THERMAL LOAD OF A DUAL-FAN GRAPHIC PROCESSING UNIT (GPU) SYSTEM ON A SERVER RACK

MUHAMMAD HAZIM BIN KAMARUDIN

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

THERMAL LOAD OF A DUAL-FAN GRAPHIC PROCESSING UNIT (GPU) SYSTEM ON A SERVER RACK

MUHAMMAD HAZIM BIN KAMARUDIN

This report is submitted
in fulfillment of the requirement for the degree of
Bachelor of Mechanical Engineering (Thermal Fluid)

Faculty of Mechanical Engineering
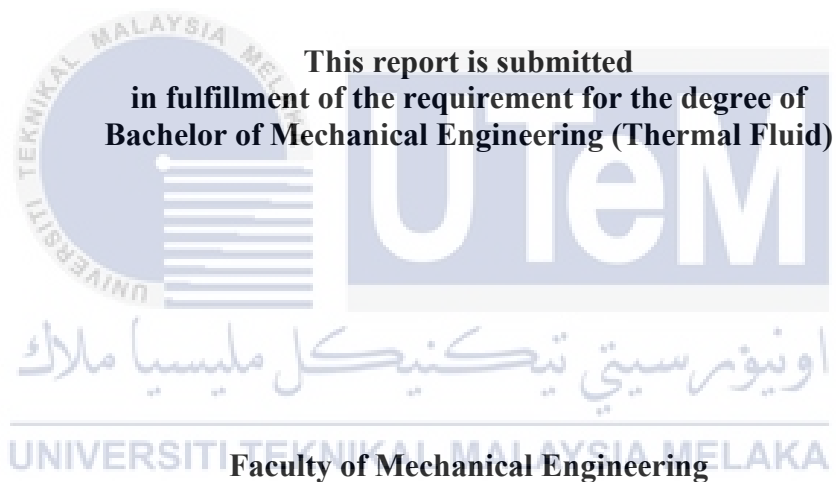
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

JANUARY 2022

DECLARATION

I declare that this project report entitled "Thermal Load of a Dual-Fan Graphic Processing Unit (GPU) System on a Server Rack" is the result of my own work except as cited in the references

Signature   :   ....

Name   :   ....

Date   :   ....

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   : ....

DEDICATION

This project I dedicated to my dearly loved family and supervisor upon their non-stop support and motivation in every part of my life.

ABSTRACT

Graphics Processing Unit (GPU) has gone through many revolutionary changes throughout the decade. GPU's processing power can 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 certain core clock and memory clock via the optimization tool of Design-Expert software. The response was recorded while applying constant amount of power supplied to the GPU. The GPU used is in the category of dual fan which is Gigabyte GeForce GTX 1070 WINDFORCE OC 8G. The relationship between the clocking (core and memory) and GPU responses (fan speed, core temperature, hash rate) were 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. For the validation and confirmation process, the best one core and memory clock was selected. The results from the validation process prove that the predicted response was precise with less than 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. *Tidak balas* GPU *telah direkodkan sepanjang kuasa yang malar dibekalkan kepada* GPU. GPU *yang digunakan dalam penyelidikan ini adalah* GPU *dua kipas yang bernama* GeForce GTX 1070 WINDFORCE OC 8G. *Hubungan 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%.*

vi

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

GPU	Graphic Processing Unit
CPU	Central Processing Unit
RSM	Response Surface Methodology
BBD	Box-Behnken Design
CCD	Central Composite Design
LAN	Local Area Network
WAN	Wide Area Network
VPU	Visual Processing Unit
AI	Artificial Intelligence
CUDA	Compute Unified Device Architecture
VRAM	Video Card Random-Access Memory
VRM	Voltage Regulator Module
PSU	Power Supply Unit
PCI	Peripheral Component Interconnect
ANOVA	Analysis of Variance
DOE	Design of Experiment
OC	Overclocking
RAM	Random Access Memory
PC	Personal Computer
LOF	Lack of fit

C.V   Coefficient of variation

LIST OF SYMBOLS

%	=	Percent
V	=	Voltage
W	=	Watts
Qmax	=	Maximum heat transfer capacity
$Q_{conduction}$	=	Rate of heat conduction, W
k	=	Thermal conductivity, W/m.K
ΔT	=	Temperature difference, K
Δx	=	Thickness, m
Q	=	Heat capacity, J
m	=	Mass of substance, kg
cp	=	Specific heat, J/kg.K
ΔT	=	Temperature difference, K
$Q_{convection}$	=	Rate of heat convection, W
h	=	Convection heat transfer coefficient, $W/m^2.K$
A	=	Surface area for heat transfer, m^2
Q_{rad}	=	Rate of heat radiation, W
ε	=	emissivity
σ	=	Stefan-Boltzmann constant, 5.67×10^{-8} $W/m^2.K^4$

T_s	= Absolute Temperature of Surface, K	
T_{surr}	= Absolute Temperature of Surrounding, K	
CFM	= Cubic Feet Per Minute	
rpm	= Revolutions per minute	
°C	= Degree Celsius	
R^2	= Coefficient of determination	
MH/s	= Mega hash per second	

CHAPTER 1

INTRODUCTION

1.1 Background

A server is a computer program that provides data or services to another computer, such as computer software and hardware. Servers are created to perform multiple essential data processing for client servers with access through the internet or local networks. A server does not have a screen or keyboard, and it is a computer that transfers data from one computer to another. The technology used in delivering the data can be employed to operate on a local area network (LAN) or over a wide area network (WAN). Site, mail, and other clients are all different forms of servers, and there are several options for each.

Server has a sensitive part and would lead to technical issues to the entire system and have many wire connections. A server rack is a storage and organizing system for electronics equipment such as GPU. Server racks help them organize the wire and protect the server and their parts from external damage. Stacking servers and other electronics equipment in a rack makes it easier to keep things organised and monitor airflow. The design of the server racks also plays an important role in increasing the server's efficiency (Gao et al., 2016).

The server's major component is the motherboard which connects all other components. Next is a processor and it was divided into two types: the Central Processing Unit (CPU) and the Graphic Processing Unit (GPU). GPU have become more programmable in recent years, allowing them to be used in many different fields. Since they're faster and more energy-efficient than a conventional CPU, GPUs can perform even larger tasks with greater effectiveness (Erik et al., 2014). To maximise the speed of the computer, parallel processing is used. In parallel computing, GPU are being used frequently. GPUs have much more parallelization capacities than CPUs as they have many cores compared to CPU (Ebubekir et al., 2018). Servers usually use several GPUs to have enhanced processing power. GPU was developed to simulate 3D graphics such as for AutoCAD, CATIA and SolidWorks software (Sadrieh et al., 2011). When time went by, they were more malleable and more competent, and they improved their skills. With their enhanced capabilities, graphic programmers could use advanced lighting and shadowing methods to improve the look of visuals. The field of deep learning has used GPUs to speed up other workloads. Also, a GPU can do several calculations simultaneously, which helps it advance the overall efficiency of the server and process large numbers of numbers quickly.

A GPU can be found on servers connected to the circuit where the CPU is located or integrated into the motherboard. There are multiple big brands in the GPU industry, including the NVIDIA, AMD, INTEL, and ARM franchises. At present, NVIDIA and AMD are known as the two most prominent GPU vendors. In the case of NVIDIA and AMD, they design and supply GPUs, which are passed on to third-party vendors such as MSI, ASUS, and Sapphire, who can make their customised changes without altering the GPU's chip in some way (Stewart., 2021). A GPU is a so-called multi-processor that performs multiple tasks simultaneously, while a conventional

processor does not enable this. This is because the GPU processor comprises hundreds or thousands of small cores or units that handle multiple tasks simultaneously for complicated graphics processing.

The speed of the GPU's VRAM is determined by the memory clock, whereas the core clock determines the speed of the GPU's chip. The GPU's core clock may be compared to a gaming PC's CPU and RAM clocks. The core clock often has a greater impact on gaming performance than the memory clock. Games' visual effects are temporarily stored in virtual memory (VRAM) on the graphics processing unit (GPU). More VRAM means the graphics card can process images faster, and faster VRAM means the users can store more assets. As a result, the games will render more quickly if the memory clock speed is greater. The clock speed of the graphics processor's GPU core determines how quickly it can process graphics.

1.2 Problem Statement

Recently, keeping electronic chips cold has been one of the biggest concerns due to the infinite number of systems requiring more efficient and less power-consuming technology (Jose-Carlos et al., 2018). The research is primarily looking for the best way to cool a GPU in the server rack since it releases a lot of heat. A significant heat source in electronic devices is the GPU (Siricharoenpanich et al., 2021). An issue related to the GPU heat of the system is semiconductor chips. If the GPU is assigned a workload including gaming and highly intensive apps, it can produce a great deal of heat. There has been a rise in interest in GPU as researchers realise, they need to find the cooling solution of the devices to satisfy the market's needs (Al-Rashed et al.,

2016). Several GPUs server designs have been upgraded to ensure good air circulation and prevent overheating. There are several methods of cooling systems such as passive and active to ensure the graphics card temperatures remain low. In active cooling, the fan is used, but for passive cooling, the heat sink is used (Svasta et al., 2017). However, server chassis has its limitations such as construction costs and it needed excessive room within for placement of the system.

Thus, to deal with this major difficulty, an analysis was carried out to keep the GPU cool using a dual-fan cooling system. Priority was given to some factors to ensure that the GPU maintains its thermal efficiency for maximum performance at an ideal temperature. A wide variety of fans was employed, but only a specific type of fans was being analysed. Any other points must be pointed out regarding the cooling component in GPU are the design of heat sink and heat pipe. While in this research, core clock and memory clock of a Dual-Fan Graphic Processing Unit (GPU) on a server rack were examined to determine the optimal fan speed, core temperature and the highest possible efficiency (hashing power) via optimization tool.

1.3 Objectives

The core objectives of this project are as follow:

1. To analyse the relationship between clocking (core and memory) and GPU responses (fan speed, core temperature, power consumption, and hash rate).

2. To locate the optimal fan speed, core temperature and the highest possible efficiency (hashing power) at certain core clock and memory clock via optimization tool.

1.4 Scope of Project

The scopes of this project are:

1. Gathering literature review of GPUs component.

2. Getting the Dual-Fan Graphic Processing Unit GTX 1070 response (fan speed, core temperature and hash rate) at a certain core and memory clock using Afterburner and PhoenixMiner software.

3. Finding solutions using numerical optimization using Design-Expert software.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

This chapter aimed to explain in detail and get a better understanding related to the project title. In this Literature Review, the Graphic Processing Unit (GPU) and the thermal management are the main keywords. The evidence and findings on this project will be investigated to their fullest to ensure complete understanding. All of the fundamentals and equations that will be applied to this project can be found in this chapter. This research aims to locate the optimal fan speed, core temperature, and the highest possible efficiency (hashing power) at certain core clock and memory clock via optimization tool.

2.2 Introduction of GPU

In today's world, maximum processing speed is needed. While the progress made by the Central Processing Unit (CPU) over the last two decades has been enormous, it has now reached a halt. To address this, NVIDIA in 1999 introduced the Graphics Processing Unit (GPU) or the Visual Processing Unit (VPU), a modern

highly parallel and multithreading processor designed for high degrees of computation. A (GPU) is a single-chip processor that manages and improves video and graphics output.

The most critical feature of a computer server is the graphics processing unit. In video and computer games, the GPU is essential for better picture rendering and many other applications as well. This is because the GPU is an extremely proficient processor at processing and rendering images. These GPUs are growing in popularity in non-gaming and artificial intelligence (AI). The GPU is a processor with the same concept as the CPU, but CPU does not have anything to do with the device's graphics kit. GPU is basically a CPU programmed and designed for graphics control. Video games have become more computationally intensive with hyper-realistic visuals and massive, complex in-game environments. With the growth of virtual reality technology and new monitor devices such as 4K displays and quick refresh speeds, graphics processing demands rapidly increase. Until now, GPU has been utilised as an academic exercise, improving performance in game computers, and as computational biophysics in terms of protein folding simulation, scalable molecular dynamics, and generating electrostatic potential maps (Owens et al., 2008). Figure 2.1 below shows the picture of dual-fan GPU, GeForce RTX™ 3060 Ti EAGLE OC 8G.

Figure 2.1: GeForce RTX™ 3060 Ti EAGLE OC 8G

2.3 The Cause of GPU Temperature Rises

Some electronic components may produce a large amount of heat when in operation. Failure to efficiently dissipate this heat away from the system and the device will result in durability issues and shorter operating lifetimes (Zambri et al., 2018). In today's processors, the number of cores within a chip rises significantly. NVIDIA's GTX280, for example, has 30 streaming multiprocessors with 240 CUDA cores, and NVIDIA Fermi GPUs have 512 CUDA cores. The most recent AMD processors have 12 cores in multi-core processors. Because of the large number of cores, it is difficult to build efficient power and temperature-controlled architectures (Dasgupta et al., 2012). Many researchers are studying and improving heat dissipation processes due to the continuous reduction in the physical size of electronic devices, the high heat flux emitted by the chip, and the limited space for installing air cooling systems (Nugroho et al., 2019). Overclocking is a good way to achieve additional GPU efficiency. It's

also a perfect way to put pressure on the machine, making it operate harder and, as a result, causing the GPU to overheat.

2.4 Components That Contribute Heat for GPU

Any part in the Graphic Processing Unit (GPU) can contribute to the system's heat. The first part is the processor or the chipset in the GPU. Other components inside the GPU, Video Card Random-Access Memory (VRAM) and connectors also generate large amounts of heat. Integrated circuits (ICs) produce heat as they are working. When it comes to graphics cards, there are several thousands of them located in a small space, which ensures that the heat they generate is difficult to manage (Evanson, 2021).

The processor is crucial since it serves as the computer's control unit and core. A control unit uses incoming data to transform and then transmits the resulting control signals to the central processor. Once the computer's processor receives instructions from the attached hardware, it tells the hardware what operations to do. A component's or circuit's capacitance refers to its ability to collect and store energy in the form of an electrical charge. Due to the required charging and discharging of capacitances, GPU produces heat via current flow that enters transistor channel resistances.

The more Video Random Access Memory (VRAM) on a card, the more complex textures or 3D meshes it will load. The more VRAM required, the higher the resolution of the picture to render. If the GPU's VRAM is overloaded with texture and image data, the overflow goes to the device RAM, negatively influencing performance. Essentially, the video will continue to render, but frame rates will drop. The capacitors, transistors, wire, and other components in VRAM serve as resistance, causing RAM

to heat up. VRAM produces heat as the transistors change states. In addition, when the electricity is turned on, the components will generate heat. Electricity is converted to heat energy when it flows through them or is blocked within.

In addition, the GPU system's power supply unit (PSU) and power connector also generate heat. The heat emitted by the power supply is the internal energy loss of the power supply. The device is able to transform AC current coming from a power outlet into a steady low voltage DC current for powering a GPU component. Capacitors in power supply units can hold strong electrical charges, and because of that it will produce some heat. The power connectors are usually used on mid-range to high-end graphics cards, since these cards need more power to operate (Akshat Verma, 2015). PCI stands for Peripheral Component Interconnect, and it is a standard interface for connecting electronic devices to the motherboard. In general, all cards below the midrange levels only pull power from the PCI Express x16 slot. Power supply connectors are needed for cards with the highest performance and specifications. If the card has a higher maximum performance, it needs to use either a 6-pin or 8-pin PCI Express power connector (Aleksandar Cosic, 2021). Using the 6-pin connector, the GPU will be supplied with an additional 75 watts of electricity. 150 watts will be provided through the 8-pin power connector. For the more popular graphics cards such as Nvidia's GeForce RTX 3060, a 6-pin and an 8-pin connector will be required so it can provide 300W at its maximum power draw. An increase in power in PSU generally increases the chance of overheating.

