e-Prime - Advances in Electrical Engineering, Electronics and Energy

A Neural Network-Based Wind Turbine Power Curve Models Using Several Wind Farms' Influencing Parameters and Topography

--Manuscript Draft--

Manuscript Number:	PRIME-D-22-00235
Full Title:	A Neural Network-Based Wind Turbine Power Curve Models Using Several Wind Farms' Influencing Parameters and Topography
Article Type:	Full Length Article
Section/Category:	ENERGY
Keywords:	Artificial neural networks, Quantile based filtering, Radial basis function, Multi-layer perceptron, Wind energy, Wind turbine power curve
Corresponding Author:	Olayinka Soledayo Ohunakin Covenant University Otta, Ogun NIGERIA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Covenant University
Corresponding Author's Secondary Institution:	
First Author:	Olayinka Soledayo Ohunakin
First Author Secondary Information:	
Order of Authors:	Olayinka Soledayo Ohunakin
	Emerald U. Henry
	Victor U. Ezekiel
Order of Authors Secondary Information:	
Abstract:	Wind turbine power curve (WTPC) modelling is of great importance for energy assessment and forecasting. In previous works, WTPC models were developed based on wind speed only. However, in this research, we developed modelling methods that represent actual WTPC by extensively considering wind farms' topography, and several field conditions (other than wind speed only) that are found to influence the power output of wind turbines such as climate variability, the effect of neighbouring wind turbines, turbulence intensity, wake effect, ambient temperature, atmospheric pressure, wind direction, and terrain conditions. We analyze the radial basis function (RBF) and multi-layer perceptron (MLP) architectures for sensitivity and modelling accuracy. A filtered dataset is passed into the models and fitting accuracies are computed alongside sensitivity analysis. The best-performing models are compared with numerous parametric and non-parametric WTPC modeling schemes. It is found that the quantile filtering (QF-NN) models outperforms all other models in terms of fitting accuracy, and outperforms all selected hybrid models in terms of computation time.
Suggested Reviewers:	Muyiwa S. Adaramola, Ph.D Professor, Norwegian University of Life Sciences muyiwa.adaramola@nmbu.no He has worked extensively with the subject area.
	Tunde Bello-Ochende, Ph.D Professor, University of Cape Town tunde.bello-ochende@uct.ac.za He has worked extensively in the subject area.
Additional Information:	
Question	**Response**

To complete your submission you must select a statement which best reflects the availability of your research data/code. IMPORTANT: this statement will be published alongside your article. If you have selected "Other", the explanation text will be published verbatim in your article (online and in the PDF). (If you have not shared data/code and wish to do so, you can still return to Attach Files. Sharing or referencing research data and code helps other researchers to evaluate your findings, and increases trust in your article. Find a list of supported data repositories in Author Resources, including the free-to-use multidisciplinary open Mendeley Data Repository.)	Other (please explain: e.g. 'I have shared the link to my data/code at the Attach File step').
You have selected other. Please explain: as follow-up to "To complete your submission you must select a statement which best reflects the availability of your research data/code. IMPORTANT: this statement will be published alongside your article. If you have selected "Other", the explanation text will be published verbatim in your article (online and in the PDF). (If you have not shared data/code and wish to do so, you can still return to Attach Files. Sharing or referencing research data and code helps other researchers to evaluate your findings, and increases trust in your article. Find a list of supported data repositories in Author Resources, including the free-to-use multidisciplinary open Mendeley Data Repository.)"	I have shared a link to the dataset.
Does your submission include Data in Brief? If so, please upload all Data in Brief files (completed Word template and any relevant data files) as a single zip file, and select "Data in Brief" as File Type.	My submission does not include Data in Brief and click Save Item.

COVERING LETTER

Subject: **Submission of manuscript**

Dear Editor,

I would like to submit the following manuscript for possible evaluation

Manuscript Title: **A Neural Network-Based Wind Turbine Power Curve Models Using Several Wind Farms' Influencing Parameters and Topography**

Name and address of corresponding author:

1. **Olayinka S. OHUNAKIN**

   The Energy and Environment Research Group (TEERG), Mechanical Engineering Department, Covenant University, P.M.B 1023, Ota, Ogun State, Nigeria.

   Email: olayinka.ohunakin@covenantuniversity.edu.ng

Names and addresses of co-authors:

2. **Emerald U. Henry**
   The Energy and Environment Research Group (TEERG), Mechanical Engineering Department, Covenant University, P.M.B 1023, Ota, Ogun State, Nigeria. Email: emerald.henry@stu.cu.edu.ng
3. **Victor U. Ezekiel**
   The Energy and Environment Research Group (TEERG), Mechanical Engineering Department, Covenant University, P.M.B 1023, Ota, Ogun State, Nigeria.
4. **Olaniran J. Matthew**
   Institute of Ecology and Environmental Studies, Obafemi Awolowo University, Ile-Ife, Osun State, Nigeria
5. **Damola S. Adelekan**
   The Energy and Environment Research Group (TEERG), Mechanical Engineering Department, Covenant University, P.M.B 1023, Ota, Ogun State, Nigeria.
6. **Mutali Nepfumbada**
   Harmattan renewables, 24 Cradock Ave, Rosebank, Johannesburg, South Africa

I affirm that the manuscript has been prepared in accordance with your "e-Prime" guide for authors.

I have read the manuscript and I hereby affirm that the content of this manuscript or a major portion thereof has not been published in a refereed journal, and it is not being submitted fully or partially for publication elsewhere.

Thank you.

Olayinka S. OHUNAKIN

Declaration of competing interest

Authors declare that there are no known competing interests.

A Neural Network-Based Wind Turbine Power Curve Models Using Several Wind Farms' Influencing Parameters and Topography

Olayinka S. Ohunakin[a,b], Emerald U. Henry[a], Victor U. Ezekiel[a], Olaniran J. Matthew[c], Damola S. Adelekan[a], Mutali Nepfumbada[d]

[a]The Energy and Environment Research Group (TEERG), Mechanical Engineering Department, Covenant University, Ogun State, Nigeria

[b]Faculty of Engineering & the Built Environment, University of Johannesburg, South Africa

[c]Institute of Ecology and Environmental Studies, Obafemi Awolowo University, Ile-Ife, Osun State, Nigeria

[d]Harmattan renewables, 24 Cradock Ave, Rosebank, Johannesburg, South Africa

Abstract

Wind turbine power curve (WTPC) modelling is of great importance for energy assessment and forecasting. In previous works, WTPC models were developed based on wind speed only. However, in this research, we developed modelling methods that represent actual WTPC by extensively considering wind farms' topography, and several field conditions (other than wind speed only) that are found to influence the power output of wind turbines such as climate variability, the effect of neighbouring wind turbines, turbulence intensity, wake effect, ambient temperature, atmospheric pressure, wind direction, and terrain conditions. We analyze the radial basis function (RBF) and multi-layer perceptron (MLP) architectures for sensitivity and modelling accuracy. A filtered dataset is passed into the models and fitting accuracies are computed alongside sensitivity analysis. The best-performing models are compared with numerous parametric and non-parametric WTPC modeling schemes. It is found that the quantile filtering (QF-NN) models outperforms all other models in terms of fitting accuracy, and outperforms all selected hybrid models in terms of computation time.

Keywords: Artificial neural networks, Quantile based filtering, Radial basis function, Multi-layer perceptron, Wind energy, Wind turbine power curve

Corresponding author: *olayinka.ohunakin@covenantuniversity.edu.ng* ; emerald.henry@stu.cu.edu.ng

1 Introduction

The demand for energy rises astronomically as countries develop [1, 2]. Industrialized nations making up only a fourth part of the world's population, utilize four-fifths of the world's energy. In most developed countries, fossil fuels are the primary source of energy; nevertheless, the usage of renewable energy has increased, and there are on-going plans to replace fossil fuels with renewable energy. Due to environmental consequences associated with fossil fuels utilization (greenhouse gas emissions), rising energy demands with the ever-increasing population, the need to switch to alternative energy sources is inevitable [3-6]. Actions must also be taken to improve the energy efficiency of production and consumption patterns, as well as to encourage the adoption of low-carbon technologies. The deployment of renewable energy technologies has been found to be a significant alternative for reducing GHG emissions, which contribute to global warming [7]. They currently meet 14% of total global energy demand [8]. Wind energy is one of the fastest-growing renewable sources of energy. In the last decade, the number of wind farms increased significantly [9]. The world installed wind energy capacity has improved from 1.29GW in 1995 [10] to 837GW at the end of 2015 [11]. This progressive increase in deployment of the technology, calls for the need for an efficient method of wind energy assessment.

The theoretical power output (P) of a wind turbine (WT) is given by the mathematical expression below:

$$P = \frac{1}{2} C_P(\lambda, \beta) \rho A \upsilon^3 \tag{1}$$

where $C_P(\lambda, \beta)$ is the power coefficient, λ and β are the tip-speed ratio and blade pitch angle respectively, ρ is the air density, A is the rotor swept area, and υ is the wind speed [12]. The main factor influencing the power generated is the wind speed, which exhibits a cubic relationship with the power, and varies significantly with height. Other factors may include wind direction, blade pitch angle, rotor dimensions, air density etc. The wind turbine power curve (WTPC) (Figure 1), is a graph that shows the relationship between a given wind speed, and the electrical power generated for a specific wind turbine; the corresponding electrical power delineates the performance of the wind turbine. The WTPC is specific to a particular brand of wind turbines. The minimum speed at which the turbine delivers useful power is regarded as the cut-in wind speed (V_c). The rated wind speed (V_r) is the wind speed at which the rated power is obtained; this is also the maximum output power of the electrical generator. The cut-out wind speed (V_f) is the highest speed at which a wind turbine may produce power,

as determined by engineering design and safety considerations [13]. The pitch, yaw, and stall control are three methods that are used to control the speed of wind turbine in order to prevent damage [14].

The theoretical power curve of a wind turbine as provided by the manufacturer, is calculated under the ideal, ignoring actual field conditions such as climate variability, effect of neighboring wind turbines and turbine inertia [12]. The actual power curve differs from a single line power curve due to the variables associated with power production on the field. Some of these variables are: turbulence intensity, wake effect, ambient temperature, atmospheric pressure, and wind direction.

1.1 General Objectives of WTPC Modelling

In general, the objectives of wind turbine power curve modelling are briefly discussed below:

1. Choice of wind turbines: Wind generators have their different power output at a given wind speed; hence, WTPC models are a useful tool that will assist wind farm developers to select appropriate generators for maximum efficiency. An important finding in literature is that judicious choice of wind turbine generator that yields higher accuracy can be achieved by utilizing the normalized power curve in the selection process [15].
2. Wind energy assessment and prediction: The wind potential of any site is a function of the prevailing wind speed at a given hub height [14]. Accurate assessments of wind resource are therefore integral to the successful development of wind farms. If the wind speed data of the site is obtained, a WTPC can be utilized for the estimation of the wind energy generated over a period of time [9]. Accurate WTPC models also assist the planning and expansion of wind farms [16].
3. Monitoring and troubleshooting: WTPC can be used for monitoring the performance of wind turbines in a wind farm [17]. A power curve obtained from available data or some methods that represents the performance of the wind turbines under normally operating conditions is used as the ground truth. The actual curve of wind turbines in the wind farm is compared with the ground truth curve. A reasonable deviation between the two curves indicate underperformance and faults [18]. Quality control charts are created from the data obtained to aid the detection and diagnosis of faulty functioning of the wind turbine, or the wind farm [13].
4. Predictive control and optimization: Wind turbines' performances can be assessed using WTPC, and thereafter developing reliable indicators for component diagnostics

and prognostics. Hence, using condition-based monitoring rather than hour-based monitoring can result in higher reliability and cheaper maintenance costs [19].

1.2 Approaches to WTPC Modelling

There are generally four approaches to WTPC modelling. These include discrete approach, stochastic approach, non-parametric or physics-based approach, and parametric or data-driven approach.

1. Discrete methods consist of modelling a continuous process by making discrete approximations [20]. One method is by dividing wind speed into intervals and obtaining a generic power value for each interval. Within this category falls the IEC-61400-12-1 standard. It discretises the wind speed range into bins of 0.5 m/s in size, all the υ and P for each data point within a bin is averaged to obtain the (υ, P) pair for each bin [21]. In this evaluation technique, air density is implicitly considered as input with wind speed and power being the output. This method takes into account the non-linear relationship between υ and P but a large number of measured data is required for improved accuracy [22]. In Llombart et al. [23] modifications were made to the IEC 61400-12 bins method with the aim of improving the accuracy; when compared to the initial bins method, their work achieved improved accuracy.
2. The Markov chain theory is the prevailing stochastic model for the analysis of a wind turbine power output. In Gottschall and Peinke [24], the dynamic behavior of the wind turbine was analyzed with respect to a stochastic signal comprising wind speed and turbulence intensity, thereby yielding power curves independent of turbulence intensity. Another advantage is that the power curve can be obtained within a few days; its disadvantage is that no other parameters are considered other than wind speed and turbulence intensity [20].
3. Parametric models or physics-based models are developed from a set of mathematical equations that includes a set of parameters that must dynamically adapt through a set of continuous data. They are generally developed from linear, non-linear, polynomial and differential equation [20]. In Shokrzadeh [25], a penalized spline regression was used in modelling the WTPC, it resulted in a better performance relative to the other tested models. In Marčiukaitis [26], a non-linear regression functions are used to model the WTPC; the models input is the wind speed while the wind direction is used to vary the curve. Kusiak and Verma [27], in their work scales this method for implementation in the entire wind farm by the utilization of wind direction specific operational curves.

In Pinson et al., [28], a simple and adjustable double exponential function was used in modelling WTPC. A 3-parameter [29], 4-parameter [30], 5-parameter [31], and 6-parameter [32] logistic functions have been used in modelling WTPC. The accuracy of a logistic function was found to improve with increasing parameters. A 9-parameter hyperbolic tangent function has also been employed in Taslimi-Renani et al. [33], in a bid to improve the accuracy. Though parametric models may contain many free parameters that can fit the data, they have a few limitations, they have a fixed form and are rigid, they are complex to use, and they form less accurate wind power curve models compared to their non-parametric counterpart [12].

4. Due to the availability of supervisory control and data acquisition (SCADA) data that enable the access to tremendous amount of data, models that consider the non-linear relationship between features regarded as input parameters and the output power can thus be developed. Hence, rather than a mathematical approach, a data driven approach is required. There are currently a few data approaches available to researchers. From available research, the cubic spline method [34] and computational intelligent (CI) approaches such as the Gaussian process [35], Bayesian-based methods [36], support vector regression methods [10], hybrid relevant vector methods [37], monotonic regression [38], heuristic and metaheuristic methods [31, 39] and artificial neural network (ANN) methods [12, 20, 22, 35]. Generally, non-parametric models are more flexible, precise, and are computationally expensive. Quite a large number of hybrid models have been implemented due to the need for outlier filtration before utilizing dataset for model creation, because non-parametric models are only as good as the available data [35].

   Least square and cubic spline interpolation methods were utilized in modelling the actual power of a wind turbine utilizing wind speed as input in Thapar et al. [34]. This method achieves comparably moderate accuracy and neglects other input parameters that influence the actual power curve. In Manobel et al. [35], a hybrid model based on Gaussian filtering and ANN was employed in modelling the actual power curve. In general, ANN models yield comparably higher accuracies than other CI models because they try to reduce prediction losses through back propagation; however, only wind speed is used as input parameter, and the utilization of hybrid models invariably increase computational cost. The Bayesian method and multi-kernel regression were employed in Wang et al. [36]. The aim of the work was to achieve both a deterministic

and probabilistic power output. A merger of both the deterministic and probabilistic power curve is the ideal power curve for power prediction; however, only the wind speed was utilized in the model creation. In Pandit and Infield [10], a dataset filtered by binning was fed into a support vector regression model with moderate accuracy. Monotonic regression was utilized for WTPC modelling in Mehrjoo et al. [38]. A temperature considerate power curve was obtained in Rodríguez-López et al. [7], utilizing fuzzy logic and ANN. The input parameters are wind speed and ambient temperature yielding a more realistic power curve. However, other influencing parameters were not considered. In Xu et al. [22], a quantile based power curve is obtained by using ANN. Power curves for nine user defined quantiles are obtained and plotted, with the collective defining the possible range of power output. This method is relevant for wind farm monitoring, but not suitable for energy assessment and prediction. It also only takes wind speed as input. A Tabu search non-symmetric fuzzy mean (TS-NSFM) approach combined with a RBF neural network was used in modelling the WTPC in the work of Karamichailidou et al. [12]. Four input parameters were utilized (wind speed, wind direction, blade pitch angle, and ambient temperature), cluster centres and weights were obtained from the TS-NSFM algorithm, and then fed into the RBF network. Modelling accuracy is greatly improved; however, the triple hybrid model is extremely computationally expensive.

