The International Journal of Engineering and Science (IJES)
|| Volume || 13 || Issue || 8 || Pages || PP 29-51 || 2024 ||
ISSN (e): 2319 – 1813 ISSN (p): 2319 – 1805

Design of Experiments Applications and Case Study on Vertical Axis Wind Turbine

Htet Yan Kyaw
[1]*The University of Newcastle, Australia*
Corresponding Author

--*ABSTRACT*---
Design of Experiments (DOE) is a methodology used for problem-solving in various situations that applies multiple notions and techniques to data collected to generate usable, and reliable engineering conclusions, inclusive to optimizing the yield of products and minimizing the costs of materials consumed in the process. Initiated by Dr. Genichi Taguchi, 'DOE' became heavily influenced in industrial manufacturing where it introduces efficient methods of studying the relationship between multiple input and output variables. The primary objective of the project is applying (DOE) onto wind turbines by using 'Randomized Complete Block Design (RCBD)' as a method of analyzing the experiment using numerical results. The principle of (RCBD) is to group specific factors into 'Blocks', which are influenced by a set of 'Treatments' to control the variability of the output. Implementing the experimental hypothesis allows the experimenter to conclude the significance of such factors in an enigma. Predicting where the yield will progress by applying 'Multiple Linear Regression' to create a model based on the results in the pre-experimentation. Such processes are undergone by 'Microsoft Excel' and 'Python' to generate clarity in studying systems and problems. In the case of VAWTs, factors including Tip Speed Ratio, Blades and Pitch angle can mainly affect the Power coefficient of the wind turbine. Thus, 'Stepwise Selection' is implemented to filter any significant factors and generate predictors to elucidate the inclination related to the dependent variable. Due to the significance of concluding solutions by repetition, (DOE) can be applied to any heavy industry where a single factor may affect the overall outcome of their product.

KEYWORDS; -*Design of Experiments, Random Complete Block Design, Multilinear regression, Stepwise Selection, Tip Speed Ratio, Blades, Degrees,VAWTs*

Date of Submission: 03-08-2024 Date of acceptance: 14-08-2024

I. INTRODUCTION

Design of Experiments (DOE) has portrayed a major contribution in science and technology since the time Sir Ronald Aylmer Fisher introduced this concept in the 1940s. [1]

Scientists and engineers alike made extensive use of this idea for both new product development and enhancements. Increasing the yield of information from each experiment that is conducted is its primary goal. By implementing "Design of Experiments (DOE)" in a product cycle, development time and expense can be significantly decreased. This enables improved process and product design, as well as increased consumer reliability. The experimental design philosophy was widely introduced to the chemical and process industries by George Edward Pelham Box and his associates in 1950. This approach made it possible to examine and track the systems and procedures of a specific product. Typical examples are the production of wafers in the electronics industry, manufacturing of engines in the car industry and synthesis of compounds in the pharmaceutical industry. [18]

Over the past decades, after Dr. Genichi Taguchi's influence, there had been a tremendous increase in demand and application of experimental designs techniques in the industrial environment. While applying the experimental designs, the 'Total Quality Management' had improved and ultimately increased the company's revenue.

The observation of a system or a process while it is in operation can be one of the most important parts of the learning process and is an integral part of understanding how the systems work. However, to understand not only the process and the cause-and-effects of the system when changing the input factors but also to observe the changes in the system output for which such inputs are produced. Specifically, performing experiments on the system will eventually catalyze into theories or hypotheses about what makes the system works.

Investigators normally perform experiments in a virtual environment, usually to discover the particularities about the process or system to confirm whether their previous experience and theories uphold. Each of these experiments or runs, are formally defined as a test in which purposefulness changes are made to input variables of a process or a system so that observation and identification of the reasons for changes that occurred in the output response. From the analysis of the changes, models relating to the response to the impute variables can be developed and used to either improve or derive the system by its model.

Experimentation has demonstrated to be an important role in 'Technological Commercialization' and 'Product Realization' activities such as new product design, formulation, manufacturing process development and process improvement. Its main aim in many cases is to develop a 'Robust' system, allowing minimal variability of external sources affecting the process. Not only such applications may be known but also applications namely marketing, service operations, general business operations can be used in a nonmanufacturing or non-product-developed settings.

Explicating that Experimentations is a vital section of a 'Scientific Engineering Method' can be portrayed as a concept for scientists and engineers for the process development and improvement of yield in product design. However, there are circumstances for which the scientific singularities are well understood allowing useful mathematical models to be developed by applying these well-understood principles. 'Mechanical Models' provides physical mechanism which follows the scientific principle. For instance, 'Ohm's Law' works as a familiar equation for the current flow in an electrical circuit as shown in Equation (1.1).

$$V = IR \tag{1.1}$$

Especially, in problems as such nowadays in science and engineering requires observations of the said system and several experimentations to clarify information on the reasons on the mechanisms. As such 'Well-Designed' experiments can often lead to a model of system of performance that allows manipulation of variables and achieving a desired result is called 'Empirical Model'.

Establishing the goals of an experiment and the variables that affect a system or process are the first steps in the process of conducting experimental designs. A planned experiment necessitates creating a thorough plan for carrying out the experiment, which leads to a more efficient approach to the data gathering phase. The yield of information that may be collected for a given amount of experimental work is maximized when an appropriate experimental design is used. In general, there are three types of experimental design: the One-Factor-at-a-Time (OFAT) approach, Factorial Experimental design, and Orthogonal array.[4]

II. MAIN OBJECTIVE

The main motivation of this project is to employ the main methodology of DOE in solving problems related to data collection and analysis. As there may be less data to be collected if the problem proves to be minuscule to impact the user, it will take less time to solve the problem. Although, a more complex analysis and an extended time is needed to unravel the adversity as it has more elements to consider, including the environmental, ethical, or physical aspects. The goal is to apply DOE onto heavier projects such as product manufacturing and interpretation. This will allow industrial related processes to loosen their difficulties and produce much more efficient methods of production. The same can be said for the improvement of existing products allowing concrete showcase of the breakthrough made by the methods of DOE. In this case, the problem resides on the how efficient the Vertical Axis Wind Turbine is in the presence of multiple factors such as 'Tip Speed Ratio', 'Number of Blades', and 'Angle of Blades' and create a model defining its prediction on where the pattern of efficiency is leading to.

Figure 1: Darrieus-type Vertical Axis Wind Turbine.

III. LITERATURE REVIEW

Figure 2: Doctor Genichi Taguchi

Dr. Genichi Taguchi, a Japanese engineer and statistician born on January 1, 1924, is primarily known for his development of a statistical method to improve the quality of manufactured goods. His initial interest in textile engineering, which led to a family kimono business, was disrupted by the military escalation of World War II. This led to his conscription into the ‘Department of Astronomy of the Navigation Institute of the Imperial Japanese Army’.After the war, in 1948, he joined the ‘Ministry of Public Health and Social Welfare’, serving under the influential politician Matosaburo Masuyama, who fostered Taguchi’s interest in experimental design. During this period, he studied at the ‘Institute of Mathematical Statistics’ and contributed to the ‘Morinaga’ pharmaceutical company’s experimental work on penicillin production.In 1950, he collaborated with the ‘Electrical Communication Laboratory of the Nippon Telegraph and Telephone Corporation’ on his advancements in ‘Design of Experiment’. This was influenced by Dr. W. Edwards Demming and the Japanese Union of Scientists and Engineers, as statistical quality control was gaining popularity in Japan.Dr. Taguchi spent 12 years developing strategies to enhance the quality and reliability of the ‘Electrical Communications Laboratory’. This led to his consultancy role at ‘Toyota’. In 1962, he completed his Ph.D. at the ‘University of Kyushu’ while continuing his work at the ‘Electrical Communication Laboratory’.From 1964 to 1982, Dr. Taguchi served as a professor at ‘Aoyama Gakuin University’. During his professorship, he introduced the ‘Quality Loss Function’, a concept that quantifies quality in monetary terms via a loss function, where greater deviation from the nominal value results in higher financial loss to the consumer.

While he published two volumes of ‘Design of Experiments’ in the 1950s, he added a third book to the series later. Despite his technique being largely unrecognized in the West at the time, he had won multiple ‘Deming’ medals and served as the Director of the ‘Japanese Academy of Quality’.Dr. Taguchi’s technique was extensively adopted by American corporations such as ‘Ford Motor Company,’ ‘Boeing,’ ‘Xerox,’ and ‘ITT Corporation’ during his subsequent visit to the United States in 1980.[9]

A. Basic Principles

The Statistical design of experiments refers to the process of planning the experiment for which the appropriate data will be collected and analyzed by statistical methods, resulting in compelling and objective closing stages. The statistical approach to experimental design is obligatory for drawing meaningful conclusions from the data. When the problem involves data that are subject to experimental errors, statistical methods are the only objective approach to analysis. Thus, there are two aspects to any experimental problem: the design of the experiment and the statistical analysis of the data.

The three basic principles of experimental design are ‘Randomization’, ‘Replication’, and ‘Blocking’, while using the factorial principle into them. First, ‘Randomization’ is defined as a used of statistical method in experimental design. Randomization occurs when both the allocation of the experimental material and the order of the randomized performance of the experimenter. This principle usually makes the statement of having statistical methods requiring independently distributed random variables to be observed on, to be valid. By properly randomizing the experiment, the process of averaging out the effects of numerous factors in the experiment can be presented accordingly. Computer software allows the assistance for experimenters in the selection and construction of experimental designs. Programs with similar usability enable the presentations of tests in the experimental design in random order, based on ‘Random Number Generator (RNG)’. It is still necessary to assign the named units of experimental materials, experimenter, gauges, or measurement devices used in an experiment.

On the other hand, 'Randomization' can lead to some difficult situations during experimentation, where the controllable factors such as, ambient temperature, shouldn't be randomized or changed less than other factors in a chemical process.

'Replication', or repeated runs, consists of frequent tests of many combinations of factors. Utilizing different factors to be combined with each other and repeated multiple times achieving multiple results observed. 'Replication' has two important properties. First, the principle allows the experimenter to obtain an estimate of the experimental error, becoming a basic unit of measurement for determining whether the observations of difference in the data can be statistically different. Second, replication allows the experimenter to obtain a precise estimate of the parameter of the data given. Although, the difference between 'Replication' and 'Repeated Measurements' have similar meanings but different implications to experimentation. For example, etching a silicon wafer in a single wafer plasma etching process and a 'Critical Dimension (CD)', which is defined as precise measurements of dimensions to a decimal point of a drawing, on the silicon wafer is measured three times. The measurements are not replicates but a form of repeated measurements and can be observed the variability in the three repeated measurements. This is a direct reflection of the variation in the measurement system or gauge and possibly change in the variability in the 'Critical Dimensions' at different areas on the wafer where the measurements were taken.