A voltage regulator module (VRM) is a word used to describe a device that regulates the voltage. A VRM transforms a machine power supply's 12 V output to a lower voltage (from 1.1 V to 3.3 V) that a GPU may access. Modern GPUs with demanding power and current requirements use VRMs to monitor and decrease the

voltage sent to the components constantly. As processors gain in power, heat generated by the VRM also increases. Overclocking a GPU necessarily involves the use of VRM. Since VRMs have to keep the voltage at a certain level, they may get very hot when doing their jobs. (Scharon Harding, 2018). Figure 2.2 below shows the basic components in GPU.
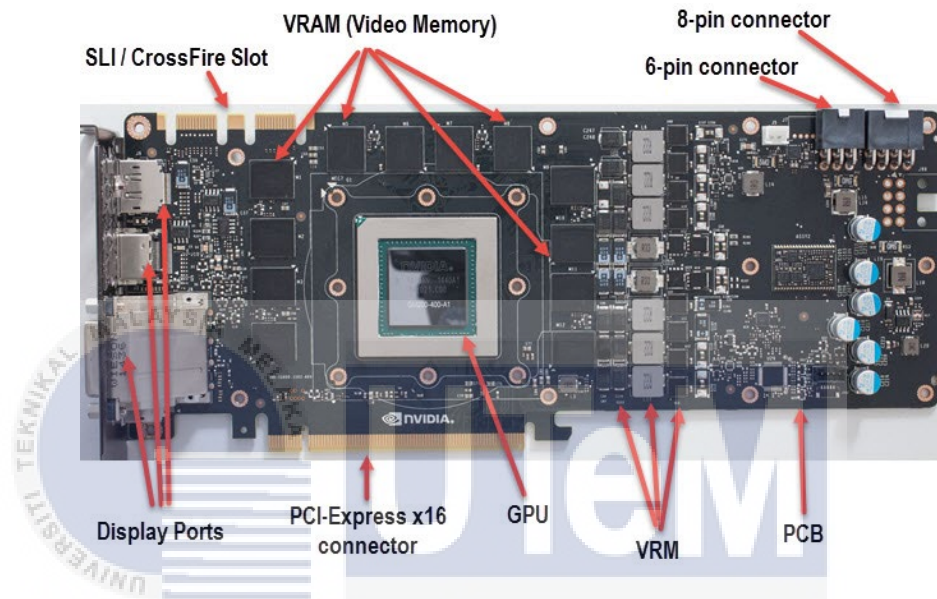
Figure 2.2: Components in GPU with labels

2.5 Components That Help GPU Cooling Process

The components that produce heat have been described in the previous section. Some crucial parts are allocated to keep the GPU temperature within range to preserve the GPU temperature. The lower GPU temperature is very important for efficiency and performance. Graphic cards are growing stronger. However, heat production levels are increasing. That is why an excellent cooling system is essential. Any graphic cards currently on the market are equipped with complete cooling systems and solutions. Overheating the GPU would cause the lifetime of the GPU would be shorter. As a

result, the cooling system in a graphics card cannot be neglected. An overheated GPU usually functioned and performs well for a short time, and then the performance will decrease substantially. Overheating GPUs may be caused by various factors, including poor ventilation, dust, and insufficient cooling. Many components perform as cooling components, including a heat sink, fan, and heat pipes. These cooling components will be explained in detail on how they function, the component type and the component issue.

2.5.1 Heat sink

If the GPU's temperature rises, the graphical performance will be degrading and, in an extreme case, it will harm the computer's performance. A heat sink is employed to solve the problem of GPU overheating. Conduction happens as one or more electric components transfer heat from the heat source to the heat sink due to forced fluid movement, while convection occurs when fluid moves through and out of the heat sink through the inlet and exhaust ports (Jie Ng et al., 2018). A heatsink is a tiny metal component used to absorb heat from the neighbouring part. Heat sinks operate by absorbing heat from the connected components, especially the component that generates heat. Most heatsinks are made out of fins similar to a radiator. Heat will be transferred to the fins when the GPU heats up.

Low heat removal efficiency of the heat sink might harm the electronic component as the temperature increases. This challenge has led computer manufacturers to use advanced technologies to boost the speed of electrical parts by improving heat removal. On the other hand, the smallest computer size increases the

total flux resistance of the system and ultimately limits fluid flow between heat sink fins (Hamdi et al., 2017). Heat sinks are widely used because they are simple to manufacture, cost-effective and heat-dissipation reliable. Heat sinks are an important component for a healthy and efficient compute especially when overclocking a GPU. Overclocking really increases the amount of heat produced by the GPU.

Nevertheless, GPUs may generate enough heat to cause internal damage during regular usage without a heatsink. Heat sinks are used for cooling electronic equipment by expanding the surface heat dissipation (Ibrahim et al.,2017). The higher the surface area of the heat sink the higher the thermal conductivity of it. In order to build a decent heatsink, factors that impact the performance and efficiency of heat sinks such the material, the measurement and the total surface heat transfer coefficient must be highlighted. Moreover, because convection occurs in the heat sink component, simpler heat sinks might replace bulky and complicated designs with advancements in air cooling.

2.5.2 Heat pipe

Heat pipes are acknowledged as one of the most efficient solutions for passive heat transfer because of their superior heat transmission characteristics, high efficiency, and design simplicity (Jouhara et al, 2018). Heat pipes are metal tube pipes filled with a fluid or coolant that transfers heat into a continuous cycle through evaporation and condensation (Mahboobe et al., 2017). The heat pipe is filled with liquids that flow to cool the GPU, while the heat sink removes heat from the liquid. The thermal plate on top of the processor will be connected to the heat pipe. In

addition, the heat pipe will remove the heat generated from the processor and transferred to the heat sink. Many aspects must be taken into consideration when designing a heat pipe for example, material compatibility, operating temperature range, diameter, power constraints, thermal and operational orientation. Heat pipes are utilized when heat is to be transported more freely into the environment from one side to the other.

According to Mahboobe et al., (2017), a heat pipe is a vacuum-sealed container charged with a suitable working quantity of fluid. The container inside surface is coated with a wick. When heat is supplied to a container end called the evaporator, the fluid is then evaporated. The vapour then goes to the opposite end of the container, which is called the condenser. The vapour condenses the condenser part of the heat pipe and transfers energy to a heatsink. The generated liquid is forced by the capillary pressure of the structure of the wick back to the evaporator. Using such processes, heat pipes may exhibit heat transfer rates that are larger in magnitude than in highly conducting metals.

2.5.3 Fan

Fan is a crucial airflow control component. It is utilised to remove heat from the heat sink and other components. The fan's principal function is to dissipate heat from the GPU until an ideal operating temperature is reached. The fan draws cold air from the surrounding environment, also known as ambient temperature, and distributes it to the heatsink, which is already removing the heat from the electronic equipment. The effectiveness of heatsinks depends mainly on the fan being pushed to transfer air

(Choi et al., 2012). The increased fan speed leads to noise, vibration issues and higher power consumption. Therefore, an ideal speed and number of fans must be considered. According to Nugroho et al., (2018), the cooling operation using a fan is called "air cooling". The heat produced by the electrical components is transferred to a conventional fan, which blows or inhales the heat. Because the capability of current conventional fan technology is limited, new cooling methods are required. Many researchers have been drawn to study and enhance thermal dissipation efficiency because of the continuous reduction in the physical size of electronic equipment, which has resulted in significant heat flux produced by the chip and limited space limitations for air cooling systems.

2.6 Design of Fan

2.6.1 Angle of Blade

The angle between the chord of the blade and the plane of rotation, also known as the angular position, is measured at a given location along the length of the blade. It is a great idea to use the blade angle to change the fan's angle of attack. Whether the fan is running at a constant or variable speed, the fan blade angle must be adjusted to provide the most efficient angle of attack. According to Huang and Gau (2012)'s research, redesigned optimum fan blades can improve fan airflow, therefore increasing the performance of the axial-flow fan. The optimum fan blade has been redesigned to incorporate the blade's setting angle, blade root chord, and blade end chord. During the experimental setup, the number of blades, fan speed, and fan diameter are all kept

constant. Figure 2.3 below shows the system variables for the original fan blade and the redesign 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 (θ,°)	47	59	33.18	52.22
Number of blades	7	7	7	7
Fan speed (ω, rpm)	2500	2500	2500	2500
Hub diameter (D_{Hub}, mm)	40.0	40.0	40.0	40.0
Fan diameter (D, mm)	110.0	110.0	110.0	110.0
Hub height ($H_u+H_l=H$, 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
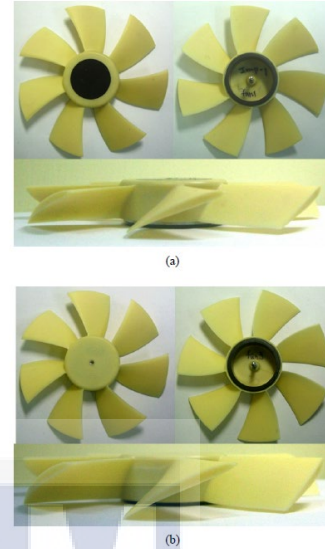
Figure 2.3: The system variables for various fan blades and fabricated original and optimal fan

We can see from the Figure 2.3 above that the setting angle of the blade influences the air volume and flow rate. At the setup with a setting angle of 59 degrees, air volume flow rate was at its greatest at 88.84 CFM, and decreased in smaller settings, ending up at 63 CFM. According to Pirunkaset et al. (2008), blade angles influence the performance of a small cooling tower. At the fan blade angles of 59°, 67°, 75°, and 83°, the research showed that the setting fan blade angles in the cooling tower were 5 tonnes of refrigeration. Based on the Figure 2.4 below, a blade angle of 59° has the greatest CFM, followed by 67°, 75°, and 83°. It's possible that the absence of performance curves at different blade angles is due to manufacturers failing to properly tune the fan blades during installation. As a result, determining the optimal

angle of the blade is critical since it influences the performance and efficiency of the cooling system.
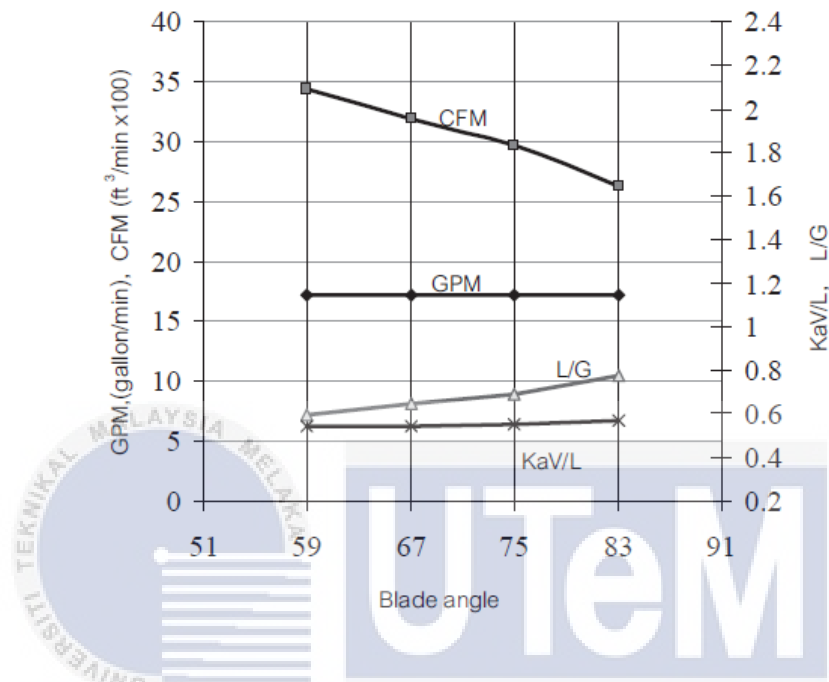
Figure 2. 4: The effect of blade angle on CFM

3.6.2 Design of Blade

An axial fan is a kind of fan that draws in fresh air by spinning an axial flow of gas around an axis. This fan has a circular flow entrance and exit. The flow generating function of the fan may be aided by generating a pressure differential and, consequently, a force that propels air through the fan. According to Wu & Huang (2019), increasing the lift-drag ratio by optimising the chord length and installation angle of the blade along the blade height. Using orthogonal optimization, three design alternatives (straight blades, C-type blades, and forward swept blades) were investigated in this study as shown in Figure 2.5 below.

Straight blade fan	Fan with a 16º forward-sweep blade	C-type blade fan
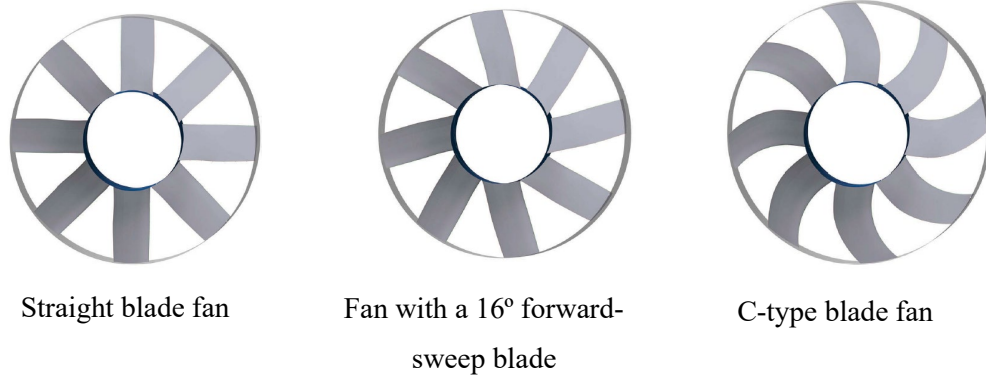
Figure 2.5: Design of blade fan

In this comparison, the forward-swept blade fan outperformed the C-type blade fan by a quantitative comparison as shown in Figure 2.6 below. Meanwhile, the straight blade fan was the least efficient, despite having a perfectly straight blade. C-type blade aerodynamics is better when blades are straight compared to when the blades are swept forward. This variant features a forward swept blade, and C-type, which is less noisy.

	Straight blade	Forward-sweep blade	C-type blade	Units
Speed	985	985	985	[r min^{-1}]
Volume flow rate	50.9628	50.6329	51.1943	[m^3 s^{-1}]
Pressure rise	1496.05	1520.47	1405.6	[Pa]
Total pressure efficiency	93.5621	94.8116	94.0066	[%]

Figure 2.6: Blade fans simulation results
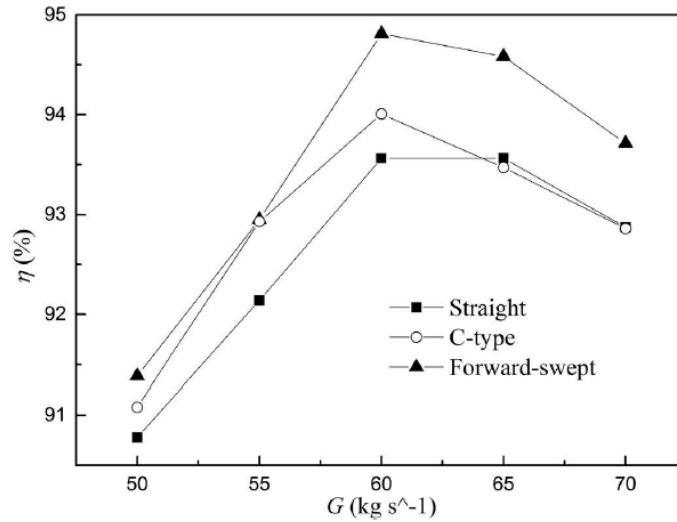
Figure 2.7: Mass flow rates using three kinds of fan total pressure efficiency curve

2.6.3 Numbers of Blade

The outside portion of the blade is mainly responsible for carrying out the job of the fan. Moreover, having a high number of blades really enhances the air flow rate and static pressure. Zhang and Jin (2011) claim that several design factors, including number of blades, hub ratio, blade angle, and blade number, impact small axial fans' aerodynamic and noise performance.
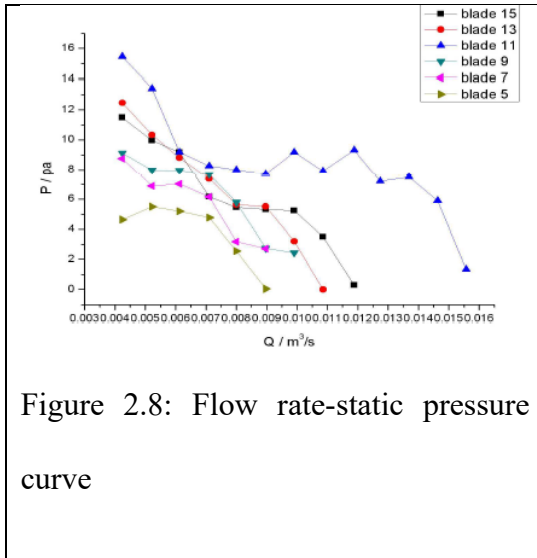
| Figure 2.8: Flow rate-static pressure curve | Figure 2.9: Flow rate-efficiency curve |
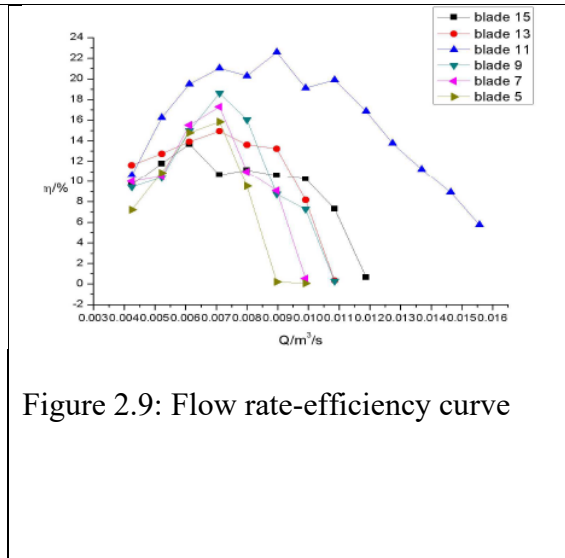
In general, total pressure and efficiency increase with the number of blades based on Figure 2.8 and Figure 2.9 above. When the number of blades is 11, total pressure and efficiency are optimal. Therefore, in the minimum flow point, flow loss is minimal, and efficiency is at its maximum since the static pressure gradient is at its minimum.
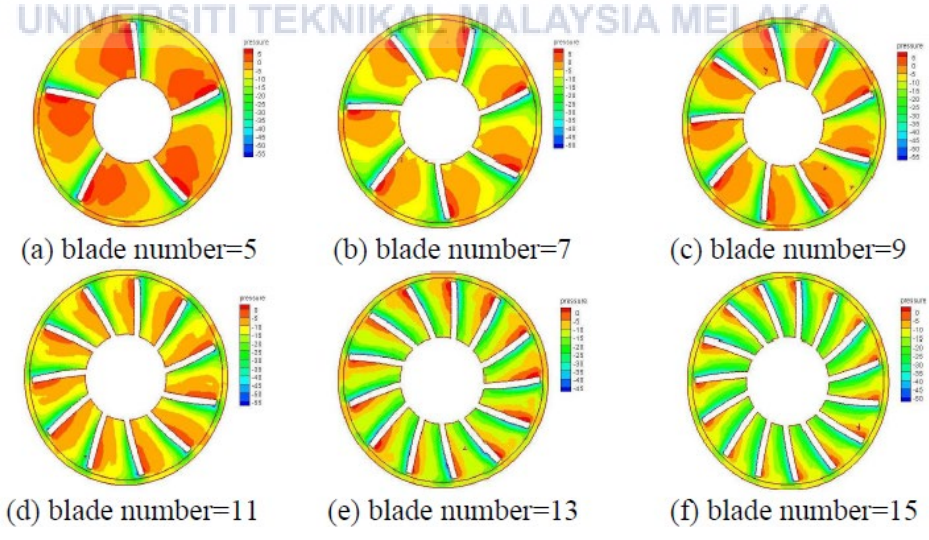
Figure 2.10: The static pressure field distribution