We found all models utilized in literature to generally develop power curves without considerations to wind farm topography and field conditions. However, several factors have been found to influence the power output of wind turbines such as climate variability, effect of neighbouring wind turbines, turbulence intensity, wake effect, ambient temperature, atmospheric pressure, wind direction and terrain conditions. This work therefore aims at developing a power curve model that accurately represents the actual field conditions of wind turbines in wind farms by considering all mentioned influencing factors. This work considers climate variability by analysing a time series plot of wind speed over four years to capture all climate conditions. The effect of neighbouring wind turbines, turbulence intensity and wake effect, is achieved by analysing the relative position of turbines on the wind farms to capture the various flow characteristics, while an assessment of the terrain map was carried out to account for disparate land formation. In addition, other than using wind speed as the only variable for the WTPC model as done in past works, we utilized several other variables

alongside wind speed including ambient temperature, wind direction, and atmospheric pressure to develop the model in this work. We further provided an assessment of the accuracy and sensitivity of the two non-parametric approaches including RBF and MLP, that we utilized for the WTPC modelling. The developed model is wind farm specific, generic to all turbines within the wind farm regardless of local orography, and invariant with geographical seasons. Lastly, the accuracy of our models was compared with various parametric and non-parametric models in literature [1, 12, 35, 55], to validate their performances.

2. Material and Methods

2.1 Data Description

Two wind farms were considered for this work including: Kelmarsh wind farm located near Haselbach, Northamptonshire in the UK at latitude 52^{o} 24’ 5.8”, longitude -0^{o} 56’ 34.6” (comprising six 2.05MV Senvion MM92 turbines [40] [see Supplementary data [A], for the dataset]), and Penmanshiel wind farm located near Grantshouse in the Scottish Borders, UK at latitude 55^{o} 52’ 16.2”, longitude -2^{o} 21’16.2” (comprising 14 Senvion MM82’ turbines [41] [see Supplementary data [B], for the dataset]). A 10-minutes SCADA data, ranging from mid-2016 to mid-2021, was obtained for six turbines from the two wind farms (three turbines from each wind farm) for modelling. The turbines were selected based on their relative position on the wind farms: top, midway, and bottom, and based on the terrain in which they fall. Six different turbine SCADA data were utilized for modelling.

A total of four parameters were required for modelling. These include: density normalised wind speed, wind direction, blade pitch angle (three input parameters), and output power (one output parameter). The effect of turbulence intensity (TI) and wake effect was a primary consideration in our turbine selection process. Turbines situated midway and bottom of a wind farm will usually experience incoming turbulence due to wind exiting from preceding wind turbines. However, this phenomenon depends on the topological spacing of wind turbines in the wind farm. The terrain of a wind farm also influences the nature of on-coming experienced by each turbine. Turbines situated in hilly sites will experience more laminar wind at higher speeds while those situated at valley sites will experience less laminar winds at lower speeds. Due to this asymmetry, care must be taken in the selection process to include all terrain conditions.

2.2 Artificial Neural Networks

Artificial neural networks (ANN) are one of the most essential computational intelligence tools, and they're utilized for a number of tasks like function approximation and pattern recognition [12]. ANNs are inspired by biological neural systems. They are made up of a collection of simple processing pieces known as neurons or nodes that are linked together by links, each having a value regarded as their weights. Each neuron can receive information from the neurons that are inputs to them and any other external input, its output information is received as input information into another neuron. Using information inputted into each neuron, they create an output, which is a linear combination of the inputs. Consider Equation (2), y_i is the output of the neuron i, x_j is the input value to the output neuron, indexed as j, w_{ij} is the weight of the link between the i_{th} and j_{th} neurons, θ_i is the bias of the neurons, and f_i is the activation function. This can be linear or non-linear. A non-linear activation function (sigmoid, Gaussian, etc.) is preferable as it allows the network to act more universally [7].

$$y_i = f_i \sum_{j=0}^{n} w_{ij}.x_j - \theta_i \quad (2)$$

Neural networks are divided into feed forward neural networks (unidirectional) or back propagation (recursive) neural networks. For the feed forward type, neurons are connected in a single direction from input to output. In back propagation neural network, the link between neurons is in both directions. The arrangement of the neurons determines the structure of the neural network. Neurons are arranged in a sequence regarded as layers. A typical neural network will comprise of an input layer, 1 to ∞ number of hidden layers, and an output layer. The most peculiar property of a neural network is their ability to learn from the data. The speed at which this learning occurs (steps taken to arrive at a local minimum) is influenced by the learning rate. Learning means adjusting the weights of the connection between neurons in order to improve accuracy [7].

2.3 Radial Basis Function Neural Networks

The RBFNN architecture was first proposed by Broom head, Lowe in their work captioned ‘Multivariate Functional Interpolation and Adaptive Networks’ 1988 [42]. RBFNNs consist of three layers by design: the input layer, the hidden layer, and the output layer [43]. Figure 2 represents the typical structure of a RBF network with a single output node for single value tasks (e.g., regression etc.). The input node distributes the *k* input variables to the *m* nodes of the hidden layer. In the hidden layer, each node has a center with the same dimensions as the number of input variables. The hidden layer applies a non-linear transformation to the input

space, transforming it into a higher-dimensional space [44]. The activity $\mu_l(\mathbf{x}(\mathrm{f}))$ of the lth node is the Euclidean normal of the difference between the f-th input vector and the node center and is given in Equation (3) as:

$$\mu_l\big(\mathbf{x}(\mathrm{f})\big) = ||\boldsymbol{x}(f) - \dot{\mathrm{x}}_{\boldsymbol{l}}|| = \sqrt{\sum_{i=1}^{k} (\boldsymbol{x}(f) - \dot{\mathrm{x}}_{\boldsymbol{l,i}})^2}\,, \quad f = 1, \dots, f \tag{3}$$

where f is the total number of available data, $\boldsymbol{x}^T(f) = [x_1(f), x_2(f), \dots, x_k(f)]$ is the input vector, and $\boldsymbol{x}_l^T = [\dot{\mathrm{x}}_{\mathbf{1}\,\boldsymbol{l}}, \dot{\mathrm{x}}_{\mathbf{2}\,\boldsymbol{l},\dots,\dot{\mathrm{x}}_{\boldsymbol{k}\,\boldsymbol{l}}}]$ is the centre of the lth node.

The activation function for each node is a radially symmetric function. In this work, we employ the sigmoid function (Equation (4)).

$$g(\mu) = \frac{1}{1 + e^{-\mu}} \tag{4}$$

The hidden node response is denoted by $\mathbf{z}(\mathrm{f})$ (Equation (5)):

$$\mathbf{z}(\mathrm{f}) = [g\left(\mu_1\big(\boldsymbol{x}(f)\big)\right), g\left(\mu_2\big(\boldsymbol{x}(f)\big)\right), \dots, g\left(\mu_m\big(\boldsymbol{x}(f)\big)\right)] \tag{5}$$

The output of an RBF network contains y unit, where y is the singular possible output value. The numerical output y(*f*) is produced by a linear combination of the hidden nodes' response (Equation (6)):

$$y(f) = \mathbf{z}(\mathrm{f}).\mathbf{w}_n = \sum_{i=1}^{m} w_{l,n} g\left(\mu_1\big(\boldsymbol{x}(f)\big)\right) \tag{6}$$

where $\mathbf{w}_n = [w_{1,n}, w_{2,n}, \dots, w_{m,n}]^T$ is a vector containing the synaptic weights corresponding to the output *n*.

The synaptic weights are commonly determined using linear regression of the hidden layer outputs to the real measured output after the RBF centers and non-linearities in the hidden layer have been fixed. In most cases, linear least squares in matrix form can be used to solve the regression problem [44].

$$\mathbf{W} = (\boldsymbol{Z}^T.\mathbf{Z})^{-1}.\boldsymbol{Z}^T.\boldsymbol{Y} \tag{7}$$

where $\boldsymbol{Z} = [\boldsymbol{z}(1), \boldsymbol{z}(2), \dots, \boldsymbol{z}(F)]^T$ is a matrix containing the hidden layer responses for all input vectors. $W = [w_1, w_2, \dots, w_n]$ is a matrix containing all the synaptic weights for the output layer and converges to a scalar containing the target vector. The target vector y(f) carries the information of the value predicted by the f-th input vector [44].

2.4 Multi-Layer Perceptron (MLP) Neural Networks

A MLP is a supplement of feed forward neural network. It consists of the basic 3-layers of a neural network, just as in the RBF neural network [45]. They are utilized for generic approximation because they can model any continuous function. A perceptron receives n features as input $x = x_1, x_1, \dots, x_n$, each of these features has an associated weight. The features inputted into the network must be numeric, hence, all non-numeric features must be first converted into numbers before being inputted into the network. The input features are passed on to an input function *u,* the function *u* computes the weighted sum of the input features [46].

$$u(\boldsymbol{x}) = \sum_{i=1}^{n} w_i x_i \tag{8}$$

The result *u*(**x**) is passed onto an activation function *f*, this function assists in producing the output of the perceptron. The activation function utilized in this step is a RELU.

$$y(x) = MAX(\mathbf{0}, x) \tag{9}$$

$$y(x) = \begin{cases} 0 \; for \; x < 0 \\ x \; for \; x \geq 0 \end{cases} \tag{10}$$

Learning in MLPs, as in RBFs consist of adjusting the weights in order to reduce the error in predicting the training data. Learning is a back propagation task achieved by a back propagation algorithm (optimizer), that attempts to minimize the loss in predicting the ground truth.

2.5 Optimizers

These are methods and algorithms utilized by a neural network for back propagation, they assist in adjusting the weights based on a user defined learning rate with the aim of reducing prediction losses. There are varieties of optimizers. We will focus on three types [47]:

1. Stochastic Gradient Descent (SDG): This is a revised version of gradient descent, where the model parameters are updated on every iteration. Gradient descent is an optimization algorithm for finding the local minimum of differentiable function. The basic representation is $\theta_j \leftarrow \theta_j - \alpha \frac{\delta}{\delta \theta_j} j(\theta)$. In gradient decent, we take the whole data for each iteration. However, in SGD we randomly select batches of data. The procedure is to select the initial parameters *v* and learning rate *r*, and thereafter randomly shuffle the data at each iteration to reach an approximate minimum [48].

$$v \coloneqq v - \eta \nabla Q_i(v) \tag{11}$$

2. Adam: It is also called Adaptive Moment Estimation algorithm. It combines root-mean-square propagation and momentum-based gradient descent to achieve better optimization. It introduces two hyper-parameters $\beta_1 \; and \; \beta_2$; their optimal values have been found to be 0.9 and 0.99 respectively [48]. The root-mean-square prop. [$V_{dw} = \frac{V_{dw}}{(1-\beta_1^t)}$] is combined with the corrected momentum-based gradient descent, [$S_{dw} = \frac{S_{dw}}{(1-\beta_2^t)}$] to yield the Adam optimizer.

$$W^{new} = W - \eta \frac{V_{dw}}{\sqrt{S_{dw} + \varepsilon}} \tag{12}$$

3. RMSProp: This is short for root-mean-square propagation. This is an adaptive learning rate method proposed by Geoffrey Hinton. The algorithm focuses on accelerating the optimization process by decreasing the number of function evaluations required to reach the local minima. It retains the moving average of squared gradients for every weight and divides the gradients by the square root of the mean square [49].

$$v(w,t) \coloneqq \gamma v(w,t-1) + (1-\gamma)(\nabla Q_i(w))^2 \tag{13}$$

where gamma is the forgetting factor, an optimal value determined experimentally to be 0.95 weights, are updated by the formula below.

$$w \coloneqq w - \frac{\eta}{\sqrt{v(w,t)}} \nabla Q_i(w) \tag{14}$$

2.6 Evaluation Metrics

These metrics statistically compares the models' output distribution with the ground truth distribution. In this work, we have considered mean absolute error (MAE), root-mean square error (RMSE), coefficient of determination (R^2) given in Equations (15), (16), and (17) respectively:

$$MAE = \frac{1}{n}\sum_{i=1}^{N}(|\acute{p}_i - P_i|) = \frac{1}{n}\sum_{i=1}^{N} AE(i) \tag{15}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\acute{p}_i - P_i)^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} AE^2(i)} \tag{16}$$

$$R^2 = \frac{(a\sum P) + (b\sum XP) - n\acute{P}^2}{(\sum P^2) - n\acute{P}^2} \quad (17)$$

2.7 Methods and Algorithms

A quantile defines a particular part of a data set by assigning each point within a distribution to be either above or below a certain limit [50]. In Figure 3, the distribution for q-quantile plots for all values $a \epsilon S$; the probability that *x* falls within quantile q is given by $P[X < x] \leq k/q$ (where *x* is a *k*-th *q*-quantile for a variable X), and the probability that *x* falls without the quantile q is given by $P[X < x] \geq 1 - k/q$ considering also that x is the *k*-th *q*-quantile for a variable *X*. The distribution is represented mathematically in Equation (18).

$$P[X < x] = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{x} e^{-\frac{t^2}{2}}\, dt \quad (18)$$

The α-th quantile $\theta_\gamma(\propto), 0 <\propto< 1$ of a finite population vector $y = (y_1, \ldots, y_N)$ is defined as

$$\theta_\gamma(\propto) = \inf\{t: F_\gamma(t) \geq \alpha\} \quad (19)$$

where $F_\gamma(t)$ is the distribution function γ. In case $\dot{F}_\gamma(t)$, an estimator of $F_\gamma(t)$, is a monotonic non-decreasing function of *t*, the customary estimator of $\theta_\gamma(\propto)$ is obtained as

$$\theta_\gamma(\propto) = \inf\{t: \dot{F}_\gamma(t) \geq \alpha\} \quad (20)$$

Let $\dot{F}_\gamma(t)$ be the customary estimator of $F_x(t)$. In case the population α-th quantile $\theta_x(\propto)$ of *x* is known, the ratio estimator of $\theta_\gamma(\propto)$ is given by

$$\ddot{\Theta}_{r\gamma}(\propto) = \frac{\ddot{\Theta}_\gamma(\propto)}{\ddot{\Theta}_x(\propto)} \theta_x(\propto) \quad (21)$$

Similarly, a difference estimator of $\theta_\gamma(\propto)$ is given by:

$$\ddot{\Theta}_{d\gamma}(\propto) = \ddot{\Theta}_\gamma(\propto) - R\{\ddot{\Theta}_x(\propto) - \theta_x(\propto)\} \quad (22)$$

where $R = \frac{\sum_{i\in S}\frac{y_i}{\pi_i}}{\sum_{i\in S}\frac{x_i}{\pi_i}}$ is a consistent estimator of the population ratio $R = Y/X$.

Both estimators $\ddot{\Theta}_{r\gamma}(\propto)$ and $\ddot{\Theta}_{d\gamma}(\propto)$ reduce to $\theta_\gamma(\propto)$ if $y_i \propto x_i \forall i \in U$. In this case, the variance become zero [51].

Algorithm 1: Quantile Filtering

Input: **[x_1, x_2, x_3]**: from dataset of initial size, in a Data Frame.

[y]: from dataset of initial size, in the same Data Frame.

Output: **[x_1, x_2, x_3]**: returned Data Frame of filtered size.

[y]: returned in same Data Frame, of filtered size.

1. Divide Data Frame into sub-Frames up to 50.
2. Define a single Data Frame, and set the power equal to the max power.
3. Define the distribution.
4. Apply quantiles to the distribution for each sub-Frame to remove outliers
5. Merge all Data Frames.
6. **END**

2.8 Network Architecture

The base parameters of the neural networks are specified for the network creation. A total of three neural networks are developed and compared for the best performing networks. It comprises of an RBF network with a single hidden layer, two variations of MLP network, one with 4-hidden layers of varying nodes, the other with 6-hidden layers of varying nodes. A link to the code is provided for model visualization [see Supplementary data [C]]. Figure 4 articulates the entire modelling process. The cleaned data is passed into a neural network containing an optimizer that assists in the back propagation of the model. It evaluates the output of the forward process by some user-defined metrics, it then employs the user specified activation function in an attempt to reduce loss.

2.9 Case Study

2.9.1 Data Selection

A temporal 10-minutes interval dataset from mid-2016 to mid-2022, was acquired from Kelmarsh and Penmanshiel wind farms in the UK (see Supplementary data [A], and [B], for the dataset). Out of these years, only a single year with time (dynamic year) can be used in developing the model. A plot of wind speed with time for all available years are plotted to aid visual identification of the most dynamic year (Figure 5). The choice of the year to be used depends on how randomly distributed wind speed to time appeared for all available years. The year with the highest standard deviation of wind speed, found to be 2020, is thus taken to be the most dynamic year (Figure 5a). Figure 5b shows the variation of wind speed to time for the year 2020.

