TIME SERIES ANALYSIS OF WIND ENERGY PRODUCTION IN TÜRKİYE

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

BACHELOR OF SCIENCE

in Energy Systems Engineering

by
Kutlay KIZIL

June 2025
İZMİR

ACKNOWLEDGMENTS

I extend my deepest thanks to my supervisor, Assoc. Prof. Dr. Ferhat BİNGÖL, whose expert guidance was paramount to this thesis. His dedication to my academic growth not only provided clarity in complex areas but also instilled a deeper passion for the research. I am very grateful to him for allowing me to be a part of project number 120N498 TÜBİTAK as a scholar, which is also an ERA-NET Smart Energy System project.

I am also grateful to the invaluable assistance provided by the research assistants: Faruk Tuna, Osmancan Denizli, and Ahmetcan Yetiş, whose support significantly facilitated various aspects of this project.

I would also like to express my gratitude to the Turkish Wind Energy Association (TÜREB) for providing valuable data concerning wind turbine specific data of the wind power plants in Türkiye, which was instrumental in developing certain aspects of the Wind Farm Database.

For further research, the underlying database created and analyzed during this study, as well as the project code, are available from the author upon request.

I am also deeply grateful for the financial support received from the Scientific and Technological Research Council of Türkiye (TÜBİTAK) through the 2209-A University Students Research Projects Support Program under project number 1919B012332867.

Finally, I would also like to acknowledge the assistance of Google's Gemini large language model in various aspects of the project's coding, which aided in the development and refinement of certain modules.

ABSTRACT

TIME SERIES ANALYSIS OF WIND ENERGY PRODUCTION IN TÜRKİYE

This study presents a comprehensive time series analysis of wind energy production in Türkiye, focusing on two primary objectives: quantifying the performance degradation of aging wind turbines and developing a robust short-term power forecasting model. A national-scale Wind Farms Database (WFD) was compiled, integrating publicly available data from Turkish institutions (EPIAS, EPDK, TUREB) with ERA5 meteorological reanalysis data.

The aging analysis, conducted on a filtered dataset of 74 wind farms over a four-year period (2020-2023), revealed a statistically significant mean annual capacity factor degradation of -0.90% (95% CI [-1.04%, -0.76%]). This degradation rate did not show a significant correlation with turbine manufacturer, installation year, or other physical plant characteristics within the analyzed timeframe.

For short-term forecasting, a Long Short-Term Memory (LSTM) network was developed to be used with autoregressive inference. The model demonstrated high predictive accuracy, achieving an average 1-hour-ahead normalized Mean Absolute Error (nMAE) of approximately 6% and an R^2 score of 0.91.

The study concludes by providing a crucial performance degradation benchmark for the Turkish wind fleet and a validated, efficient LSTM model for forecasting. It also highlights the critical performance gap between idealized academic studies using reanalysis data and real-world applications, suggesting future work on techniques like scheduled sampling to improve model robustness against the uncertainties of autoregressive inference.

ÖZET

TÜRKİYE GENELİNDE RÜZGAR ENERJİ ÜRETİMİNİN ZAMAN SERİSİ ANALİZİ

Bu çalışma, Türkiye'deki rüzgar enerjisi üretimine yönelik kapsamlı bir zaman serisi analizi sunmakta ve iki temel hedefe odaklanmaktadır: yaşlanan rüzgar türbinlerinin performans düşüşünü nicelemek ve güvenilir bir kısa vadeli güç tahmin modeli geliştirmek. Türk kurumlarından (EPİAŞ, EPDK, TÜREB) kamuya açık olarak temin edilen veriler ile ERA5 meteorolojik yeniden analiz verilerini entegre eden ulusal ölçekli bir Rüzgar Enerji Santralleri Veritabanı (WFD) oluşturulmuştur.

Dört yıllık bir dönemi (2020-2023) kapsayan ve filtrelenmiş 74 rüzgar santralinden oluşan bir veri seti üzerinde yürütülen yaşlanma analizi, istatistiksel olarak anlamlı, yıllık ortalama -%0.90'lık bir kapasite faktörü düşüşü olduğunu ortaya koymuştur (%95 GA [-1.04, -0.76]). Bu düşüş oranının, analiz edilen zaman dilimi içinde türbin üreticisi, kurulum yılı veya diğer fiziksel santral özellikleri ile anlamlı bir korelasyon göstermediği tespit edilmiştir.

Kısa vadeli tahmin için bir Uzun Kısa Süreli Bellek (LSTM) ağı geliştirilmiştir. Optimize edilmiş otoregresif model, 1 saatlik ileriye dönük tahminde yaklaşık %6'lık bir normalize Ortalama Mutlak Hata (nMAE) ve 0.91'lik bir R^2 skoru elde ederek yüksek bir öngörü doğruluğu sergilemiştir.

Çalışma, Türkiye'de rüzgar enerjisi için kritik bir performans düşüşü referans noktası ve verimli bir LSTM modeli sunarak sonuçlanmaktadır. Aynı zamanda, yeniden analiz verileri kullanan idealize edilmiş akademik çalışmalar ile gerçek dünya uygulamaları arasındaki kritik performans farkını vurgulamakta ve modelin otoregresif modelin dezavantajlarına karşı direncini artırmak için 'scheduled sampling' gibi teknikler üzerine gelecekteki çalışmalar için önerilerde bulunmaktadır.

TABLE OF CONTENTS

TABLE OF CONTENTS iii
LIST OF FIGURES v
LIST OF TABLES vii
CHAPTER 1. INTRODUCTION 1
1.1. Background And Literature Review 2
1.1.1. Wind Turbine Aging and Performance Degradation 2
1.1.1.1. Mechanisms of Wind Turbine Aging 2
1.1.1.2. Capacity Factor as a Performance Indicator 3
1.1.1.3. Geographic Focus & Limitations of Existing Studies 4
1.1.2. Wind Power Forecasting 4
1.1.2.1. Overview of Forecasting Methods 5
1.1.2.2. Machine Learning in Wind Power Forecasting 5
1.1.2.3. LSTM Network Architecture 6
1.1.2.4. Evaluation Metrics 7
CHAPTER 2. METHODOLOGY 9
2.1. Wind Farms Database 10
2.1.1. Data Sources 10
2.1.2. Database Structure 10
2.2. Turbine Aging Analysis 15
2.2.1. Data Preprocessing 15
2.2.2. Capacity Factor Calculation 17
2.2.3. Trend Analysis 18
2.3. Short-term Forecasting 19
2.3.1. LSTM Architecture 19
2.3.2. Input Features and Shape 21
2.3.3. Forecasting Methods 23
2.3.3.1. Single-Shot Forecast 24
2.3.3.2. Autoregressive Forecast 24
2.3.4. Evaluation Metrics 26
2.3.4.1. R-squared (R2) 26
2.3.4.2. Mean Absolute Error (MAE) 26
2.3.4.3. Root Mean Squared Error (RMSE) 27
2.3.4.4. Mean Absolute Percentage Error (MAPE) 27

2.3.4.5. Normalised MAE (nMAE) and Normalised RMSE (nRMSE)..... 27
CHAPTER 3. RESULTS..... 28
3.1. Aging of the Turbines..... 28
3.1.1. Overview of Turkish Wind Power Plant Fleet and Capacity..... 28
3.1.2. Analysis of Capacity Factor Degradation Trends..... 31
3.1.3. Influence of WPP Characteristics on Capacity Factor Trends..... 32
3.1.3.1. Influence of Turbine Brand on CF Trends..... 32
3.1.3.2. Influence of Installation Year on CF Trends..... 33
3.1.3.3. Influence of Other Factors on CF Trends..... 34
3.2. LSTM Model Performance..... 37
3.2.1. Justification of the Autoregressive Forecasting Preference..... 37
3.2.2. Hyperparameter Optimization and Model Selection..... 42
3.2.2.1. Analysis on the Representative Wind Farm..... 43
3.2.2.1.1. Time Steps (Lookback Period)..... 43
3.2.2.1.2. Network Layer Cells..... 43
3.2.2.1.3. Learning Rate..... 44
3.2.2.1.4. Batch Size..... 45
3.2.2.1.5. Number of Epochs..... 46
3.2.2.1.6. Dropout Rate..... 47
3.2.2.2. Adaptive Training Strategy..... 48
3.2.2.3. Multi-Farm Validation and Globally Optimal Configuration..... 49
3.2.2.3.1. Time Steps..... 49
3.2.2.3.2. Network Layer Cells..... 50
3.2.2.3.3. Learning Rate..... 50
3.2.2.3.4. Batch Size..... 51
3.2.3. Quantitative Performance of the Optimized Model..... 53
3.2.3.1. Overall Predictive Accuracy and Error Dynamics..... 53
3.2.4. Qualitative Analysis of Prediction Case Studies..... 55
3.2.4.1. Performance at Fixed Forecast Horizons..... 55
3.2.4.2. Analysis of an Individual Inference..... 59
3.2.4.2.1. Case Study A: Forecasting a Major Ramp-Up Event..... 59
3.2.4.2.2. Case Study B: Forecasting Through Volatility..... 60
CHAPTER 4. DISCUSSION..... 63
CHAPTER 5. CONCLUSION..... 65
REFERENCES..... 67
APPENDIX A CALCULATED TRENDS FOR THE WIND PARKS..... 71
APPENDIX B REASONS FOR WIND FARM REMOVAL FROM THE STUDY 73
APPENDIX C WFD ID-PLANT NAME MATCHES..... 74

LIST OF FIGURES

Figure **Page**

Figure 2.1. Illustration of DBSCAN……16

Figure 3.1. Scatter Plot of the Wind Farms Trends…….28

Figure 3.2. Box Plot of the Wind Farm Trends and the Individual Data Points (before outlier removal)……...29

Figure 3.3. Box Plot of the Wind Farm Trends and the Individual Data Points (after outlier removal)……...30

Figure 3.4. Distribution of Capacity Factor Trends……..31

Figure 3.5. Quantile-Quantile Plot of CF Trends and Normal Distribution……32

Figure 3.6. Box Plot of the Trend Distributions Grouped by Turbine Brands…..……..33

