Wind Power Forecasting using Machine Learning

Kartik Anant Kulkarni, Roll No. 210493
Mathematical Engineering and Physics using Machine Learning (EE798Z)

Paper Summary

Wind power, among the renewables, strikes the best balance between cost and output. Accurately forecasting wind power will greatly simplify the task for companies, traders and on the longer scale, governments as well.

This forecasting has been traditionally carried out using statistical models and recently, using machine learning. Due to the high non-linearity and the strong dependence on weather, this problem has been frequently tackled with a huge ensemble of predictors and lot of simulated data.

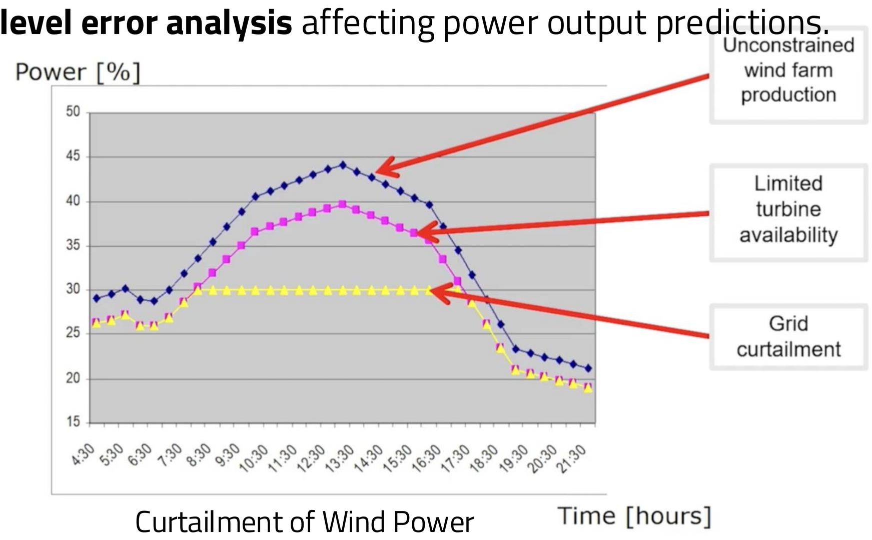
This project aims to explore the very recent open-source release of the SDWPF dataset, and build a robust model for forecasting wind power, in the light of the new features.

Background

- Wind power forecasting is essentially a **time-series** forecasting problem.
- Physically, **Betz's law** limits the maximum efficiency of a wind turbine, given by,

$$\eta = \frac{16}{27}\left(1 - \left(\frac{v}{v_{rated}}\right)^3\right)$$

where v is the wind speed and v_{rated} is the rated wind speed of the turbine.
- Forecasts require **tailored horizons** and **resolutions**. They are either **deterministic** or **probabilistic**. **Ramp predictions** and user-specific constraints makes the problem complex.
- **Finance** (curtailments, LMPs) plays a crucial role in actual wind throughput and must be considered separately.
- Balancing **quality, reliability**, and **price** in forecasting is difficult. Additional challenges like costly maintenance of **multi-step models** and rise of **energy storage, phase and level error analysis** affecting power output predictions.



Curtailment of Wind Power

Project Details

- **Platform** –
  OSX-ARM Architecture in a Python 3.11 Miniforge-based environment
- **Dataset** –
  Recently released (July 2024) **SD-WPF dataset** contains data for *134 turbines at a frequency of 25 min, for a duration of around 2 years,* for a windfarm (owned by Longyuan Power Group corp. Ltd.) in China. It has the **geolocations with elevation** of each turbine, **weather and ERA5 reanalysis data** spans the "**real"** values (not simulated) of almost 4 out of the 5 categories (SCADA data is missing)

Metrics and Evaluation

- The choice of metric significantly affects the output. **Mean absolute error (MAE)**, is used for maintaining generalisability.
- A forecast horizon of **24 hours with a frequency of 1 hour** was chosen as balance between dataset size and computation limitation.