It was discussed in Karamichailidou et al., [12], that the variations in TI due to field conditions such as relative position of turbines, local orography, and wake effects of neighbouring wind turbines, have a significant effect on the power output. Hence, for a generic wind turbine model, the effect of turbulence intensity must be considered. We utilized SCADA data from three turbines selected from each of the two wind farms. The three turbines were

selected with the aim of capturing power variations resulting from TI, as influenced by the relative positions of the turbines, and topography of the respective farms. Based on the most prevalent wind direction for the year, turbines at the top of the wind farm will experience little or no TI, thus resulting in limited/no reduction in power output, while wind turbines at midway of the wind farm will experience the most power reduction due to TI and wake effect, irrespective of the wind direction. However, wind turbines at the bottom of the wind farm will experience some power reductions due to TI. An area view (by satellite) of the wind farms and turbine numbers used in the selection of the three turbines is given in Figure 6.

Terrain variation is also an important factor to be considered in wind turbine model development. As height increases, wind speeds tend to be faster and more laminar. At hill sites, the wind will be laminar because there is little obstruction to air flow, whereas at valley sites there tends to be slower and more turbulent wind speed. Hence, topographic consideration is critical to selecting the best turbine data from SCADA, needed for modelling. Figure 7 represents the terrain maps of the selected wind farms, indicating hilly and valley portions needed for the dataset selection process. It can be seen that the contour maps in Figure 7 show very little variation in land topography. The topographical variations will constitute only a slight variation in wind speed from turbine to turbine within the selected wind farms.

2.9.2 Data Visualization

A plot of wind speed to power generated is the basis for this analysis. Figure 8 shows a scatter plot of wind speed to power output for the six turbines (from the two wind farms) used for model development. It is evident from Figure 8 that the actual curve of these wind turbines deviates from the standard manufacturer's power curve. Generally, the actual power curve of a wind turbine will contain abnormal values (outliers), especially from parameters such as wind speed values close to cut-in or cut-out speed, negative power values, non-operational periods, invalid, missing or corrupted data due to sensor and pitch malfunction, dirt or icing on blades, etc. [52]. Hence, to aid the development of an accurate model, a data cleaning process must be employed to filter out all erroneous data.

2.9.3 Data Cleaning

There are several methods employed in literature for the filtration of erroneous data. However, most are computationally expensive due to their utilization of complex mathematics. In this work, we proposed a filtering approach which is based on quantiles set on a distribution. This is achieved by dividing data into wind speed bins and setting user defined quantiles. Quantile values are obtained experimentally and may differ from one wind farm to another. A

link to the database containing the code has been provided to aid the visualization of the set quantiles [see Supplementary data [C]]. The output (filtered power curve) from the filtering algorithm is shown in Figure 9.

A test of the efficiency of any filtering algorithm is the quantity of data points it regards as outliers in its process of cleaning the dataset. A more efficient filtering algorithm will filter out fewer data points in order to achieve a clean dataset, while a less efficient filtering algorithm will filter out more data points in order to achieve a clean dataset. In literatures, models utilized for data cleaning are static, in that upon implementation, the data points filtered out is constant. However, in this work, outputs of the quantile filtering (QF) method vary the amount of filtered data points contingent on user pre-set quantiles. The amount of data points cleaned out based on the quantiles set developed in this research, is shown in Table 1.

After filtering, the dataset passes a stage of batch normalization. This standardizes the dataset between 0 and 1. The efficiency and speed of a neural network is enhanced when we convert the variable data to values between 0 and 1 for the fitting process. Normalized data are then split into training, testing and validation datasets. In this research, the dataset from each turbine was divided randomly using 65% for training, 25% for validation, and 20% for testing. This is shown in Table 2.

2.9.4 Implementation of Algorithm

The visual analysis of the actual wind speed to power plot of a wind turbine is the basis for the choice of the most optimum quantile for filtering. An assessment of the distribution of the dataset must be carried out during the specification of quantiles. A skewed distribution of the dataset will constitute poor filtering even with the most generic quantile; it thus requires a rather unique quantile values that should be selected based on the nature of its distribution. For the quantiles used in this research, a link to the database containing the code is provided [see Supplementary data [C]]. The filtered data comprising three input parameters (density normalized wind speed, wind direction, and blade pitch angle) are fed into three neural networks, with the aim of accurately predicting the power output. These input parameters were found to improve the modelling accuracy of power curve models, as each constitute an influence on the power output of a wind turbine [53, 54]. A RBF network of 1-hidden layer, a MLP network of 4-hidden layers, and a MLP network of 6-hidden layers were used for modelling. There is an increase in the number of hidden layers for the second MLP network; this was done with the aim of inspecting the performance of MLP network with more hidden layers. The architectures are indicated in Table 3.

3. Results and Discussion

From Table 4, data 1 to 3 are obtained from Kelmarsh wind farm, while data 4 to 6 are from Penmanshiel wind farm. The differences in the accuracies between datasets from the same wind farm is due to the variation of the data distribution, while the differences recorded between datasets from different wind farms are as a result of the differences in field conditions at both farm sites. An analysis of the data distribution was performed within our code using the seaborn library.

The field conditions at Kelmarsh wind farm are generally more dynamic than that at the Penmanshiel wind farm. Hence, the fitting accuracy is slightly less for the Kelmarsh wind farm than the Penmanshiel wind farm. The test for a good model is its generalization ability. The standard deviation of the datasets from both wind farms is minimal, thus indicating that the model generalizes appropriately. There is a slight variation in performance when various optimizers are applied on the RBF network. The best optimizer for this network is the *adam* optimizer as it has shown better performance on all the datasets compared to the other optimizers (though not far behind). All optimizers performed generally well on the RBF network. In this work, the RBF network is generally appropriate for modelling WTPC regardless of field conditions. Data 1 and 4 are obtained from wind turbines at the top of the wind farms where TI has negligible effect on power output. Data 2 and 5 are obtained from wind turbines midway the wind farms, and are mostly affected by TI and wake effect. Data 3 and 6 are obtained from wind turbines at the bottom of the wind farms; these are somewhat influenced by TI and wake effect. In this work, it is worth noting that we obtained similar fitting accuracies regardless of the position of the wind turbine.

We ran the models on Google Collaboratory (12.63GB GPU), for 50 epochs. Figure 10 depicts the loss to epoch plot for each optimizer. The model arrived minimum loss within just a few epochs, hence, more time could thus be saved by reducing the number of epochs. The MLP model is generally sensitive to optimizers, only the *adam* optimizer performed appropriately. It can be seen from Figure 11 that the *rmsprop* and *sgd* optimizers worked to increase the loss rather than to decrease it. Therefore, care should be taken in choosing the right optimizer for an MLP model.

The variations found in the dataset due to field conditions, do not affect the accuracy of the MLP model. Taking the values of the *adam* optimizer as a case study, the values of both MAE and R^2 vary only slightly across *data* 1 to 6. This is an indication that the MLP network generalizes properly for all turbines within a wind farm, and for turbines in multiple wind

farms. The best performing case of the 4-input MLP network outperforms all cases in the RBF network. Generally, an increase in the number of hidden layers of an MLP network reduces the sensitivity of the model to change in optimizer. However, it also slightly increases the models' sensitivity to dataset distribution. This can be observed in Table 6 with the standard deviation of the 6-hidden layers of the MLP network being relatively more than that of the 4-input MLP network. The 6-hidden layers MLP network generally requires more time to train, but results in higher fitting accuracy than the other models. From Figure 12, it can be seen that the network accommodates the use of the *rmsprop* optimizer but not the *sgd* optimizer.

Table 7 summarises the results of the application of QF-NN algorithm in comparison with the best performing parametric and non-parametric models found in literature, and with comparison to the manufacturer's power curve. From Table 7, QF-NN models outperform slightly better than all other parametric and non-parametric models in terms of fitting accuracy. Generally, the standard deviation of the QF-NN's fitting accuracy, is found to be relatively small. It is also evident from Table 7, that hybrid non-parametric models utilizing computational intelligence (CI), generally perform better than parametric or other non-parametric models. However, they are computationally expensive and will take a longer time to compute. QF-NN proves to be an excellent alternative to other hybrid non-parametric models, in that it outperforms them in terms of fitting accuracy and also computes in very little time regardless of the size of the data. This is important if these models are to be implemented for real time power prediction and forecasting. The best performing model is the QF-MLP with 4-hidden layer using adam optimizer, having MAE of 15.3 and R^2 of 0.996. Though it is the best performing model, other NN models are also found to be close in performance.

An overall analysis on the QF-NN algorithm is necessary in order to justify the research objectives. The input variables of the models are density, wind speed, temperature, and blade pitch angle. In the literature, very few works have utilized more than one input variable, this was mainly due to the methods employed for power curve modelling that accommodates just one input variable. It is worth noting that models developed under such conditions cannot accurately represent actual power curves at field conditions. The four variables considered in this research are those that majorly contribute to the power output of a wind turbine.

Quantile filtering only considers wind speed and power output in the cleaning process because it is a two-dimensional cleaning process requiring 2-parameters to define the plane. The two-dimensional cleaning is sufficient to filter out outliers, and it does that in less time than a higher dimensional cleaning algorithm. Passing the cleaned data into a shallow neural

network harnesses the powers of the network for efficient fitting of the input parameters to the output power in a supervised framework. We have also compared the accuracy and sensitivity of RBF and MLP networks used for the WTPC modelling.

4. Conclusion

In this work, we introduced a new method of obtaining a very accurate WTPC model, by taking into account several variables (found to affect power output from a wind turbine), other than wind speed only. These variables are density, wind speed, temperature, and blade pitch angle. The datasets were obtained from Kelmarsh and Penmanshiel wind farms in Haselbach and Grantshouse, UK, respectively. Since wind turbine performance vary with TI, we considered the topography of the locations to aid the selection of appropriate turbines to be selected for the modelling.

We utilized quantile filtering approach, which is a data cleaning method where a set point or quantile represents relevant data points (right of the set point), and outlier data points (left of the set point). The outlier data are cleaned out. The filtered data is passed to three models namely: RBF, QF-MLP with 4-hidden layers, and QF-MLP with 6-hidden layers. Their results were compared in order to obtain the best performing model for WTPC modelling, based on shallow neural network. We also compared the performance of the newly developed model with parametric and non-parametric models found in literatures [see ref. 1, 12, 35, 55]. We found that the QF-MLP with 6-hidden layers, gave the best performance and also the most sensitive model to dataset distribution. The QF-MLP networks with 4- and 6-layers are sensitive to the selected optimizer. The RBF network is the most stable network because it has the least standard deviation between the datasets. The three models outperform all parametric and non-parametric models in terms of fitting accuracy, and they also compute in tremendously reduced time than the other hybrid models (i.e., TS-NSFM, SFM).

We have developed a WTPC model that is generic to all turbines in wind farms, taking topography and other field conditions into consideration during the dataset selection and modelling.

Declaration of competing interest

Authors declare that there are no known competing interests.

Supplementary materials

Supplementary material for this this article is found at:

[A] Kelmarsh wind farm data. https://zenodo.org/record/5841834#.YsGnSHbMLIU

[B] Penmanshiel Wind Farm Data. https://zenodo.org/record/5946808#.YsGnSXbMLIU

[C] Link to the code: https://github.com/henrii1/wind-turbine-power-curve-modelling.git

References

[1] Wang Y, Hu Q, Li L, Foley AM, Srinivasan D. Approaches to wind power curve modeling: A review and discussion. Renew. Sustain. Energy Rev 2019; 116 (109422). doi: 10.1016/j.rser.2019.109422.

[2] Ohunakin O. S. Adaramola M. S. Oyewola O. M. Wind energy evaluation for electricity generation using WECS in seven selected locations in Nigeria. Applied energy 2011; 88 (9): 3197-3206.

[3] Ohunakin O. S. Akinnawonu O. O. Assessment of wind energy potential and the economics of wind power generation in Jos, Plateau State, Nigeria. Energy for sustainable Development 2012; 16 (1): 78-83.

[4] Ohunakin O. S. Wind resource evaluation in six selected high-altitude locations in Nigeria. Renewable Energy, 2011; 36 (12): 3273-3281.

[5] Flint L. *Wind energy.* 1979; 29 (5).

[6] G. Bölük G, Mert M. Fossil & renewable energy consumption, GHGs (greenhouse gases) and economic growth: Evidence from a panel of EU (European Union) countries. *Energy 2014;* 74: 439–446.

[7] Rodríguez-López M, Cerdá E, del Rio P. Modeling wind-turbine power curves: Effects of environmental temperature on wind energy generation. *Energies 2020;* 13 (18): 1-21.

[8] United Nations Development Programme. Energy *and The Challenge of Sustainability*. 2000. Available online: https://www.undp.org/sites/g/files/zskgke326/files/publications/World%20Energy%20Assessment-2000.pdf. Accessed [03/07/2022].

[9] Mehrjoo M. Using Machine Learning Methods for Wind Turbine Power Curve Modeling. 2021.Available online: https://mspace.lib.umanitoba.ca/handle/1993/35697. Accessed [03/07/2022].

[10] Pandit R. Infield D. Comparative analysis of binning and support vector regression for wind turbine rotor speed based power curve use in condition monitoring. *Proc. - 2018 53rd Int. Univ. Power Eng. Conf. UPEC 2018*; 1–6. doi: 10.1109/UPEC.2018.8542057.

[11] Global Wind Energy Council. Global Wind Report 2022. Available online: https://gwec.net/global-wind-report-2022:text:Total global wind power capacity, carbon emission of South America. Accessed [03/03/2022]

[12] Karamichailidou D. V. Kaloutsa V. Alexandridis A. Wind turbine power curve modeling using radial basis function neural networks and tabu search. *Renew. Energy 2021;* 163: 2137–52.

[13] Lydia M. Kumar S. S. Selvakumar A. I. Prem Kumar G. E. A comprehensive review on wind turbine power curve modeling techniques. *Renew. Sustain. Energy Rev. 2014 2014;* 30: 452–460.

[14] Manwell J. F. McGowan J. G. Rogers A. L. Wind Energy Explained (Theory, Design and Application). Second Edition. 2010. ISBN: 978-0-470-01500-1

[15] Jangamshetti S. H. Guruprasada Rau V. Normalized power curves as a tool for identification of optimum wind turbine generator parameters. *IEEE Trans. Energy Convers. 2001;* 16 (3): 283–8.

[16] Norgaard P. Holttinen H. A Multi-Turbine Power Curve Approach. *Nord. Wind Power Conf.* 2004; 1–2. Available Online: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:A+Multi-Turbine+Power+Curve+Approach#0.

[17] Sohoni V. Gupta S. C. Nema R. K. A Critical Review on Wind Turbine Power Curve Modelling Techniques and Their Applications in Wind Based Energy Systems. *J. Energy 2016*; 4: 1–18. doi: 10.1155/2016/8519785.

[18] Kusiak A. Zheng H. Song Z. On-line monitoring of power curves. *Renew. Energy 2009;* 34 (6): 1487–93.

[19] Uluyol O. Parthasarathy G. Foslien W. Kim K. Power curve analytic for wind turbine performance monitoring and prognostics. *Proc. Annu. Conf. Progn. Heal. Manag. Soc. 2011;* 435–442.

[20] Pelletier F. Masson C. Tahan A. Wind turbine power curve modelling using artificial neural network. *Renew. Energy 2015;* 89: 207–214.

[21] IEC 61400-12-1:2017 RLV. Wind energy generation systems - Part 12-1: Power performance measurements of electricity producing wind turbines. 2005; 1–8.

[22] Xu K. Yan J. Zhang H. Zhang H. Han S. Liu Y. Quantile based probabilistic wind turbine power curve model. *Appl. Energy 2021;* 296: 116913. doi: 10.1016/j.apenergy.2021.116913.

[23] Llombart A. Watson S. J. Llombart D. Fandos J. M. Power curve characterization I: Improving the bin method. J. *Renew. Energy Power Qual. 2005;* 1 (3): 367–71.

[24] Gottschall J. Peinke J. Stochastic modelling of a wind turbine's power output with special respect to turbulent dynamics. *J. Phys. Conf. Ser.* 2007; 75 (012045): 1-8.

[25] Shokrzadeh S. Jafari Jozani M. Bibeau E. Wind turbine power curve modeling using advanced parametric and nonparametric methods. *IEEE Trans. Sustain. Energy 2014;* 5(4): 1262–69.

[26] Marčiukaitis M. Žutautaitė I. Martišauskas L. Jokšas B. Gecevičius G. Sfetsos A. Non-linear regression model for wind turbine power curve. *Renew. Energy 2017;* 113: 732–741.

[27] Kusiak A. Verma A. Monitoring wind farms with performance curves. *IEEE Trans. Sustain. Energy 2013;* 4(1): 192–9.

[28] Pinson P. Nielsen H. A. Madsen H. Robust estimation of time-varying coefficient functions-Application to the modeling of wind power production. *Math. Model 2007;* 57 (1) 1–36. Available Online: http://www2.compute.dtu.dk/~ppin/docs/robust_IntProgProject_v4.pdf.

[29] Hernández-Escobedo Q. Manzano-Agugliaro F. Zapata-Sierra A. The wind power of Mexico. *Renew. Sustain. Energy Rev. 2010;* 14 (9): 2830–40.

