

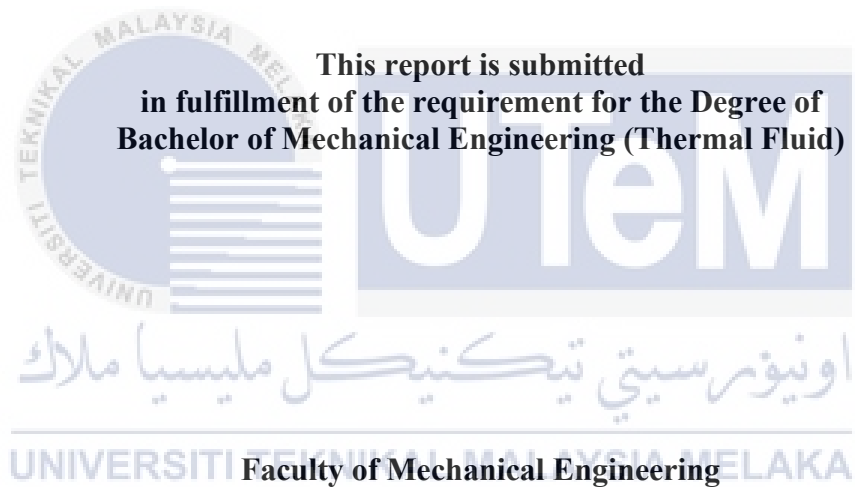
**THERMAL LOAD OF A SINGLE-FAN GRAPHIC PROCESSING UNIT (GPU)
ON A SERVER CONFIGURATION**



UNIVERSITI TEKNIKAL MALAYSIA MELAKA

**THERMAL LOAD OF A SINGLE-FAN GRAPHIC PROCESSING UNIT (GPU)
ON A SERVER CONFIGURATION**

MUHAMMAD SUFI RIDHWAN BIN MOHD KHIR




UNIVERSITI TEKNIKAL MALAYSIA MELAKA

JANUARY 2022

DECLARATION

I declare that this project report entitled “Thermal Load of A Single-Fan Graphic Processing Unit (GPU) on a Server Configuration” is the result of my own work except as cited in the references

	Signature	:
	Name	:
	Date	:

اونيورسيتي تیکنیکل ملیسيا ملاک
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

APPROVAL

I hereby declare that I have read this project report and in my opinion this report is sufficient in terms of scope and quality for the award of the degree of Bachelor of Mechanical Engineering (Thermal Fluid).



Signature :

Name of Supervisor :

Date :

UTeM

اونيورسيتي تيكنيكل مليسيا ملاك

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

DEDICATION

This project work is dedicated tom my beloved family and friends for always support and give me the strength throughout my life.



ABSTRACT

Graphics Processing Unit (GPU) has gone through many revolutionary changes throughout the decade. GPU's processing power is able to challenge the existing Central Processing Unit (CPU) in the task of running high-profile software. However, any electronics that flow current into it will generate heat. The thermal problem is the major problem for GPUs. The optimal operating GPU requires minimum fan speed and core temperature with the highest efficiency (hashing power). This research focused on obtaining the optimal response at a particular core clock and memory clock via the optimization tool of Design-Expert software. The responses were recorded while applying a constant amount of power supplied to the GPU. The GPU used is in the Single-Fan Graphic Processing Unit, ASUS NVIDIA® GeForce TURBO-GTX1070-8G. The relationship between the clocking (core and memory) and GPU responses (fan speed, core temperature, hash rate) was observed. By that, Central Composite Design (CCD) generated one equation for each fan speed, core temperature, and hash rate. Next, the optimization process suggests several new clocks settings that give the best performance than the current setting. Next, the optimization process offers several new clocks settings that give the best performance than the current setting. The best core and memory clock was selected for the validation and confirmation process. The results from the validation process prove that the predicted response was precise with less than a 2% deviation.

ABSTRAK

Unit Pemprosesan Grafik (GPU) telah melalui banyak perubahan revolusioner sepanjang dekad. Kuasa pemprosesan GPU mampu mencabar Unit Pemprosesan Pusat (CPU) yang sedia ada dalam tugas menjalankan perisian berprofil tinggi. Walau bagaimanapun, apabila komponen elektronik mengalirkan arus elektrik ke dalamnya ia akan menghasilkan sejumlah haba. Masalah haba adalah masalah utama untuk GPU. GPU yang beroperasi secara optimum memerlukan kelajuan kipas dan suhu teras yang minimum dengan kecekapan tertinggi (kuasa hashing). Penyelidikan ini akan memberi tumpuan untuk mendapatkan tindak balas optimum pada jam teras dan jam memori yang tertentu menggunakan alat pengoptimuman perisian Design-Expert. Tindak balas GPU telah direkodkan sepanjang kuasa yang malar dibekalkan kepada GPU. GPU yang digunakan dalam penyelidikan ini adalah Kipas Tunggal GPU yang bernama ASUS NVIDIA® GeForce TURBO-GTX1070-8G. Hubung kait antara jam (teras dan memori) dan tindak balas GPU (kelajuan kipas, suhu teras, kadar hash) direkodkan. Hasil daripada ini, Reka Bentuk Komposit Pusat (CCD) telah menjana satu persamaan untuk setiap kelajuan kipas, suhu teras dan kadar hash. Seterusnya, proses pengoptimuman telah mencadangkan beberapa tetapan jam baharu yang memberikan prestasi yang lebih baik daripada tetapan semasa. Bagi proses validasi dan pengesahan pula, satu tetapan jam (teras dan memori) terbaik baharu telah dipilih. Hasil proses validasi membuktikan bahawa respons yang diramalkan adalah tepat dengan sisihan kurang daripada 2%.

ACKNOWLEDGEMENTS

In the Name of Allah, the Most Merciful

First, I'd like to express my gratitude and praise to Allah the Almighty, my Creator, and Sustainer, for everything I've received throughout my life. Additionally, I'd like to express my gratitude to Universiti Teknikal Malaysia Melaka (UTeM) for giving me an opportunity to pursue the Final Year Project for 2 semesters as partial fulfillment of my requirement for my course, Bachelor of Mechanical Engineering with Hons.

I would like to thank my family, especially my mother and father, who have been my biggest supporters until my research was fully finished. My family has encouraged me attentively with their fullest and truest attention to accomplish my work with truthful self-confidence. My sincere appreciation to Dr. Muhammad Zulfattah Bin Zakaria as my supervisor for his encouragement, counsel, and inspiration. His unwavering patience in guiding and imparting priceless insights will be cherished for eternity. Additionally, I would like to express my gratitude to my project colleague at Universiti Teknikal Malaysia Melaka (UTeM), Mr. Hazim, Mr. Aizuddin and Mr. Haziq for their continuous support throughout my journey. Finally, I want to convey my appreciation to everyone who aided, encouraged, and motivated me to complete my education.

TABLE OF CONTENT

CHAPTER	CONTENT	PAGE
	ABSTRACT	vi
	ABSTRAK	vii
	ACKNOWLEDGEMENTS	vii
	TABLE OF CONTENT	ix
	LIST OF FIGURES	xii
	LIST OF ABBREVIATIONS	xv
	LIST OF SYMBOLS	xvii
CHAPTER 1	INTRODUCTION	1
	1.1 Background	1
	1.2 Problem Statement	3
	1.3 Objective	4
	1.4 Scope of Project	4
CHAPTER 2	LITERATURE REVIEW	5
	2.1 Overview	5
	2.2 Introduction of Graphic Processing Unit (GPU)	5
	2.3 Components Contributing Heat in Graphic Processing Unit	8
	2.4 Components of Cooling GPU System	11
	2.4.1 Heat Sink	12
	2.4.2 Heat Pipe	13
	2.4.3 Fan	14
	2.4.4 Thermal paste	15

2.5	Design of Fan	16
	2.5.1 Angle of Blade	16
	2.5.2 Design of Blade	18
	2.5.3 Numbers of Blade	22
	2.5.4 Fan Outer Diameter	26
2.6	Equipment and Software	30
	2.6.1 Graphics Processing Unit	30
	2.6.2 Afterburner Software	32
	2.6.3 Design Expert	33
	2.6.4 Response Surface Method	35
	2.6.5 Central Composite Design	36
	2.6.6 PheonixMiner Software	39
	2.7 Fundamental of Heat Transfer	41
CHAPTER 3	METHODOLOGY	44
	3.1 Introduction	44
	3.2 Flowchart	45
	3.3 Schematic Diagram	46
	3.4 Design of Experiment	46
	3.5 Experimental Setup	49
	3.6 Statistical Analysis and Optimization of Clocking and GPU Response	51
CHAPTER 4	RESULT AND DISCUSSION	52
	4.1 Result	52
	4.2 CCD and ANOVA (Fit)	54
	4.2.1 Fan Speed Response	54
	4.2.2 Core Temperature Response	58
	4.2.3 Hash Rate Response	62
	4.3 Effect of Independent Variables on Response Variables (Contour Plot and Equation)	65
	4.3.1 Fan Speed	65
	4.3.2 Core Temperature	67

	4.3.3 Hash Rate	69
	4.4 Optimization of Single-Fan Graphic Processing Unit (GPU)	71
CHAPTER 5	CONCLUSION AND RECOMMENDATION	76
	5.1 Conclusion	76
	5.2 Recommendation	78
	REFERENCES	79
	APPENDIX	85



LIST OF TABLES

TABLE	TITLE	PAGE
2.1	Specification of of Single-Fan Graphic Processing Unit (GPU)	17
2.2	Blade fans simulation results	20
2.3	Factors contribution in optimum combination of the fan for analytical results	27
2.4	Factors contribution in optimum combination of the fan for experimental results	28
2.5	Specification of Single-Fan Graphic Processing Unit (GPU)	30
3.1	Independent variables and their corresponding levels for GPU clocking	47
4.1	Experiment design and response value obtained by the GPU	53
4.2	ANOVA for fan speed response	54
4.3	Fit statistics for fan speed response	55
4.4	ANOVA for core temperature response	58
4.5	Fit statistics for core temperature response	58
4.6	ANOVA for hash rate response	62
4.7	Fits statistics for hash rate response	62
4.8	Optimization of Single-Fan Graphic Processing Unit (GPU)	71
4.9	Numerical optimization solution	73

LIST OF FIGURES

FIGURE	TITLE	PAGE
2.1	Gigabyte GeForce RTX™ 3080 TURBO 10G	7
2.2	Components in Graphic Processing Unit	11
2.3	Heat pipe mechanism	14
2.4	The system variables for various fan blades and fabricated original and optimal fan	17
2.5	The effect of blade angle on CFM	18
2.6	Design of blade fan	19
2.7	Mass flow rates using three kinds of fan total pressure efficiency curve.	20
2.8	Comparison of blade profile before and after optimization	21
2.9	Fan performance curve comparison	21
2.10	Verticity diagram before and after optimization	22
2.11	Flow rate-static pressure curve	23
2.12	Flow rate-efficiency curve	23
2.13	The static pressure field distribution	23
2.14	Fan blade number against overall sound pressure level	24
2.15	Changes in number of blades	25
2.16	Characteristics curves of the fans with different number of blades	26
2.17	Experimental - Responses of parameters by mean effects plot for SN ratios	29

	Analytical - Responses of parameters by mean effects plot for	29
2.18	SN ratio	
2.19	Afterburner software	33
2.20	Design Expert Software	34
2.21	Generation in a central composite design of points	36
2.22	Central Composite Design (CCD) Flow Diagram	38
2.23	PhoenixMiner Software	40
3.1	Flowchart of methodology	45
3.2	Schematic diagram	46
3.3	Classic central for central composite design for 2 factors	48
3.4	Central Composite Design layout in Design Expert	50
3.5	Design CCD (Actual) in Design Expert	51
4.1	Predicted vs Actual fan speed	57
4.2	Fan speed normal plot	57
4.3	Predicted vs Actual Core Temperature	61
4.4	Core temperature normal plot	61
4.5	Predicted vs Actual Hashrate	64
4.6	Contour plot for the combined effect of core clock (A), memory clock (B), and fan speed	65
4.7	Contour plot for the combined effect of core clock (A), memory clock (B), and core temperature	67
4.8	Contour plot for the combined effect of core clock (A), memory clock (B), and hash rate	69

LIST OF ABBREVIATIONS

GPU	Graphic Processing Unit
CPU	Central Processing Unit
RAM	Random Access Memory
VRAM	Video Random Access Memory
PSU	Power Supply Unit
CFD	Computational Fluid Dynamic
ALU	Arithmetic Logic Unit
TIMs	Thermal Interface Materials
NACA	National Advisory Committee for Aeronautic
HS	Heat Sink
SN	Signal to Noise
PCIe	Peripheral Component Interconnect Express
SIMT	Single-Instruction Multiple-Thread
HPC	High Performance Computing

RAMDAC Random Access Memory Digital-To-Analogue Converter

HDD Hard Disc Drive

SSD Solid-State Drive

VRM Voltage Regulator Module

CCD Central Composite Design

RSM Response Surface Methodology

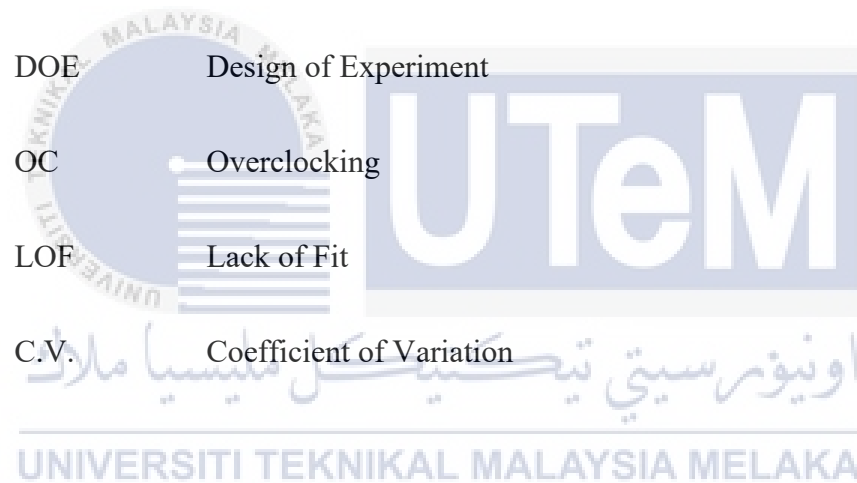
ANOVA Analysis of Variance

DOE Design of Experiment

OC Overclocking

LOF Lack of Fit

C.V. Coefficient of Variation



LIST OF SYMBOLS

Q	= Heat transferred
W	= Watt
J	= Joule
K	= Kelvin
V	= Voltage
m	= Mass, kg
CFM	= Cubic Feet per Minute
c_p	= Specific heat, $J/kg \cdot K$
ΔT	= Temperature difference, K
$Q_{conduction}$	= Rate of heat conduction, W
k	= Thermal conductivity, $W/m \cdot K$
A	= Cross sectional area, m^2
Δx	= Thickness, m
$Q_{convection}$	= Rate of heat transfer, W
h	= Convective are for heat transfer coefficient $W/m^2 \cdot K$
Pa	= Pressure
%	= Percent

$Q_{emit,max}$	= Rate of heat transfer radiation, W
ε	= Emissivity
σ	= Stefan-Boltzmann constant ($5.670 \times 10^{-8} \text{W/m}^2 \cdot \text{K}^4$)
rpm	= Revolution per minute
MH/s	= Mega hash per second
$^{\circ}\text{C}$	= Degree Celsius



CHAPTER 1

INTRODUCTION

1.1 Background

A graphics processing unit (GPU) is a computer that is designed specifically to produce the image seen on the monitor. To the more significant market, it is recognized for offering consistent, more beautiful video and game graphics that users demand today. The GPU was released in 1999. In the early 1990s, all application was entirely under the control of the Central Processing Unit (CPU). A GPU combined with a CPU can assist in tasks that involve a large amount of computational work, such as rendering. Because of this, many machines, including entry-level computers, are fitted with GPUs (Jablin et al., 2012). Since the GPU can do several calculations simultaneously, this speed-up application processing. Also, advances in information technology (such as faster internet and cell phone access) have resulted in an extraordinary increase in the need for GPU for computers and computer components.

Nowadays, current technology does not meet the expectations of the majority of consumers. In addition, high processing power is also needed to build complex applications. For example, the games and higher-performance applications like Solid-Work, Catia and other programs like it are all dependent on powerful technology such as

mining cryptocurrency need greater processing power. On the other hand, CPU-bound applications would not support real-time multimedia. Because of its current design, it often cannot handle multimedia applications in real-time for the vast majority of personal computers, and video processing is not quick enough to be done by a CPU.

According to Ganesh Iyer & Dipakumar Pawar (2018) in mining cryptocurrency, GPU mining is slightly profitable because CPUs cannot handle computational stresses where GPUs have the upper hand. Moreover, the use of GPUs has the following benefits instead of these technologies: (1) GPUs allow for consolidation that simplifies design (2) due to cost savings and allow the addition of a GPU to production lines to be cheap (3) because of their programmability (Muyan-Özçelik et al., 2011).

GPUs greatly enhance programmability. Purpose-designed graphics cards were designed to handle computation- and timing-demanding functions, allowing threads to be performed on a greater number of the processing unit. The efficiency of today's multicore CPUs can be surpassed when it has a huge number of threads. GPU computing has been successfully used in diverse engineering and scientific problems requiring numerical analysis (Martínez-Frutos et al., 2017). This is because GPUs blend logic and processing units, it can be built from cores, each of which has a unique feature called arithmetic logic units (ALUs). One functional unit processes each processor thread of execution in this analysis. Each thread in the application is assigned to one functional unit for processing, and those are referred to as thread processors. All thread processors in a GPU core perform the exact instructions, as they share a control unit (Smistad et al., 2015).

1.2 Problem Statement

Additionally, semiconductor devices have caused an issue with the GPU. When a GPU is intensive, it generates much heat for video and scientific simulations. A cooling system is needed to disperse the heat. However, with the other methods considered, assume the cooling system does not cooperate or is inefficient. If this is enabled, the GPU will get too hot, which will lead to a chip failure. This action will impact the graphic card's results. An item that would allow the electronics to be damaged would exist in the worst-case scenario. According to Choi et al., (2012) the typically permissible operating temperature of a CPU is below 70°C. The reliability of the chips then decreases by 10% for every 2°C above the allowable operating temperature. Also, Vargas-Vazquez et al., n.d. (2018) had claimed that electronic chip cooling is one of the most challenging challenges as the demand for increasingly powerful computer systems grows exponentially.

To overcome this issue, a study of heat dissipation techniques for single-fan graphics processing units (GPU) and GPU controllers is ongoing. Several external variables and factors and a few device factors must be accounted for to keep the GPU operating at its peak performance and achieve an ideal temperature. Heat sink, heat pipe and fans help the GPU processor maintain a stable temperature. While, in this particular research, core clock and memory clock were analyzed to find the optimum fan speed, low core temperature and highest efficiency (hashing power) of the Single-Fan Graphic Processing Unit (GPU) on a server configuration.

1.3 Objectives

The objectives of the project are:

1. To validate the relationship between clocking (core and memory) and GPU responses (fan load, core temperature, power consumption, and hash rate).
2. To determine the optimal fan speed, core temperature and the highest possible efficiency (hashing power) at a specific clock and memory clock via optimization tool.