Project Organization

The project was divided into **3 phases**:
1. Data Analysis, Baseline and Initial Approach
2. Exploiting ERA5: Turbine-Level Forecast Refinement
3. Spatio-Temporal Wind Farm Output Modelling

Phase 1

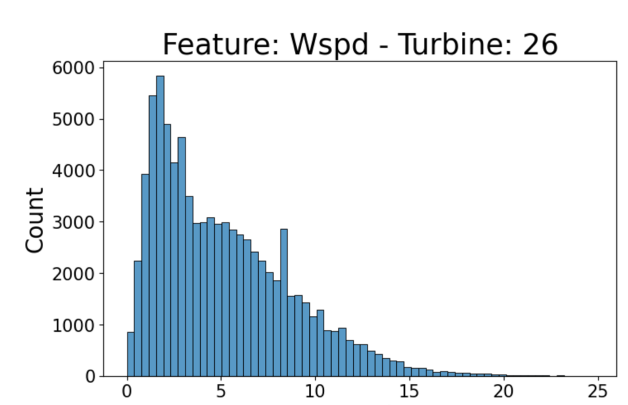
The goal was to analysis of the dataset and built a baseline model using an LSTM-CNN architecture for forecasting wind power from a single turbine.

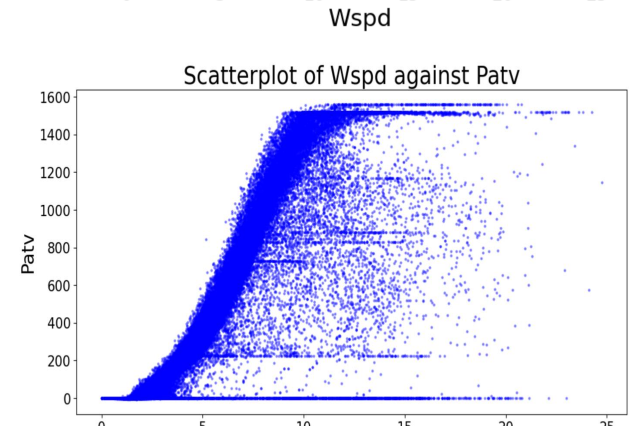
Key Insights:
- **Wind speed** has a strong correlation with wind power.
- **Missing values** were better filled with subsequent values rather than linear interpolation.
- **Wind Speed Detection**: Wind speed roughly follows Betz Law with noise due to curtailment. A Poisson-like behaviour in the distribution of wind speeds was observed, guiding further modelling strategies.
- **Feature Transformation**: Angle-based features were converted to sine/cosine representations to capture cyclical nature, and abnormal temperature values were corrected.
- **Time of Day Feature** was added to help recognize temporal patterns.

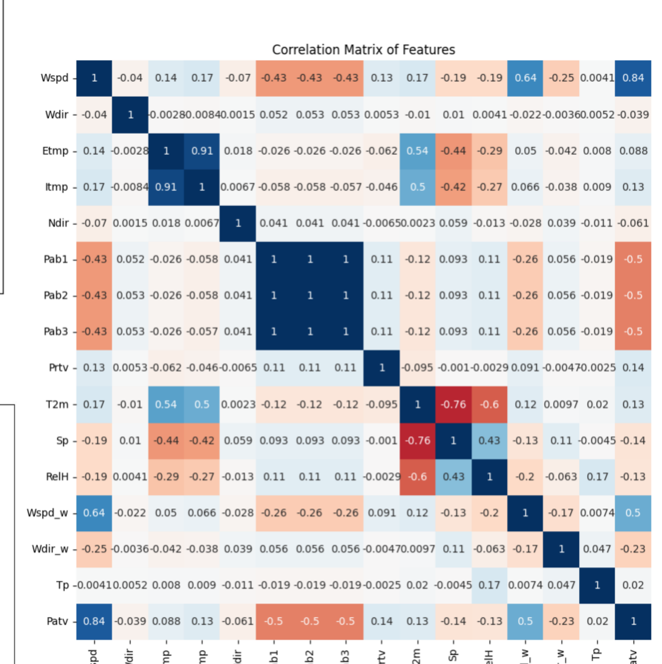
Model	MAE	% Improvement
Baseline Models		
Random Forecast	468.37 KW	0.0%
Replicate Yesterday	412.99 KW	11.82%
Moving Average	369.20 KW	21.17%
LSTM-CNN Models		
Target Variable History	420.91 KW	10.13%
All Feature Variables	434.48 KW	7.24%
"Perfect" Wind Speed Forecasts	232.51 KW	50.36%
"Noisy" Wind Speed Forecasts	264.80 KW	43.47%

