Turbine Wake Analytics Project: Predicting Power Loss Using Neural Networks (Shafkat Ibrahimy)

GitHub Project Link: https://github.com/shafkat22/windwakeloss

0. Executive Summary

This study developed three artificial neural network models to analyse the power loss caused by wake effects as an alternative approach to conventional statistical and physics-based models, such as the Jensen Wake Model and Computational Fluid Dynamics models (CFD). By analysing six datasets on wind turbine pairs, this project developed Convolutional Neural Networks (1D CNN), Deep Artificial Neural Networks (Deep ANN), and Long Short-Term Memory (LSTM) to uncover complex relationships affecting wake loss. After hyperparameter tunning, this study found CNN to be the overall best-performing model, with the lowest average RMSE, and Deep ANN to be the most consistent in predicting power loss. Future studies can explore the development of hybrid systems utilising both physics-based and artificial neural network models to optimise wind farms and reduce the effects of wake loss.

1. Background and Project Aims

When the blades of a wind turbine rotate, they affect the natural wind flow and disrupt the wind going into downstream turbines. This occurs due to the upstream turbine extracting kinetic energy from the wind stream, reducing the energy generated by the downstream turbine because of a lower wind velocity. This is known as a wake. It is therefore important to analyse these losses, optimising designs to maximise energy yield.

Figure 1 below has been created to illustrate how the direction of the incoming wind affects the wake loss experienced by the downstream turbine (Turbine B), where θ is an acute angle between the wind stream and the alignment line of the turbine pair. The maximum energy loss occurs when θ = 0, and the loss in wind velocity is more pronounced at locations closer to the upstream turbine (Turbine A).

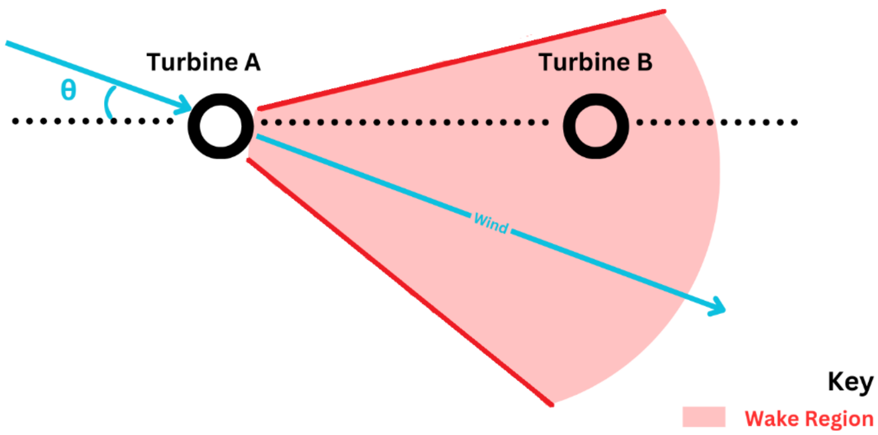

Figure 1: Wake Loss

Wind analysts use wake models and simulations with wind farm layouts to optimise the general efficiency of wind farms. Wake models range from simple statistical models, such as the Jensen Wake Model, to complex CFD (Computational Fluid Dynamics) physics models based on RANS equations (Reynolds-Averaged Navier-Strokes). An example is shown below in *Equation 1: Jensen Wake Model,* where ΔV is the velocity deficit, V_0 is the upstream wind velocity, C_T is the thrust coefficient, σ is the wake decay constant, and d is the distance from the upstream turbine to the downstream turbine.

$$\Delta V = V_0\left(1 - \sqrt{1 - \frac{C_T}{(8\sigma^2/d^2)}}\right)$$

Equation 1: Jensen Wake Model

This study aims to leverage the strength of complex artificial neural networks to capture the nonlinear relationships between features and wake loss in wind farms, predicting the power loss due to wake effects. Every wake model has its strengths and weaknesses, and utilising advanced data-driven models in combination with conventional models can potentially improve the energy output of wind farms by offering a different perspective. However, it is important to note that this study is limited by the quality of the available datasets used and more accurate predictions can be made with access to additional datasets containing meteorological, geographical, and operational data. This study is still useful as the framework can be reproduced and applied to other case studies with additional datasets. Beyond this case study, advanced hybrid systems blending both conventional and deep learning models can be further explored.

