Hochschule Flensburg

University of Applied Science

Wind Energy Technology Institude

Research on wind Energy

Power Prediction and Energy Arbitrage Modelling for Wind Farms Using Machine Learning

Market Integration and Commercial Optimization, WiSe 2025/26

Master's degree in Wind Engineering, Flensburg, Germany

Submitted by

Karan Soni (760153)
Madhvesh Gorakhiya (730251)

Supervising Professors:

Prof. Dr. Andreas Heinen

Date of submission:

26.01.2025

1. Abstract

Objective, methodology, and key findings

Highlight ML approach and arbitrage results

2. Introduction

Wind energy is one of the fastest-growing renewable industries in Germany. It plays a critical role in reducing carbon emissions and achieving sustainability targets. [1] However, wind power generation is variable because wind speeds fluctuate continuously, making it difficult to predict power output accurately. This variability creates significant challenges for electricity markets, where prices also fluctuate strongly. Wind production is high during low-price periods, selling electricity immediately can reduce profits for wind-farm operators. (Add reference at here) Storing energy for later sale may increase revenue, but deciding when to store or sell is complex.

Machine Learning (ML) is good suited for improve forecasting and operational decisions. [2] ML models can capture nonlinear patterns in wind speed and power output, enabling more accurate predictions compared to traditional statistical methods. Forecasting is essential not only for grid stability but also for implementing energy arbitrage strategies, which allow operators to maximize profits during price fluctuations. [3] However, uncertainty in electricity prices and wind power output makes revenue prediction challenging. Selling at the wrong time or storing energy unnecessarily can lead to financial losses.

This report investigates two critical research questions:

1. Can machine-learning models accurately predict wind-farm power output and identify the optimal hours to STORE or SELL energy?
2. If low-price energy is stored as hydrogen, can the hybrid wind-hydrogen system remain economically profitable?

The primary objectives of this study are to develop and evaluate ML-based models for wind power forecasting and to assess the economic evaluation of hydrogen-based energy storage in arbitrage scenarios.

3. Literature Review

3.1 Wind Power Prediction

Traditional methods (physical models, statistical)

3.2 Energy Arbitrage

Market-based strategies, storage integration

3.3 Machine Learning in Wind Energy

Common algorithms (Regression, Neural Networks, Ensemble)

3.4 Gap Analysis

What existing studies lack and how your work addresses it

4. Data and Preprocessing

Data sources (wind speed, power output, market prices)

Data cleaning and normalization

Feature selection and engineering

5. ML Models

Machine Learning (ML) is increasing using in renewable energy systems to improve forecasting and operational decision-making. In this report, ML models help predict power output and support strategies for energy arbitrage, such as deciding whether to store energy or sell it. This study focuses on two models: Regression Model, which predicts wind power output based on input parameters and Classification Regression Model, which classifies operational decisions into STORE or SELL based on input parameters. These models will be explained in detail in the following subsections.

5.1 Regression model

Before this chapter, we have already described the source of the CSV file and the location of wind farm. In this study, we implemented two regression models, which are Linear Regression and Polynomial Regression. These models were chosen because they are simple and easy to implement. It is suitable for starting models for this research.

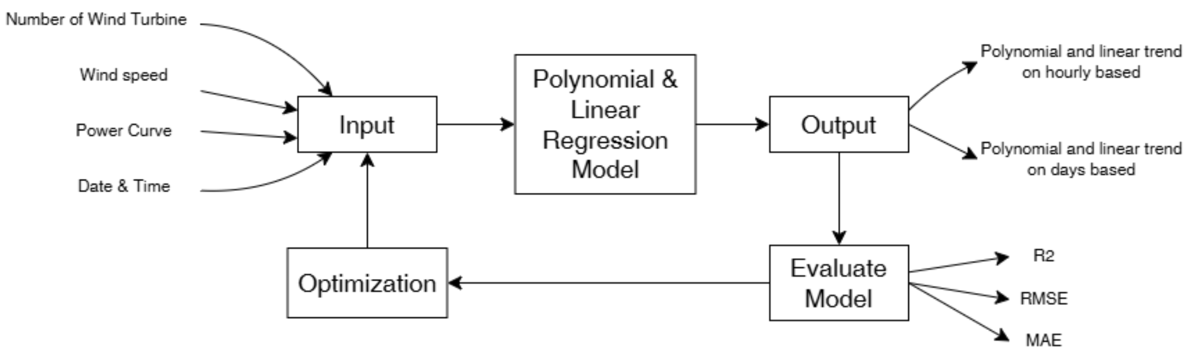

Figure 5. 1: Block Diagram of Wind farm Model (own)

In the flowchart, the modelling process considers four key input parameters are the number of wind turbines, wind speed, the power curve, and date/time information. These features were selected because they have a direct influence on wind power generation and are critical for accurate forecasting. After applying the ML models, we focused on generating two types of outputs. First, the polynomial and linear trends on an hourly basis. Second, , the polynomial and linear trends on a daily basis.