Challenges & Observations:
- **LSTM-CNN** improved predictions by ~45% over the random baseline. However, MAE was still high, with challenges faced in overfitting and capturing trends with limited data.
- **Forecasts** greatly improved model performance.
- Turbine location data and additional behaviours like **seasonality** were not explored in this phase.

Phase 1 - Visualizations







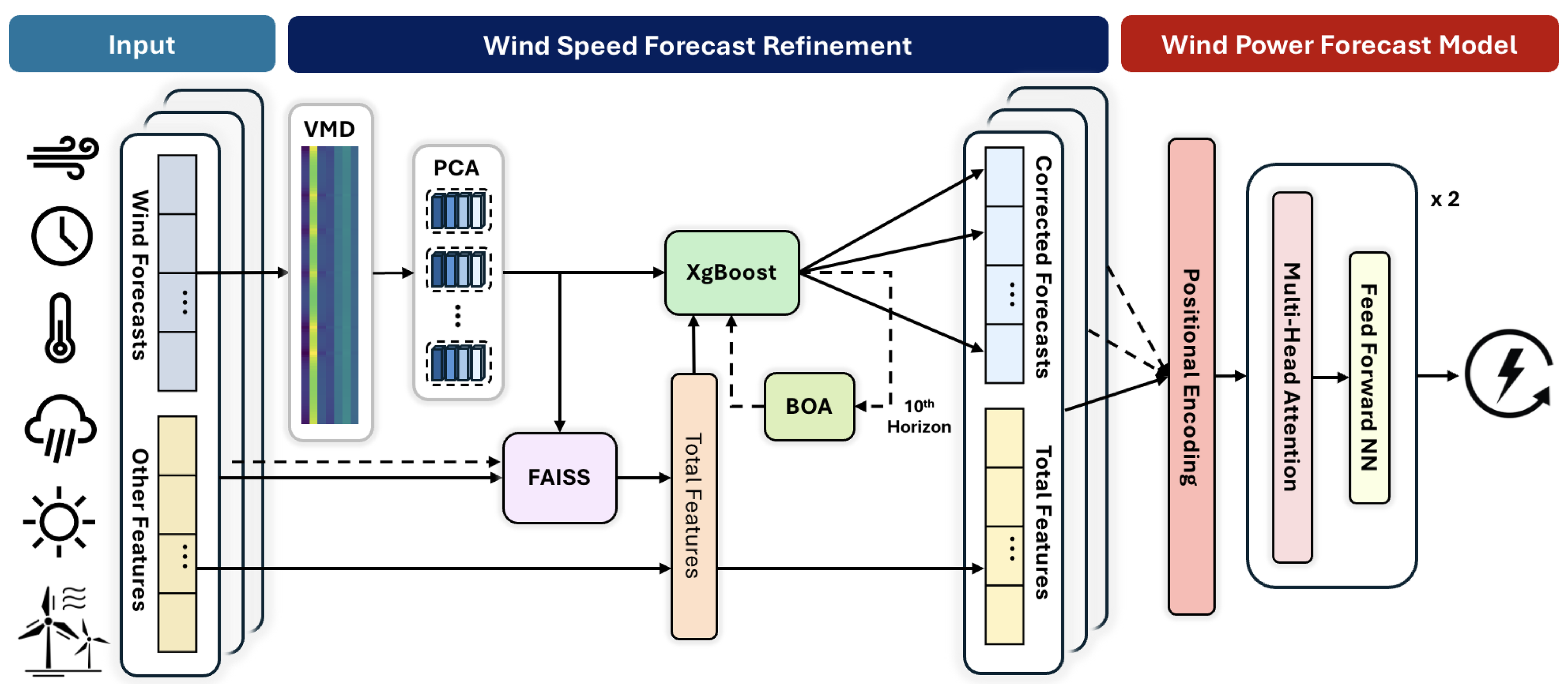
Phase 2

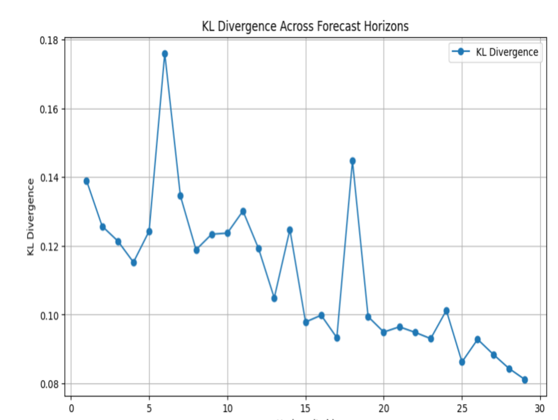
The goal was to refine wind power forecasts using ERA5 climate re-analysis data for a single turbine.

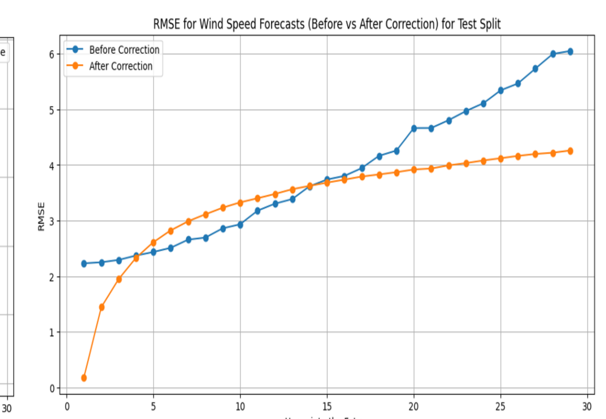
Key steps:
- **FAISS** based similarity search was carried out to obtain closest historical data.
- **ERA5** wind speed forecasts, along with additive noise, improved the model by **reducing reliance** on turbine-specific data and increasing generalization. Farm-level weather features like temperature and humidity helped reduce noise in the predictions.
- **Dimensionality Reduction**: After applying Variable Mode Decomposition (**VMD**) on the forecast data, dimensionality was reduced using **PCA**, simplifying while preserving key information.
- A **multi-regressor XgBoost** model was trained using the reduced PCA features and additional weather data to predict wind speed with a **count:Poisson** objective. **Bayesian Optimization** was used to fine-tune the model's hyperparameters around the chosen pivot point at 10^{th} horizon from the **KL divergence** stability.
- The model showed **robustness to noise**, leading to improved ultra-short-term forecast accuracy.
- A **transformer** model with multi-headed attention was also introduced due to its position-based encoding.
- The refined model achieved an MAE of **146.57 KW**, showing a **68.7% improvement** over the baseline model.

