarked hard disk drive performance optimization. The first experiment concerned the experimental design and Analysis of Variance (ANOVA). The performances of three Western Digital hard disk drives: Mars-RE, Viking, and Mercury benchmarked using PCMark05 and PCMark Vantage as functions of performance parameters: data rate, seek time, and overhead were characterized using a two-level full factorial design. The second experiment dealt with the study of regression analysis which an equation was derived to describe the statistical relationship between factors: data rate, seek reads, seek writes, overhead and benchmarked Mars-RE disk performance and to predict new observations. A three-level full factorial design was applied. In the third case, the same experimental design was analyzed with additional replicates through Response Surface Methodology (RSM) for response optimization. The RSM models for Mars-RE disk drive performance benchmarked using both PCMark05 and PCMark Vantage were verified with Atlantis-RE disk drive, a prior generation hard drive of Mars-RE disk drive. It was reasonable to use Atlantis-RE disk drive for the confirmatory experiment since it is within the same model family as Mars-RE disk drive. The two hard disk drives therefore have the same technical specifications for most uncontrollable parameters that may have impacts on benchmarked performance.

3.1 Experiment 1

The objective of the Design of Experiments (DOE) application during the initial phase was supposed to be a screening experiment used to screen out the few significant main effects among all factors. However, the factors: short seek reads, overhead , and data rate were proposed by the engineers at Western Digital who participated in this study as already known significant to benchmarked disk performance in PCMark scores. This experiment is therefore focused on identifying the characteristics of short seek reads, overhead, and data rate and better appreciating their impacts on the disk performance of Mars-RE, Viking, and Mercury disk drives benchmarked using PCMark software.

3.1.1 Methodology

A full factorial design is used to investigate the effects of each factor and the interactions between them on a defined response by conducting all possible combinations of the factor levels. The factor settings are varied simultaneously rather than one-at-a-time in order to detect the important interactions between the factors.

In a two-level full factorial design, each factor has only two levels. The two levels are often referred as upper level and lower level and denoted as "+1" and "-1" respectively. The experimental runs therefore include 2^n combinations of factor levels for n factors. Even though a two-level full factorial design is only able to explore within a limited factor space, it indicates the direction of major trends with relatively few runs for small number of factors. This information is extremely useful for optimization in a wider region.

3.1.2 Experimental Design

The general procedure of DOE usually begins with experimental design, which defines the objective and selects controllable factors which the settings can be adjusted accordingly and are independent of other factors. Often the experiment also has to account for a number of uncontrollable factors which are held as close to constant as possible. The experiments are replicated in randomized order for execution to compensate the impact of uncontrollable factors on the response. The DOE approach is planned out in a way such that the changes in response can be observed and identified efficiently.

In this experiment, the three Western Digital hard disk drives under investigation were with aliases: Mars-RE, Viking, and Mercury. The effects of the three factors: short seek reads, overhead, and data rate were investigated on the response identified as disk performance benchmarked using PCMark software. The upper and lower levels adopted for the controllable factors were translated to their numerical values with corresponding units as indicated in Table 3.1 for each hard disk drive. The hard drives Mars-RE (1TB), Viking (300GB), and Mercury (500GB) were partitioned to 100GB, 20GB, and 25GB respectively for the test runs. The low level setting for short seek reads was determined based on its lowest default setting. The high level setting was then determined by taking the largest integer multiple of the low level setting which is less than or equal to 127 wedges.

Table 3.1: Factor levels for the two-level full factorial design.

Factor	Unit	Symbol	Mars-RE		Viking		Mercury	
Short Seek Reads	wedge	x_1	14	126	6	126	13	117
Overhead	μs	x_2	0	5	0	3	0	5
Data Rate	MB/s	x_3	65.00	112.00	82.30	122.40	41.75	80.00

This two-level full factorial design of three factors includes $2^3 = 8$ test runs. The test runs are carried out at upper and lower experimental levels denoted using + and - respectively as listed in Table 3.2, usually referred as a design matrix.

An interaction effect indicates that the effect of one factor is dependent upon the level of one or more other factors. The algebraic signs for the interaction effects are obtained by multiplying the corresponding factor columns row by row.

Table 3.2: Design matrix for the two-level full factorial design.

x_1: Short Seek Reads				x_2: Overhead		x_3: Data Rate	
Run	x_1	x_2	x_1x_2	x_3	x_1x_3	x_2x_3	$x_1x_2x_3$
1	-	-	+	-	+	+	-
2	-	-	+	+	-	-	+
3	-	+	-	-	+	-	+
4	-	+	-	+	-	+	-
5	+	-	-	-	-	+	+
6	+	-	-	+	+	-	-
7	+	+	+	-	-	-	-
8	+	+	+	+	+	+	+

The two-level full factorial design with three factors can be displayed geometrically as the cube shown in Figure 3.1 for a better understanding. Each axis represents the range of a factor, and the two ends indicate the high and low levels of this factor. The joint effects occur at the eight corners represent all unique combinations of factor levels in this design.

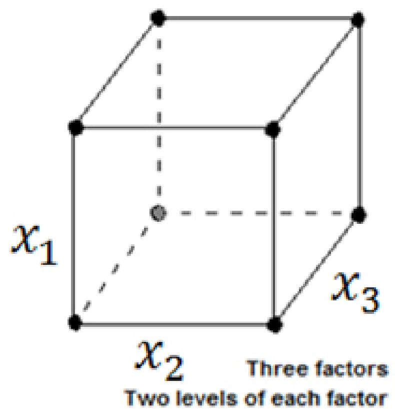

Figure 3.1: Design cube for the two-level full factorial design [Minitab Help, 2007].

3.1.3 Results and Discussion

The design was replicated three times, resulting in a total of 24 runs (i.e. 8 runs per replicate). The test runs were carried out in randomized order. Table 3.3-3.5 show the benchmarked disk performance in PCMark05 and PCMark Vantage scores for the 8 combinations of short seek reads, overhead, and data rate settings to be done in this two-level full factorial design for Mars-RE, Viking, and Mercury disk drive respectively. The experimental runs are listed in standard order as demonstrated in the design matrix in Table 3.2. For each combination of factor settings, the responses of the three replications are denoted as y_1, y_2, and y_3 respectively. $\bar{y}$ is the average and S is the standard deviation.

The experimental data in Table 3.3-3.5 were analyzed individually using Analysis of Variance (ANOVA) with a general linear model which includs terms up through order 3 (three-way interaction). The ANOVA is a commonly used analysis for statistical sensitivity which computes the contributions of individual factor effects and their interactions and determines which effects can be regarded as significant to the response. The name indicates that the analysis is based on estimates of variance within responses of all combinations of factor levels in the design.

The ANOVA decomposes the total variability in the experimental data into sources of variation along with their degrees of freedom, total sum of squares, mean squares, F-statistics, and P-values. These statistics were computed using Minitab and arranged in ANOVA tables as shown in Figure 3.2-3.7 for each experiment. The F-statistics and P-values are useful for determining whether an effect is significantly related to the response. Typically, the F-statistics are compared against the critical F found in F-distribution table and the P-values are compared against $\alpha = 0.05$. It is common to declare an effect statistically significant if the P-value is less than $\alpha = 0.05$ or the F-statistic is greater than the critical F. The smaller the P-value or the larger the F-statistic, the more significant is the effect.

Table 3.3: Benchmarked Mars-RE disk drive performance.

Mars-RE	PCMark05					Mars-RE	PCMark Vantage				
Run	y_1	y_2	y_3	$\bar{y}$	S	Run	y_1	y_2	y_3	$\bar{y}$	S
1	6373	6389	6345	6369	22.27	1	3810	3840	3821	3824	15.18
2	7419	7433	7423	7425	7.21	2	4542	4563	4550	4552	10.60
3	5931	5999	5942	5957	36.50	3	3585	3602	3592	3593	8.54
4	6760	6807	6842	6803	41.15	4	4244	4304	4308	4285	35.85
5	6008	6040	5947	5998	47.25	5	3661	3689	3690	3680	16.46
6	7024	6976	7037	7012	32.13	6	4357	4360	4390	4369	18.25
7	5691	5712	5760	5721	35.37	7	3490	3531	3512	3511	20.52
8	6544	6496	6583	6541	43.58	8	4127	4133	4182	4147	30.17

General Linear Model: PCMark05 versus ShortSeekReads, Overhead, DataRate

```
Factor          Type   Levels  Values
ShortSeekReads  fixed       2  14, 126
Overhead        fixed       2  0, 5
DataRate        fixed       2  65, 112


Analysis of Variance for PCMark05, using Adjusted SS for Tests

Source                           DF   Seq SS   Adj SS   Adj MS        F      P
ShortSeekReads                    1   616001   616001   616001   493.39  0.000
Overhead                          1  1191267  1191267  1191267   954.16  0.000
DataRate                          1  5233202  5233202  5233202  4191.59  0.000
ShortSeekReads*Overhead           1    30459    30459    30459    24.40  0.000
ShortSeekReads*DataRate           1     1717     1717     1717     1.38  0.258
Overhead*DataRate                 1    61307    61307    61307    49.10  0.000
ShortSeekReads*Overhead*DataRate  1      100      100      100     0.08  0.781
Error                            16    19976    19976     1249
Total                            23  7154030


S = 35.3341   R-Sq = 99.72%   R-Sq(adj) = 99.60%
```

Figure 3.2: ANOVA table for Mars-RE disk drive benchmarked using PCMark05.

Mars-RE

The ANOVA in Figure 3.2 indicates that the effect of data rate ($F = 4191.59$) contributes the most to the total variance and overhead ($F = 954.16$) is the second most significant effect. The influences of short seek reads and data rate interaction ($P = 0.258$) and short seek reads, overhead, and data rate interaction ($P = 0.781$) on Mars-RE disk performance benchmarked using PCMark05 are low and not statistically significant within the confined experimental domains.

The main effects plot and interaction plot are the graphical methods commonly used in conjunction with the ANOVA. A main effects plot connects the response mean for each factor level to visually investigate how the response is affected by different levels of a factor. The steeper slope represents the greater magnitude of the main effect. An interaction plot compares the relative strength of the effects across factors. Nonparallel lines on the interaction plots represent high interactions between the factor pair and indicate that the effect of one factor on response is dependent upon the level of another factor. The greater the difference in slope between the lines, the higher is the significance of interaction.

Figure 3.3 shows the individual effects of all factors on main effects plots. Short seek reads and overhead appear to affect Mars-RE disk performance benchmarked using PCMark05 with a negative correlation. Higher disk performance can be reached with smaller level settings of short seek reads and overhead. Data rate affects Mars-RE performance with a positive correlation. Higher disk performance is observed at higher level setting of data rate.

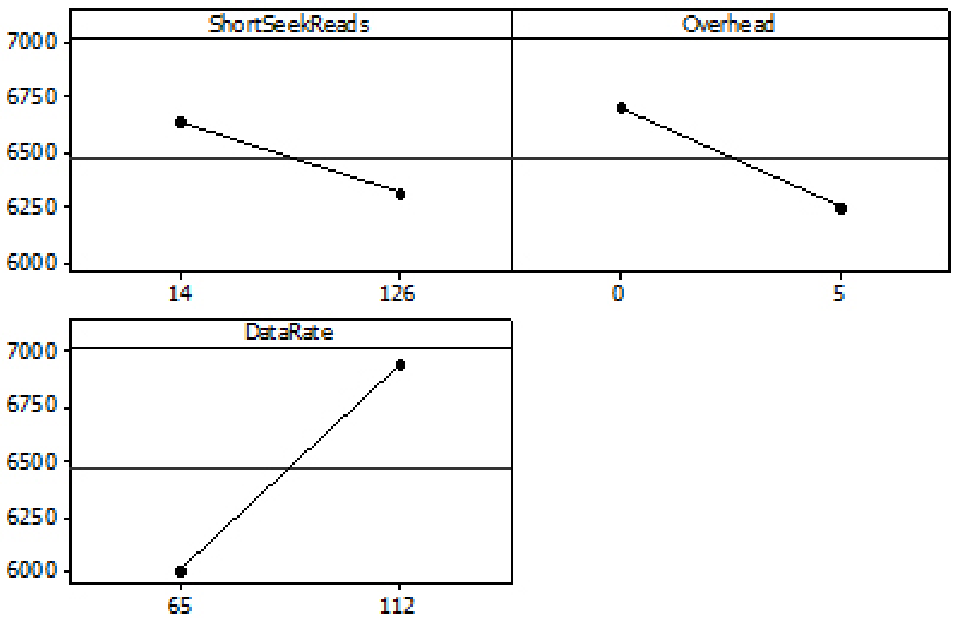

Figure 3.3: Main effects plots of Mars-RE disk drive benchmarked using PC-Mark05.

Figure 3.4 indicates significant interactions between short seek reads and overhead and between overhead and data rate. Parallel lines indicate no significant

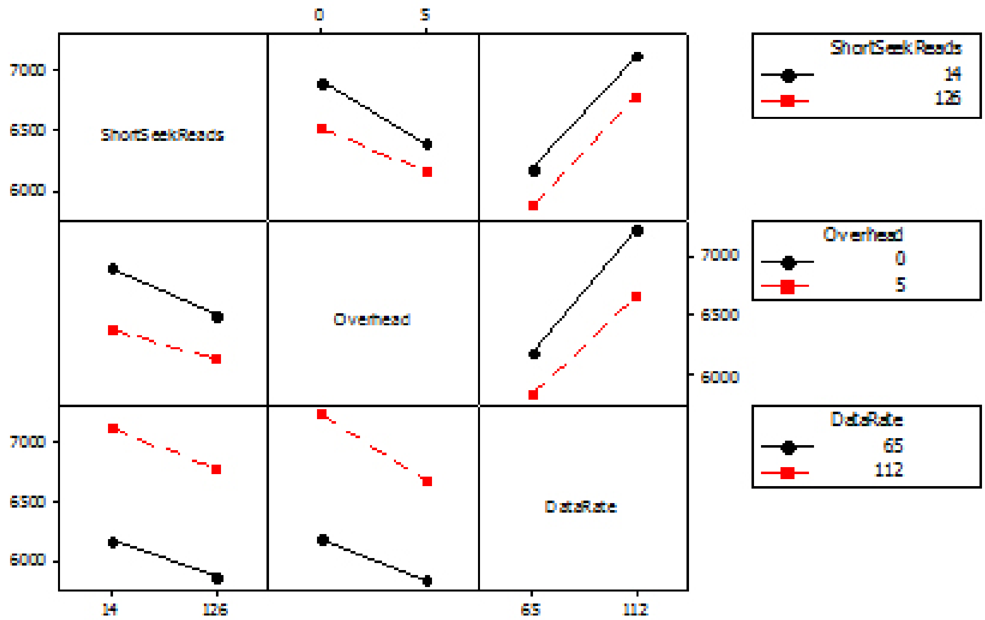

Figure 3.4: Interaction plots of Mars-RE disk drive benchmarked using PCMark05.

interaction between the short seek reads and data rate. As long as the factor levels are kept within the confined ranges, the lower level of short seek reads in wedges, lower level of overhead in microseconds, and higher level of data rate in megabytes per second are always the better choices for achieving higher benchmarked Mars-RE performance in PCMark05 scores.

Figure 3.5 shows that data rate ($F = 6219.47$) participates actively to the benchmarked Mars-RE disk performance in PCMark Vantage scores. Overhead ($F = 650.07$) is the second most significant effect. The interaction of short seek reads, overhead, and data rate ($P = 0.632$) is the only insignificant term on Mars-RE performance benchmarked using PCMark Vantage in the ANOVA table.

The individual effects of all factors in Figure 3.6 are similar to the main effects plot for Mars-RE benchmarked using PCMark05 in Figure 3.3 except that the benchmarked performance scoring scales are different. Similar conclusions on the correlations between each factor and response are reached.

Figure 3.7 shows that there are significant interactions between short seek reads and overhead, between short seek reads and data date, and between overhead and

General Linear Model: PCMarkVantage versus ShortSeekReads, Overhead, DataRate

Factor	Type	Levels	Values
ShortSeekReads	fixed	2	14, 126
Overhead	fixed	2	0, 5
DataRate	fixed	2	65, 112

Analysis of Variance for PCMarkVantage, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
ShortSeekReads	1	111930	111930	111930	246.25	0.000
Overhead	1	295482	295482	295482	650.07	0.000
DataRate	1	2827007	2827007	2827007	6219.47	0.000
ShortSeekReads*Overhead	1	4240	4240	4240	9.33	0.008
ShortSeekReads*DataRate	1	3384	3384	3384	7.45	0.015
Overhead*DataRate	1	2926	2926	2926	6.44	0.022
ShortSeekReads*Overhead*DataRate	1	108	108	108	0.24	0.632
Error	16	7273	7273	455		
Total	23	3252351				

S = 21.3200 R-Sq = 99.78% R-Sq(adj) = 99.68%

Figure 3.5: ANOVA table for Mars-RE disk drive benchmarked using PCMark Vantage.