2. Methodology

2.1 Data Collection and Preprocessing

Six dataset CSV files containing data on six pairs of wind turbines were used in this study (Hwangbo, H.). Each dataset includes data from 2010-2011 containing power output, wind velocities, and wind direction of a pair of turbines measured at the turbine and at the mast (3 masts in total; mast 1 covers pair 1 and 2; mast 2 covers pair 2 and 3; mast 3 covers pair 5 and 6) at 10-minute intervals. All turbine pairs have a northwest-to-southeast orientation, and each pair is positioned such that no pair of turbines other than the turbine in the pair is in close range (10 times the rotor diameter). In addition, the relative bearings of the turbine pairs are shown in section 2.2.

The data preprocessing involved the removal of rows containing missing values or outliers for the feature Va (free-stream or ambient velocity) and the creation of a new feature, differential velocity, derived from two measured velocity variables V1 and V2 (velocity at turbines). Moreover, records of data were labelled based on which turbine in a pair was considered the downstream turbine, which was determined by the wind direction.

2.2 Data Visualisation and Exploration

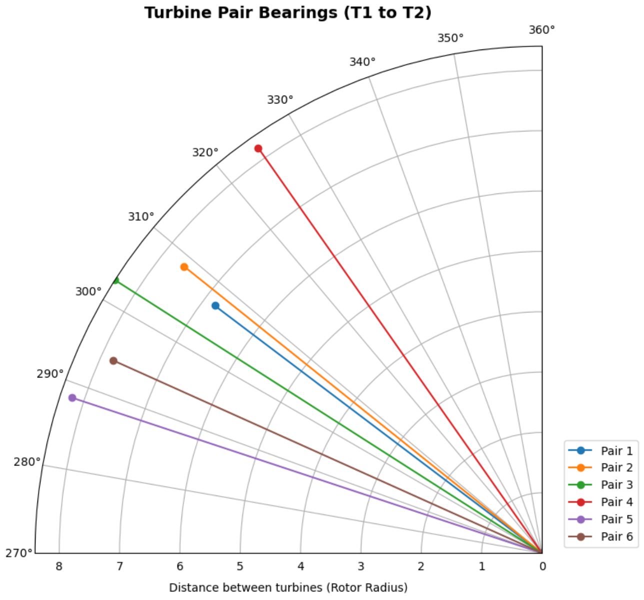

Figure 2: Relative Positions of Turbine Pairs

Figure 2 above shows the relative bearings and distance between the two turbines in a pair. All turbines in a particular pair are within 10 times the rotor radius of each other, and the bearings from turbine 2 to turbine 1 are 307.1, 308.7, 302.6, 325.0, 288.3, 294.2 degrees for Pairs 1-6 respectively. Figure 3 below illustrates a scatter plot of the power difference (wake loss; PW1-PW2) across different wind directions. As expected, the maximum power loss occurs when the angle θ = 0, where the wind direction relative to the turbine alignment is parallel, causing the downstream turbine to be in full wake of the upstream turbine.