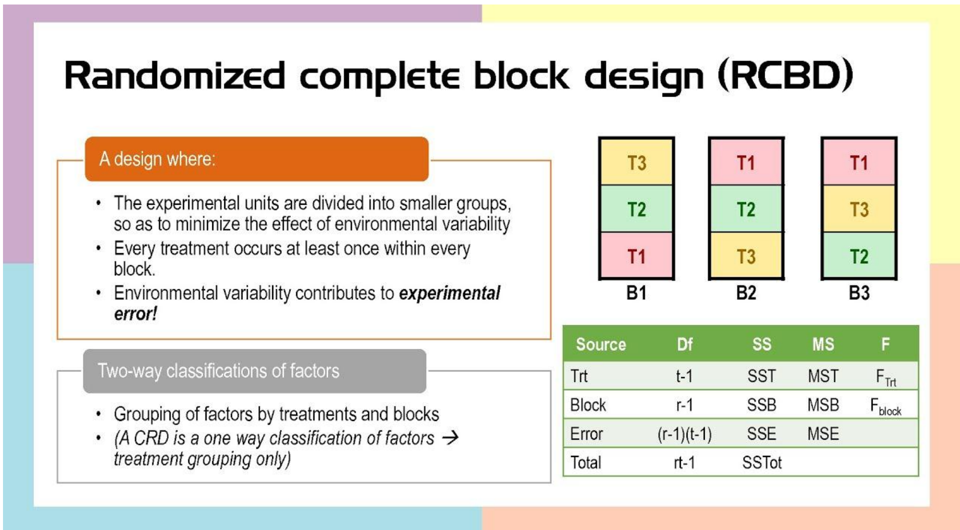
Another instant provides that when four semiconductor wafers being manufactured simultaneously in an oxidation furnace at a particular gas flow rate and time and measurements were taken on the thickness of the oxide of each wafer. The measurements taken were not replicates but repeated measurements, which meant that repeated measurements reflect the differences among the wafers and other sources of variability within the furnace testing. However, replication reflects on the sources of variability both between tests and within each test.

'Blocking' is a design technique used in improving the precision with comparisons among the factors of interest are made. Blocking often is used to reduce (or eliminate) the variability being transmitted from 'Nuisance Factors', for which factors that may influence the experimental response but not related to the experiment. For instance, in an experiment for mixing two batches of raw materials in a chemical mixing process is ran multiple times. Due to the differences in batches of raw materials which had supplier-to-supplier variability was ignored and remained uninterested leaving this batches of raw material as a 'Nuisance Factor'.

Each of the batches of raw material are considered as a block, because the variability within the batch was expected to be smaller than the variability between each batch.[9]

B. Random Complete Block Design (RCBD)

Figure 3: Illustration of the use of RCBD.



The 'Random Complete Block Design (RCBD)' generally is a statistical technique that is widely used in various fields of research, such as agriculture, medicine, and social sciences. The RCBD is a type of experimental design that aims to reduce the effect of extraneous variables that may affect the outcome of the study. These variables are also known as confounding factors or nuisance factors, and they can introduce unwanted variability and bias in the results.

The RCBD works by dividing the experimental units (the subjects or objects that receive the treatments) into homogeneous groups, called 'Blocks'. The blocks are formed based on factors that is known or suspected to influence the outcome of the study, such as age, gender, location, time, etc. The idea is to create blocks that are as similar as possible within each block, but as different as possible between blocks. This way, the variability within each block is minimized, while the variability between blocks is maximized as shown in Figure 2.[19]

Within each block, all the treatments are randomly assigned to the experimental units. This means that each treatment is tested within each block, and each block contains a complete set of treatments. The randomization ensures that the treatments are distributed evenly across the blocks, and that any bias caused by the blocking factor is balanced out. The randomization also helps to prevent any systematic errors or confounding effects from other factors.

The RCBD has several advantages over other experimental designs. Some of these advantages are:

- Increased precision: By controlling for the effect of the blocking factor, the RCBD reduces the variability in the experimental error, which increases the precision of the experiment. This means that the RCBD can detect smaller differences between treatments with a higher level of confidence.
- Reduced bias: By ensuring that each treatment is tested across different blocks, the RCBD reduces the bias caused by the blocking factor. This means that the RCBD can provide more accurate estimates of the treatment effects and their standard errors.
- Efficient use of resources: By reducing the variability in the experimental error, the RCBD allows for more efficient use of resources. This means that the RCBD can achieve the same level of precision with fewer experimental units or fewer replications than other designs [10].

The RCBD is commonly used in various industries for different purposes. For example:

- In agriculture, the RCBD is often used to test different crop varieties, fertilizers, pesticides, irrigation methods, etc. The blocks are usually based on soil type, fertility level, slope, aspect, etc. The RCBD helps to account for the spatial variability and environmental heterogeneity in agricultural fields.

In conclusion, the randomized complete block design is a useful statistical technique that can help researchers reduce the effect of extraneous variables in their studies. It has several benefits over other experimental designs in terms of precision, bias reduction, and resource efficiency. It is widely used in various industries for different purposes and applications. However, it also has some limitations and challenges that need to be considered before using it [10].

C. MULTIPLE LINEAR REGRESSION

Multilinear Regression commonly is a statistical technique that uses two or more independent variables to predict the outcome of a dependent variable. The independent variables are also known as explanatory variables, predictor variables, or regressors, while the dependent variable is also known as the response variable or the

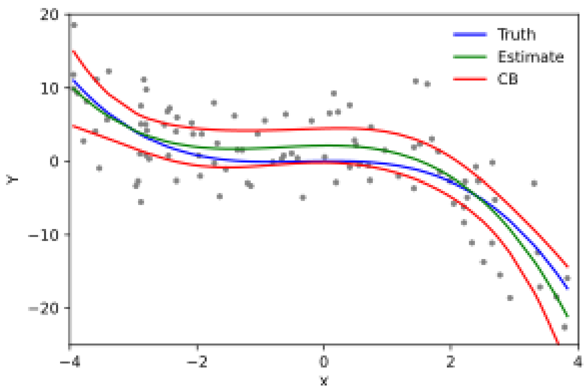


Figure 4: Multiple Linear Regression model example.

regressed. The multilinear regression model assumes that there is a linear relationship between the dependent variable and each of the independent variables, and that the effects of the independent variables are additive.[23]

The 'Multilinear Regression Model' can be defined as an Equation (1.2) or an empirical model relating to the independent variable, represented as the coefficient.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon \quad (1.2)$$

Where y is the dependent variable, representing the output of the experiment, $x_1 and\ x_2$ shows the named independent variables. The parameters, $\beta_0, \beta_1, \beta_2$ are called the 'Regression Coefficients'.

The 'Multilinear Regression Model' can be used to analyze the relationship between one dependent and multiple independent variables. The dependent variable is the main outcome or the response variables for which it is predicted or explained, while the independent variables are main factors that influences the dependent variable, affecting it positively or negatively.

Several features of the 'Multilinear Regression Model' can be made useful in various purposes and applications in different industries. Some of such features are:

- Being able to handle more independent variables simultaneously and estimate their relative importance and contribution to the dependent variable.
- Provides a quantitative measure of how befitting the model fits the data and how much variation in the dependent variable that can be explained by the independent variables [23].

The drawbacks of using the 'Multilinear Regression Model' need to be considered when applying to certain scenarios, giving some limitations and challenges. Some of which are:

- When independent variables, that are usually unrelated to each other, shows a highly correlation with each other and cause instability and redundancy in the model. This is called, 'Multicollinearity'.
- Referring from 'RCBD', the 'Multilinear Regression model' requires a large amount of data to ensure sufficient power and accuracy of the resultant model.

The multilinear regression model can be used for various purposes and applications in different industries. For example, it can be used to analyze how different factors affect the sales, revenue, profit, demand, supply, or cost of a product or service; how different physical or chemical properties affect the performance, efficiency, reliability, or safety of a system or process; how different biological or environmental factors influence the health status, disease risk, or treatment outcome of a patient or population; and so on.

Singapore is a leading manufacturing hub in Asia and globally. It has a strong and diverse manufacturing base, with leadership positions in sectors such as electronics, biomedical sciences, precision engineering, and chemicals. Singapore has embraced Industry 4.0, which is a term for the fourth industrial revolution that involves digitalization, automation, artificial intelligence, and internet of things. Singapore has established a strong base of leading technology and solutions providers that support Industry 4.0 adoption among manufacturers. Singapore has also embarked on a series of initiatives to ensure its manufacturing sector is prepared for the future.[17]

One of these initiatives is to promote multilinear regression as a tool for advanced manufacturing. Multilinear regression can help manufacturers to achieve various objectives such as:

- Improving product quality: Multilinear regression can help manufacturers to identify and control the critical factors that affect product quality, such as material properties, machine parameters, process conditions, and human factors. Multilinear regression can also help manufacturers to monitor product quality in real time using sensors and feedback systems.
- Reducing production cost: Multilinear regression can help manufacturers to optimize their production processes by finding the optimal combination of inputs that minimize cost while meeting quality standards. Multilinear regression can also help manufacturers to reduce waste and energy consumption by identifying inefficiencies and anomalies in their production systems.
- Increasing production efficiency: Multilinear regression can help manufacturers to increase their production output by maximizing their utilization of resources such as machines, materials, labour, and time. Multilinear regression can also help manufacturers to improve their production flexibility by adapting to changing customer demands and market conditions.

- Enhancing production innovation: Multilinear regression can help manufacturers to explore new possibilities for product design, development, and improvement by testing different scenarios and hypotheses using data analysis. Multilinear regression can also help manufacturers to discover new insights and patterns from their production data that can lead to breakthrough innovations.[2]

To illustrate how multilinear regression is used in Singapore manufacturing and industrial systems, here are some examples from different sectors:

- Electronics: In electronics manufacturing, multilinear regression can be used to optimize chip design and fabrication by analysing how various factors such as temperature, voltage, current, pressure, speed, etc., affect chip performance and reliability.
- Biomedical sciences: In biomedical sciences manufacturing, multilinear regression can be used to improve drug development and production by analysing how various factors such as chemical composition, dosage, formulation, delivery method, etc., affect drug efficacy and safety.
- Precision engineering: In precision engineering manufacturing, multilinear regression can be used to enhance machine performance and maintenance by analyzing how various factors such as vibration, noise, wear, tear, etc., affect machine functionality and durability.
- Chemicals: In chemicals manufacturing, multilinear regression can be used to optimize chemical processes and products by analyzing how various factors such as temperature, pressure, flow rate, concentration, etc., affect chemical reactions and properties.