In flow fields, high and low pressure occur. High pressure is often greater than low pressure at first and low pressure gradually increasing as the number of blades grows as shown in Figure 2.10 above. When blade number 11 is used, the flow pressure distribution is the most even
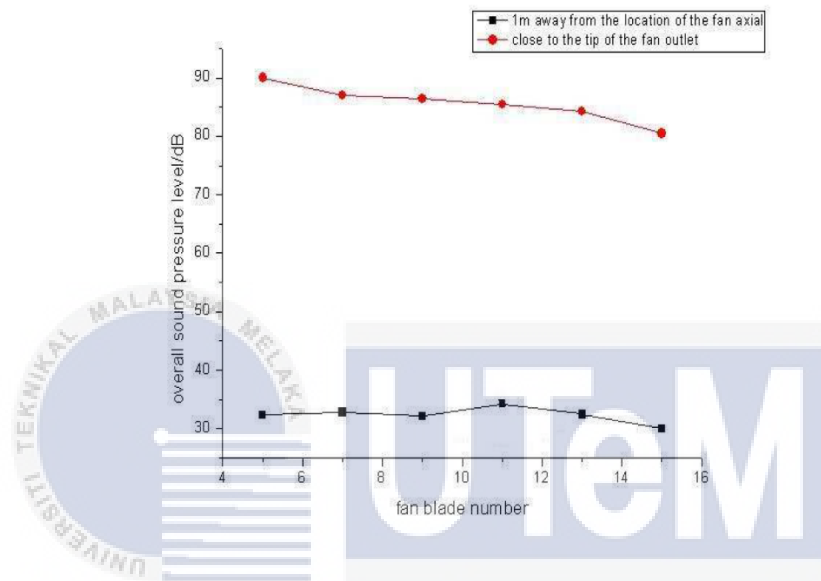
Figure 2.11: Fan blade number against overall sound pressure level

One of the primary causes of broadband noise in fan steady flow fields is Eddy current, and effective Eddy currents largely exist in the tip clearance. As fan blade numbers increase, sound pressure level (SPL) decreases at the tip of the fan blades. However, at 1 metre from the impeller centre point, the sound pressure level rises again for fan models with 11 blades as shown in Figure 2.11 above. The reason might be related to the fan blade clearance generating static electricity.

According to Rajabi et al. (2017), as the number of blades rises, so does the maximum flow rate. Though certain blades have been changed, the rotational speed has remained constant at 2900 revolutions per minute. In this instance, the research

divides fans into three categories based on the number of blades: four, five, and six as shown in Figure 2.12 below.

Figure 2.12: Changes in number of blades

The characteristic curves in the Figure 2.13 below indicate that when the flow rate increases from 200 CFM to 400 CFM, the pressure will rise. The static pressure is significantly decreased when the flow rate is raised from 400 CFM to 800 CFM. The use of six-bladed propellers improves maximum pressure at 400 CFM by 32%. When the number of blades increases, the surface area of the blades expands, increasing the amount of energy transferred when they collide with fluid. As a consequence, the outflow pressure rises. When there are high flow rates, the most efficient separations are utilised, and increasing the number of blades has little effect on the output pressure. The graph below indicates that at the maximum flow rate of 800 CFM, the static pressure generated by the different number of blades is almost equal to 42 Pa. Increasing the number of blades has no effect on the flow rate range. While increasing the exit static pressure, it also increases the input flow pressure.
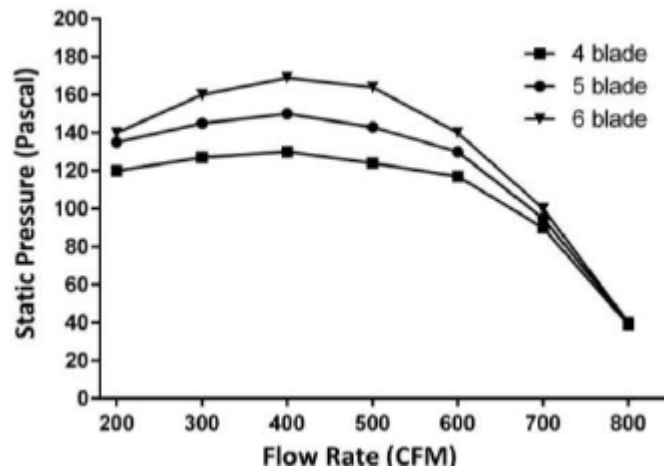
Figure 2.13: Characteristics curves of the fans with different number of blades

2.7 Heat transfer

Heat is the kind of energy that is related to the movement of molecules or atoms. Heat is transferred from high to low temperatures. Heat can be transferred in three ways which are conduction, convection and radiation. Heat transfer is the transfer of energy caused by a temperature differential. The basic equation in heat transfer is shown below:

$$Q = mc_p\Delta T$$

Where,
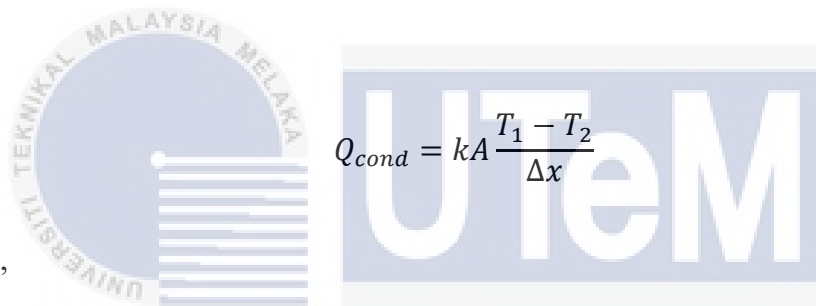
Q = Heat capacity, W

m = Mass of substance, kg

c_p = Specific heat, J/kg.K

ΔT = Temperature difference, K

Conduction is the transmission of heat between objects that are in touch with each other. The more efficient the conductor, the faster heat transfer will occur. When a material is heated, particles acquire more energy and vibrate more. As these molecules bump into nearby particles, they transmit energy to them. This flows from the heated end of the material through the cold end, thereby combining the heat from the hot end with the rest of the substance. For conduction occurs in solids, heat transfer is done through lattice vibrations of molecules and by free electrons through energy transport (Ghassemi et al., 2021). The equation for conduction is shown below:

$$Q_{cond} = kA \frac{T_1 - T_2}{\Delta x}$$

Where,

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

k = Thermal conductivity, W/m.K

A = Cross sectional area, m^2

ΔT = Temperature difference, K

Δx = Thickness, m

Convective heat transfer is the transfer of heat from one body to another through a moving gas or fluid. The fluid motion in natural convection is caused by natural forces such as buoyancy. Because the fluid velocity associated with natural convection is very low, the heat transfer coefficient experienced in natural convection

24

is also quite low. The fluid motion in forced convection is generated by a source such as a pump or a fan. Because huge quantities of thermal energy can be transferred effectively, it is one of the most common types of heat transfer utilised by engineers. Because it maintains a high temperature differential between the body and the surrounding air or water, convection is a particularly effective mode of heat transmission (Sokolova, 2019). The faster the liquid flow, the greater the convection rate. The equation for convection heat transfer is as shown below:

$$Q_{convection} = hA\Delta T$$

Where,

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

h= Convection heat transfer coefficient, W/m^2.K

A= Surface area for heat transfer, m^2

ΔT= Temperature difference, K

The energy that is released by matter in the form of photons or electromagnetic waves is referred to as radiation heat transfer (Shahidian et al., 2020). When bodies are not in direct physical touch with one other or separated in space, the heat transfer is termed heat radiation. The movement of molecules and atoms in a solid, liquid, or gaseous state causes all substances to release energy in the form of electromagnetic radiation due to vibrational and rotational movement (Ganji et al., 2018). The equation for radiation heat transfer is as shown below by Stefan–Boltzmann law:

$$Q_{rad} = \varepsilon\sigma A_s (T_s{}^4 - T_{surr}{}^4)$$

Where,

Q_{rad} = rate of heat radiation, W
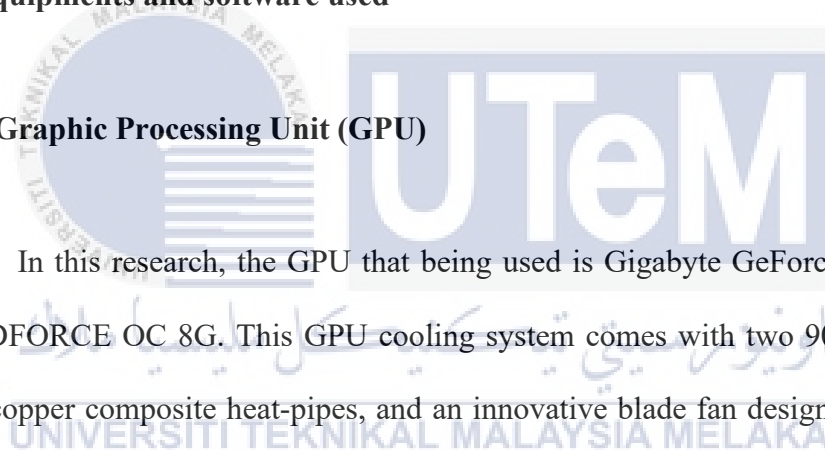
ε = emissivity

σ = Stefan-Boltzmann constant, 5.67 x 10^{-8} W/m^2.K^4

T_s = absolute temperature of surface, K

T_{surr} = absolute temperature of surrounding, K

2.8 Equipments and software used

2.8.1 Graphic Processing Unit (GPU)

In this research, the GPU that being used is Gigabyte GeForce® GTX 1070 WINDFORCE OC 8G. This GPU cooling system comes with two 90mm fans, two pure copper composite heat-pipes, and an innovative blade fan design that improves heat dissipation by allowing heat pipes to directly touch the GPU. The specification is shown in Table 2.1. However, the claimed performance is based on the highest possible theoretical interface values from the various Chipset manufacturers or organisations that developed the interface standard. System setup may have an impact on actual performance. The picture of the GPU is shown in Figure 2.14.

A better performance is possible with the aid of GPUs Overclocking, a useful feature for improving the speed and quality of desktop computing. While it may increase device performance, it will reduce the device's longevity. Choosing a GPU with a high-quality cooling system because it can help boost airflow and temperature

controls. On top of that, the fan speed can be adjusted according to the user's needs. GPU clock speed is a measure of how quick a GPU's cores are. These cores are responsible for creating and displaying visuals. As a result, quicker processing is achieved by increasing the clock speed of the GPU. The clock speed of the GPU is equivalent to the number of processing cycles per second that the GPU can carry out.

Table 2.1: Specification of the GPU

Category	Products specification
Graphics Processing	Gigabyte GeForce GTX 1070
Core Clock	Boost: 1771 MHz / Base: 1582 MHz in OC mode Boost: 1746 MHz / Base: 1556 MHz in Gaming mode
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 x 16
Digital max resolution	7680x4320
Multi-view	4
Dimension	H= 37mm, L= 280mm, W= 131mm
PCB Form	ATX

DirectX	12
OpenGL	4.5
Recommended PSU	500W
Power connectors	8 pin*1

Figure 2.14: Gigabyte GeForce GTX 1070 WINDFORCE OC 8G

2.8.2 Response Surface Methodology (RSM)

Response Surface Methodology (RSM) is an optimization tool that can identify interrelationship between variables (Anwar et al., 2021). Muthukumaran et al., 2017 stated that the effects of independent input process factors on output variables might be shown using RSM. This technology helps researchers to discover better findings by reducing the number of experiments they need to do. It is commonly utilised by various researchers to minimise cost, effort, and time with a small number of trials by varying the input independent variables (Milano et al., 2018). According to Mehmood et al.,

2017, in regard to testing the impact of many independent variables and their interaction on response variables, RSM is an efficient mathematical and statistical methodology.

RSM may be defined as a technique that includes complex computation for optimization process. By using this experimental design, all of the independent variables are integrated and experimental data is then used to derive a set of equations which, theoretically, provides the value of an output. A regression analysis that incorporates controlled variables is used to get the results (Anwar et al., 2021).

The name RSM stands for a group of statistical and mathematical methods used for creating and exploiting empirical models. RSM aims to correlate a response to the amounts of various input variables that affect it by using suitable design and analysis of trials (Eyjolfsson, 2015). Instead of reducing the number of experimental runs, the RSM results are claimed to have a low enough probability of being incorrect to meet requirements for statistical significance (Mitra et al., 2020).

RSM implementation has three stages; the first is designing experiments using Box Behnken and Central Composite Design (CCD). Fewer trials are needed when utilising the Central Composite or Box Behnken design to create response surfaces (Wagner et al.,2021). Second is the statistical and regression analysis to develop response surface models. Finally, model equation parameters and variables are optimised using the response surface models which called optimization process (Wang et al., 2020). In this research, RSM was employed to determine the optimal fan speed, core temperature and the highest possible efficiency (hashing power) of GPU.

In Design-Expert software, there are various response surface method optimization such as Central Composite Design (CCD), Box-Behnken Design (BBD)

and optimal design. For both CCD and BBD, they are developed for estimating a quadratic model. CCD is created from a two-level factorial design, and augmented with centre points and axial points, meanwhile BBD requires only three levels for each factor. Optimal Design is customized for fitting a linear, quadratic or cubic model. When enhanced as recommended by the software, it can generate a large number of levels, but these can be restricted by choosing the discrete factor option.

2.8.3 Central Composite Design (CCD)

Central composite design (CCD), also known as a Box-Wilson central composite design, is often used to construct a second-order polynomial for the response variables without using a whole full factorial design of trials in response surface technique. A polynomial with quadratic terms needs at least three levels of each component in the experimental design to determine its coefficients. Factorial points, central points, and axial points are the three types of points in a CCD system. Points in the factorial cube are n-dimensional cube vertices that originate from the full or fractional factorial design, where the factor levels are coded as –1, +1. The central point is the location in the middle of the design area where everything else is centred. Axial points are placed at a distance from the design centre on the axes of the coordinate system symmetrically with regard to the central point (Sahoo and Barman, 2021).

In the case of a limited number of prepared experiments, the central composite design is a better technique for modelling different technological processes than the "one variable at a time" approach (Savic et al., 2021). An important component of

response surface theory is the CCD model. In order to create a more accurate second-order quadratic model, this optimization approach eliminates the requirement for a three-level factorial experiment which made this method have a great advantage (Bhattacharya, 2021). A lower number of trials is required for the CCD, thus models of the CCD were employed for theoretical study of experimental space (Rakić et al., 2014).

As shown in Figure 2.15 below, estimating the response surface's curvature is made feasible by using the 'star' points outside the experimental domain and the 'centre' points within the experimental domain.
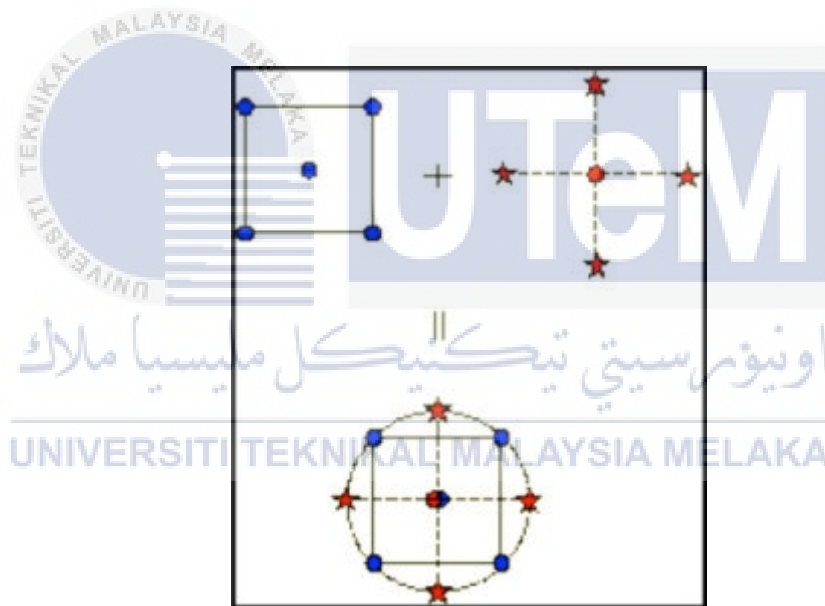
Figure 2.15: Generation in a central composite design of points

Factorial point levels are (+1, -1), whereas star point levels are (+a,-a), where 'a' must be greater than or equal to 1. The star points represent lower and higher extreme values. The parameter value is determined by the computation capabilities and the necessary accuracy of the response surface estimation. The placement of the points determines the quality of the estimate. The setting of value and the number of

trials in the domain's centre have an impact on estimation accuracy (Ait-Amir et al., 2020).