[30] Sohoni V. Gupta S. Nema R. A comparative analysis of wind speed probability distributions for wind power assessment of four sites. *Turkish J. Electr. Eng. Comput. Sci.* 2016; 24 (6): 4724–35.

[31] Lydia M. Selvakumar A. I. Kumar S. S. Kumar G. E. P. Advanced algorithms for wind turbine power curve modeling. *IEEE Trans. Sustain. Energy 2013;* 4(3): 827–835.

[32] Villanueva D. Feijóo A. Comparison of logistic functions for modeling wind turbine power curves. *Electr. Power Syst. Res. 2018;* 155: 281–8.

[33] Taslimi-Renani E. Modiri-Delshad M. Elias M. F. M. Rahim N. A. Development of an enhanced parametric model for wind turbine power curve. *Appl. Energy 2016;* 177: 544–552.

[34] Thapar V. Agnihotri G. Sethi V. K. Critical analysis of methods for mathematical modelling of wind turbines. *Renew. Energy 2011;* 36(11): 3166–77.

[35] Manobel B. Sehnke F. Lazzús J. A. Salfate I. Felder M. Montecinos S. Wind turbine power curve modeling based on Gaussian Processes and Artificial Neural Networks. *Renew. Energy 2018;* 125: 1015–1020.

[36] Wang Y. Hu Q. Meng D. Zhu P. Deterministic and probabilistic wind power forecasting using a variational Bayesian-based adaptive robust multi-kernel regression model. *Appl. Energy 2017;* 208: 1097–1112.

[37] Jing B. Qian Z. Wang A. Chen T. Zhang F. Wind Turbine Power Curve Modelling Based on Hybrid Relevance Vector Machine. *J. Phys. Conf. Ser. 2020;* 1659 (1): doi: 10.1088/1742-6596/1659/1/012034.

[38] Mehrjoo M. Jafari Jozani M. Pawlak M. Wind turbine power curve modeling for reliable power prediction using monotonic regression. *Renew. Energy 2020;* 147: 214–222.

[39] Goudarzi A. Davidson I. E. Ahmadi A. Venayagamoorthy G. K. Intelligent analysis of wind turbine power curve models. *IEEE Symp. Comput. Intell. Appl. Smart Grid, CIASG 2015;* doi: 10.1109/CIASG.2014.7011548.

[40] Plumley C. Kelmarsh wind farm data (0.0.3) [Data set]. 2022, doi: https://doi.org/10.5281/zenodo.5841834.

[41] Plumley C. Penmanshiel Wind Farm Data (0.0.2) [Data set]. 2022, doi: https://doi.org/10.5281/zenodo.5946808.

[42] IRWINE S. J. C. WARD M. C. MULLIN J. B. Royal Signals and Radar Establishment. *Emerg. Technol. Situ Process 2012;* 139 (4148): 221.

[43] Yang Y. Wang P. Gao X. A Novel Radial Basis Function Neural Network with High Generalization Performance for Nonlinear Process Modelling. *Processes 2022;* 10 (1): 1–16.

[44] Alexandridis A. Chondrodima E. A medical diagnostic tool based on radial basis function classifiers and evolutionary simulated annealing. *J. Biomed. Inform. 2014;* 49: 61–72.

[45] Abirami S. Chitra P. Chapter Fourteen-Energy-efficient edge based real-time healthcare support system. Advances in Computers 2020; 117 (1): 339-368.

[46] Menzies T. Kocagüneli E. Minku L. Peters F. Turhan B. Using Goals in Model-Based Reasoning. *Shar. Data Model. Softw. Eng.2015;* 1: 321–353. 2015, doi: 10.1016/b978-0-12-417295-1.00024-2.

[47] https://www.kdnuggets.com/2020/12/optimization-algorithms-neural-networks.html#:~:text=Optimizers are algorithms or methods,problems by minimizing the function.

[48] https://medium.com/analytics-vidhya/this-blog-post-aims-at-explaining-the-behavior-of-different-algorithms-for-optimizing-gradient-46159a97a8c1.

[49] https://www.analyticsvidhya.com/blog/2021/10/a-comprehensive-guide-on-deep-learning-optimizers/.

[50] https://www.statista.com/statistics-glossary/definition/356/quantile/#:~:text=A quantile defines a particular,) and percentiles (hundredth).

[51] Arnab R. Estimation of Distribution Functions and Quantiles. *Surv. Sampl. Theory Appl. 2017;* 747–772. doi: 10.1016/b978-0-12-811848-1.00023-6.

[52] Park J. Y. Lee J. K. Oh K. Y. Lee J. S. Development of a novel power curve monitoring method for wind turbines and its field tests. *IEEE Trans. Energy Convers. 2014;* 29 (1): 119–128.

[53] Alexandridis A. Chondrodima E. Giannopoulos N. Sarimveis H. A fast and efficient method for training categorical radial basis function networks. *IEEE Trans. Neural Networks Learn. Syst. 2017;* 28(11): 2831–36.

[54] Shetty R. P. Sathyabhama A. Pai P. S. Comparison of modeling methods for wind power prediction: a critical study. *Front. Energy 2020;* 14(2): 347–358.

[55] T. Li, X. Liu, Z. Lin, and R. Morrison, “Ensemble offshore wind turbine power curve modelling – An integration of isolation forest, fast radial basis function neural network, and metaheuristic algorithm,” Energy, vol. 239, p. 122340, 2022, doi:

10.1016/j.energy.2021.122340.

Figure

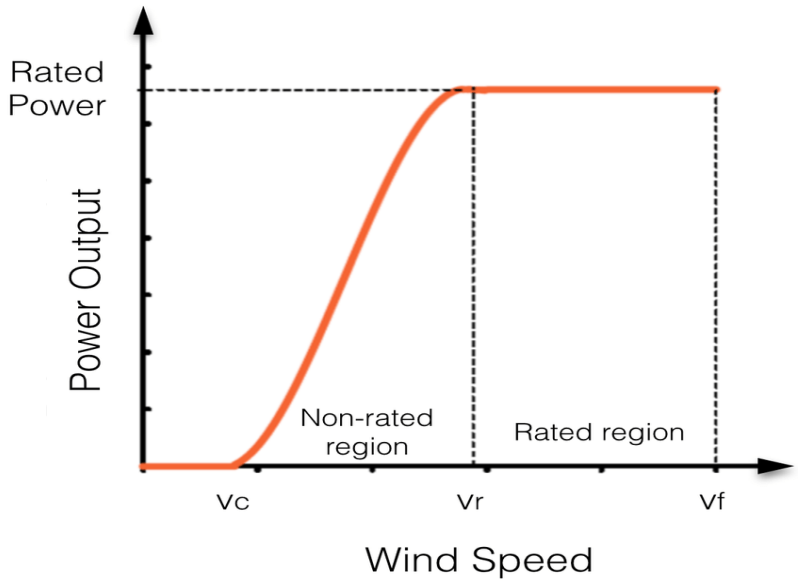


Figure 1: A Typical Wind Turbine Power Curve

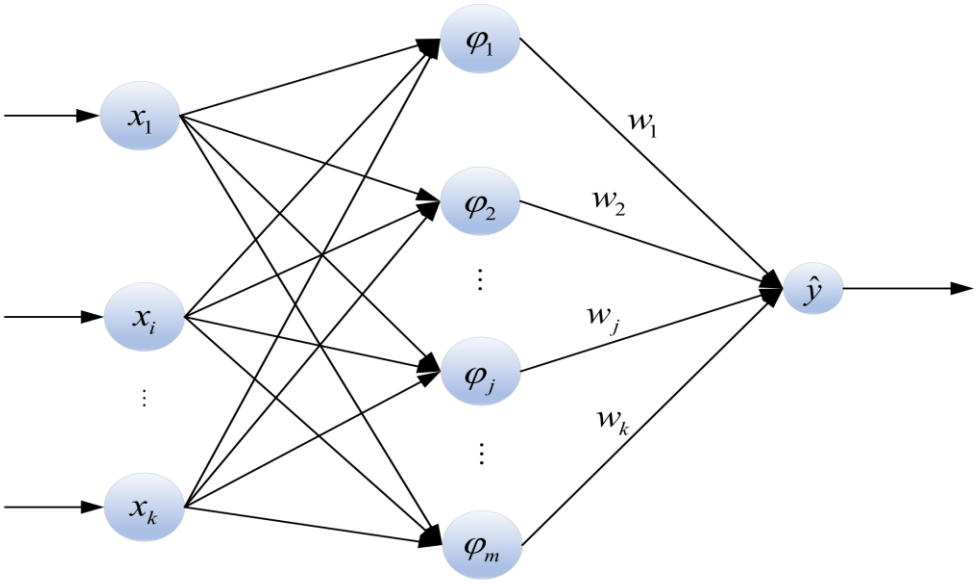


Figure 2: Typical Structure of an RBFNN

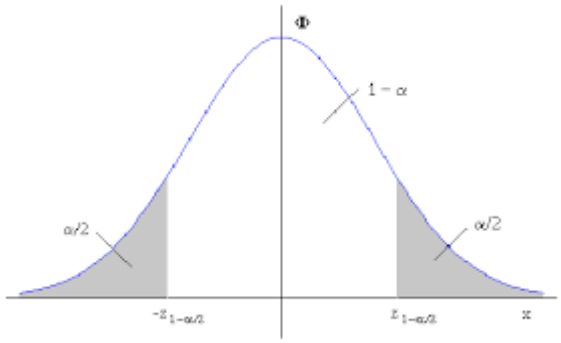


Figure 3: Normal Distribution for Quantile Specification

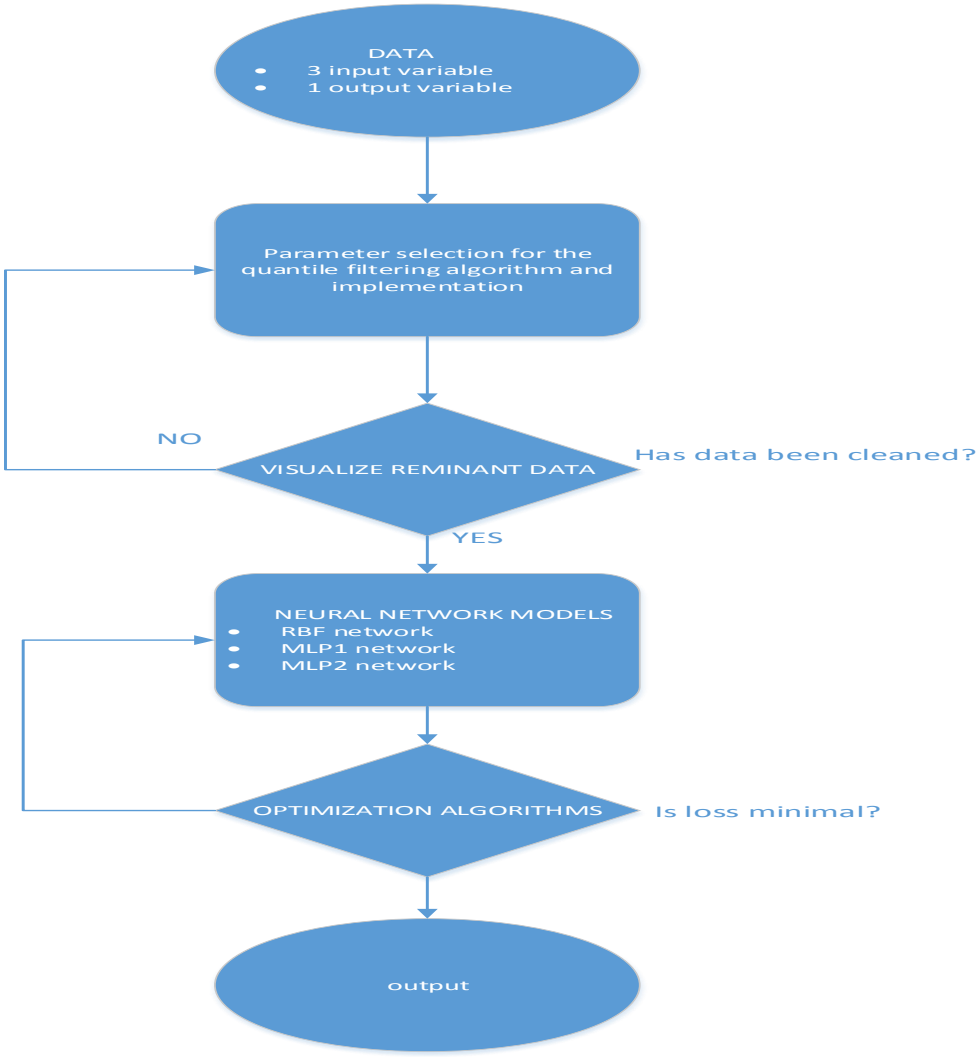


Figure 4: QF-NN Algorithm

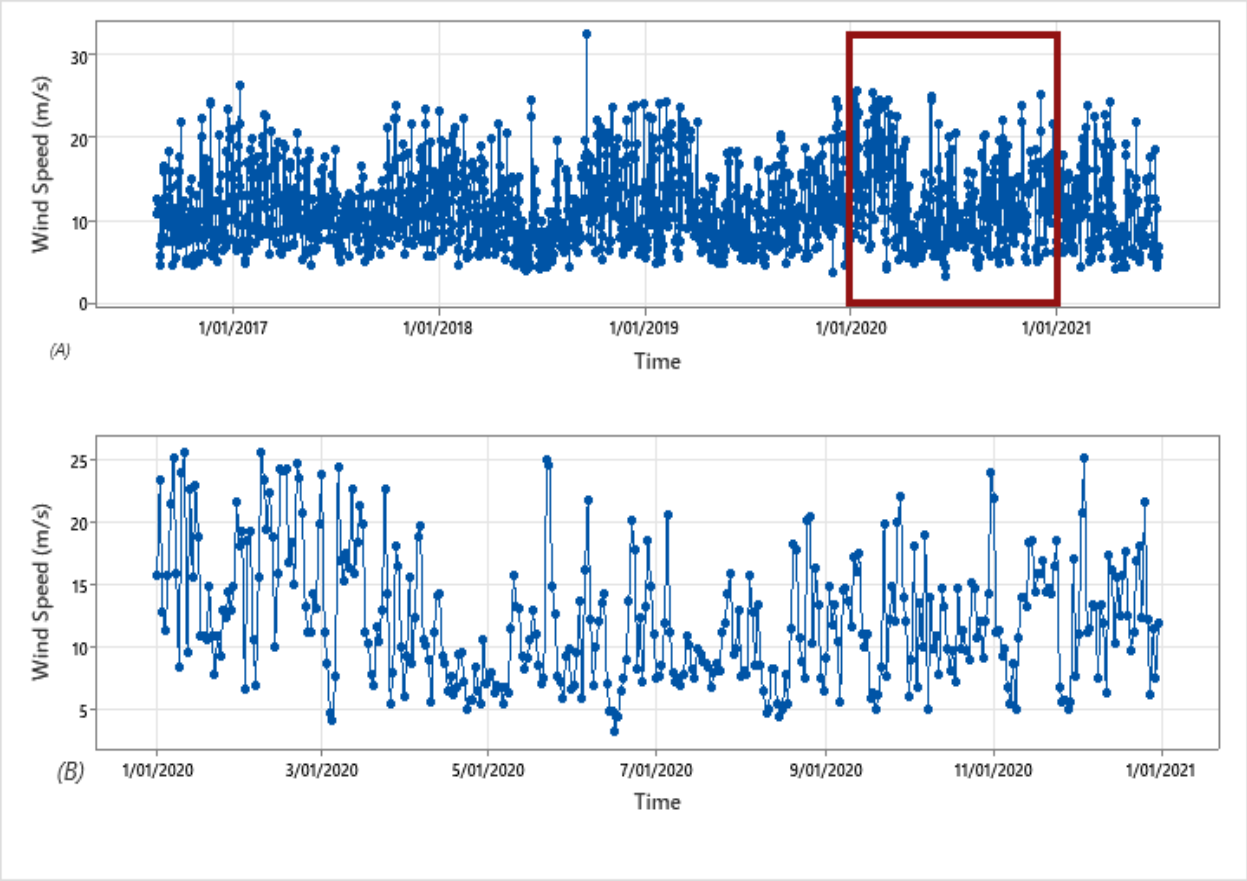
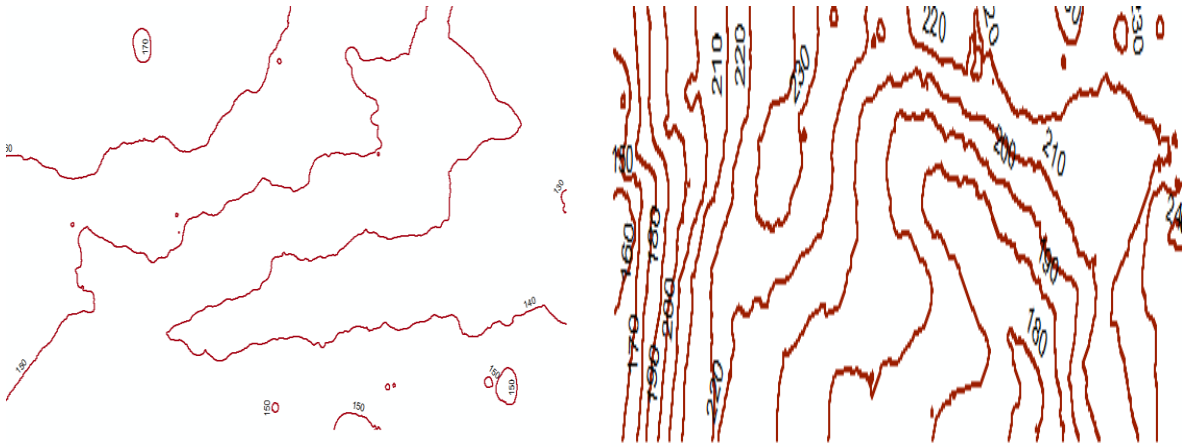