1.4 Scopes of Project

The scopes of this project are:

1. Gathering literature review of GPUs component.
2. Getting the Single-Fan Graphic Processing Unit (GPU) response (fan speed, core temperature, power consumption and hash rate) at a certain core and memory clock by using Afterburner and PhoenixMiner software.
3. Finding numerical optimization solutions by using Design-Expert software.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

This chapter mainly describes getting a clearer and better understanding of the title. This literature review will be briefly divided into four main keywords: Graphic Processing Unit (GPU), Thermal Management, Design of Fan, Equipment and Software. This combination of these four keywords will be explained thoroughly to understand this project fully. All equations and fundamental knowledge that apply in this project will be demonstrated in this section. The ultimate goal for this research is to find optimal fan speed, core temperature and the highest possible efficiency (hashing power) at a certain clock and memory clock.

2.2 Introduction of Graphic Processing Unit (GPU)

Graphic Processing unit has been a key player in both business and consumer applications. The GPU is designed for parallel processing and is commonly used in graphics and video rendering applications. Though graphics cards are well known for their

gaming capability, they are increasingly used in advanced AI (Artificial Intelligence). GPU consists of multiple components: graphics processor, video memory, Random Access Memory digital-to-analog converter (RAMDAC), and driver software. GPUs aim to speed up 3D rendering. As time went on, their software and hardware became more capable, they gained flexibility and generalization abilities. According to Muyan-Özçelik et al., (2011), As seen in previous studies, graphics processing units (GPUs) are multiplying in application areas outside of traditional graphics that are well suited for their capabilities. This made it possible for the more experienced graphics programmers to achieve more striking visual effects and complex lighting and shading in the modelling software such as AutoCAD, Solidworks, and Catia. Several other software developers have also discovered the High-Performance Computing (HPC) and profound learning potential of GPUs. Other significant benefits are increased performance for games and other tasks that take advantage of parallel data access and banking memory, caches, and single-instruction multiple-thread (SIMT) execution, among others (Kerr et al., 2012).

GPU is divided into Integrated Graphics Processing Units and Discrete Graphics Processing Units. An integrated GPU does not exist independently and is combined with the CPU on the same circuit board. While for discrete GPU, it has its own boards and circuit and is commonly placed on PCIe (Peripheral Component Interconnect Express) slots. Nowadays, most GPUs on the market today are currently on the integrated level of GPUs. For integrated graphics, both the graphics processing unit and the CPU work together to render the monitor's image. A motherboard CPU that includes fully integrated GPU results in thinner and lighter and more energy-efficient systems, and less powerful CPUs lowers the cost. For high resource-intensive applications, such as 3D games, it is

superior to use a discrete GPU (usually called a 'dedicated graphics card' or 'high-performance graphics card.' When processing capacity increases, energy consumption and heat generation expenses will also increase. With GPU-accelerated rendering, optimum performance generally comes with extra cooling dedicated GPUs. Maroosi et al. (2013) show that simulation speed increases by 15 to 38 times when using shared memory relative to sequential simulation.



Figure 2. 1: Gigabyte GeForce RTX™ 3080 TURBO 10G

2.3 Components Contributing Heat in Graphic Processing Unit

As a matter of fact, GPUs are engineered to work in the worst-case scenario in terms of the process, temperature, and voltage change (Zamani et al., 2020). A GPU naturally gets heated, and becoming that warm is entirely natural. This is because some components might be contributing to heat generation in the system. The first component is Control Processing Unit (CPU). CPU is a crucial component in the GPU as it works as a control unit and the core of the computer. Control unit means it describes any instructions in the computer and processes the data in the computer. CPUs are constructed using billions or even trillions of transistors, these parts have finite, non-negligible resistance and are electrically connected to the rest of the circuitry through current flow, which depends on the operation and state of the circuit they are part of. As the CPU uses electricity as its main power, the CPU can switch on the flow of electric signals by allowing the microscopic transistors to operate or off by stopping them.

As a result, electrical current passes through the CPU, or is held up and turns into heat. Thermoelectric materials cause temperature variations by running electricity through them Annapragada et al., (2012). When current flows through the processor material, a small amount of heat is generated, which reduces the functionality of thermoelectric modules (Krishnan, Garimella, Chrysler, & Mahajan, 2007). Overclocking GPUs has also contributed to an increase in CPU heat generation. While increasing the CPU's frequency to a higher level can help reduce execution time, it could also harm the CPU. This will lead to the damage of the CPU because of an increase in heat-generating capacity.

The second component contributing to the power supply unit (PSU) heat generation in the GPU. The graphics card is one of the most power-hungry electric elements when it comes to powerful PCs. Using a graphics card with a budget or entry-level GPU, the power requirements will be less. However, due to the higher GPU requirements of the mid-range and higher-end GPUs, they will require slightly more power. While mid-range and high-end graphics cards obtain their power from a PCI-E x16 slot on the motherboard, entry-level or budget cards receive power from the slot itself and are limited to a maximum of 75W.

Therefore, power-hungry cards require power from the PSU, even if the system only uses a small amount of power (6-pin and 8-pin PCIe connectors). A system failure and potential component damage can occur due to a power supply overheating due to overloading, fan failure, poor airflow outside and inside the system. The expansion slots have traditionally been the leading source of power supply overload problems. A typical scenario for overloading the system power supply is having several hard drives, CD-ROM drives, and floppy drives all connected to the same power supply.

Video Random Access Memory (VRAM) is another component contributing to heat generation. GPUs' VRAM is essential because the operating system, gaming textures, and lighting effects are all located in VRAM, so the processor in the device can use that data without too much hassle. In addition, random-access storage, such as a hard disc drive (HDD), solid-state drive (SSD), or optical drive, is often easier to read and write from than to read and write from other forms of storage, such as a hard disc drive (HDD), solid-state drive (SSD), or optical drive. VRAM also produces heat because it holds a chip of a transistor and capacitor when electricity flows over these components.

The last component is Voltage Regulator Module (VRM). The first responsibility of a VRM is to transform the 12-volt power coming from the computer's power supply to a usable voltage. Usually, processors run at 1.1V to 1.3V. A voltage that is too high can damage delicate electronics. Powering a processor often requires precise voltage delivery, and the supplied voltage must be exact to the exact voltage specification. VRM is always described as underappreciated parts, do not think they are unimportant. Because without a proper VRM, CPU or GPU will not receive power at a consistent voltage. An underperforming VRM will decrease a processor's overall output and inhibit it from handling increased workloads. Overclocking GPU can cause VRM to become so hot indirectly to unintended system shutdowns.

By end large, it can be concluded that heat dissipation of GPU comes CPU, VRAM, VRM and power supply. To ensure that optimum temperature can be maintained, the CPU, VRAM and VRM are located at the heart of the GPU so that a broad range of cooling agents that are also components in GPU can be supported. These components' location is typically in the centre of the GPU to maintain the GPU at the optimum temperature to prevent damage.

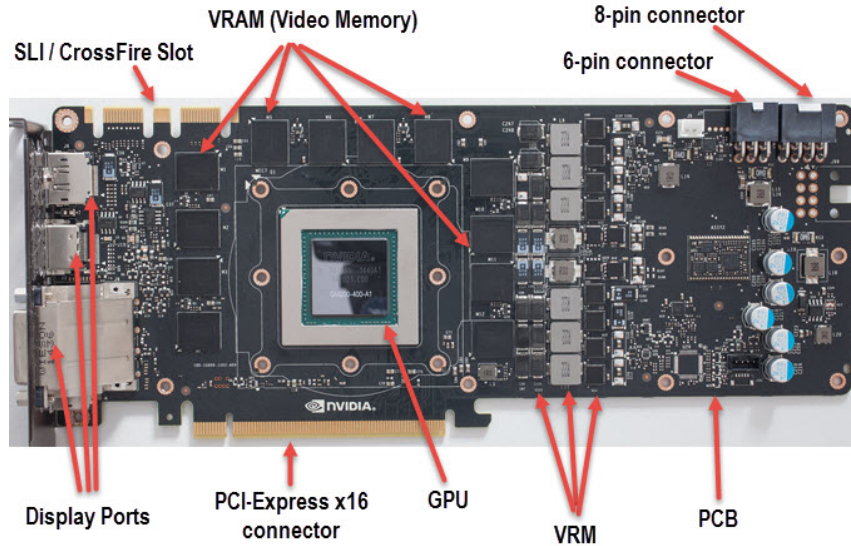


Figure 2. 2: Components in Graphic Processing Unit

2.4 Components of Cooling GPU System

Due to the inherent characteristics of computer technology, it is unavoidable that heat will be generated. However, too much heat can cause computer hardware to slow down. An example of this is when the CPU gets too hot, in which case a mechanism automatically reduces the system's efficiency to protect the GPU. So, to keep the environment stable, heat must be dissipated quickly from the system or make sure the components are always in optimum temperature.

Overheating is dangerous hence the system must be cooled at whatever cost, to avoid damage from overheating. (Huang & Gau, 2012). Many factors can be caused by overheating GPUs such as poor ventilation, dust and malfunction or inadequate cooling. Therefore, ensuring the optimum GPU temperature is critical and cannot be considered a

simple task affecting GPU efficiency and lifespan. Many components are heat sinks, fans, heat piping components, and thermal paste. The cooling components are detailed on their functions, the component form and the component problem. Since the operating temperature increases, electrical components undergo a steep decline in durability, performance, and dependability. Consequently, electronic devices need to contain adequate thermal management techniques (Feng et al., 2018).

2.4.1 Heat sink

A heat sink is a component that increases heat flow away from a hot system. It does this by increasing the working surface of the system and the volume of low-temperature fluid moving through its extended surface. Based on the configuration of each unit, heat sinks are mostly aesthetics, architecture, and ultimate capabilities. Heat sink operates as one of the heat removing agents in the GPU. The GPU transfers heat to the heatsink by conduction. The primary heat transmission mechanism from the heatsink is convection, but radiation has a modest effect.

There are two different heat sink types which are passive heat sinks. Using heat sinks to improve passive cooling performance is a common practice (Feng et al., 2018). Passive heat sinks rely on natural convection and have no mechanical components. They are hence 100% dependable. Passive heat sinks consist of a radiator finned in aluminium which distributes the heat by convection. A continuous airflow must go through the fins for passive heat sinks to work. Next is, active heat sink.

Forced convection occurs in this active heat sink where a further device such as a fan or bladder causes the movement of the fluid particles. Active heat sinks use the computer's power supply and may incorporate a fan. A heat sink is a device composed of conductive metal and is used to absorb heat from elements that operate at high temperatures and then distribute that heat elsewhere (Dhaiban et al., 2020). The fins increase and maximize the area contact with the electronic component.

2.4.2 Heat Pipe

A heat pipe is a tube whereby a heat transfer liquid circulates. The heat pipe plays a vital role in managing heat flow away from the GPU's components. Heat pipes use evaporative-condensation cycles to convey vast amounts of energy (Mahdavi et al., 2018). A heat pipe is generally added to the heat sink. A typical heat pipe consists of a vessel with a winding structure inside walls. The vessel is first emptied, then loaded and sealed with a working fluid. The working fluid evaporates from liquid to vapor when the heat pipe is heated at one end (phase change). According to Sanhan et al., (2020), the heat pipe's thermal performance can also be increased by including nanofluid particles in the base fluid. The vapour travels to the other end of the thermal pipe at close sonic speed utilizing pressure created by the temperature difference through the hollow core, where the heat is dissipated using a heat sink or other means.

The vapour condenses to liquid and simultaneously releases heat. The fluid then returns into liquid form, and the helical structure pumps the fluid with capillary force back

to the original end. Indirectly shows that heat pipes use natural convection to dissipate heat. Heating through a heat pipe can transfer up to 1000 times more thermal energy than copper, as it is the best conductor, but only when the heat pipe's temperature is less than - 17°C per foot (The Heat Pipe Advantage, n.d.). There are several advantages of using heat pipes: passive, long-lasting, minimum maintenance, and flexibility in size.

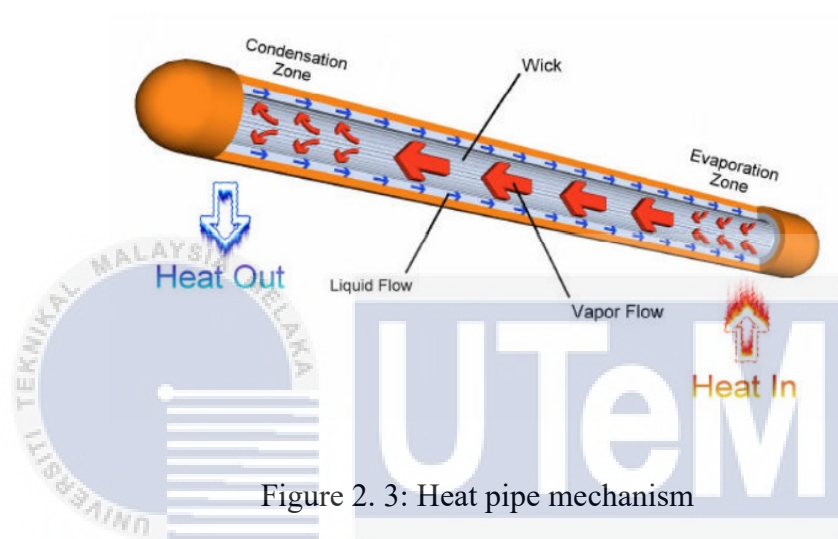


Figure 2. 3: Heat pipe mechanism

2.4.3 Fan اونيورسيتي تيكنيكل مليسيا ملاك

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Today, as the GPU is compacter, more functional and located in a small room, the heat from the GPU system needs to dissipate more seriously and the thermal problem cannot be solved by only using the heat sink. Generally, in a blower type GPU, a fan acts as an active cooler, a physical device that keeps the entire computer or computer cool by air moving from or to the computer or the component. Ventilators like fans usually linked to components are used in connection with a heat sink for increasing the area of the heated air surface, thereby boosting the cooling efficiency. The fan's speed can be controlled by software or manually adjusted, and it is not necessarily an automatic procedure.

A BIOS computer can automatically adjust the computer speed of the integrated fan system. In high-end systems such as computers, the amount of power dissipation continues to develop fast, while the description of silicon die-temperature is still constant, approximately 450°C . This electrical equipment is frequently cooled by forced airflow with the help of axial fans or centrifugal blowers. Electronic gadgets and servers often utilize the axial fan for cooling, reported in Rajabi et al.,(2017).

The design of blower type GPU is like a turbine fan that sucks the air and distributes it laterally, very differently from a standard propeller-blade fan. Some air hits the front ventilator housing, while around one-third goes to the heatsink of metal attached to the GPU core.

2.4.4 Thermal paste

Thermal paste is a highly high heat conductive material that improves heat conduction between two objects. Thermal Interface Materials (TIMs) can be heat conductors up to 100 times higher than air. Thermal paste is normally placed directly on the CPU or other IC to ensure a more direct heat transmission from the heat sink to the chip. It generates a uniform surface area covering 100% of the CPU surface area.

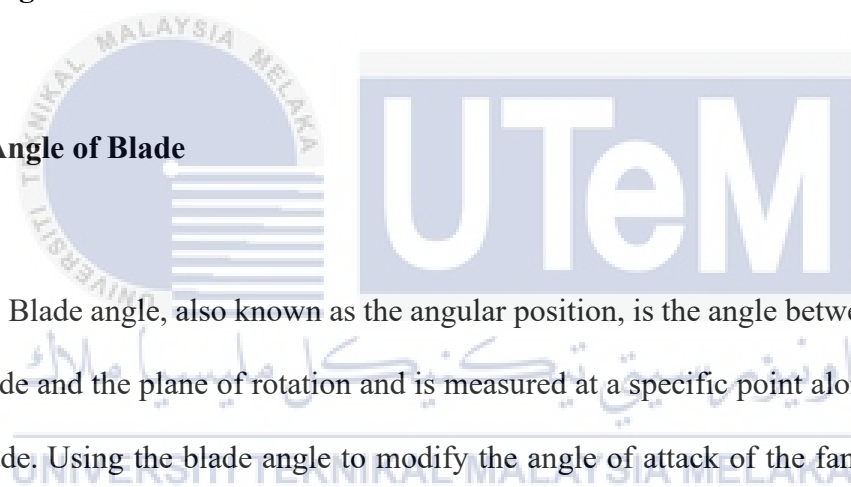
Thermal paste also enables the heat to be transferred from the CPU to the heat sink and keeps the CPU cooled and nominal. It also prevents the creation and loss of air gaps between the CPU and the heat sink, as the air is a poor heat conductor. This chemical is a

cost-effective option and is sold in any place. The lack of thermal paste between the CPU and the heat sink will eventually result in overheating and failure.

The electrically insulative thermal-conductive property of thermal paste makes it possible for manufacturers and builders to use this in nearly all electronic applications that require connections between thermal management systems. However, it requires certain chemical compositions to achieve these features.

2.5 Design of Fan

2.5.1 Angle of Blade



Blade angle, also known as the angular position, is the angle between the chord of the blade and the plane of rotation and is measured at a specific point along the length of the blade. Using the blade angle to modify the angle of attack of the fan is an excellent way. Adjusting the fan blade angle is required to offer the most efficient angle of attack, whether the fan is running at a constant speed or is at varying speeds. Based on research done by Huang & Gau (2012), redesigned optimal fan blade can increase fan airflow, thereby improving the performance of the axial-flow fan. Redesigned optimal fan blade includes the setting angle of blade, blade root chord, blade end chord. The number of blades, fan speed and fan diameter, are kept constant during the experimental setup. The system variables for the original fan blade and redesign fan blade are shown in the Figure 2.4.

Table 2. 1 : Specification of original, optimal and designed fan blade

	Original fan blade	Optimal fan blade	Designed fan blade #1	Designed fan blade #2
NACA airfoil (mpta)	4609	5608	4506	4609
Blade root chord (Lr. mm)	18.0	19.38	17.59	18.17
Blade end chord (Le. mm)	33.0	37.86	30.64	34.18
Setting angle (0.°)	47	59	33.18	52.22
Number of blades	7	7	7	7
Fan speed (o. rpm)	2500	2500	2500	2500
Hub diameter mm)	40.0	40.0	40.0	40.0
Fan diameter (D. mm)	110.0	110.0	110.0	110.0
Hub height mm)	15.0	15.0	15.0	15.0
Fan gap (mm)	2.0	2.0	2.0	2.0
Air volume flow rate (Q. CFM)	78.19	88.84	63.0	83.0

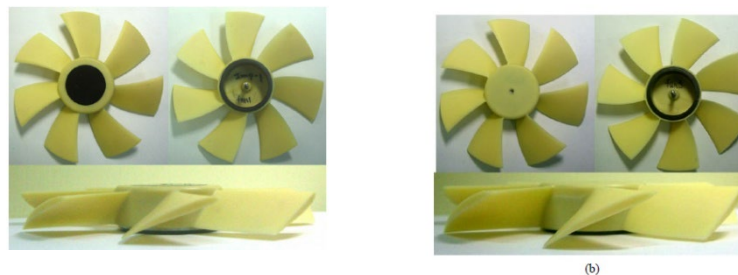


Figure 2. 4 : The system variables for various fan blades and fabricated original and optimal fan

By analyzing the Table 2.1 above, we can see that the setting angle of the blade affects the air volume of flow rate. When the setting angle was set at 59 degrees, air volume flow rate was highest at 88.84 CFM, while the smaller the setting angle, the lower the air volume flow rate, which is around 63 CFM. Pirunkaset et al. (2008) claim that blade angles also influence small cooling tower performance.