A significant amount of contemporary pharmaceutical research has utilised CCD as a research tool. According to Krishna Veni et al., (2020), CCD model may be used to create and prepared environmentally sensitive Eudragit coated solid lipid nanoparticles. Central composite design was employed in another work by Ye Qingzhuo et al., (2020) to manufacture puerarin nanostructured lipid carriers by predicting response variables and constructing 3D plots. Design-Expert software eventually can explain the CCD model greatly. It helps the researcher to optimize the results of the experiment using the software (Bhattacharya, 2021). Figure 2.16 below described the various procedures involved in CCD.
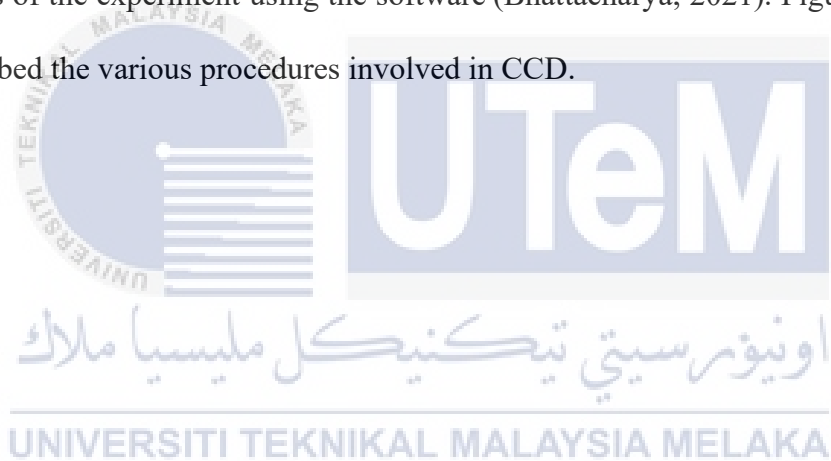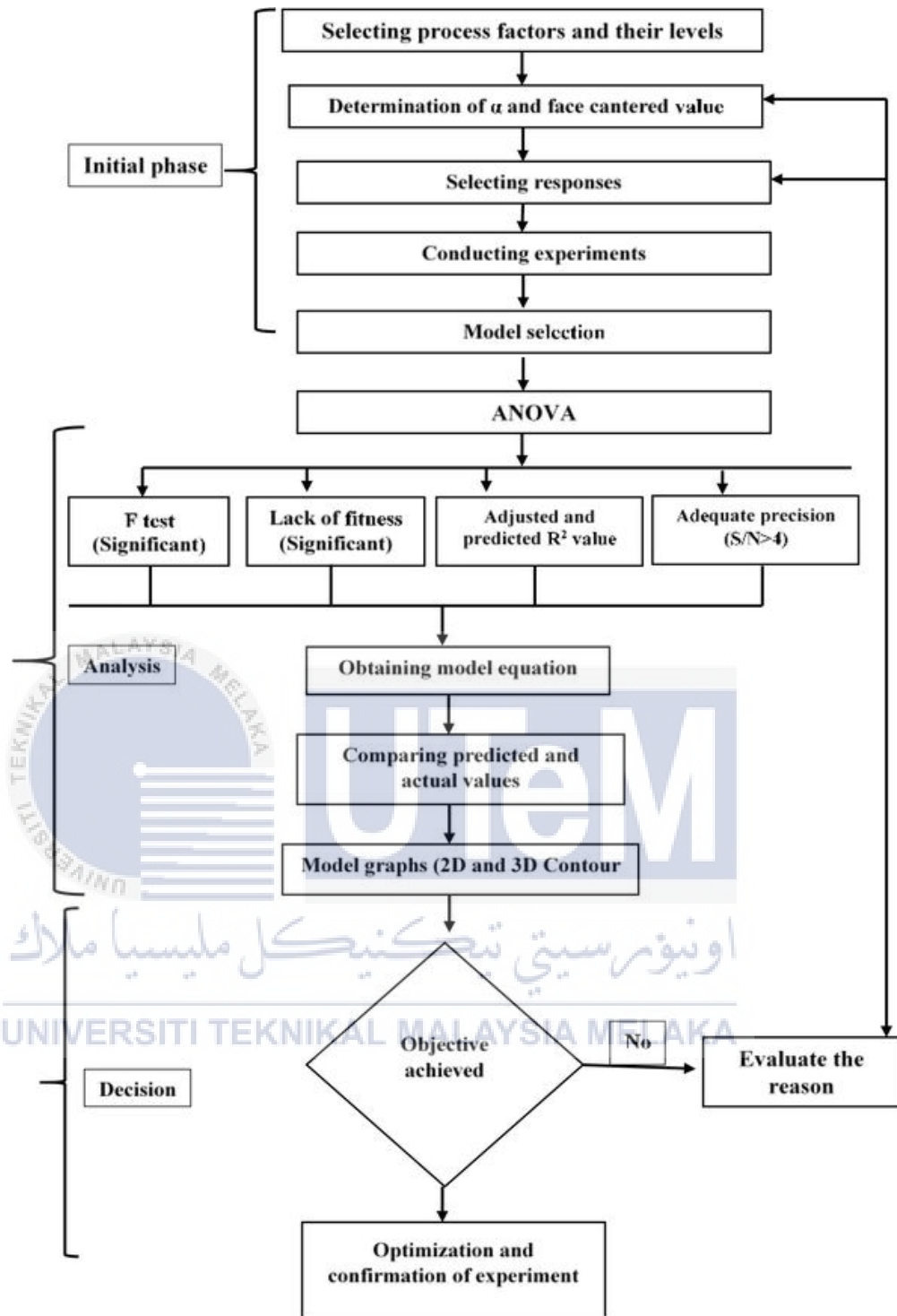
Figure 2.16: CCD flow diagram

2.8.4 Design-Expert Software

Design–Expert is a Stat-Ease Inc. statistical software programme devoted only to the design of experiments (DOE). In terms of testing, Design–Expert provides screening, characterisation, optimization, resilient parameter design, mixture designs, and combination designs. To screen for as many variables as possible, Design–Expert offers test matrices. Analysis of variance (ANOVA) is used to prove the statistical significance of these variables. Analyzing the data using graphic tools reveals any abnormalities and shows the influence of each element on the intended results.

In order to design an optimal experiment on the process, mixture, or combination of variables and components, Design-Expert offers strong tools. Design-Expert makes it simple to assess and to represent the findings most accurately. It provides a variety of graphs to assist consumers in identifying significant impacts and visualising the findings. Figure 2.17 below shows various choices of setting in Design-Expert software such as response surface, mixture and factorial. This software will be used to conduct the experiment and research.

Figure 2.17: Design-expert interface

2.8.5 Overclocking 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 harmin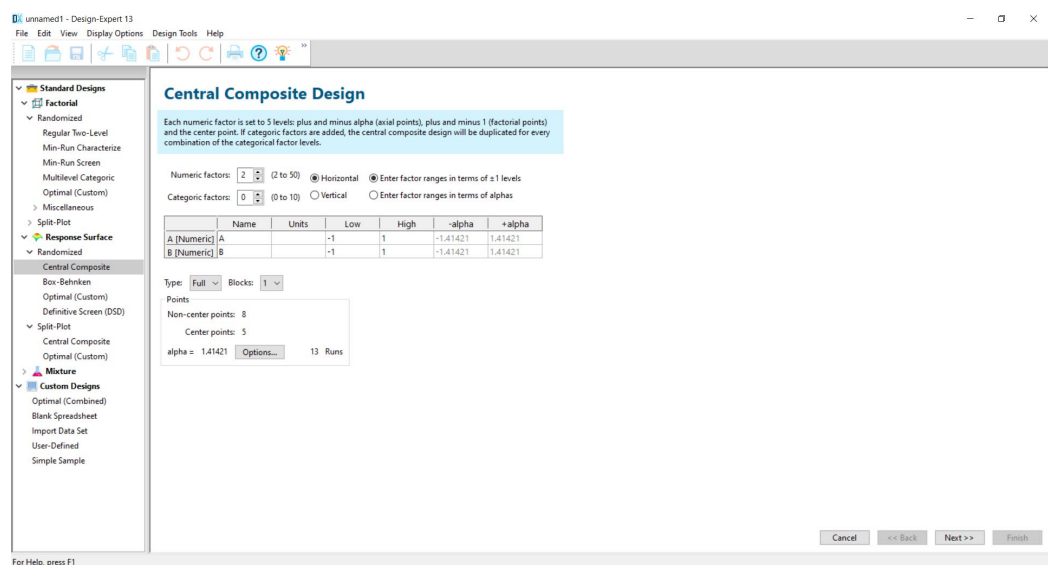g them by mistake. Overclocking software like Afterburner restricts user modifications to numbers they deem safe due to built-in safety features in most graphics' processors these days. Thermal and power consumption monitoring is done by the processors and drivers themselves to guarantee the safety of the product.

Afterburner MSI can fully control the GPU in terms of memory clock, core clock, fan speed, core voltage, temperature limit, power limit and others. Furthermore, this software can be used if the GPU is from other brands, for example, Gigabyte, Asus, and others. It is simple and precise to access the graphics card settings using MSI Afterburner Overclocking 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 utilises OC Scanner, an Nvidia-developed technology, to identify any GPU model available for usage with Afterburner. Once the OC Scanner has identified the video card, it will calculate the best overclocking settings for that particular card. This software will be used to conduct this experiment and research as they can control the GPU. Figure 2.18 below shows the Afterburner software interface.
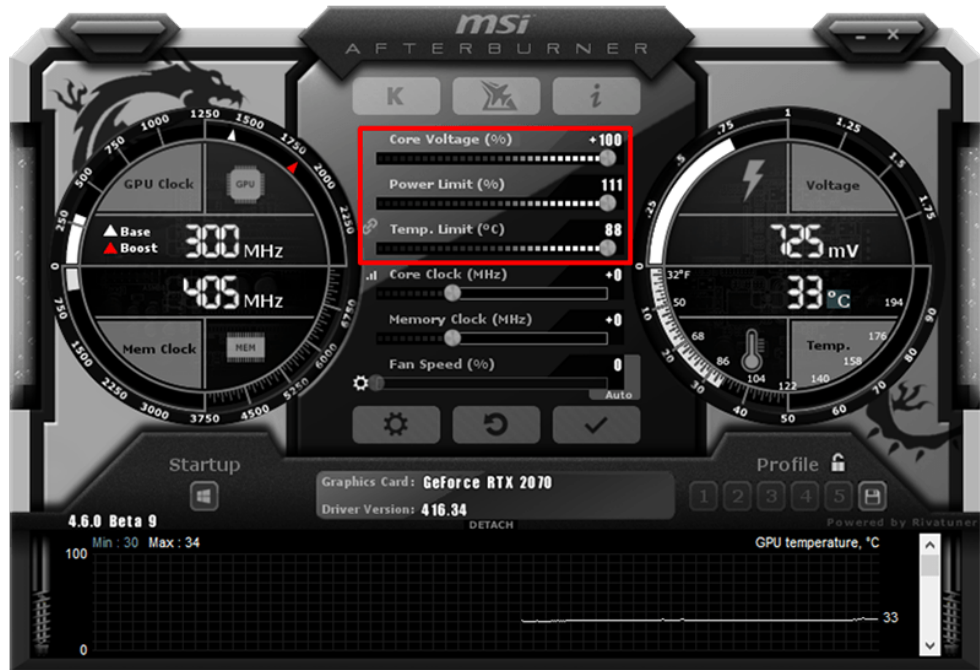
Figure 2.18: Afterburner interface

2.8.6 PhoenixMiner Software

PheonixMiner is mining software that is used to generate new cryptocurrencies and add new blocks to existing blockchains. The mining party receives newly mined cryptocurrency as a reward for their contribution to the blockchain. A computer's graphics processing unit (GPU) is used to assist with cryptocurrency mining. In today's world, the majority of mining is done via a mining pool, which distributes the reward among its members via a network of computers. Figure 2.9 below shows the PhoenixMiner software interface.

PhoenixMiner, a cryptocurrency miner, uses the Dagger Hashimoto (Ethash) algorithm to generate cryptocurrency. This list contains a number of well-known cryptocurrencies, including Ethereum, Ethereum Classic, and MOAC. The programme is compatible with both Nvidia and AMD graphics cards. According to the creators, it

36

is the fastest and most cost-effective Ethereum on Ethash miner on the market right now. However, there is a 0.65% service charge, so the user will only "work" for developers for 35 seconds out of every 90 minutes of GPU mining. The software is used to report the GPU's hash rate (MH/s) value in this research. The term hash rate refers to the total amount of computational power used to process transactions and mine new coins.

Figure 2.19: PhoenixMiner software

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter focuses on the research methodology employed to accomplish the project's objectives. For example, this project is about finding the optimal fan speed, core temperature and the highest efficiency (hashing power) at certain core and memory clocks of the Graphic Processing Unit (GPU). A flowchart will be used to show the steps in the process so that the goals can be met. The methodology, software, procedures, and processes used in the study will all be discussed in detail in this section. Keeping track of all plans and actions is the goal of this chapter, which aims to avoid any mistakes that may be made due to a lack of planning. A method will be selected with limitations to implement it, and a second and third plan will be considered as well. The findings from the literature review and primary research have been used to build this comprehensive analysis.