D. Stepwise Selection Process

'Stepwise Selection' is a system of fitting a regression model which involves selecting the best subset of predictor variables from a larger set of potential variables. Regression models themselves are used to describe and predict the relationship between response variables and one or more predictor variables, based on the observed data. However, having variables that are more than fingers can count, 'Stepwise Selection' can help reduce the complexity and improve the accuracy of the regression models by eliminating the irrelevant or redundant variables [24-31].

'Stepwise Selection' can be applied to various types of regression models, such as linear regression and polynomial regression. Each type of regression model has different applications and assumptions, depending on the nature of the response variable and the predictor variables. 'Stepwise Selection' has many real-world applications which can be practiced in different fields and domains, namely business, medicine, agriculture, psychology, economics, and manufacturing. Examples of each domain for which 'Stepwise Selection' is adapted.

- Businesses can use 'Stepwise Selection' to build linear regression models that measure how advertising spending affects revenue. They can also use it to optimize their budget allocation and maximize their profits.
- Agricultural researchers employ stepwise selection to create polynomial regression models. These models help quantify the impact of fertilizer and water on crop yields and assist in determining the optimal resource allocation for crop growth.
- Psychologists employ stepwise selection to construct ridge regression models. These models analyze the impact of various personality traits on happiness and mitigate the influence of multicollinearity, thereby enhancing the accuracy of their estimates.
- Economists utilize stepwise selection to construct time series regression models [32-41]. These models analyze the dynamics of economic indicators over time, enabling them to forecast future trends and identify potential disruptions.[24]
- Manufacturers can improve their suppliers' performance and predict whether to determine the strategic developments based on the model analysis, allowing them to benchmark operation efficiencies and importance in components.[24]

'Stepwise Selection' proves to be a useful technique for building regression models to become simple, accurate and interpretable. However, it also has some limitations and challenges. The computational cost can be high, especially when there are many possible variables to consider. The variable selection process can depend on the sequence, the criterion, and the significance level of the variables. The model can fit the training data too well and fail to generalize to the new data, which is called overfitting. The model can disregard the theoretical or domain knowledge that may inform the model selection process. The model can produce estimates and confidence intervals that are not accurate and do not reflect the uncertainty of the model selection.[24]

E. Vertical Axis Wind Turbine Background Information

1. Introduction

The 'Wind' energy has been a promising source of power generation in recent years due to its potential to be a 'Carbon' free power generation. Most of the wind energy technology is based around on HAWTs (Horizontal Axis Wind Turbines), which are the most recognizable utilization of power generation. The HAWTs are

Figure 5: Vertical Axis Wind Turbine.

described as a typical wind turbine spinning around the horizontal axis, parallel to the wind direction. Large scale projects such as windfarms on flatlands, offshore and mountain terrains.[13]

However, VAWTs, which are wind turbines that have blades rotating around a vertical axis, perpendicular to the wind direction, also have its advantages. For example, VAWTs can capture wind from any direction, have a lower center of gravity, and are easier to install and maintain. They can also be placed near the ground, where the generator is located, which reduces the need for a tall tower and makes maintenance easier. VAWTs are more suitable for urban areas, where they can reduce the transmission losses by being closer to the demand center, as shown in Figure 3. In addition, VAWTs can also be used in remote areas, streetlights, and households, because they have an independent power generation system. As well as providing power for portable devices, emergency signs in disaster events. Therefore, there is a growing interest in deploying VAWTs in urban areas. However, compared to HAWTs, very few VAWTs are commercially available.[13]

Figure 3 shows the force on the blade in the flap direction for a straight-bladed VAWT. The figure indicates that the rotation not only causes a centrifugal force, but also a bending force from the pivot point. The design of a VAWT mainly consists of the bending strength of the blade. Moreover, the flow field characteristics of a VAWT are different from those of a HAWT. As for a HAWT, the wind flows uniformly into the rotor surface and the torque is constant. But for a VAWT, the wind velocity that flows into the rotor surface is disturbed by the downstream region. This results in a large fluctuation of torque and fatigue loads on the support structure.

1.1. Factors Needed to Consider

'Tip Speed Ratio (TSR)' is generally defined as the ratio of tangential speed of the tip of the blade to the actual velocity of the wind applied. TSR is an important parameter which affects the performance and the efficiency of a typical wind turbine. Its optimistic setup depends on the design of the blade and the wind conditions. Generally, in a relationship to the wind turbine in this case, the VAWTs allows itself to operate in low and turbulent wind, while being an easier assembly in installation and maintenance. While having lower visual and acoustic impact, VAWTs also have some disadvantages such as having a lower power coefficient, being subjected to cyclic stresses, and having complex aerodynamics.[15] The TSR's relationship with the VAWTs is dependent on the shape and the number of the blades installed on the wind turbine.[8] Compared to HAWTs, the VAWTs usually have a lower TSR value as they have higher drag and lower lift coefficients.[12]

The 'Blades' are the one of the main components of the wind turbine which rotates in the wind and captures the kinetic energy of the wind. Being shaped like aerofoils, which creates lift and drag forces when the wind flows over them. The blades' effect on the wind turbine can influence its performance and efficiency in several ways.[7]

The 'Degree' or the 'Pitch Angle' of the blades is the angle between the blade chord and the plane of rotation. The 'Degree' affects the aerodynamic performance and the output of the wind turbine, which determines the angle of attack of the wind applied on the blade and the blade chord. The angle of attack affects the torque and rotational speed of the turbine by influencing the lift and drag forces acting on the blade. The ideal pitch angle is determined by the blade design, wind speed, and tip speed ratio. The ratio of the blade tip speed to the wind speed is known as the tip speed ratio. The form, quantity, length, and profile of the blades are all part of the blade design.Various studies investigated that the 'Degree's effect on the VAWT affects the aerodynamics performance overall [11]. The changes in the 'Pitch Angle' with a constant number of blades on a VAWT changes with a 60° and 120°, found that a 60° 'Pitch Angle' shows the highest 'Power Coefficient' [5].

The 'Power Coefficient' is the ratio of the power extracted by the wind turbine to the total power generated by the wind. It reflects the efficiency and performance of the wind turbine. Several factors influence the power coefficient, such as the wind speed, the tip speed ratio, the pitch angle, and the blade design. The 'Power Coefficient' was optimized using the numerical and experimental methods in a study made by 'Gokhale and Dhatrak'. Using the open source CFD solver, they simulated a 2D VAWT with different pitch angles. It was observed that the 'Power Coefficient' was at its optimal level of 0.18 where with the pitch angle at 0.35 rads. (Lee, 2020) The 'Power Coefficient' is an important factor for it indicates how well the wind turbine converts the wind energy into mechanical energy. It also influences how much energy other wind turbines can generate in a wind farm.[14]

IV. THE APPROACH

The 'Statistics' is a quantitative variable that is derived from a set of data, which is collected and analyzed. This also contains both measure of 'Location' and 'Spread'. The 'Measures of Location' is defined as quantitative values about the center or concentration of data of 'Mean', 'Medium' and 'Mode'.

The 'Mean' can be emphasized as that the average observations (or data) are drawn from the 'Population of all possible measurements.

The 'Medium' is a location of the middle value of a data set. This depends on whether the number of data points is odd or even number. Rearranging the data set in an ascending or descending order, the data points can be easily observable to find the medium.

The 'Mode' is a type of measurement describing the most common or frequent value in the data set collected. The 'Mode' can be used in both numerical and categorical data sets. In detecting the mode of the data set, the specific data set is counted and identified as the highest frequent value.

The 'Measure of Spread' is defined as the variability on the data as it spreads on the model. This shows the importance of technicality on the modelling of the data and presents the trend, showing the difference between them.

The 'Range' is the difference between the largest and smallest value in the data set.The 'Mean Absolute Deviation (MAD)' is the average absolute deviation from the sample mean.The 'Standard Deviation' is the method of measuring the typical distance that values are from the mean.Equation (4.1) shows the formula for 'Standard Deviation'.

$$\sigma = \sqrt{\frac{\sum(x_i - \mu)^2}{N-1}} \tag{4.1}$$

Where μ is the population mean, $x_i$is the i^{th} element from the population and the N presents the population size. At the essential level, the 'Standard Deviation' shows how dispersed the data values are in a dataset along with their corresponding standard deviations.

The 'Variance' is straightforwardly the squared of the 'Standard Deviation' giving the Equation (4.2). It is a statistical concept that measures how much a set of numbers differs from the mean value. It is also a method of quantifying the dispersion of the data.

$$\sigma^2 = \frac{\sum(x_i - \mu)^2}{N-1} \tag{4.2}$$

A. Random Complete Block Design

The Random Complete Block Design is generally used to reduce the effect of having irrelevant variables in an experiment. Dividing the experimental units into blocks and applying each treatment is assigned to them in a random order within each block. Based on one factor, the blocks are created which affect the outcome of the study, and the main aim of using the RCBD application is to ensure each treatment is tested within each block.

In 'Random Complete Block Design', first identify the research problem and experiment's objective. Then,define the dependent (response output) and independent (predictors) factors. Next, determine the treatment levels and number, which are single independent factors for end response comparison and decide the number and size of blocks, which are similar experimental units' groups. Assign randomized treatments to each block's experimental units using Microsoft Excel. This ensures each treatment is tested within each block, balancing out any 'Blocking Factor' bias. Finally, record and analyze the experiment data using 'Two-way Analysis of Variance (ANOVA)' to test for significant differences among treatments, blocks, and between treatment and blocks.
Sum of Squares: Treatments and Blocks

$$\sum_{j=1}^{a} y_{ij} = y_{i.} \qquad i = 1,2,3, \dots \dots, a$$

$$\sum_{i=1}^{b} y_i = y_{.j} \qquad j = 1,2,3, \dots \dots, b$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij} = y_{..} \tag{4.3}$$

Similarly, the $\bar{y_i}.$ averageof the observations taken under treatment i, $\overline{y_{.J}}$, is the average of the observation in block j. $\bar{y}_{..}$, is the grand average of all the observations. That is,

$$\bar{y_i}. = \frac{y_{i.}}{b} \overline{y_{.J}} = \frac{y_{.j}}{a} \bar{y}_{..} = \frac{y_{..}}{N} \tag{4.4}$$

Expressing the total corrected sum of squares as:

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\left(y_{ij} - \bar{y}_{..}\right)^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\left[(y_{i.} - \bar{y}_{..}) + \left(y_{.j} - \bar{y}_{..}\right) + \left(\bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}\right)^2\right] \tag{2.17}$$

This represents the partition of the total sum of squares. This is a fundamental ANOVA equation for the RCBD. Showing the sums of squares in Equation (4.5) symbolically, having

$$SS_{Total} = SS_{Treatments} + SS_{Blocks} + SS_{Error} \tag{4.5}$$

Since there are N number of observations, SS_T has $N-1$ degrees of freedom. There are a treatments and b blocks, so the $SS_{Treatments}$ and SS_{Blocks} have $a-1$ and $b-1$ digress of freedom, respectively. The 'Error' sum of squares is a sum of squares between the cells excluding the sum of squares for treatments and blocks, giving $(a-1)(b-1)$.