Once the models were trained and predictions obtained, we conducted an evaluation to determine which model performed best for the given dataset, which are discussed in detail in the Results section. Although optimization techniques could further improve model accuracy such as tuning or more input parameters or use different machine learning models, these were not implemented due to time constraints.

5.2 Classification Regression Model

The next model focuses on classification, specifically deciding whether to sell or store energy. For this task, we used a classification approach, and among the various available models, we selected Logistic Regression. This choice was made because Logistic Regression is simple, easy to implement, and requires fewer input parameters. The future, if the number of input parameters increases, more advanced models such as neural networks could be considered, as they offer greater predictive power for complex datasets.(give reference here)

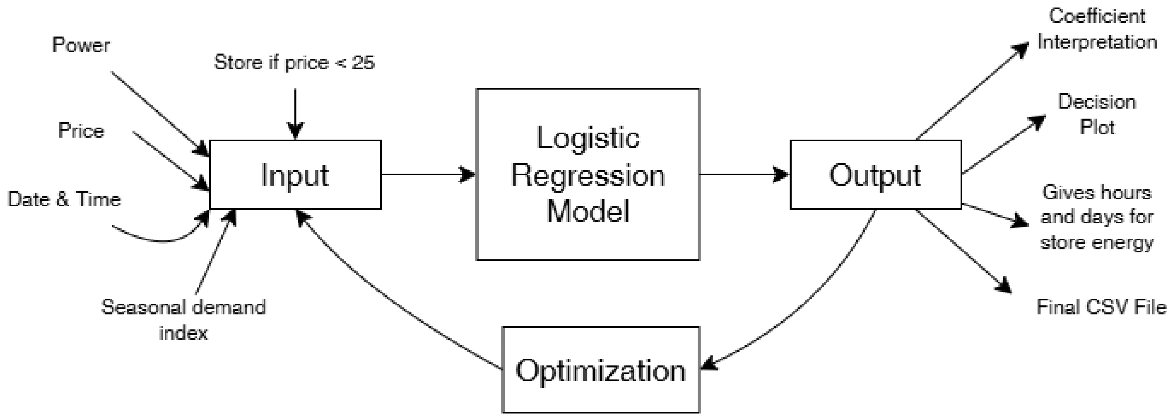

Figure 5. 2: Decision Arbitrage Model

The inputs for this classification model include the predicted power on an hourly basis (obtained from the regression models), electricity price from the CSV file, date and time, and seasonal demand (assumed for this study (see table 5.1)). It is important to note that real hourly seasonal demand data in future work would improve accuracy significantly. In figure 5.2, the Logistic Regression model processes these inputs to produce several outputs. First, coefficient interpretation, which indicates the influence of each factor on the decision. Second, a decision plot showing whether to sell or store energy. Third, recommended hours and days for storing energy. Last but not least, a final CSV file containing date, hour, predicted power, price, and the store/sell decision.

Season	Months	Factor
Winter	12, 1, 2	1.2
Spring	3, 4, 5	0.9
Summer	5, 7, 8	0.8
Autumn	9, 10, 11	1

Table 5. 1: Seasonal Demand

The evaluation of this model and its performance metrics will be discussed in the Results section, while additional details and supporting figures are provided in the Appendix.

6. Hydrogen Storage concept

- Role of hydrogen in energy storage
- Integration with wind energy and arbitrage
- Benefits and limitations

7. Results

8. Discussion

- Interpretation of results.
- Limitations of the study.
- Practical implications for wind farm operators.

9. Conclusion

- Summary of findings.
- Future work (e.g., real-time prediction, hybrid models).

10. References

- All cited papers, datasets, and tools.

11. Appendices

- Additional graphs, tables, or code snippets.

GitHub Link

On our GitHub homepage, we have attached our report, presentation, land leasing contract, and economic calculation Excel file. If you want to gain more knowledge about the topic, please check out our GitHub.