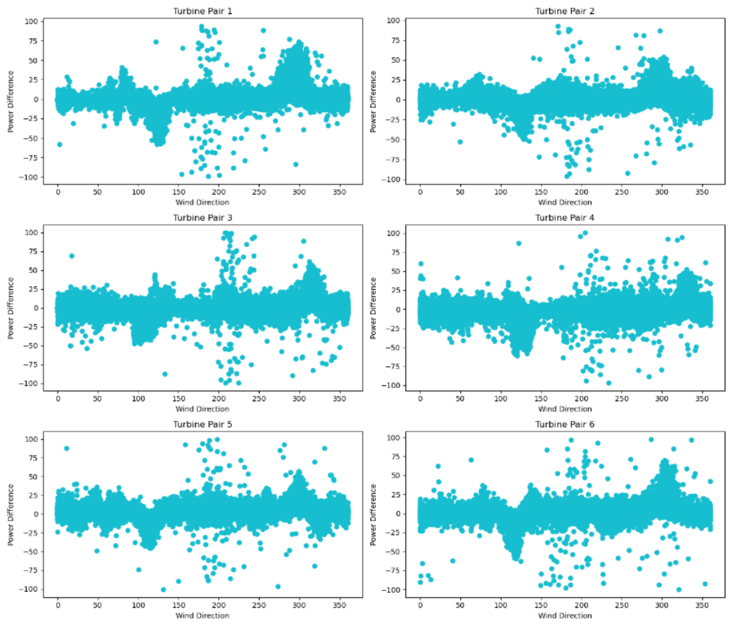

Figure 3: Power Loss vs Wind Direction

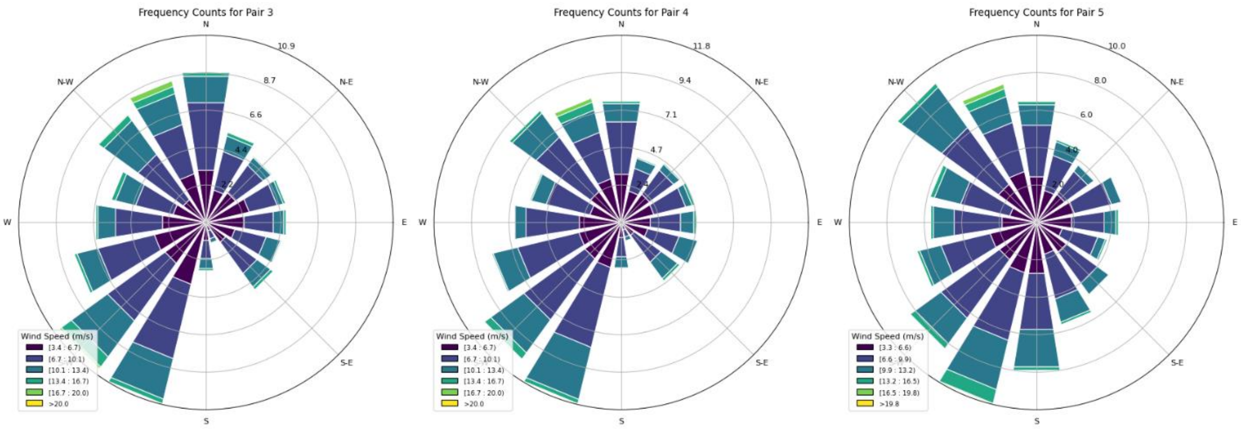

Figure 4: Wind Rose Frequency Count Diagrams.

Figure 4 illustrates several wind-rose diagrams, which show the frequency of wind directions and wind velocity in the datasets. For Pairs 3-5, the most frequent wind direction was southwest. Although Figure 4 only contains the diagrams for Pairs 3-5, the frequency counts for the other pairs were also explored. For Pairs 1-2, the most common wind direction was south, followed by southwest. Lastly, for pair 6, the most common directions were south-southwest and northwest. For all pairs, the most frequent wind velocities ranged from 3.4 to 10 m/s, with the highest counts in the 6.7-10 m/s range.

2.3 Model Development and Tuning

Three neural network models were developed, including Convolutional Neural Networks (CNN), Deep Artificial Neural Networks (Deep ANN), and Long Short-Term Memory Networks (LSTM). The justification behind these model selections was based on the individual strengths of the models. CNNs are effective in processing spatial data, which can relate to the layout or design of wind farms. Deep ANNs are good at feature extractions and predicting using high-dimensional data. LSTMs are effective with time-series data, which can capture temporal changes in wind direction and velocities. Lastly, an ensemble method, Random Forest Regressor, was also tested on the data for comparison purposes.

Figure 5 illustrates the general architecture of the artificial neural network models developed in this study. The architecture involves an input layer, several processing layers, and an output layer.