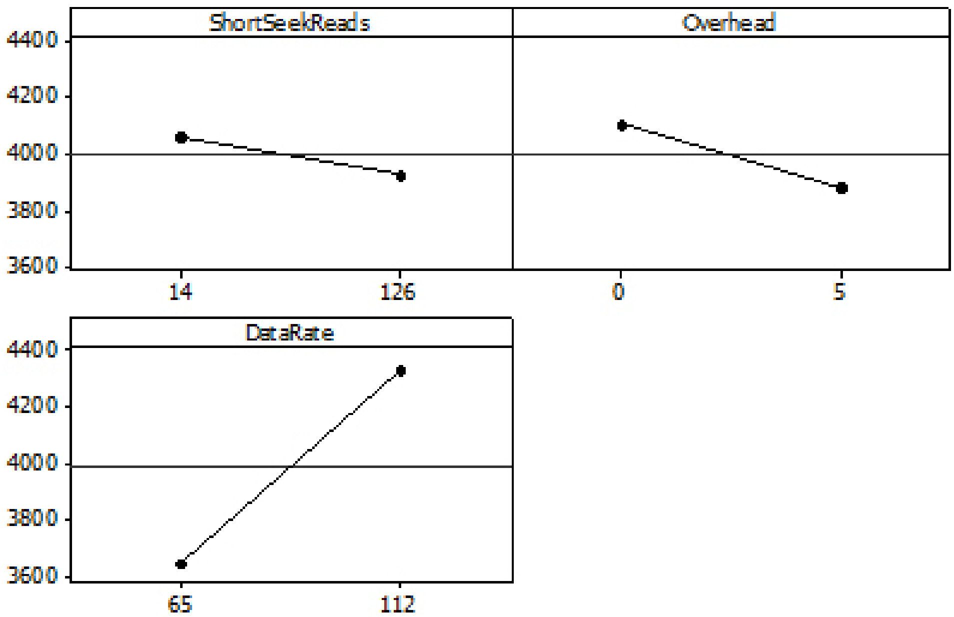

Figure 3.6: Main effects plots of Mars-RE disk drive benchmarked using PCMark Vantage.

data rate. Short seek reads and overhead both have greater effects at the low level than the high level; data rate has greater effect at its high level than low level.

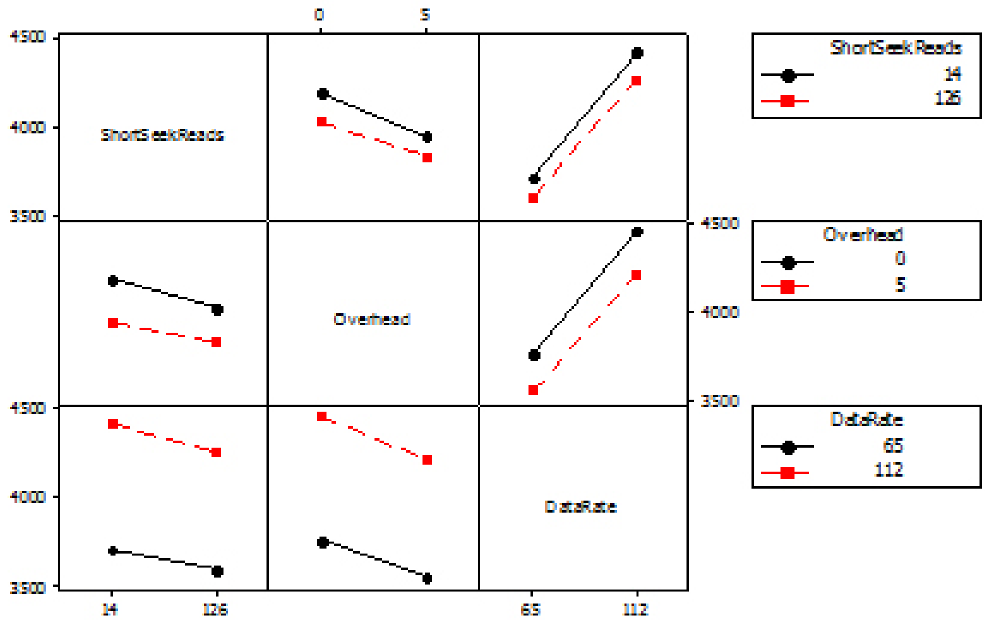

Figure 3.7: Interaction plots of Mars-RE disk drive benchmarked using PCMark Vantage.

Viking

Table 3.4: Benchmarked Viking disk drive performance.

Viking	PCMark05					Viking	PCMark Vantage				
Run	y_1	y_2	y_3	$\bar{y}$	S	Run	y_1	y_2	y_3	$\bar{y}$	S
1	8002	8062	8042	8035	30.55	1	5444	5501	5513	5486	36.86
2	8914	8936	9003	8951	46.36	2	6062	6027	6015	6034	24.42
3	7121	7065	7144	7110	40.63	3	4853	4897	4799	4849	49.08
4	7668	7714	7751	7711	41.58	4	5322	5334	5361	5339	19.97
5	7613	7620	7632	7621	9.61	5	5202	5210	5252	5221	26.86
6	8590	8524	8540	8551	34.43	6	5843	5857	5859	5853	8.72
7	6920	6929	6887	6912	22.11	7	4830	4771	4823	4808	32.23
8	7573	7601	7580	7584	14.57	8	5252	5273	5286	5270	17.16

Figure 3.8 shows the ANOVA of the factor effects and interactions on benchmarked Viking disk performance in PCMark05 scores. A significant factor effect of overhead ($F = 5245.05$) is notably found. Data rate ($F = 3457.33$) also contributes significantly to the total variance. There is no significant evidence for the effects of short seek reads and data rate interaction ($P = 0.126$) and short seek

reads, overhead, and data rate interaction ($P = 0.293$) on benchmarked Viking disk performance in PCMark05 scores.

General Linear Model: PCMark05 versus ShortSeekReads, Overhead, DataRate

Factor	Type	Levels	Values
ShortSeekReads	fixed	2	6, 126
Overhead	fixed	2	0, 3
DataRate	fixed	2	82.30, 122.38

Analysis of Variance for PCMark05, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
ShortSeekReads	1	485357	485357	485357	459.98	0.000
Overhead	1	5534401	5534401	5534401	5245.05	0.000
DataRate	1	3648060	3648060	3648060	3457.33	0.000
ShortSeekReads*Overhead	1	89670	89670	89670	84.98	0.000
ShortSeekReads*DataRate	1	2752	2752	2752	2.61	0.126
Overhead*DataRate	1	122551	122551	122551	116.14	0.000
ShortSeekReads*Overhead*DataRate	1	1247	1247	1247	1.18	0.293
Error	16	16883	16883	1055		
Total	23	9900922				

S = 32.4833   R-Sq = 99.83%   R-Sq(adj) = 99.75%

Figure 3.8: ANOVA table for Viking disk drive benchmarked using PCMark05.

The main effects plots on Figure 3.9 show very strong negative correlation between overhead and benchmarked performance of Viking disk drive in PCMark05 scores. This result confirms that overhead has the most significant factor effect among the three factors as indicated in the ANOVA table. The positive correlation between data rate and benchmarked performance of Viking disk drive is also strong; short seek reads appears to influence the disk performance with decent negative correlation.

It is easy to observe from Figure 3.10 that when short seek reads setting is lower, its interaction with overhead is larger and a lower overhead setting has higher interaction with data rate. Parallel lines indicate that there is no statistically significant interaction between short seek reads and data rate on performance of Viking disk drive benchmarked using PCMark05.

The ANOVA in Figure 3.11 indicates that all three factor effects are found significant to benchmarked Viking disk performance in PCMark Vantage scores especially the effects of overhead ($F = 2358.28$) and data rate ($F = 1977.90$).

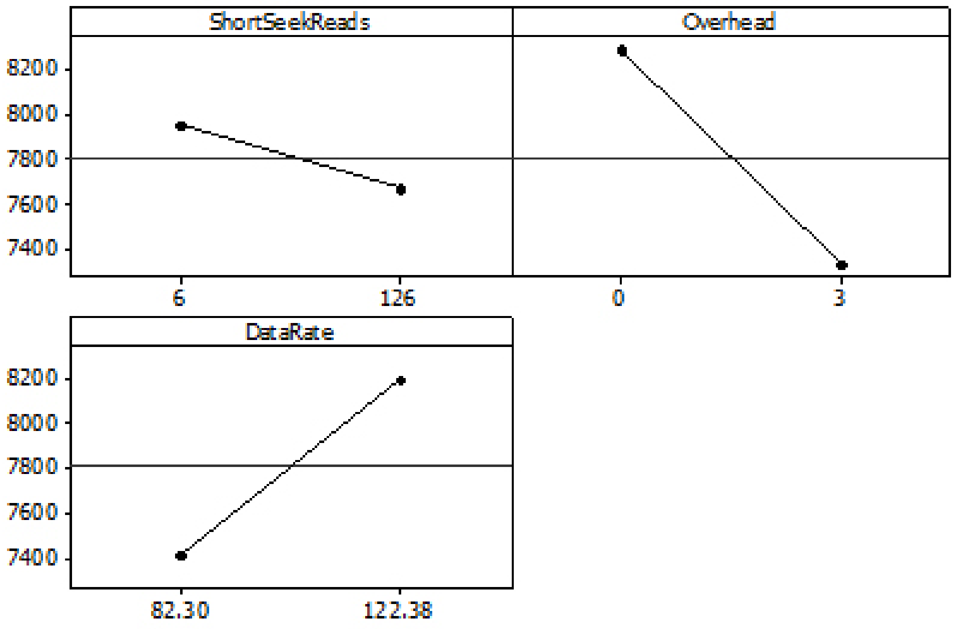

Figure 3.9: Main effects plots of Viking disk drive benchmarked using PCMark05.

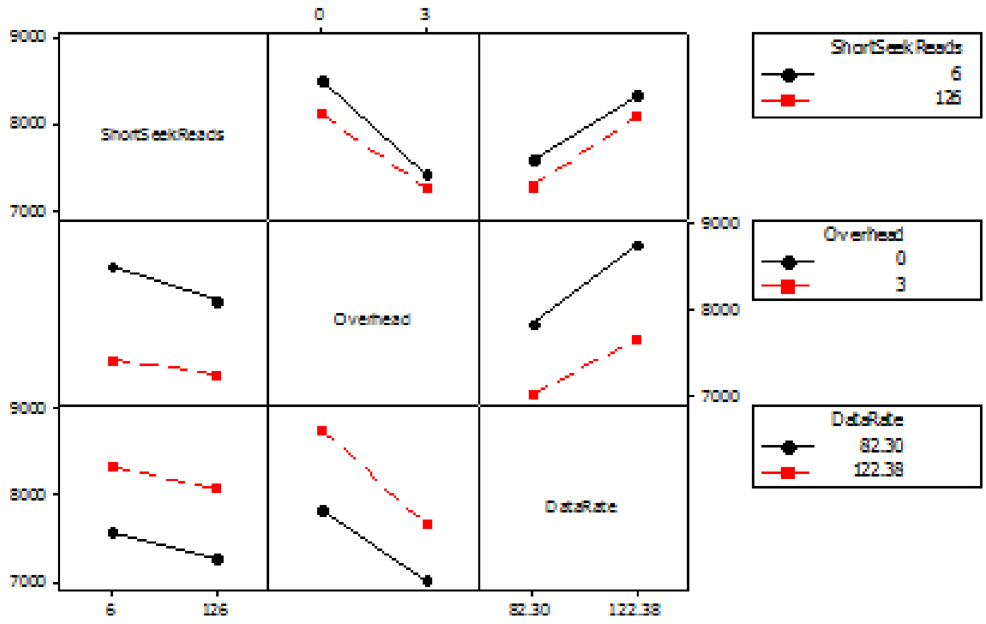

Figure 3.10: Interaction plots of Viking disk drive benchmarked using PCMark05.

Within the factor domains, the interaction between short seek reads and data rate ($P = 0.260$) is the only insignificant term to the performance of Viking disk drive benchmarked using PCMark Vantage.

Main effects plots in Figure 3.12 concludes that the correlations between each

General Linear Model: PCMarkVantage versus ShortSeekReads, Overhead, DataRate

Factor	Type	Levels	Values
ShortSeekReads	fixed	2	6, 126
Overhead	fixed	2	0, 3
DataRate	fixed	2	82.30, 122.38

Analysis of Variance for PCMarkVantage, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
ShortSeekReads	1	116204	116204	116204	134.84	0.000
Overhead	1	2032344	2032344	2032344	2358.28	0.000
DataRate	1	1704534	1704534	1704534	1977.90	0.000
ShortSeekReads*Overhead	1	42336	42336	42336	49.13	0.000
ShortSeekReads*DataRate	1	1176	1176	1176	1.36	0.260
Overhead*DataRate	1	19608	19608	19608	22.75	0.000
ShortSeekReads*Overhead*DataRate	1	4537	4537	4537	5.27	0.036
Error	16	13789	13789	862		
Total	23	3934529				

S = 29.3563 R-Sq = 99.65% R-Sq(adj) = 99.50%

Figure 3.11: ANOVA table for Viking disk drive benchmarked using PCMark Vantage.

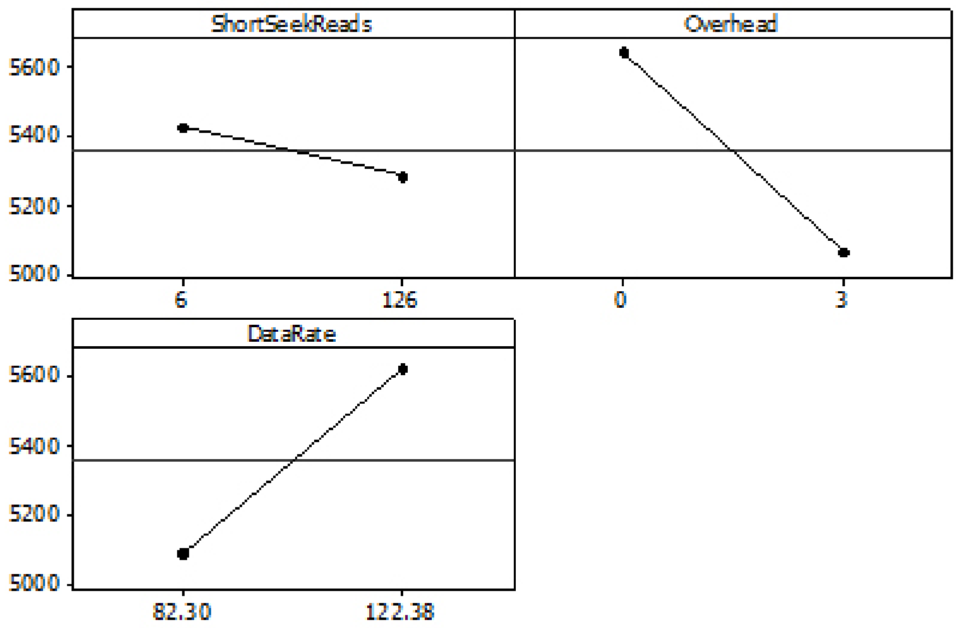

Figure 3.12: Main effects plots of Viking disk drive benchmarked using PCMark Vantage.

factor and benchmarked performance of Viking disk drive in PCMark Vantage scores are similar to the results from Figure 3.9, which the Viking disk performance was benchmarked using PCMark05.