3.2 Flowchart
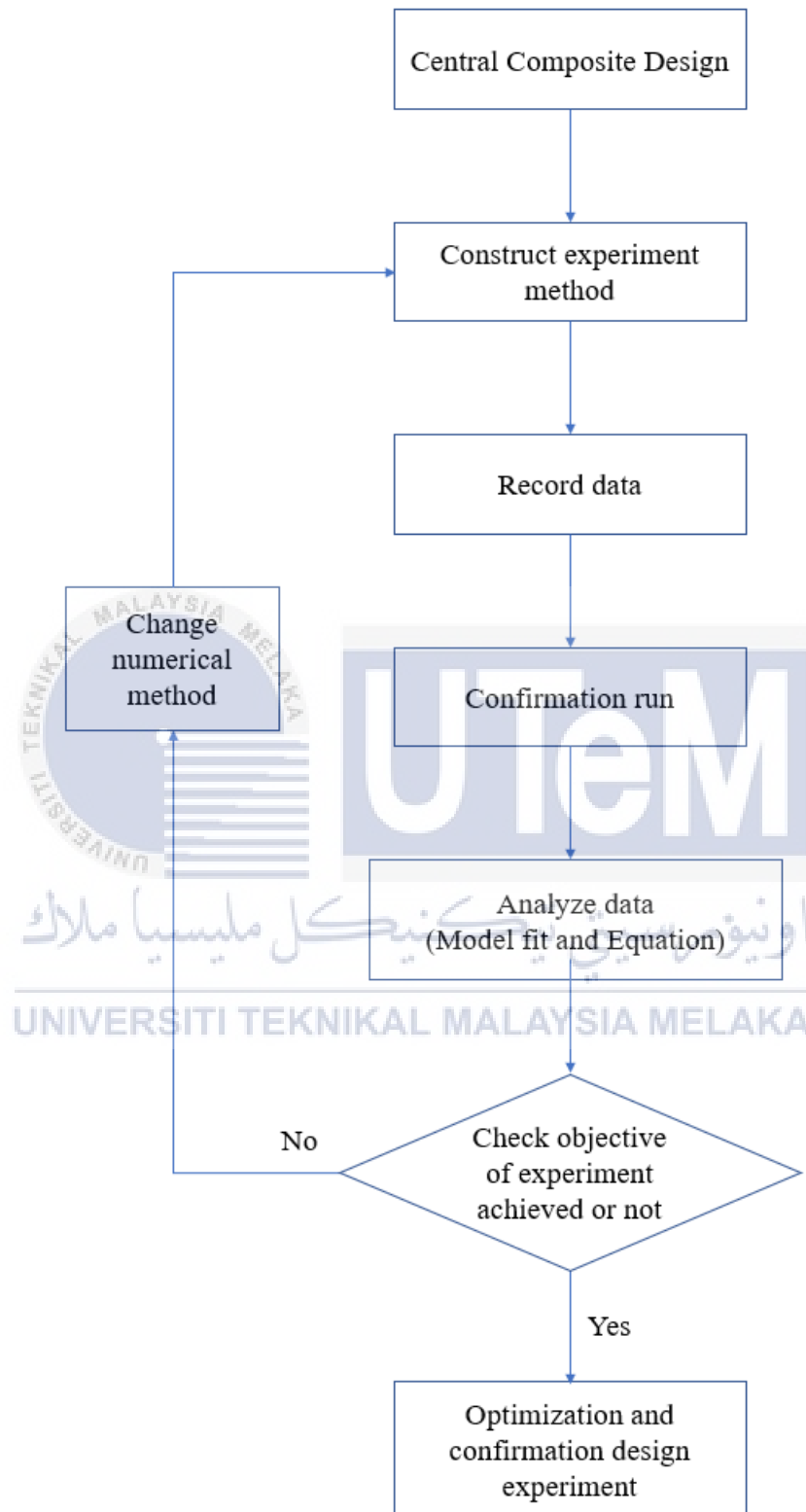
Figure 3.1: Flowchart of methodology

3.3 Schematic Diagram
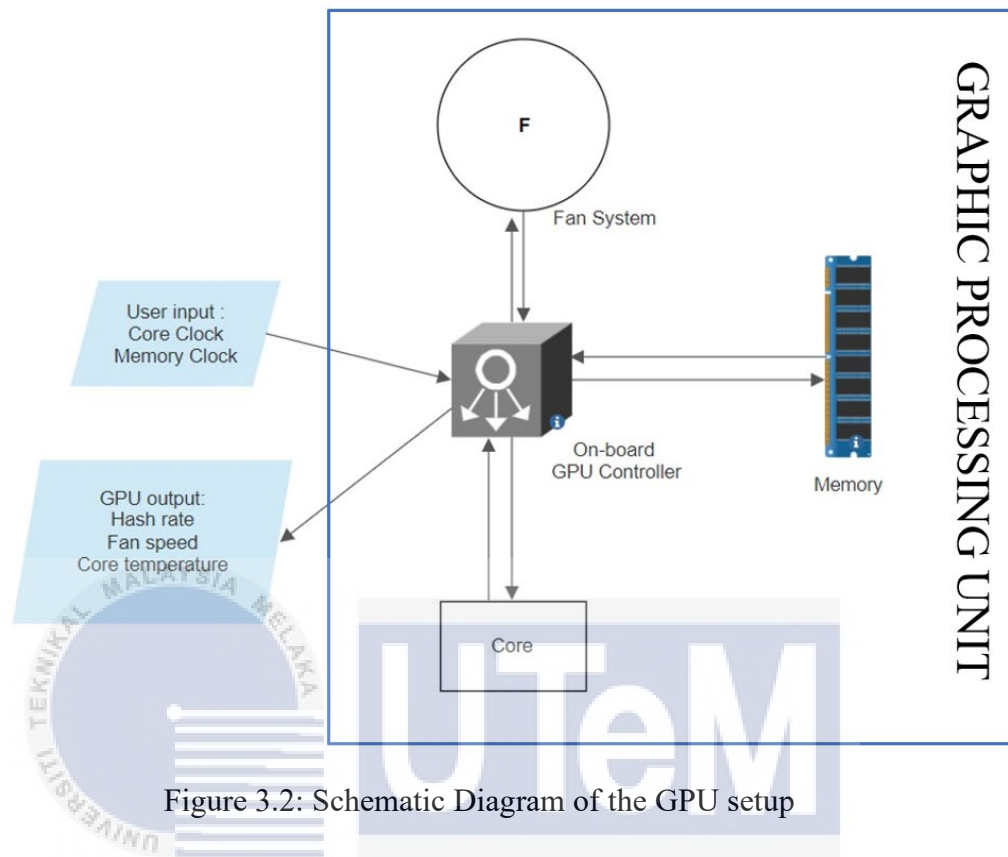
Figure 3.2: Schematic Diagram of the GPU setup

3.3 Design of experiment

Response Surface Methodology (RSM) is a statistical and mathematical approach for modelling and analysing response or output variables that are impacted by several input parameters (Rai et al., 2016). The Central Composite Design (CCD) has a setting where the numeric factor can be filled with 2 to 50 variables, while the Box-Behnken Design (BBD) can be filled with 3 to 21 variables. Numeric factors represent the manipulated variables of the experiment. Since this research has only 2 numeric factors, CCD was chosen to be used to optimise the data of the experiments. RSM based on the CCD in Design-Expert version 13.0 software was used to design

the experiments to analyse the relationship between clocking (core clock, memory clock) and GPU response (fan speed, core temperature, hash rate). In this research, the GPU that being used is Dual-Fan GeForce GTX 1070 WINDFORCE OC 8G. This GPU cooling system comes with two 90mm fans.

Analysis of experimental data was done statistically with the help of Design-Expert software. Various statistical parameters (such as lack-of-fit, predicted and adjusted multiple correlation coefficients, and coefficient of variation) of various polynomial models were evaluated to find the most accurate one. Analysing variance and computing p-value at probabilities of higher than 0.05 revealed that it is significant. Design-Expert Software was used to create response plots to get the better understanding regarding the relationship between emulsifying conditions and response variables. When optimising variables and responses, the CCD technique helps determine the number of tests that should be run. The minimum, intermediate, and maximum values of each variable are labelled as -α, −1, 0, +1 and +α as shown in the Table 3.1 below.

Table 3.1: Independent variables and their corresponding levels for GPU clocking

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

Table 3.1 above shows that the core clock setting is between -200MHz and 200MHz, meanwhile the core clock setting is between 300MHz and 900MHz. since the optimum values can be obtained from the Minerstat website, which are 100 MHz and 300 MHz for GTX 1070. The values obtained from the Minerstat website will be used as a reference to conduct this research. Next, the factorial +1 and -1 design points are represented by the square's four corners as shown in Figure 3.3 below. The parameters +1 and -1 are used to specify the boundaries of the study region, where it is thought that the optimal situation exists. Axial points, on the other hand, will often fall outside of this range. The axial +alpha and -alpha design points are represented by four-star points. For even the most severe axial runs, there are alphas to make sure everything still works. The focus must be on something that can be operated.
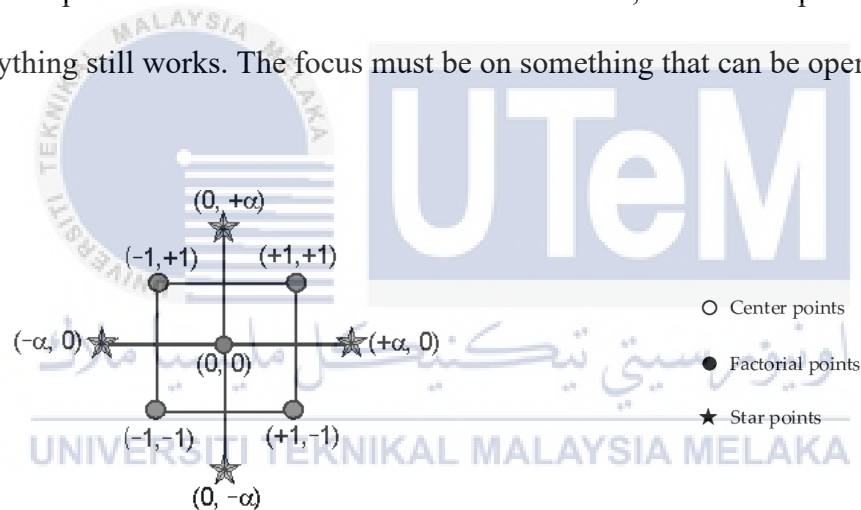
Figure 3.3: Classic CCD for 2 factors square's four corners

3.4 Experimental Setup (Design-Expert and GPU)

Using Afterburner software, core clock and memory clock will be controlled according to the values set by Design-Expert software. The GPU response such as fan speed (rpm), core temperature (°C) can be recorded by using Afterburner software while hash rate (MH/s) can be observed using PhoenixMiner software. The software

is connected to the GPU to be investigated. The schematic diagram of the GPU setup for this research is shown in the Figure 3.2 above. The GPU data response will be included in the Design-Expert for the analysis and optimization process to get the optimal fan speed and the highest efficiency (hashing power) at certain core clock and memory clock. The optimum core clock and memory clock value can be obtained from the Minerstat website, which are 100 MHz and 300 MHz for GTX 1070. The values obtained from the Minerstat website will be used as a reference to conduct this research. So, to conduct this research, the range of core clock and memory clock is being set to (-200 to 200 MHz) and (300 to 900MHz). Figure 3.4 below shows the options available to set the numeric factors. The Design-Experts allow the user to choose whether to enter the factor ranges in terms of (+1, -1) levels or alphas. In this research, the factor ranges are filled in terms of alphas because the maximum and minimum are already being set for the core and memory clock.

Figure 3.4: Design-Expert layout for CCD

13 runs, including four axial points, four fractional factorial points and five central points were randomly performed according to Central Composite Design (CCD), which are shown in Figure 3.5 below. The experiment was run continuously and repeated several times to get a constant response (GPU output).

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: The design layout screen

3.3 Statistical Analysis and Optimization of Clocking and GPU Response

The regression study of GPU response (fan speed, core temperature and hash rate) was carried out using Design-Expert software. The software will do an analysis of variance (ANOVA) on the models that have been built. In order to evaluate the degree of disparity among a set of data, an ANOVA is commonly utilized. It demonstrates the model's significance by evaluating the model's statistical fit. It is then possible to use the optimization constraint option to establish the numerical setting to achieve the optimal fan speed, core temperature and the highest possible efficiency (hashing power) at certain core and memory clock. CCD is a good way to fit a quadratic surface model consecutively. A lot of the design options that does not fit are taken out of CCD's evaluation process.
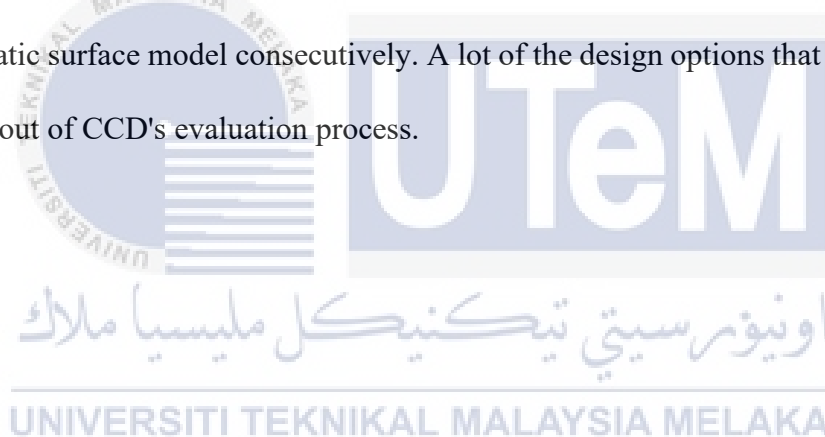
CHAPTER 4

RESULTS AND DISCUSSION

4.1 Results

The experimental sequence, based on the randomized Central Composite Design (CCD) of Response Surface Methodology (RSM), was random to eliminate the systematic errors caused by extraneous factors (Raymond H. Myers and Anderson-Cook, 2016). The memory clock and core clock are being set using the Afterburner software as suggested in Figure 3.5 Design of Experiment (DOE). The results of the response which are fan speed (rpm), core temperature(°C) and hash rate (MH/s) were collected based on the actual design of experiments and presented in Table 4.1 to get the optimal fan speed, core temperature and the highest efficiency at certain core clock and memory clock.

Table 4.1: Experimental design and response values 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	1935	58.4	21.8
2	Factorial	141.421	387.868	1915	58.1	24.7
3	Factorial	-141.421	812.132	1914	58.4	21.6
4	Factorial	141.421	812.132	1900	57.9	24.5
5	Axial	-200	600	1917	58.5	21.1
6	Axial	200	600	1892	57.9	25.2
7	Axial	0	300	1946	58.4	23.3
8	Axial	0	900	1920	58.1	23
9	Center	0	600	1915	58	23.2
10	Center	0	600	1921	58.2	23.2
11	Center	0	600	1916	58.1	23.2
12	Center	0	600	1917	58.1	23.2
13	Center	0	600	1915	58.1	23.2

4.2 CCD and ANOVA (Fitting Model)

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	2085.46	5	417.09	72.58	< 0.0001	significant
A-Core Clock	601.27	1	601.27	104.62	< 0.0001	-
B-Memory Clock	661.93	1	661.93	115.18	< 0.0001	-
AB	9.00	1	9.00	1.57	0.2510	-
A²	325.23	1	325.23	56.59	0.0001	-
B²	382.23	1	382.23	66.51	< 0.0001	-
Residual	40.23	7	5.75	-	-	-
Lack of Fit	15.43	3	5.14	0.8295	0.5429	not significant
Pure Error	24.80	4	6.20	-	-	-
Cor Total	2125.69	12	-	-	-	-

Std. Dev.= 2.40, Mean= 1917.15, C.V. %= 0.1250, R²= 0.9811, Adjusted R²= 0.9676, Predicted R²= 0.9302, Adeq Precision= 32.9260

In this project the validations for the result are based on Lack of Fit (LOF), R², Adjusted R², Predicted R² and Coefficient of Variation (C.V.) %. The LOF for fan speed responses is not significant, indicating that the model is adequate and implying that the built models were mathematically well fit. P-value of 0.5429 means there is a 54.29% chance that a Lack of Fit F-value this large could occur due to noise. Besides, except for the lack of fit, a small p-value indicated a close relationship between calculated responses and independent inputs. As seen in the table above, the model F-

value of 72.58 indicated that the models are statistically significant when the p-value is less than 0.05. This is critical for confirming the significance of input variables in relation to responses, since components with a larger p-value (more than 0.10) signify less significance, and vice versa.

The correlation between R^2 and adjusted R^2 is the second factor in verifying the constructed models. Fan speed values for R^2 and adjusted R^2 are greater than 0.9 indicate that 90% of the variations in the responses (results) were due to the independent input factors (memory clock and core clock). In other words, the models generated are highly reliant on the input variables, and all factors in this experiment are significant. The statistical analysis findings (ANOVA) indicated that the experimental data could be well represented by a quadratic polynomial model, with R^2 coefficient of determination values of 0.9811 for fan speed. The Predicted R^2 of 0.9302 is reasonably close to the Adjusted R^2 of 0.9676, and the difference is less than 0.2, indicating that the model fits the data and can be used to interpolate successfully. Thirdly, cross-validation of the models' adequate precision values of 32.9260 and coefficient of variation (CV) % of 0.1250 proved the models' flexibility and dependability.

Finally, visual diagnostics were performed by comparing the experimentally determined values against the predicted values derived by the programme using the modelled equations. As seen in Figure 4.1 below, both projected and actual fan speed values are near the centerline, with no apparent outliers. The plot as shown in Figure 4.1 indicates that the experimental and predicted values are closer together, indicating that the CCD model is capable of developing a good correlation between the input

components and the expected response values (fan speed). The fan speed models were

validated and may now be utilised for the experimental design.
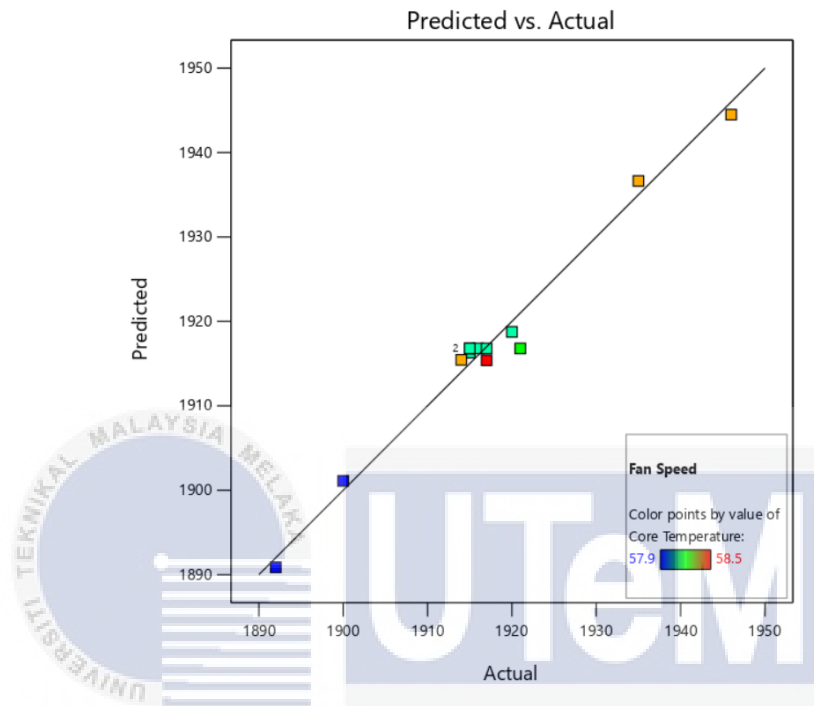
Figure 4.1: Graph of predicted (calculated) versus actual (experimental) for fan speed

4.2.2 Core Temperature Response

Table 4.3: ANOVA for core temperature response

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	0.4399	5	0.0880	22.13	0.0004	significant
A-Core Clock	0.3397	1	0.3397	85.44	< 0.0001	-
B-Memory Clock	0.0487	1	0.0487	12.25	0.0100	-
AB	0.0100	1	0.0100	2.52	0.1568	-
A²	0.0133	1	0.0133	3.35	0.1099	-
B²	0.0329	1	0.0329	8.27	0.0238	-
Residual	0.0278	7	0.0040			-
Lack of Fit	0.0078	3	0.0026	0.5221	0.6898	not significant
Pure Error	0.0200	4	0.0050	-	-	-
Cor Total	0.4677	12	-	-	-	-

Std. Dev.= 0.0631, Mean= 58.17, C.V. %= 0.1084, R²= 0.9405, Adjusted R²= 0.8980, Predicted R²= 0.8141, Adeq Precision= 14.0187

Lack of Fit (LOF), R^2, Adjusted R^2, Predicted R^2 and Coefficient of Variation (C.V.) %. are used to validate the results in this project. LOF indicates that the model is adequate for the core temperature responses, and the models were mathematically well-fit since the responses are not significant. A low p-value suggested a strong correlation between calculated responses and independent inputs, except for the lack of fit. In addition, Table 4.3 above shows that the models are statistically significant when the p-value is less than 0.05. This is important for confirming the importance of input variables in relationship to responses because components with a higher p-value (greater than 0.10) indicate less significance and vice versa.