The 'Mean Square' is given by the division of the sum of squares of the 'Treatments', 'Blocks' and 'Errors' by their degrees of freedom, as shown in Equation (4.6)

$$MSS = \frac{[SS_{Treatments},\ SS_{Blocks},\ SS_{Error}]}{(N-1)} \tag{4.6}$$

Using the Analysis of Variance for a Randomized Complete Block Design,

The Sum of Squares for Treatments formula shows,

$$SS_{Treatment} = \frac{1}{b}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{N} \tag{4.7}$$

The Sum of Squares for 'Blocks' formula shows.

$$SS_{Blocks} = \frac{1}{a}\sum_{j-1}^{b} y_{.j}^2 - \frac{y_{..}^2}{N} \tag{4.8}$$

The Sum of Square for Total formula gives,

$$SS_{Total} = \sum_{i=1}^{a}\sum_{j-1}^{b} y_{ij}^2 - \frac{y_{..}^2}{N} \tag{4.9}$$

The Sum of Square for Error can be presented as

$$SS_{Error} = SS_{Total} - SS_{Treatment} - SS_{Block} \tag{4.10}$$

The Mean Square for Treatments, Blocks, and Error gives

$$\frac{SS_{Treatment}}{a-1} \tag{4.11}$$

$$\frac{SS_{Block}}{b-1} \tag{4.12}$$

$$\frac{SS_{Error}}{(a-1)(b-1)} \tag{4.13}$$

The F value for Treatment gives,

$$\frac{\left(\frac{SS_{Treatment}}{a-1}\right)}{\left(\frac{SS_{Error}}{(a-1)(b-1)}\right)} \tag{4.14}$$

The same can be applied to the F value for Blocks gives,

$$\frac{\left(\frac{SS_{Block}}{b-1}\right)}{\left(\frac{SS_{Error}}{(a-1)(b-1)}\right)} \tag{4.15}$$

The F-critical value comes from utilizing the F distribution table or using the in-built function from Microsoft Excel, which corresponds to a given level of significance and degrees of freedom for both 'Treatments' and errors variation. This is further used to test for null hypothesis, whether there is no difference among the 'Mean' values of 'Treatments'. The F-critical value depends on the chosen level of significance (usually 0.05 or 0.15), the number of 'Treatments' and 'Blocks'.

The 'Coefficient of Determination' or the 'R^2' is a measure of the conditions of the good fit of the model. In terms of regression analysis, it is a statistical measure of how well the regression line approximates the actual data. This is important to predict the future outcomes or testing the hypotheses. The Equation (4.16) shows the formula of R^2.

$$R^2 = 1 - \frac{Sum\ of\ Squares\ Error\ (SS_{Error})}{Total\ Sum\ of\ Square\ (SS_{Total})} \tag{4.16}$$

The 'Adjusted R-Squared' or 'R^2adj' is the corrected model accuracy of measure for linear models that accounts for predictors that are not significant in a regression model. Deriving from R^2, the adjusted R^2increases the effects of the numbers that were overestimated which can lead to having values lower than R^2 if the effect doesn't improve the model. Normally, adjusted R^2 is always less than or equal to R^2, lying between values of 1 to zero with the model having a perfect prediction or no prediction at all respectively. The Equation (4.17) shows the formula to calculating the adjusted R^2.

$$Adjusted\ R^2 = 1 - \frac{(1 - R^2)\cdot(N-1)}{(N-P-1)} \tag{4.17}$$

Where N is the amount of data in a predictor's column, P is the number of coefficients for which the regression model contains.

The P value is the probability value for which the obtained results of the observed F-value, assuming the null hypothesis, saying that there is no difference among the 'Treatments' is true. The P value is calculated based on the F-value along with the degrees of freedom for 'Treatments' and errors. The smaller the P value, the stronger the statement where the alternative hypothesis is true. In order to find the P value, a 'T' test statistics is required. The test statistics 'T' generally is a number where the number of standard deviations is determined on a probability distribution which utilizes a table shown in the formula of the 'T' test statistic can be found in Equation (4.18).

$$T = \frac{\hat{x} - x_0}{\frac{\sigma}{\sqrt{N}}} \tag{4.18}$$

The $\hat{x}$ is the sample mean and x_0 is the sample value that is estimated by the experimenter, the σ is the standard deviation and the *N* is the number of samples.

B. Multilinear Regression model

When two or more variables are either inversely or directly related, modelling, and exploring the relationship of the said variables, where they can be described. For example, a chemist desired to increase the chemical yield of the product, relating to the operation temperature. The chemist may build a model relating yield to temperature and apply the models for prediction, process optimization or process control. It is more of a statistical technique where two or more independent variables are used to predict the outcome or the product of a dependent variable.

The single dependent variable or the response variable solely depends on the independent variable or the regressor variable which can be variables $x_1, x_2, x_3, \ldots, x_k$. Such relationships can be described by a mathematical model called the 'Regression Model'. 'Regression Model' is used to fit a set of sample data. Using the model as a functional relationship between the response and input variables, the experimenter can select an appropriate function to approximate the regressor since in most cases, the true functional relationship is unknown.

Multilinear regression models create a more detailed analysis with multiple independent variables and their interactions. A standard model can be shown in Equation (4.19).

$$y = x_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon \tag{4.19}$$

The Equation (4.19) is a multilinear regression model with k as regressor variables. This model describes the plane in the k-dimensional space of the regressor variables (x_1). Models that are more complex in appearance compared to Equation (4.20) can sometimes be analyzed by multiple linear regression techniques. The addition of an interaction of the term to the first order in two variables, which leads to the Equation (4.20).

$$y = x_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon \tag{4.20}$$

The Equation (4.20) can be rewritten as Equation (4.21).

$$y = x_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon \tag{4.21}$$

Equations (4.20) and (4.21) are a fixed multilinear regression model with three regressors. β_3 is the regressor coefficient of β_1 and β_2. The same can be said for the x_3. Although with the interactions of the regressor variables can direct to the second order 'Response Surface Model' with two variables, presenting the Equation (4.22).

$$y = x_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \epsilon \tag{4.22}$$

And letting the regressor coefficients and variables to be replaced into Equation (4.23).

$$y = x_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \epsilon \tag{4.23}$$

The method of the least squares is generally the estimation of the regression coefficients in a multiple linear regression model. Suppose there are N observations on the response variable are available, for instance, y_1, y_2, …., y_i. Along with each observed y_i, there is an observation on each regressor variable and let x_{ij} to denote as the i^{th} observation or level of variable x_j. The error term is assumed to be ϵ in the model. The model can be written in the Equation (4.24).

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_N x_{iN} + \epsilon \tag{4.24}$$

The method of least squares chooses the β's in the Equation (4.24) so that the sum of squares of the errors, ϵ_i, is the minimised.

y	x_1	x_2	$\cdots$	x_k
y_1	x_{11}	x_{12}	$\cdots$	x_{1k}
y_2	x_{21}	x_{22}	$\cdots$	x_{2k}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
y_n	x_{n1}	x_{n2}	$\cdots$	x_{nk}

Figure 6: Matrix of the example

In a matrix notation, the development of the normal equation that parallels the development of Equation (4.25), which can be rewritten as

$$y = X\beta + \epsilon \tag{4.25}$$

Where it is expanded as,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \text{ and } \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Figure 7: Expanded Matrix

In general, $\boldsymbol{y}$ is the single column vector of the observations, $\boldsymbol{X}$ is the stated matrix with ($N \times K$) matrix of levels of the independent variables. $\boldsymbol{\beta}$ is also a single column vector of the regression coefficients, and $\boldsymbol{\epsilon}$ is a single column vector of random errors.

Finding the vector of the 'Least Squares Estimators', $\hat{\beta}$, that minimizes

$$L = \sum_{i=1}^{n} \epsilon_i^2 = \epsilon'\epsilon = (y - X\beta)'(y - X\beta) \tag{4.26}$$

Which can be expressed as

$$L = y'y - \beta'X'y - y'X\beta + \beta'X'X\beta$$

$$= y'y - 2\beta'y + \beta'X'y + \beta'X'X\beta \tag{4.27}$$

Due to $\beta'X'y$ is a (1 × 1) matrix, and its 'Transpose' where $(\beta'X'y)' = y'X\beta$ is the same scaler matrix which simplifies to

$$X'X\hat{\beta} = X'y \tag{4.28}$$

Equation (4.28) shows the matrix form of the least square normal equations. Solving the equations, in Equation (4.28), requires multiplying both sides by the inverse of $X'X$, giving the least squares estimator of β in Equation (4.29)

$$\hat{\beta} = (X'X)^{-1}X'y \tag{4.29}$$

This can be alternatively rewritten as the predicted responses as:

$$\hat{y} = \hat{\beta}X$$
$$\hat{y} = X(X'X)^{-1}X'y \tag{4.30}$$

Utilizing the predicted values gained using the model Equation (4.30) the model must undergo 'Model Adequacy Checking', which is an important part of the data analysis procedure. The 'Residual Plots' is used to determine to examine the fitted model to ensure an adequate approximation and verify that the least squares regression assumptions are not violated. This safeguards that the regression model doesn't give poor or misleading results.

The predicted responses can be obtained by multiplying the single column of matrix y in Equation (4.30), gaining the observed responses of the matrix H in Equation (4.31), where H is equivalent to $(X'X)^{-1}X'$.

$$\hat{y} = \hat{\beta}X$$

$$\hat{y} = X(X'X)^{-1}X'y$$

$$\hat{y} = Hy \tag{4.31}$$

The H matrix is called the 'Hat' matrix due to mapping the vectors of the observed values into a vector of fitted value. The residuals (e_i) from the fitted model can be fittingly written in a matrix notion and the covariance matrix of the residuals is shown in Equation (4.32):

$$e_i = y_i - \hat{y}_i$$

$$Cov(e_i) = \sigma^2(I - H) \tag{4.32}$$

The matrix $(I - H)$ isn't normally a diagonal, which the residuals have different variances and correlated. Thus, the variance of the i^{th}residual gives.

$$V(e_i) = \sigma^2(1 - h_{ii}) \tag{4.32}$$

The h_{ii} or the 'Leverage' is the diagonal elements of the H matrix with values between $0 \le h_{ii} \le 1$. The main properties of the 'Leverage' is that it is a measure of the location of the *i^th^* in the *X* space, where the variance of e_i depends on where the values of x_i are located and the sum of all the h_{ii} equals to the number of the parameters or p. Taking the inequality of variance into account when scaling the residuals, plotting the 'Studentized Residuals' is recommended, using the Equation (4.33).

$$r_i = \frac{e_i}{\sqrt{MSS_{error} \cdot (1 - h_{ii})}} \tag{4.33}$$

The 'Studentized Residuals' have constant variance of 1 regardless of the place of the x_i in the matrix. In large data sets, the variance stabilises, which in the case of the 'Standardised Residuals' have the least influence on the least squares fit.