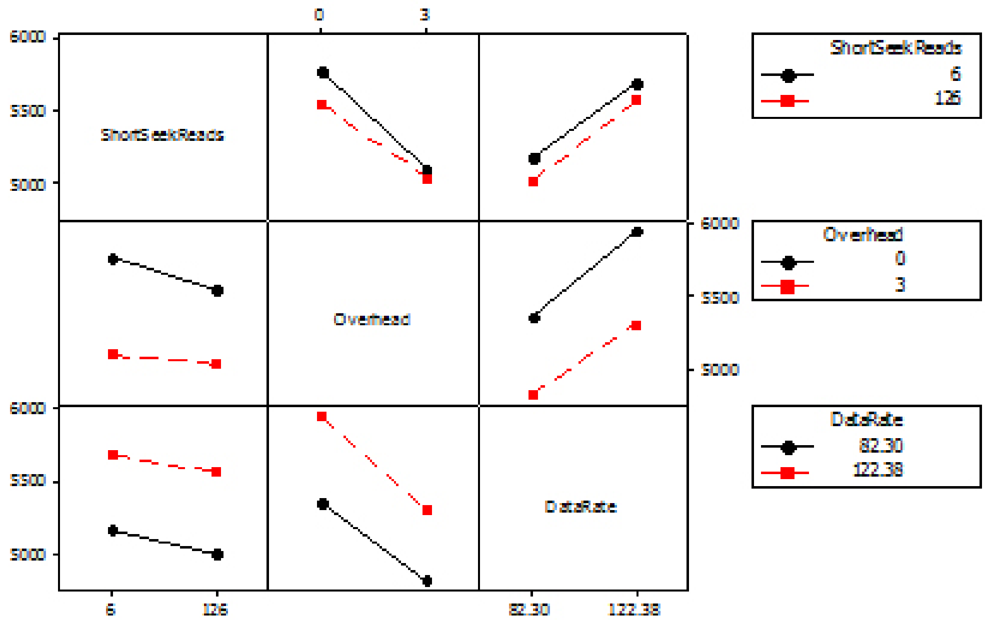

Figure 3.13: Interaction plots of Viking disk drive benchmarked using PCMark Vantage.

The interaction between short seek reads and data rate is again found not statistically significant on benchmarked performance of Viking disk drive in PCMark Vantage scores on interaction plots in Figure 3.13. Lower short seek reads level leads to larger interaction effect with overhead and lower overhead level has a larger interaction effect with data rate.

Mercury

Figure 3.14 shows that the factor effects of data rate ($F = 6760.30$) has dominant influence to benchmarked performance of Mercury disk drive in PCMark05 scores. Short seek reads ($F = 281.92$) and overhead ($F = 79.16$) both also participate the contributions to the total variance; however, their factor effects are nowhere closer to data rate's. The only interaction found statistically significant to Mercury disk performance is between short seek reads and data rate ($P = 0.008$).

The Main effects plots in Figure 3.15 confirms that data rate is the most dominant factor effect among all three factors to benchmarked Mercury disk performance in PCMark05 scores by showing that data rate has the greatest magnitude

Table 3.5: Benchmarked Mercury disk drive performance.

Mercury	PCMark05					Mercury	PCMark Vantage				
Run	y_1	y_2	y_3	$\bar{y}$	S	Run	y_1	y_2	y_3	$\bar{y}$	S
1	4192	4200	4210	4200	9.02	1	2391	2389	2405	2395	8.72
2	5028	5035	4998	5020	19.66	2	2662	2690	2682	2678	14.42
3	4098	4117	4115	4110	10.44	3	2338	2340	2358	2345	11.02
4	4960	4917	4918	4931	24.54	4	2695	2708	2622	2675	46.36
5	4048	4062	4054	4054	7.02	5	2349	2334	2348	2343	8.39
6	4867	4784	4854	4835	44.64	6	2663	2641	2640	2648	13.00
7	3978	3997	3997	3990	10.97	7	2280	2299	2309	2296	14.73
8	4764	4699	4744	4735	33.29	8	2608	2644	2568	2606	38.02

General Linear Model: PCMark05 versus ShortSeekReads, Overhead, DataRate

```
Factor          Type   Levels  Values
ShortSeekReads  fixed       2  13, 117
Overhead        fixed       2  0, 5
DataRate        fixed       2  41.75, 80.00


Analysis of Variance for PCMark05, using Adjusted SS for Tests

Source                          DF   Seq SS   Adj SS   Adj MS        F      P
ShortSeekReads                   1   156817   156817   156817   281.92  0.000
Overhead                         1    44033    44033    44033    79.16  0.000
DataRate                         1  3760417  3760417  3760417  6760.30  0.000
ShortSeekReads*Overhead          1       96       96       96     0.17  0.683
ShortSeekReads*DataRate          1     5046     5046     5046     9.07  0.008
Overhead*DataRate                1      417      417      417     0.75  0.400
ShortSeekReads*Overhead*DataRate 1      523      523      523     0.94  0.347
Error                           16     8900     8900      556
Total                           23  3976247


S = 23.5850   R-Sq = 99.78%   R-Sq(adj) = 99.68%
```

Figure 3.14: ANOVA table for Mercury disk drive benchmarked using PCMark05.

of main effect. The performance of Mercury disk drive benchmarked using PCMark05 increases significantly when data rate moves from its low level to high level. Both short seek reads and overhead at their low levels have higher disk performance means than at their high levels.

In Figure 3.16, the interaction plots of short seek reads versus data rate ambiguously show nonparallel lines which indicate a slightly significant interaction to benchmarked disk performance of Mercury in PCMark05 scores. The interactions between short seek reads and overhead and between overhead and data rate are

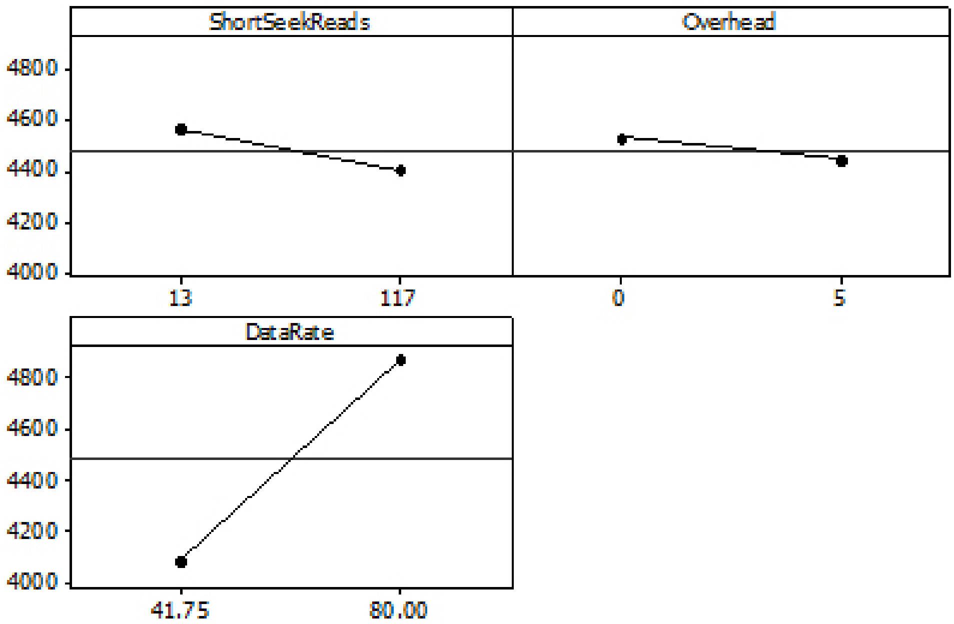

Figure 3.15: Main effects plots of Mercury disk drive benchmarked using PC-Mark05.

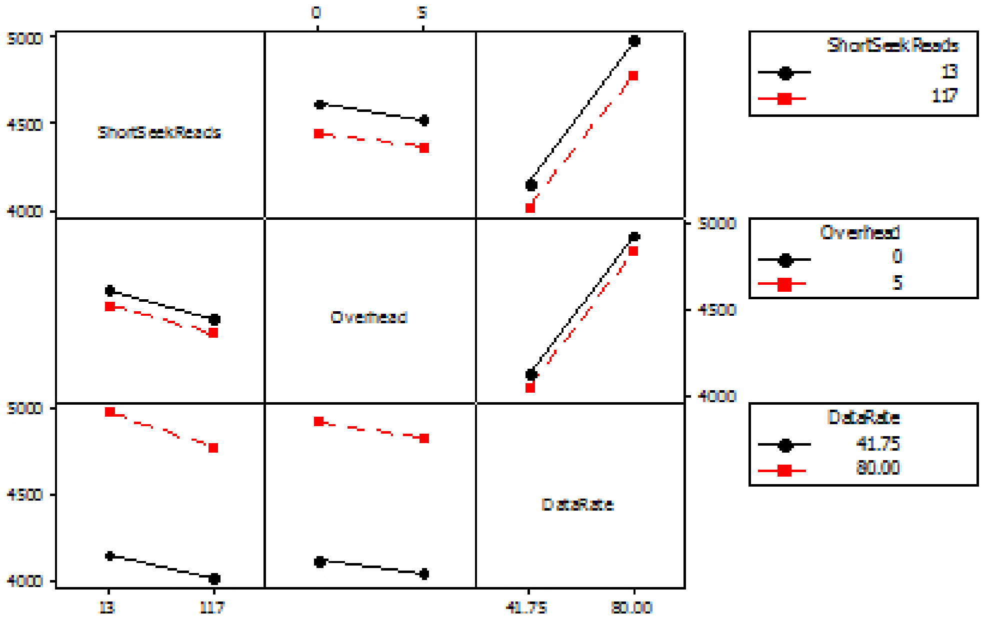

Figure 3.16: Interaction plots of Mercury disk drive benchmarked using PC-Mark05.

not significant as the two lines remain parallel over all confined factor levels.

The ANOVA of benchmarked performance of Mercury disk drive in PCMark Vantage scores in Figure 3.17 shows similar results as benchmarked using PC-

General Linear Model: PCMarkVantage versus ShortSeekReads, Overhead, DataRate

Factor	Type	Levels	Values
ShortSeekReads	fixed	2	13, 117
Overhead	fixed	2	0, 5
DataRate	fixed	2	41.75, 80.00

Analysis of Variance for PCMarkVantage, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
ShortSeekReads	1	14850	14850	14850	26.66	0.000
Overhead	1	7526	7526	7526	13.51	0.002
DataRate	1	565187	565187	565187	1014.70	0.000
ShortSeekReads*Overhead	1	495	495	495	0.89	0.360
ShortSeekReads*DataRate	1	2	2	2	0.00	0.952
Overhead*DataRate	1	1053	1053	1053	1.89	0.188
ShortSeekReads*Overhead*DataRate	1	610	610	610	1.10	0.311
Error	16	8912	8912	557		
Total	23	598636				

S = 23.6008   R-Sq = 98.51%   R-Sq(adj) = 97.86%

Figure 3.17: ANOVA table for Mercury disk drive benchmarked using PCMark Vantage.

Mark05 in Figure 3.14. The F-statistics of data rate ($F = 1014.70$) indicates that it has extremely strong effect to benchmarked disk performance of Mercury in PCMark Vantage scores while factor effects of short seek reads ($F = 26.66$) and overhead ($F = 13.51$) are also significant but contribute very little to the total variance. There is no significant evidence for any of the interactions within the confined factor ranges as all the corresponding P-values are greater than $\alpha = 0.05$.

Main effects plots in Figure 3.18 show strong influence to benchmarked performance of Mercury disk drive in PCMark Vantage scores from data rate. Data rate at its high level has a higher disk performance mean than at its low level, which indicates a positive correlation between data rate and benchmarked disk performance. On the other hand, benchmarked Mercury disk performance increases when either short seek reads or overhead moves from its high level to low level.

Figure 3.19 shows that there are significant interactions between short seek reads and overhead and between overhead and data rate. However, the factor effect of data rate alone to disk performance of Mercury benchmarked using PCMark Vantage is extremely dominant as indicated in the ANOVA table. None of the

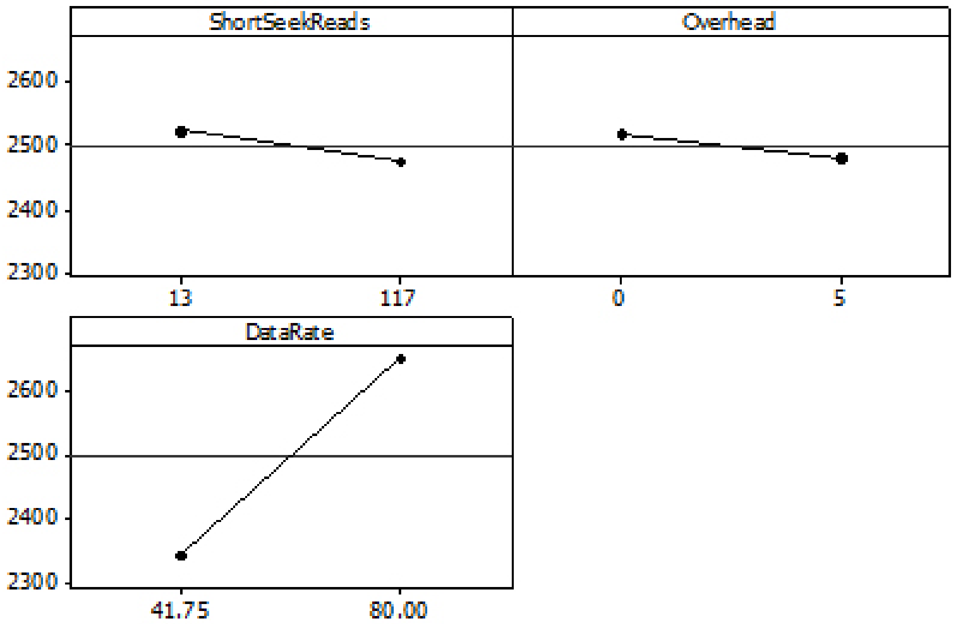

Figure 3.18: Main effects plots of Mercury disk drive benchmarked using PCMark Vantage.

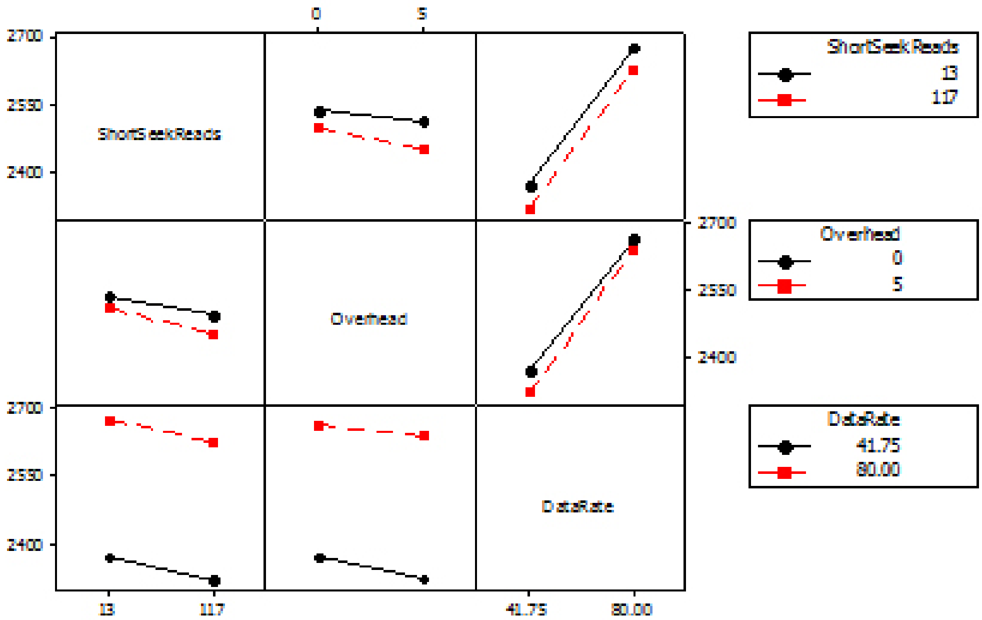

Figure 3.19: Interaction plots of Mercury disk drive benchmarked using PCMark Vantage.

interactions contributes enough to total variance to be considered significant to disk performance.

The two-level full factorial design with three factors was analyzed with the

ANOVA using Minitab. The total variability in the experimental data was decomposed into three factor effects (x_1, x_2, x_3), three two-way interactions (x_1x_2, x_1x_3, x_2x_3), and one three-way interaction ($x_1x_2x_3$). The F-statistic and its probability value (P-value) provide useful information for determining the statistical significance of all given effects to the benchmarked disk performance. Main effects plot and interaction plot help visually judge the presence of main effects and interactions over confined factor levels and compare their strength. The information provided by the DOE techniques quantified some common sense to make well-founded assumptions for further investigation. The analysis results of Mars-RE, Viking, and Mercury disk drives are summarized in Table 3.6 for comparison.