A setting fan blade angle of 59°, 67°, 75°, and 83° has been seen in the study for 5 tonnes of refrigeration in a cooling tower. Taking Figure 2.5 into consideration, it can be determined that the CFM for a 59° blade angle has the greatest flow, followed by 67°, 75°, and 83°. The poor performance may be attributed to fan blades that have not been properly installed by the manufacturer. So, it is crucial to find the optimum angle of blade as it affects the cooling system's performance and efficiency.

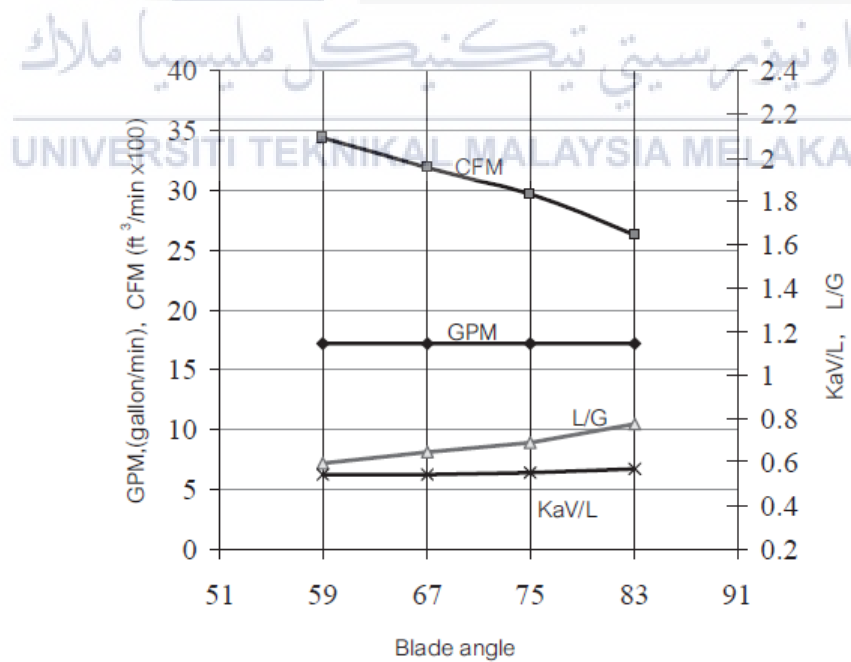


Figure 2. 5 : The effect of blade angle on CFM

2.5.2 Design of Blade

An axial fan is a type of fan that uses an axial flow of gas, rotating about an axis, to pull in the fresh air. Flow entry and exit of this fan is circular. The fan's flow producing function can be facilitated by creating a pressure difference and, as a result, a force which drives the flow of air through the fan. Wu & Huang (2019) state that, to improve lift-drag ratio, it is possible to optimize the chord length and installation angle of the blade. This study evaluated three design options: straight blades, C-type blades, and forward-swept blades as shown in Figure 2.6.



Figure 2. 6 : Design of blade fan

Table 2.2 shows a quantitative comparison between the three fan designs showed that the forward-swept blade fan was the most efficient overall, with the C-type blade fan being just slightly lower. Meanwhile, the straight blade fan was the least efficient overall. When operating under non-design conditions, forward-swept blade and C-type blade aerodynamics are better than straight blades. Forward swept blade, and C-type which is low noise are the main advantages of this model.

Table 2. 2 : Blade fans simulation results

	Straight blade	Forward-sweep blade	C-type blade	Units
Speed	985	985	985	$\left[\frac{r}{min}\right]$
Volume flow rate	50.9628	50.6329	51.1943	$\left[\frac{m^3}{s}\right]$
Pressure rise	1496.05	1520.47	1405.6	[Pa]
Total pressure efficiency	93.5621	94.8116	94.0066	[%]

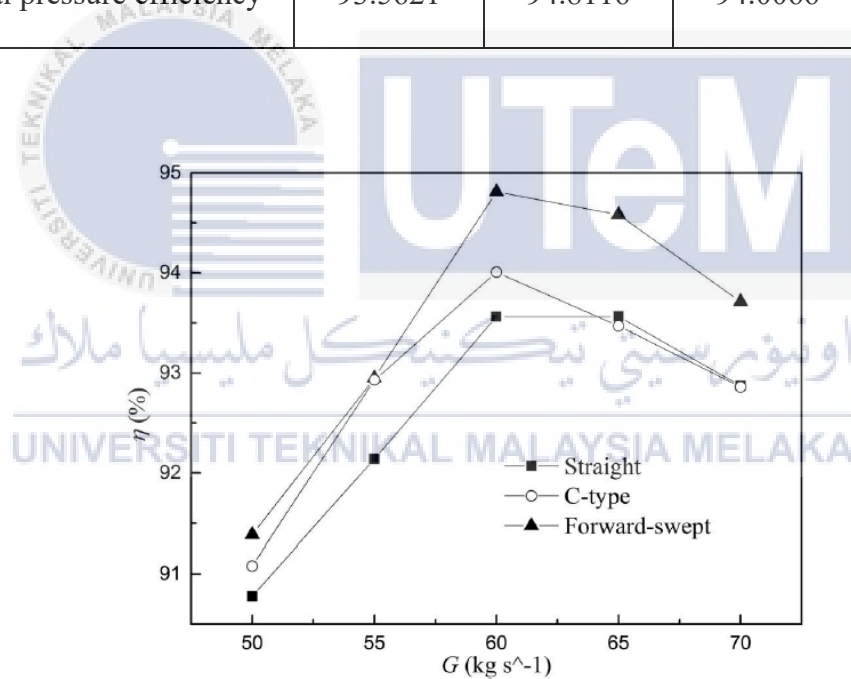


Figure 2. 7 : Mass flow rates using three kinds of fan total pressure efficiency curve.

Research published by Zhou et al. (2021) found that the blade design of a centrifugal fan can also improve its aerodynamic performance. Based on the research, blade fans for forward-curved centrifugal are designed using three different methods:

using NACA airfoils instead of typical single-arc blades, using double-arc or multi-arc blades, and using impeller structure size criteria to select the ideal combination.

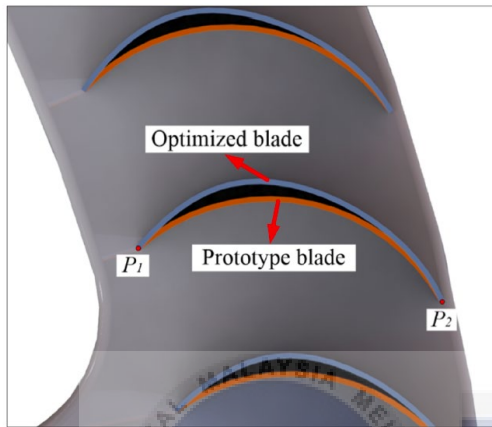


Figure 2. 8 : Comparison of blade profile before and after optimization

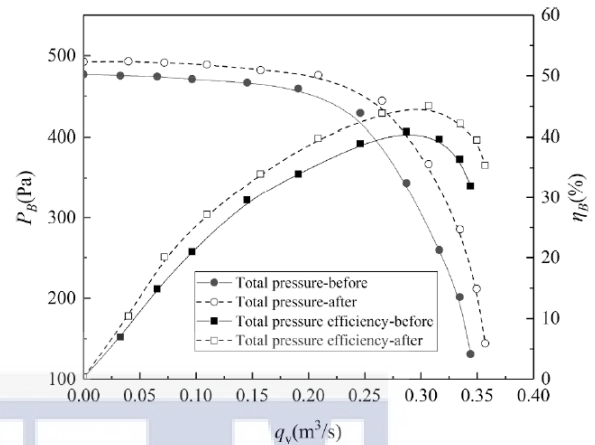


Figure 2. 9 : Fan performance curve comparison

Figure 2.8 is derived from Figure 2.9, and it indicates that a substantial change in the blade profile happens at the intake side, with the blade profile at the exit changing less. With this inlet and outlet installation angle reduction, the inlet installation angle has decreased by 4.2° , while the outlet installation angle has increased by 1.8° . When the total air volume increased by $1.18 \text{ m}^3/\text{min}$ at the design operating point, the pressure efficiency increased by 4.21% . Improvement in efficiency and flow rate can be used to reduce energy usage. Reducing the vortex size and dispersion range at the leading edge of the fan enhances the internal flow structure. To reduce the noise generated by fans, vorticity must be avoided.

The pattern shown in Figure 2.10 is the result of an optimization. The internal flow structure of the fan has been strengthened since the vortex size and range in this area have been decreased. Also, the reduction of vorticity might be accompanied by a reduction in fan noise. In this case, the divergence of the Coriolis acceleration caused by the flow (an apparent circulation) causes the noise. By analyzing vortex size trends, we may infer that decreasing vortex size also leads to a decrease in the divergence of the Coriolis acceleration the airflow gets, which decreases aerodynamic noise.

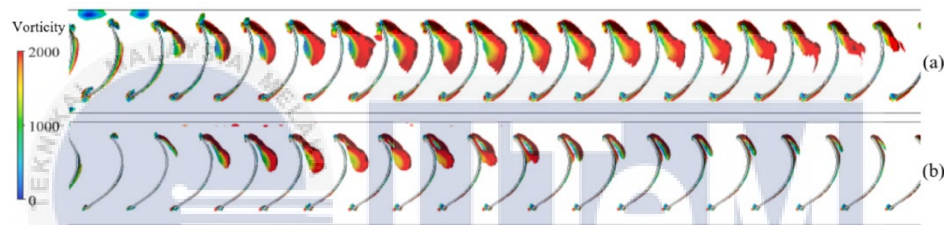


Figure 2. 10 : Vorticity diagram before and after optimization

2.5.3 Numbers of Blade

The outer portion of the blade does the most work for the fan. The number of blades is therefore important. According to Zhang and Jin (2011), there are several influences on the aerodynamic and noise performance of tiny axial fans, including blade number, hub ratio, blade angle, and blade number.

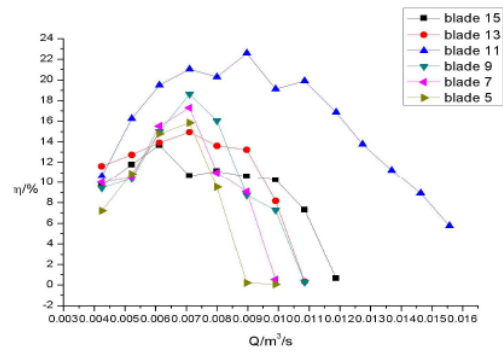
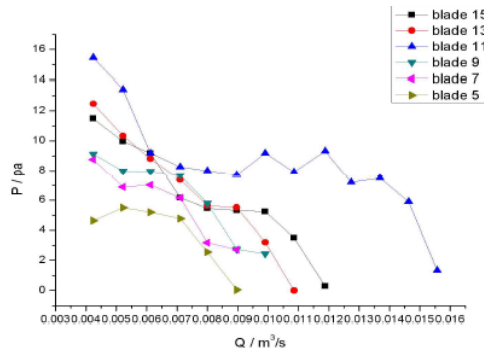


Figure 2. 11 : Flow rate-static pressure curve

Figure 2. 12 : Flow rate-efficiency curve

Based on Figure 2.11 and Figure 2.12, total pressure and efficiency are often higher when the number of blades increases. Flow range should be large, and total pressure and efficiency are at their peak when the number of blades is 11. Therefore, the flow loss is located in a point where the pressure gradient is least, and efficiency is maximum since the static pressure gradient is the smallest.

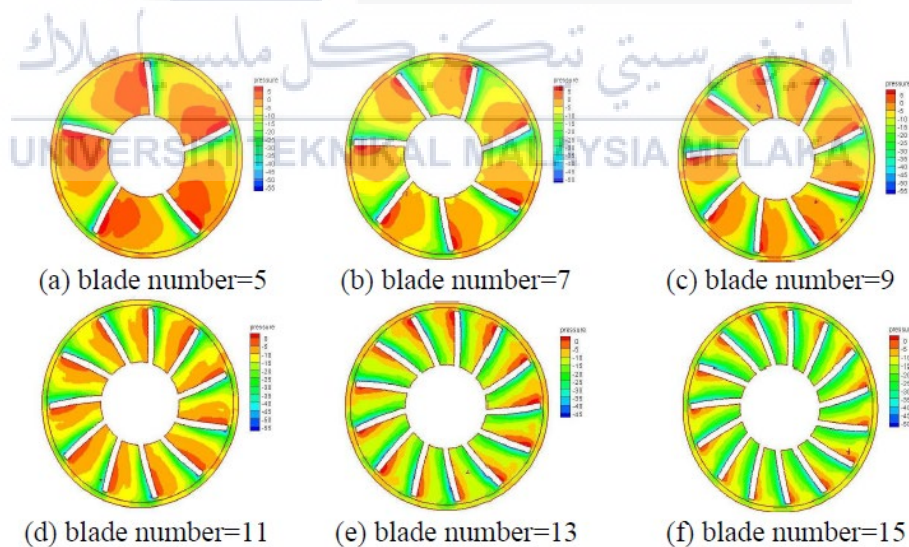


Figure 2. 13 : The static pressure field distribution

Figure 2.13 shows high pressure and low pressure exist in flow fields. In general, high pressure is larger than low pressure initially and low pressure gradually increases as blade numbers grow. Flow pressure distribution is most even when blade number is 11.

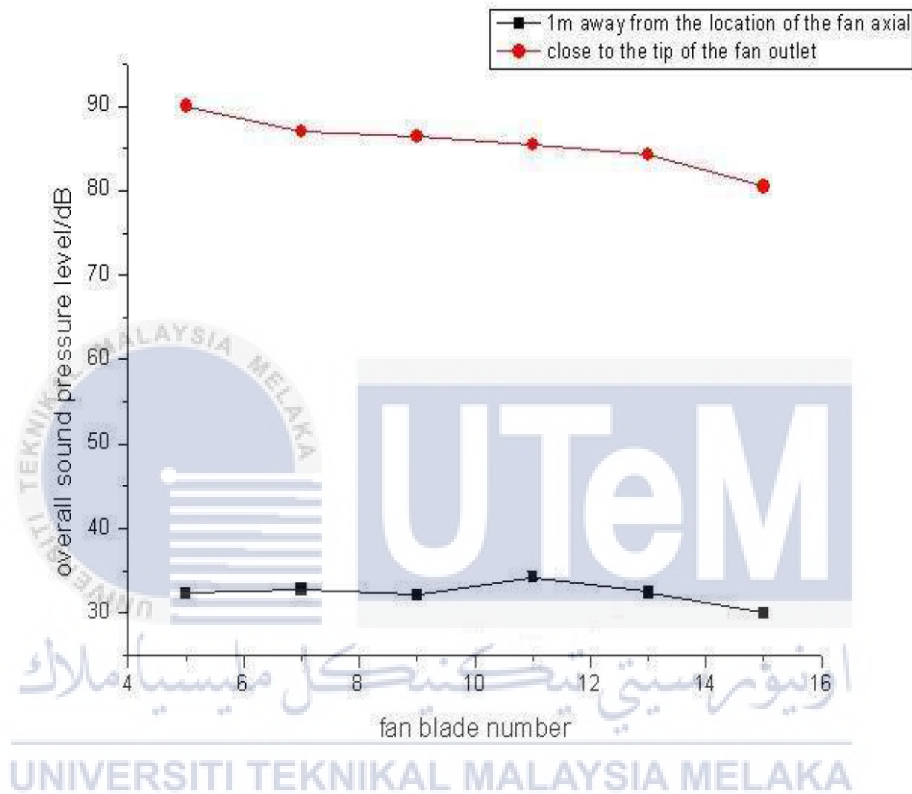


Figure 2. 14 : Fan blade number against overall sound pressure level

In fan constant flow fields, Eddy Current is one of the key causes to generate broadband noise and the effectiveness of Eddy Current mostly occurs in the tip clearance, which then influences the overall broadband noise. Sound pressure level of aerodynamic noise, which is in the proximity of the tip of the fan blades, drops as blade numbers rise. However, for fan model of 11 blade number, at 1 meter distant from the impeller center point, the sound pressure level increases again as shown in Figure 2.14. This may be due to the static electricity generated by the fan blade clearance.

According to Rajabi et al., (2017), as the number of blades increases, the maximum flow rate increases proportionally. The rotational speed has been steady at 2900 revolutions per minute, though some blades have been modified. In this case, the study classifies fans into three types: 4, 5, and 6 blades, respectively as shown in Figure 2.15.

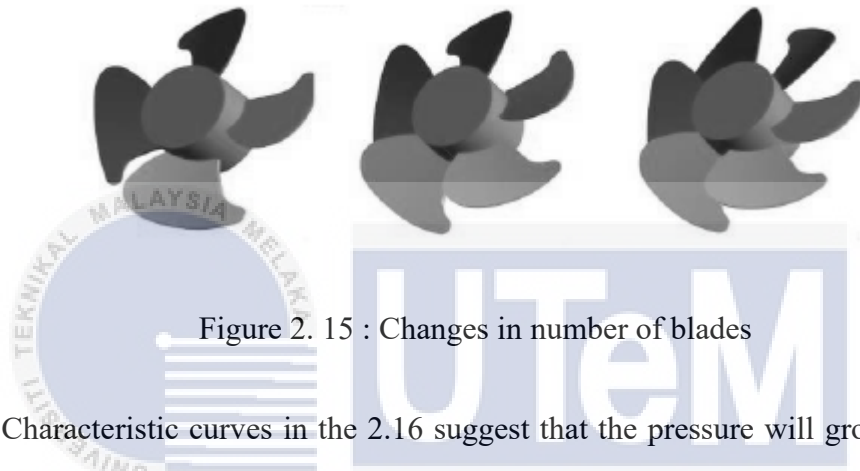


Figure 2. 15 : Changes in number of blades

Characteristic curves in the 2.16 suggest that the pressure will grow as the flow rate is increased from 200 CFM to 400 CFM. When the flow rate is increased from 400 CFM to 800 CFM, the static pressure is greatly reduced. Applying six bladed propellers increases maximum pressure at 400 CFM by 32 percent. An increase in the number of blades causes the surface of the blades to become larger, increasing the amount of energy transferred when they meet fluid.

This increases outflow pressure. When large flow rates are present, the most efficient separations are used and increasing the number of blades has limited impact on the outlet pressure. The Figure 2.16 shows that at the maximum flow rate of 800 CFM, the static pressure produced by the various number of blades is almost equivalent to 42 Pa. Adding more blades cannot impact the flow rate range. While it will raise the outlet static pressure, it will also raise the inlet flow pressure.