The 'PRESS Residuals' or the prediction error of sum of squares delivers a helpful residual scaling as it finds the error for the i^{th} for each observation and produces an estimate data set. The formula can be seen as:

$$PRESS = \sum_{i=1}^{n} \left(\frac{e_i}{1 - h_{ii}}\right)^2 \tag{4.34}$$

Data points for which h_{ii} are large will have large PRESS residuals, having high influence on the observations. The difference between the ordinary residuals and the PRESS residuals is obviously indicated between the model that fits the data properly and model that predicts poorly.

Finally, the PRESS can be used to compute an approximate Residual, Re^2 for prediction giving:

$$Re^2 = 1 - \frac{PRESS}{SS_T} \tag{4.35}$$

The 'Studentized Residual' isthe outlier diagnostic which is referred to as internal scaling of the residual because MSS_{error} is generated of the estimate of variance obtained from fitting the model to all observations. Another approach would be used to estimate the variance based on the data set with the i^{th} observation removed. The estimate of the variance obtained by S_i^2 in Equation (4.36).

$$S_i^2 = \frac{(n - p) \cdot MSS_{error} - \frac{e_i^2}{(1 - h_{ii})}}{n - p - 1} \tag{4.36}$$

Using the MSS_{error} to produce an externally studentised residual, gives the 'R-Student', *Rs.*

$$Rs_i = \frac{e_i}{\sqrt{S_i^2 (1 - h_{ii})}} \tag{4.37}$$

The R-Student has nearly the same values compared to the studentized residuals but if the i^{th} observation is significant, then the S_i^2 can differ greatly from the MSS_{error} thus making the R-student to be more sensitive to use under standard assumptions which offers a more formal procedure for outlier detection via hypothesis testing

C. Stepwise Selection Regression

1. INTRODUCTION

Stepwise Selection is a procedure where experimenters are interested in the composition of the dataset on which the regressor's predictors are good candidates for the first to show the best relationship either between the response output, y or input predictor variables x_1, x_2, x_3, x_4. A strong correlelation also exists between the predictors x_1 and x_4.

2. THE PROCEDURE

The initial process is to set a significant level for deciding when to enter a predictor into the stepwise selection process. The significant level is denoted as α.

3. Step #1

Fit one of the predictor models or the independent variables, which is regressing the y on x_1 or y on x_2 or y on x_3. Undergoing the RCBD process, finding the P value can compare with the significant level. If it is less than the significant level, then the first predictor is put into the stepwise model. If the predictor has the P value less than the significant level, proceed to step #2.

4. Step #2

Suppose that the x_1 had the smallest P value below α and thus, deemed to be the best single predictor occurring from Step #1. Secondly, fitting each of the two-predictor models that includes x_1 as one of the predictors, which is regressing y on $x_1$and x_2 or y on $x_1$and x_3 and so on. These second predictors whose P value is the smallest are put into the stepwise model. If there are predictors of P value more than the significant level, the process of fitting is stopped. The model with one predictor obtained from Step #1 is the final model. However, consider that x_2 was regarded to be the best second predictor and therefore inputted into the stepwise model. It is now apparent that x_1 is the first predictor in the model, and when the second predictor x_2 enters the stepwise model, the regression model's significance will be affected. This is checked by observing whether the P value on $x_1$is equal to zero and the P value greater than the significant level on x_2.

5. Step #3

Assume both $x_1$and x_2 have reached int the two-predictor stepwise mode and remained without any interactions. Next, fitting the three-predictor model which includes the first two predictors $x_1$and x_2 and regressing y on x_1, x_2 and x_3. If there are more than three predictors, fitting the predictor models with y on x_1, x_2 and x_4 or y on x_1, x_2 and x_5 and so on. Repeating the process on both Step #1 and Step #2, the P value can be gained for the three-predictor model. Of those predictors whose P value is less than the significant level, it is placed in the stepwise model with its P value being the smallest. The opposite when it is higher than the significant level, Step #3 can be disregarded. However, assuming that x_3 was held to be the best third predictor and hence placed into the stepwise model. Since x_1 and x_2 were regarded as the first predictors in the model before initiating Step #3, stepping back and observe whether entering x_3 into the stepwise model affects the significant level of both x_1 and x_2 predictors. Then, examining the P values for both x_1 and x_2 at the three-predictor model, where the third predictor has the P value greater than the significant level, repeat the three-predictor model with another predictor until it is less than the significant level.

6. Step #4

This step includes stopping the entire process after continuously repeating the previous three steps by adding predictors and creating a larger and more complex regression model. This also includes the set number of predictors that can be included during the stepwise selection process.

V. CONDITIONS AND ASSUMPTIONS

To have the methodology undergo the experiment, which is the case study of the 'Effects on the Number of Blades on Aerodynamic Forces on a Straight-bladed Vertical Axis Wind Turbine', it is necessary to have multiple assumptions and consistency throughout the experiment to leave out any discrepancies.
The first variables used to undertake the experiment are 'Tip Speed Ratio', 'Number of Blades', and 'Angle of Blades (Degrees)' where the output is the 'Power Coefficient'.

The first independent variable consists of the 'Number of Blades' includes a definite quantity of the number of blades ranging from 2 to 5 blades. The second independent variable is the 'Angle of Blades (Degrees)' ranging from 4 to 14. The 'Degrees' here are all interacted with the number of blades which are the main conditions to tabulate the 'Tip Speed Ratio'.

In the interaction of both the 'Blades' and 'Degrees', it is important to note that in the number of blades interacting with the number of degrees in the experiment differs greatly as the experimenter had compared the effects of studying the regression models as they had best assumed that the some of the interactions can be left untouched. The interactions shown in Table (1) are where 'Blades' and 'Degrees' relate with each other three times before moving on another iteration. The colored blocks are the main factors which are used to experiment.

Table 1: Interaction of Blades and Degrees

	Degrees					
Number of Blades	4	6	8	10	12	14
2						
3						
4						
5						

VI. BEST RESULTS

Predictor models, useful for finding parameter relationships and making predictions for new data, are based on an equation estimating a dependent variable's value from independent variables. These models are categorized by ascending linear total order, from first to third order and so on. In the 'First order plus Interaction', parameters are multiplied in combinations without squaring. Residual models diagnose predictor model quality and accuracy, improving them by observing residual data set variation in relation to the regressed dataset.

The 'Third Order Only' dataset extends the 'Second Order plus Interaction' with additional parameters '$(TSR)^3$', '$(B)^3$', and '$(D)^3$', reaching cubic levels of 'First Order Only' parameters. This results in a more complex non-linear pattern than simple linear regression, potentially leading to overfitting due to more parameters. Overfitting occurs when the function aligns too closely with the dataset, reducing prediction accuracy. Equations (6.1) and (6.2) illustrate the functional and observational forms of the 'Third Order Only' equations.Figures (9) and (8) establish the 'Third Order Only' dataset which underwent the Stepwise selection and Multiple linear regression processes.

$$Cp = f\begin{pmatrix} TSR, B, D, TSRB, TSRD, BD, TSR^2, B^2, D^2, (TSRB)^2, \\ (TSRD)^2, (BD)^2, (TSR)^3, (B)^3, (D)^3 \end{pmatrix} \quad (6.1)$$

$$\begin{aligned}\hat{y} = \; & x_0 + TSRx_1 + Bx_2 + Dx_3 + TSRBx_4 + TSRDx_5 + BDx_6 + TSR^2x_7 + B^2x_8 + D^2x_9 + (TSRB)^2x_{10} \\ & + (TSRD)^2x_{11} + (BD)^2x_{12} + (TSR)^3x_{13} + (B)^3x_{14} \\ & + (D)^3x_{15} \end{aligned} \quad (6.2)$$

Figure (9) shows significant differencewith residues scattering from start to end. The model's shape, a downward-opening convex parabola, indicates substantial variance in the fitted value for predictions. Residual plots along the 'LOESS' line show the 'Third Order Only' dataset's independent variables improved the model fit.

Figure (8) details the predictor coefficients for the 'Third Order Only' regression model. With 7 degrees of freedom, it's the largest, most complex model. It has the best 'R-Squared Adjusted' of 0.91, accounting for 91% variation in the dependent variable by the 15 independent variables. All predictors have a P value of zero, indicating statistical significance and high correlation with the response variable. This suggests the model interprets all variations in the response variable with no residual error. Despite the unlikelihood of a zero P value, the large datasets fit the regression model perfectly. 'TSR' is the top predictor, followed by '$(TSR)^3$'. 'TSR' interacts only with 'B', making 'Number of Blades' the second-best dataset for variability and correlation growth. Equation (6.3) produces the non-linear regression model for the 'Third Order Only' dataset.

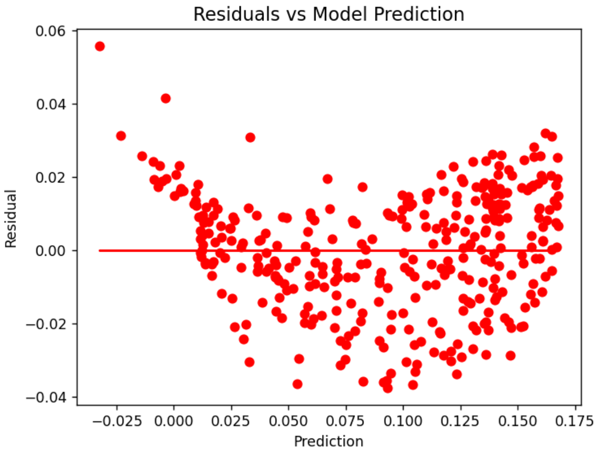


Figure 9: Residual Graph on Third Order only

	Coeff	df	P	R2	R2 Adj
Regression		7	0.00000	0.91	0.91
Intercept	0.5688				
TSR	-0.9556	1	0.00000	0.23	0.21
(TSR)^3	-0.1120	1	0.00000	0.50	0.48
(TB)^2	-0.0150	1	0.00000	0.56	0.55
TB	0.2173	1	0.00000	0.76	0.76
T2	0.5392	1	0.00000	0.83	0.83
Blades	-0.2404	1	0.00000	0.83	0.83
B2	0.0230	1	0.00000	0.91	0.91
Error		337			
Total Error		344			

Figure 8: Diagnosis on Third Order Only.

$$\hat{y} = -0.2113 + (0.2179)(TSR) + (-0.0593)(T^2) + (-0.006)((TSRB)^2) + (0.061)(TSRB) + (0)((BD)^2) \quad (6.3)$$

Figure (10) showcases the experimental data of the ‘Tip Speed ratio’ and ‘the Power Coefficient’ from the experimenting with the wind turbine made of 4 ‘Blades’ with the ‘Angle of Blades’ of 8 degrees with the 5 regression models plotted.