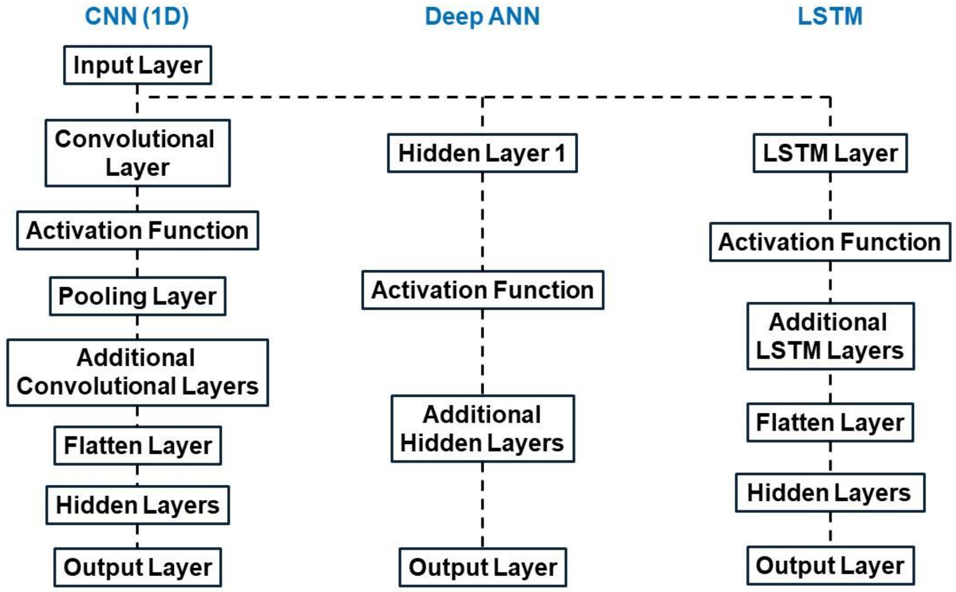

Figure 5: Artificial Neural Network Models Architecture.

The target variable of the model output was the power output relative to rated power at turbine 2 (PW2) in a particular pair, which can range from 0-100. The predictors (features) were V1, V2, Va (velocities measured at the turbines and at the masts), D (wind direction), differential velocity, and power of turbine 1 (PW1). It is important to note that depending on wind direction, either turbine 1 or turbine 2 would be considered the downstream turbine which is affected by wake loss.

2.3.1 Deep Artificial Neural Networks (Deep ANN)

The study's model architecture consisted of 3 hidden layers, each containing 80 neurons, and used the Rectified Linear unit (ReLu) activation function. The training was conducted for 400 epochs with a validation size of 0.1, the 'Adam' optimiser, and Mean Squared Error (MSE) as the loss function. These hyperparameter values and training configurations were based on minimising the RMSE (Root Mean Squared Error) by testing a range of values detailed below.

Firstly, a baseline 3-layer ANN model was developed, and a range of neuron values were tested. Neurons in a layer have weights and biases, which are parameters that are changed during training depending on patterns in the data. Figure 6 below shows the effect of the number of neurons in the hidden layer on performance. From the plot, after 80 neurons, increasing the number of neurons has a negative impact on the RMSE value, which is potentially due to overfitting.

Next, a range of epochs were tested. Figure 7 shows the effect of epochs on performance, and 400 was found to be the optimal value. Epochs refer to the number of iterations where the model sees the whole dataset for training. If the number of epochs is too high, it will lead to overfitting, but if it is too low, it will lead to underfitting the model. Overfitting causes failure in being able to generalise the model for unseen data while underfitting the model fails to find sufficient patterns in the data to make accurate predictions.

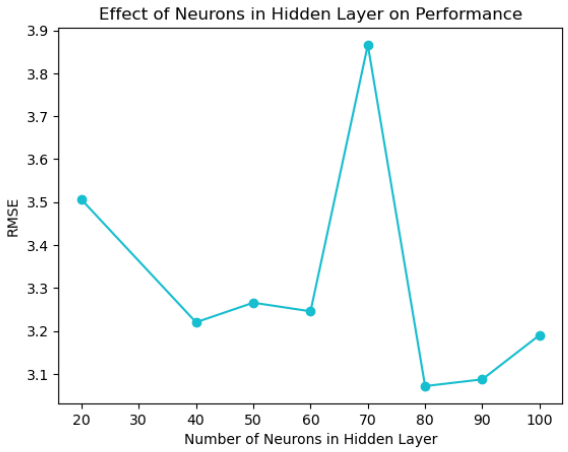

Figure 6: Number of neurons vs RMSE.