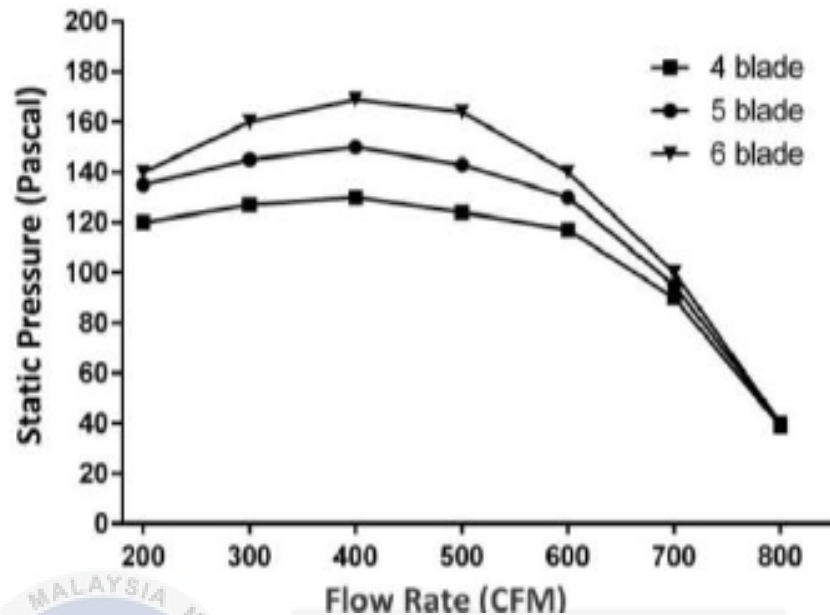


Figure 2. 16 : Characteristics curves of the fans with different number of blades

2.5.4 Fan Outer Diameter

Centrifugal blower is made up of three critical components: a fan, a volute casing, and an inlet duct. The fan revolves, thereby creating a low-pressure zone at the inlet. This causes the fan to pull air from the surrounding atmosphere into the intake. In Jayapragasan and Reddy (2017) research, fan performance depends on fan outer diameter, number of blades, and fan blade angle. The researcher used Taguchi's orthogonal array method and Minitab to examine the effectiveness of several variables: fan outer diameter, fan blade angle, and the number of blades to achieve an optimal set of variables.

In Table 2.3, the CFD results data show that the fan outer diameter is the most influential when it comes to static pressure, while the number of blades has the least impact. While

for experimental, static pressure is also largely dependent on the fan outer diameter, whereas the number of blades effects it the least as shown in Table 2.4.

Table 2. 3 : Factors contribution in optimum combination of the fan for analytical results.

Factors	DOF	Sum of squares	Mean squares	Variance	% Contribution
Fan outer diameter	3	1028.385	342.795	35.445	75.018
Fan blade angle	3	179.101	59.700	6.173	13.065
Number of blades	3	163.351	54.450	5.630	11.916
Total	9	1370.852			-

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Table 2. 4 : Factors contribution in optimum combination of the fan for experimental results.

Factors	DOF	Sum of squares	Mean squares	Variance	% Contribution
Fan outer diameter	3	336.830	278.943	69.767	80.619
Fan blade angle	3	102.031	34.010	8.506	9.828
Number of blades	3	99.271	33.090	8.276	9.562
Total	9	1038.133	-	-	-
Error	6	23.989	3.998	-	-

Figure 2.18 demonstrates that an optimal combination for analytical results has an outside diameter of 190 mm, a blade angle of 60 degrees, and eight blades. In Figure 2.17, the optimal set of variables for experiments is found to be an outer diameter of 190 mm, an angle of 80°, and a number of blades of 8. This is due to the assumptions made in CFD like steady state air flow, implicit solver, constant temperature, constant air density and viscosity unlikely happened in the real life.

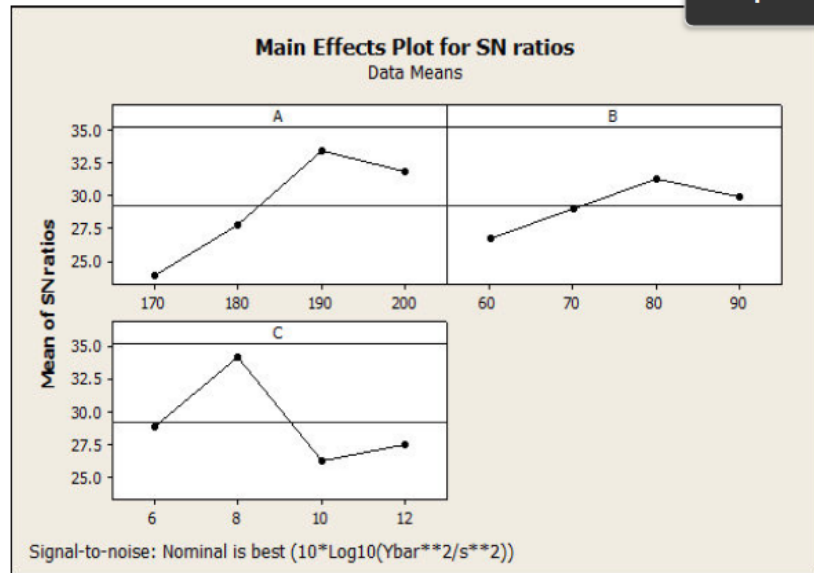


Figure 2. 17 : Experimental - Responses of parameters by mean effects plot for SN ratios

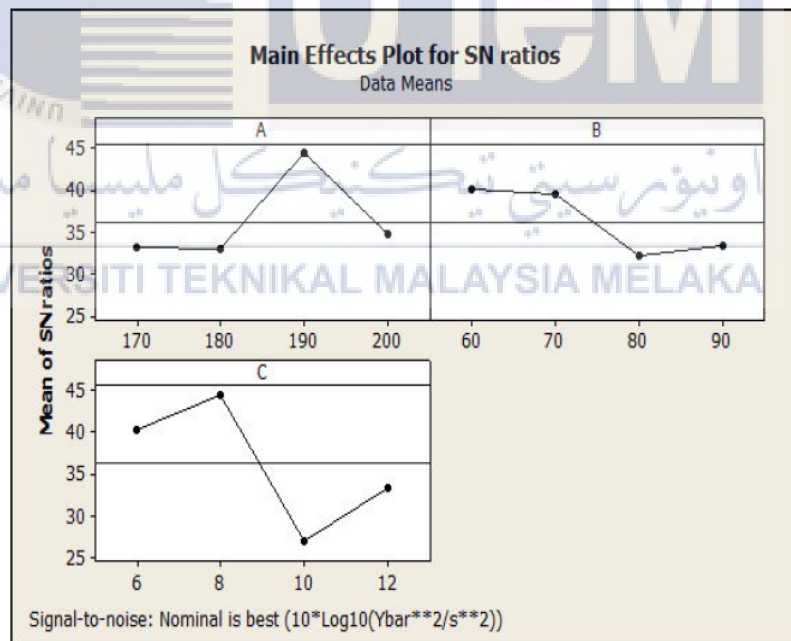


Figure 2. 18 : Analytical - Responses of parameters by mean effects plot for SN ratio

2.6 Equipment and Software

2.6.1 Graphics Processing Unit

Nowadays, gamers looking to upgrade their graphics cards will find the ASUS NVIDIA® GeForce TURBO-GTX1070-8G to be an excellent value. Most games will run well on this card, even with the graphical settings set to the lowest level. It will confidently exceed the required system requirements for today's titles. ASUS NVIDIA® GeForce TURBO-GTX1070-8G cooling system comes with dual-ball bearing fan as it runs smoother by reducing friction. Furthermore, it is also capable to improve 4 times longer card lifespan and cooling efficiency. Asus claimed that this GPU was made Industry Only Auto-Extreme Technology with Super Alloy Power II, which can deliver premium quality and the best reliability among the GPU in the market. More specifications of the GPU are shown in the Table 2.5 below.

Next, overclock a GPU is a valuable option for enhancing desktop computing speed and quality, making it possible to perform better. It may improve equipment performance, but it will also shorten its useful life. It's best to go with a GPU with a good cooling system to better manage airflow and temperature. Furthermore, the owner has complete control over the fan speed. Graphics card clock speed measures the speed of the GPU's processing units (the "cores"). The cores in question are in charge of generating and showing visual information on screen. As a result, boosting the GPU's clock speed results in faster processing. The number of processing cycles per second that a GPU can do is expressed in terms of clock speed.

Table 2. 5 : Specification of Single-Fan Graphic Processing Unit (GPU).

Category	Products specification
Graphics Processing	ASUS NVIDIA® GeForce TURBO-GTX1070-8G
Core Clock	Boost: 1683 MHz / Base: 1506 MHz
CUDA® Cores	1920
Process Technology	16nm
Memory clock	8008MHz
Memory Size	8GB
Memory Type	GDDR5
Memory Bus	256 bit
Card Bus	PCI-E 3.0
Digital max resolution	7680x4320
Multi-view	4
Dimension	H= 26.67cm, L= 11.12cm, W= 3.81cm
PCB Form	ATX
DirectX	12
OpenGL	4.5
Recommended PSU	500W
Power connectors	8 pin*1

2.6.2 Afterburner Software

Chip manufacturers seem to have taken notice of consumer concerns about the potential damage to their goods. Manufacturers have been designing their goods in recent years to reduce the possibility of harming them by mistake. Overclocking software like Afterburner restricts user modifications to numbers they deem safe due to built-in safety features in most graphics' processors these days. Thermal and the processors and drivers themselves do power consumption monitoring to guarantee the product's safety.

Afterburner is a powerful graphics card overclocking tool from MSI that allows complete control over graphics cards' performance. Extra capabilities include modifying fan profiles, benchmarking, and video recording. It gives full control of the GPU of the memory clock, core clock, fan speed, core voltage, temperature limit, power limit, etc. MSI Afterburner can work with any brand of graphics card, for example, Gigabyte, Asus, and others, is free to download and use. It is simple and precise to access the graphics card settings using MSI Afterburner Overclocking (OC) software. The GPU's clock frequency and voltage may be increased while the fan speed can be controlled to achieve the ideal balance between performance and temperature.

Afterburner utilizes OC Scanner, an Nvidia-developed technology, to identify any GPU model available for usage with Afterburner. Once the OC Scanner has determined your video card, it will calculate the best overclocking settings for that particular card. This software will be used to conduct this experiment and research to control the GPU. Figure 2.19 below shows the Afterburner software.

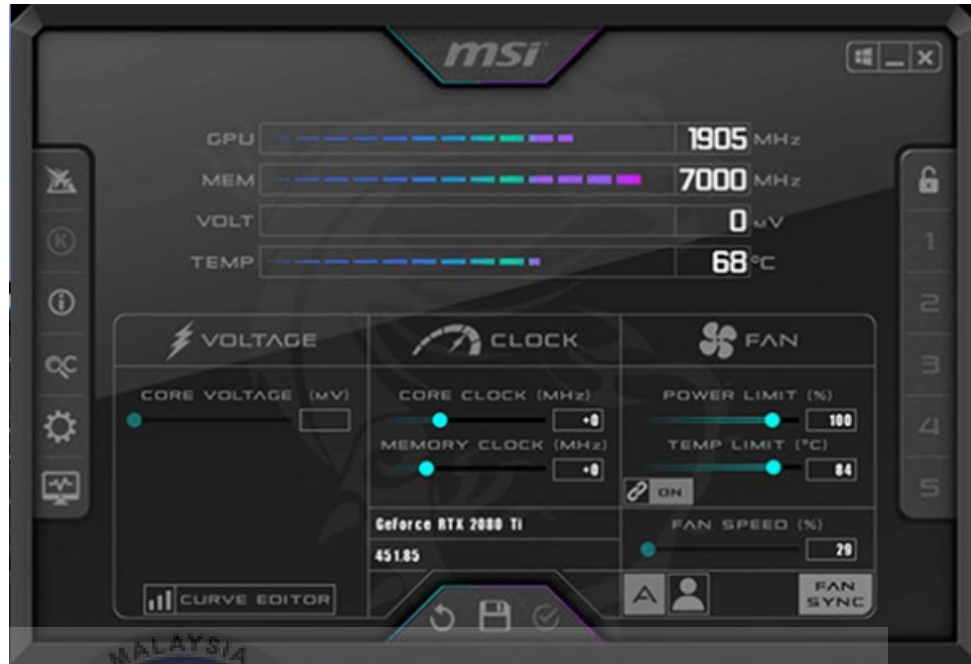


Figure 2. 19 : Afterburner software

2.6.3 Design Expert

Design-Expert is a Stat-Ease Inc. statistical software tool dedicated only to experiment design (DOE). Design-Expert offers screening, characterization, optimization, resilient parameter design, mixture designs, and combination designs as testing services. Design-Expert provides test matrices to screen for as many variables as feasible. ANOVA is used to establish the statistical significance of various variables. Analyzing the data graphically identifies any anomalies and demonstrates the effect of each ingredient on the expected output.

Design-Expert provides powerful tools for designing an ideal experiment on a process, mixture, or combination of factors and components. It makes a simple way to

analyze and appropriately depict the results. Design-Expert also includes several graphs to aid users in identifying and visualizing key consequences. Figure 2.20 below illustrates the many settings available in Design-Expert software, including response surface, mixture, and factorial. The experiment and investigation will be conducted using this programme.

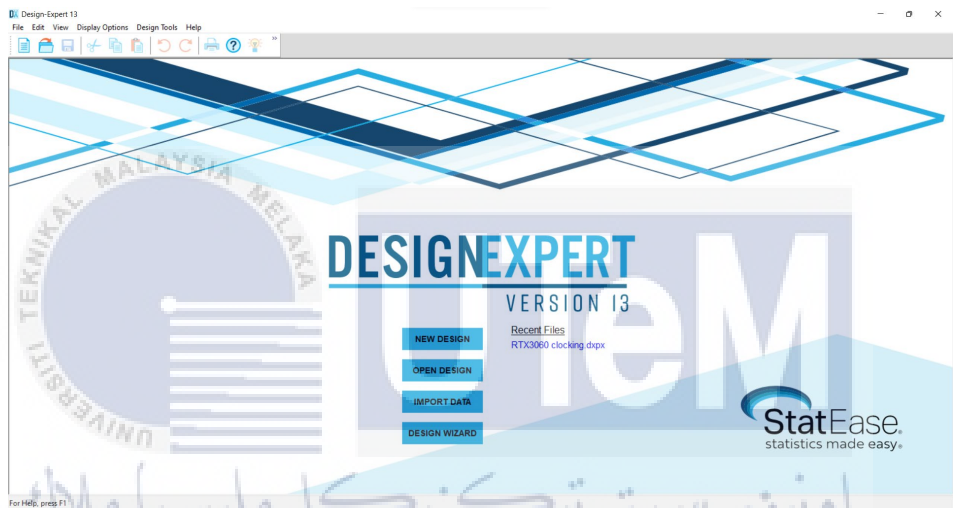


Figure 2. 20 : Design Expert Software
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2.6.4 Response Surface Method

Response Surface Method (RSM) is a regression technique that has been pioneered by Box and Wilson, which is used to investigate the associations between many explanatory factors and one or more response variables. According to Zeng et al., 2009 the purpose of RSM is to estimate a complicated model using data derived at various places in the design space. Furthermore, RSMs that use low-order polynomial functions (second-order is frequently implemented) can efficiently describe low-order issues, and the related calculation of a response surface model is rapid and inexpensive.

RSM also refers to a collection of mathematical and statistical methodologies for modelling the link between input factors and output variables. This methodology will reduce the overall number of tests necessary in trials, hence saving money and time. Additionally, RSM assists by accurately defining and detecting the influence of multiple independent variables interacting on the answer when they are altered concurrently (Campatelli et al., 2014). FNIDES et al., 2011, also stated that the RSM technique offers considerably more precise estimation of the factor's ideal levels; it is possible to determine which of the investigated levels is the best and the exact number that optimizes the design.

According to Field et al., 2017, RSM should be conducted after carefully selecting factors with a significant influence on the results. This may be accomplished through the use of screening studies, such as factorial designs. These first-order designs approximate the linear functions of the output variables. These designs do not attempt to measure curvature. Second-order techniques such as Central Composite, Box-Behnken, and

Doelhart estimate the variables' curvature-interaction and present it as a quadratic equation.

2.6.5 Central Composite Design

Central composite design (CCD), also known as a Box-Wilson central composite design, is frequently used in response surface techniques to create a second-order polynomial for the response variables without utilizing a full factorial design of trials. According to Roy et al., (2017), the mathematical link between the response and changeable process parameters is established using response surface modelling (central composite design). To identify the polynomial coefficients with quadratic terms at least three variables of levels of each component must be included in the experimental design. There are three sorts of points in a CCD system: factorial points, central points, and axial points. The factorial cube's points are n-dimensional cube vertices derived from the full or fractional factorial design, where the factor levels are denoted by the symbols -1 , $+1$. The central point is the spot in the design area in the centre of which all other elements are centred. Axial points are positioned symmetrically to the central point on the coordinate system's axes (Sahoo and Barman, 2021).

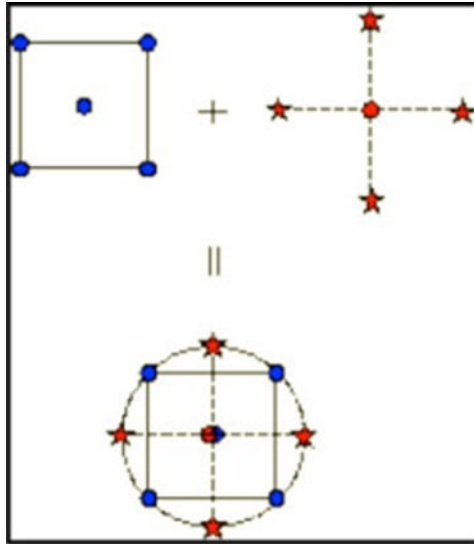


Figure 2. 21 : Generation in a central composite design of points

As seen in Figure 2.21, measuring the curvature of the response surface is made possible by utilizing "star" points located outside the experimental domain and "center" points situated in the experimental domain. The levels of factorial points are +1, -1, but the levels of star points are +a,-a, where a must be higher than or equal to 1. The star points denote the extreme values at the lower and upper ends. The parameter value is decided by the calculation capability and the required precision in estimating the response surface. The order in which the points are placed dictates the quality of the assessment. The value set and the number of trials conducted in the domain's center influence the estimation accuracy (Ait-Amir et al., 2020). There was a discussion of the many phases involved in central composite design (CCD) in Figure 2.22.