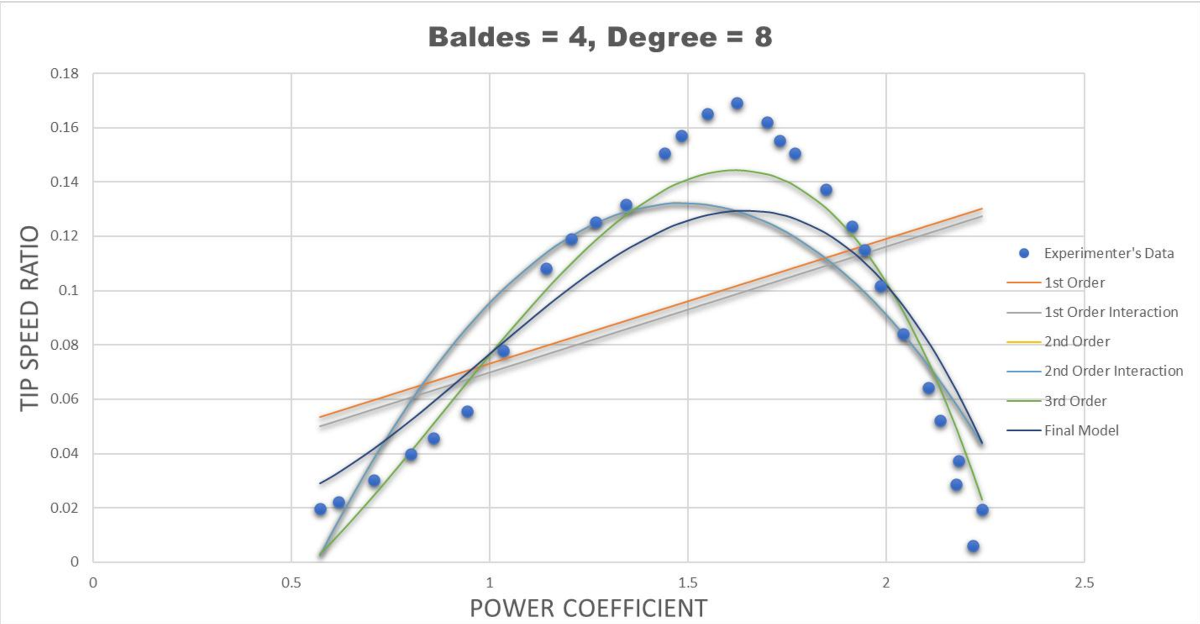


Figure 10: Multiple Linear Regression Model.

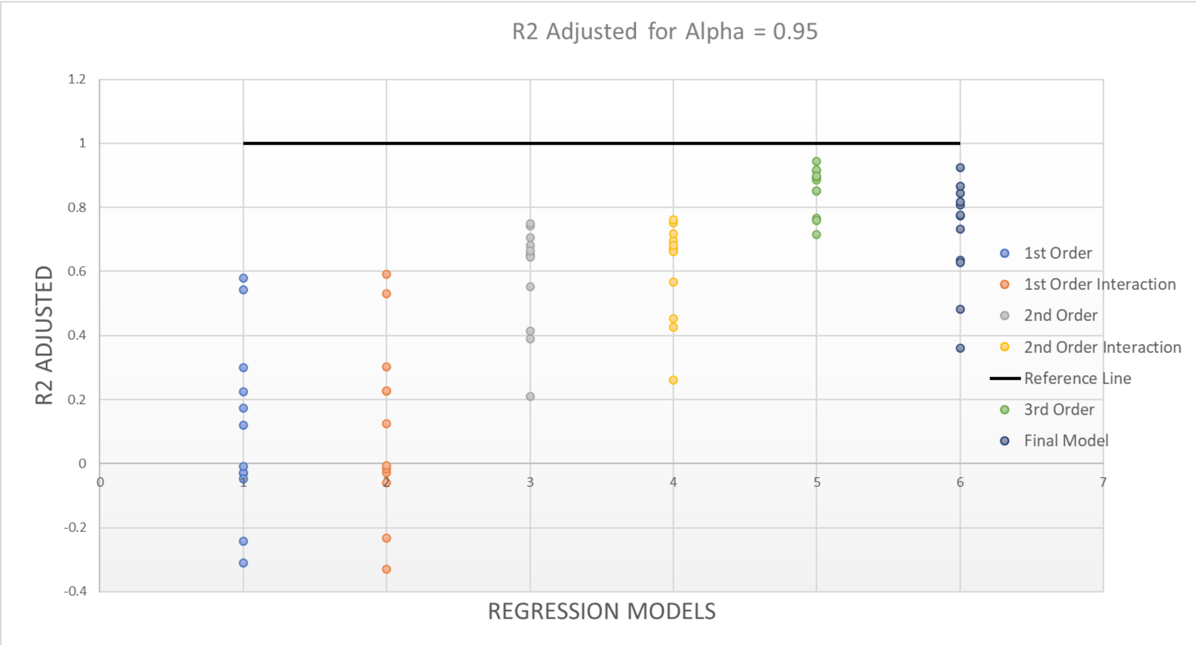


Figure 11: Scatter Plot of the R2 Adjusted data.

Figure (10) shows the multiple linear regression model from the seventh experiment using 4 'Blades' and 8 'Degrees'. The 'Experimenter's Data' is represented as blue dots. The data plots show a steep increase towards a 'Tip Speed Ratio' of 0.18 and a gradual decline towards 0. The '1st Order' and '1st Order Interaction' models progress linearly and parallel to each other. The '2nd Order' and '2nd Order Interaction' models overlap, suggesting a similar non-linear regression model and the same 'R-squared Adjusted'. Their initial regression plot starts above the negative axis, indicating a potential best fit. The '3rd Order' model best fits the 'Experimenter's Data' without any negative values, but it's the second-best fit compared to the 'Final Model', which also stays in the positive quadrant and is closer to the 'Experimenter's Data'. All models stay in the positive quadrant, suggesting reliability for predictions.

	R^2 adj Data					
	α = 0.05	α = 0.05	α = 0.05	α = 0.05	α = 0.05	α = 0.95
	1st Order	1st Interaction	2nd Order	2nd Interaction	3rd Order	Final Model
B2D4	0.224653702	0.227347349	0.649476037	0.664081202	0.852007028	0.774026448
B2D6	0.579410156	0.590745944	0.681669202	0.694932986	0.944310351	0.844378848
B2D8	0.543413793	0.530302438	0.551755043	0.566214558	0.913872662	0.635708239
B3D6	0.299281622	0.302035813	0.74209753	0.752413629	0.889993038	0.924323793
B3D8	-0.029968843	-0.0284986	0.654736921	0.670430697	0.88500903	0.731443519
B3D10	0.119232122	0.125144053	0.705411697	0.717195229	0.850851421	0.864874519
B4D6	-0.028527171	-0.015229329	0.64531453	0.662204314	0.895308021	0.80873705
B4D8	-0.242064124	-0.232285494	0.749331547	0.760725567	0.91726155	0.775084028
B4D10	-0.047674861	-0.058949277	0.663899948	0.68070495	0.897153408	0.816579824
B5D10	-0.009380741	-0.006271685	0.389083365	0.425019638	0.767121588	0.48238008
B5D12	0.172059454	0.227328285	0.208718376	0.261470484	0.758600285	0.627472565
B5D14	-0.310930149	-0.328826098	0.414016102	0.453081696	0.714396019	0.360916133

Figure 12: R2 Adjusted values of different number of Blades and Degrees.

Figure (11) appears to be the scatter plot of the ‘R-Squared Adjusted’ data for all the regression models. Here, it is evident that the ‘First Order Only’ and ‘First Order plus Interaction’ regression models have the similar linear regression models due to having nearly the same ‘R-Squared Adjusted’. However, the ‘R-Squared Adjusted’ values of having lower values being below 0, indicates the regression models do not fit the ‘Experimenter’s Data’ well and it explained the little variance in the ‘Power Coefficient’ by the ‘Tip Speed Ratio’. The First Order Only’ and ‘First Order plus Interaction’ regression models contains at least two or three terms and two or three coefficients each, making linear lines. This entails that the relationship between the ‘Power Coefficient’ and the ‘Tip Speed Ratio’ is not linear and requires more than three terms and coefficients to acquire the appropriate data pattern. These ‘R-Squared Adjusted’ values established that their regression models may be too simple and inadequate to fit the data, as they missed out many important features and traces of the data.

‘Second Order Only’ and ‘Second Order plus Interaction’ regression models have similar analyses to their first order counterparts. However, their ‘R-Squared Adjusted’ values differ by 0.2. The highest ‘R-Squared Adjusted’ values of ‘Second Order Only’ models are slightly less than those of ‘Second Order plus Interaction’. These second order models lack the potential to fit larger datasets due to insufficient terms and coefficients. Despite the unfit relationship between ‘Power Coefficient’ and ‘Tip Speed Ratio’ with quadratic terms and coefficients, the ‘R-Squared Adjusted’ for all second order models are more applicable for predicting less complex, non-linear datasets.

The ‘Third Order Only’ regression, with the highest ‘R-Squared Adjusted’ values, best fits the data but is complex due to its eight parameters. It includes all linear, quadratic, and cubic terms and coefficients, differing in the highest power of independent variables. This collection creates indefinite correlation, increasing model complexity and interpretability, making it hard to understand the ‘Power Coefficient’ and ‘Tip Speed Ratio’ relationship. The numerous predictor coefficients risk overfitting and potential multicollinearity, leading to unreliable predictions. Despite this, ‘Third Order Only’ models are the best alternatives to the ‘Final Model’.

The ‘Final Model’ regression has the second highest ‘R-Squared Adjusted’ values, close to the variance in the ‘Power Coefficient’ by the ‘Tip Speed Ratio’. It’s a complex equation with six coefficients and five parameters, including one or two factors from each linear order, resulting in an ‘R-Squared Adjusted’ value of 0.81. This suggests a highly non-linear relationship between the ‘Power Coefficient’ and ‘Tip Speed Ratio’, requiring a higher curvature or additional parameters for accurate data patterns. With a high significance level of 0.95, this model can fit the data, balancing fit and complexity.

In summary, the ‘Final Model’ regression best fits the ‘Experimenter’s Data’. The ‘Third Order Only’ model has the highest ‘R-Squared Adjusted’ value but struggles to explain the ‘Power Coefficient’ and ‘Tip Speed Ratio’ interactions. Second order models best fit less complex, non-linear data, while first order models poorly fit the current data due to limited explanation of variable relationships.