Figure 3.7. Box Plot of the Trend Distributions Grouped by Turbines' Installation Year……..34

Figure 3.8. Scatter Plot of Trend Values vs Mean WPP Elevation..……35

Figure 3.9. Scatter Plot of Trend Values vs Number of Turbines.……...35

Figure 3.10. Scatter Plot of Trend Values vs WPP's Capacity……36

Figure 3.11. Scatter Plot of Trend Values vs WPP's Capacity Density……...36

Figure 3.12. Error Metrics for 24 Hour Forecast Horizon (Autoregressive Inference)...38

Figure 3.13. Error Metrics for 24 Hour Forecast Horizon (Single-Shot Inference).…...39

Figure 3.14. Forecast Examples of a Single Inference (for Single-Shot)...…….40

Figure 3.15. Forecast Examples of a Single Inference (for Autoregressive)……...41

Figure 3.16. Error Metrics for Various Time Steps (single WPP)……...43

Figure 3.17. Error Metrics for Various Cell Layer Configurations (single WPP)……..44

Figure 3.18. Error Metrics for Various Learning Rates (single WPP)…….45

Figure 3.19. Error Metrics for Various Batch Sizes (single WPP)……..46

Figure 3.20. Error Metrics for Various Epochs (single WPP)……47

Figure 3.21. Error Metrics for Various Dropout Rates (single WPP)…….48

Figure 3.22. Error Metrics for Various Time Steps (All WPPs)……..49

Figure 3.23. Error Metrics for Various Cell Layer Configurations (All WPPs)……..50

Figure 3.24. Error Metrics for Various Learning Rates (All WPPs)……...51

Figure 3.25. Error Metrics for Various Batch Sizes (All WPPs)…….52

Figure 3.26. Averaged Error Metrics for All WPPs in Study……..53

Figure 3.27. Box Plot of the Average Error Metrics for All WPPs is Study……...54

Figure 3.28. Time Series View of the Forecast Horizons (t+1 to t+4)……56

Figure 3.29. Time Series View of the Forecast Horizons (t+5 to t+8)……56

Figure 3.30. Time Series View of the Forecast Horizons (t+9 to t+12)……..57

Figure 3.31. Time Series View of the Forecast Horizons (t+13 to t+16)……57

Figure 3.32. Time Series View of the Forecast Horizons (t+17 to t+20)……58

Figure 3.33. Time Series View of the Forecast Horizons (t+21 to t+24)……58

Figure 3.34. Examples of Singled-Out Inferences (Autoregressive)……...59

Figure 3.35. Examples of Singled-Out Inferences (Autoregressive)……...61

LIST OF TABLES

Table **Page**

Table 2.1. Structure of the Wind Farms Database……...12

Table 2.2. LSTM Parameters Selected as Initial Reference Point……...19

Table 2.3. LSTM Hyperparameters From Literature……...20

Table 3.1. Basic Capacity Factor Trend Statistics……...30

Table 3.2. Selected Hyperparameters Based on The Experimental Hyperparameter Optimization……52

CHAPTER 1.

INTRODUCTION

Every day, the world's need for renewable energy increases more than ever. Wind energy, being one of the major sources of renewable energy, plays a big role in satisfying this need. By the end of 2023, the global total capacity of renewables has reached 43% (3.9 TW), and 26.3% (1,017 GW) of this capacity was from wind energy (International Renewable Energy Agency, 2024). For Türkiye, the installed capacity for renewable energy sources accounts for 57.24% (63.16 GW) of the total capacity, and the wind energy's percentage among other renewables is 19.3% (12.19 GW) as of May 2024 (Turkish Ministry of Energy and Natural Resources, 2024).

Wind and solar energy have a big difference from other major sources of energy. Unlike the conventional power plants which can adjust their outputs on demand, wind and solar energy varies with certain things, which poses a significant challenge for grid operators. For wind energy, this is mostly based on the wind speed, while for solar energy, it's the irradiance. This fluctuating nature of wind and solar energy can create significant challenges, like grid balancing and frequency control. If the stability of the grid is not achieved, it can cause power outages or damage to the equipment, especially in regions where a smart grid does not exist yet, like Türkiye. The unpredictability of these energy sources can also negatively affect the electricity market operations, with the common trading markets being the Day-Ahead and Intraday markets. To briefly explain the difference between these markets, the Day-Ahead market allows electricity to be bought and sold for the following day, while the Intraday market enables trading of electricity within the same day to adjust for real-time fluctuations.

As a solution for the problem of unpredictability, one of the important methods is the forecasting solutions. Accurate forecasts are crucial for grid operators, because they help balance the supply and demand while optimizing the use of renewable resources. In addition to balancing, forecasting also allows for optimizing maintenance

schedules for wind farm owners. This way they can schedule the maintenance for the low periods, therefore maximizing both the productivity and profits.

It is known that the efficiency of wind turbines declines throughout the years of operation, due to mechanical reasons (Staffell & Green, 2014). This aging induced decline is an important factor to consider for investors, during the cost analysis stage of a wind turbine project. It is important to confirm this by analyzing the actual production data of the wind turbines and finding the aging trends. This analysis, together with implementing a wind power forecasting model, are the two major objectives of this research.

1.1. Background And Literature Review

1.1.1. Wind Turbine Aging and Performance Degradation

The increasing global installed wind capacity has led to a growing number of aging wind turbines, making the analysis of their performance decline with age a crucial area of research (Byrne et al., 2020). Similar to most technical systems, wind turbines experience a degradation in efficiency over their operational lifespan (Astolfi et al., 2021). Understanding the mechanisms driving this performance degradation, quantifying its extent, and assessing its economic impact are essential for optimizing the lifecycle management of wind energy infrastructure and ensuring the continued viability of wind power as a key renewable energy source.

1.1.1.1. Mechanisms of Wind Turbine Aging

Wind turbines are complex electromechanical systems operating under demanding and variable environmental conditions, leading to a multitude of aging mechanisms affecting their performance. These mechanisms can be broadly categorized into mechanical wear and tear, component degradation, environmental factors, and operational practices.

Mechanical wear and tear affect numerous rotating components. Gearboxes, for instance, experience failures due to stress cycles accumulated over years of operation

(Astolfi et al., 2020). The original gearbox of a Vestas V52 turbine analyzed by Byrne et al. (2020) operated for thirteen years before replacement.

Studies highlight that the generator insulation degradation is caused by prolonged exposure to temperature fluctuations and vibrations. Furthermore, the pitching and yaw systems, crucial for optimal power extraction, can experience wear in their mechanical and hydraulic components, leading to slower or less accurate adjustments(Byrne et al., 2020). The hydraulic station controlling the blade pitching system, if not well maintained, can have oil pressure issues affecting blade pitch response (Astolfi et al., 2021).

Environmental factors play a significant role in accelerating the aging process. Exposure to harsh weather conditions, such as high winds, extreme temperatures, humidity, and salt spray in coastal environments, can degrade materials and accelerate corrosion. Studies note that blade fouling by dirt, ice, and insects can impede aerodynamic performance (Staffell & Green, 2014).

Operational practices also influence the rate of aging. Suboptimal control strategies, frequent start-stop cycles, and inadequate maintenance can exacerbate wear and tear on turbine components (Byrne et al., 2020). The decision to run a gearbox until failure, as was the case in the study by Byrne et al. (2020), represents a specific operational strategy with implications for performance degradation.

1.1.1.2. Capacity Factor as a Performance Indicator

The capacity factor (CF) is a key metric used to quantify the actual energy production of a wind turbine or wind farm relative to its maximum potential energy production over a specific period, typically a year (Xu et al., 2023). It is calculated as the ratio of the actual energy production of the turbines to the energy they would have produced if it had operated at its rated power continuously during the same period. Expressed as a percentage, a higher capacity factor generally indicates better performance and efficiency.

In the context of wind turbine aging, the capacity factor provides a valuable indicator of performance decline over time. Analyzing trends in capacity factors allows for the assessment of the overall health and efficiency of wind assets as they age. However, it's important to note that changes in capacity factor can also be influenced by factors other than aging, such as variations in wind resources and technological upgrades (Xu et al., 2023). Therefore, methodologies that account for these confounding factors are crucial for accurately attributing changes in capacity factor to aging.

1.1.1.3. Geographic Focus & Limitations of Existing Studies

While the studies by Byrne et al. (2020) and Astolfi et al. (2020, 2021) focused on Ireland and Italy, offering some relevance to the Mediterranean region in the case of the Italian turbines, direct studies on wind turbine aging specifically focused on Turkiye are lacking in the literature. However, the findings from studies in regions with similar wind resources or environmental challenges could offer some comparative insights.

Several limitations are inherent in the current body of research on wind turbine aging:

- Data Availability and Quality: Access to long-term, high-resolution production data and detailed maintenance records can be limited. The quality and consistency of the available data can also impact the accuracy of the analysis.
- Turbine and Wind Farm Heterogeneity: Aging rates and patterns can vary significantly depending on the turbine model, manufacturing quality, site characteristics, operational practices, and maintenance history. Fleet-level studies may mask these individual variations, while single-turbine studies may lack generalizability.

1.1.2. Wind Power Forecasting

The increasing integration of wind energy sources into power grids necessitates reliable and accurate forecasting of wind power generation. The variability and intermittency of wind resources pose a significant challenge to maintaining the grid stability, optimizing energy trading, and planning operational schedules. Consequently,

substantial research efforts have been directed towards developing and refining wind power forecasting methodologies across various temporal scales and using diverse modeling techniques (YÜREK et al., 2021).