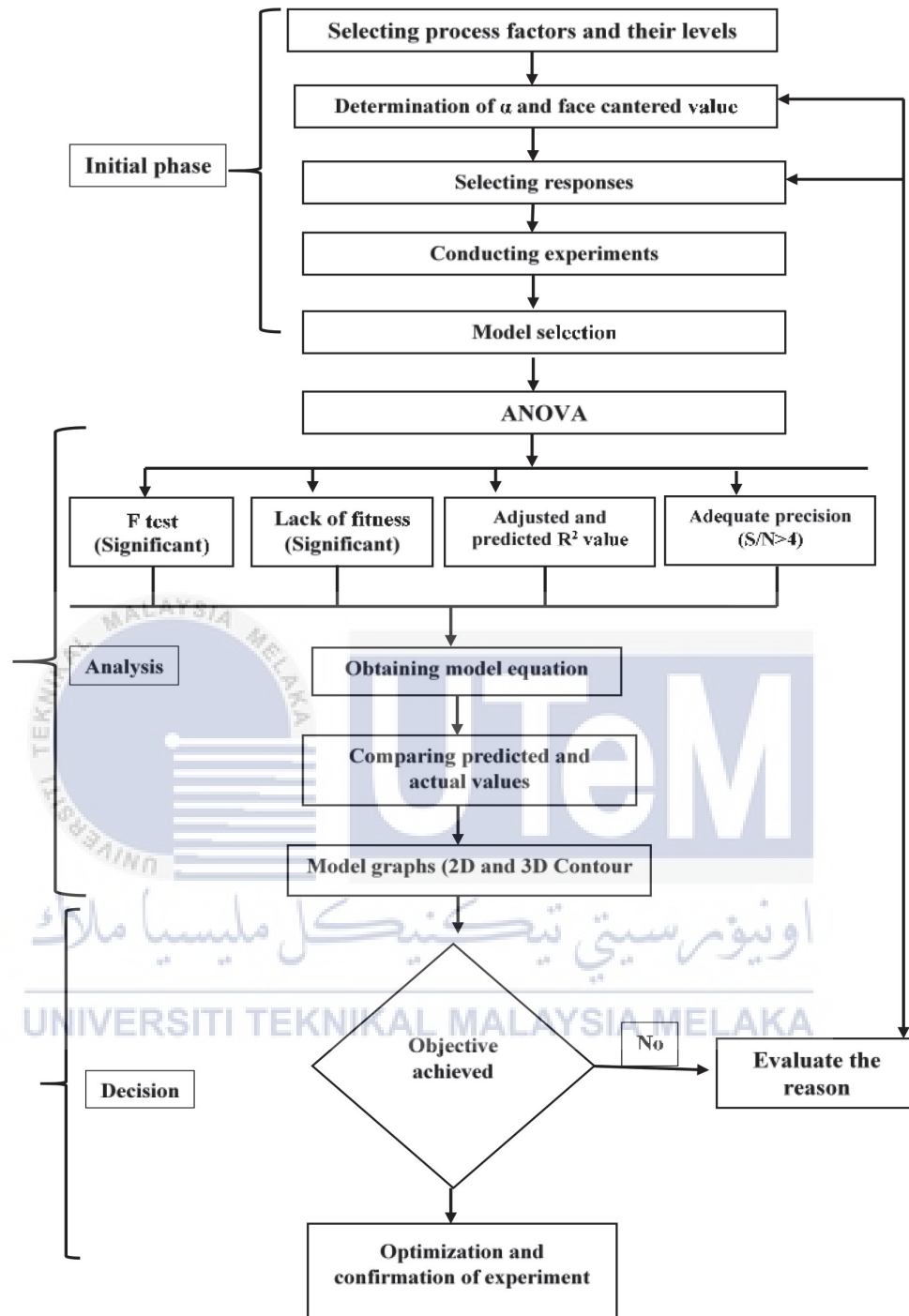


Figure 2. 22 : Central Composite Design (CCD) Flow Diagram

2.6.6 PhoenixMiner Software

PhoenixMiner as shown in Figure 2.23 is mining software that is used to create new cryptocurrencies and to add new blocks to an existing blockchain. As a reward for their contribution to the blockchain, newly mined cryptocurrency belongs to the mining party. The graphics processing unit (GPU) in a computer is exploited to help with cryptocurrency mining. In today's world, most mining is done through a mining pool, which divides the reward between its members via a network of computers.

Dagger Hashimoto (Ethash) algorithm is used by PhoenixMiner, a cryptocurrency miner, to generate cryptocurrency. There are a number of well-known cryptocurrencies in this list, including Ethereum, Ethereum Classic, and MOAC. The application works with both Nvidia and AMD graphics cards. According to the creators, it is the quickest and most cost-effective Ethereum/Ethash miner currently on the market. However, there is a 0.65% service charge, which means that the user will only "work" for developers for 35 seconds out of every 90 minutes of GPU mining. Claymore's miner development fee is one percent, as a comparison. The GPU response (hash rate) will be determined by this software.

```
C:\Windows\system32\cmd.exe
Eth: Incorrect shares 0 (0.00%), est. stales percentage 0.00%
Eth: Average speed (5 min): 31.490 MH/s
Eth: Effective speed: 0.00 MH/s; at pool: 0.00 MH/s

Eth speed: 31.432 MH/s, shares: 0/0/0, time: 0:03
Eth: New job #33dbc875 from eth.2miners.com:2020; diff: 8600MH
Eth speed: 31.428 MH/s, shares: 0/0/0, time: 0:03
Eth speed: 31.423 MH/s, shares: 0/0/0, time: 0:03
Eth: New job #393815e8 from eth.2miners.com:2020; diff: 8600MH
Eth: GPU1: ETH share found!
Eth: Share actual difficulty: 149.9 GH (!)
Eth: Share accepted in 38 ms
GPU1: 62C 53%
Eth speed: 31.400 MH/s, shares: 1/0/0, time: 0:03
Eth speed: 31.387 MH/s, shares: 1/0/0, time: 0:03
Eth: New job #7dd09d95 from eth.2miners.com:2020; diff: 8600MH
Eth speed: 31.303 MH/s, shares: 1/0/0, time: 0:03
Eth speed: 31.428 MH/s, shares: 1/0/0, time: 0:03
Eth: New job #8f331afe from eth.2miners.com:2020; diff: 8600MH
Eth: New job #74fde7d0 from eth.2miners.com:2020; diff: 8600MH
Eth: New job #e2233fa9 from eth.2miners.com:2020; diff: 8600MH
Eth speed: 31.455 MH/s, shares: 1/0/0, time: 0:03
Eth speed: 31.523 MH/s, shares: 1/0/0, time: 0:03
GPU1: 63C 54%

*** 0:03 *** 7/25 23:23 *****
Eth: Mining ETH on eth.2miners.com:2020 for 0:03
Eth: Accepted shares 1 (0 stales), rejected shares 0 (0 stales)
Eth: Incorrect shares 0 (0.00%), est. stales percentage 0.00%
Eth: Maximum difficulty of found share: 149.9 GH (!)
Eth: Average speed (5 min): 31.473 MH/s
Eth: Effective speed: 38.22 MH/s; at pool: 38.22 MH/s

Eth: New job #6799cccc from eth.2miners.com:2020; diff: 8600MH
Eth speed: 31.444 MH/s, shares: 1/0/0, time: 0:03
Eth: New job #1af5a861 from eth.2miners.com:2020; diff: 8600MH
Eth: New job #1c705393 from eth.2miners.com:2020; diff: 8600MH
Eth speed: 31.408 MH/s, shares: 1/0/0, time: 0:03
Eth: New job #2604504d from eth.2miners.com:2020; diff: 8600MH
```

Figure 2. 23 : PheonixMiner Software



2.7 Fundamental of Heat Transfer

Heat transfer is a dynamic process involving transferring heat from a cooler body to a warmer body. There are three fundamental mechanisms in which heat is transmitted are: convection, conduction, and radiation. When temperature differentials exist between two things, heat is transferred from a hot object to a cold one. The basic equation in heat transfer is given:

$$Q = m \times c_p \times \Delta T$$

Where,

Q = Heat transferred, W

m = Mass, kg

c_p = Specific heat, J/kg · K

ΔT = Temperature difference, K

Conduction is the transmission of thermal energy from the high-energy particles to the low-energy particles of a stationary medium (solid, liquid, or gas) due to interactions between the particles is known as thermal conduction. The flow of internal thermal energy that takes place through collisions of tiny particles and the movement of electrons within a body is known as heat conduction (Bi, 2018). During random motion, molecules collide and diffuse, allowing conduction to occur.

$$Q_{conduction} = k A \frac{\Delta T}{\Delta x}$$

Where,

$Q_{conduction}$ = Rate of heat conduction, W

k = Thermal conductivity, $W/m \cdot K$

A = Cross sectional area, m^2

ΔT = Temperature difference, K

Δx = Thickness, m

In heat convection, energy transmission happens in heat convection, due to fluid motion (Sheng, 2013). Heat convection is a kind of heat transmission in which heat is moved through the fluid by the mass motion of the fluid itself. When the fluid of an object is heated, heat is transferred to the surrounding environment, which moves the energy away from the source of heat. Also, convective heat transfer happens when the surface temperature differs from that of the fluid around it (Bi, 2018).

Convection heating is demonstrated by cooking in a jacketed pan. If the pan does not have a stirrer, natural convection causes density changes, which causes heat transfer. However, if the pan has a stirrer, the natural convection is pushed, and the temperature rises. The fundamental relation for convection heat transfer is:

$$Q_{convection} = h \times A_s \times \Delta T$$

Where,

$Q_{convection}$ = Rate of heat transfer, W

h = Convective are for heat transfer coefficient, $W/m^2 \cdot K$

A_s = The surface area through which convection heat transfer takes place, m^2

ΔT = Temperature difference, K

The transfer of heat energy by electromagnetic radiation is called radiation heat transfer. The energy that is emitted by matter in the form of photons or electromagnetic waves is referred to as radiation heat transfer (Shahidian et al., 2020). Heat-generating and heat-absorbing materials radiate energy independently of the medium in which they exist and depend on the material's temperatures, geometric arrangements, and surface patterns.

$$Q_{emit,max} = \varepsilon\sigma A_s T_s^4$$

Where,

$Q_{emit,max}$ = Rate of heat transfer radiation, W

ε = Emissivity

σ = Stefan-Boltzmann constant ($5.670 \times 10^{-8} \text{W/m}^2 \cdot \text{K}^4$)

A_s = Surface area, m^2

T_s = Surface temperature, K



CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter will focus on the strategy employed to complete the research project's objectives. For example, this project is about finding the optimal fan speed, core temperature, and the highest possible efficiency (hashing power) at certain core and memory clock via optimization tools, and the highest possible efficiency (hashing power) at certain core and memory clock via optimization tools. Next, a flowchart will demonstrate the process to achieve the objectives step by step. This chapter will go over every study component, including the methodology, software, processes, and procedures used. This chapter aims to guarantee that all plans and activities are correctly handled to avoid errors resulting from a poor planning strategy. The appropriate and adequate methodology will be used to accomplish the study goal. A method will be chosen with constraints to implement it, having a second and third plan would be considered too. This comprehensive analysis has been built on the foundation of findings gained from sources such as the literature review and the primary research.

3.2 Flow Chart

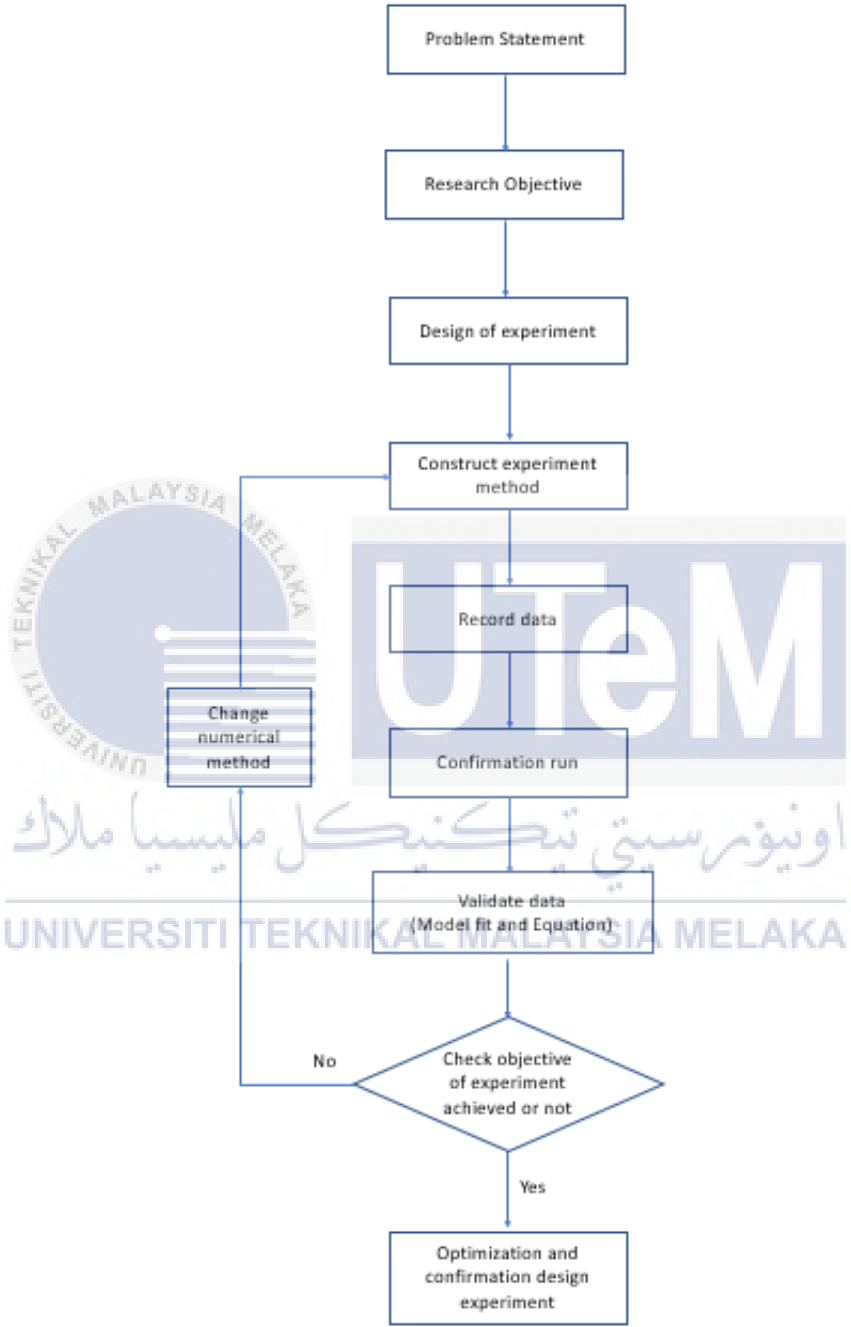


Figure 3. 1: Flowchart of methodology

3.2 Schematic Diagram

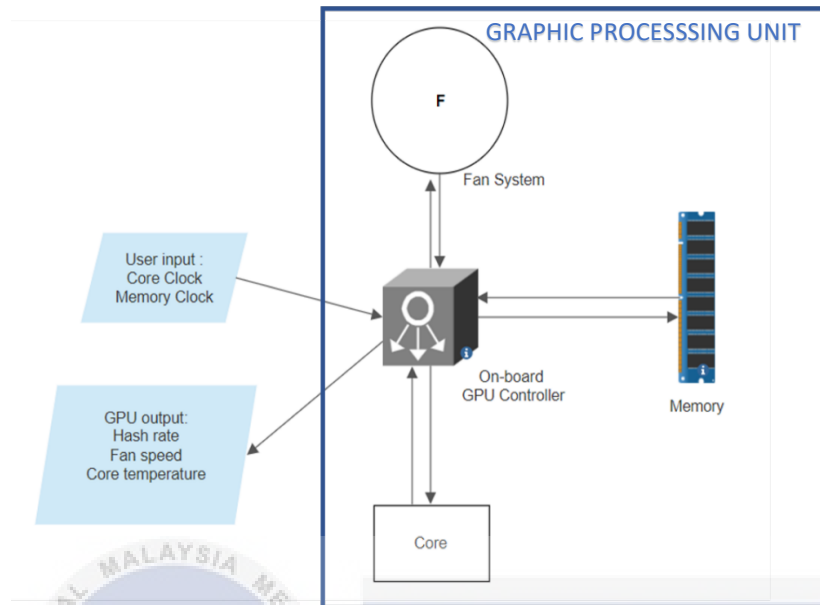


Figure 3. 2 : Schematic diagram

3.3 Design of Experiment

The primary advantage of RSM over the traditional time-consuming one-variable-at-a-time approach is the reduced number of experimental runs required for a more systematic investigation of the processing variables, including their simultaneous interaction and modelling of the selected response parameters (Ghelich et al., 2019). Response surface methodology (RSM) employs statistics, theory, and mathematics to maximize the number of independent variables. The numeric factor in the Central Composite Design (CCD) can be filled with two to 50 variables, whereas the Box-Behnken Design (BBD) can be filled with three to twenty-one variables. The experiment's manipulated variables are represented by numerical factors. Because there are only two

numeric factors in this study, BBD was chosen to optimize the data from the experiments. The experiments were designed using RSM based on the Central Composite Design in Design-Expert version 13.0 software to examine the relationship between clocking (core clock, memory clock) and GPU response (fan speed, core temperature, hash rate).

The experimental results were analyzed statistically using Design Expert (version 13). Numerous statistical parameters (such as goodness-of-fit, predicted and adjusted multiple correlation coefficients, and coefficient of variation) of various polynomial models were evaluated in order to determine the most accurate one. A significant difference was discovered after analyzing variance and computing the F-value at probabilities of 0.5, 0.1, and 0.01. To gain a better understanding of the relationship between emulsifying conditions and response variables, response plots were created using Design Expert Software. When variables and responses are optimized, the CCD technique assists in determining the number of tests to run. Each variable's minimum, intermediate, and maximum values are denoted by the symbols -a, 1, 0, +1, and +a, respectively, as shown in Table 3.1.

Table 3. 1 : Independent variables and their corresponding levels for GPU clocking

Independent variable	Symbol	Coded levels				
		-a	-1	0	+1	+a
Core Clock (MHz)	A	-200	-141.421	0	141.421	200
Memory Clock (MHz)	B	300	387.868	600	812.132	900

According to the Minerstat website, the optimal core clock and core clock value for GPU GTX 1070 are 100MHz and 300MHz. In this project the core clock was set ranges between -200MHz and 200MHz, while the memory clock was set ranges between 300MHz and 900MHz. The factorial +1 and -1 design points are depicted in Figure 3.3 below. The values +1 and -1 are used to define the study region's limits, which are believed to include the optimal scenario. On the other hand, axial points frequently fall outside of this range. The axial +alpha and -alpha design points are depicted in the illustration by four-star points. Even for the most severe axial runs, there are alphas to ensure that everything continues to work properly. The emphasis must be on something that is operable.

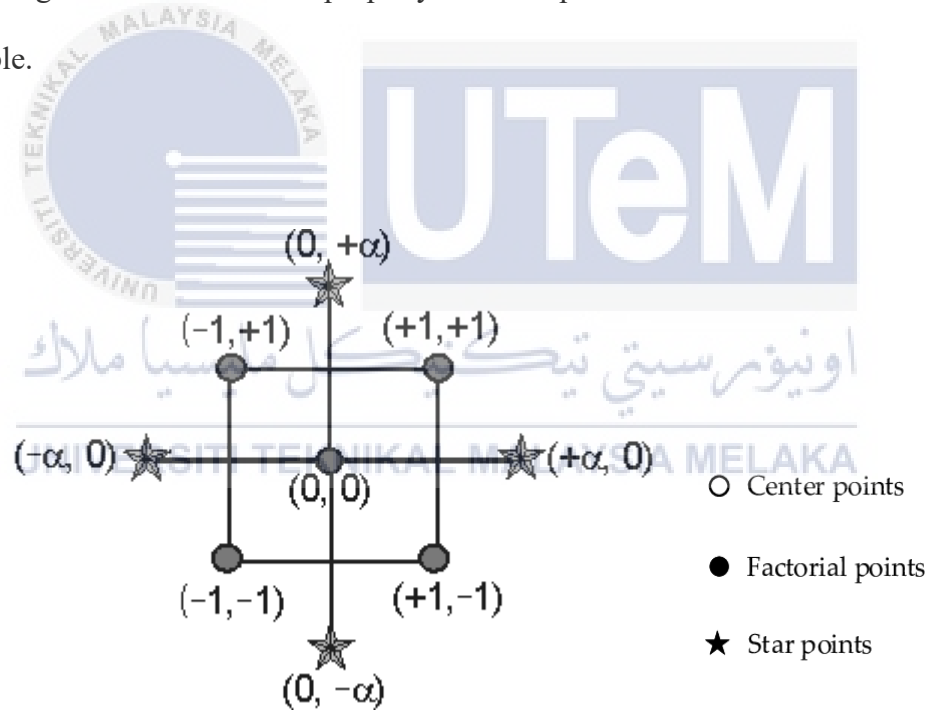


Figure 3. 3 : Classic central for central composite design for 2 factors

3.5 Experimental Setup

The experiment was set up as shown in Figure 3.2. Afterburner software will be used as user input control of the core clock and memory clock in accordance with the values specified by Design Expert software. The GPU response, such as fan speed (rpm), core temperature ($^{\circ}\text{C}$) and hash rate (MH/s) can be observed using the Afterburner and PheonixMiner software. Then, the experimental were run continuously and repeated several times to get a constant response (GPU output). Next, the responses will be transferred into the Design Expert software for analysis and optimization process to get the optimal fan speed and the highest efficiency (hashing power) at certain core clock and memory clock. The optimal core clock and core clock value for GPU GTX 1070 are 100MHz and 300MHz, respectively, according to the Minerstat website. The core clocks were set to a range of (-200 to 200 MHz) and memory clock (300 to 900MHz) in this study. The options for configuring the numeric factors are shown in Figure 3.4. The Design Experts provide the user with the option of entering factor ranges as +1,-1 levels or alphas. This research defines the factor ranges in terms of alphas because the core and memory clocks' maximum and minimum values have already been established.