Table 3.6: Summary of ANOVA results in Experiment 1.

	x_1: Short Seek Reads	x_2: Overhead	x_3: Data Rate
Disk Drive	Benchmarking Software	Significant Main Effect(s)	Significant Interaction Effect(s)
Mars-RE	PCMark05	x_1, x_2, x_3	x_1x_2, x_2x_3
	PCMark Vantage	x_1, x_2, x_3	x_1x_2, x_1x_3, x_2x_3
Viking	PCMark05	x_1, x_2, x_3	x_1x_2, x_2x_3
	PCMark Vantage	x_1, x_2, x_3	x_1x_2, x_2x_3, $x_1x_2x_3$
Mercury	PCMark05	x_1, x_2, x_3	x_1x_3
	PCMark Vantage	x_1, x_2, x_3	None

The significant interactions between factors identified in this experiment indicate that the change in benchmarked disk performance as the setting of one factor moves between its low and high levels is dependent upon the level of one of more other factors. However, different benchmarking software evaluates the performance of the same disk drive with different tests which may be influenced specifically by certain interactions between factors.

The benchmarked disk performance of Mars-RE, Viking, and Mercury disk drives cannot be compared with each other due to the fact that they are all in different model families. Impacts on benchmarked disk performance due to uncontrollable parameters such as clock speed, revolutions per minute, architecture, etc., are usually different from one model family to another. The significance of the same interaction onto the performance of different hard disk drive would therefore be different even though they are evaluated with the same benchmarking software.

3.2 Experiment 2

In Section 3.1, the two-level full factorial design provided quick identification of major trends within a small number of experimental runs. The ANOVA evaluated the significant contributions of the factor effects and their interactions to benchmarked hard disk drive performance. The objective of Experiment 2 was to describe the statistical relationship between the factors and a response by a mathematical equation for new observation predictions through regression analysis.

Rather than continuing to use the original three factors proposed by engineers at Western Digital Corporation, the interrelationship between seek time subfactors: short seek reads, long seek reads, short seek writes, and long seek writes was studied from engineering aspects for a better understanding as described in Section 2.1.2. Since the partitioned hard disk drive capacity also determines which distance group (zone) is covered, it is reasonable to consider short seek reads (first three zones) and long seek reads (next five zones) as a single factor, seek reads, by setting all eight zones with the same level. Similar settings are applied for seek writes.

Mars-RE, known as a mainstream desktop disk drive manufactured by Western Digital Corporation was the only hard disk drive studied during this experimental phase and throughout the rest of this research due to time constraint. A three-level full factorial design was used to obtain more information about the main effects of newly defined factors.

3.2.1 Methodology

The data set observed from a full factorial design can be described in terms of a functional relationship between the factors and a response by fitting a regression model. The mathematical form of a multiple linear regression model, which depends on n independent factors x_1, x_2, ..., x_n, can be expressed as in Equation 3.1, where $\hat{y}$ is a single, dependent response prediction and the regression coefficient. β_0, is the y-intercept of the regression model and is defined as the grand average of all response observations and β_i's $(0 < i \leq n)$ are linear coefficients correspond to the expected change in response y per unit change in x_i, when all the remaining independent variables are held constant.

$$\hat{y} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i \tag{3.1}$$

Fitting a linear multiple regression model essentially involves using the experimental observations to estimate the regression coefficients such that the critical terms can be determined and included in the final model. Lower-order regression models are often used to provide good approximations of the relationship between factors and response.

A linear regression model shows a steady rate of increase or decrease in the response observations. Regression generally uses the least squares method which derives the equation for model fitting by minimizing the sum of the square of the residuals. A residual represents the difference between an observed response y and its corresponding fitted value $\hat{y}$ as in Equation 3.2.

$$e = y - \hat{y} \tag{3.2}$$

3.2.2 Experimental Design

During this experimental phase, the factors under investigation were data rate, seek reads, seek writes, and overhead. The four factors are designated as x_1, x_2,

x_3, and x_4 respectively, each has the low, middle, and high levels as indicated in Table 3.7 for the three-level full factorial design. Unit conversions were performed for a more meaningful interpretation of results. The units of seek reads and seek writes are converted from wedges to microseconds by multiplying a specified Mars-RE wedge-to-wedge time of 28.82 μs/wedge. The unit of overhead is converted from microseconds to megahertz as described in Equation 2.3.

Table 3.7: Factor levels for the three-level full factorial design.

Factor	Unit	Symbol	Level		
Data Rate	MB/s	x_1	65	80	110
Seek Reads	μs	x_2	144.1	1729.2	3458.4
Seek Writes	μs	x_3	144.1	1729.2	3458.4
Overhead	MHz	x_4	75	96	126

The three-level full factorial design consists of $3^4 = 81$ possible factor combinations for the four factors each with three levels. Each treatment was run three times in randomized order. A total of $81 \times 3 = 243$ test runs were included in the full factorial design for Equation 3.3.

The regression analysis indicates the significance and direction of the statistical relationship between the factors and response by estimating the regression coefficients of the fitted model through the T-statistics and their P-values. A P-value can be found from Student's T-distribution table and compared against $\alpha = 0.05$ after the T-statistic is calculated.

3.2.3 Results and Discussion

The statistical relationship between the four factors and the response is derived by substituting the value of 4 for n in Equation 3.1 since there are four parameters under investigation in this experiment. Equation 3.3 predicts the Mars-RE disk performance benchmarked using PCMark05 given specified factor settings of data rate, seek reads, seek writs, and overhead.

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 \tag{3.3}$$

The regression analysis relates the overall mean effect (β_0) and the effects of the factors (β_i's, $0 < i \leq n$) at different levels. The integration of the linear multiple regression model in Equation 3.4 gives an approximation of the response based on the experimental results.

$$\hat{y} = 3557 + 21.1x_1 - 0.157x_2 - 0.039x_3 + 8.75x_4 \tag{3.4}$$

Figure 3.20 summarizes the significance of each regression coefficient in the linear multiple regression model in Equation 3.4. The low probability values ($P = 0.000$'s) of all linear effects indicate their significances to Mars-RE disk performance benchmarked using PCMark05. The application of T-test determines that data rate ($T = 49.26$) has the most dominant effect and seek writes ($T = -6.64$) has the least effect among the four parameters to Mars-RE disk performance benchmarked using PCMark05.

The adjusted coefficient of multiple determination (R^2_{Adj}) is the percentage of response variable variation, adjusted for the number of factors in the model. This measure is explained by its correlation with one or more factors for testing the goodness of fit of the regression equation. The R^2_{Adj} of 93.9% in Figure 3.20 indicates how close the data points will fall along the fitted regression line. The four linear effects account for 93.9% of the variance of Mars-RE disk performance benchmarked using PCMark05 by this linear multiple regression model.

The sign of each regression coefficient indicates the direction of the relationship between the factors and response. Positive regression coefficients of data rate and overhead suggest increases in these factors for achieving maximal Mars-RE disk performance benchmarked using PCMark05. On the other hand, seek reads and seek writes should follow the opposite direction.

The regression coefficients represent the mean change in a response per unit

Regression Analysis: Score versus DataRate, SeekReads, SeekWrites, Overhead

Predictor	Coef	SE Coef	T	P
Constant	3556.61	55.14	64.50	0.000
DataRate	21.0728	0.4278	49.26	0.000
SeekReads	-0.157100	0.005913	-26.57	0.000
SeekWrites	-0.039236	0.005913	-6.64	0.000
Overhead	8.7524	0.3824	22.89	0.000

S = 124.760   R-Sq = 94.0%   R-Sq(adj) = 93.9%

Figure 3.20: Regression analysis for Mars-RE disk drive benchmarked using PCMark05.

change in a factor while holding the other factors constant. Specifically, Mars-RE disk performance benchmarked in PCMark05 scores was expected to increase by 21.073 per megabyte increase in data rate, to decrease by 0.157 per microsecond increase in seek reads, to decrease by 0.039 per microsecond increase in seek reads, and to increase by 8.752 per megahertz increase in overhead. The adequacy and significance of this model fit to the experimental data is tested in the form of Analysis of Variance (ANOVA) summarized in Figure 3.21.

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	57592682	14398170	925.03	0.000
Residual Error	238	3704500	15565		
Total	242	61297181			

Figure 3.21: ANOVA table for Mars-RE disk drive benchmarked using PCMark05.

This multiple regression model is found highly significant, according to the low probability value of $P = 0.000$. This result suggests that the factors identified in this model are in fact of significant importance to the Mars-RE disk performance benchmarked using PCMark05. The model is concluded to be adequate for response prediction within the factor domains employed.

3.3 Experiment 3 - PCMark05

A linear multiple regression model that predicts Mars-RE disk performance benchmarked using PCMark05 as a function of relevant controllable factors: data rate, seek reads, seek writes, and overhead was obtained in Section 3.2. The Response Surface Methodology (RSM) is a collective use of Design of Experiments (DOE) techniques, Analysis of Variance (ANOVA), and regression models employed during this experimental phase to optimize the derived model equation for gaining precise operating conditions of all significant factors within the observation space for maximum benchmarked Mars-RE disk drive performance in PCMark scores.

3.3.1 Methodology

The Response Surface Methodology (RSM) involves a combination of computation and visualization for optimizing a response in a desired direction by iteratively adjusting parameter settings. The response surface model is generalized as in Equation 3.5, where the notations remain the same as the linear multiple regression model described in Section 3.2.

$$\hat{y} = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i<j}^{n} \beta_{ij} x_i x_j + \sum_{i=1}^{n} \beta_{ii} x_i^2 \tag{3.5}$$

This frequently used quadratic model consists of the linear regression model in the first half of the equation and additional interactions and squares in the second half of the equation. Higher-order terms would follow if necessary. The response surface of a model with linear terms alone, such as the linear multiple regression model derived in Equation 3.4, represents a two-dimensional plane within the three-dimensional factor space. Additional interactions and higher-order terms describe the local shape of the response surface such that the interactions allow for warping of the plane and the squares indicate an optimal response that is either the maximum or minimum on the response surface. An adequate response surface model helps optimize the experimental design with much ease through statistical and graphical analyses.

The results of the linear multiple regression analysis in Section 3.2 showed that all the linear effects of the parameters are significant to Mars-RE disk performance benchmarked using PCMark05.

A three-level full factorial design and a general full factorial design were selected with the four factors proposed in Section 3.2: data rate, seek reads, seek writes, and overhead. Even though a full factorial design might not be the best design choice for the application of RSM, it was a more suitable design specifically for this study due to the fact that the levels of proposed factors were with associated constraints such as the adjustment step sizes and unit conversions that limited level selections.

3.3.2 Experimental Design

Variable designations from Section 3.2 continued to be used during this experimental phase. The level selections for the applications of RSM remained the same as for the linear multiple regression analysis discussed in Section 3.2. The low, middle, and high levels of each controllable factor were indicated in Table 3.8. The application of RSM was planned and conducted with two additional replications for each test run in randomized order.

Table 3.8: Factor levels for the three-level full factorial design (PCMark05).

Factor	Unit	Symbol	Level		
Data Rate	MB/s	x_1	65	80	110
Seek Reads	μs	x_2	144.1	1729.2	3458.4
Seek Writes	μs	x_3	144.1	1729.2	3458.4
Overhead	MHz	x_4	75	96	126

The three-level full factorial design conducted with four controllable factors has $3^4 = 81$ combinations of factor settings each with five replicates. A total of 405 (i.e. 81 runs per replicate) runs were conducted in this experiment.

The response observations benchmarked under different factor conditions were first fitted to a linear model which was the same as the linear regression model obtained in Section 3.2. However, higher precision was expected since two more replicates were taken into account during this experimental phase.

The adequacy of the fitted linear response surface model can be verified by examining the least squares assumptions in the ANOVA and regression analysis through residual plots, which are commonly used for examining the goodness of model fit.

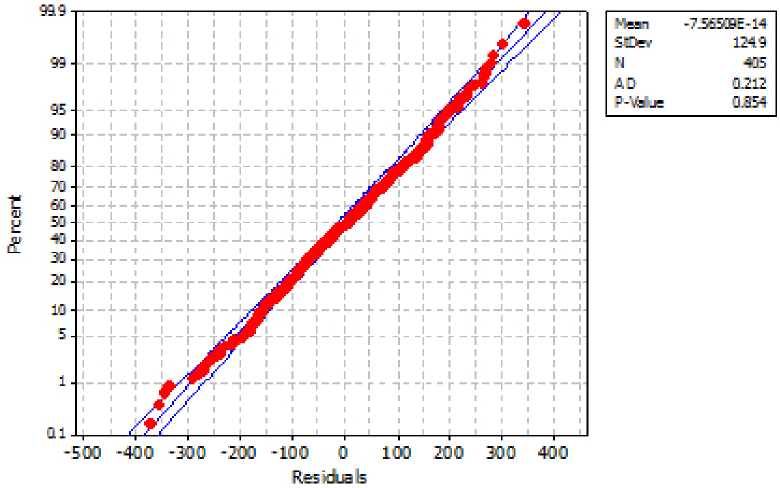

Figure 3.22: Normal probability plot of residuals of the Mars-RE linear response surface model (PCMark05).

The three assumptions to be fulfilled are: normal distribution, independence, and constant variance of the residuals. A normal probability plot can be used to detect the normality by verifying how the points fall along the distribution line and between the 95% confidence intervals as shown in Figure 3.22. A summary table with distribution parameter estimates along with the Anderson-Darling (AD) statistic and P-value helps evaluate the distribution fit statistically. The practical interpretation of the small Anderson-Darling statistic and P-value over $\alpha = 0.05$, statistically confirmed that the model fitted the experimental data set adequately. The residuals on a residuals versus order plot should exhibit no clear pattern to detect time-independence of residuals as shown in Figure 3.23. A residuals versus fitted values plot should show no recognizable pattern of residuals on both sides of zero to detect constant variance as shown in Figure 3.24.

Since the three statistical assumptions are all satisfied for the linear response surface model fit to the data, this model is expected to produce unbiased coefficient estimates with minimum variance. A response surface model with higher degrees

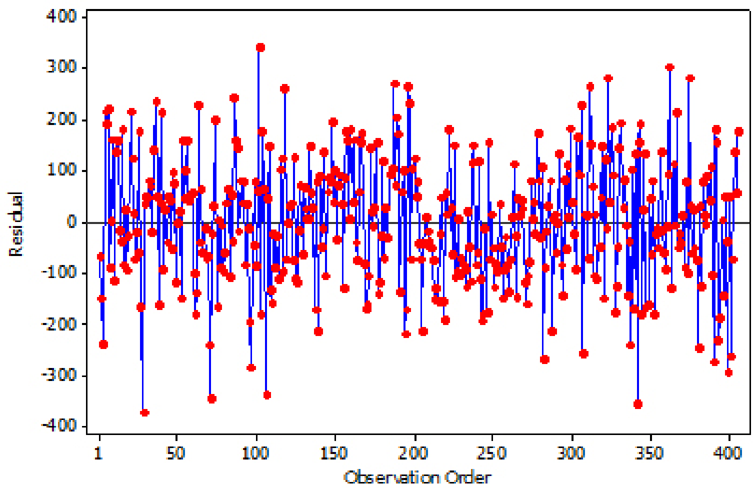

Figure 3.23: Residuals versus order plot of the Mars-RE linear response surface model (PCMark05).

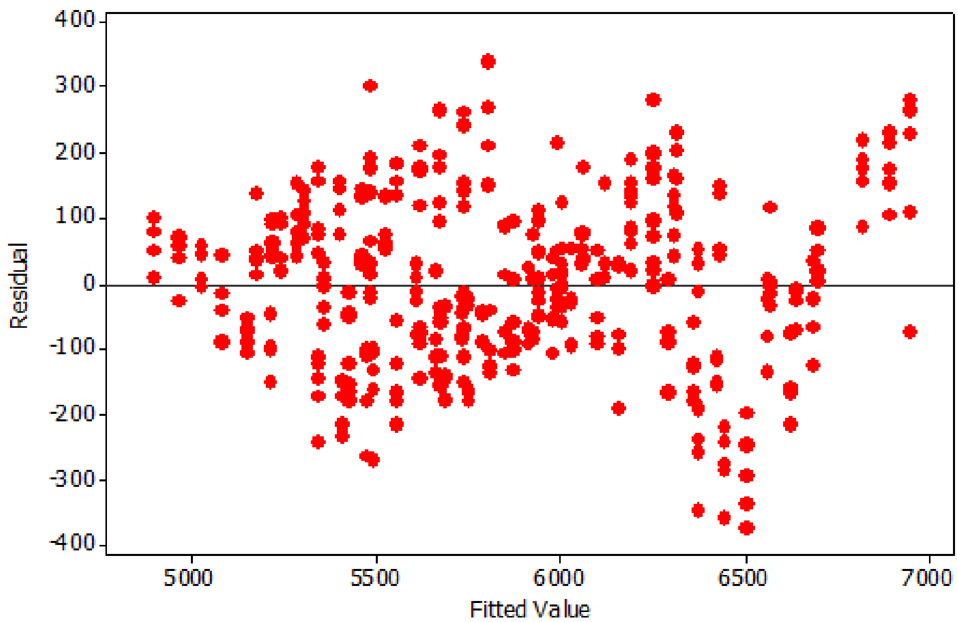

Figure 3.24: Residuals versus fitted values plot of the Mars-RE linear response surface model (PCMark05).

or additional interactions may not necessarily fit the data set better than this linear response surface model; therefore it is futile to seek a more complex model which may end up with infinite number of inadequate solutions. It is usually the best to choose a model with the lowest possible degree to keep it as simple as possible but also without being under-specified to avoid misleading conclusions.