1.1.2.1. Overview of Forecasting Methods

Wind power forecasting methods can be broadly categorized into three main approaches: physical models, statistical models, and machine learning models (Foley et al., 2010). Physical models, such as Numerical Weather Prediction (NWP), utilize meteorological data and physical laws governing atmospheric conditions to simulate wind patterns and predict future wind speeds, which are then converted to power output based on wind turbine characteristics (Foley et al., 2012). These models excel in capturing the underlying physical dynamics but can be computationally intensive and require detailed site-specific information. Statistical models employ historical wind power data and time series analysis techniques, such as Autoregressive Integrated Moving Average (ARIMA) and exponential smoothing, to identify patterns and extrapolate future power generation based on past trends and seasonality (Wang et al., 2018). While computationally efficient and relatively easy to implement, statistical models may struggle to capture the complex non-linear relationships in the data (Taslimi Renani et al., 2016). Machine learning models, on the other hand, utilize algorithms that can learn complex non-linear relationships between the input features (e.g., meteorological data, historical power data) and the target variable (wind power output) through data-driven approaches (YÜREK et al., 2021). These models, including artificial neural networks (ANNs), support vector machines (SVMs), and decision trees, have increased popularity because of their abilities to handle high-dimensional data and capture complex dependencies.

1.1.2.2. Machine Learning in Wind Power Forecasting

Machine learning (ML) techniques have become increasingly prevalent in wind power forecasting due to their ability to model complex, non-linear relationships present in wind data. A variety of ML algorithms have been successfully applied to this domain, including:

- Artificial Neural Networks (ANNs): These models, inspired by the structure of the human brain, consist of interconnected nodes (neurons) organized in layers and can learn complex patterns from data (Memarzadeh & Keynia, 2020). Various ANN architectures, such as Multi-Layer Perceptrons (MLPs) and Recurrent Neural Networks (RNNs), have been used for wind power prediction (Oh et al., 2022).
- Support Vector Machines (SVMs): SVMs are powerful supervised learning algorithms that can be used for both classification and regression tasks. In wind power forecasting, SVMs can model non-linear relationships between wind parameters and power output by mapping the input data into a high-dimensional space using kernel functions (Zhang et al., 2019).
- Recurrent Neural Networks (RNNs): RNNs are specifically designed to process sequential data by maintaining an internal state that captures information from previous time steps. While traditional RNNs can struggle with learning long-range dependencies due to the vanishing gradient problem, their variants, particularly LSTMs and Gated Recurrent Units (GRUs), have shown significant promise in time series forecasting (Kisvari et al., 2021).

1.1.2.3. LSTM Network Architecture

Long Short-Term Memory (LSTM) networks are a specialized type of RNN architecture. They are designed to overcome the vanishing gradient problem and effectively learn long-range dependencies in sequential data (Zhang et al., 2019). An LSTM unit, or cell, consists of several key components:

- Cell State: The cell state acts as a memory bank, carrying relevant information across multiple time steps. It is updated and modified by the gating mechanisms.
- Input Gate: The input gate determines what new information from the current input should be stored in the cell state. It comprises a sigmoid layer that decides which values to update and a tanh layer that creates a candidate vector of new values.

- Forget Gate: The forget gate decides which information from the previous cell state should be discarded. It consists of a sigmoid layer that outputs values between 0 and 1 for each number in the cell state, where 1 means "keep this" and 0 means "get rid of this."
- Output Gate: The output gate controls what information from the cell state should be outputted at the current time step. It involves a sigmoid layer that decides which parts of the cell state to output and a tanh layer that creates a vector of values from the cell state, which is then filtered by the output of the sigmoid layer.

LSTMs are particularly well-suited for time series forecasting, including wind power prediction, due to several key characteristics, these are:

- Handling Non-linear Dependencies: Wind power generation is influenced by a multitude of factors, including wind speed, wind direction, temperature, and potentially other meteorological variables, exhibiting complex non-linear relationships. LSTMs, with their non-linear activation functions and layered architecture, can effectively learn and model these intricate dependencies.
- Capturing Long-Range Dependencies: Wind patterns and power generation can exhibit temporal dependencies that span across significant time intervals, such as diurnal and seasonal variations. The memory cells and gating mechanisms of LSTMs enable them to retain relevant information from distant past time steps, allowing them to capture these long-range dependencies more effectively than traditional RNNs, which are susceptible to the vanishing gradient problem.

1.1.2.4. Evaluation Metrics

The performance of wind power forecasting models is typically assessed using several common evaluation metrics. These metrics below provide quantitative assessments of the forecasting accuracy and they are used to compare the performance of different models and methodologies in wind power prediction. More detailed explanations and formulas of these metrics are given in the methodology section.

- Mean Absolute Error (MAE): It measures the average magnitude of the errors in a forecast, without considering their direction.
- Root Mean Square Error (RMSE): It calculates the square root of the average of the squared errors, giving a bigger weight to larger errors compared to MAE.
- Mean Absolute Percentage Error (MAPE): Expresses the average absolute error as a percentage of the actual values, providing a relative measure of forecast accuracy.
- R-squared (R^2): It represents the proportion of the variance in the dependent variable that is predictable from the independent variables, indicating the goodness of fit of the model.
- Normalized Mean Absolute Error (nMAE): The MAE value divided by a normalization factor (e.g., the range or mean of the target variable), allowing for comparison across different datasets or scales.
- Normalized Root Mean Square Error (nRMSE): The RMSE value divided by a normalization factor (e.g., the range or mean of the target variable), providing a normalized measure of the prediction error.

CHAPTER 2.
METHODOLOGY

At the core of this study lies the database created over the previous years, independently from this study, by us, as part of the TUBITAK project number 120N498. This database, which will be called the Wind Farms Database, or WFD from now on, consisted of the every possible publicly available data for the licensed wind farms located in Turkiye. WFD consisted of the following data before the start of this study.

1. General information about wind farms/turbines
2. Installed capacity changes for the wind farms throughout the years
3. Time series data of the plant level productions

During this study, weather re-analysis data for the same timeframes as the available production data of the wind plants is incorporated to the database, as well as some minor enchantments, like automation of the new data retrieval, and fixing the encountered mistakes for the general information of the wind farms.

For the weather re-analysis data, ECMWF Reanalysis v5 (ERA5) dataset is used. ERA5 is a re-analysis dataset, developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), provides hourly data for atmospheric, ocean and land cover variables, from 1940 to the present day (Hersbach et al., 2020). During the preliminary study of this research, the New European Wind Atlas (NEWA) was planned for use. NEWA, compared to ERA5, has a better resolution, but it lacks the globality of ERA5 and its data availability ends at the end of 2018. During the later analysis of the production data of the wind turbines, it is seen that the production data before 2018 was lacking considerably, and the reliability of the available data was questionable. For this reason, the NEWA dataset is removed from the scope of this study.

2.1. Wind Farms Database

2.1.1. Data Sources

The data present in the database comes from various sources. Below are all of the sources given in alphabetical order, the details about the exact data retrieved and their place in the database will be explained afterwards.

- Republic of Turkey Energy Market Regulatory Authority (EPDK)
- The Ministry of Energy and Natural Resources
- Turkish Wind Energy Association (TUREB)
- Energy Markets Operation Inc. (EPIAS)
- The ERA5 global reanalysis

2.1.2. Database Structure

WFD has the following tables (apart from the tables automatically created by SpatiaLite extension) with their brief explanation given. The full list of columns in these tables are also given at the table below.

- “wf”: Mostly consist general information about the plant like license number and license dates, name, address, city/district, capacity information (from EPDK), unique identifiers for EPİAŞ
- “wf_turbine_coordinates”: Coordinates of the turbines found at EPDK
- “wf_border_coordinates””: Border point coordinates of the plants found at EPDK
- “wf_turbine”: Information about installed power, turbine brand, model and power information and the project investor, all of which is from TUREB
- “productions”: Production data from EPIAS, which consist of two different types of productions called ‘real time’ (wind_rt) and ‘amount of data to be used as basis for reconciliation’ (wind_uevm), as well as solar production counterparts. There is also the availability information, also taken from EPIAS, for both the wind and solar parts of the plants. In this study, the “wind_uevm” data is used, because it was observed that the realtime data was not as reliable,

frequently exhibiting more noise and missing values compared to the reconciled data.

- “ministry_of_energy”: Consist of the broken down by date version of the given licenses, as well as the unit number and powers of the given licenses.
- “timestamp_id”: Consist of the unique identifiers given to each timestamp, which are stored as text data in ISO 8601 format. This helps us reduce the size of the database by storing the timestamps as integers in other tables.
- “era5_raw_date”: This table contains the ERA5 data taken directly from the downloaded ERA5 dataset.
- “era5_calculated_data”: This table contains the weather data calculated using the data from the ERA5 dataset.

Tables	wf	wf (continued)	wf (continued)	wf_turbines
Columns	wf_id	installed_power_mechanical	epias_plant_name	wf_id
	license_number	installed_power_electrical	epias_plant_shortname	investor
	plant_name	capacity_mechanical	epias_organization_etso_code	company_name
	start_date	capacity_electrical	epias_organization_id	project_name
	end_date	email	epias_organization_name	installed_power
	license_holder	geom_polygon	epias_organization_shortname	turbine_brand
	tax_no	license_status	epias_info_lock	turbine_model
	contact_address	battery_capacity_mwh	epias_uevm_plant_id	turbine_power
	phone_number	battery_power_mwe	epias_uevm_plant_name	start_date_of_operation
	plant_type	epias_region	epias_uevm_plant_eic	turbine_number
	city	epias_plant_id	epias_uevm_plant_shortname	
	district	epias_plant_eic		

Table 2.1. (cont. on next page): Structure of the Wind Farms Database

Tables	ministry_of_energy	productions	wf_border_coordinates	wf_turbine_coordinates
Columns	wf_id	wf_id	wf_id	wf_id
	license_date	ts_id	point_index	turbine_index
	license_number	wind_rt	k	u
	company_name	wind_uevm	km	um
	plant_name	wind_eak	kx	ux
	city	solar_rt	ky	uy
	unit_power_electrical	solar_uevm	index	geom
	unit_number	solar_eak	geom	
	additional_unit_power_electrical			
	acceptance_date			

Table 2.1. (cont): Structure of the Wind Farms Database

Tables	era5_raw_data	era5_raw_data (cont')	era5_calculated_data	timestamp_id
Columns	ts_id	sea_surface_temperature	ts_id	ts_id
	latitude	u10n	latitude	timestamp
	longitude	v10n	longitude	
	temperature	fg10	ws100	
	pressure	i10fg	wd100	
	dew_point	surface_latent_heat_flux	ws10	
	surface_sensible_heat_flux	boundary_layer_dissipation	wd10	
	u100	boundary_layer_height	air_density	
	v100	charnock	relative_humidity	
	u10	forecast_surface_roughness		
	v10	friction_velocity		
	mean_sea_level_pressure	land_sea_mask		

Table 2.1. (cont): Structure of the Wind Farms Database

2.2. Turbine Aging Analysis

2.2.1. Data Preprocessing

The production data of a wind farm is often not reliable as its own. There are several reasons for this; it could be that some of the turbines in a farm might require maintenance, as any mechanical devices do, or it could be a malfunction of a component, or it could be simply a curtailment from the grid operator. Whatever the reason is, these events create unreliable data for doing an analysis.