Figure 5: Variation in wind speed with time for: (a) various years, and (b) selected year

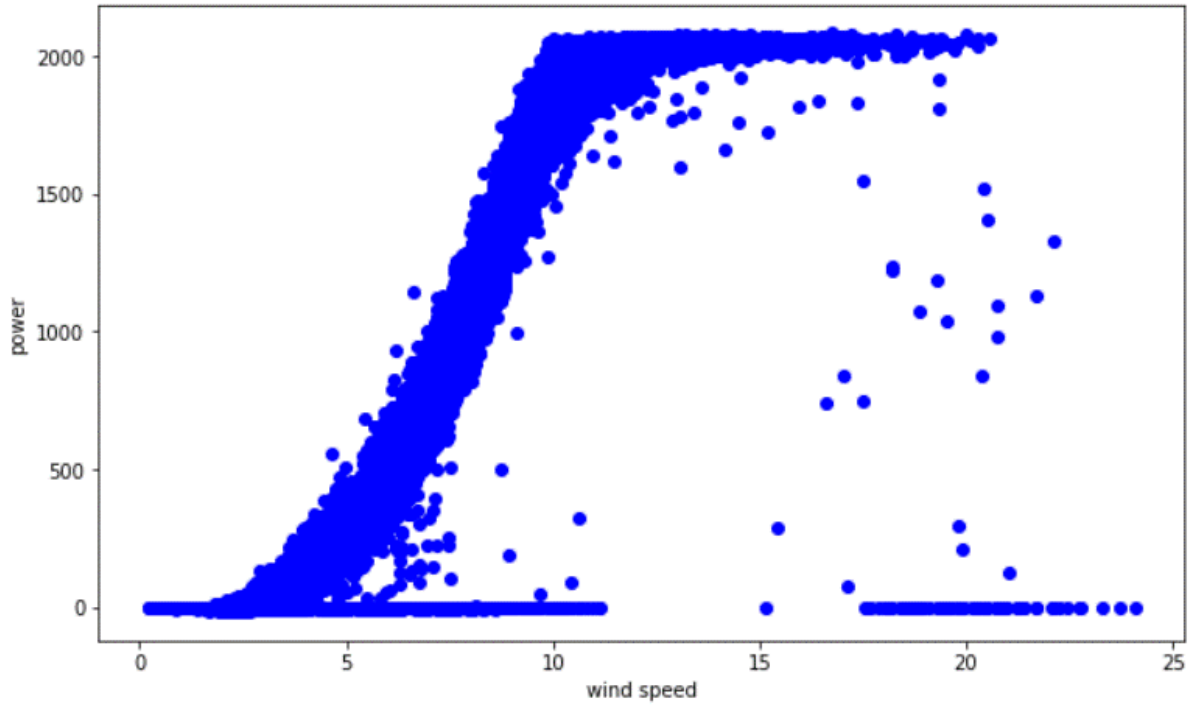
(a) (b)

Figure 6: Satellite Image of (a) Kelmarsh and (b) Penmanshiel Wind Farms showing the turbine numbers (Source: ref. [37, 38])

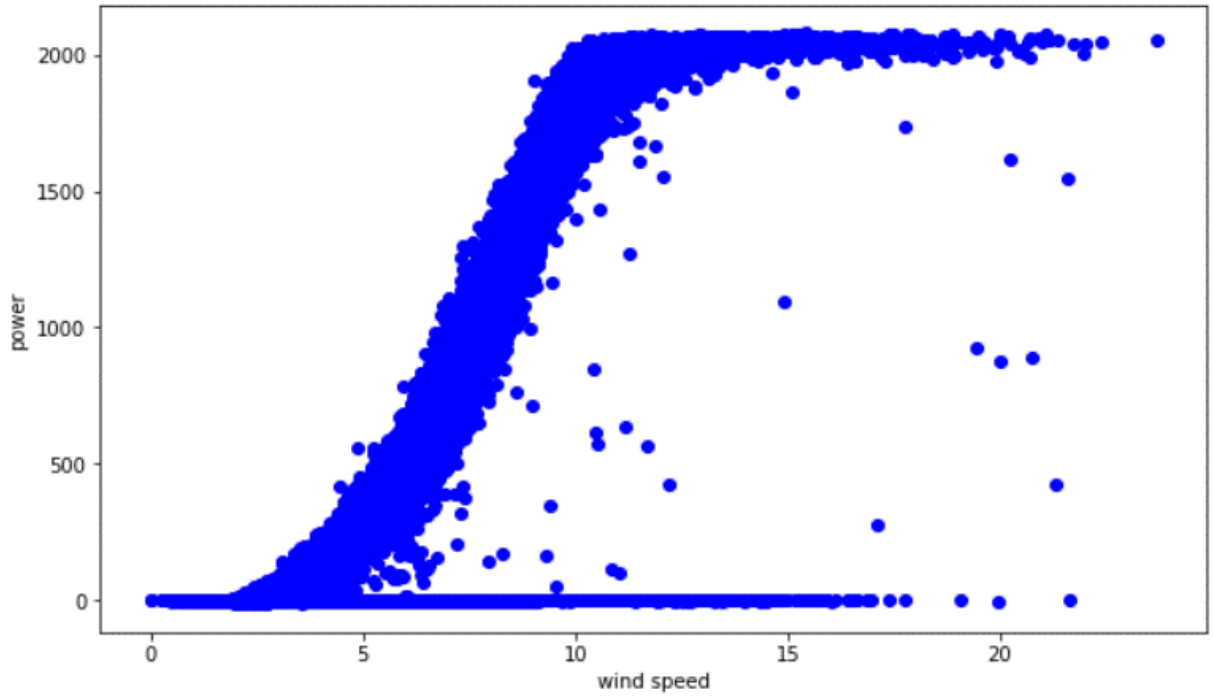


(a) (b)

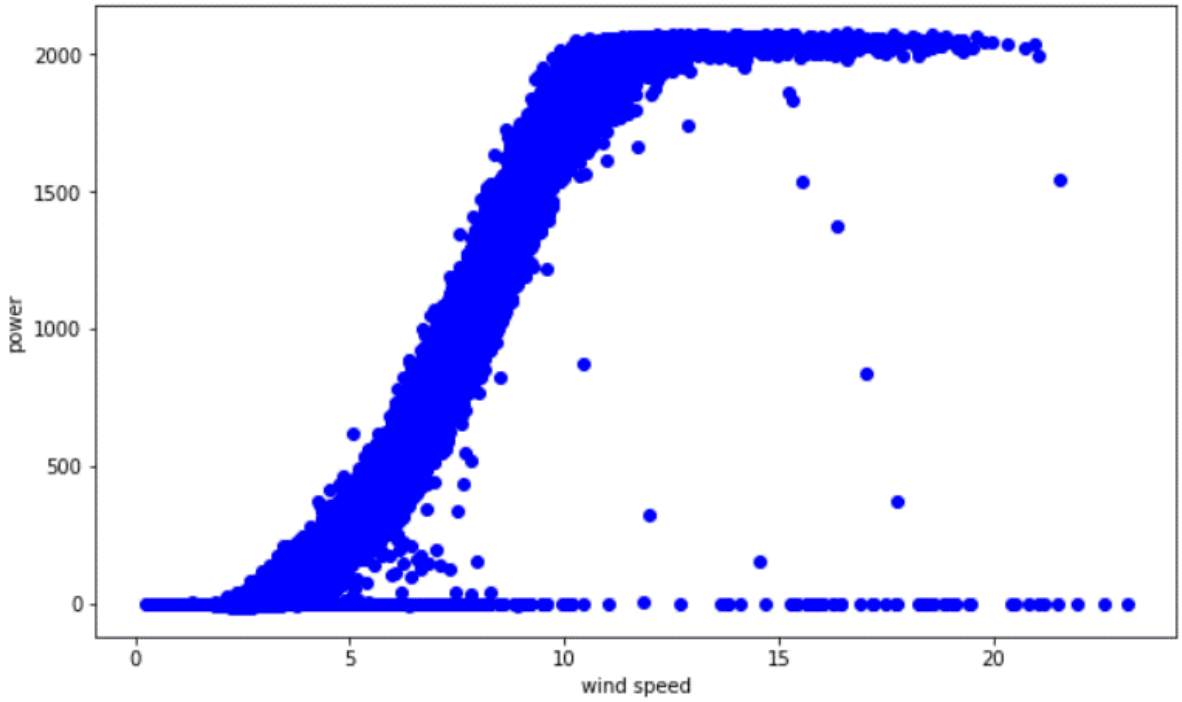
Figure 7: The Developed Contour Maps for (a) Kelmarsh, and (b) Penmanshiel Wind Farms



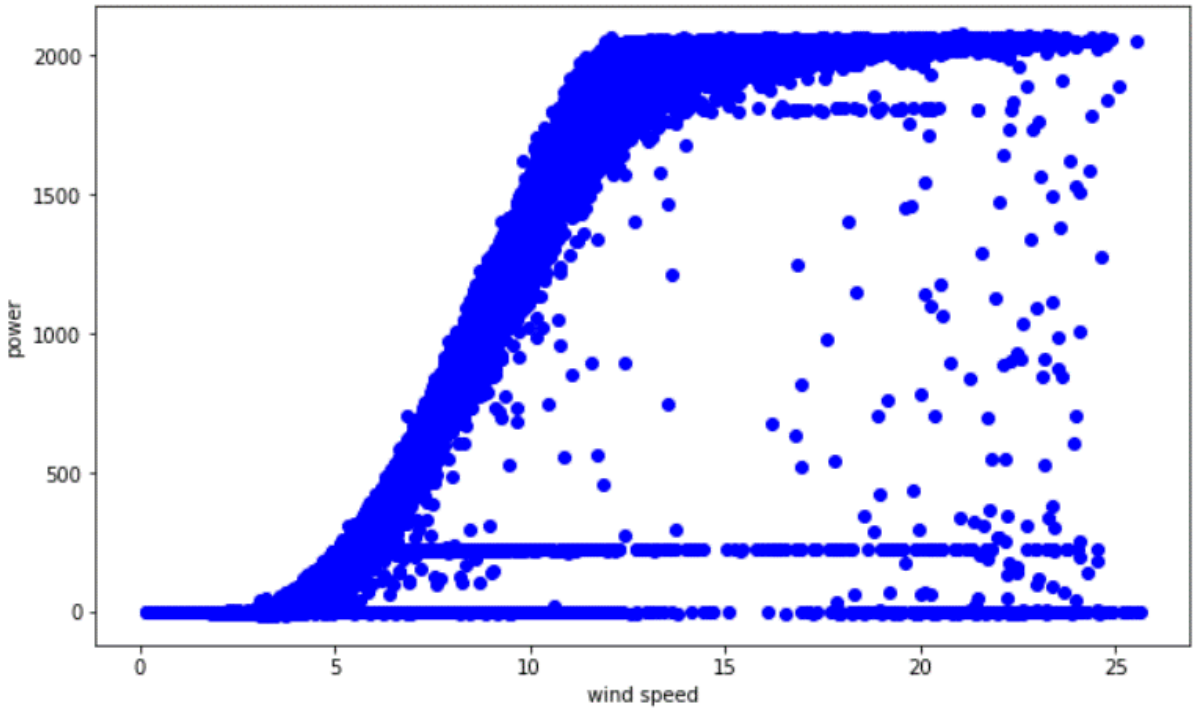
a)


power
wind speed
0
500
1000
1500
2000
0
5
10
15
20

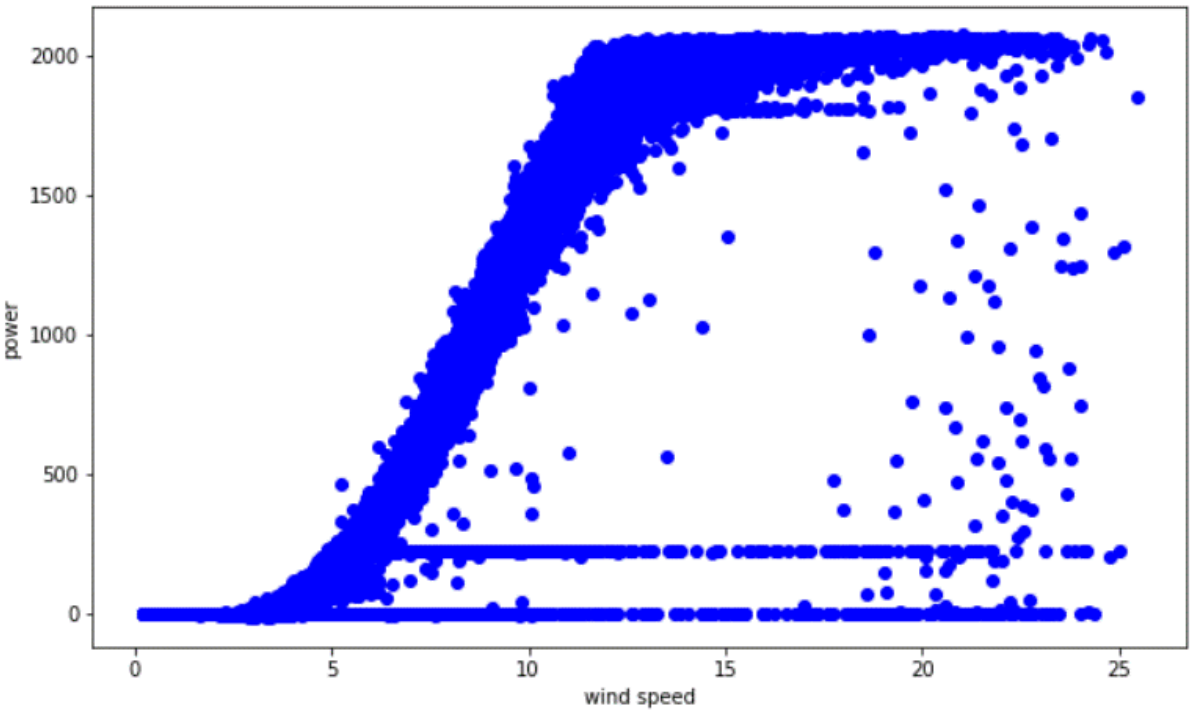
b)


power
wind speed
0
500
1000
1500
2000
0
5
10
15
20

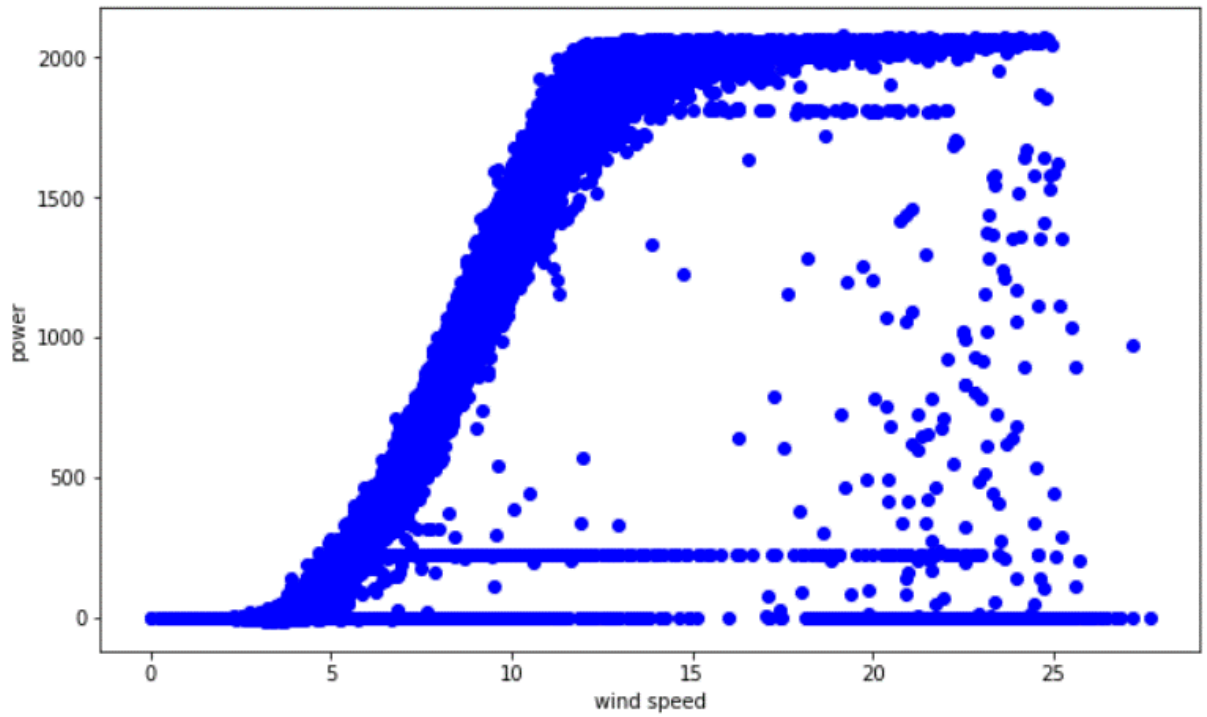
c)


power
wind speed
2000
1500
1000
500
0
0
5
10
15
20
25

d)