3.3.3 Results and Discussion

The linear response surface model consists of only the linear terms of Equation 3.5, where $\hat{y}$ is the predicted response benchmarked using PCMark05. Four controllable factors were involved in this study and hence by substituting the value 4 for n in Equation 3.5 as described in Equation 3.6, where x_1, x_2, x_3, and x_4 designated the controllable factors: data rate, seek reads, seek writes, and overhead respectively.

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 \qquad (3.6)$$

The summary of the response surface coefficients of Equation 3.7 and their significance is listed in Figure 3.25. The adjusted coefficient of multiple determination, R^2_{Adj} for Equation 3.7 is 98.06% which indicates a high-degree of correlation between the observed and predicted responses. Only about 2% of the total variations are not explained by this linear response surface model. The applications of T-values and P-values are used to check the significance of each coefficient.

Response Surface Regression: PCMark05 versus DataRate, SeekReads, SeekWrites, Overhead

Estimated Regression Coefficients for PCMark05

Term	Coef	SE Coef	T	P
Constant	3552.86	42.9612	82.699	0.000
DataRate	21.09	0.3333	63.267	0.000
SeekReads	-0.16	0.0046	-34.496	0.000
SeekWrites	-0.04	0.0046	-8.395	0.000
Overhead	8.74	0.2979	29.348	0.000

S = 125.488 PRESS = 6463299
R-Sq = 93.87% R-Sq(pred) = 93.71% R-Sq(adj) = 93.81%

Figure 3.25: Response Surface Methodology for Mars-RE disk drive benchmarked using PCMark05.

Figure 3.25 shows that the coefficients of all the linear effects are statistically significant, with P-values less than $\alpha = 0.05$. Equation 3.7 is the linear response surface model fitted to the data set obtained from the three-level full factorial design.

$$\hat{y} = 3553 + 21.1x_1 - 0.159x_2 - 0.039x_3 + 8.74x_4 \tag{3.7}$$

The summary of the ANOVA for the linear response surface model is shown in Figure 3.26. The ANOVA table indicates that this linear response surface model is highly significant within the range of the four controllable factors employed, as is evident from the large F-statistic (1531.12) and its low probability value ($P = 0.000$).

Analysis of Variance for PCMark05

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	4	96444454	96444454	24111114	1531.12	0.000
Linear	4	96444454	96444454	24111114	1531.12	0.000
Residual Error	400	6298941	6298941	15747		
Lack-of-Fit	76	5317973	5317973	69973	23.11	0.000
Pure Error	324	980967	980967	3028		
Total	404	102743395				

Figure 3.26: ANOVA table for Mars-RE disk drive benchmarked using PCMark05.

However, the small probability value ($P = 0.000$) of lack-of-fit statistic in Figure 3.26 suggests significant inadequacy of model fit to the experimental data observed in this three-level full factorial design. This obviously happens when the model does not well describe the experimental data but it can also arise when the model adequately represents the data but the precision of the replicates is so high that their variance is very small. Therefore, a significant lack-of-fit statistic does not necessarily mean that the model is unusable.

A better understanding of what could cause the lack-of-fit statistic to be significant may help determine the adequacy of the model. The formula for lack-of-fit statistic is described in Equation 3.8 [26], where MS is the mean square value.

$$\text{Lack-of-fit F-test} = \frac{\text{Lack-of-fit MS}}{\text{Pure Error MS}} \tag{3.8}$$

The numerator in this equation is the variation between the actual values and the values predicted from the model. The denominator is the variation between the replicates that is an estimate of the normal variation that cannot be accounted for by any model. By definition, a significant lack-of-fit statistic suggests that the variation of the replicates about their mean values is less than the variation of the actual points about their predicted values.

In these experiments, the replicates were actually run more like repeated measurements on a single disk drive of the same factor combination rather than running on multiple of the same disk drives independently under the same experimental conditions. The pure error of the natural variation could have been underestimated and that possibly led to a small denominator in the lack-of-fit formula.

Although the predicted linear response surface fits the model points well as indicated by the distributional properties of the residuals in Figure 3.22, the greater differences between the actual data points and the response plane than the differences between the replicates triggers the significant lack-of-fit statistic. In this case, the lack-of-fit statistic is no longer valid to the model. Decisions about whether this model is a good fit or not can be made based on how well the residuals are normally distributed and falls within 95% confidence intervals. During this experimental phase, the low lack-of-fit statistic can be ignored since the distribution of the residuals is quite satisfactory.

3.3.4 Response Optimizer

Minitab provides a response optimizer tool that calculates an optimal solution for user-defined response and draws the corresponding parameter levels on the plot based on the derived model equation. By interactively adjusting the factor settings, the response optimizer shows how different parameter combinations affect the predicted model responses on the optimization plot. In this research, the response optimizer was used to search for factor settings with a maximum response of disk performance benchmarked using PCMark05. Each factor effect on the PCMark05 response is shown on the optimization plot. The vertical red lines correspond to the current factor settings indicated in red at the top column.

The horizontal blue dashed-line represents the response for current factor level combination as indicated in blue at the left column.

The response of disk performance benchmarked using PCMark05 is maximized when data rate and overhead are at their highest levels, 110 MB/s and 126 MHz respectively and seek reads and seek writes are at their lowest levels, both 144.1 μs.

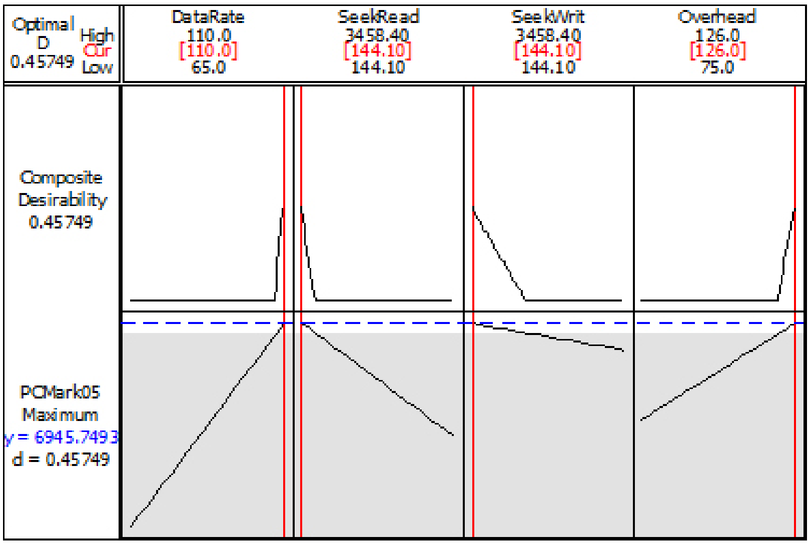

Figure 3.27: Optimization plot for the linear response surface model (PCMark05).

The optimization plot in Figure 3.27 indicates the corresponding response prediction to the parameter settings when moving the red factor level lines for each factor. This tool is extremely useful for exploring the sensitivity of response to significant factor changes, predicting the response of a specific factor level combination, searching for lower-cost factor settings nearby the optimal, or discovering the neighborhood of a local solution. These optimal factor settings for the maximal response of disk performance are required for graphical analysis in the next section.

3.3.5 Graphical Analysis

The surface plot and contour plot are the graphical representations of the response surface model generated from different perspectives using Minitab. These visual representations are useful for establishing desirable response predictions and optimizing parameter conditions. No curvature was formed in the underlying response surface since the linear model does not include any interactions or squares in the equation.

A surface plot provides a clear three-dimensional view of the linear response surface model in the factor space. The predicted response is plotted on the z-axis versus the selected pair of factors on the x-axis and y-axis. The two factors on the surface plot are inspected through various characterization sequences within the defined factor domains. The remaining factors are held constant at their optimal levels, as suggested by Minitab's response optimizer in previous section.

The projection of the response surface onto the selected factors is represented on the two-dimensional contour plot. The two axes denote the selected pair of parameters being inspected while the remaining factors are held constant at their optimal levels as specified on a surface plot. Each contour line of constant response connects all points that have the same response prediction. No curved contour line was drawn in the plane due to the fact of the flat surface on surface plot.

Figure 3.28 shows the effect of data rate and seek reads while keeping the level settings of seek writes and overhead at their optimal values of 144.1 μs and 126 MHz respectively. It can be seen that disk performance benchmarked using PCMark05 is increased with increase in data rate and reduction in seek reads. The maximum predicted response is reached at a data rate of 105-110 MB/s and a seek reads of 144.1-1000 μs. Both the surface plot and contour plot agree that the maximum Mars-RE disk performance benchmarked using PCMark05 under the stated conditions falls within the 6750-7000 score range.

Apparent improvement in response prediction is observed over an increase for data rate in comparison to moderate improvement over a reduction for seek reads. The response stability indicates various combinations of data rate and seek reads settings attributed to the same response prediction along a contour line on contour plot.

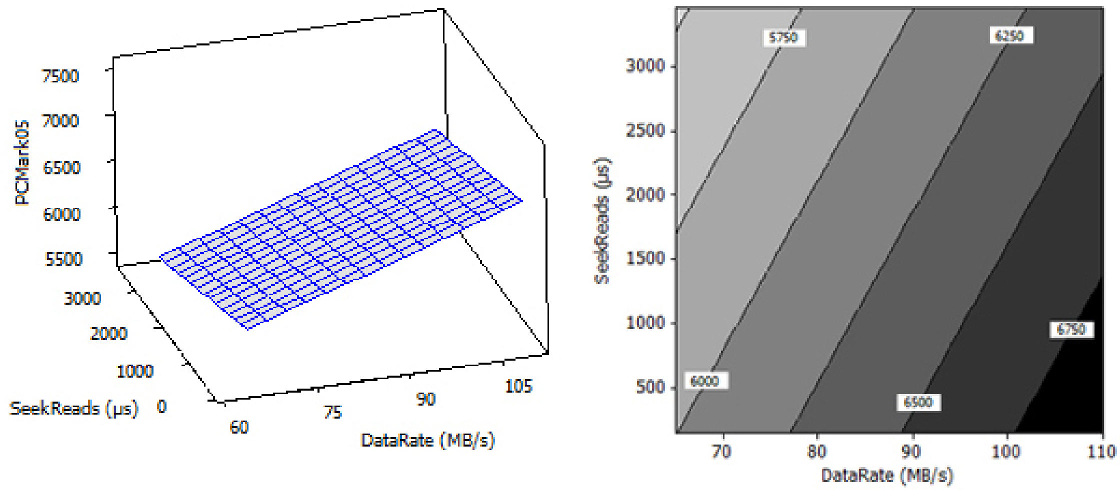

Figure 3.28: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark05 versus data rate and seek reads.

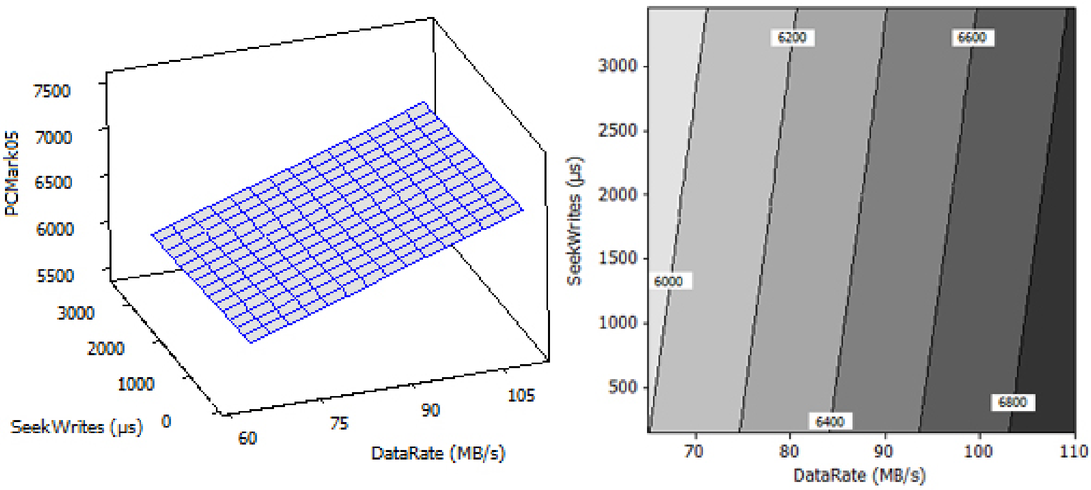

Figure 3.29: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark05 versus data rate and seek writes.

Figure 3.29 shows how Mars-RE disk performance benchmarked using PCMark05 varies with data rate and seek writes at the optimal seek reads setting of 144.1 μs and overhead setting of 126 MHz. Apparent disk performance improvement in PCMark05 scores is observed over an increase for data rate from 65 MB/s to 110 MB/s at a constant level of seek writes. On the other hand, a trivial response improvement in PCMark05 scores is observed over a reduction for seek writes from 3458.4 μs to 144.1 μs at a constant data rate level. The maximal response in PCMark05 scores occurs within 6800-7000 score range at a data rate of 105-110 MB/s and a seek writes of 144.1-1000 μs.

The contour lines in Figure 3.29 are roughly parallel to the seek writes-axis which suggests that seek writes has a relatively weak effect within the factor domains. According to the ANOVA summarized in Table 3.18, data rate (T=63.267, P=0.000) is much more dominant to Mars-RE disk performance benchmarked using PCMark05 when compared with seek writes (T=-8.395, P=0.000). Therefore, the disk performance of Mars-RE tends to be more consistent within this factor space.

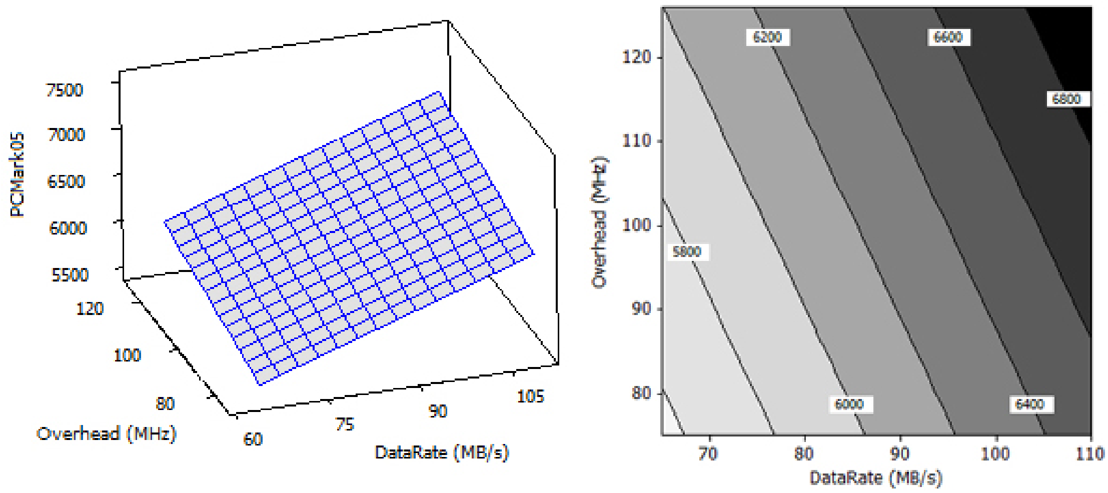

Figure 3.30: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark05 versus data rate and overhead.

The surface plot and contour plot in Figure 3.30 describe Mars-RE disk performance benchmarked using PCMark05 as a function of data rate and overhead at fixed levels of seek reads and seek writes, both 144.1 μs.

The response prediction improves over the increases for both data rate and overhead. The maximal Mars-RE disk performance is reached at a data rate of 105-110 MB/s and an overhead of 120-126 MHz. The targeted Mars-RE disk performance benchmarked using PCMark05 is between the scores of 6800 and 7000 as indicated on both surface plot and contour plot.