For the reasons above, we have implemented a filtering method using the installed capacity (data from the ministry_of_energy table) and the availability information (data from the productions table) we have, followed by an outlier detection and removal method by using the wind speed data for the wind farm that is being analyzed, and a density based clustering algorithm called Density-Based Spatial Clustering of Applications with Noise, or simply DBSCAN (Ester et al., 1996).

The DBSCAN algorithm mainly uses two parameters called “esp” and “min_samples”. The “esp” parameter determines the maximum radius length that is considered when scanning the neighbors of a point. And the “min_samples” parameter determines the number of points present inside the area determined by the “esp” parameter. There are of course other, optional, parameters, but they are left as the default values for this study.

The DBSCAN algorithm has the following procedure for detecting clusters:

- Initially, all points are labeled as ‘unvisited’
- For each of the ‘unvisited’ points, the algorithm iteratively goes through all of them and do:
    - Finds all the other points within the radius “esp”
    - If the amount of points (other than the visited point) is equals or greater “min_samples”, then this point is labeled as a “core point”
- For each “core point”, assign them into a cluster if they are not already.
- Expand the cluster recursively by adding all of the density-reachable points.

- Points left out are marked as “noise”

The figure below is a great example of how the points are marked after the algorithm finishes. The points named ‘A’ are the core points, the ‘B’ and ‘C’ points, while not being a core point (they do not satisfy the “min_samples”), are still considered inside the cluster, and the point ‘N’ is either outside of the ‘esp’ distance and also not a core point.

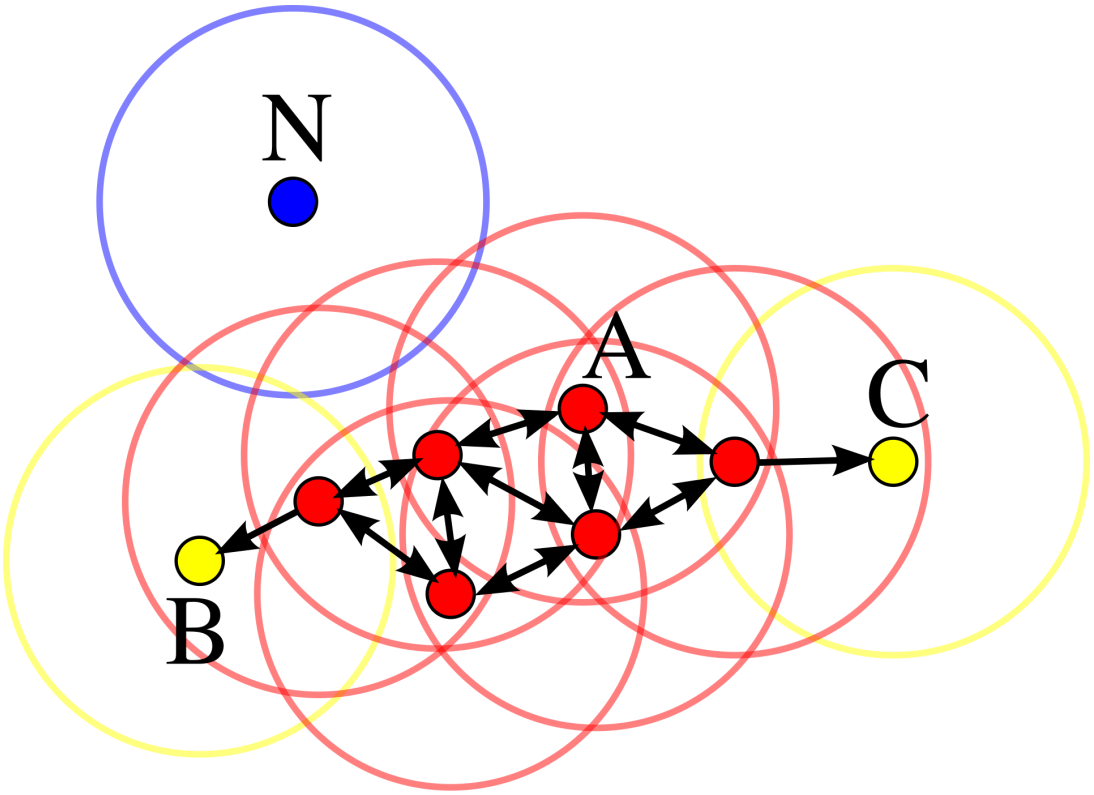

Figure 2.1. Illustration of DBSCAN, by Chire, 2021, CC BY-SA 3.0, https://commons.wikimedia.org/wiki/File:DBSCAN-Illustration.svg

With the preprocessing pipeline in place, each wind farm in the database is carefully investigated for the quality of the data they have. The oldest available data has the start timestamp of 2019-01-01, and during our observations, we chose to disregard the data from 2019, and instead used the data starting from the start of 2020 until the end of 2023, a period of four years.

2.2.2. Capacity Factor Calculation

Capacity factor (CF) is an important metric in the energy industry. It is the ratio of the actual energy produced to the maximum energy that can be produced if the plant was able to. In layman's terms, the capacity factor tells us how a power plant produces energy compared to its potential. CF is calculated with the formula below:

$$Capacity\ Factor\ (\%) = \frac{Actual\ Energy\ Generated\ (MWh)}{Installed\ Capacity\ (MW) \times Time\ (h)} \quad (1)$$

We have the actual energy generation data from EPIAS for every hour, and the installed capacity information comes from the ministry of energy, where they provide the acceptance day of the production license for each wind farm. This information is used to calculate the capacity factor at each hour, for every available timestamp in the production data. We chose not to save this information back into the database since the data from the ministry of energy was not easy to match the up-to-date list of wind energy licenses, and occasionally required us to manually assign the acceptance information onto the current list of farms. Also, calculating the capacity factors on the fly does not require a notable processing power, and enables us to resample the data into hourly, daily, monthly, etc. resolutions with ease.

We also experimented with calculating the capacity factor based on the availability data (instead of removing the data where the availability is lower than a set percentage), but the availability data was not reliable and had more frequent problems where the availability was showing lower than the production for a given time, and there was many time frames where the data was showing a low availability, but the production was even lower or the plant was even not producing any power. The frequency of this issue returned us back into using the production license of the plant.

It is also important to note that some wind plants have a bigger installed power than the production license they have. This is usually distinct by calling the former mechanical power (MWm) and the latter the electrical power (MWe). In an ideal scenario, the mechanical power is what we are looking for, but for our situation, the production data we have is limited by the electrical power of the wind plant. If we use

the mechanical power, we would never see the production reach that highest levels, and when we use the electrical power, the capacity factor is calculated higher than what it should be. For the wind plants where the mechanical and electrical power is equal to each other, this is obviously not a problem, but for many others, this problem exists, and using the mechanical power is unfortunately not an option. Since the capacity factor is used to calculate the trend throughout the years, as will be explained in the following section, this is not a major problem.

2.2.3. Trend Analysis

The primary goal of the trend analysis is to examine how much the wind farms degrade in their performance over the years. The trend is calculated using a simple, first-degree linear regression model. Before calculating the trend, the hourly capacity factors are resampled into a quarter-end frequency. By doing this we aimed to smooth out the short term fluctuations while retaining enough data points. The equation for the linear regression model is the following:

$$\widehat{CF} = a + bt \tag{2}$$

Where:

- $\widehat{CF}$ is the average capacity factor,
- a is the y-intercept,
- b is the slope,
- t is the time in years.

And the trend percentage is calculated as:

$$Trend\ (\%) = \frac{CF_{end} - CF_{start}}{t_{end} - t_{start}} \times 100 \tag{3}$$

Where:

- CF_{end} is the capacity factor at the end of the time period,
- CF_{start} is the capacity factor at the beginning of the time period,

- t_{end} is the time value at the end of the time period,
- t_{start} is the time value at the beginning of the time period.

2.3. Short-term Forecasting

In this study, a multi-layered LSTM network was developed for short-term wind power forecasting. The model developed, uses a sequential architecture based on the Keras deep learning library.

2.3.1. LSTM Architecture

In this study, the core of the LSTM model is left unmodified, in other words, the vanilla LSTM model is in use. That is, the activation functions of the model's gates are left unchanged.

The main changes to the model are made by adjusting the hyperparameters and the model's layers. For the number of the layers and which layers to use, we started by researching what parameters the other studies for wind power forecasting uses, which can be seen in Table 2.2. Based on this, we choose the following parameters to be used initially as a reference point.