51

The second step in the model-verification process is to look at the correlation between R^2 and adjusted R^2. A core temperature values for R^2 and adjusted R^2 greater than 0.89 indicate that 89 percent of the responses (results) were influenced by independent input sources (memory and core clock). As a result of this experiment, all the input variables have a major impact on any models that are produced. For core temperature, the statistical analysis results (ANOVA) showed that a quadratic polynomial model could adequately reflect the experimental data, with an R^2 coefficients of determination of 0.9405. The Predicted R^2 of 0.8141 is close to the Adjusted R^2 of 0.8980, with a difference of less than 0.2, indicating that the model matches the data and can be used to successfully interpolate. Thirdly, cross-validation of the models' adequate precision values of 14.0187 and coefficient of variation (CV) % of 0.1084 demonstrated the models' flexibility and reliability.

Finally, visual diagnostics were carried out by comparing the experimentally measured values with the predicted values derived by the programme using the modelled equations. There are no obvious outliers in the predicted and actual core temperature values, as shown in the Figure 4.2 below. According to the plot shown in Figure 4.2 below, the CCD model's ability to generate a good correlation between the input components and the expected response values is demonstrated (core temperature). For the experiment's design, the core temperature models have been proven to be accurate and can be used.
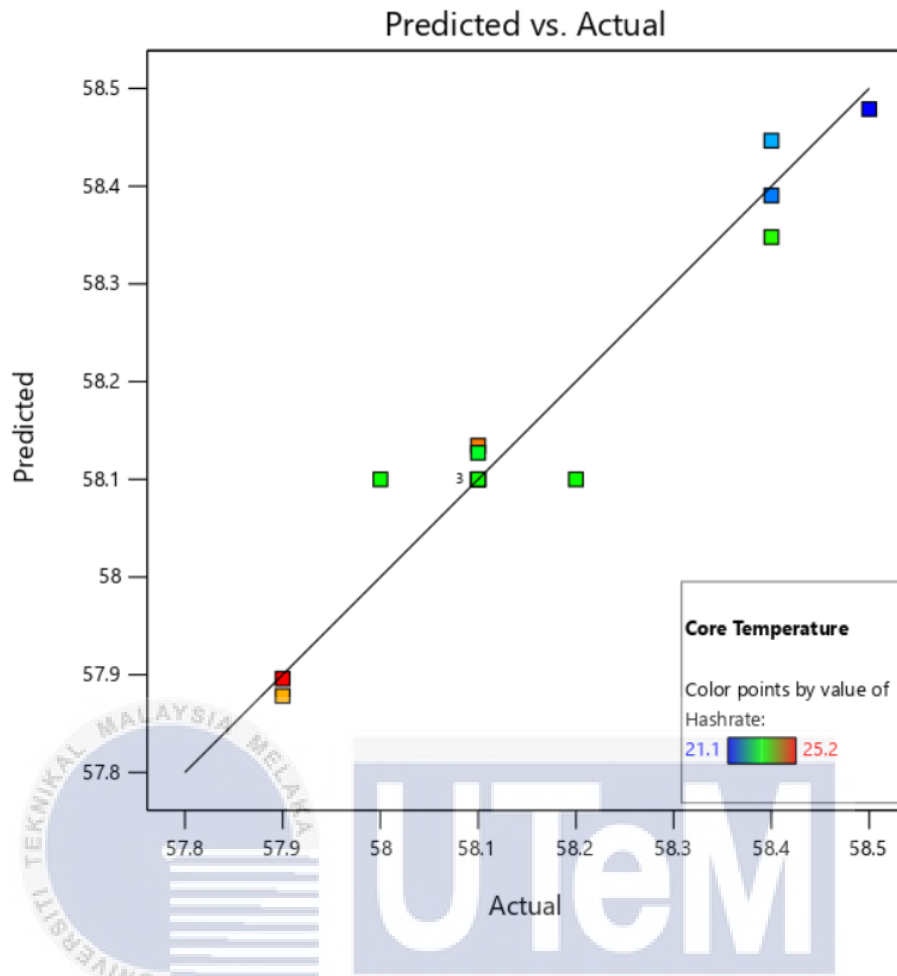
Figure 4.2: Graph of predicted (calculated) versus actual (experimental) for core temperature

4.2.3 Hash Rate Response

Table 4.4: ANOVA for core hash rate response

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	16.91	5	3.38	3.200E+05	< 0.0001	significant
A-Core Clock	16.81	1	16.81	1.591E+06	< 0.0001	-
B-Memory Clock	0.0849	1	0.0849	8037.40	< 0.0001	-
AB	0.0000	1	0.0000	0.0000	1.0000	-
A²	0.0043	1	0.0043	411.48	< 0.0001	-
B²	0.0043	1	0.0043	411.48	< 0.0001	-
Residual	0.0001	7	0.0000	-	-	-
Lack of Fit	0.0001	3	0.0000	-	-	-
Pure Error	0.0000	4	0.0000	-	-	-
Cor Total	16.91	12	-	-	-	-

Std. Dev.= 0.0033, Mean= 23.17, C.V. %= 0.0140, R^2= 1.0000, Adjusted R^2= 1.0000, Predicted R^2= 1.0000, Adeq Precision= 1856.8674

R^2, Adjusted R^2, Predicted R^2 and Coefficient of Variation (C.V%) are used to validate the results in this project. Table 4.4 above contains the findings of the suggested quadratic model. In ANOVA, the significance of the variables was considered only when the p-values were less than 0.05. The Model p-value of less than

0.0001 demonstrates the model's significance. This suggests that there is only a 0.01% possibility that a significant "Model F- Value" may occur due to noise. Factors A, B, and all of their interactions are all regarded statistically significant based on the ANOVA Table 4.4 above.

The correlation between R^2 and adjusted R^2 is the next aspect to consider when evaluating the constructed models. R^2 and adjusted R^2 values for hash rate is 1 indicate that the independent input variables (memory clock and core clock) were responsible for 100% of the changes within the responses (results). To put it another way, the built models are heavily dependent on the input variables, and all factors employed in this experiment are significant. Although there is a little difference between the adjusted and predicted R^2 values (less than 0.2), this model fits the data well and may be relied upon for interpolation.

A diagnostic plot comparing actual values to predicted values might be used to validate results because the programme does not provide the p-value for Lack of Fit. According to the plot shown in Figure 4.3 below, the CCD model's ability to generate a good correlation between the input components and the expected response values is demonstrated (hash rate). Since the graph reveals no outliers, the validity of the hash rate models has been confirmed and they may now be employed in the experimental design.
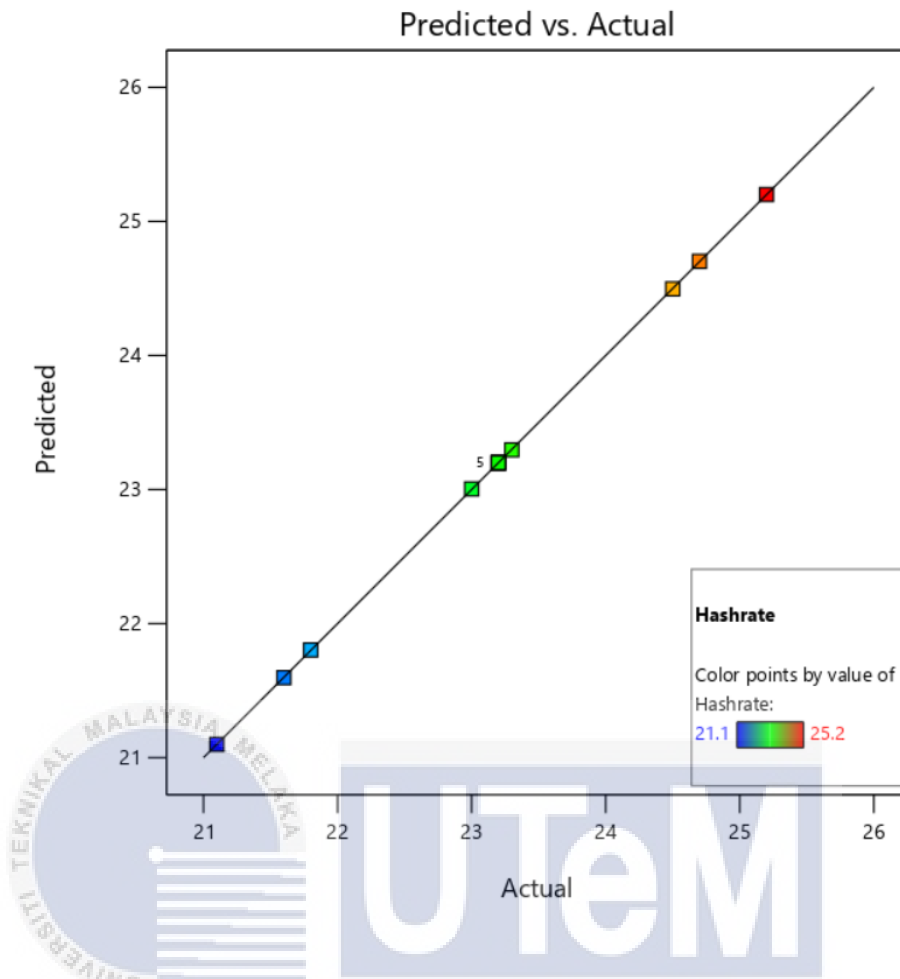
Figure 4.3: Graph of predicted (calculated) versus actual (experimental) for hash rate

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

4.3.1 Fan speed
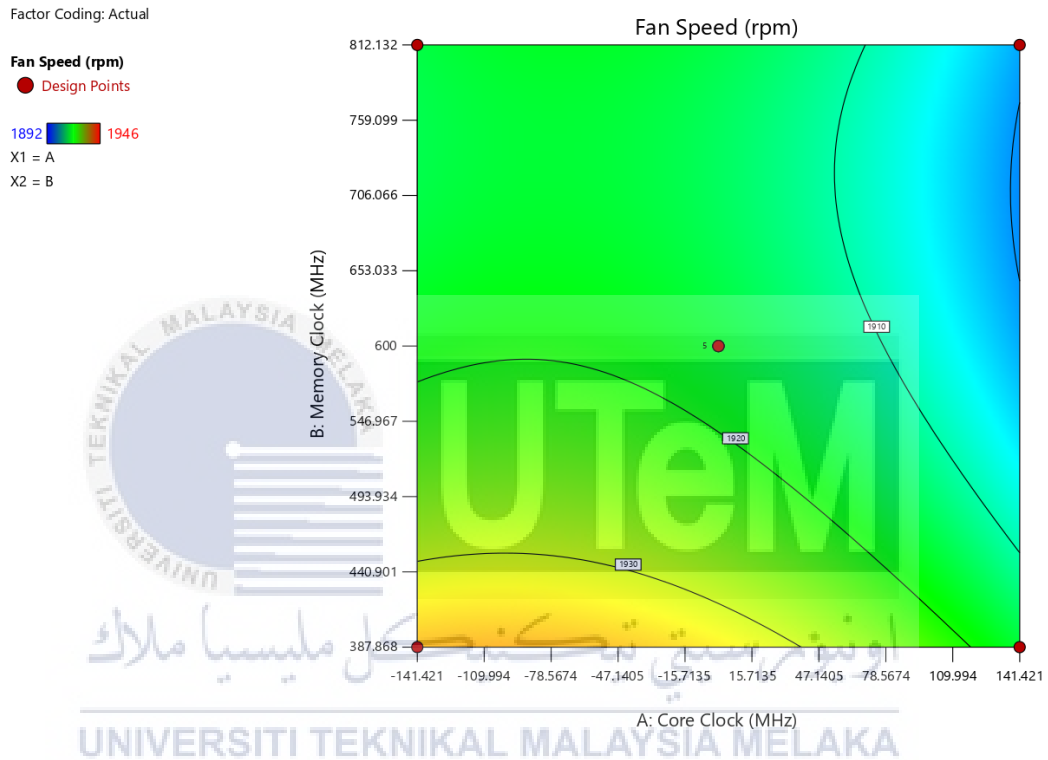
Figure 4.4: Contour plot for the combined effect of core clock (A), memory clock (B) and fan speed

Fan speed = 1916.8 + -8.66942 * A + -9.09619 * B + 1.5 * AB + -6.8375 * A^2 + 7.4125 * B^2 – (Eq.1)

Figure 4.4 above represents the relationship between fan speed and the frequency of the memory and core clocks. With an increase in core clock to a maximum of 141.421MHz and an increase in memory clock to a maximum of 812.132 MHz, the fan speed decreased. The red colour indicates high fan speeds region (higher

than 1930 rpm), whereas the blue colour shows the region with low fan speeds (lower than 1910 rpm). Core clock values between (-141.421 to 15.7135 MHz) and memory clock (387.868 to 440.091 MHz) are yellow and red, indicating a high fan speed in those regions. Additionally, the core clock values (109.994 to 141.421 MHz) and memory clock values (493.934 to 812.132 MHz) in the graph above are in blue colour, indicating that these clocking levels result in a low fan speed when compared to the values of the other independent variables.

Due to the fact that a small p-value (less than 0.05) implied a close relationship between computed responses and independent inputs, except for the case of Lack of Fit. Fan speed is largely affected by A (the core clock), B (the memory clock), and A^2 and B^2 in this scenario, as indicated in Table 4.2. Due to the negative sign in front of the components in Eq. (1), it is evident that A component may lower fan speed with a multiplier of -8.66942, B component with a multiplication of -9.09619, and A^2 component with a multiplier of -6.8375. Based on the equation, it can be concluded that the largest multiplier value in component B, which is the memory clock, is the primary cause of decreased fan speed. Additionally, the programme validates this component since the p-value for the memory clock is less than 0.000.

4.3.2 Core Temperature
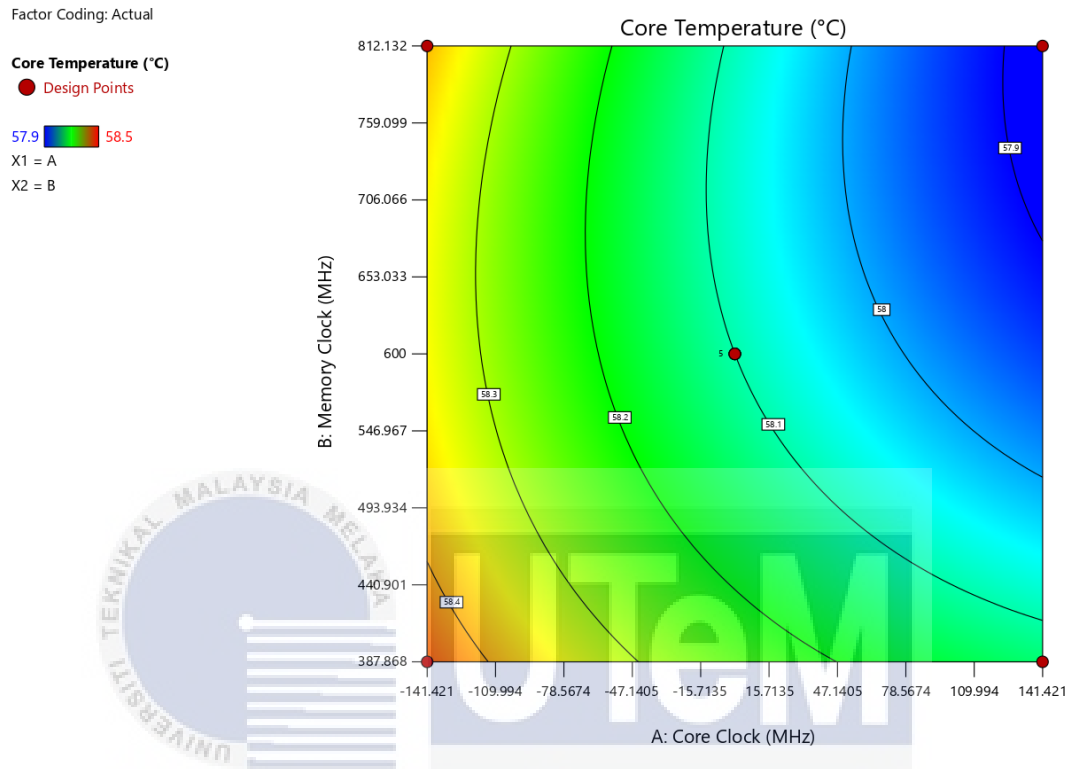
Figure 4.5: Contour plot for the combined effect of core clock (A), memory clock (B) and core temperature

Core temperature = 58.1 + -0.206066 * A + -0.078033 * B + -0.05 * AB + 0.04375 * A^2 + 0.06875 * B^2 - (Eq.2)

Figure 4.5 above depicts the effect of the memory clock and core clock on core temperature. It can be seen that increasing the core clock up to a maximum limit of 141.421 MHz and increasing the memory clock up to a maximum limit of 812.132 MHz causes the core temperature to decrease. The colour red represents areas with a high core temperature (higher than 58.4 °C), whereas blue represents areas with a low core temperature (lower than 58.1°C). The areas are yellow and red in colour for core

clock values of (-141.421 to -78.5674 MHz) with memory clock values of (387.868 to 546.967 MHz), indicating that the core temperature of the GPU is relatively high in those areas. Furthermore, the core clock values (-141.421 to -109.994 MHz) and memory clock values (759.099 to 812.132 MHz) appear in yellow and red, indicating that the core temperature is high at those clocking values. The graph shows blue for core clock values of (15.71 to 141.421MHz) and memory clock settings of 493.934MHz and higher, indicating that the GPU's core temperature is relatively low at those settings.