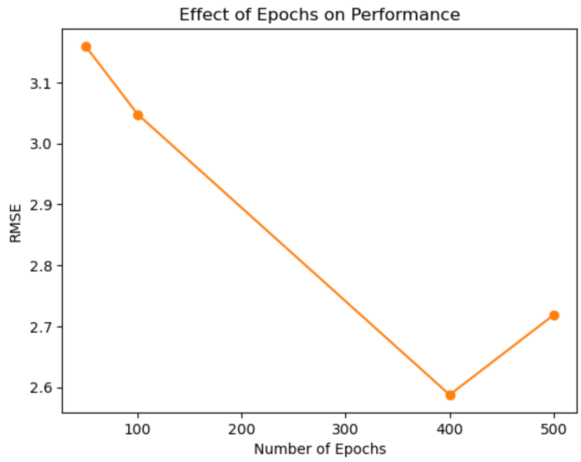

Figure 7: Number of epochs vs RMSE.

Figure 8 below, shows the effect of the number of hidden layers on performance. The hidden layers are where the network learns from the input data and applies transformations for predictions. Each layer contains neurons (or nodes) which carry out calculations on the data. The function of the hidden layers is to find complex relationships in the data.

Figure 9 below, shows the performance of different activation functions, and in this study, 'ReLu' was the best-performing function. The activation function is a mathematical transformation that captures non-linear patterns and affects the network's ability to learn and generalise the model to work well on unseen data. These hyperparameters affect the convergence speed, gradient stability, and prediction performance.

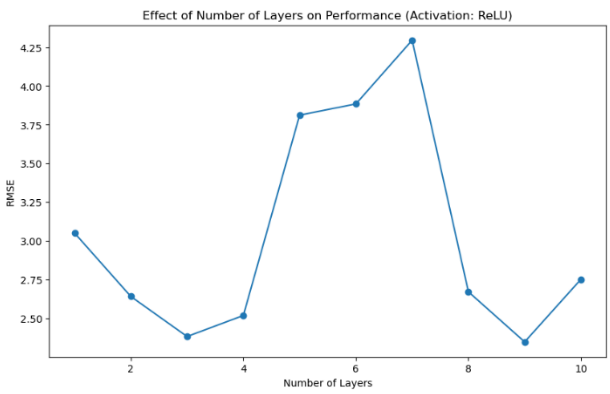

Figure 8: Number of hidden layers vs RMSE.

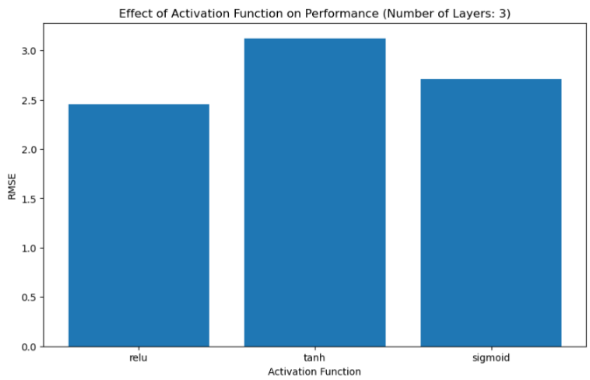

Figure 9: Activation Functions vs RMSE.

2.3.2 Convolutional Neural Networks (CNN-1D)

The 1D CNN model's architecture used in this study consists of 4 convolutional layers with a kernel size of 3, 32 filters per layer, and 128 neurons. Then followed by pooling layers and 3 hidden layers. The model used ReLu as the activation function. To ensure compatibility, the training and testing data were reshaped before being processed by the model as the 1D model requires the input signal to be a particular format. 1-dimensional CNNs expect the input to be a 3D array, where each dimension refers to a characteristic of the data. The first dimension (axis 0) is the samples in the dataset, where a sample is an instance of data (row). The second dimension (axis 1) is the time steps of a sample, which in this study is represented by the timestamp. The third dimension (axis 2) is the features of each sample (columns).

Several hyperparameter groups were tested, and the results of these can be seen in Figure 10. Each hyperparameter group consisted of different values of the model's parameters, adjusting the kernel size, number of filters, number of neurons, and number of hidden and convolutional layers.