Parameter	Value
Train/Test Ratio	0.9/0.1
Layers & Neuron	128/64/32 + Dense layer
Optimization Function	Adam
Epochs	0.001
Batch Size	128
Dropout	0.2

Table 2.2. LSTM Parameters Selected as Initial Reference Point

Length of the Input Data	Train/Test ratio	Number of Neurons	Optimization Function	Epochs	Learning Rate	Batch Size	Number of Layers	Dropout	Lookback	Article
1200	0.8/0.2	200, 100	Not mentioned	Not mentioned	0.0001-0.1	10-256	2			(Memarzadeh & Keynia, 2020)
Not mentioned	0.7/0.3	500, 100, 50	Rmsprop	10	0.001	30	3			(Shahid et al., 2021)
2908 days	Not mentioned	4/8/16/32/64/72/128/256	Adam/SGD/RMSProp	200/250/500/1000	0.005	128	3			(Demirtop & Sevli, 2024)
8759	0.8/0.2	64	Adam	50	Not mentioned	128	4	0.25	1, 5, 10	(Shao et al., 2021)
17500	0.7/0.3	75, 30, 30	Adam	10	Not mentioned	64	3		1	(Yoldaş, 2022)

Table 2.3. LSTM Hyperparameters From Literature

2.3.2. Input Features and Shape

For the training of the LSTM model, the weather data which affects the wind production are needed to be selected. The wind speed is the first variable that is a very obvious one, but it's known that the temperature, pressure and wind direction also have the potential to improve the forecast accuracy (Geerts, 1984).

We chose to include wind speed and wind direction, as well as the other weather parameters, but increasing the number of features would result in the training taking much longer time, and it was harder to create a forecast model that could be considered good-enough. For this reason, air density, which is calculated with the temperature, pressure and the dew point temperature, is calculated and used. To calculate the air density, the relative humidity would need to be calculated first, both are calculated with the following formulas for relative humidity and air density, respectively.

$$RH = 100 \times exp(\frac{17.625 \times dew\_temp - 273.15}{243.04 + dew\_temp - 273.15}) \div exp(\frac{17.625 \times temperature - 273.15}{243.04 + temperature - 273.15}) \quad (4)$$

$$X_v(t, P, RH) = \frac{RH}{100} \times (\alpha + \beta + P + \gamma \times t^2) \times \frac{exp(C_1 \times tempK^{\ 2} + C_2 \times tempK + C_3 + \frac{C_4}{tempK})}{P} \quad (5)$$

$$Z(t, P, RH) = 1 - \frac{P}{tempK} \times [\alpha_0 + \alpha_1 \times t + \alpha_2 \times t^2 + (b_0 + b_1 \times t) \times X_v + (c_0 + c_1 \times t) \times X_v^{\ 2}) + \frac{P^2}{tempK^2} \times (d + e \times X_v^{\ 2})] \quad (6)$$

$$\rho(t, P, RH) = \frac{P \times M_a}{Z(t,P,RH) \times R \times tempK} \times (1 - X_v \times (1 - \frac{M_v}{M_a})) \quad (7)$$

Where;

- $\alpha = 1.00062$
- $\beta = 3.14 \times 10^{-8}$
- $\gamma = 5.6 \times 10^{-7}$
- $C_1 = 1.2378847 \times 10^{-5}$
- $C_2 = -1.9121316 \times 10^{-2}$

- $C_3 = 33.93711047$
- $C_4 = -6.3431645 \times 10^{3}$
- $\alpha_0 = 1.58123 \times 10^{-6}$
- $\alpha_1 = -2.9331 \times 10^{-8}$
- $\alpha_2 = 1.1043 \times 10^{-10}$
- $b_0 = 5.707 \times 10^{-6}$
- $b_1 = -2.051 \times 10^{-8}$
- $c_0 = 1.9898 \times 10^{-4}$
- $c_1 = -2.376 \times 10^{-6}$
- $d = 1.83 \times 10^{-11}$
- $e = -0.765 \times 10^{-8}$
- $R = 8.314472$
- $M_a = 28.96546 \times 10^{-3}$
- $M_v = 18.01525 \times 10^{-3}$

For the wind speed and wind direction, instead of calculating each with the u and v wind vectors, it was decided to use the wind vectors as it is. Also, among the two different data for the wind vectors in ERA5, that is the wind vectors at 10 meters and 100 meters (the name of which are u10 & v10, and u100 & v100), the one at 100 meters elevation are used, since the wind turbines are typically closer to the 100 meters than they are to 10 meters elevation.

Apart from these three weather data, u100, v100 and air density, the production data from the previous hour(s) is also used as an input feature for the model. This is a common practice in neural network based prediction models, and it's used to help the model understand the temporal dependencies, it's also called a lagged feature, and can be set to any other feature in the input. While this can improve the forecast errors significantly, it comes with a drawback of skewing the predictions to the value of the real production input value, therefore resulting in bad predictions for the next predictions. The details of this are explained further in the Forecasting Methods section.

In a short summary, the features used are the following:

- Wind component u at 100 meters
- Wind component v at 100 meters
- Air density
- Production value of the present time

For the LSTM model, the input array needs to be 3 dimensional. The first dimension is called “samples”, which is the number of sequences fed into the model. The second dimension is called the “timestep”, and it determines the length of each sequence. And the third and the last dimension is called “features”, which is the number of features in each timestep.

The practical meaning of these dimensions changes based on the characteristics of the input data. For example, when using LSTM for predicting the next word in a sequence of words, setting the input data so that the time step dimension is 5, would mean that the model is predicting the next word using the last 5 words in a sequence. But for weather data, knowing the weather a few more hours before the prediction does not play a significant role *compared to* word prediction.

In this study, for the similar reasons given at the aging analysis part, only (up to) 4 years of data is used for each wind farm, which corresponds to roughly 35,040 hours of data points. Which then split into train and test with respect to the 0.8/0.2 split ratio. With the four features chosen, this corresponds to an array shape of “[28032, 1, 4]”. If the time steps value is wanted to be increased, the values of the features from the time-step amount of samples before needs to be put into the second dimension of the input array, which then would become a shape of “[28032, time-step, 4]”, the important thing to notice here is that the sample size is not reduced and the created time-steps are just the copies of the earlier samples.

2.3.3. Forecasting Methods

There are two major methods to use when making a prediction using a neural network, single-shot forecasting and autoregressive forecasting. These methods differ in how they use trained LSTM model’s output to put together a multi-step prediction. As the wording ‘multi-step’ suggests, this is only applicable if a forecast of more than an hour is needed, which is most often wanted and how power forecasting happens in real world applications. If only the next step ahead is predicted, then it is called a single step prediction.

2.3.3.1. Single-Shot Forecast

When predicting with the single-shot method, the prediction model is needed to be trained for the amount of steps ahead as the length of the prediction wanted. If the next 3 hours is what is desired, then the model needs to be able to predict the production in all 3 hours into the future, i.e. (t+1), (t+2), and (t+3). If the production of the future 24 hours is needed, then the model needs to predict all 24 hours at once, and so on.

Arguably, the most distinctive feature of the single-shot method is, it requires a different model, trained specifically for the wanted forecast horizon, so that the model 'learns' during the training, that it must predict the specified horizon all at once, using the input it's given. This *can* make the single-shot model better at predicting the fixed horizon future if the model is designed correctly, but the need for training different models for different needs also increases the cost of making a prediction.

One more possible downside of the single-shot method is, when the forecast horizon is too long, the model's quality may be worse because of the difficulty in directly mapping the input to distant future productions during the training, therefore forcing the model to 'generalize' more.

2.3.3.2. Autoregressive Forecast

Also called iterative forecasting, this method requires a more simple but also accurate-at-what-it-does model. This method often only predicts the next time step (t+1) based on the current input sequence, and does a very good job at it. The working principle of the autoregressive method can be explained as:

- Initial Input: The model starts with the initial input sequence (e.g., meteorological data and lagged production value(s)) for the current time step (t).
- First Prediction: The model predicts the output for the next time step, t+1.
- Input Update: The (known) meteorological values of the time t+1 is combined with the first prediction.
- Iterative Prediction: The model uses this updated input sequence to predict the output for time step t+2

- Repeat: Steps 3 and 4 are repeated until predictions for all desired time steps (up to the forecast horizon) are generated.

  This process can also be represented mathematically as:

  $$\hat{y}_{t+k} = f(X_{t+k-1}, \hat{y}_{t+k-1}),\ for\ k = 1, 2, 3..., H \quad (8)$$

  Where:

  - $\hat{y}_{t+k}$ is the predicted production at the time step $t + k$.
  - $f()$ represents the LSTM model.
  - X_{t+k-1} represents the meteorological input features at the time step $t + k - 1$.
  - H represents the forecast horizon.

The biggest disadvantage of the autoregressive method is the error accumulation. While the LSTM model itself, trained for predicting only the next hour, is good at what it does, it will still make errors. And since the LSTM model is trained with the known, real production values during the training phase, feeding the prediction back into the input will be seen as the ground truth for the model. There are of course methods for overcoming this, such as data augmentation to introduce noise to the input data during training, using separate error correction models, or using techniques like professor forcing or scheduled sampling, which the latter is discussed at the discussion part.

In this study, while the Python code is designed to be used with both single-shot and autoregressive prediction, the main focus is the iterative prediction. The reason for selecting this is mostly because of the computational limitations. To be able to evaluate the single-shot method with for/with different forecast horizons, it is needed to train a different model for each horizon, for many hyperparameter combinations (because we can not justify using the same ideal hyperparameters used for autoregressive), and for many different wind farms.

2.3.4. Evaluation Metrics

To quantitatively evaluate the performance of the LSTM model, a range of standard evaluation metrics were employed. These metrics provides a comprehensive assessment of the model's accuracy and also its ability to capture the characteristics of wind power production.

2.3.4.1. R-squared (R^2)

R-squared, which is also known as the coefficient of determination, calculates the proportion of the variance in the dependent variable (wind power production) that is predictable from the independent variables (input features).

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y})^2} \quad (9)$$

Where:

- y_i is the actual production.
- $\hat{y}_i$ is the predicted production.
- $\bar{y}$ is the mean of the actual production values

The R^2 provides a measure of the overall fit of the model. It ranges from 0 to 1, with higher values indicating a better fit.

2.3.4.2. Mean Absolute Error (MAE)

MAE measures the mean magnitude of the errors in a set of prediction values, without considering their direction. It is the average of the absolute differences between predicted and actual values.

$$MAE = \frac{1}{N}\Sigma\left|y_i - \hat{y}_i\right| \quad (10)$$

MAE is expressed in the same units as the data itself. Lower values indicate better performance and a value of zero means no error at all.