Central Composite Design

Each numeric factor is set to 5 levels: plus and minus alpha (axial points), plus and minus 1 (factorial points) and the center point. If categorical factors are added, the central composite design will be duplicated for every combination of the categorical factor levels.

Numeric factors: (2 to 50) Horizontal Enter factor ranges in terms of ± 1 levels

Categorical factors: (0 to 10) Vertical Enter factor ranges in terms of alphas

	Name	Units	Low	High	-alpha	+alpha
A [Numeric]	Core Clock	MHz	-141.421	141.421	-200	200
B [Numeric]	Memory Cloc	MHz	387.868	812.132	300	900

Type: Blocks:

Points

Non-center points: 8

Center points: 5

alpha = 1.41421 13 Runs

Figure 3. 4 : Central Composite Design layout in Design Expert

Based on Figure 3.5 below, 13 runs were performed randomly according to CCD, including four axial points, four fractional factorial points, and five central points.

Std	Run	Space T...	Factor 1 A:Core Clock MHz	Factor 2 B:Memory Cl... MHz	Response 1 Fan Speed rpm	Response 2 Core Temper... °C	Response 3 Hashrate MH/s
6	1	Axial	200	600			
9	2	Center	0	600			
5	3	Axial	-200	600			
3	4	Factorial	-141.421	812.132			
13	5	Center	0	600			
2	6	Factorial	141.421	387.868			
4	7	Factorial	141.421	812.132			
11	8	Center	0	600			
12	9	Center	0	600			
10	10	Center	0	600			
8	11	Axial	0	900			
7	12	Axial	0	300			
1	13	Factorial	-141.421	387.868			

Figure 3. 5 : Design CCD (Actual) in Design Expert

3.6 Statistical Analysis and Optimization of Clocking and GPU Response

Design-Expert software was used to do a regression analysis of the GPU responses (fan speed, core temperature, and hash rate). The software will do an analysis of variance (ANOVA) on the constructed models. ANOVA is frequently used to determine the degree of dissimilarity between sets of data. It establishes the model's significance by an analysis of the model's statistical fit. The optimization constraint option can then be used to configure the numerical parameter to reach the ideal fan speed, temperature, and Hash rate. CCD is an effective method for fitting a quadratic surface model sequentially. Numerous design alternatives that do not fit are eliminated from CCD's evaluation process.

CHAPTER 4

RESULT AND DISCUSSION

4.1 Result

The Box-Wilson second-order central composite design (CCD), and particularly its original version, the central composite circumscribed (CCC), is a well-established, successful, and widely used RSM design. Through a reduced number of design points and an accurate curvature estimation, the CCD is perfect for assigning operation specific variables to a range of evaluations in order to gather a fair quantity of information for assessing lack-of-fit (Ghelich et al., 2019). The results were obtained based from two inputs. The afterburner software is used to set the memory clock and core clock, as shown in Table 3.1 DOE, and based on the real design of experiments, the response findings, fan speed, core temperature, and hash rate, were gathered and shown in Table 4.1 This will help in finding the ideal fan speed and the maximum efficiency at specific core clock and memory clock by using numerical method CCD.

Table 4. 1 : Experiment design and response value obtained by the GPU

Standard order	Space type	Independent input variable (MHz)		Response (experiment)		
		A: Core clock	B: Memory clock	Fan speed (rpm)	Core temperature (C)	Hash rate (MH/s)
1	Factorial	-141.421	387.868	2814	67.5	19.5
2	Factorial	141.421	387.868	2688	64.6	20.7
3	Factorial	-141.421	812.132	2814	67.8	19.6
4	Factorial	141.421	812.132	2688	64.7	20.6
5	Axial	-200	600	2859	68.1	19.6
6	Axial	200	600	2683	64.5	21.2
7	Axial	0	300	2731	65.4	19.8
8	Axial	0	900	2730	65.5	19.8
9	Center	0	600	2730	65.2	19.8
10	Center	0	600	2731	65.3	19.8
11	Center	0	600	2730	65.3	19.8
12	Center	0	600	2730	65.2	19.8
13	Center	0	600	2730	65.5	19.8

4.2 CCD and ANOVA (Fit)

4.2.1 Fan Speed Response

Table 4. 2 : ANOVA for fan speed response

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	34318.55	5	6863.71	20229.71	< 0.0001	significant
A-Core Clock	31362.80	1	31362.80	92436.93	< 0.0001	
B-Memory Clock	0.2500	1	0.2500	0.7368	0.4191	
AB	0.0000	1	0.0000	0.0000	1.0000	
A ²	2912.79	1	2912.79	8585.00	< 0.0001	
B ²	0.3141	1	0.3141	0.9259	0.3680	
Residual	2.38	7	0.3393	-	-	
Lack of Fit	1.58	3	0.5250	2.63	0.1871	not significant
Pure Error	0.8000	4	0.2000	-	-	
Cor Total	34320.92	12	-	-	-	

Table 4. 3 : Fit statistics for fan speed response

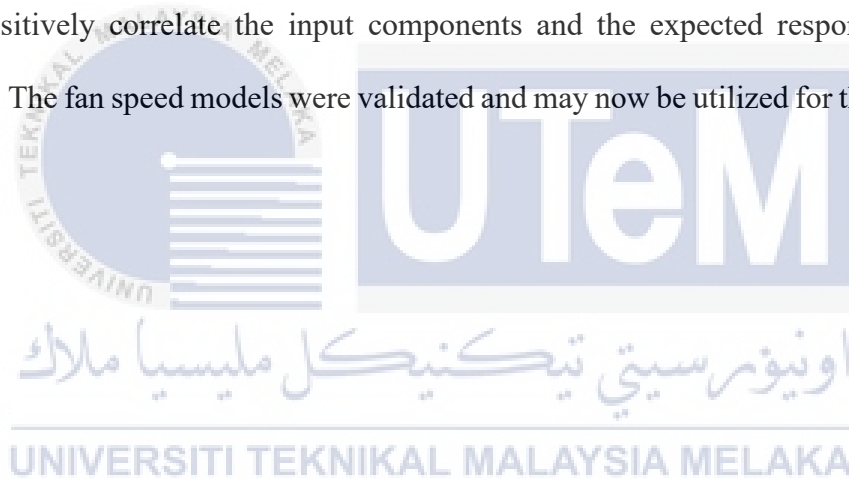
Std. Dev.	0.5825	R ²	0.9999
Mean	2742.92	Adjusted R²	0.9999
C.V. %	0.0212	Predicted R²	0.9996
		Adeq Precision	447.5266

In this project the validations for the result are based on Lack of Fit (LOF), R², Adjusted R², Coefficient of Variation (C.V.) and Predicted R². The LOF is insignificant for fan speed responses, indicating that the model is adequate. The model F-value of 20229.71 indicates that the models are significant when the p-value is less than 0.05, as shown in the Table 4.2. A small p-value indicated a close link between computed responses and independent inputs except for the lack of fit. This is important for assessing the significance of input factors toward responses, since higher p-values (greater than 0.10) indicate the lesser significance and vice versa. The models' lack of fit was likewise not statistically significant, implying that both built models were mathematically well fit.

The correlation between R² and adjusted R² is a second element in determining the validity of the generated models. In, fan speed R² values greater than 0.99 indicate that more 99% of the changes in the responses (results) were driven by the independent input factors (memory clock and core clock). In other words, the built models are highly dependent on the input variables, and all of the variables in this experiment are significant. ANOVA findings showed that a quadratic polynomial model could describe the experimental data adequately with R² coefficient of determination values of 0.9999 for Fan speed. Although there is a slight difference between the adjusted and predicted R² values (less than 0.2), this model fits the data well and may be relied upon for

interpolation. Thirdly, cross-validation of the models' adequate precision values of 447.5266 and coefficient of variation (CV) % of 0.0212 proved the models' flexibility and dependability.

Finally, visual diagnostics were performed by comparing the experimentally determined values against the predicted values derived by the programmed using the modelled equations. As seen in Figures 4.1 and 4.2, both figures show almost all values are near the centerline, with no apparent outliers. The plot in Figure 4.1 indicates that the experimental and predicted values are in the best fit line, indicating that the CCD model can positively correlate the input components and the expected response values (fan speed). The fan speed models were validated and may now be utilized for the experimental design.



Fan Speed

Color points by value of Hashrate:

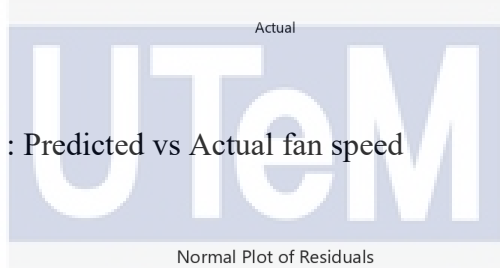
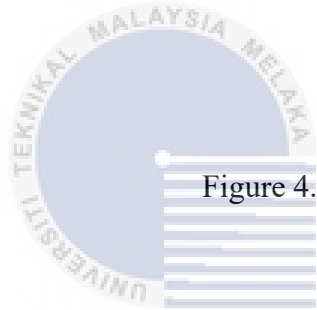
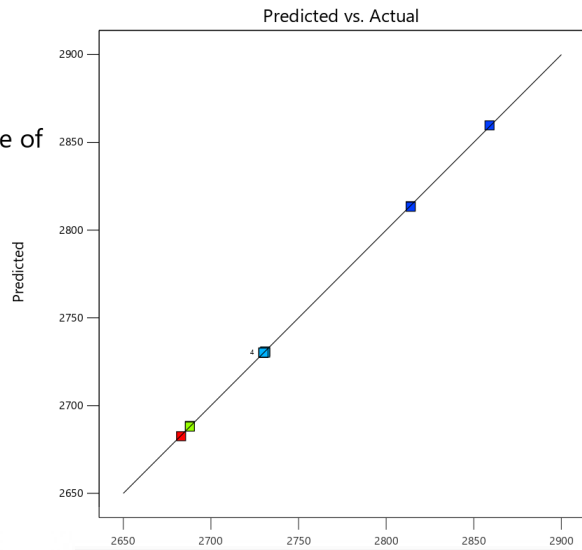
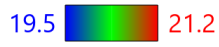


Figure 4. 1 : Predicted vs Actual fan speed

Fan Speed

Color points by value of Hashrate:

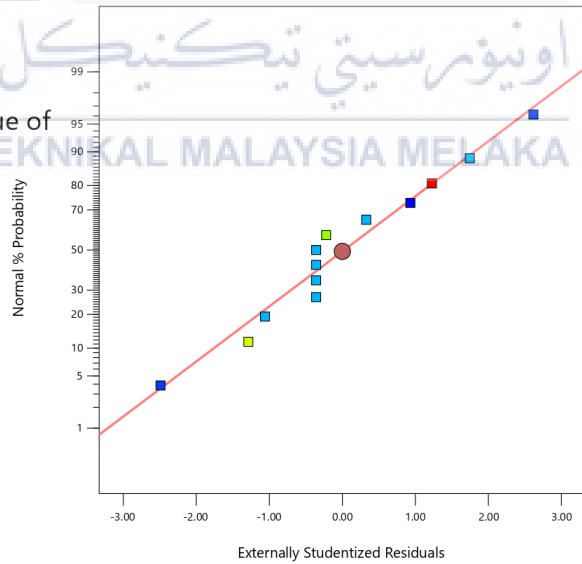
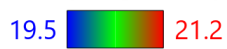


Figure 4. 2 : Fan speed normal plot

4.2.1 Core Temperature Response

Table 4. 4 : ANOVA for core temperature response

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	17.71	5	3.54	76.79	< 0.0001	significant
A-Core Clock	15.38	1	15.38	333.39	< 0.0001	
B-Memory Clock	0.0366	1	0.0366	0.7945	0.4024	
AB	0.0100	1	0.0100	0.2168	0.6556	
A ²	2.25	1	2.25	48.79	0.0002	
B ²	0.1438	1	0.1438	3.12	0.1208	
Residual	0.3229	7	0.0461			
Lack of Fit	0.2629	3	0.0876	5.84	0.0606	not significant
Pure Error	0.0600	4	0.0150	-	-	
Cor Total	18.03	12	-	-	-	

Table 4. 5 : Fit statistics for core temperature response

Std. Dev.	0.2148	R ²	0.9821
Mean	65.74	Adjusted R ²	0.9693
C.V. %	0.3267	Predicted R ²	0.8911
		Adeq Precision	26.8766

The validations for the core temperature in this project are based on the Lack of Fit (LOF), R^2 , Adjusted R^2 , Predicted R^2 , and Coefficient of Variation (C.V.). The LOF is not significant for core temperature responses, which indicates that the model is acceptable. An extremely small p-value indicated that the computed responses and independent inputs were tightly linked. This suggests that the models are significant when the p-value is below 0.05, as seen in the table above. This is critical for measuring the validity of input factors toward responses, as larger p-values (greater than 0.1) reflect lower significance and vice versa. The models' lack of fit was also not statistically significant, meaning that both created models were mathematically fit.

A second factor in evaluating the validity of the created models is the correlation between R^2 and adjusted R^2 . R^2 for core temperatures greater than 0.9821 imply that more than 98% of the responses (results) variations were influenced by independent input sources (memory clock and core clock). For better or worse, the output models are very dependent on the input variables, and all of the variables in this experiment matter. With R^2 coefficients of 0.9821 for fan speed, a quadratic polynomial model accounted for a significant portion of the experimental data. Even though the modified and anticipated R^2 values differ by less than 0.2, this model fits the data well and can be used for interpolation. Thirdly, cross-validation of the models' appropriate precision values of 26.8766 and coefficient of variation (CV) percent of 0.3267 demonstrated adaptability and dependability.

Finally, visual diagnostics were done by comparing the experimentally determined values to the expected values derived by the programme using the modelled equations. Figures 4.3 and 4.4 show no apparent outliers in core temperature projections

or actual measurements. Experimental values are more closely aligned with those expected by the CCD model. The experimental design can now use the validated core temperature models.



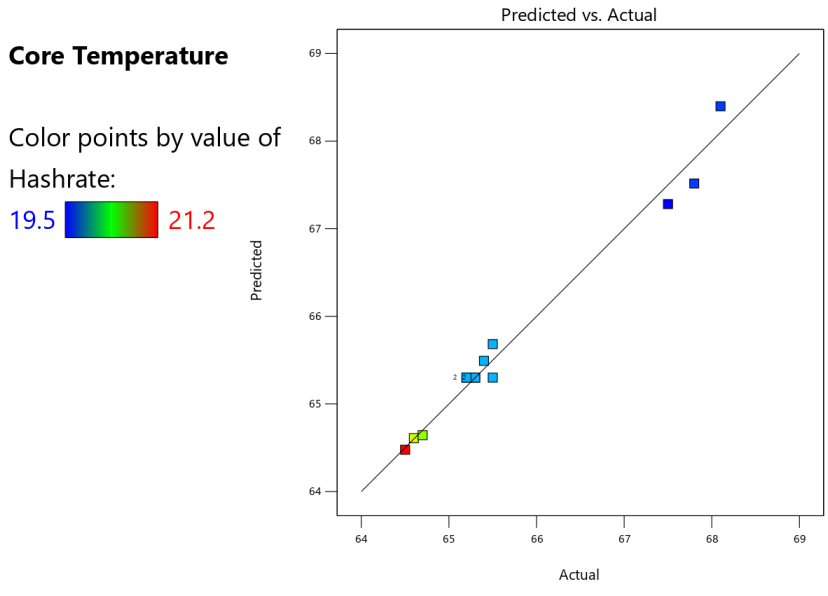


Figure 4.3 : Predicted vs Actual Core Temperature

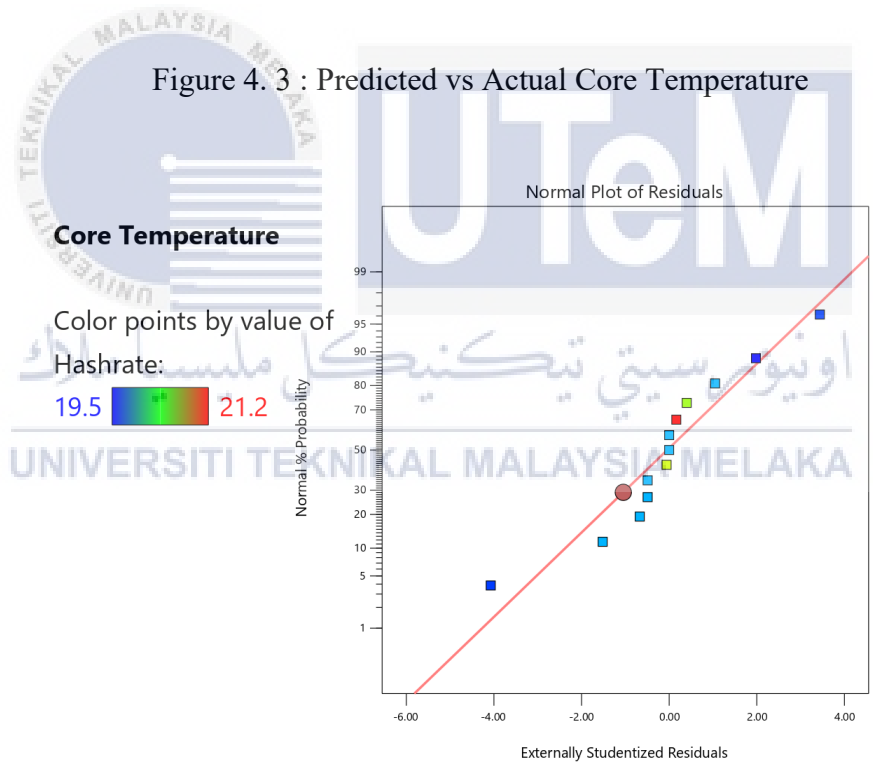


Figure 4.4 : Core temperature normal plot

4.2.3 Hash Rate Response

Table 4. 6 : ANOVA for hash rate response

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	3.14	5	0.6273	8923.62	< 0.0001	significant
A-Core Clock	2.49	1	2.49	35415.14	< 0.0001	
B-Memory Clock	0.0000	1	0.0000	0.0000	1.0000	
AB	0.0100	1	0.0100	142.26	< 0.0001	
A ²	0.6261	1	0.6261	8906.56	< 0.0001	
B ²	0.0000	1	0.0000	0.0000	1.0000	
Residual	0.0005	7	0.0001	-	-	
Lack of Fit	0.0005	3	0.0002	-	-	
Pure Error	0.0000	4	0.0000	-	-	
Cor Total	3.14	12	-	-	-	

Table 4. 7 : Fits statistics for hash rate response

Std. Dev.	0.0084	R ²	0.9998
Mean	19.98	Adjusted R ²	0.9997
C.V. %	0.0420	Predicted R ²	0.9989
		Adeq Precision	297.8873

To determine the model's relevance and fitness, an ANOVA based on CCD was developed. It also determines the effects of the various variables on the hash rate value and the relationships between them. Only when the p-values were less than 0.0001 were the results of an ANOVA considered significant. Results for the quadratic model proposed in Table 4.6 above are listed. For a model with an F-value of 8932.62, the model is significant. A "Model F- Value" this large has only a 0.01 percent chance of occurring due to the noise. The ANOVA table shows that A, B, and all interactions are statistically significant factors.