Figures 3.31-3.33 show surface plots and contour plots that describe the predicted response variation of this linear model versus seek reads and seek writes, seek reads and overhead, and seek writes and overhead respectively.

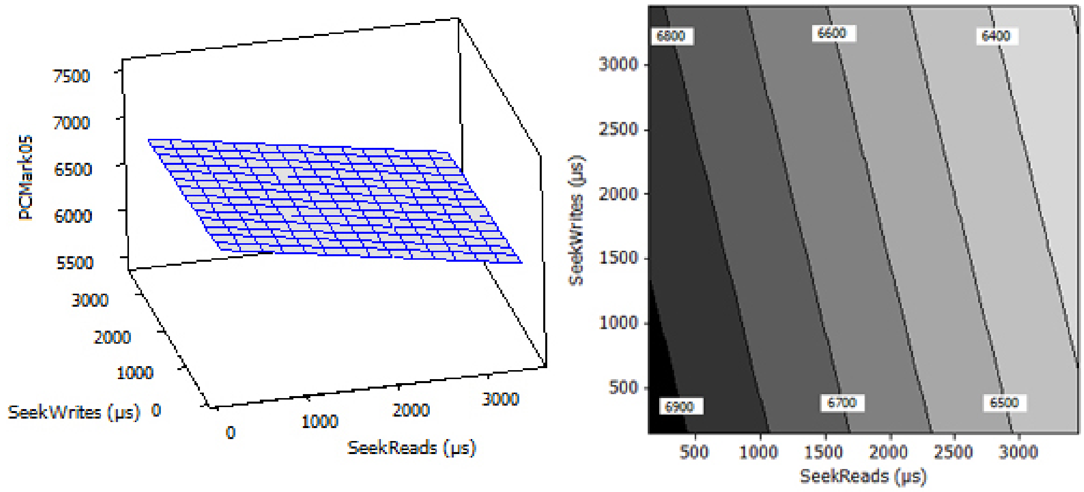

Figure 3.31: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark05 versus seek reads and seek writes.

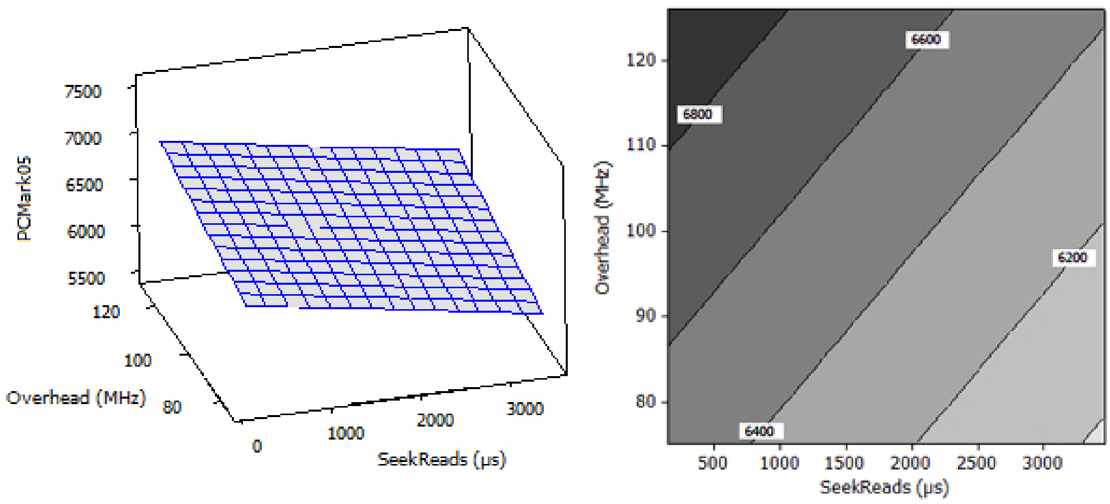

Figure 3.32: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark05 versus seek reads and overhead.

Although the surface plot and contour plot serve as useful analysis tools, they are not as effective if the optimal operating conditions of the parameters have not been determined. If the optimal parameter settings are not known beforehand, the settings of the factors not being inspected on the plots will be held constant at the central conditions, which are their mid-levels. However, these plots will only indicate the response prediction in the neighborhood of the central conditions of the parameters. Based on the results from the surface plot and contour plot under the central conditions of parameter settings, the optimal range of each parameter

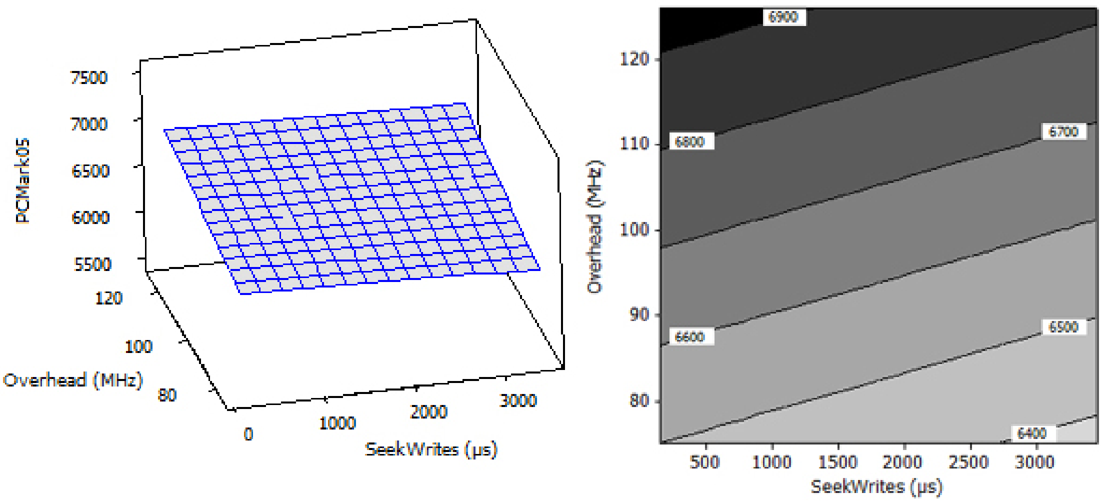

Figure 3.33: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark05 versus seek writes and overhead.

can be determined on the plots for a desired response prediction. The plots can then be generated again with the settings of factors not on the plots held constant at levels selected within the determined optimal ranges.

When the central conditions of data rate at 87.5 MB/s, seek reads at 1801.25 μs, seek writes at 1801.25 μs, and an overhead of 100.5 MHz are used as a starting point, the improved Mars-RE disk performance benchmarked using PCMark05 can be yielded by increasing the levels of data rate and overhead and reducing the seek reads and seek writes settings. Specifically, the maximal response can be achieved by optimizing the parameter conditions at a data rate of 105-110 MB/s, a seek reads and a seek writes both of 144.1-500 μs, and an overhead of 120-126 MHz.

The relationship among these four parameters can be described by the linear response surface model as a significant synergistic effect when analyzing a targeted Mars-RE disk performance benchmarked using PCMark05 within this factor space. As a consequence, such a targeted response may possibly be realized under several data rate, seek reads, seek writes, and overhead level combinations. Increasing one of the factor levels gives different opportunities to increase or decrease the other factor settings in order to reach the same targeted response. All the surface plots and contour plots in Figures 3.28-3.33 suggest that the maximal Mars-RE disk performance benchmarked using PCMark05 falls within 6900-7000 score range and is closer to the 6900 contour line. This maximal response can be reached at optimal data rate of 105-110 MB/s, seek reads and a seek writes both of 144.1-500 μs, and overhead of 120-126 MHz. The results from these graphical analyses agree with the optimal settings of data rate at 110 MB/s, seek reads at 144.1 μs, seek writes at 144.1 μs, and overhead at 126 MHz obtained using Minitab's response optimizer to reach maximal Mars-RE disk performance of 6945.75 in PCMark05 scores.

The main purpose of the graphical analysis is to provide a visual representation of the response predictions in the neighborhood of the indicated parameter settings. A large amount of surface plots and contour plots is required to be analyzed when there are quite a few significant factors to the response due to the fact that surface plot and contour plot only analyze the response variation of one pair of factors at a time. However, these plots are useful for confirming the optimal values of response predictions or parameter ranges suggested by the model equation in an efficient manner.

3.3.6 Verification of Model

A confirmatory experiment is usually conducted to confirm the accuracy and stability of the fitted model. Typically, this is done by running the experiment with optimal settings to make sure that the observed response value is reasonably close to the predicted value. The confirmatory experiment conducted for this study is different from a standard confirmatory test. Rather than validating the linear response surface models with the same Mars-RE disk drive, a technically similar hard disk drive, Atlantis-RE was chosen for model verification. Atlantis, 3.5", 500 GB, SATA 3 Gb/s, 16 MB Cache, 7,200 RPM is a prior generation disk drive from the same model family as Mars-RE but with different capacity, buffer size, and drive GB/platter.

Overhead setting of this Atlantis-RE disk drive was not adjustable. The level of overhead was held constant at a delay of 0 μs, equivalent to 114 MHz. The parameter settings of data rate, seek reads, and seek writes were randomly selected to compensate the factor variations with prerequisites for compromising the technical specifications of Atlantis-RE disk drive.

Settings of data rate were chosen based on the sequential read speed plot of the Atlantis-RE disk drive measured using HD Tach RW version 3.0.1.0. Each disk partition under investigation was 12.5 GB. The settings of seek reads and seek writes were randomly selected with the lowest boundary of 150.95 μs and highest boundary of 3441.66 μs that are within the factor space of the original Mars-RE experiments.

Table 3.9: Validation results of the linear response surface model (PCMark05).

Run	x_1	x_2	x_3	x_4	y	$\hat{y}$
1	104	150.95	150.95	114	6875	6713
2	104	150.95	3441.66	114	6665	6586
3	104	3230.33	633.99	114	6147	6205
4	104	1720.83	2686.91	114	6418	6365
5	104	603.80	3320.90	114	6664	6518
6	104	3441.66	150.95	114	6032	6190
7	104	3441.66	3441.66	114	6021	6063
8	87	150.95	150.95	114	6568	6354
9	87	150.95	3441.66	114	6439	6227
10	87	3139.76	1267.98	114	5811	5836
11	87	483.04	3381.28	114	6495	6177
12	87	1509.50	2294.44	114	6284	6056
13	87	3441.66	150.95	114	5813	5831
14	87	3441.66	3441.66	114	5732	5704
15	72	150.95	150.95	114	6090	6038
16	72	150.95	3441.66	114	5902	5911
17	72	2777.48	2052.92	114	5431	5547
18	72	1811.40	3019.00	114	5688	5663
19	72	2143.49	1901.97	114	5649	5654
20	72	3441.66	150.95	114	5409	5515
21	72	3441.66	3441.66	114	5238	5388

The confirmatory experiment was carried out under 21 different parameter combinations. Each parameter combination was run only once. The uncoded parameter settings and their actual and predicted responses of this confirmatory experiment were analyzed using Minitab and listed in Table 3.9.

Figure 3.34 shows the observed response versus predicted response plot of this confirmatory experiment. The experimental points were roughly distributed around the ratio 1:1 diagonal line on the plot which indicated that this linear response surface model makes decent performance predictions for hard disk drive within Mars-RE model family when benchmarked using PCMark05.

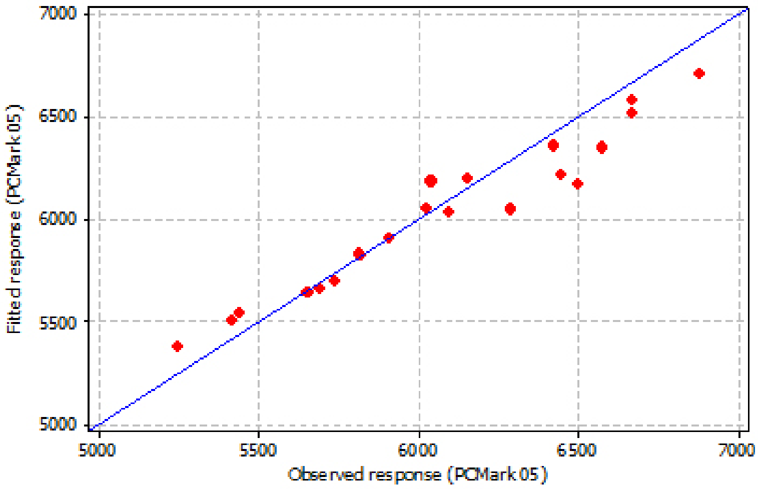

Figure 3.34: Fitted response versus observed response plot for model verification (PCMark05).

The DOE characterized Mars-RE disk drive with a linear response surface model. Although this model equation was empirical, results of the confirmatory experiment indicated that the experimental values were found to be significantly in agreement with the predicted responses. The validation of this linear model using Atlantis-RE, a prior generation disk drive of Mars-RE within the same model family was quite satisfactory.

3.4 Experiment 3 - PCMark Vantage

3.4.1 Experimental Design

Another RSM experiment was conducted to evaluate the effects of data rate, seek reads, seek writes, and overhead on the Mars-RE disk performance benchmarked using PCMark Vantage. This experiment was similar to the previous one where the benchmarked Mars-RE disk performance in PCMark05 scores was studied. However, there were some differences such as only two levels of seek writes were investigated and each treatment had only three replicates due to time constraint. Table 3.10 indicates the levels of each controllable factor.

Table 3.10: Factor levels for the general full factorial desgin (PCMark Vantage).

Factor	Unit	Symbol	Level		
Data Rate	MB/s	x_1	65	80	110
Seek Reads	μs	x_2	144.1	1729.2	3458.4
Seek Writes	μs	x_3	144.1		3458.4
Overhead	MHz	x_4	75	105	126

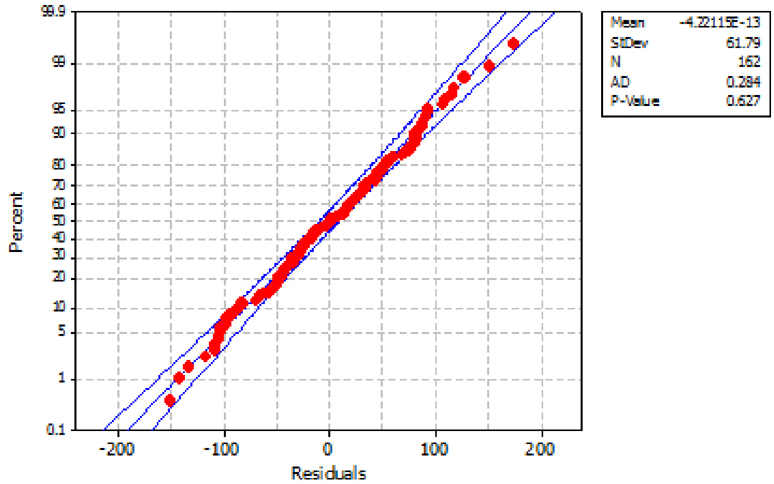

Figure 3.35: Normal probability plot of residuals of the Mars-RE linear response surface model (PCMark Vantage).

A total of 162 runs (i.e. $3 \times 3 \times 2 \times 3 = 54$ runs per replicate) were conducted in this general full factorial design with four controllable factors and three replicates. The adequacy of the fitted linear response surface model can be verified by residual plots as shown in Figure 3.35-3.37. The linear response surface model is concluded as an adequate fit to the obtained data set since all three statistical assumptions are all satisfied.

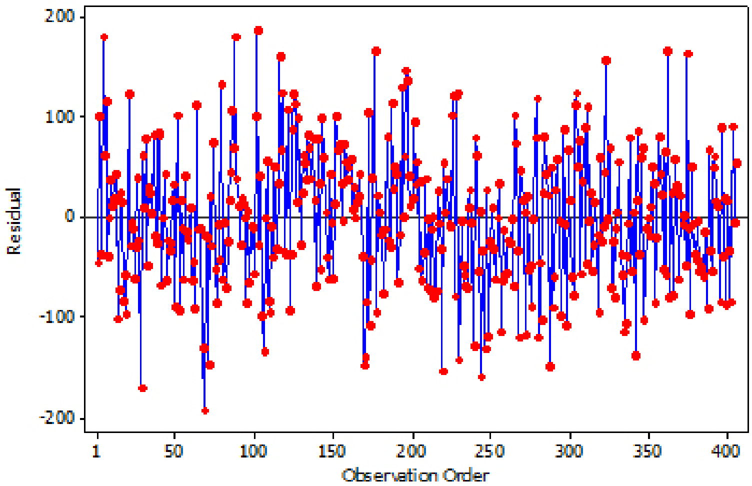

Figure 3.36: Residuals versus order plot of the Mars-RE linear respons surface model (PCMark Vantage).