The relationship between the ‘Power Coefficient’ and ‘Tip Speed Ratio’ can be predicted by the having highly non-linear regression models and requires a higher order polynomial model to gain the predicted data pattern. As shown in Figure (12), the ‘R-Squared Adjusted’ values dropped down dramatically when reaching 5 Blades with a higher angle of blades. This means that with a higher number of blades, the more it affects the power output. The 3 and 4 blades with the same degrees gave the most power output with less power consumption.

Aside from the negative ‘R-Squared Adjusted’ values, one of the conclusions that can be made is that the ‘Tip Speed ratio’ has the most effect on the models, having interacted with that parameter in its quadratic and cubic forms. This shows that the ‘Tip Speed Ratio’ is the most relevant factor affecting the VAWTs as of whole. The ‘Number of Blades’ follows next by its interaction with the ‘Tip Speed Ratio’ parameter. Next, the ‘Angle of Blades (Degrees)’ became irrelevant to the regression model in the stepwise selection process only having an effect in one or two regression models, specifying that the ‘Degree’ factor isn’t too related to the dependent variable.

VII. CONCLUSION

Design of experiments (DOE) is a scientific and systematic approach to planning, conducting, and analyzing experiments. It is used to optimize the performance, quality, and efficiency of a product, process, or system. DOE is also a methodology for problem-solving in various situations, employing complex techniques to collect data and generate reliable engineering solutions. Industries use DOE to reduce costs, increase profits, and grow product yield. By discovering the optimal combinations of factors and levels that affect the output or response of interest, DOE helps gain a competitive edge. For instance, DOE can be applied to increase the yield of wheat during a sunny climate where there is more moisture to help grow. This can be achieved by adding fertilizers, increasing the iteration of watering before it dries up, or adding scarecrows to avoid natural predators from preying on the wheat.

DOE uses the 'Randomized Complete Block Design' (RCBD) to group experimental units into 'Treatments' and 'Blocks'. This method reduces variation and increases precision of response variables. Each 'Block' contains all the 'Treatments' or factor levels, with each 'Treatment' assigned to one experimental unit within each 'Block'. The 'Blocks' are obtained by known sources of variation such as location and time. By blocking, variation in the 'Blocks' is reduced and accounted for in analysis. ANOVA is used to test the significance and estimate the confidence of the 'Treatments' and 'Blocks' effects, as well as compare the contrasts of the treatments. RCBD provides valid and reliable recommendations for modern industries.

Multiple linear regression model is a statistical model that explains the relationship between two or more independent variables and a dependent variable. It fits linear, quadratic, or cubic equations to estimate coefficients from data. The stepwise selection process is used to select variables iteratively based on their significance level, R-Squared, or R-Squared Adjusted values. Its purpose is to maximize the fit and minimize the complexity of the model. Multiple linear regression and stepwise selection processes are used in various fields such as economics, engineering, medicine, and social sciences. They can study factors influencing demand and supply of a product, performance, and efficiency of a mechanical system, diagnostic and treatment of a disease, or behavioral and attitude patterns in a group of people. In short, they help identify relevant variables, test relationships between variables, make accurate predictions, and optimize response variables.

In conclusion, 'Design of Experiments Applications' can change the industry and reality by enhancing the design and optimization of products and processes that can fully deliver a better value, quality, and performance to the customers as well the society. Utilizing the applications, the methods of solving problems lie on statistical notions and prove which solutions can bring the best outcomes. Future work include incorporating machine learning [40-44] and artificial intelligence techniques [45-50]to improve the project outcomes.

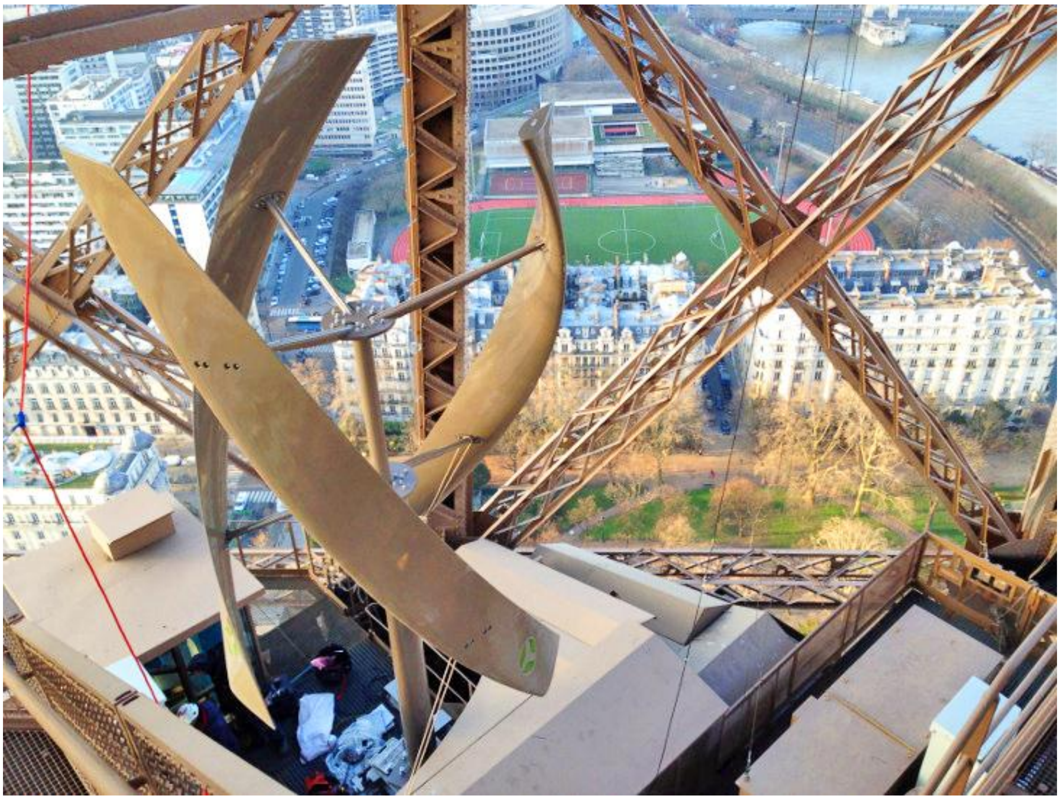


Figure 13: Vertical Axis Wind Turbine Employed in Urban Areas.

REFERENCE

[1]. A.Fisher, R. (1971). The Design of Experiments, The Ninth Edition. New York: Haffner Press.

[2]. Bevans, R. (2020, February 20th). Multiple Linear Regression | A Quick Guide (Examples). Retrieved from Scribbr: https://www.scribbr.com/statistics/multiple-linear-regression/

[3]. Bevans, R. (2023, November 8th). Hypothesis Testing | A Step-by-Step Guide with Easy Examples. Retrieved from Scribbr: https://www.scribbr.com/statistics/hypothesis-testing/

[4]. C.Montgomery, D. (2017). Design and Analysis of Experiments, Ninth Edition. United States: Arizona State University.

[5]. Colin, D. (2017, November 9). Effect of chord-to-diameter ratio on vertical-axis wind turbine wake development. Retrieved from SpringerLink: https://link.springer.com/article/10.1007/s00348-017-2451-6

[6]. Craig, B. A. (2020). Randomized Completed Block Design (RCBD). Indiana: Department of Statistics Purdue University. Retrieved from https://www.stat.purdue.edu/~bacraig/notes514/topic11a.pdf

[7]. Crossley, S. (2012). Wind Turbine Blade Design Review. Nottingham: University of Nottingham. Retrieved from https://research.usq.edu.au/download/2ee7aace5bb507562ff26b3fbdef42ed741397de16927bf64ce7d15a93747def/912186/Blade%20design%20review_Wind%20Engineering.pdf

[8]. Gokhale, G. (2021). Performance analysis of vertical axis wind turbines by varying tip-speed ratio using open source CFD solver. 2ND INTERNATIONAL CONFERENCE ON MANUFACTURING, MATERIAL SCIENCE AND ENGINEERING 2020: ICMMSE 2020, 18-19. Retrieved from https://pubs.aip.org/aip/acp/article-abstract/2358/1/110005/1000837/Performance-analysis-of-vertical-axis-wind?redirectedFrom=fulltext

[9]. Graham Wilson, &. R. (1993). Taguchi methodology within total quality. Bedford: IFS.

[10]. Haq, R. u. (2022, April 10). Randomized Complete Block Design: A Comprehensive Guide. Retrieved from Learn Statistics: http://www.learn-stat.com/randomized-complete-block-design/

[11]. Jan Wisniewski et al. (2021). Wind tunnel comparison of four VAWT configurations to. Poland: Warsaw University of Technology. Retrieved from https://wes.copernicus.org/articles/6/287/2021/wes-6-287-2021.pdf

[12]. Langlang et al. (2019). Maximum Power of Horizontal and Vertical Wind Turbine to Changes on Wind Speed and Number of Blade. Denpasar, Indonesia: IEEE. Retrieved from https://ieeexplore.ieee.org/document/8981448/authors#authors

[13]. Lee, K. (2020, November 09). Sciencing. Retrieved from How Much Poer Does a Wind Turbine Generate?: https://sciencing.com/much-power-wind-turbine-generate-6917667.html

[14]. Oscar et. al. (2023, March 16). MDPI. Retrieved from Comparison of Power Coefficients in Wind Turbines Considering the Tip Speed Ratio and Blade Pitch Angle: https://www.mdpi.com/1996-1073/16/6/2774

[15]. Parker, C. M. (2016). The effect of tip speed ratio on a vertical axis wind turbine at high Reynolds numbers. Springer Link, 74.

[16]. Ross, P. J. (1996). Taguchi techniques for quality engineering : loss function, . New York: McGraw-Hill.

[17]. Singapore: A leading Manufacturing hub. (2018, May 21). Retrieved from EDB Singapore: https://www.edb.gov.sg/en/business-insights/insights/singapore-a-leading-manufacturing-hub.html

[18]. Smith, A. (2015). George Edward Pelham Box. London: Royal Society.

[19]. Songwang. (2023, Auguest 10). Randomized Complete Block Design with and Without Subsamples. Retrieved from Pages.stat.wisc.edu: https://pages.stat.wisc.edu/~songwang/RCBD.pdf

[20]. ST&D. (2023, November 25). Topic 6. Randomised Complete Design (RCBD). Retrieved from Psfaulty.plantsciences: https://psfaculty.plantsciences.ucdavis.edu/agr205/Lectures/2011_Lectures/L6_RCBD.pdf

[21]. UK, N. U. (2023, Novmber 29th). Numercy, Maths, and Statistics - Academic Skill Kit. Retrieved from ncl.ac.uk: https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/regression-and-correlation/coefficient-of-determination-r-squared.html#:~:text=6%20See%20Also-,Definition,line%20approximates%20the%20actual%20data.