The correlation between R^2 and adjusted R^2 is the next step in verifying the developed models. Value R^2 hash rate is 0.9998, indicating that the independent input variables were responsible for all of the changes in the responses (results) (memory clock and core clock). In other words, the built models are highly dependent on the input variables, and all of the variables employed in this experiment are significant in this context. Accordingly, a model with a Predicted R^2 of 0.9989 and an Adjusted R^2 of 0.9997 is in reasonable agreement because the difference is less than 0.0008 points. A diagnostic plot comparing actual values to anticipated values might validate results because the software does not display the p-value for Lack of Fit. A graph as in Figure 4.5 showing no outliers demonstrates that hash rate models can be employed for further experimentation, based on the rationale given above.

Hashrate

Color points by value of Hashrate:

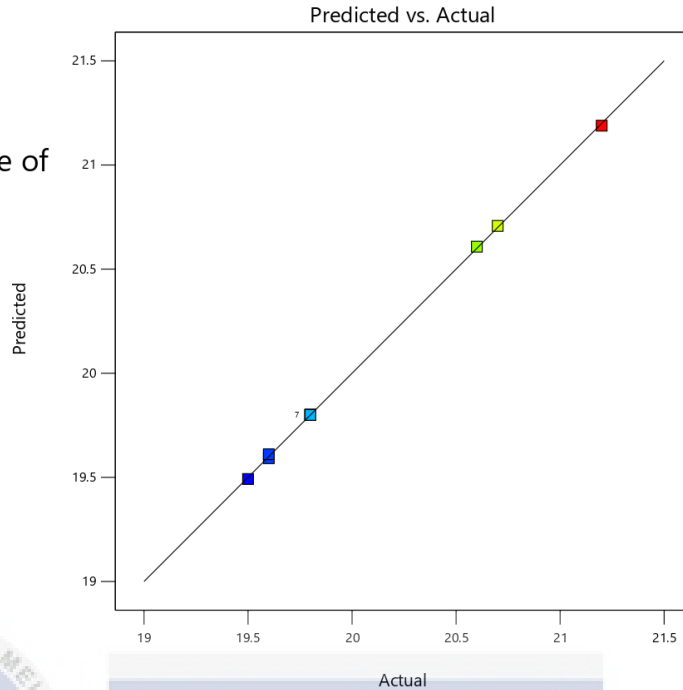


Figure 4. 5 : Predicted vs Actual Hashrate

اونيورسيتي تیکنیکل ملیسيا ملاک

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

4.3 Effect of Independent Variables on Response Variables (Contour Plot and Equation)

4.3.1 Fan Speed

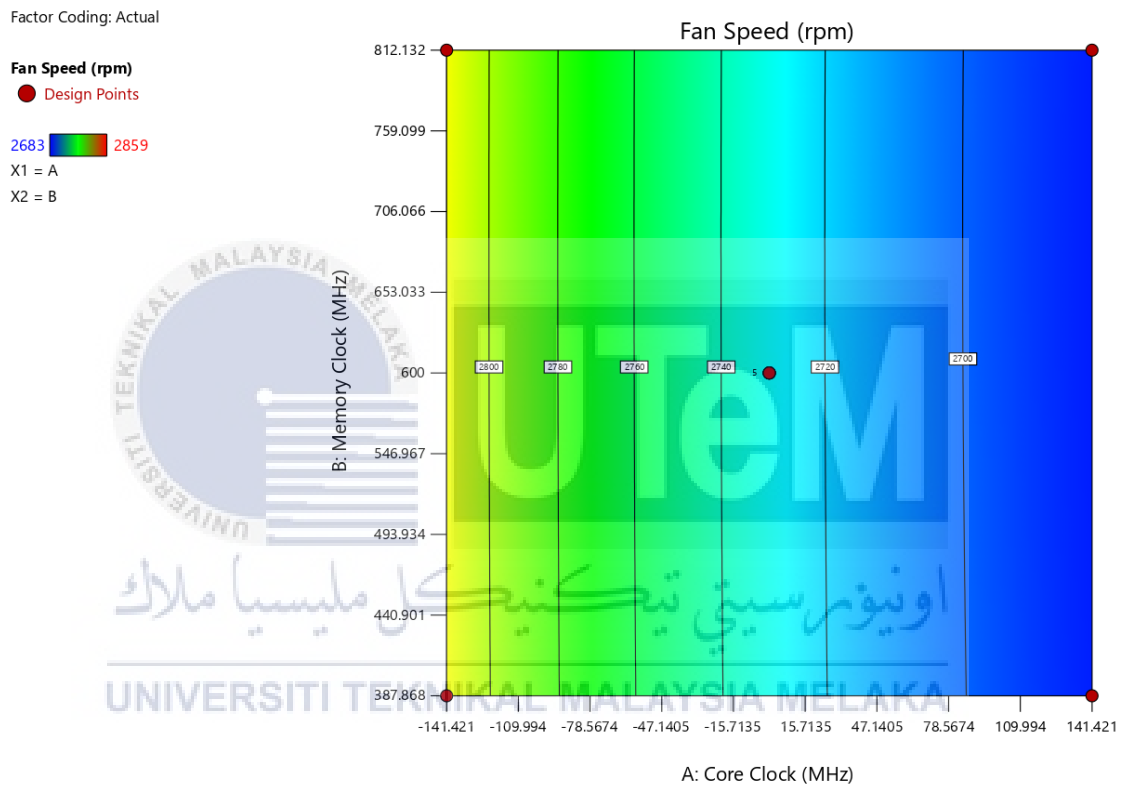


Figure 4. 6 : Contour plot for the combined effect of core clock (A), memory clock (B), and fan speed

Figure above depicts the effect of the memory clock and core clock on fan speed. It can be seen that the fan speed (in the white box) decreases from 28 rpm to 27 rpm as the core clock increases up to a maximum of 141.421MHz. Yellow denotes places with high fan speeds, whereas blue denotes areas with low fan speeds. The fan speed is relatively high for the core clock values (-141.421 to -130.223MHz), as indicated by the

yellow areas. Furthermore, the core clock values (32.869 to 141.421 MHz) in the graph above are highlighted in blue, meaning that those clocks employ a low fan speed compared to other independent variable values.

$$\text{Fan speed} = 2730.2 + -62.6127 * A + -0.176777 * B + 5.39661e-13 * AB + 20.4625 * A^2 + 0.2125 * B^2 \text{ (Eq. 1)}$$

Except for Lack of Fit, a small P-value (less than 0.05) indicated a close relationship between calculated responses and independent inputs. As shown in Table 4.2, the parameters influencing fan speed the most in this situation are A (Core clock), B (Memory Clock), A² and B². Based on Eq. (1), the negative sign in front of the components in the equation, can prove that fan speed can be decreased by A component with a multiplier of -62.6127, B component with a multiplier of -0.176777. This is also supported by the components' lowest p-values (after A, B, A² and B²), as shown in Table 4.2. Based on the equation Eq. (1), it is possible to assume that the A component, which is the memory clock, is the most significant contributor to the fan speed because it has the highest multiplier value. The software confirms this component because the p-value for the memory clock is less than 0.0001.

4.3.2 Core Temperature

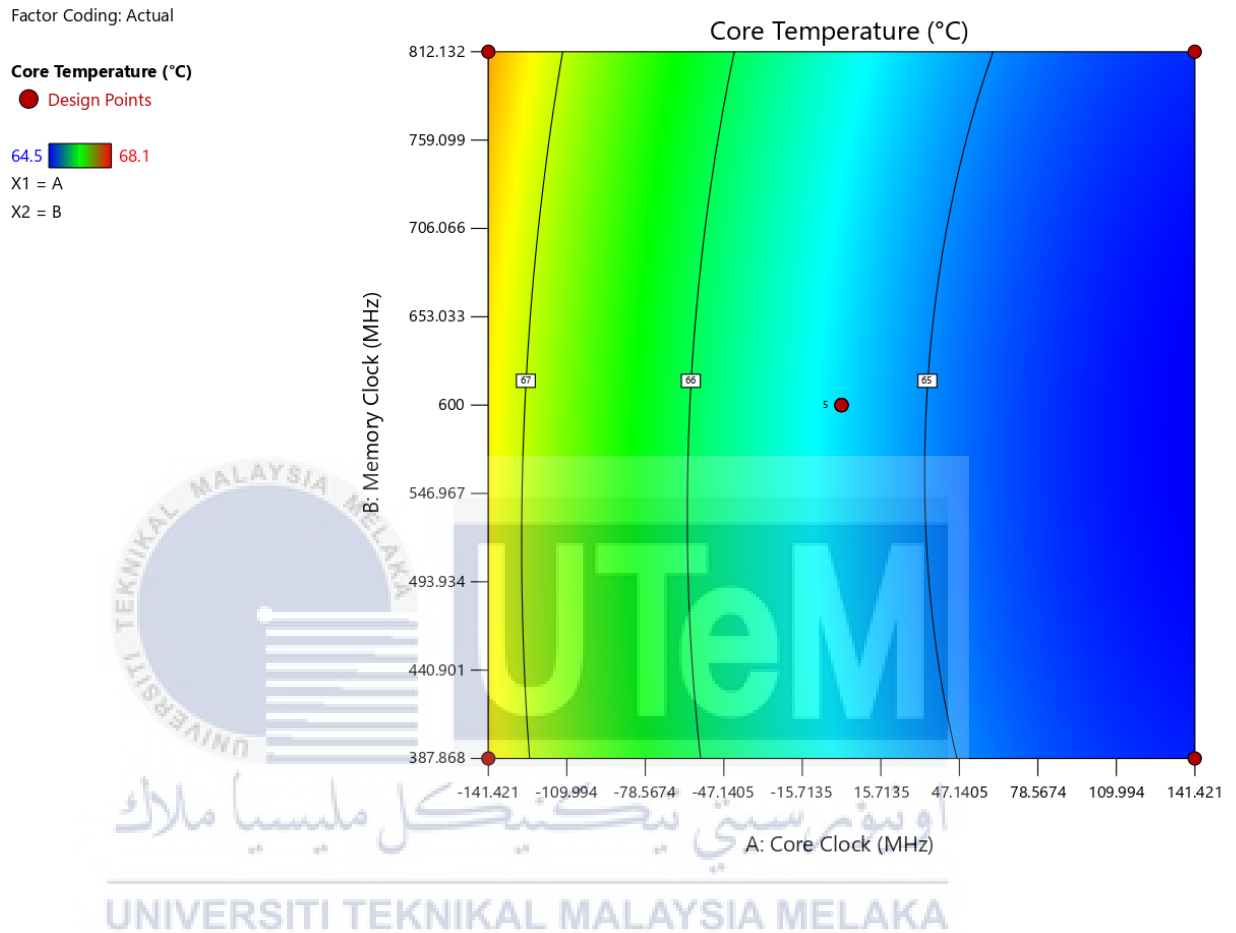


Figure 4. 7 : Contour plot for the combined effect of core clock (A), memory clock (B), and core temperature

The Figure 4.7 illustrates the effect of the memory clock and core clock on core temperature. Increasing the core clock up to a maximum limit of 141.421MHz and increasing the memory clock up to a maximum limit of 812.132 MHz causes a decrease in core temperature from 67 °C to 65 °C. The red color denotes locations with a high core temperature, while the blue denotes areas with a low core temperature. The areas are yellow and red for core clock values ranging from (-141.421 to -117.312 MHz), indicating

that the GPU's core temperature is relatively high in those areas. The core temperature of the GPU is relatively low at these settings of 15.7135 to 141.421MHz of core clock and the graph is in blue colors.

$$\text{Core temperature} = 65.3 + -1.3864 * A + 0.0676777 * B + -0.05 * AB + 0.56875 * A^2 + 0.14375 * B^2 \text{ (Eq. 2)}$$

A small P-value (less than 0.05) implies a close association between calculated responses and independent inputs, except for Lack of Fit. The most critical parameters influencing core temperature, as shown in Table 4.3, are A (Core clock) and B. (Memory clock). The negative sign in front of the components in Eq. (2) shows that A component with a multiplier of -1.3864 and AB component with a multiplier of -0.05 may reduce core temperature. According to the equation, the A component, which is the core clock, contributes the most to the GPU's core temperature because it has the highest multiplier value. The software also confirms this component because the p-value for core clock is less than 0.0001.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

4.3.3 Hash Rate

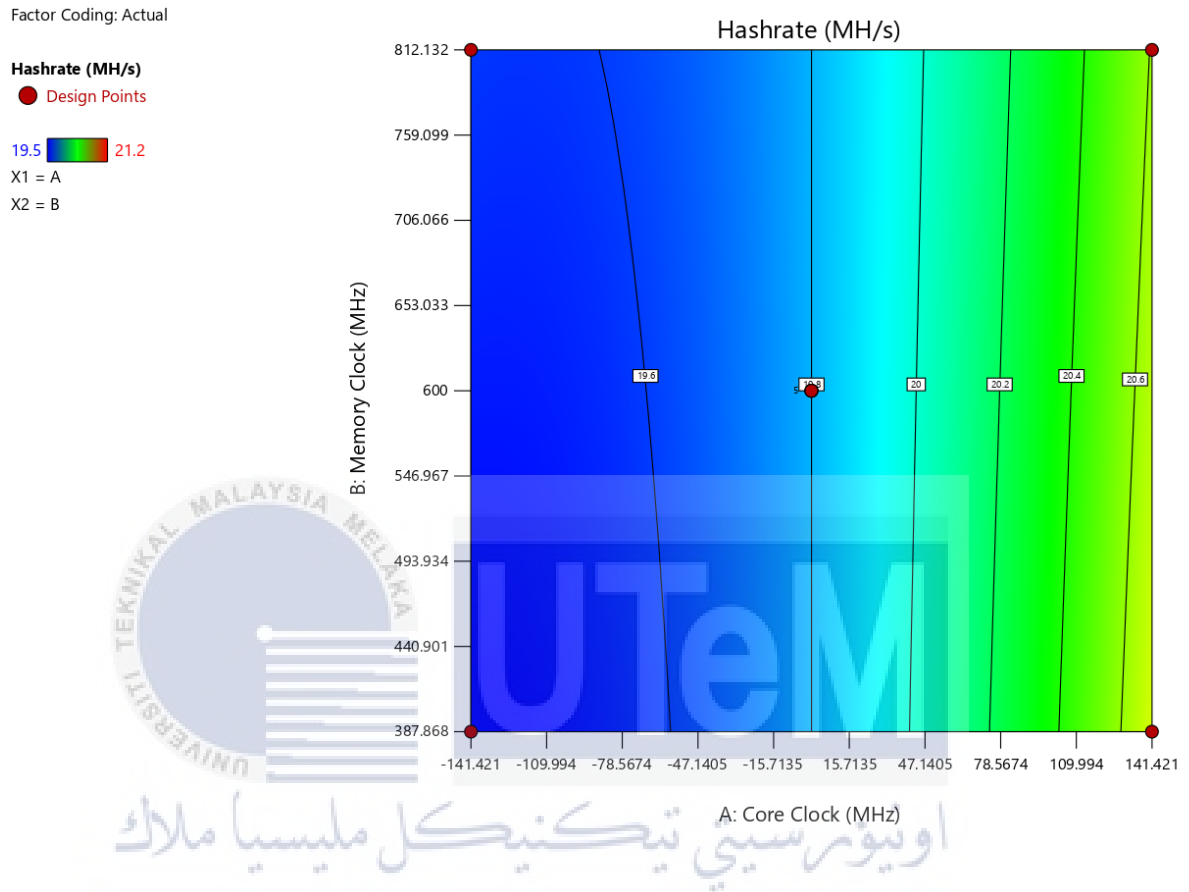


Figure 4. 8 : Contour plot for the combined effect of core clock (A), memory clock (B), and hash rate

The figure above illustrates the effect of the memory and core clock on the hash rate. The hash rate increased from 19.6 MH/s to 20.6 MH/s with the rise in core clock up to a maximum of 141.421MHz. Yellow denotes places with a high hash rate, whereas blue denotes areas with a low hash rate. For the core clock values of (126.74 to 141.421 MHz), the locations are colored yellow, indicating that the GPU's hash rate is relatively high in those areas. Furthermore, the core clock values (-141.421 to 17.558 MHz) in the graph

above are blue, indicating that specific clocking settings have a low hash rate value than other independent variable values.

$$\text{Hash rate} = 19.8 + 0.00494454 * A + -5.07744e-17 * B + -1.66667e-06 * AB + 1.5e-05 * A^2 + 3.59548e-20 * B^2 \text{ (Eq.3)}$$

Except for Lack of Fit, a small P-value (less than 0.05) indicated a statistically significant between calculated responses and independent inputs. As indicated in Table 4.6, the elements influencing the core temperature the most in this situation are A (Core clock) and B (Memory Clock), A2 and B2. As the positive sign of components in the (Eq.3), it is evident that core temperature can be increased by A component with a multiplier of 0.00494454, A2 component with a multiplier of 1.5e-05, and B2 component with a multiplier of 3.59548e-20. So, it is possible to conclude that the A component, which is the core clock, is the primary contributor to the hash rate values because it has the highest multiplier value. The software confirms this component because the p-value for the core clock is less than 0.0001.

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

4.4 Optimization of Single-Fan Graphic Processing Unit (GPU)

Table 4. 8 : Optimization of Single-Fan Graphic Processing Unit (GPU)

Name	Goal	Lower Limit	Upper Limit	Lower Weight	Upper Weight	Importance
A: Core Clock	is in range	-141.421	141.421	1	1	3
B: Memory Clock	is in range	387.868	812.132	1	1	3
Fan speed	minimize	2683	2859	1	1	3
Core Temperature	minimize	64.5	68.1	1	1	3
Hash rate	maximize	19.5	21.2	1	1	3
Ambient Temperature	none	28.7	28.8	1	1	3
Hot Spot Temperature	none	75.6	79.6	1	1	3

Based on the analyses, numerical optimization was performed by the desirability function using Design Expert Software. The core and memory clock values acquired from the Minerstat website are within the optimal range. As a result, the core clock has been set to -141.421 MHz for the lower limit and 141.421 MHz for the maximum limit. Meanwhile, the lower limit for the memory clock is 387.868 and the upper limit is 812.132. Furthermore, the response variables' objectives are minimal fan speed, minimum core temperature, and maximum hash rate. The numerical optimization produced 21 alternative solutions, with only one chosen option that is ideal for GPU operation. Table

4.8 lists the anticipated and experimental fan speed, core temperature, and hash rate for each of the twenty-one suggested option.