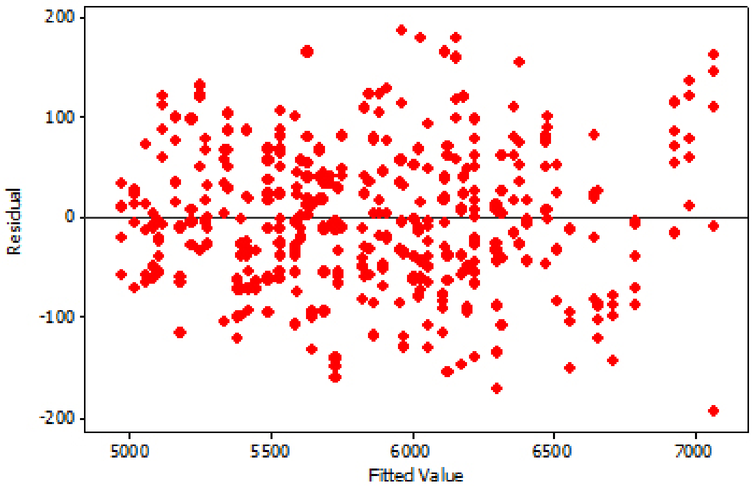

Figure 3.37: Residuals versus fitted values plot of the Mars-RE linear respons surface model (PCMark Vantage).

3.4.2 Results and Discussion

The linear response surface model fitted to the data set obtained from the general full factorial design is derived in Equation 3.9. A summary of the Analysis of Variance (ANOVA) for this linear response surface model is shown in Figure 3.38.

$$\hat{y} = 2463 + 13.28x_1 - 0.121x_2 - 0.100x_3 + 4.988x_4 \qquad (3.9)$$

The ANOVA table indicates that this linear response surface model is highly

Analysis of Variance for PCMarkVantage

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	4	20597623	20597623	5149406	1315.41	0.000
Linear	4	20597623	20597623	5149406	1315.41	0.000
Residual Error	157	614606	614606	3915		
Lack-of-Fit	49	561574	561574	11461	23.34	0.000
Pure Error	108	53032	53032	491		
Total	161	21212228				

Figure 3.38: ANOVA table for Mars-RE disk drive benchmarked using PCMark Vantage.

significant, as is evident from the large F-statistic (1315.41) and its very low probability value ($P = 0.000$). This model was found to be adequate for prediction within the range of the four controllable factors employed. Since the residuals are normally distributed and falls within 95% confidence intervals as shown in Figure 3.35, the low probability value ($P = 0.000$) of lack-of-fit can be ignored.

A summary of the response surface coefficients of Equation 3.9 and their significance is listed in Figure 3.39. The adjusted coefficient of multiple determination, R^2_{Adj} is 97.03%, indicates a high-degree of correlation between the observed and predicted responses. Only about 3% of the total variations are not explained by this linear response surface model. Figure 3.39 shows that the coefficients of all the terms are statistically significant, with P-values less than $\alpha = 0.05$.

Response Surface Regression: PCMarkVantag versus DataRate, SeekReads, SeekWrites, Overhead

Estimated Regression Coefficients for PCMarkVantage

Term	Coef	SE Coef	T	P
Constant	2462.99	34.1644	72.092	0.000
DataRate	13.28	0.2628	50.541	0.000
SeekReads	-0.12	0.0036	-33.336	0.000
SeekWrites	-0.10	0.0030	-33.839	0.000
Overhead	4.99	0.2349	21.234	0.000

S = 62.5674 PRESS = 658204
R-Sq = 97.10% R-Sq(pred) = 96.90% R-Sq(adj) = 97.03%

Figure 3.39: Response Surface Methodology for Mars-RE disk drive benchmarked using PCMark Vantage.

3.4.3 Response Optimizer

Minitab's response optimizer suggests that the benchmarked Mars-RE disk performance in PCMark Vantage scores is maximized when data rate and overhead are at their highest settings, 110 MB/s and 126 MHz respectively and seek reads and seek writes are at their lowest settings, both 144.1 μs as shown on the optimization plot in Figure 3.40. These optimal factor settings are the same as for the maximal Mars-RE disk performance benchmarked using PCMark05 and will be used for graphical analysis in the next section.

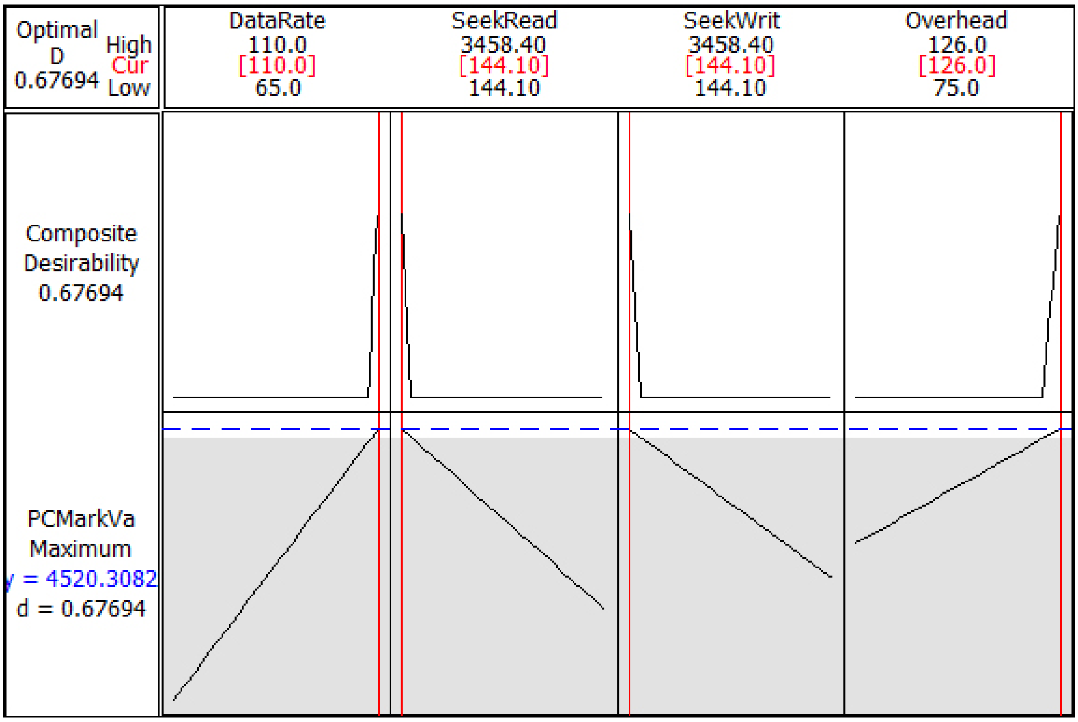

Figure 3.40: Optimization plot for the linear response surface model (PCMark Vantage).

3.4.4 Graphical Analysis

The response surface of a linear model alone represents a two-dimensional flat plane within the three-dimensional factor space without any indication of local shape or warp. Figures 3.41-3.46 show surface plots and contour plots that describe the response variations of this linear model with data rate and seek reads, data rate and seek writes, data rate and overhead, seek reads and seek writes, seek reads and overhead, and seek writes and overhead respectively. These plots demonstrate how benchmarked Mars-RE disk performance in PCMark Vantage scores varies with selected factor pair while the other two factors are held at their optimal settings (data rate at 110 MB/s, seek reads at 144.1 μs, seek writes at 144.1 μs, and overhead at 126 MHz) derived from Minitab's response optimizer.

The contour lines on these contour plots are all straight and parallel to each other since the response surface is a flat plane without any curvature. The contour lines are parallel to a directions near one of the two diagonals of the plots which indicate that the parameters have relatively similar effect to disk performance benchmarked using PCMark Vantage due to the simplicity of this linear model. According to the response surface analysis summarized in Figure 3.39, data rate ($T = 50.541$, $P = 0.000$) has the most dominant effect, overhead ($T = 21.234$, $P = 0.000$) is the least dominant, and seek reads ($T = -33.336$, $P = 0.000$) and seek writes ($T = -33.839, P = 0.000$) have almost the same effects.

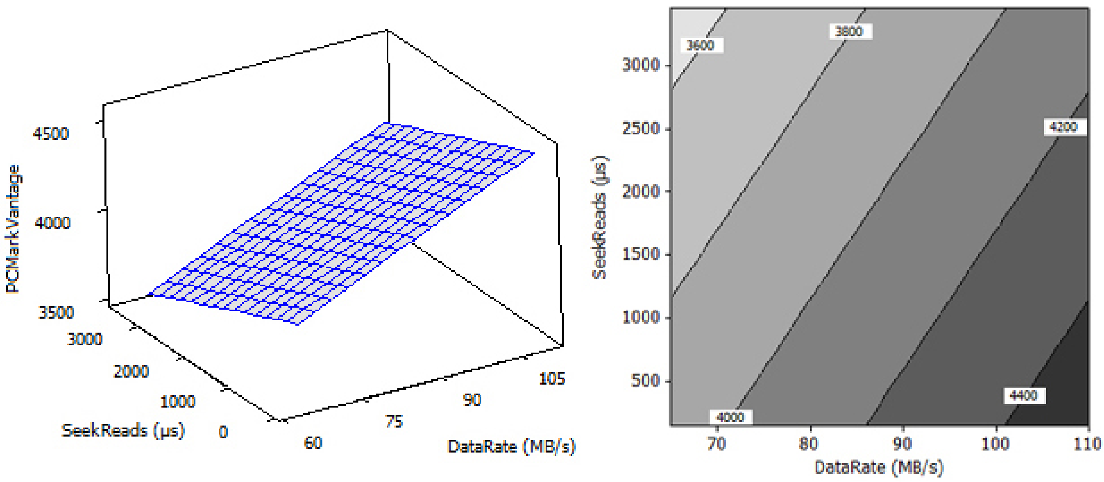

Figure 3.41: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark Vantage versus data rate and seek reads.

Figure 3.43 describes the benchmarked Mars-RE disk performance in PCMark Vantage scores as a function of data rate and overhead, the most and least dominant effects, respectively. The contour lines are parallel to a direction that is near the overhead-axis which suggests that overhead has a relatively weaker effect within the factor domains. Apparent improvement in Mars-RE disk performance

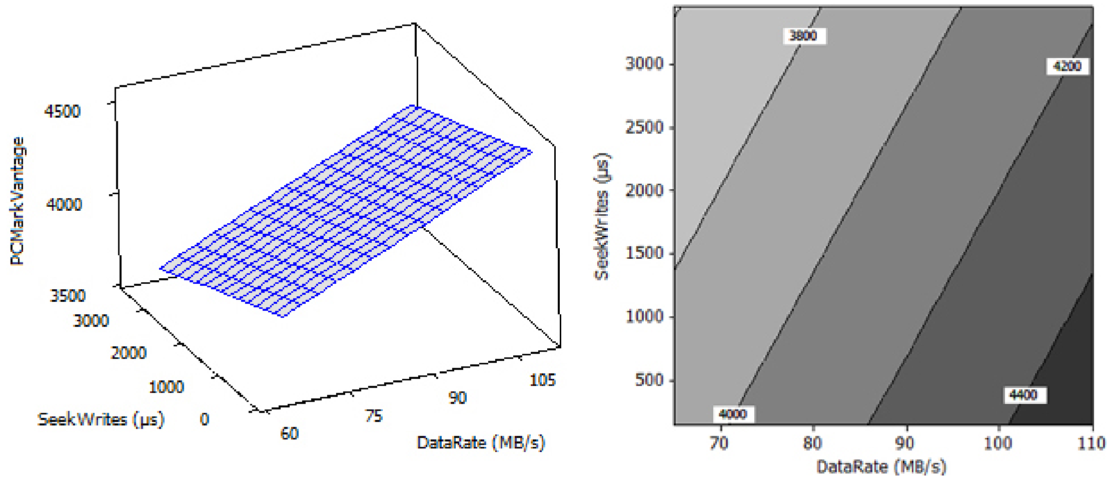

Figure 3.42: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark Vantage versus data rate and seek writes.

of roughly 600 in PCMark Vantage scores is observed over an increase for data rate from 65 MB/s to 110 MB/s at a constant level of overhead in comparison to moderate improvement of roughly 200 in PCMark Vantage scores over an increase for overhead from 75 MHz to 126 MHz at a constant data rate level.

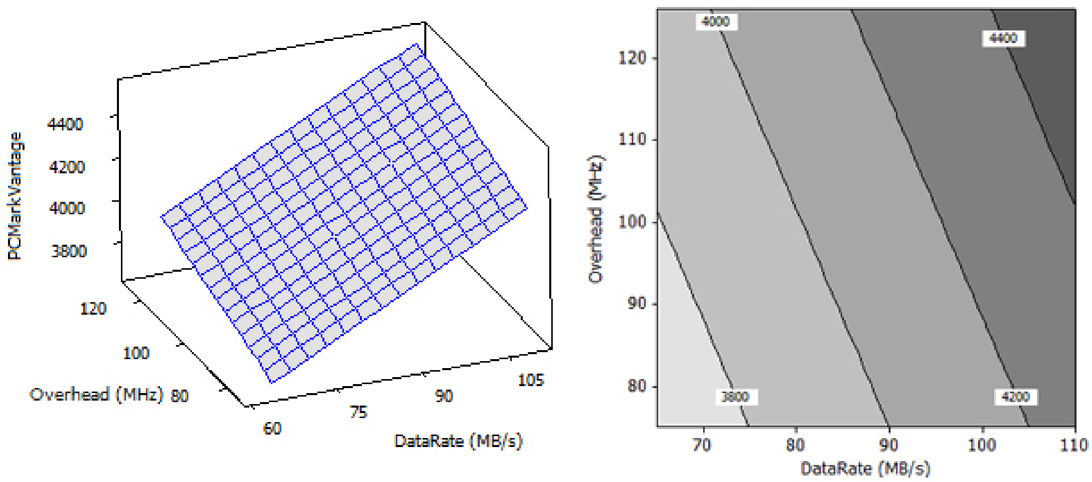

Figure 3.43: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark Vantage versus data rate and overhead.

When the central conditions of data rate at 87.5 MB/s, seek reads at 1801.25 μs, seek writes at 1801.25 μs, and an overhead of 100.5 MHz are used as a starting point, the improved Mars-RE disk performance benchmarked using PCMark Vantage can be yielded by increasing the levels of data rate and overhead and reduc-

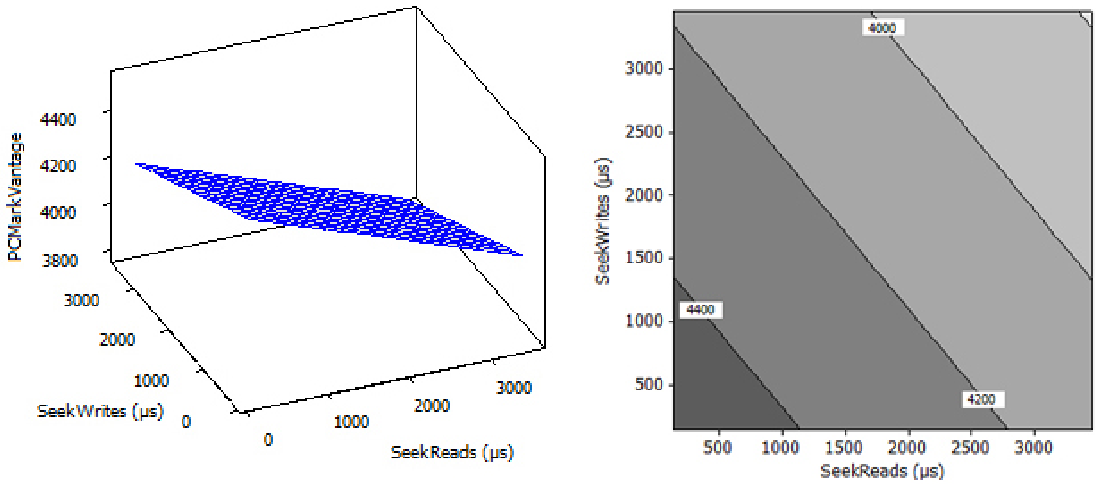

Figure 3.44: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark Vantage versus seek reads and seek writes.