2.3.4.3. Root Mean Squared Error (RMSE)

RMSE is the square root of the average of squared differences between the predicted and actual values.

$$RMSE = \sqrt{\frac{1}{N}\Sigma(y_i - \hat{y}_i)^2} \quad (11)$$

RMSE is also expressed in the same unit as the data itself, and lower values also indicate a better performance. RMSE penalizes higher errors more than lower errors, because the errors are being squared before being averaged.

2.3.4.4. Mean Absolute Percentage Error (MAPE)

MAPE expresses the accuracy as a percentage of the error.

$$MAPE = \frac{100}{N}\Sigma\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (12)$$

MAPE is expressed in percentage, lower values mean better performance. MAPE gives a relative measure of error, which means that it's useful for comparing the performance of different datasets (i.e. different wind plants).

However, it is often problematic for the datasets used in this study. That is because, MAPE can be very large or become undefined when actual values are close to zero. Since we are dealing with hourly data, which has many hours where the production is zero or very close to it, MAPE is often not useful and it is not calculated for the results.

2.3.4.5. Normalised MAE (nMAE) and Normalised RMSE (nRMSE)

nMAE and nRMSE are the normalized versions of MAE and RMSE, allowing for comparison across different scales. In this study. Both are normalized using the capacity of the respective wind farm.

CHAPTER 3.

RESULTS

3.1. Aging of the Turbines

3.1.1. Overview of Turkish Wind Power Plant Fleet and Capacity

To assess the overall trend of wind turbine performance degradation in Türkiye, we analyzed the aggregated capacity factor data from a final set of 75 wind farms, spanning the period from the start of 2020 to the end of 2023. The set of wind farms with their corresponding reasons for removal is given at the end in Appendix-2.

The Figure 3.1 below shows the trend data for each of the remaining wind farms, with the mean and three standard deviation lines. While there are several farms with a trend higher than zero, one of them near the plus three standard deviation line does raise some suspicions.

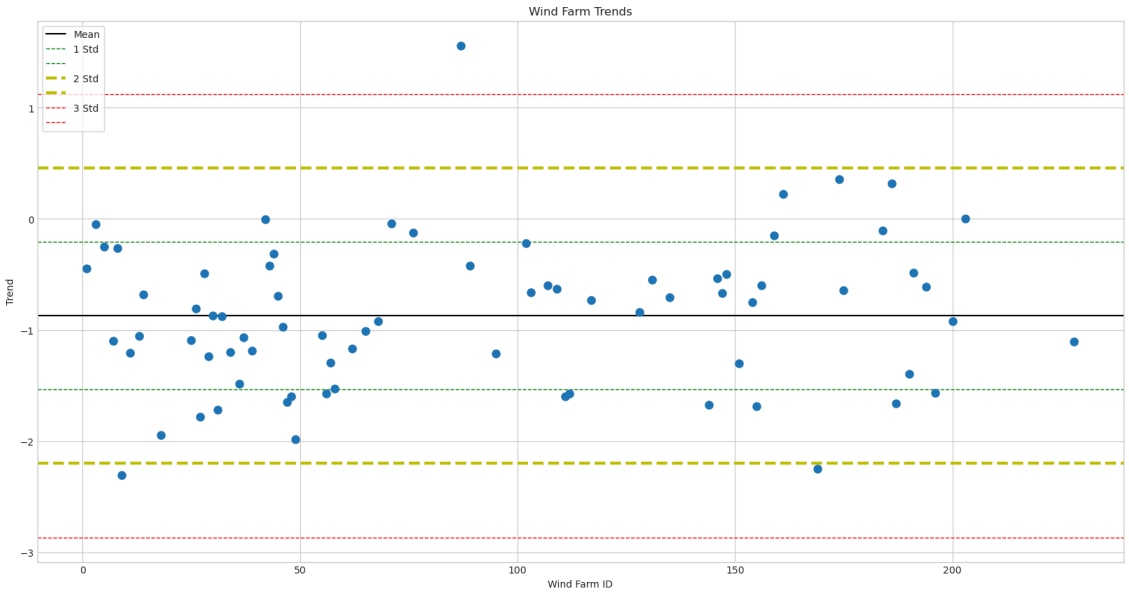

Figure 3.1. Scatter Plot of the Wind Farms Trends

While the production data of these wind farms are filtered based on their quality and visually inspected, there is always the possibility of some other reason behind this deviation for this farm, which is why for outlier detection Tukey's method or more

generally know, the 1.5*IQR rule is applied (Tukey, 1977). The Figure 3.2 below shows the box plot and the distribution of the trends for the wind plants.

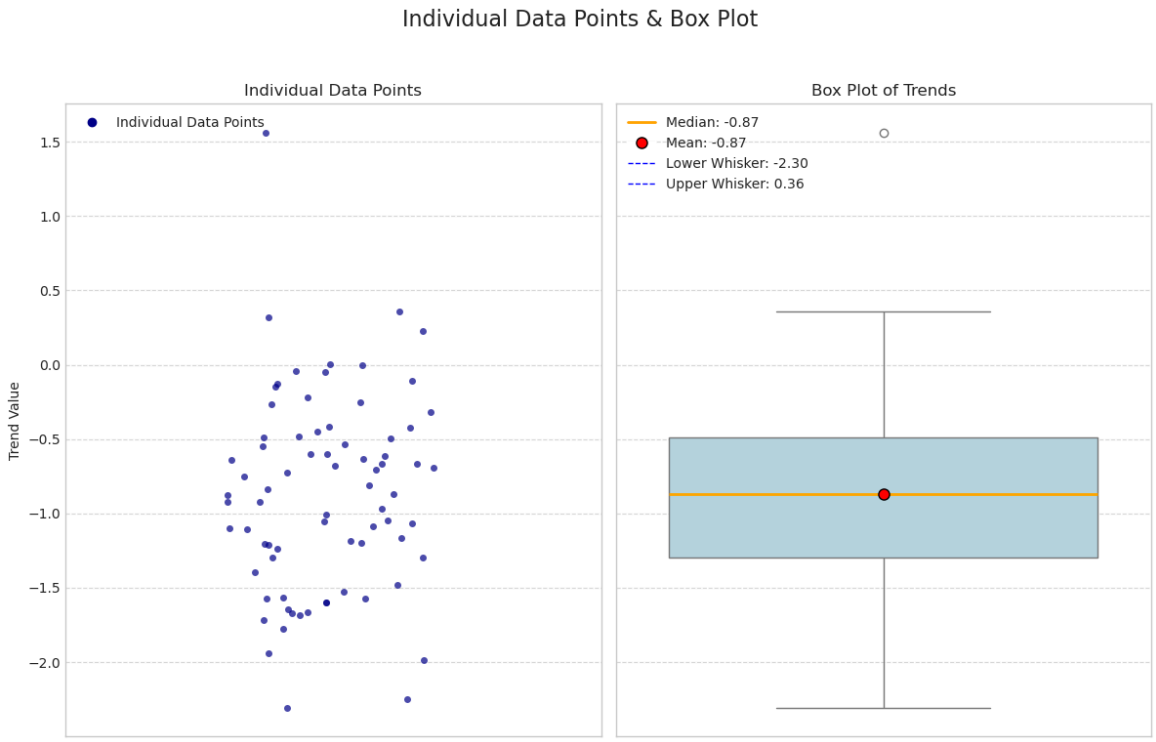

Figure 3.2. Box Plot of the Wind Farm Trends and the Individual Data Points (before outlier removal)

After removing the outlier, there are no trend values left outside of the 1.5 IQR range and all data points are left within the whiskers, as can be seen in the Figure 3.3 below.

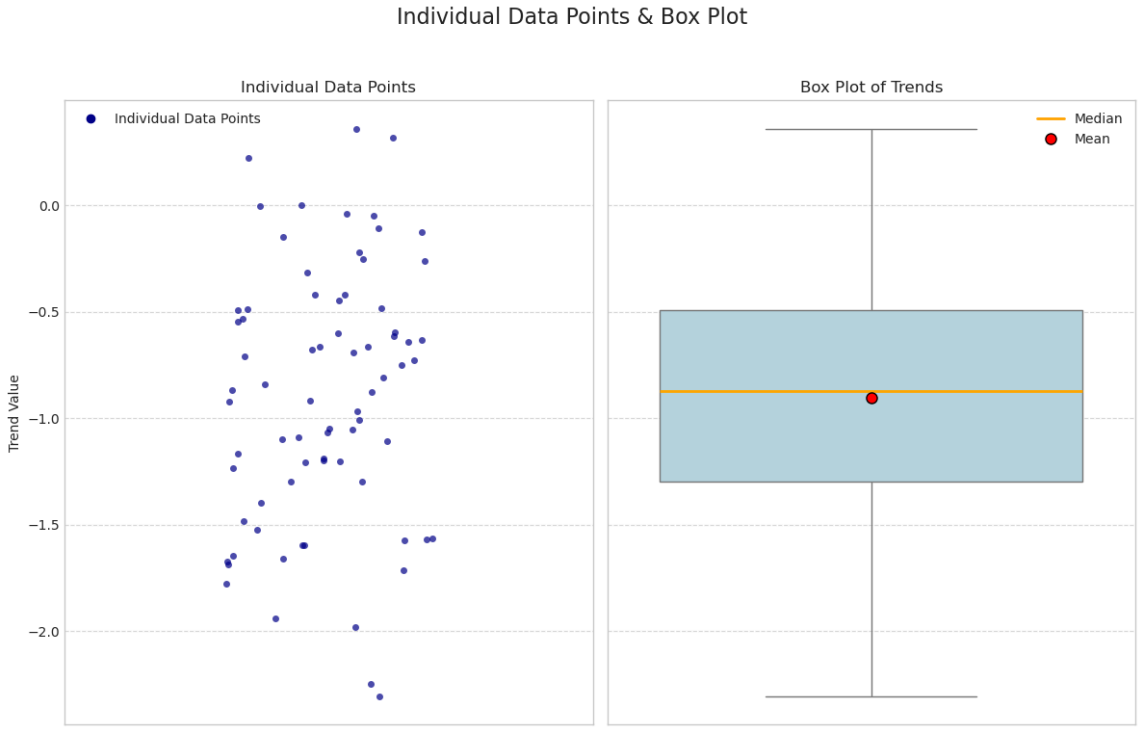

Figure 3.3. Box Plot of the Wind Farm Trends and the Individual Data Points (after outlier removal)

Statistics of the trend data for the 75 wind farms, before and after the outlier removal, are also given on the table 3.1. below. The complete list of trends for these wind farms (including the outliers) can be found in the Appendix-A.