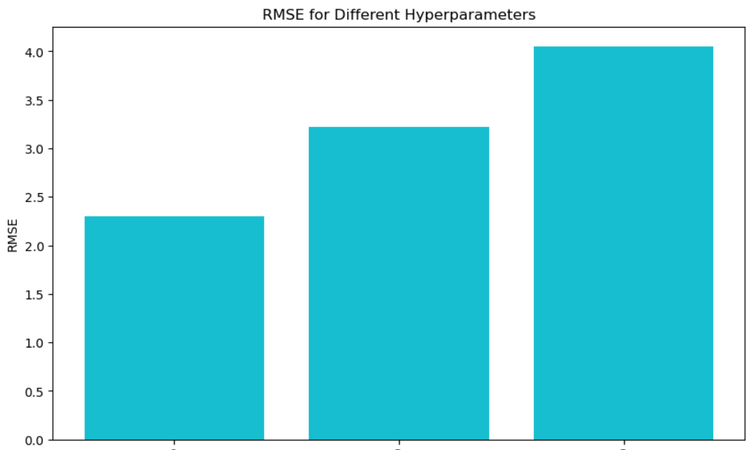

Figure 10: Hyperparameter Groups vs RMSE.

2.3.3 Long Short-Term Memory (LSTM)

The LSTM model architecture in this study consisted of 2 LSTM layers, 2 hidden layers, and 128 neurons. The activation function used was the hyperbolic tangent (tanh). The model was trained for 400 epochs with the 'Adam' optimiser, which is a similar configuration to the previous models.

LSTMs are a kind of Recurrent Neural Network architecture which was created to better understand the long-term dependencies that exist in sequential data. It uses gating mechanisms and memory cells together to let gradients backpropagate easily, updating and forgetting information over time.

3. Results

Table 1: Results (RMSE) of predicting the turbine power 2 (PW2)

	Pair 1	Pair 2	Pair 3	Pair 4	Pair 5	Pair 6	Average
CNN	2.15	3.18	2.36	2.44	1.96	2.56	2.44
Deep ANN	2.39	2.42	2.39	2.42	2.62	3.00	2.54
LSTM	2.35	3.20	2.22	2.27	2.90	2.60	2.59
RF	2.70	3.13	2.38	3.53	2.51	2.38	2.78

Table 1 shows the RMSE metric values for predicting the turbine power (PW2) of turbine 2 in a pair. The Deep ANN model performed consistently for all pairs at around 2.4-3, with an average RMSE of 2.54. Although the Deep ANN model has the second lowest average RMSE, it had the least variability (standard deviation of 0.26), making it the most consistent. The LSTM model had an error rate ranging from 2.2 to 3.2. It performed relatively well for Pair 1, 3 and 4, but it also had inconsistent performance with a high variability (standard deviation of 0.37). Random Forest had the worst average RMSE, with the error ranging from 2.4 to 3.5.

The CNN model had the lowest average RMSE of 2.44, which makes it the best-performing model in this table for predicting turbine 2's power (PW2). CNN has good potential in capturing the spatial relationship within a dataset, therefore with additional datasets containing data on turbine layouts, it may be able to improve its ability to analyse wake loss.

4. Limitations and Future Work

As mentioned previously, the datasets used in this study lack information on environmental conditions such as air pressure, temperature, rainfall, and other relevant conditions which may affect wake loss. With more information regarding the atmospheric conditions, topographical data, and turbine construction data, the models may exhibit better prediction capabilities for wake loss, which is useful for optimising wind farm configurations.

This study had a systematic method for hyperparameter tuning, as described in section 3. However, tuning neural network models is a complex process and identifying the optimal parameters may require further experimentation on a wider range of parameter values. Other optimisation techniques and feature engineering techniques may improve the models' performance.

Future studies can explore developing hybrid systems which take advantage of artificial neural networks' capabilities in finding complex relationships affecting wake loss in combination with advanced physics-based models. By leveraging the strength of multiple approaches, optimising wind farms to maximise energy yield can be further analysed.

References

Ding, Y., 2019. Data science for wind energy. Chapman and Hall/CRC.

Hwangbo, H., Johnson, A.L. and Ding, Y., 2018. Spline model for wake effect analysis: Characteristics of a single wake and its impacts on wind turbine power generation. IISE Transactions, 50(2), pp.112-125.