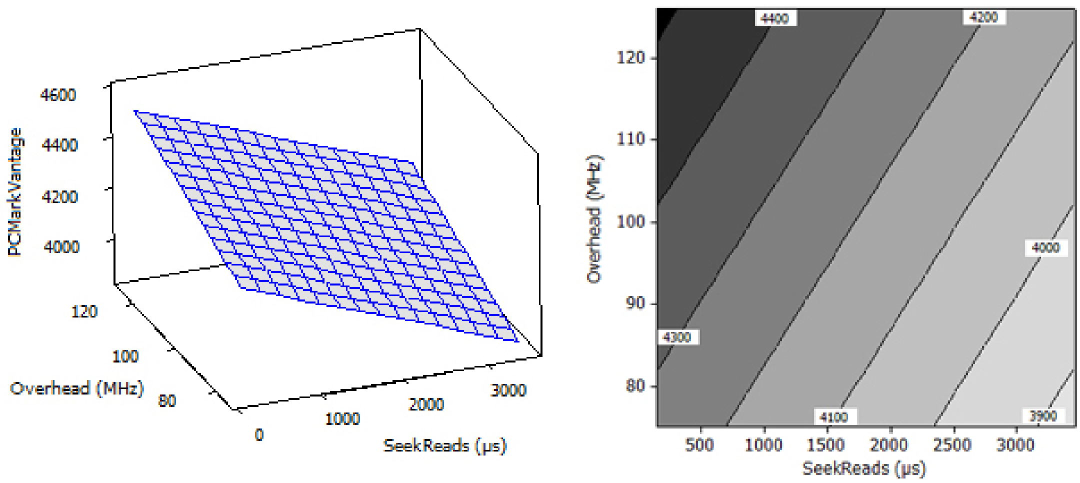

Figure 3.45: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark Vantage versus seek reads and overhead.

ing the seek reads and seek writes settings. Specifically, these graphical analyses suggest that the maximal response can be achieved by optimizing the parameter conditions at a data rate setting of 105-110 MB/s, seek reads setting and seek writes setting both of 144.1-500 μs, and an overhead setting of 120-126 MHz which agrees with the optimal settings of data rate at 110 MB/s, seek reads at 144.1 μs, seek writes at 144.1 μs, and overhead at 126 MHz derived from Minitab's response optimizer. All the surface plots and contour plots in Figures 3.41-3.46 suggest that the maximal response prediction in PCMark Vantage scores falls within 4500-4600

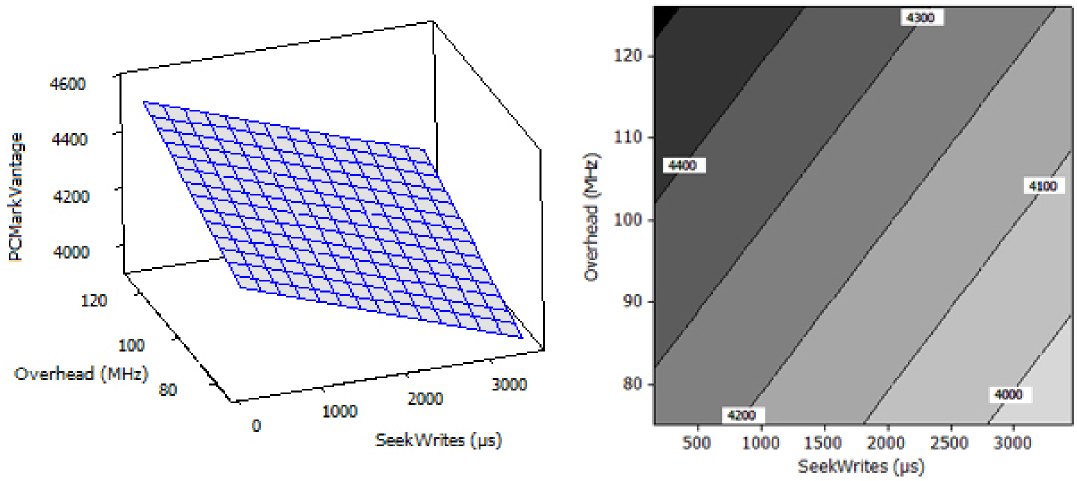

Figure 3.46: Surface plot and contour plot of Mars-RE disk performance benchmarked using PCMark Vantage versus seek writes and overhead.

range and is close to the 4500 contour line which also agrees with the estimated PCMark Vantage score of 4520 by Minitab's response optimizer.

3.4.5 Verification of Model

Atlantis-RE disk drive was again used for the comparison of the 21 predicted responses and their corresponding observed responses. Each factor combination was run only once. The settings of data rate and overhead remained the same as the settings of the confirmatory experiment of the model benchmarked using PCMark05. Seek reads and seek writes settings were randomly selected within the domains of the original Mars-RE experiments. Table 3.11 lists the uncoded factor settings and their corresponding actual and predicted responses analyzed using Minitab.

The experimental points on the observed response versus predicted response plot in Figure 3.47 were roughly distributed around the diagonal line on the plot. The result indicated that this empirical model equation as a function of data rate, seek reads, seek writes, and overhead makes decent prediction for benchmarked Mars-RE disk performance in PCMark Vantage scores.

Table 3.11: Validation results of the linear response surface model (PCMark Vantage).

Run	x_1	x_2	x_3	x_4	y	$\hat{y}$
1	104	150.95	150.95	114	4456	4482
2	104	150.95	3441.66	114	4073	4071
3	104	301.90	1267.98	114	4318	4325
4	104	2415.20	2928.43	114	3753	3848
5	104	513.23	2988.81	114	4114	4089
6	104	3441.66	150.95	114	3895	3976
7	104	3441.66	3441.66	114	3601	3631
8	87	150.95	150.95	114	4152	4245
9	87	150.95	3441.66	114	3865	3868
10	87	422.66	301.90	114	4128	4196
11	87	1781.21	1569.88	114	3825	3885
12	87	2807.67	452.85	114	3734	3844
13	87	3441.66	150.95	114	3672	3764
14	87	3441.66	3441.66	114	3397	3452
15	72	150.95	150.95	114	3987	3979
16	72	150.95	3441.66	114	3710	3632
17	72	1751.02	3351.09	114	3527	3476
18	72	543.42	2505.77	114	3779	3693
19	72	2083.11	271.71	114	3659	3725
20	72	3441.66	150.95	114	3487	3521
21	72	3441.66	3441.66	114	3253	3238

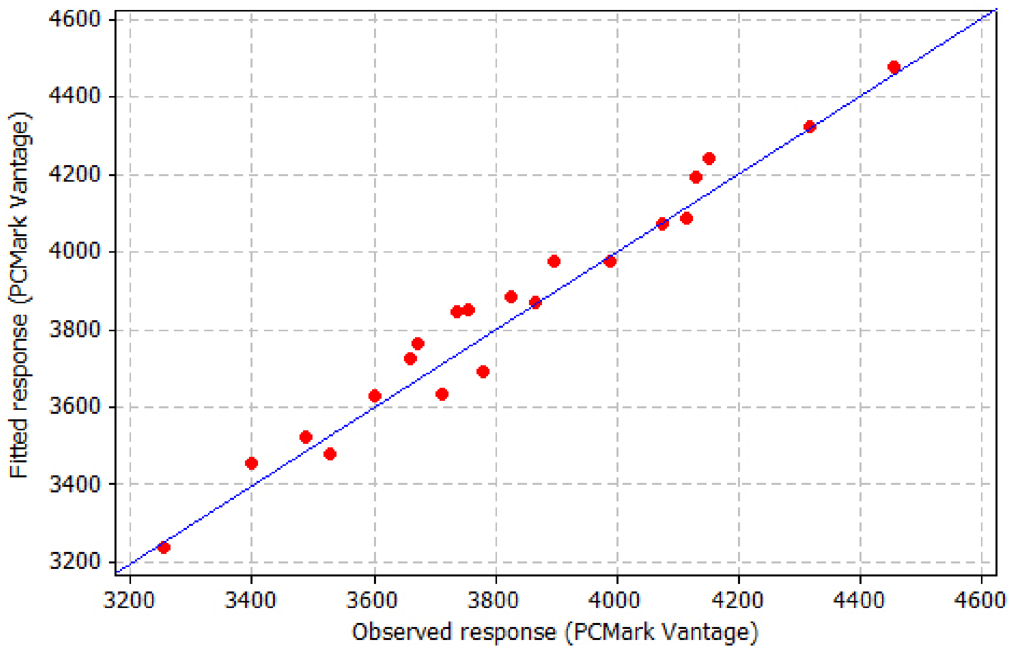

Figure 3.47: Fitted response versus observed response plot for model verification (PCMark Vantage).

Chapter 4

Conclusions

4.1 Hard Disk Drive Performance Modeling

The primary goal of this study is to model the performance characteristics of selected hard disk drives manufactured by Western Digital Corporation. This study presents a comprehensive approach which allows us to highlight the impacts of data rate, seek time, and overhead on performance of selected hard disk drives benchmarked using PCMark software. Such analysis is made possible by the deployment of Design of Experiments (DOE) methodology. The information gained from DOE analysis can easily be used for optimization regarding the different influences from the factors.

The DOE methodology offers a wide range of statistical techniques and graphical representations for planning experiments, analyzing data, and providing reproducible results. This fact-based approach is extremely useful in making informed decisions with confidence, even with very limited data, time, and resources given in development cycle. Numerous combinations of factor settings can be evaluated for the best overall combination in a small number of test runs. Significant effects on the performance variability of selected hard disk drives benchmarked using PCMark software were identified with efficient designs instead of the tranditional hit-and-miss or trial-and-error approaches. The DOE methodology is a powerful tool that gives early warning of potential problems to avoid expensive time-wasting projects, to reduce in cost due to testing, labor, and materials, and to increase in quality and reliability.

The linear response surface models were derived to fit the Mars-RE disk performance benchmarked using PCMark05 and PCMark Vantage. All the linear effects were identified to be statistically significant. Both statistical models have found that data rate is the most significant factor among all four factors under investigation.

As mentioned earlier, the "fine-tuned" performance factor settings for most disk drive products in the market today may not be at their optimal levels indicated by the model trends due to many practical concerns such as costs and time-to-market. The parameter settings of the hard disk drive being characterized determine the direction of the prediction trend and the amount of setting adjustments to achieve the goal of the next "fine-tuned" generation. The fact of the uncertainties further ahead outside the defined factor space makes it more critical to require evolutionary operations to ensure robustness of the model by continuously updating the design space of the predictive model. Each successive experiment is then designed based on results thus far and can be shifted in the direction of improvement. Under certain circumstances, a combination of factor settings that is synergistic with these factors and meets conditions that are as close to the optimal settings as possible may be chosen over others for a most economical and beneficial solution of such performance improvement.

4.2 Future Work

There is great potential for a partially or fully automated data acquisition based disk performance characterization system. Automation would remove or reduce the chances of operational error (i.e. typo in factor level adjustments) or equipment failure (i.e. SATA cable failure) caused by modifying the firmware or hardware settings repeatedly for each test run. The overall accuracy of measurements would be improved and therefore less replications would be required for each run. It would also be more time-efficient because less time would be spent on unnecessary debugging and repeating the data acquisition process, and more test runs could be completed within the time frame since no operator needs to be physically presented.

It is left to future work to compare different data acquisition designs for a higher quality or more accurate model prediction and to automate the data acquisition process for efficiency in the sense of time. Although this study did provide a complete Design of Experiments (DOE) procedure for the Western Digital Mars-RE and Atlantis-RE model family, the future experiments will still require detailed analyses during each stage to make appropriate adjustments for the design plan. It is recommended to follow the principles of this study but not to be restricted by it as outside of the empirical design space and future technology advancements are still unknown to us at this moment.

Bibliography

[1] E. Grochowski, and R.F. Hoyt, *Future trends in hard disk drives.* IEEE Transactions on Magnetics, Vol. 32, No. 3, May 1996, pp. 1850-1854.

[2] S.W. Ng, *Advances in disk technology: performance issues.* IEEE Transactions on Computer, Vol. 31, No. 5, May 1998, pp.75-81.

[3] C. Ruemmler and J. Wilkes, *An introduction to disk drive modeling.* IEEE Transactions on Computer, Vol. 27, No. 3, March 1994, pp. 17-28.

[4] D.C. Montgomery, *Design and Analysis of Experiments.* 4th Edition, Wiley, October 1996.

[5] R. Berry, *Computer benchmark evaluation and Design of Experiments, a case study.* IEEE Transactions on Computer, Vol. 41, No. 10, October 1992, pp.1279-1289.

[6] P. Triantafillou, S. Christodoulakis, and C. Georgiadis, *A comprehensive analytical performance model for disk devices under random workloads.* IEEE Transactions on Knowledge and Data Engineering, Vol. 14, No. 1, 2002, pp. 140-155.

[7] C.R. Lumb, J. Schindler, and G.R. Ganger, *Freeblock Scheduling Outside of Disk Firmware.* Proc. First Usenix Conf. File and Storage Technologies (FAST '02), January 2002.

[8] S.R. Schmidt and R.G. Launsby, *Understanding Industrial Designed Experiments.* Air Academy Press, October 1994.

[9] S.W. Schlosser, J. Schindler, S. Papadomanolakis, M. Shao, A. Ailamaki, C. Faloutsos, G.R. Ganger, *On multidimensional data and modern disks*. Third Usenix Conf. File and Storage Technologies (FAST '05), 2005.

[10] J. Gim, Y. Won, J. Chang, J. Shim, and Y. Park, *Dig: Rapid characterization of modern hard disk drive and its performance implication*. Proc. 5th International Workshop on Storage Network Architecture and Parallel I/Os (SNAPI), 2008, pp. 74-83.

[11] S.W. Ng, *Improving Disk Performance Via Latency Reduction*. IEEE Transactions on Computer, Vol. 40, No. 1, January 1991, pp. 22-30.

[12] D. Kotz, S.B. Toh and S. Radhakrishnan, *A detailed simulation model of the HP-97560 disk drive*. Tech. Rep. PCS-TR94-220, Department of Computer Science, Dartmouth College, July 1994.

[13] J. Schindler and G. R. Ganger, *Automated disk drive characterization*. Tech. Rep. CMU-CS-99-176, Carnegie Mellon University, December 1999.

[14] B.L. Worthington, G.R. Ganger, Y.N. Patt, and J Wilkes, *On-line extraction of SCSI disk drive parameters*. Tech. Rep. CSE-TR-323-96, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, December 19, 1996.

[15] D. M. Jacobson and J. Wilkes, *Disk scheduling algorithms based on rotational position*. Tech. Rep. HPL-CSP-91-7rev1, HP Laboratories, February 1991.

[16] Q. Hao, G. Guo, S. Chen, and T. Low, *A Self-tuning Robust Track-following Control of Sampled-Data Hard Disk Drive Servo System*. Proc. American Control Conference, June 2001.

[17] D. Oh, J. Koo, C. Kim, and Y. Byun, *Optimization of wide-bandwidth hard disk drive actuator design using statistical methods*. IEEE Transactions on Magnetics, Vol. 41, No. 2, February 2005, pp. 836-841.

[18] J. T. Li, Z. J. Liu, M. A. Jabbar, and X. K. Gao, *Design optimization for cogging torque minimization using Response Surface Methodology*. IEEE Transactions on Magnetics, Vol. 40, No. 2, March 2004, pp. 1176-1179.

[19] mkomo.com, *A history of storage cost*.
http://www.mkomo.com/cost-per-gigabyte.

[20] S. Mueller and Micro House International, Inc., *Micro House PC Hardware Library*. Macmillan Computer Publishing, July 1998.

[21] StorageReview.com, *Hard disk drive reference guide*.
http://www.storagereview.com/guide/.

[22] R. K. Roy, *Design of Experiments using the Taguchi approach: 16 steps to product and process improvement*. Wiley-Interscience, February 2001.

[23] R.G. Petersen, *Design and analysis of experiments*. Marcel Dekker Inc, 1985.

[24] J.S. Milton and R.H. Myers, *A first course in the theory of linear statistical models*. PWS-KENT Pub. Co., November 1998.

[25] KnowledgeTek, Inc., *Disk Drive Technology*. KnowledgeTek, Inc., 2005.

[26] S. Kraber, *Stat-Teaser, News from Stat-Ease, Inc.* Stat-Ease, Inc., May 2004.

[27] P. Sherman, *Demystifying Design of Experiments and related taboos*. Quality Digest Magazine, April 2008.
http://www.qualitydigest.com/inside/six-sigma-article/demystifying-design-experiments.