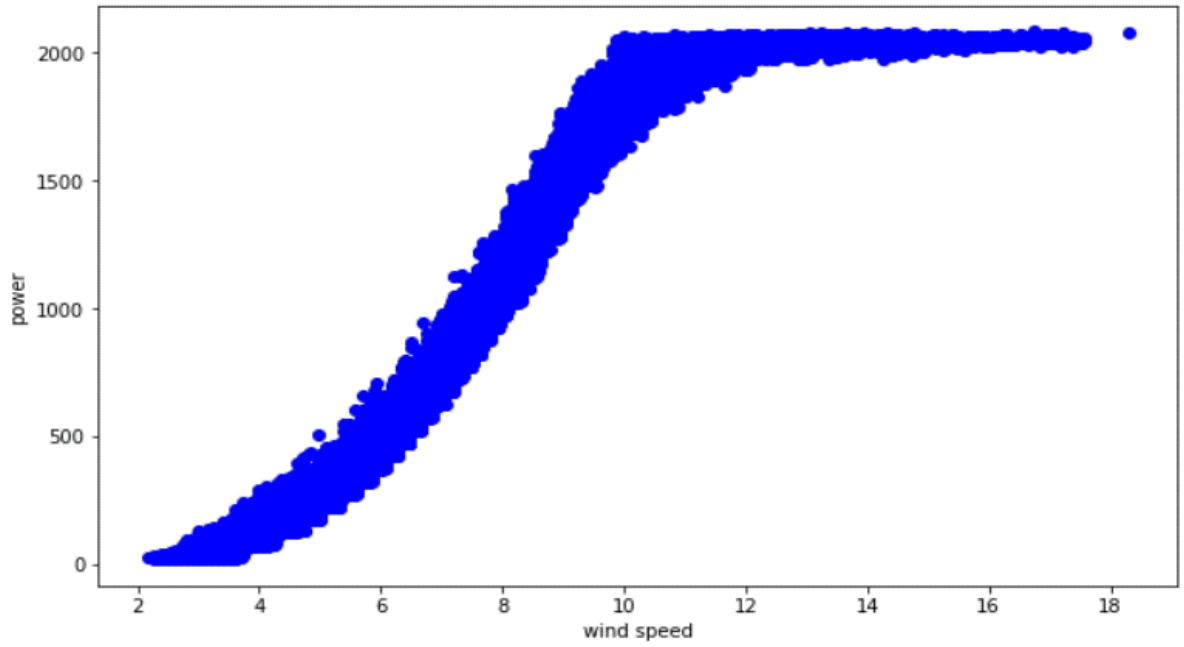

power
wind speed
2000
1500
1000
500
0
0
5
10
15
20
25

e)

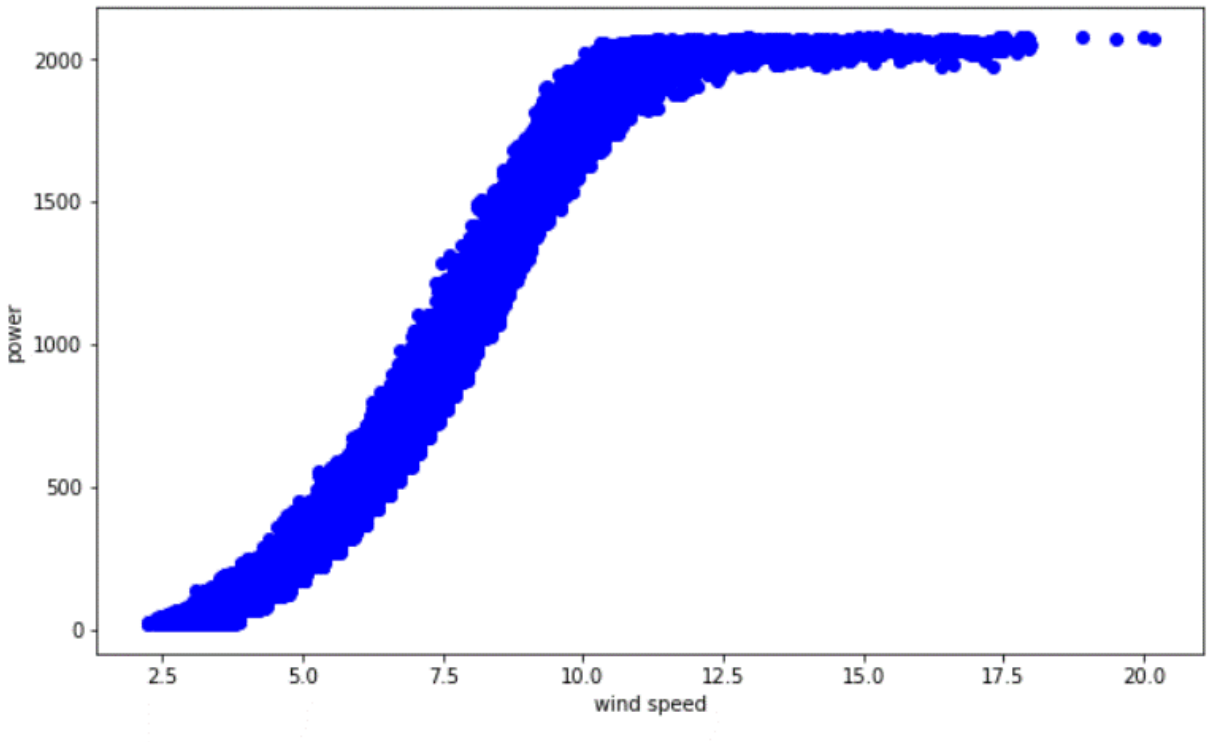


f)

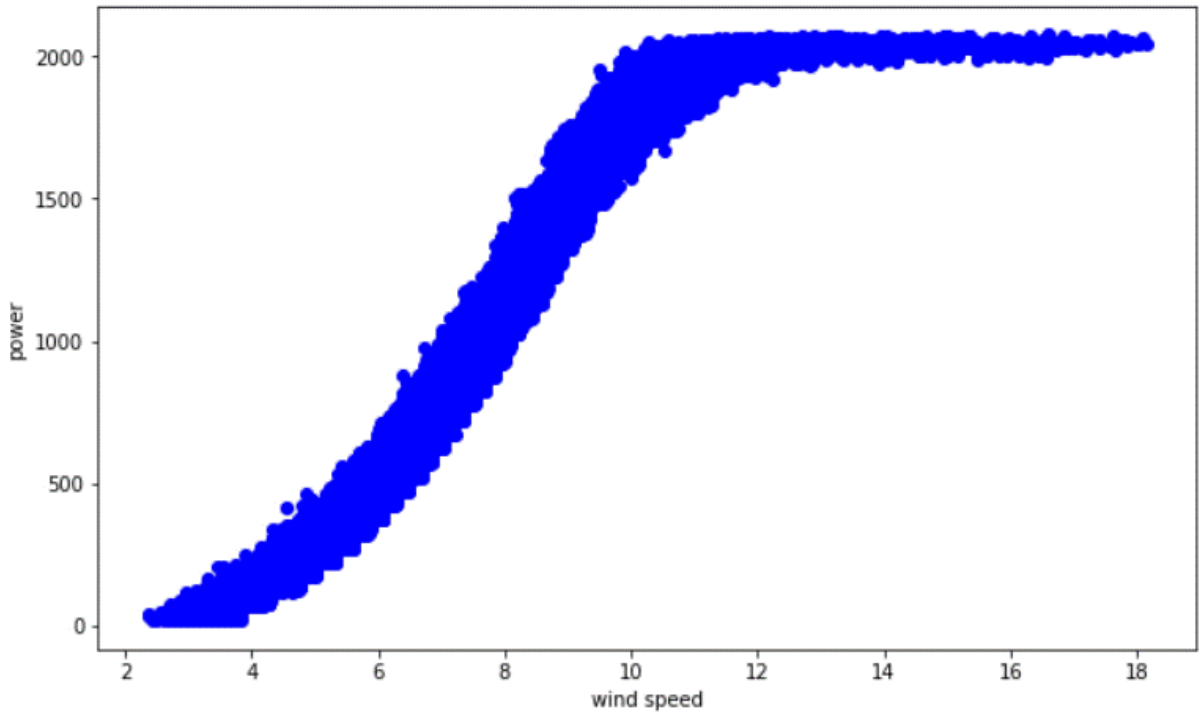
Figure 8: Actual Power output as a function of wind speed for: (a) Turbine 1, (b) Turbine 2, (c) Turbine 3, (d) Turbine 4, (e) Turbine 5, (f) Turbine 6


power
wind speed
0
500
1000
1500
2000
2
4
6
8
10
12
14
16
18

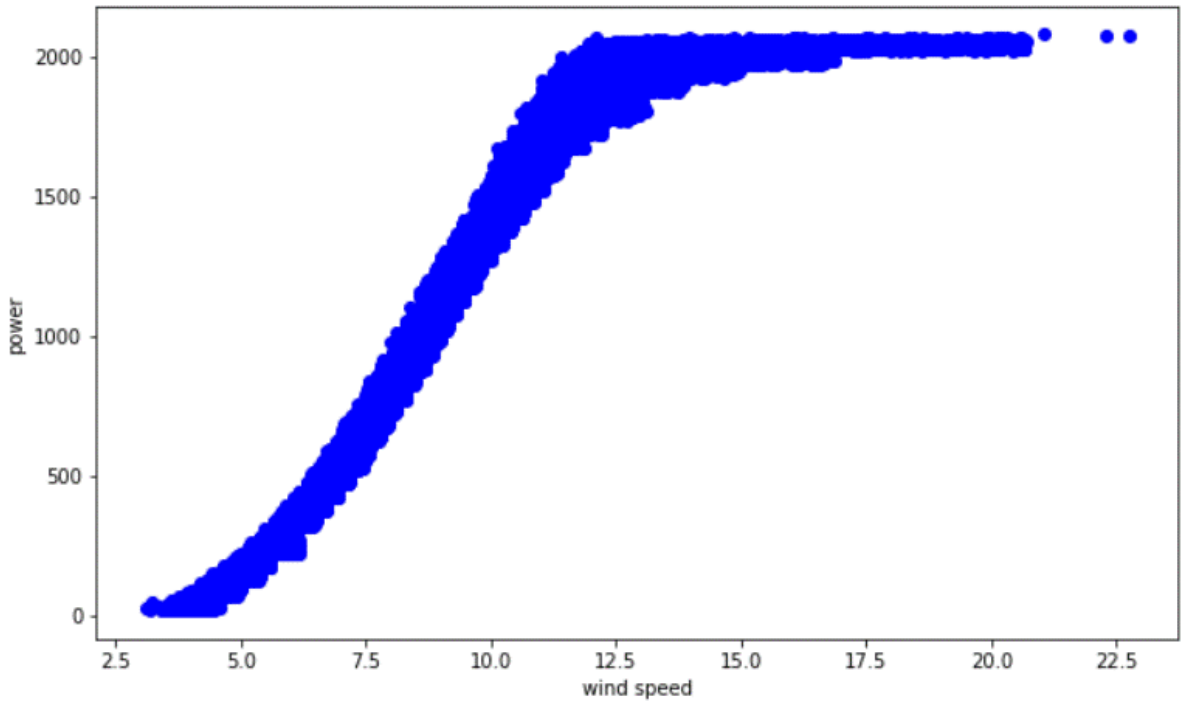
a)


power
wind speed
0
500
1000
1500
2000
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

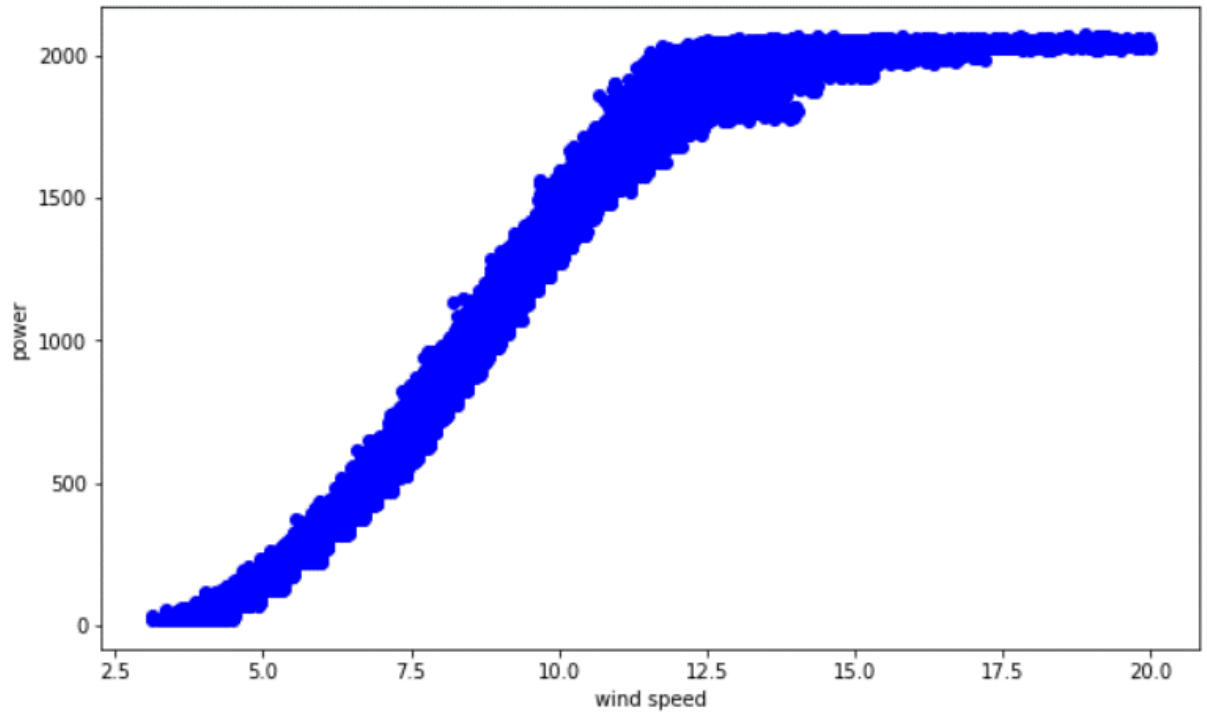
b)


2000
1500
1000
500
0
power
2
4
6
8
10
12
14
16
18
wind speed

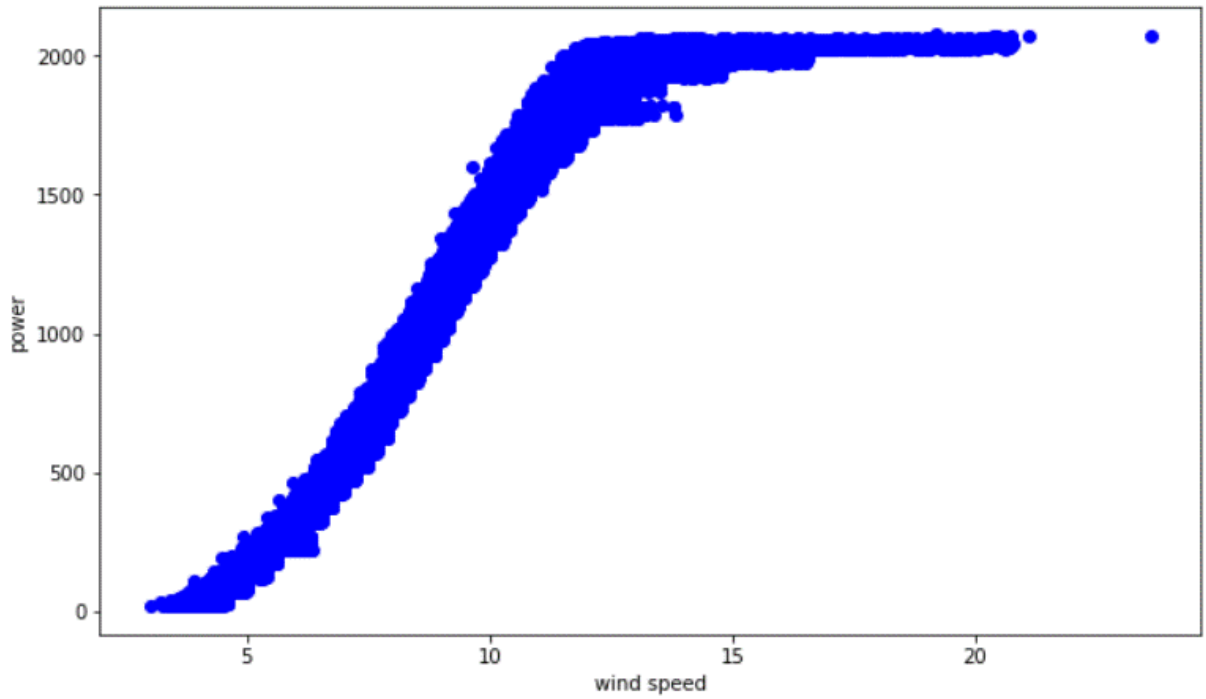
c)