Except for Lack of Fit, a small P-value (less than 0.05) indicates a close relationship between calculated responses and independent inputs. As indicated in Table 4.3, the most important parameters impacting core temperature are A (Core clock) and B (Memory clock). Due to the negative sign in front of the components in Eq. (2), it is evident that core temperature can be reduced by A component with a multiplier of -0.206066, B component with a multiplier of -0.078033, and AB component with a multiplier of -0.05. Based on the equation, the A component, which is the core clock, is the most important contributor to the GPU's core temperature because it has the largest multiplier value. The software confirms this component because the p-value for the core clock is less than 0.0001

4.3.3 Hash rate
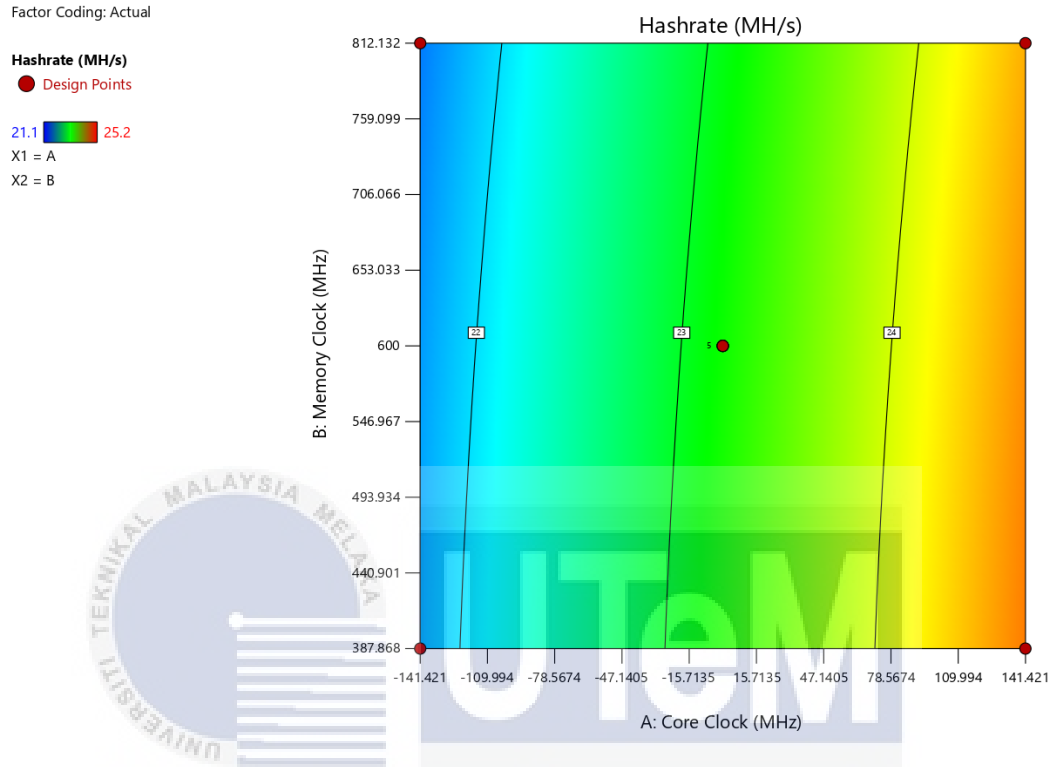
Figure 4.6: Contour plot for the combined effect of core clock (A), memory clock (B) and hash rate

Hash rate = 23.2 + 1.44978 * A + -0.103033 * B + 7.05963e-15 * AB + -0.025 * A^2 + -0.025 * B^2 - (Eq.3)

Figure 4.6 above describes the impact of the memory clock and core clock on hash rate. As the core clock increases, the hash rate increases until it reaches 141.421MHz. The colour red denotes places with a high hash rate (higher than 24 MH/s), whereas the colour blue denotes areas with a low hash rate (lower than 22 MH/s). The locations are yellow and red in colour for core clock values of (78.5674 to 141.421 MHz) with a memory clock of (387.868 to 812.312 MHz), indicating that the GPU's hash rate is relatively high in those areas. In addition, the blue core clock values

(-141.421 to -109.994 MHz) and memory clock values (387.868 to 812.132 MHz) in the graph above indicate that such clocking settings have a low hash rate value when compared to the other independent variables.

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 4.4, the factors that have the greatest impact on core temperature are A (Core clock) and B (Memory clock), A^2 and B^2. Due to the positive sign in front of the components in Eq. (3), it is clear that hash rate can be increased by A component with a multiplier of 1.44978 and AB component with a multiplier of 7.05963e-15. As a result of the equation, it can be concluded that the A component, which is the core clock, is the most important contributor to hash rate values because it has the highest multiplier value. The software also confirms this component because the p-value for the core clock is less than 0.0001.

4.4 Data optimization

Table 4.5: Numerical optimization setting

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	1892	1946	1	1	3
Core Temperature	minimize	57.9	58.5	1	1	3
Hash rate	maximize	21.1	25.2	1	1	3
Hot Spot Temperature	none	68.9	70	1	1	3
Power Consumption	none	119.4	120.7	1	1	3

After all the analysis that has been done, numerical optimization was executed by desirability function using Design-Expert Software. Twelve different solutions were found which contain different levels of independent variables. The goals selected for the optimization of the GPU settings were core clock and memory clock are in range since the value for core clock and memory clock that have been obtained from the Minerstat website are in the optimum range. Therefore, the core clock is set to -141.421 MHz for lower limit and 141.421 MHz for upper limit. Meanwhile, for the memory clock, the lower limit is 387.868 and the upper limit is 812.132. In addition, the goals for the response variables are with the minimum fan speed, minimum core temperature and the maximum hash rate. Those settings gave out 12 possible solutions with only one selected setting, which is optimal for the GPU to operate. All twelve

suggested solutions were selected for experiment with the predicted and experimental

fan speed, core temperature and hash rate are listed in Table 4.6 below.

Table 4.6: Numerical optimization solution

Number	Core Clock	Memory Clock	Fan speed	Core Temperature	Hash rate	Hot Spot Temperature	Power Consumption	Desirability	
1	141.421	685.185	1899.438	57.897	24.579	68.935	120.138	0.901	
2	141.421	683.252	1899.454	57.898	24.580	68.937	120.138	0.901	
3	141.421	687.578	1899.420	57.897	24.578	68.933	120.138	0.901	
4	141.421	681.330	1899.470	57.899	24.582	68.938	120.138	0.901	
5	141.421	689.807	1899.406	57.896	24.577	68.932	120.138	0.901	
6	141.421	678.599	1899.496	57.900	24.583	68.940	120.138	0.901	
7	141.420	692.333	1899.391	57.895	24.575	68.930	120.138	0.901	
8	141.421	708.314	1899.347	57.890	24.566	68.918	120.138	0.901	
9	141.421	712.428	1899.349	57.889	24.563	68.915	120.138	0.900	
10	141.421	761.395	1899.804	57.880	24.532	68.878	120.138	0.895	
11	**141.421**	**651.996**	**1899.877**	**57.910**	**24.598**	**68.960**	**120.138**	**0.895**	**Selected**
12	141.421	772.555	1900.019	57.879	24.524	68.870	120.138	0.893	

The experiment will be run again using 1 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 condition. This difference is due to the limitation of technology, since the sensor is not too sensitive with small changes in GPU response and the software (PhoenixMiner and Afterburner) have lower precision of 0.1 only. These small deviations can be credited to a reliable testing device as well as the proposed mathematical model which able to predict the value of GPU response. Numerical optimization based on CCD can predict an almost accurate the desired lowest fan speed, lowest core temperature and highest possible hash rate at constant power consumption without need to find manually at each core clock and memory clock. The GPU response can be predicted using any core clock, and memory clock settings with the equation obtained. Based on the validation process, it can be concluded that the accuracy of CCD is experimentally proven. In conclusion, choosing the CCD method for this experiment will give the best result.

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. Result shows that a non-harmonious core and memory clock setting effects the overall performance adversely, leads 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 13 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 2 percent deviation.

5.1 Recommendation

Based on 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 dual fans and only on the Gigabyte GeForce GTX 1070 WINDFORCE OC 8G 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

(Vargas-Vazquez et al., WATERBLOCK MODELLING for GPU LIQUID COOLING) Siricharoenpanich, A., Wiriyasart, S., & Naphon, P. (2021). Study on the thermal dissipation performance of GPU cooling system with nanofluid as coolant. Case Studies in Thermal Engineering, 25(February), 100904. https://doi.org/10.1016/j.csite.2021.100904

"Crypto Mining Monitor and Management Software." Minerstat, minerstat.com/. Accessed 9 Jan. 2022.

"Design-Expert | Stat-Ease." Www.statease.com, www.statease.com/software/design-expert/. Accessed 26 Sept. 2021.

"Response Surface Methodology by Raymond H. Myers | Perlego." Www.perlego.com, 2016, www.perlego.com/book/996444/response-surface-methodology-process-and-product-optimization-using-designed-experiments-pdf. Accessed 9 Nov. 2021.

"Thermal Paste Application Techniques." Puget Systems, 2012, www.pugetsystems.com/labs/articles/Thermal-Paste-Application-Techniques-170/. Accessed 1 April 2021.

Ahmadian-Elmi, M., Mohammadifar, M., Rasouli, E., & Hajmohammadi, M. R. (2021). Optimal design and placement of heat sink elements attached on a cylindrical heat-generating body for maximum cooling performance. Thermochimica Acta, 700, 178941. https://doi.org/10.1016/j.tca.2021.178941

Ahmed, H. E., Salman, B. H., Kherbeet, A. S., & Ahmed, M. I. (2018). Optimization of thermal design of heat sinks: A review. International Journal of Heat and Mass Transfer, 118, 129–153. https://doi.org/10.1016/j.ijheatmasstransfer.2017.10.099

Ait-Amir, Bouzid, et al. "6 - Meta-Model Development." ScienceDirect, ISTE, 1 Jan. 2020, www.sciencedirect.com/science/article/pii/B9781785481901500062. Accessed 29 Sept. 2021.

Akshat Verma. "AkshatBlog." Akshatblog.com, 2015, www.akshatblog.com/graphics-card-components-explained-in-detail/. Accessed 23 May 2021.

Aleksandar Cosic. "GPU Power Connectors Explained [Simple Answer] - GPU Mag." GPU Mag, GPU Mag, 4 Apr. 2021, www.gpumag.com/gpu-power-connectors-explained/. Accessed 23 May 2021.

Anish, M., & Kanimozhi, B. (2018). Experimental investigation and heat transfer process on longitudinal fins with different notch configuration. International Journal of Ambient Energy, 39(1), 34–37. https://doi.org/10.1080/01430750.2016.1222965

Anwar, Khairul, et al. "Overview on the Response Surface Methodology (RSM) in Extraction Processes." Journal of Applied Science & Process Engineering, vol. 2, no. 1, 2015. Accessed 22 Sept. 2021.

Asirvatham, L. G., Nimmagadda, R., & Wongwises, S. (2013). Heat transfer performance of screen mesh wick heat pipes using silver-water nanofluid. International Journal of Heat and Mass Transfer, 60(1), 201–209. https://doi.org/10.1016/j.ijheatmasstransfer.2012.11.037

Bădălan, N., & Svasta, P. (2017). Fan vs. passive heat sink with heat pipe in cooling of high power LED. 2017 IEEE 23rd International Symposium for Design and Technology in Electronic Packaging, SIITME 2017 - Proceedings, 2018-January, 296–299. https://doi.org/10.1109/SIITME.2017.8259911

Bhattacharya, Sankha. Central Composite Design for Response Surface Methodology and Its Application in Pharmacy. Www.intechopen.com, IntechOpen, 28 Jan. 2021, www.intechopen.com/online-first/74955. Accessed 25 Sept. 2021.

Buber, E., & Diri, B. (2018). Performance analysis and CPU vs GPU comparison for deep learning. 2018 6th International Conference on Control Engineering and Information Technology, CEIT 2018, October, 1–6. https://doi.org/10.1109/CEIT.2018.8751930

Chen, J. S., & Chou, J. H. (2015). The length and bending angle effects on the cooling performance of flat plate heat pipes. International Journal of Heat and Mass Transfer, 90, 848–856. https://doi.org/10.1016/j.ijheatmasstransfer.2015.06.032

Choi, J., Jeong, M., Yoo, J., & Seo, M. (2012). A new CPU cooler design based on an active cooling heatsink combined with heat pipes. Applied Thermal Engineering, 44, 50–56. https://doi.org/10.1016/j.applthermaleng.2012.03.027

Culham, J. R., Khan, W. A., Yovanovich, M. M., & Muzychka, Y. S. (2007). The influence of material properties and spreading resistance in the thermal design of plate fin heat sinks. Journal of Electronic Packaging, Transactions of the ASME, 129(1), 76–81. https://doi.org/10.1115/1.2429713

Dasgupta, A., Hong, S., Kim, H., & Park, J. (n.d.). A New Temperature Distribution Measurement Method on GPU Architectures.

De Schampheleire, S., De Kerpel, K., Deruyter, T., De Jaeger, P., & De Paepe, M. (2015). Experimental study of small diameter fibres as wick material for capillary-driven heat pipes. Applied Thermal Engineering, 78, 258–267. https://doi.org/10.1016/j.applthermaleng.2014.12.027
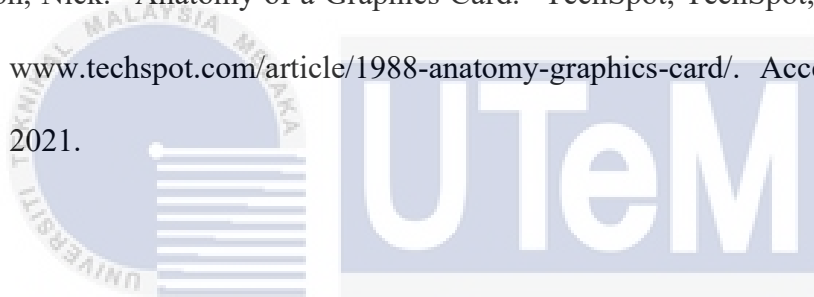
Eberly, D. H. (2020). CPU Computing. GPGPU Programming for Games and Science, 33–120. https://doi.org/10.1201/b17296-8

Evanson, Nick. "Anatomy of a Graphics Card." TechSpot, TechSpot, 16 Mar. 2020, www.techspot.com/article/1988-anatomy-graphics-card/. Accessed 23 May 2021.

Eyjolfsson, Reynir. "Introduction." Design and Manufacture of Pharmaceutical Tablets, 2015, pp. 1–28, www.sciencedirect.com/science/article/pii/B9780128021828000015, 10.1016/b978-0-12-802182-8.00001-5. Accessed 20 Feb. 2020.

Feng, S., Shi, M., Yan, H., Sun, S., Li, F., & Lu, T. J. (2018). Natural convection in a cross-fin heat sink. Applied Thermal Engineering, 132(December), 30–37. https://doi.org/10.1016/j.applthermaleng.2017.12.049

Ganji, Davood Domairry, et al. "Radiation Heat Transfer." Nonlinear Systems in Heat Transfer, 2018, pp. 105–151, 10.1016/b978-0-12-812024-8.00003-5.