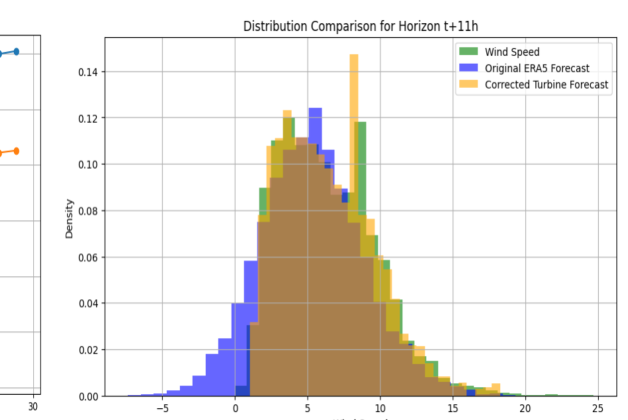
Brief Conclusion:
The model still hasn't captured **all the non-linearities**, a common issue with "real-world" time series forecasting. Despite higher architectural complexity, the improvement is is mostly in the **ultra-short term**. **More data** and more **complex models** would likely be needed to overcome these challenges.

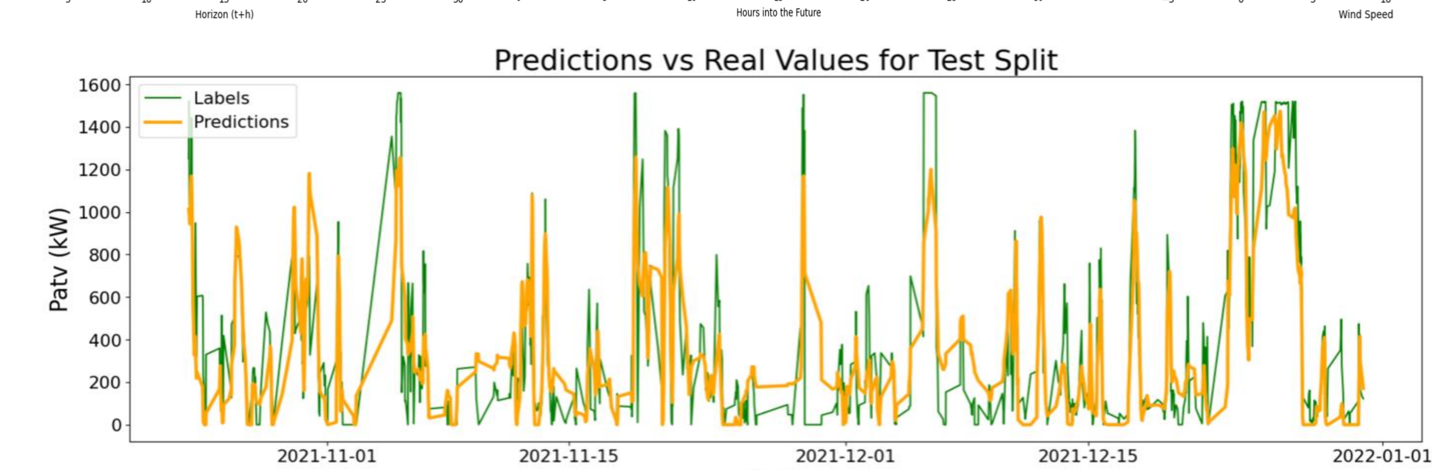
Phase 2 - Visualizations









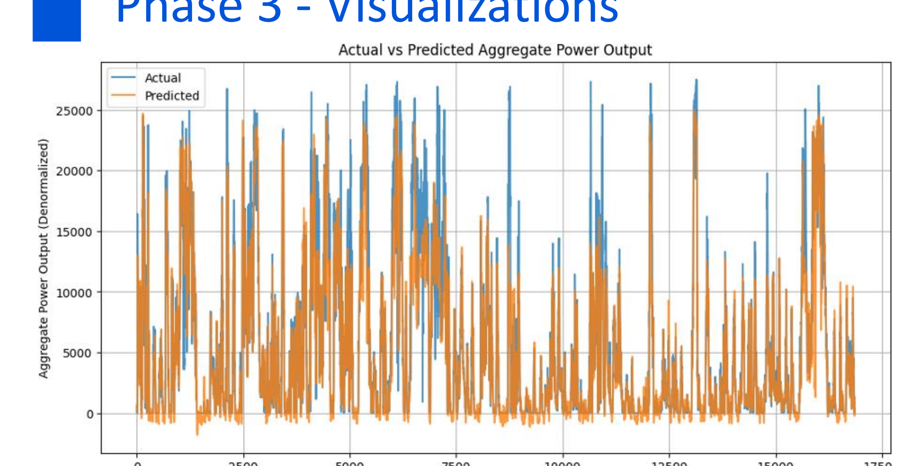


Phase 3

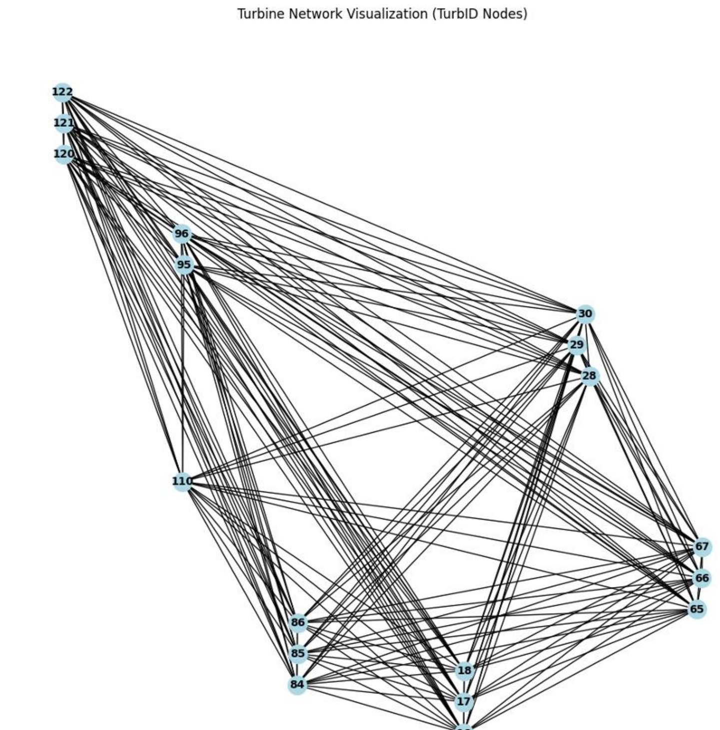
The goal was to model the spatio-temporal output of the wind farm using data from multiple turbines and their geolocations.
- The full dataset was reduced to **18 turbines**, and a **CNN-GRU** model was used to predict aggregated wind power output. Surprisingly, the model easily outperformed the single-turbine model, achieving an MAE of **1371.10 KW**.
- **Short-Term Forecasting**: The model was limited to 25-minute intervals due to time and memory constraints, with predictions independent of wind speed forecasts.
- A message-passing **graph neural network (GNN)** was used next, with **GCN** layers capturing spatial relationships and **GRU** layers handling temporal patterns. A fully connected layer mapped these to the power output.
- The model easily **overfits** due to **limited data** (1000 datapoints per turbine). The results are **promising** and using more data and further work can improve results.
- **Dynamic graph networks** with attention-based message passing, as explored in other works, may help reduce overfitting and better model turbine interactions.

Brief Conclusion: While the model showed promise, it requires more data and advanced techniques to address overfitting and improve performance.

Phase 3 - Visualizations



Turbine Network Visualization (TurbID Nodes)



Future work

Key areas for extending the project include:
- **Financial Impact Modelling**: Obtain LMP data to work with congestion, etc.
- **Ramp Forecasting**: Explore forecasting ramps and other time-series aspects.
- **False Data Injection Attacks**: Mitigate impacts of attacks on the model for real time operation.
- **Probabilistic Forecasting**: Use quantile loss (Pinball loss) for median-oriented MAE to reduce smoothing.
- **Additional Pivot Points**: Use Bayesian Optimization Algorithm (BOA) to generate more pivot points for better hyperparameter generalization.
- **Neural ODE Integration**: Integrate Neural ODEs with graph models for a hybrid physics-informed approach using Chebyshev polynomial-based graph convolution for dynamic spatial-temporal forecasting.