2000
1500
1000
500
0
power
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
22.5
wind speed

d)

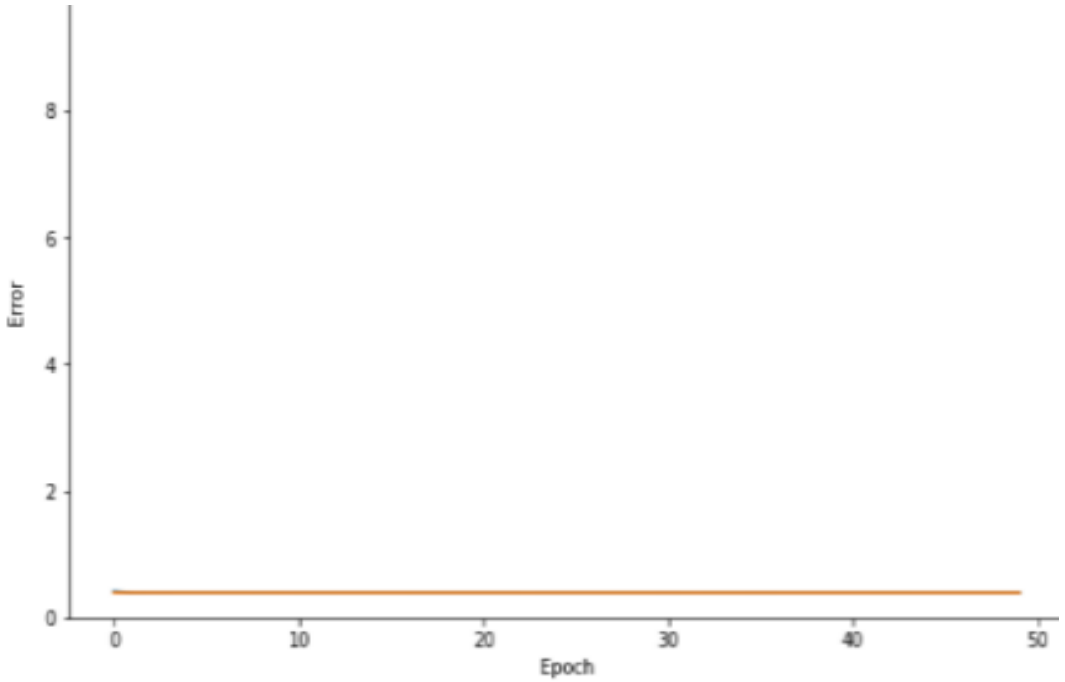


e)

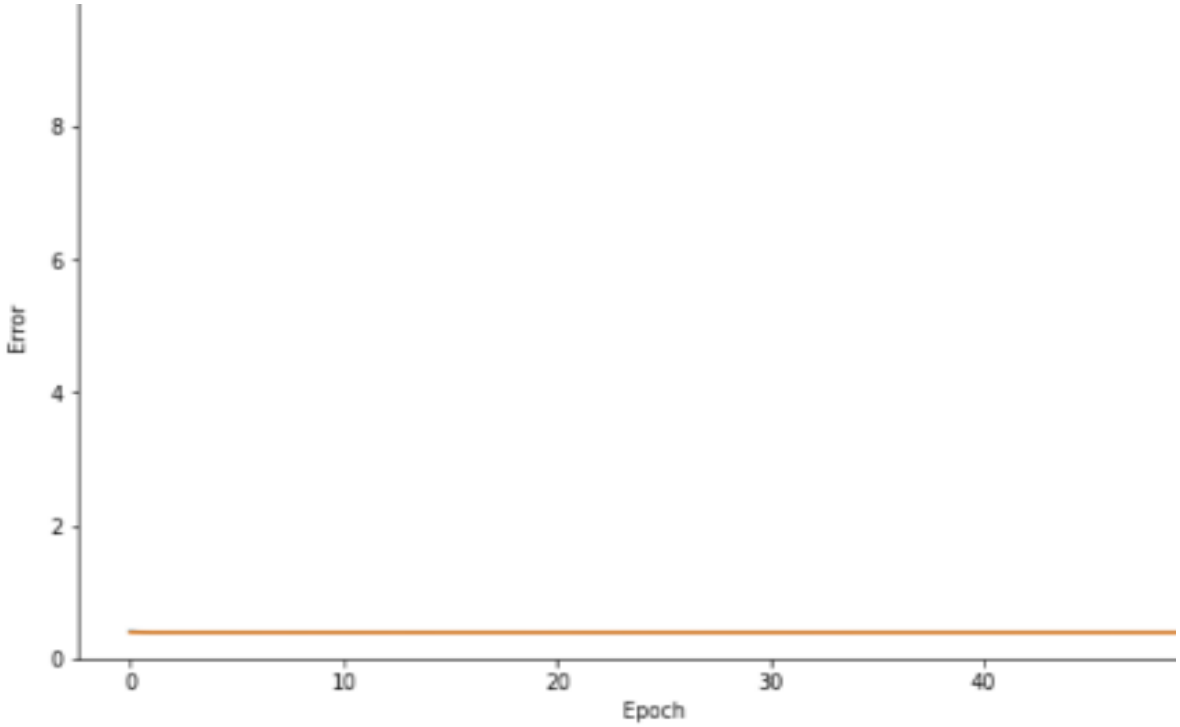


f)

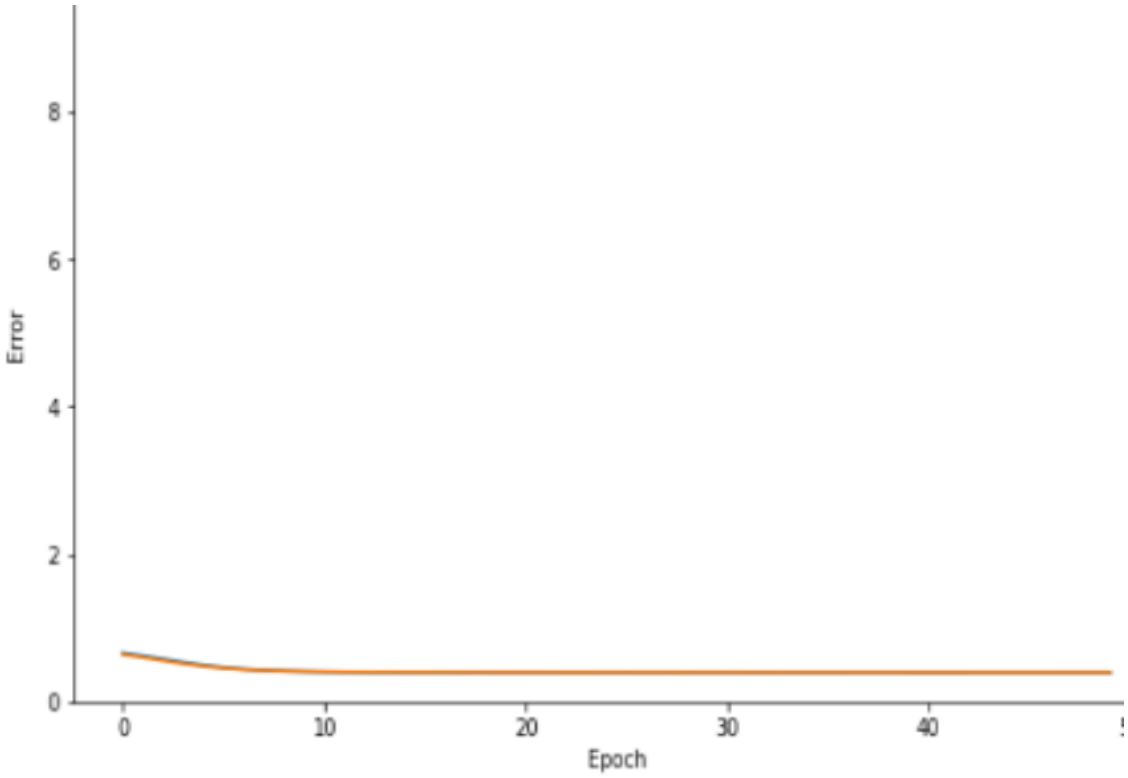
Figure 9: Filtered Power Curve for: (a) Turbine 1, (b) Turbine 2, (c) Turbine 3, (d) Turbine 4, (e) Turbine 5, and (f) Turbine 6


Error
Epoch
0
2
4
6
8
0
10
20
30
40
50

(a)

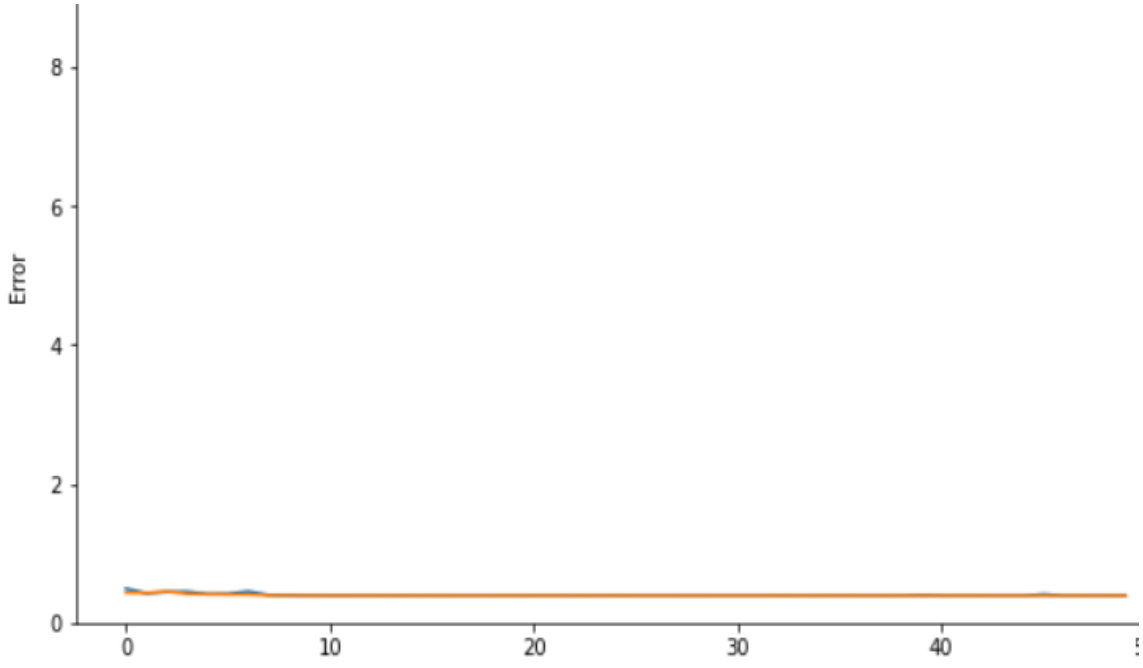

Error
Epoch
0
2
4
6
8
0
10
20
30
40

(b)

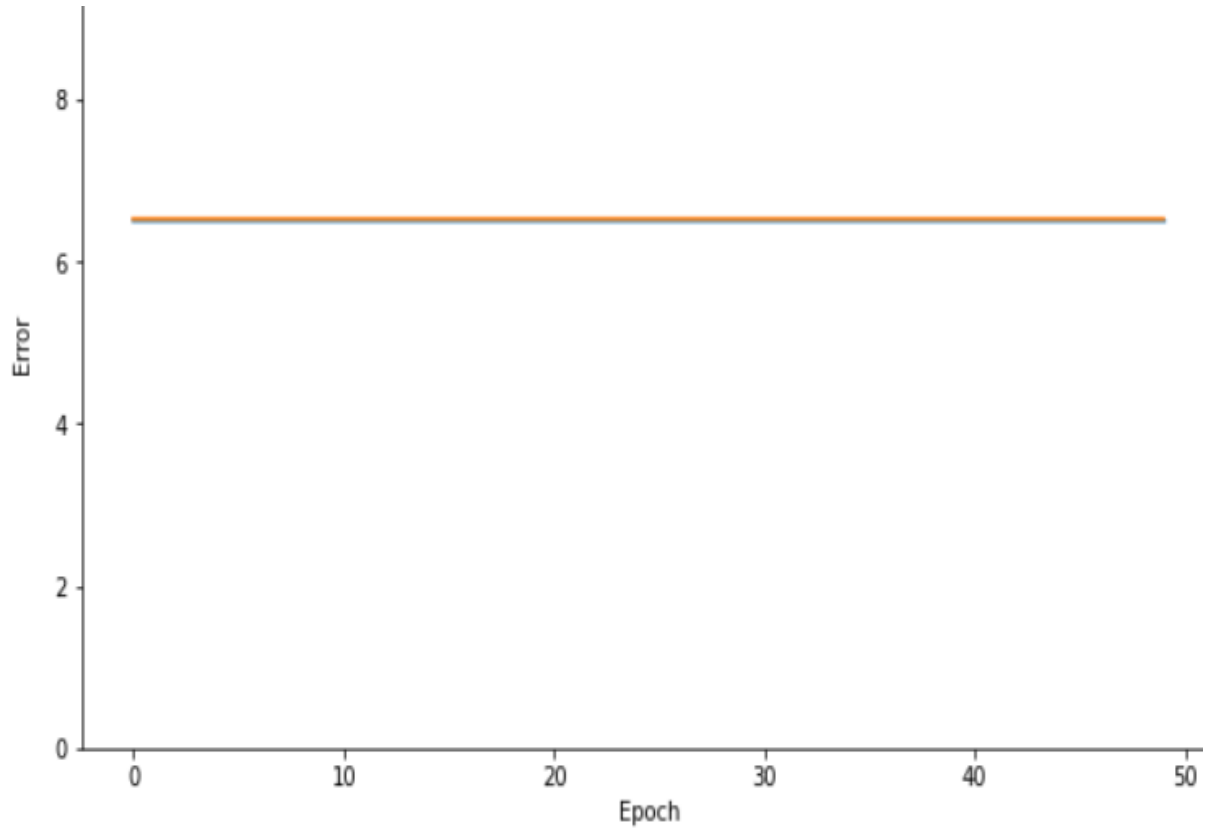


(c)

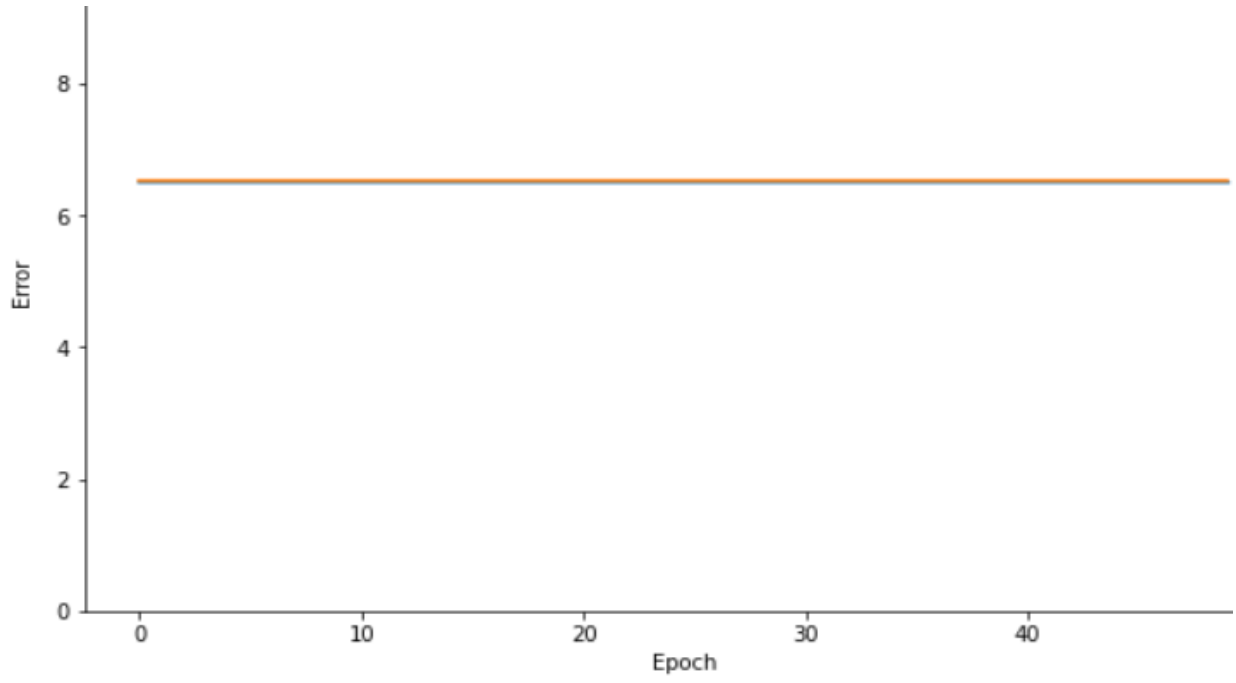
Figure 10: Plot of Loss to Epoch of the RBF Network: (a) adam, (b) rmsprop, and (c) sgd