Gao, T., Kumar, E., Sahini, M., Ingalz, C., Heydari, A., Lu, W., & Sun, X. (2016). Innovative server rack design with bottom located cooling unit. Proceedings of the 15th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, ITherm 2016, 1172–1181. https://doi.org/10.1109/ITHERM.2016.7517681

Huang, C.-H., & Gau, C.-W. (2012). An optimal design for axial-flow fan blade: theoretical and experimental studies. Journal of Mechanical Science and Technology, 26(2), 427–436. https://doi.org/10.1007/s12206-011-1030-7

Hussain, A. A., Freegah, B., Khalaf, B. S., & Towsyfyan, H. (2019). Numerical investigation of heat transfer enhancement in plate-fin heat sinks: Effect of flow direction and fillet profile. Case Studies in Thermal Engineering, 13(December 2018). https://doi.org/10.1016/j.csite.2018.100388

Jouhara, H. (2018). Heat Pipes. Comprehensive Energy Systems, 4–5(m), 70–97. https://doi.org/10.1016/B978-0-12-809597-3.00403-X

Mahdavi, M., Tiari, S., De Schampheleire, S., & Qiu, S. (2018). Experimental study of the thermal characteristics of a heat pipe. Experimental Thermal and Fluid Science, 93(August 2017), 292–304. https://doi.org/10.1016/j.expthermflusci.2018.01.003

Mehmood, Tahir, et al. "Optimization of Olive Oil Based O/W Nanoemulsions Prepared through Ultrasonic Homogenization: A Response Surface Methodology Approach." Food Chemistry, vol. 229, Aug. 2017, pp. 790–796, 10.1016/j.foodchem.2017.03.023. Accessed 23 Apr. 2020.

Milano, Jassinnee, et al. "Optimization of Biodiesel Production by Microwave Irradiation-Assisted Transesterification for Waste Cooking Oil-Calophyllum Inophyllum Oil via Response Surface Methodology." Energy Conversion and Management, vol. 158, Feb. 2018, pp. 400–415, 10.1016/j.enconman.2017.12.027. Accessed 22 Sept. 2021.

Mitra, Pranabendu, et al. "Pumpkin (Cucurbita Maxima) Seed Oil Extraction Using Supercritical Carbon Dioxide and Physicochemical Properties of the Oil." Journal of Food Engineering, vol. 95, no. 1, Nov. 2009, pp. 208–213, 10.1016/j.jfoodeng.2009.04.033. Accessed 20 Nov. 2020.

Mjallal, I., Farhat, H., Hammoud, M., Ali, S., Shaer, A. A. L., & Assi, A. (2018). Electronic Devices. 02003, 1–4.

Mujtaba, M.A., et al. "Ultrasound-Assisted Process Optimization and Tribological Characteristics of Biodiesel from Palm-Sesame Oil via Response Surface Methodology and Extreme Learning Machine - Cuckoo Search." Renewable Energy, vol. 158, Oct. 2020, pp. 202–214, 10.1016/j.renene.2020.05.158. Accessed 22 Sept. 2021.

Muthukumaran, Chandrasekaran, et al. "Process Optimization and Kinetic Modeling of Biodiesel Production Using Non-Edible Madhuca Indica Oil." Fuel, vol. 195, May 2017, pp. 217–225, 10.1016/j.fuel.2017.01.060. Accessed 22 Sept. 2021.

Ng, J. (2019). Computational fluid dynamics simulation on the heat sink performance of a graphics processing unit thermal management. Jurnal Kejuruteraan, 31(1), 139–147.
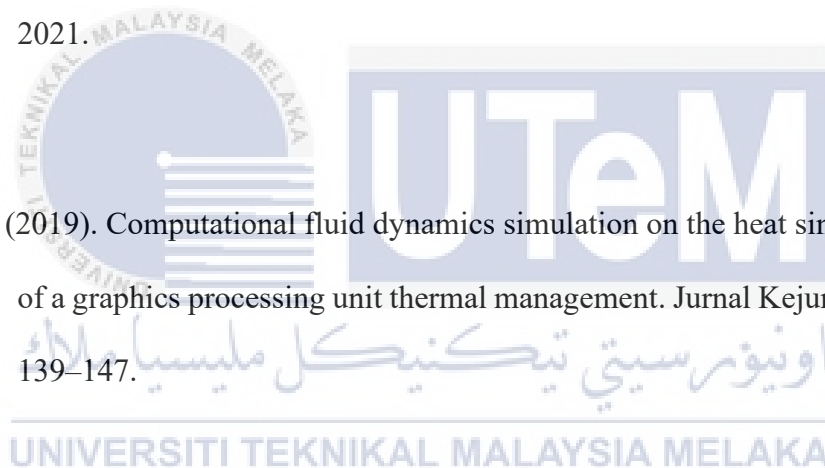
Pambudi, N. A., Bugis, H., Kuncoro, I. W., Setiawan, N. D., Hijriawan, M., Rudiyanto, B., & Basori, B. (2019). Preliminary experimental of GPU immersion-cooling. E3S Web of Conferences, 93, 1–4. https://doi.org/10.1051/e3sconf/20199303003

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

Rai, Amit, et al. "Supercritical Extraction of Sunflower Oil: A Central Composite Design for Extraction Variables." Food Chemistry, vol. 192, Feb. 2016, pp. 647–659, 10.1016/j.foodchem.2015.07.070. Accessed 28 Feb. 2020.

Rajabi, N. & Rafee, Roohollah & Frazam-Alipour, S.. (2017). Effect of blade design parameters on air flow through an axial fan. International Journal of Engineering, Transactions A: Basics. 30. 1583-1591. 10.5829/ije.2017.30.10a.20.

Rakić, Tijana, et al. "Comparison of Full Factorial Design, Central Composite Design, and Box-Behnken Design in Chromatographic Method Development for the Determination of Fluconazole and Its Impurities." Analytical Letters, vol. 47, no. 8, 2 May 2014, pp. 1334–1347, 10.1080/00032719.2013.867503. Accessed 26 Sept. 2021.

Sadrieh, A., & Bahri, P. A. (2011). Optimal control of the process systems using Graphic Processing Unit. In IFAC Proceedings Volumes (IFAC-PapersOnline) (Vol. 44, Issue 1 PART 1). IFAC. https://doi.org/10.3182/20110828-6-IT-1002.01804

Sahoo, Prasanta, and Tapan Kr. Barman. "ANN Modelling of Fractal Dimension in Machining." Mechatronics and Manufacturing Engineering, 2012, pp. 159–226, 10.1533/9780857095893.159. Accessed 25 Sept. 2021.

Savic, Ivana, et al. "Modelling and Optimization of Methylene Blue Adsorption from Aqueous Solution Using Bentonite Clay." ScienceDirect, Elsevier, 1 Jan. 2014, www.sciencedirect.com/science/article/abs/pii/B9780444634559500714. Accessed 25 Sept. 2021.

Scharon Harding. "What Is a VRM? A Basic Definition." Tom's Hardware, Tom's Hardware, 12 Dec. 2018, www.tomshardware.com. Accessed 23 May 2021.

Shahidian, Azadeh, et al. "Introduction." Bio-Engineering Approaches to Cancer Diagnosis and Treatment, 2020, pp. 1–22, 10.1016/b978-0-12-817809-6.00001-7.
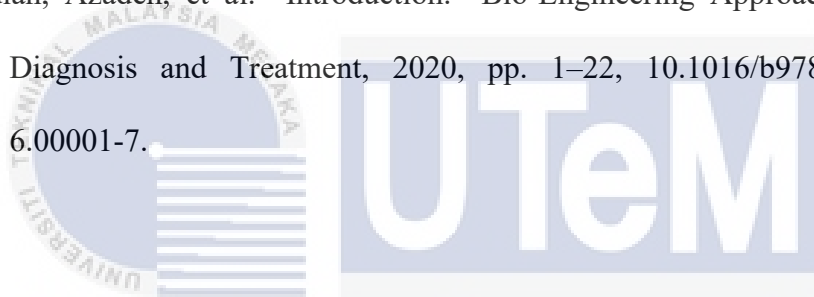
Singh, G. (2017). Natural Convection Heat Transfer from Modified 1Degree 2 Degree And 3 Degree Outward Expansion of Pin Fins. 7(8), 14567–14570.

Smistad, E., Falch, T. L., Bozorgi, M., Elster, A. C., & Lindseth, F. (2015). Medical image segmentation on GPUs - A comprehensive review. Medical Image Analysis, 20(1), 1–18. https://doi.org/10.1016/j.media.2014.10.012

Sokolova, Inna. "Temperature Regulation." Encyclopedia of Ecology, 2019, pp. 633–639, 10.1016/b978-0-12-409548-9.11129-7.

Solomon, A. B., Ramachandran, K., & Pillai, B. C. (2012). Thermal performance of a heat pipe with nanoparticles coated wick. Applied Thermal Engineering, 36(1), 106–112. https://doi.org/10.1016/j.applthermaleng.2011.12.004

Stewart, S. (2021, February 18). Best Graphics Card Brands & Manufacturers [AMD & NVIDIA]. GamingScan; GamingScan. https://www.gamingscan.com/best-graphics-card-brands-manufacturers/

Tharayil, T., Asirvatham, L. G., Cassie, C. F. M., & Wongwises, S. (2017). Performance of cylindrical and flattened heat pipes at various inclinations including repeatability in anti-gravity – A comparative study. Applied Thermal Engineering, 122, 685–696. https://doi.org/10.1016/j.applthermaleng.2017.05.007

Veni, D. Krishna, and N. Vishal Gupta. "Development and Evaluation of Eudragit Coated Environmental Sensitive Solid Lipid Nanoparticles Using Central Composite Design Module for Enhancement of Oral Bioavailability of Linagliptin." International Journal of Polymeric Materials and Polymeric Biomaterials, vol. 69, no. 7, 5 Feb. 2019, pp. 407–418, 10.1080/00914037.2019.1570513. Accessed 23 Feb. 2021.

Wagner, John R., et al. "25 - Design of Experiments." ScienceDirect, William Andrew Publishing, 1 Jan. 2014, www.sciencedirect.com/science/article/pii/B9781437734812000259. Accessed 27 Mar. 2021.

Wang, Hongwu, et al. "Application of Response Surface Methodology to Optimise Supercritical Carbon Dioxide Extraction of Essential Oil from Cyperus Rotundus Linn." Food Chemistry, vol. 132, no. 1, May 2012, pp. 582–587, 10.1016/j.foodchem.2011.10.075. Accessed 10 May 2020.

Wu, Y., & Huang, D. (2019). Optimization design of axial fan blade. Journal of the Chinese Institute of Engineers, 42(6), 473–478. https://doi.org/10.1080/02533839.2019.1611478

Ye, Qingzhuo, et al. "Development and Evaluation of Puerarin-Loaded Controlled Release Nanostructured Lipid Carries by Central Composite Design." Drug Development and Industrial Pharmacy, vol. 47, no. 1, 2 Jan. 2021, pp. 113–125, 10.1080/03639045.2020.1862170. Accessed 9 Nov. 2021.

Zhang, L., & Jin, Y. Z. (2011). Effect of Blade Numbers on Aerodynamic Performance and Noise of Small Axial Flow Fan. Advanced Materials Research, 199-200, 796–800. https://doi.org/10.4028/www.scientific.net/amr.199-200.796

Zhou, S., Zhou, H., Yang, K., Dong, H., & 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, 100971. https://doi.org/10.1016/j.seta.2020.100971
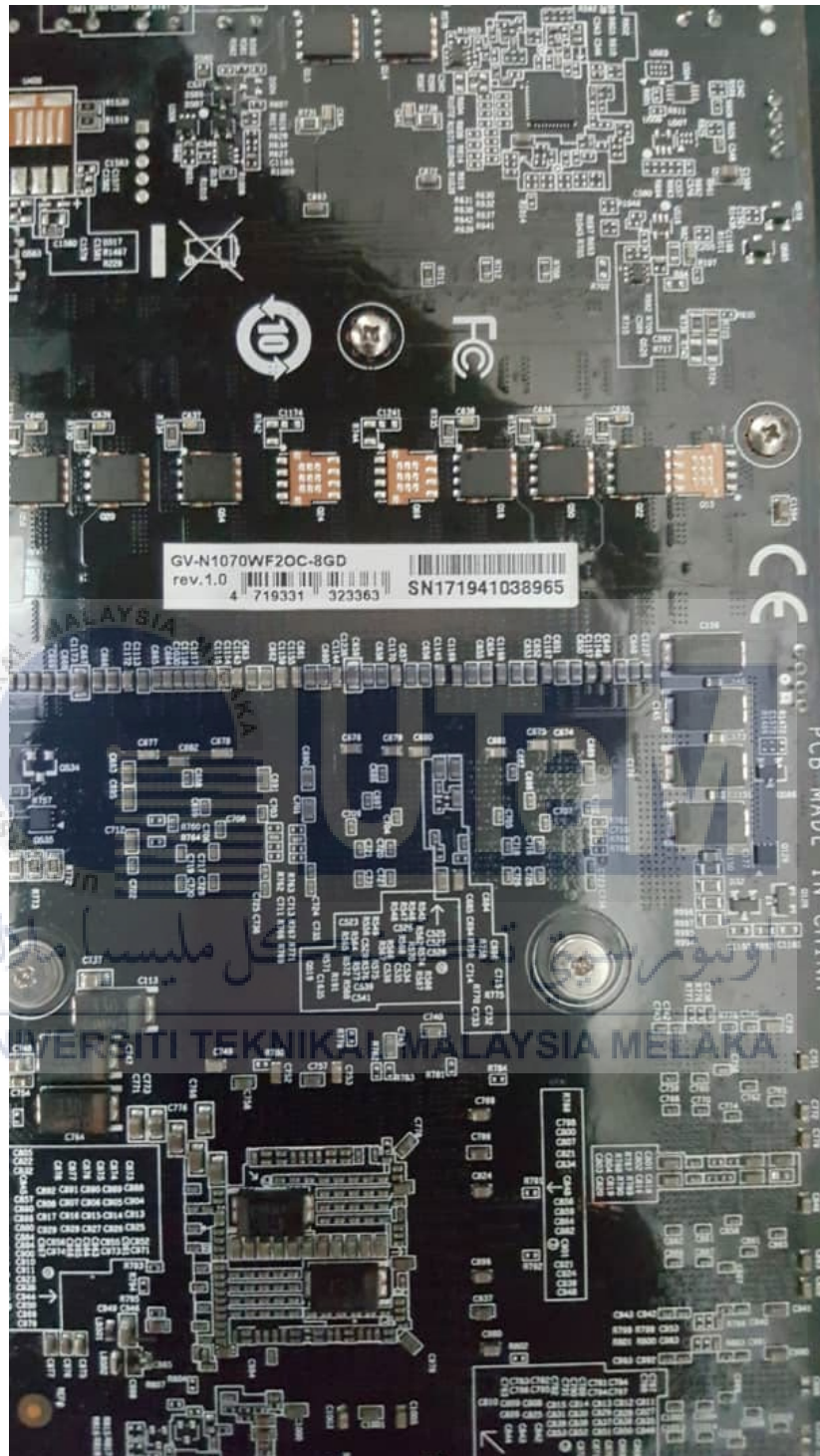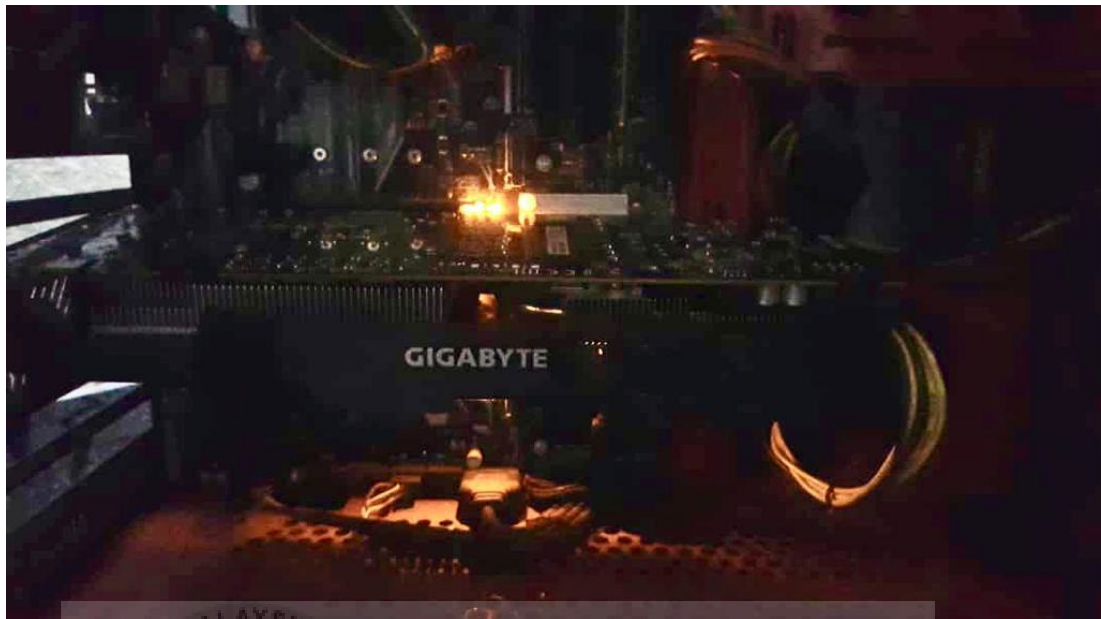
APPENDICES

Appendix A: Top view of the GPU

Appendix B: Front view of the GPU

Appendix C: GPU's board

Appendix D: GPU in the PC case (view 1)

Appendix E: GPU in the PC case (view 2)

Appendix F: GPU in the PC case (view 3)