[22]. Kok, C.L.; Siek, L. Designing a Twin Frequency Control DC-DC Buck Converter Using Accurate Load Current Sensing Technique. Electronics 2024, 13, 45. doi: 10.3390/electronics13010045

[23]. Kok, C.L.; Chia, K.J.; Siek, L. A 87 dB SNR and THD+N 0.03% HiFi Grade Audio Preamplifier. Electronics 2024, 13, 118. doi: 10.3390/electronics13010118

[24]. Kok, C.L.; Fu, X.; Koh, Y.Y.; Teo, T.H. A Novel Portable Solar Powered Wireless Charging Device. Electronics 2024, 13, 403. doi: 10.3390/electronics13020403

[25]. Aung, K.H.H.; Kok, C.L.; Koh, Y.Y.; Teo, T.H. An Embedded Machine Learning Fault Detection System for Electric Fan Drive. Electronics 2024, 13, 493. doi: 10.3390/electronics13030493

[26]. Chen, J.; Teo, T.H.; Kok, C.L.; Koh, Y.Y. A Novel Single-Word Speech Recognition on Embedded Systems Using a Convolution Neuron Network with Improved Out-of-Distribution Detection. Electronics 2024, 13, 530. doi: 10.3390/electronics13030530

[27]. Don Chua, W.F.; Lim, C.L.; Koh, Y.Y.; Kok, C.L. A Novel IoT Photovoltaic-Powered Water Irrigation Control and Monitoring System for Sustainable City Farming. Electronics 2024, 13, 676. doi: 10.3390/electronics13040676

[28]. Ullah, M. I. (2015, August 25). Randomized Complete Block Design. Retrieved from Basic Statistics and Data Analysis: https://itfeature.com/design-of-experiment-doe/randomized-complete-block-design

[29]. Young, L. S. (2020, March 19). Notation for the Population Model. Retrieved from STAT 501: https://online.stat.psu.edu/stat501/

[30]. Zach. (2020, November 6th). Statology. Retrieved from What is Stepwise Selection? (Explanation & Examples): https://www.statology.org/stepwise-selection/

[31]. Kok, C.L.; Kusuma, I.M.B.P.; Koh, Y.Y.; Tang, H.; Lim, A.B. Smart Aquaponics: An Automated Water Quality Management System for Sustainable Urban Agriculture. Electronics 2024, 13, 820. doi: 10.3390/electronics13050820

[32]. Kok, C.L.; Dai, Y.; Lee, T.K.; Koh, Y.Y.; Teo, T.H.; Chai, J.P. A Novel Low-Cost Capacitance Sensor Solution for Real-Time Bubble Monitoring in Medical Infusion Devices. Electronics 2024, 13, 1111. doi: 10.3390/electronics13061111

[33]. Kok, C.L.; Tang, H.; Teo, T.H.; Koh, Y.Y. A DC-DC Converter with Switched-Capacitor Delay Deadtime Controller and Enhanced Unbalanced-Input Pair Zero-Current Detector to Boost Power Efficiency. Electronics 2024, 13, 1237. doi: 10.3390/electronics13071237

[34]. Kok, C.L.; Tan, T.C.; Koh, Y.Y.; Lee, T.K.; Chai, J.P. Design and Testing of an Intramedullary Nail Implant Enhanced with Active Feedback and Wireless Connectivity for Precise Limb Lengthening. Electronics 2024, 13, 1519. doi: 10.3390/electronics13081519

[35]. Koh, Y.Y.; Kok, C.L.; Ibraahim, N.; Lim, C.G. Smart Water ATM with Arduino Integration, RFID Authentication, and Dynamic Dispensing for Enhanced Hydration Practices. Electronics 2024, 13, 1657. doi: 10.3390/electronics13091657

[36]. Kok, C.L.; Ho, C.K.; Dai, Y.; Lee, T.K.; Koh, Y.Y.; Chai, J.P. A Novel and Self-Calibrating Weighing Sensor with Intelligent Peristaltic Pump Control for Real-Time Closed-Loop Infusion Monitoring in IoT-Enabled Sustainable Medical Devices. Electronics 2024, 13, 1724. doi: 10.3390/electronics13091724

[37]. Kok, C.L.; Ho, C.K.; Lee, T.K.; Loo, Z.Y.; Koh, Y.Y.; Chai, J.P. A Novel and Low-Cost Cloud-Enabled IoT Integration for Sustainable Remote Intravenous Therapy Management. Electronics 2024, 13, 1801. doi: 10.3390/electronics13101801

[38]. Zach. (2020, January 24th). Statology. Retrieved from A Simple Explanation of How to Interpret Variance: https://www.statology.org/a-simple-explanation-of-how-to-interpret-variance/

[39]. Teo, B.C.T.; Lim, W.C.; Venkadasamy, N.; Lim, X.Y.; Kok, C.L.; Siek, L. A CMOS Rectifier with a Wide Dynamic Range Using Switchable Self-Bias Polarity for a Radio Frequency Harvester. Electronics 2024, 13, 1953. doi: 10.3390/electronics13101953

[40]. Kok, C.L.; Ho, C.K.; Tanjodi, N.; Koh, Y.Y. A Novel Water Level Control System for Sustainable Aquarium Use. Electronics 2024, 13, 2033. doi: 10.3390/electronics13112033

[41]. Kok, C.L.; Ho, C.K.; Tan, F.K.; Koh, Y.Y. Machine Learning-Based Feature Extraction and Classification of EMG Signals for Intuitive Prosthetic Control. Appl. Sci. 2024, 14, 5784. doi: 10.3390/app14135784

[42]. Kok, C.L.; Ho, C.K.; Teo, T.H.; Kato, K.; Koh, Y.Y. A Novel Implementation of a Social Robot for Sustainable Human Engagement in Homecare Services for Ageing Populations. Sensors 2024, 24, 4466. doi: 10.3390/s24144466

[43]. C. -L. Kok, Q. Huang, D. Zhu, L. Siek and W. M. Lim, "A fully digital green LDO regulator dedicated for biomedical implant using a power-aware binary switching technique," 2012 IEEE Asia Pacific Conference on Circuits and Systems, Kaohsiung, Taiwan, 2012, pp. 5-8, doi: 10.1109/APCCAS.2012.6418957.

[44]. C. -L. Kok, X. Li, L. Siek, D. Zhu and J. J. Kong, "A switched capacitor deadtime controller for DC-DC buck converter," 2015 IEEE International Symposium on Circuits and Systems (ISCAS), Lisbon, Portugal, 2015, pp. 217-220, doi: 10.1109/ISCAS.2015.7168609.

[45]. L. Battisti, A. A. Castelli, and M. Benini, "Optimization of a Vertical Axis Wind Turbine using DoE technique," IEEE Transactions on Energy Conversion, vol. 24, no. 3, pp. 451-458, 2009.

[46]. P. C. Migliore, W. E. Johnson, and G. E. Waltz, "Design and analysis of a vertical-axis wind turbine," IEEE Transactions on Energy Conversion, vol. 22, no. 1, pp. 69-74, 2007.

[47]. A. S. Bahaj, L. Myers, and P. A. B. James, "Fundamentals applicable to the design of vertical axis wind turbines," IEEE Transactions on Energy Conversion, vol. 21, no. 3, pp. 549-555, 2006.

[48]. M. H. Mohamed, G. Janiga, E. Pap, and D. Thévenin, "Optimization of blade design for a vertical-axis wind turbine using simulated annealing," IEEE Transactions on Energy Conversion, vol. 25, no. 4, pp. 1037-1045, 2010.

[49]. C. Menet, "Design optimization of a vertical-axis wind turbine using Taguchi method and Response Surface Methodology," IEEE Transactions on Sustainable Energy, vol. 4, no. 2, pp. 293-302, 2013.

[50]. R. Howell, N. Qin, J. Edwards, and N. Durrani, "Wind tunnel and numerical study of a small vertical axis wind turbine," IEEE Transactions on Energy Conversion, vol. 26, no. 1, pp. 133-140, 2011.

[51]. W. Tjiu, T. Marnoto, M. Mat, M. Ruslan, and K. Sopian, "Optimization of vertical axis wind turbine performance using Taguchi method," IEEE Transactions on Sustainable Energy, vol. 6, no. 1, pp. 27-34, 2015.

[52]. B. S. P. R. Rao and M. J. V. Sundararajan, "Experimental investigation on the performance of a vertical axis wind turbine with variable aspect ratio," IEEE Transactions on Energy Conversion, vol. 29, no. 1, pp. 8-15, 2014.

[53]. A. S. Archer and M. Z. Jacobson, "Evaluation of global wind power," IEEE Transactions on Geoscience and Remote Sensing, vol. 51, no. 6, pp. 3435-3444, 2013.

[54]. F. Paraschivoiu, "Aerodynamic performance of H-Darrieus vertical axis wind turbines," IEEE Transactions on Energy Conversion, vol. 29, no. 1, pp. 8-15, 2012.

[55]. C. L. Kok, T. H. Teo, Y. Y. Koh, Y. Dai, B. K. Ang and J. P. Chai, "Development and Evaluation of an IoT-Driven Auto-Infusion System with Advanced Monitoring and Alarm Functionalities," 2024 IEEE International Symposium on Circuits and Systems (ISCAS), Singapore, Singapore, 2024, pp. 1-5, doi: 10.1109/ISCAS58744.2024.10558602.

[56]. C. -L. Kok, L. Siek and W. M. Lim, "An ultra-fast 65nm capacitorless LDO regulator dedicated for sensory detection using a direct feedback dual self-reacting loop technique," 2012 IEEE International Symposium on Radio-Frequency Integration Technology (RFIT), Singapore, 2012, pp. 31-33, doi: 10.1109/RFIT.2012.6401604.

[57]. C. Wu et al., "Asymmetrical Dead-Time Control Driver for Buck Regulator," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 24, no. 12, pp. 3543-3547, Dec. 2016, doi: 10.1109/TVLSI.2016.2551321.

[58]. D. Zhu, J. Wang, L. Siek, C. L. Kok, L. Qiu and Y. Zheng, "High accuracy time-mode duty-cycle-modulation-based temperature sensor for energy efficient system applications," 2014 International Symposium on Integrated Circuits (ISIC), Singapore, 2014, pp. 400-403, doi: 10.1109/ISICIR.2014.7029502.

[59]. D. Zhu, J. Wang, C. L. Kok, L. Siek and Y. Zheng, "A new time-mode on-chip oscillator-based low power temperature sensor," 2015 IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC), Singapore, 2015, pp. 411-414, doi: 10.1109/EDSSC.2015.7285138.

[60]. C. -L. Kok and L. Siek, "A novel 2-terminal zener voltage reference," 2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS), Seoul, Korea (South), 2011, pp. 1-4, doi: 10.1109/MWSCAS.2011.6026364.