(a)

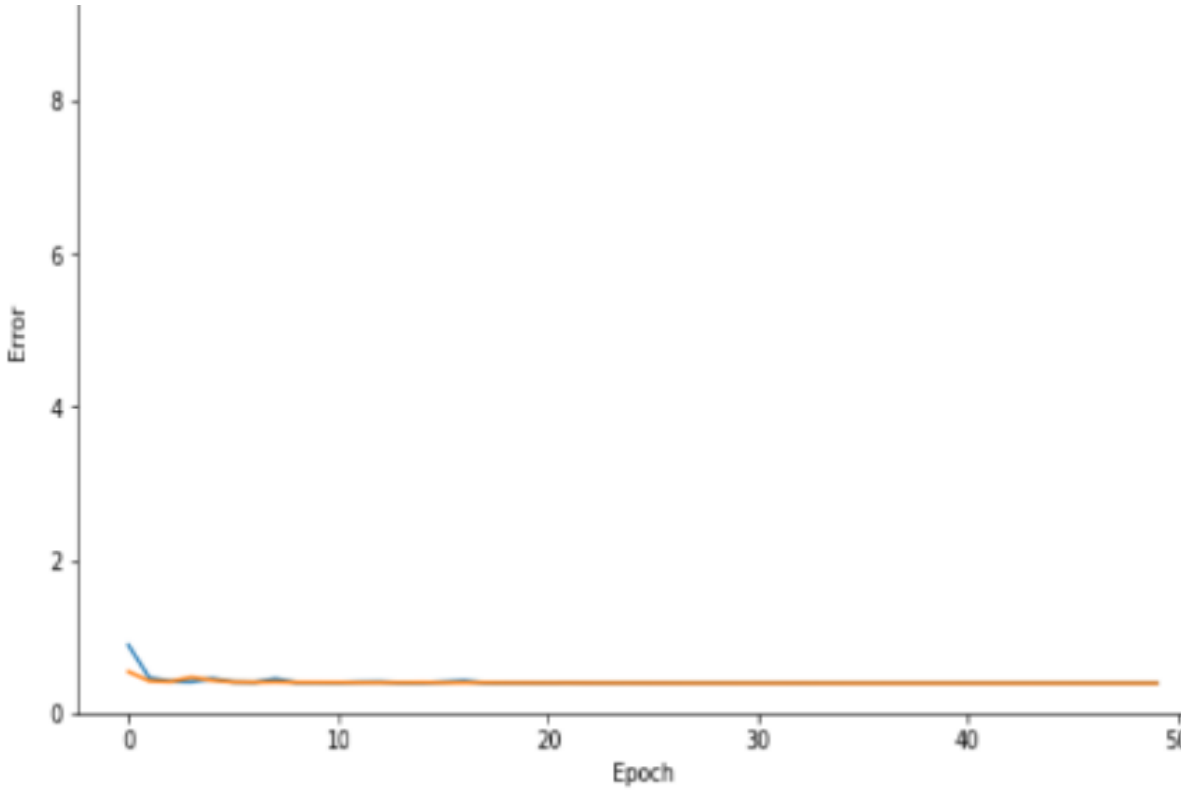


(b)

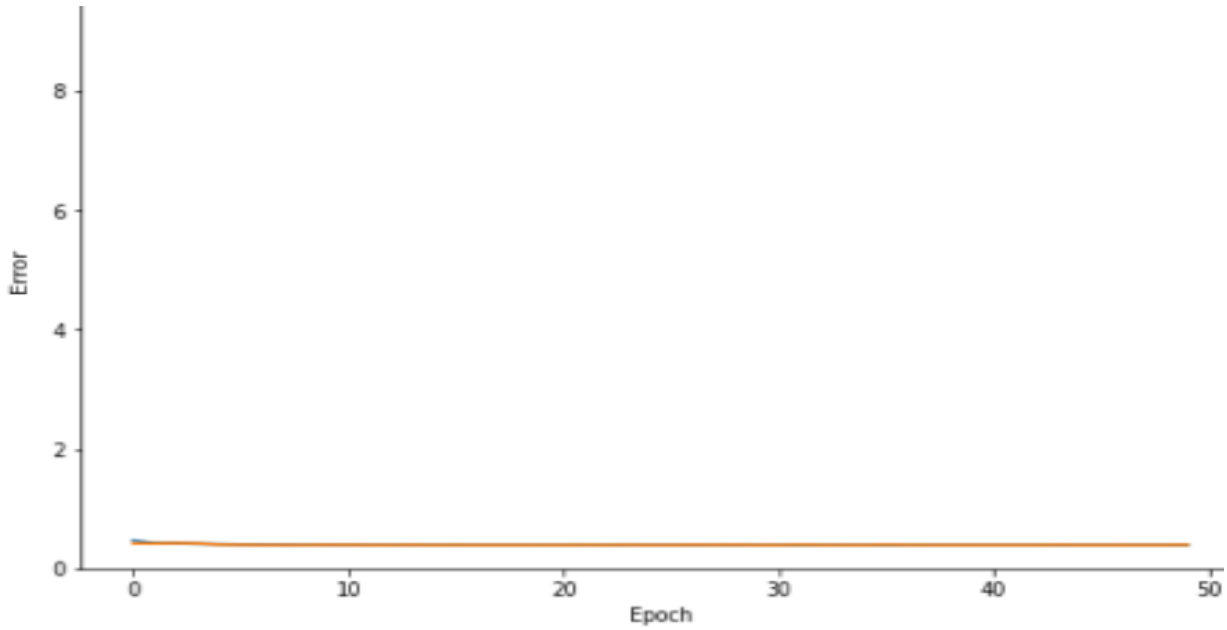


(c)

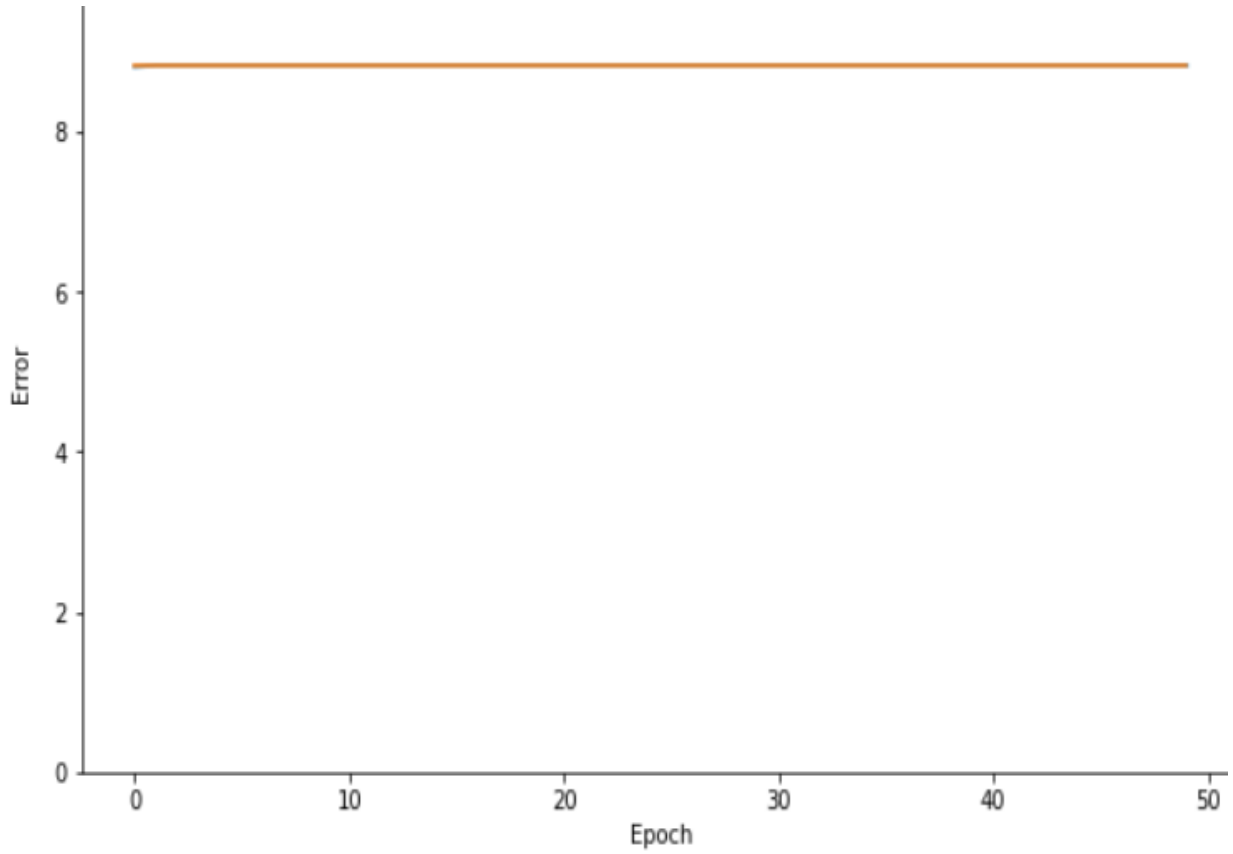
Figure 11: Plot of Loss to Epoch of the 4-Hidden layer MLP Network: (a) adam, (b) rmsprop, and (c) sgd


Error
Epoch
0
2
4
6
8
0
10
20
30
40

(a)


Error
Epoch
0
2
4
6
8
0
10
20
30
40
50

(b)



(c)

Figure 12: Plot of Loss to Epoch of the 6-Hidden layer MLP Network: (a) adam, (b) rmsprop, and (c) sgd

Table

Table 1: Percent Data Loss due to Filtering

	Before Cleaning	After Cleaning	Percent Data Loss
Dataset 1	52704	43425	17.6%
Dataset 2	52704	41196	21.8%
Dataset 3	52704	41278	21.6%
Dataset 4	52704	37247	28.1%
Dataset 5	52704	35282	29.5%
Dataset 6	52704	37760	27.9%

Table 2: Division of Dataset

	Training Size	Validation Size	Test Size
Dataset 1	28226	10856	8685
Dataset 2	26777	10299	8239
Dataset 3	26830	10319	8255
Dataset 4	24210	9311	7449
Dataset 5	22933	8820	7056
Dataset 6	24544	9440	7552

Table 3: Architecture of the Developed Models

	Model Name	Input Layer	Hidden Layer	Output Layer	Activation	Loss Function
Model 1	RBF	1	1	1	*sigmoid*	*Binary crossentropy*
Model 2	MLP1	1	4	1	*relu*	*Binary crossentropy*
Model 3	MLP2	1	6	1	*relu*	*Binary crossentropy*

Table 4: Test Result for Model 1 (RBF Model)

	optimizer	**MAE (KW)**	**RMSE (KW)**	**R^2**	**Computation time (sec)**
Data 1	adam	21.7	30	0.9924	145.94
	rmsprop	21.9	30	0.9923	202.69
	sgd	33.0	42.6	0.9833	142.50
Data 2	adam	23.7	32.1	0.9899	133.99
	rmsprop	23.6	32.7	0.9897	142.96
	sgd	35.2	46.6	0.9789	142.48
Data 3	adam	21.1	28.7	0.9917	142.59
	rmsprop	21.2	29.6	0.9913	142.475
	sgd	31.9	42.2	0.9815	142.511
Data 4	**adam**	**16.8**	**22.8**	**0.9955**	114.999
	rmsporp	17.1	23.3	0.9951	124.435
	sgd	23.9	30.1	0.9920	111.959
Data 5	adam	17.3	23.8	0.9950	142.571
	rmsprop	17.8	24.3	0.9933	142.663
	sgd	28.1	36.6	0.9880	112.231
Data 6	adam	18.5	25.6	0.9945	142.561
	rmsprop	17.3	24.7	0.9948	142.676
	sgd	25.7	32.5	0.9911	142.507

Table 5: Test Result for Model 2 (QF-MLP with 4-Hidden Layer)

	optimizer	**MAE (KW)**	**RMSE (KW)**	**R^2**	**Computation time (sec)**
Data 1	adam	21.6	28.6	0.9932	202.680
	rmsprop	50.2	60.8	-2.2444	166.041
	sgd	45.7	56.8	-1.8431	140.157
Data 2	adam	23.4	31.5	0.9902	142.671
	rmsprop	38.2	50.1	-1.4062	202.954
	sgd	98.8	100.3	-9.3963	143.103
Data 3	adam	21.7	28.8	0.9915	138.004
	rmsprop	35.4	46.77	-1.2681	159.599
	sgd	122.6	125.8	-15.0733	134.258
Data 4	**adam**	**16.1**	**22.1**	**0.9960**	142.637
	rmsporp	42.6	55.3	-1.4986	142.953
	sgd	51.3	63.2	-2.2674	115.142
Data 5	adam	16.7	23.1	0.9953	121.813
	rmsprop	36.7	49.4	-1.1796	140.560
	sgd	144.8	148.3	-19.0171	116.485
Data 6	adam	19.4	27.7	0.9940	125.826
	rmsprop	46.1	58.4	-1.7645	202.962
	sgd	49.5	61.1	-2.0359	121.506

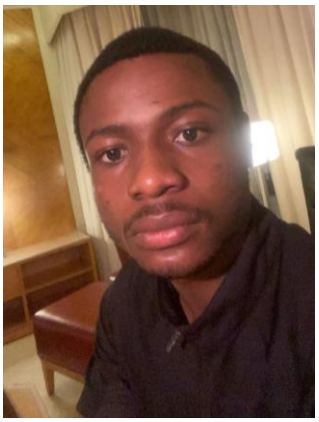
Table 6: Test Result for Model 3 (QF-MLP with 6-Hidden Layer)

	optimizer	**MAE (KW)**	**RMSE (KW)**	**R^2**	**Computation time (sec)**
Data 1	adam	25.6	35.0	0.7495	202.733
	rmsprop	43.8	55.06	-1.6645	181.260
	sgd	742.5	743.4	-486.80	202.629
Data 2	adam	40.9	52.1	-1.6036	202.737
	rmsprop	24.0	32.2	0.9898	174.694
	sgd	38.2	49.8	-1.373	140.108
Data 3	adam	22.5	30.5	0.9905	202.761
	rmsprop	34.4	46.4	-1.2357	203.162
	sgd	35.6	47.1	-1.3049	137.383
Data 4	**adam**	**15.3**	**21.5**	**0.9962**	142.753
	rmsporp	31.5	40.0	0.9871	151.888
	sgd	33755	3755	-11680	122.132
Data 5	adam	28.1	44.8	0.9825	142.739
	rmsprop	16.9	23.1	0.9925	203.155
	sgd	47.4	57.9	-1.9906	120.360
Data 6	adam	17.5	24.2	0.9953	142.737
	rmsprop	417.3	543.5	-1.4034	203.118
	sgd	421.0	547.6	-1.423	125.839

Table 7: Comparison of Results of Parametric and Non-parametric Models

	Hidden layers	**MAE (KW)**	**R^2 score**	**Computational time (s)**
QF-RBF(adam)	1	**16.8**	**0.995**	**114.999**
QF-RBF(rmsprop)	1	**17.1**	**0.995**	**124.656**
QF-MLP1-(adam)	4	**16.1**	**0.996**	**142.637**
QF-MLP2-(adam)	6	**15.3**	**0.996**	**142.753**
TS-NSFM	1	18.935	0.994	4896
NSFM	1	46.823	0.955	-
SFM	1	19.103	0.992	987
Cubic Spline	-	32.7123	0.985	1.50
SVM	-	27.345	0.989	80
GP	-	23.497	0.990	45.43
Logistic Function	-	29.956	0.986	1.89
Parametric	-	33.545	0.984	1.46
Modified IEC-BINS	-	29.453	0.989	1.32
Manufacturer power curve	-	40.234	0.979	-

Author Bio(s) with Photo(s)



Emerald U. Henry is a Mechanical Engineering graduate from Covenant University. His current research focus is renewable energy and wind turbine optimization. He is also a member of the energy and environment research group within Mechanical Engineering in Covenant University, Ota.