	CF Trends (%/Year) (including outliers)	CF Trends (%/Year) (excluding outliers)
mean	-0.87	-0.90
std	0.67	0.61
min	-2.30	-2.30
25%	-1.30	-1.30
50%	-0.87	-0.87
75%	-0.49	-0.49
max	1.56	0.36

Table 3.1. Basic Capacity Factor Trend Statistics

3.1.2. Analysis of Capacity Factor Degradation Trends

The final analysis was conducted on a filtered and outlier-removed sample of 74 wind farms, derived from a total population of 297 licensed farms present in the database. Based on standard sample size calculations, a sample of this size is sufficient to represent the entire population with a 95% confidence level and a margin of error of approximately 10%, assuming maximum variability. This robust sample provides a strong statistical foundation for the study's findings.

A one-sample t-test performed on the trend values from these 74 farms revealed a mean degradation of -0.90% per year. This result is statistically significant ($t(74) = -12.74$, $p < .001$, 95% CI $[-1.04, -0.76]$). The 95% Confidence Interval indicates that one can be 95% confident that the true average annual degradation rate for the entire fleet lies between -1.04% and -0.76%. As this interval does not include zero, it corroborates that the observed degradation is highly unlikely to be due to random chance. The validity of the t-test is further supported by the data's approximately normal distribution, as shown in the histogram in the Figure 3.4 and the Quantile-Quantile plot in the Figure 3.5 below.

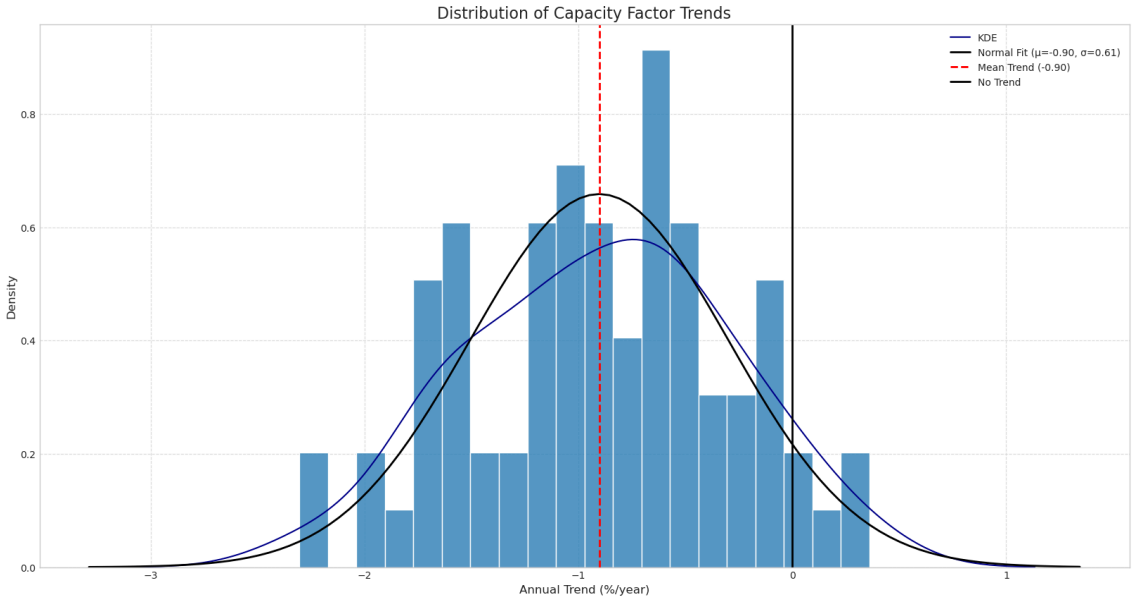

Figure 3.4. Distribution of Capacity Factor Trends

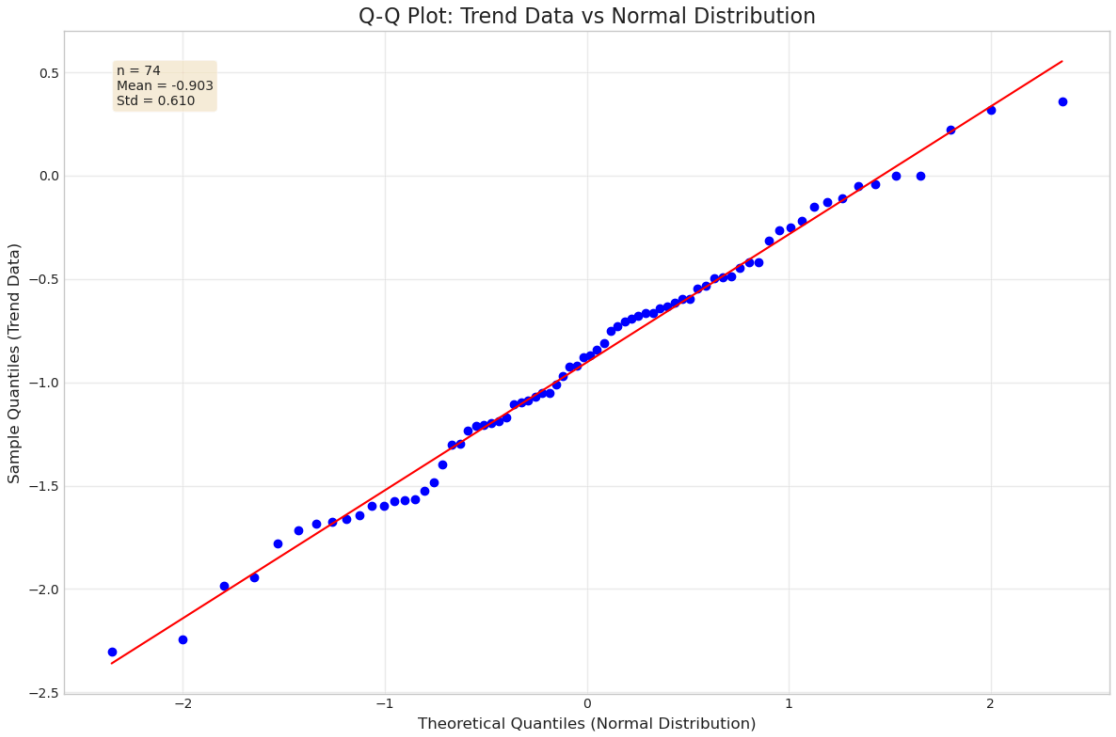

Figure 3.5. Quantile-Quantile Plot of CF Trends and Normal Distribution

3.1.3. Influence of WPP Characteristics on Capacity Factor Trends

Following the analysis of the overall degradation trends, this section investigates whether these trends are influenced by specific wind power plant characteristics. The analysis explores factors ranging from the turbine brand to the operational timeline and physical layout (elevation, scale, and density). To ensure the validity of the findings, each analysis is performed on the most appropriate subset of the data.

3.1.3.1. Influence of Turbine Brand on CF Trends

To analyze the influence of the turbine manufacturer on long-term performance, the capacity factor (CF) trends were grouped by brand. To ensure a direct and fair "apples-to-apples" comparison, this analysis was performed on the subset of 69 wind farms that exclusively utilize turbines from a single brand. This approach eliminates the confounding effects of sites with mixed-brand technologies.

The boxplot in Figure 3.6 visualizes the distribution of trends for each manufacturer. Visually, while there are minor differences in the median trends and the spread of the data between brands, no single manufacturer stands out as having a dramatically different performance degradation profile.

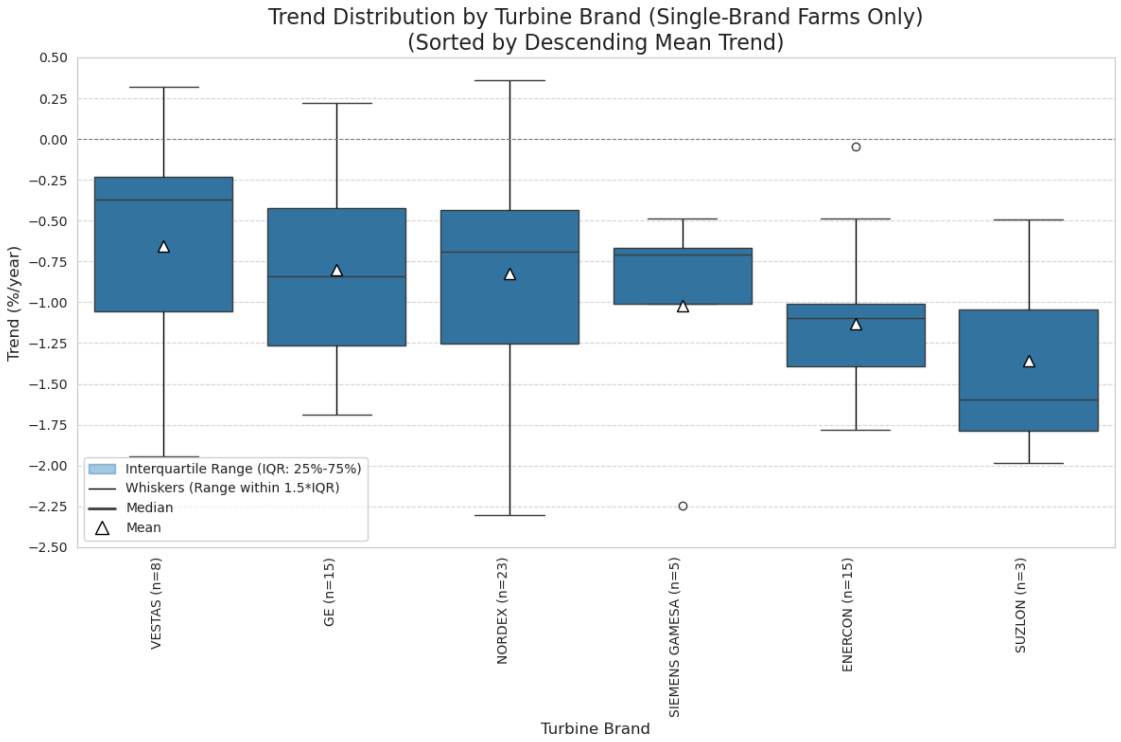

Figure 3.6. Box Plot of the Trend Distributions Grouped by Turbine Brands

To quantitatively assess if these visual differences are statistically significant, an Analysis of Variance (ANOVA) test was conducted. The results yielded an F-statistic of 1.1798 and a corresponding p-value of 0.3291. Since this p-value is substantially greater than the standard significance level of 0.05, it indicates that there is no statistically significant evidence to conclude that the mean degradation rates differ among the turbine brands in this dataset.