Table 4. 9 : Numerical optimization solution

Number	Core Clock	Memory Clock	Fan speed	Core Temperature	Hash rate	Hot Spot Temperature	Power Consumption	Desirability
1	141.421	481.556	2688.215	64.517	20.686	75.536	91.458	0.877
2	141.421	480.660	2688.217	64.518	20.686	75.537	91.455	0.877
3	141.421	483.524	2688.211	64.516	20.685	75.535	91.465	0.877
4	141.421	485.651	2688.207	64.515	20.685	75.534	91.472	0.877
5	141.421	477.378	2688.223	64.520	20.687	75.538	91.443	0.877
6	141.421	490.329	2688.198	64.512	20.684	75.532	91.489	0.877
7	141.421	471.003	2688.236	64.525	20.688	75.541	91.420	0.877
8	141.421	463.221	2688.252	64.531	20.690	75.545	91.393	0.876
9	141.421	501.416	2688.178	64.505	20.681	75.528	91.531	0.876
10	141.421	458.561	2688.262	64.534	20.691	75.548	91.378	0.876
11	141.421	507.439	2688.167	64.502	20.680	75.526	91.553	0.876

12	141.421	448.853	2688.284	64.543	20.693	75.554	91.345	0.876	
13	141.421	513.624	2688.157	64.499	20.678	75.524	91.577	0.876	
14	141.421	445.084	2688.292	64.546	20.694	75.556	91.333	0.876	Selected
15	141.421	538.071	2688.120	64.489	20.672	75.519	91.673	0.875	
16	141.421	564.070	2688.086	64.483	20.666	75.516	91.781	0.873	
17	141.421	608.609	2688.043	64.483	20.656	75.519	91.979	0.871	
18	141.421	661.874	2688.016	64.500	20.643	75.536	92.237	0.868	
19	141.421	737.667	2688.025	64.554	20.625	75.585	92.645	0.859	
20	141.421	772.139	2688.046	64.591	20.617	75.616	92.846	0.854	
21	141.421	775.811	2688.049	64.596	20.616	75.620	92.868	0.853	

In the validation and confirming process, the experiment will be rerun using one solution provided by the Design expert software using the constructed equation for the confirmation and validation process. The solution is the one that has the highest hash rate value at a certain memory and core clock. Then, the experiment is run using the selected core and memory clock and the GPU response to compare it with the predicted values by the software. The predicted and experimental value deviations were below 2% at optimization conditions. This difference is due to technology limitation. The sensor is not too sensitive with small changes in GPU response, and the software (PhoenixMiner and Afterburner) has lower precision of 0.1 only. These slight deviations can be credited to a reliable testing device and the proposed mathematical model that can predict the value of GPU response. Numerical optimization based on CCD can predict the desired lowest fan speed, lowest core temperature, and highest possible hash rate at constant power consumption without manually finding each core clock and memory clock. The GPU response can be predicted using any core and memory clock settings with the obtained equation. Based on the validation process, it can be concluded that the accuracy of CCD is experimentally proven. In conclusion, choosing the CCD method will give the best result for this experiment.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

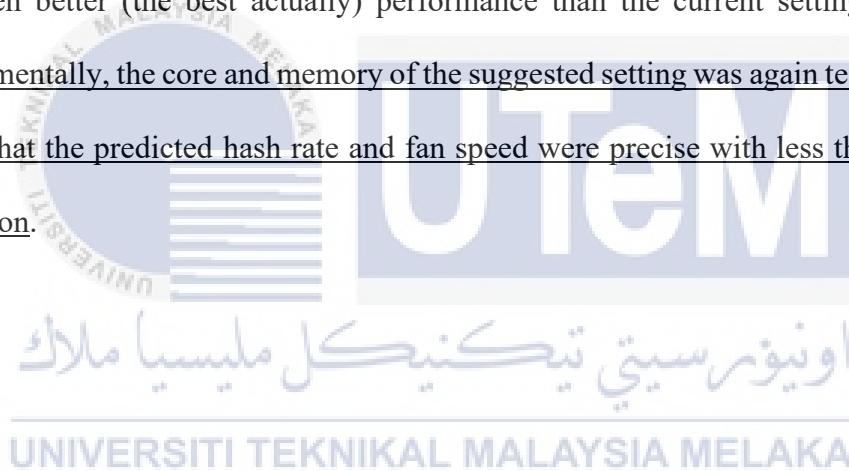
Core clock and memory clock are two main factors for an optimal performance of a graphic processing unit (GPU). Optimal performance of a GPU is not only reflected by its maximum processing power (hash rate) but more importantly - highest performance at a stable and continuous operation. This is vital especially for a GPU to be utilized under non-stop, high-power-density operation. Results show that a non-harmonious core and memory clock setting affects the overall performance adversely, leading to lower hash rate and an increase in core temperature.

This situation leads to even more energy is required to cool down the GPU. This situation fatigues the cooling system consists of fan bearing, rotor, electrical connection, and other parts. The harmonious clocking of both parameters was obtained from literature and varies from on-board chips to another. However, for GTX1070 used in this project, it was shown that the core memory clock can range from -200 to 200 MHz, while memory clock can be set from 300 to 900 MHz.

Acting upon the above-mentioned data set, a design on experiment was built. This is vital to minimize the amount of experiment as the setup and experiment running for

each clock setting is time-consuming. Optimization tool namely Design Expert was used to accommodate the process. The chosen design was Central Composite Design or CCD as literature shows that it was previously used by many researchers for an almost identical data set and CCD was shown to produce an equation which predicts the results with an accuracy of over 95 percent.

Using CCD for this project, only 21 experiments were required to be conducted, and all experiments were done in less than seven weeks. The generated equation of hash rate, fan speed and core temperature were obtained and it suggested a new clock setting for even better (the best actually) performance than the current settings. To prove it experimentally, the core and memory of the suggested setting was again tested, and results show that the predicted hash rate and fan speed were precise with less than two percent deviation.



5.2 Recommendation

Based from the literature review done during this project, it was clear that there are many factors contributing to performance of a graphic processing unit (GPU) for examples: number of fans on the GPU, rotation direction of fans either in the same direction or different, fan type (axial or radial), heat transfer fluid (air or liquid) and many more. This project only covers Single-Fan Graphic Processing Unit (GPU) and only on the GPU GTX 1070 model. There are plenty of other models which yet to be tested for an optimal operating parameter. That is on the side of the hardware.

Considering the software, the design of experiment via Central Composite Design (CCD) used in this project is also only one out of many designs available for optimization process. Response Surface Method or RSM, Box Behnken and many other optimization designs are still not yet tested. Even some literature shows that these designs were also capable to obtain the desired parameters at high accuracy. But due to time constraint, this was not done. It is highly advisable to execute the project again via these designs and compare them with CCD so that a more in-depth discussion can be produced.

REFERENCES

Annapragada, S.R., Salamon, T., Kolodner, P., Hodes, M. and Garimella, S.V. (2012). Determination of Electrical Contact Resistivity in Thermoelectric Modules (TEMs) From Module-Level Measurements. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2(4), pp.668–676.

Bhattacharya, S. (2021). *Central Composite Design for Response Surface Methodology and Its Application in Pharmacy*. [online] www.intechopen.com. IntechOpen. Available at: <https://www.intechopen.com/chapters/74955> [Accessed 8 Jan. 2022].

Bi, Z. (2018). Applications—Heat Transfer Problems. *Finite Element Analysis Applications*, pp.341–377.

Campatelli, G., Lorenzini, L. and Scippa, A. (2014). Optimization of process parameters using a Response Surface Method for minimizing power consumption in the milling of carbon steel. *Journal of Cleaner Production*, 66, pp.309–316.

Field, G., Marchel, I., Supijo, C., Pratama, H. and Najmi, A. (2017). Selection of a design for response surface You may also like Linear model analysis for ordinal response in a mixture experiment W Andani, A H Wigena and U D Syafitri - Fractional

factorial and D-optimal design for discrete choice experiments (DCE) A Z Tazliqoh, A H Wigena and U D Syafitri - Response Surface Method Using Box- Behnken Design for Probabilistic Resource Assessment: A Case Study in Atadei Recent citations Easing or tightening control strategies: determination of COVID-19 parameters for an agent-based model. *Materials Science and Engineering*.

FNIDES, B., YALLESE, M.A., MABROUKI, T. and RIGAL, J-F. (2011). Application of response surface methodology for determining cutting force model in turning hardened AISI H11 hot work tool steel. *Sadhana*, [online] 36(1), p.109. Available at: <https://link.springer.com/article/10.1007%2Fs12046-011-0007-7> [Accessed 11 Nov. 2021].

Ganesh Iyer, S. and Dipakumar Pawar, A. (2018). *GPU and CPU Accelerated Mining of Cryptocurrencies and their Financial Analysis*.

Ghelich, R., Jahannama, M.R., Abdizadeh, H., Torknik, F.S. and Vaezi, M.R. (2019). Central composite design (CCD)-Response surface methodology (RSM) of effective electrospinning parameters on PVP-B-Hf hybrid nanofibrous composites for synthesis of HfB₂-based composite nanofibers. *Composites Part B: Engineering*, 166, pp.527–541.

Gul, M., Zulkifli, N.W.M., Kalam, M.A., Masjuki, H.H., Mujtaba, M.A., Yousuf, S., Bashir, M.N., Ahmed, W., Yusoff, M.N.A.M., Noor, S., Ahmad, R. and Hassan, M.T. (2021). RSM and Artificial Neural Networking based production optimization of sustainable Cotton bio-lubricant and evaluation of its lubricity & tribological properties. *Energy Reports*, 7, pp.830–839.

Huang, C.-H. and Gau, C.-W. (2012). An optimal design for axial-flow fan blade: theoretical and experimental studies. *Journal of Mechanical Science and Technology*, 26(2), pp.427–436.

Intel. (2021). *What Is a GPU? Graphics Processing Units Defined*. [online] Available at: <https://www.intel.com/content/www/us/en/products/docs/processors/what-is-a-gpu.html> [Accessed 21 May 2021].

Jablin, T.B., Prabhu, P., Jablin, J.A., Johnson, N.P., Beard, S.R. and August, D.I. (2012). Automatic CPU-GPU communication management and optimization. *ACM SIGPLAN Notices*, 47(6), p.142.

Jayapragasan, C. and Reddy, K. (2017). DESIGN OPTIMIZATION AND EXPERIMENTAL STUDY ON THE BLOWER FOR FLUFFS COLLECTION SYSTEM. *Journal of Engineering Science and Technology*, 12(5), pp.1318–1336.

Kerr, A., Damos, G. and Yalamanchili, S. (2012). GPU Application Development, Debugging, and Performance Tuning with GPU Ocelot. *GPU Computing Gems Jade Edition*, [online] pp.409–427. Available at: <https://www.sciencedirect.com/science/article/pii/B9780123859631000307> [Accessed 21 May 2021].

Mahdavi, M., Tiari, S., De Schampheleire, S. and Qiu, S. (2018). Experimental study of the thermal characteristics of a heat pipe. *Experimental Thermal and Fluid Science*, 93, pp.292–304.

Maroosi, A., Muniyandi, R.C., Sundararajan, E.A. and Zin, A.M. (2013). Improved Implementation of Simulation for Membrane Computing on the Graphic Processing Unit. *Procedia Technology*, 11, pp.184–190.

Martínez-Frutos, J., Mar Tínez-Castejón, P. and Herrero-Pérez, D. (2017). Efficient topology optimization using GPU computing with multilevel granularity. *Advances in Engineering Software*, 106, pp.47–62.

Muyan-Özçelik, P., Glavtchev, V., Ota, J.M. and Owens, J.D. (2011). Real-Time Speed-Limit-Sign Recognition on an Embedded System Using a GPU. *GPU Computing Gems Emerald Edition*, pp.497–515.

Pirunkaset, M. and Laksitanonta, S. (2008). Study on the Effect of Blade Angle on the Performance of a Small Cooling Tower. *Nat. Sci.*, 42, pp.378–384.

Roy, N., Kuar, A.S. and Mitra, S. (2017). 7 - *Underwater pulsed laser beam cutting with a case study*. [online] ScienceDirect. Available at: <https://www.sciencedirect.com/science/article/pii/B9780857094858000073> [Accessed 7 Jan. 2022].

Sahoo, P. and Barman, T.Kr. (2012). ANN modelling of fractal dimension in machining. *Mechatronics and Manufacturing Engineering*, pp.159–226.

Sanhan, W., Vafai, K., Kammuang-Lue, N., Terdtoon, P. and Sakulchangsattajai, P. (2020). Numerical simulation of flattened heat pipe with double heat sources for CPU and GPU cooling application in laptop computers. *Journal of Computational Design and Engineering*, 8(2), pp.524–535.

Savic, I., Gajic, D., Stojiljkovic, S., Savic, I. and Gennaro, S. di (2014). *Modelling and Optimization of Methylene Blue Adsorption from Aqueous Solution Using Bentonite Clay*. [online] ScienceDirect. Available at:
<https://www.sciencedirect.com/science/article/abs/pii/B9780444634559500714>.

Shahidian, A., Ghassemi, M., Mohammadi, J. and Hashemi, M. (2020). Introduction. *Bio-Engineering Approaches to Cancer Diagnosis and Treatment*, pp.1–22.

Sheng, J.J. (2013). Steam Flooding. *Enhanced Oil Recovery Field Case Studies*, pp.361–388.

Smistad, E., Falch, T.L., Bozorgi, M., Elster, A.C. and Lindseth, F. (2015). Medical image segmentation on GPUs – A comprehensive review. *Medical Image Analysis*, 20(1), pp.1–18.

The Heat Pipe Advantage. (n.d.). [online] Available at:
<http://www.affinitas.com.sg/images/The%20Heat%20Pipe%20Advantage%20-%200%207.pdf> [Accessed 27 May 2021].

Vargas-Vazquez, J.-C., Gutierrez-Garcia, J.-A., Hernandez-Guerrero, A., Luviano-Ortiz, L., Zuñiga-Cerroblanco, J.-L., Rosas, J. and Cruz, S. (n.d.). *WATERBLOCK MODELLING FOR GPU LIQUID COOLING*.

Wu, Y. and Huang, D. (2019). Optimization design of axial fan blade. *Journal of the Chinese Institute of Engineers*, 42(6), pp.473–478.

Zamani, H., Tripathy, D., Bhuyan, L. and Chen, Z. (2020). SAOU. *Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design*.

Zeng, G., Li, S.H., Yu, Z.Q. and Lai, X.M. (2009). Optimization design of roll profiles for cold roll forming based on response surface method. *Materials & Design*, 30(6), pp.1930–1938.

Zhang, L. and Jin, Y.Z. (2011). Effect of Blade Numbers on Aerodynamic Performance and Noise of Small Axial Flow Fan. *Advanced Materials Research*, 199-200, pp.796–800.

Zhou, S., Zhou, H., Yang, K., Dong, H. and Gao, Z. (2021). Research on blade design method of multi-blade centrifugal fan for building efficient ventilation based on Hicks-Henne function. *Sustainable Energy Technologies and Assessments*, 43, p.100971.

APPENDIX



Appendix A : Top View of ASUS NVIDIA® GeForce TURBO-GTX1070-8G



Appendix B : Right side view of ASUS NVIDIA® GeForce TURBO-GTX1070-8G



Appendix C : Left side view of ASUS NVIDIA® GeForce TURBO-GTX1070-8G

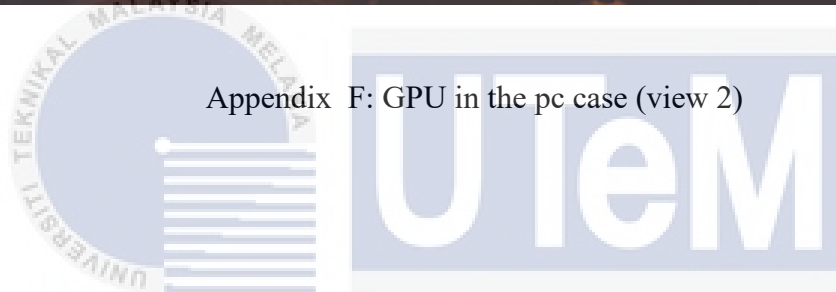
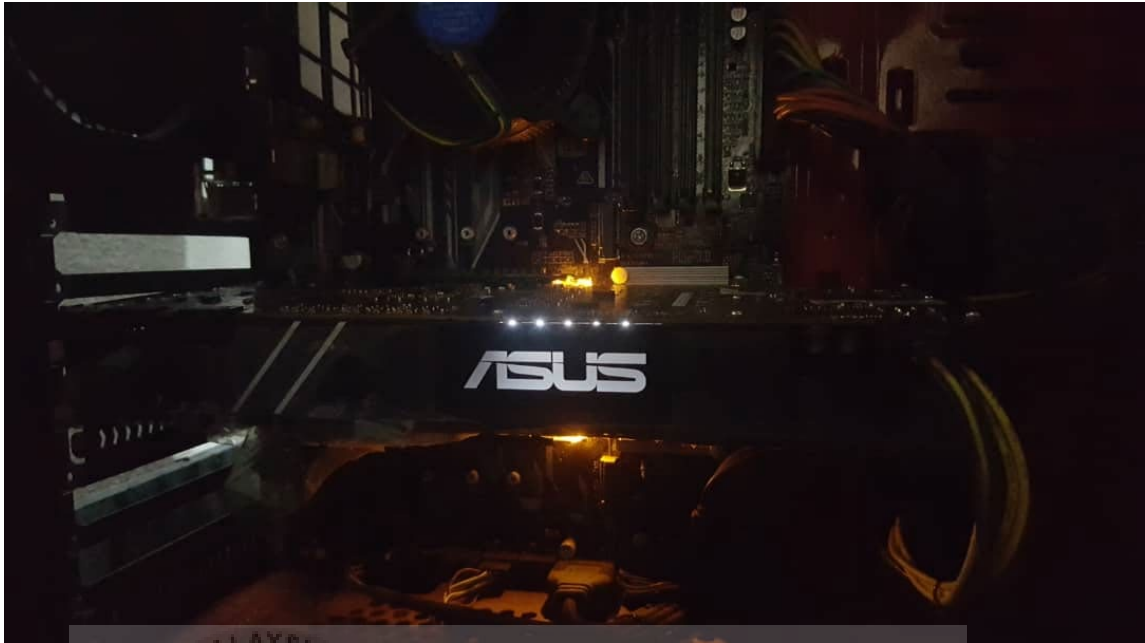


Appendix D : Back view of ASUS NVIDIA® GeForce TURBO-GTX1070-8G





Appendix E : GPU in the pc case (view 1)



Appendix F: GPU in the pc case (view 2)

GPU	Model	VRAM	Temp	Usage	Power	Mem	Min	Max	PL	MT	
GPU 0	GeForce GTX 1070	8117 MB - ASUS	24.08 MHz	76%	84%	121 W	0	200	-	800	- 120
GPU 0	Radeon RX 580	8192 MB - GIGABYTE	31.75 MHz	59%	62%	103 W	0	1225	1	500	2100
GPU 1	GeForce GTX 1070	8118 MB - ASUS	23.72 MHz	76%	77%	119 W	0	200	-	800	- 120
GPU 2	GeForce RTX 3080	10018 MB - MSI	90.15 MHz	64%	65%	207 W	0	-300	-	2600	- 240
GPU 3	GeForce GTX 1070	8119 MB - ASUS	24.10 MHz	77%	68%	118 W	0	200	-	800	- 120
GPU 1	Radeon RX 580	8192 MB - ASUS	31.63 MHz	54%	62%	90 W	0	1225	1	900	2100

Appendix G : PhoenixMiner software