3.1.3.2. Influence of Installation Year on CF Trends

To test whether a wind farm's commissioning year affects its recent performance degradation rate, the analysis used the most strictly filtered dataset available. This included only the 32 "homogeneous" farms, which have both a single turbine model and

a single installation year, providing the most unambiguous data for this specific question.

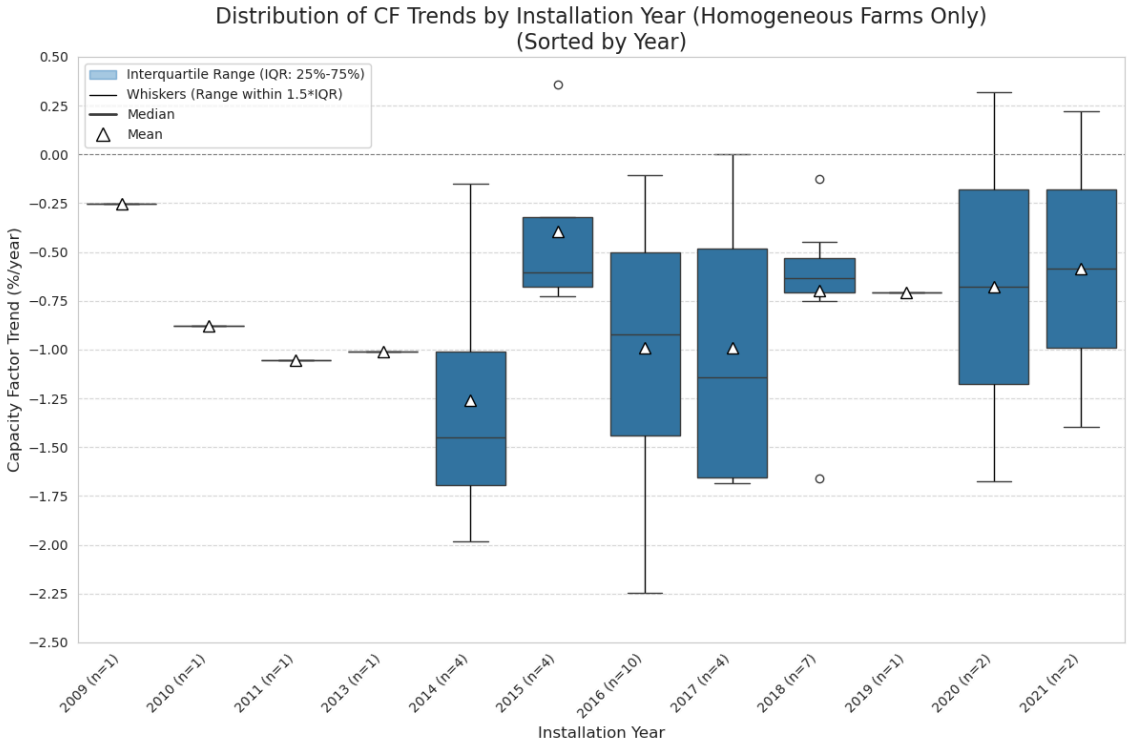

Figure 3.7. Box Plot of the Trend Distributions Grouped by Turbines' Installation Year

An ANOVA test was performed across these installation year cohorts, resulting in an F-statistic of 0.4827 and a p-value of 0.8926. This very high p-value indicates that there is no statistically significant evidence of a difference in mean degradation trends among the different installation year groups. This quantitative result reinforces the visual finding from the boxplot and supports the conclusion that, for the 2020-2024 period analyzed, the operational age of the turbines was not a driver of their yearly degradation rate.

3.1.3.3. Influence of Other Factors on CF Trends

Finally, the impact of various physical and scale-related characteristics was investigated. This analysis was performed on the full dataset of all 74 unique wind farms. The Figures 3.8, 3.9, 3.10 and 3.11 shows the scatter plots for each characteristic investigated.

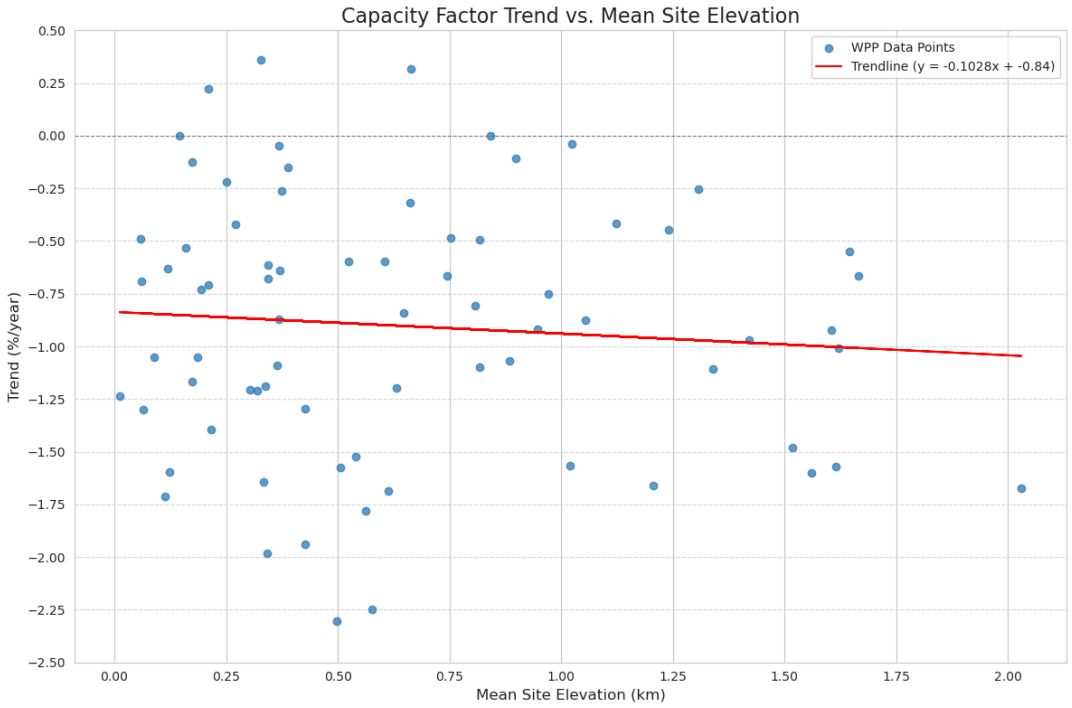

Figure 3.8. Scatter Plot of Trend Values vs Mean WPP Elevation

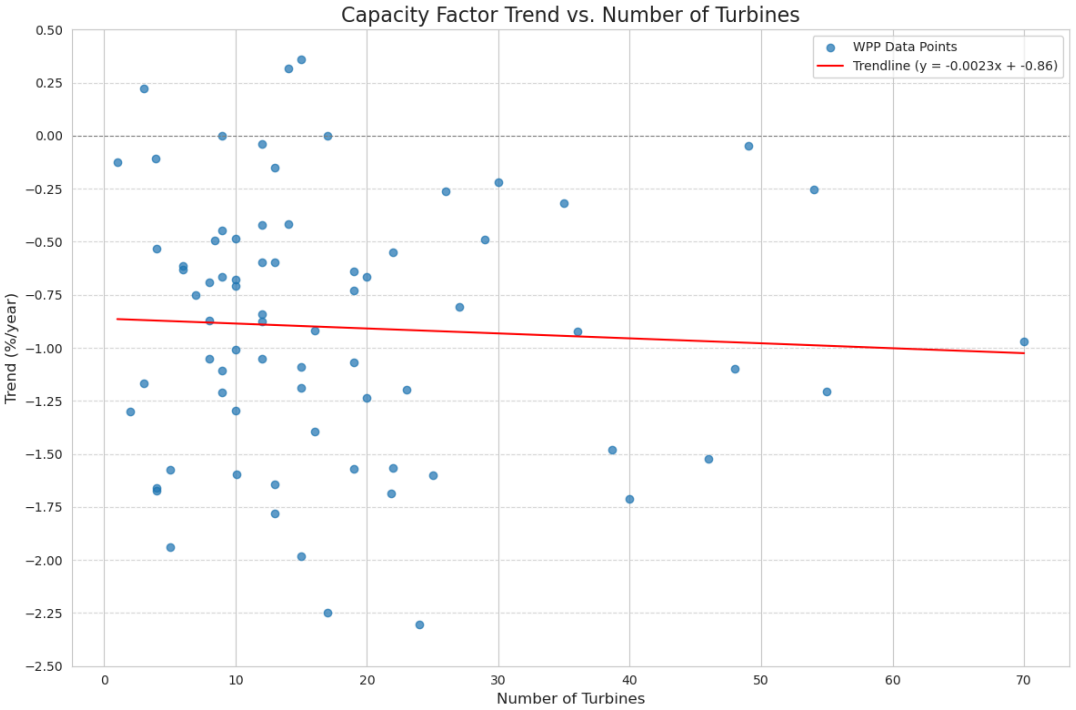

Figure 3.9. Scatter Plot of Trend Values vs Number of Turbines

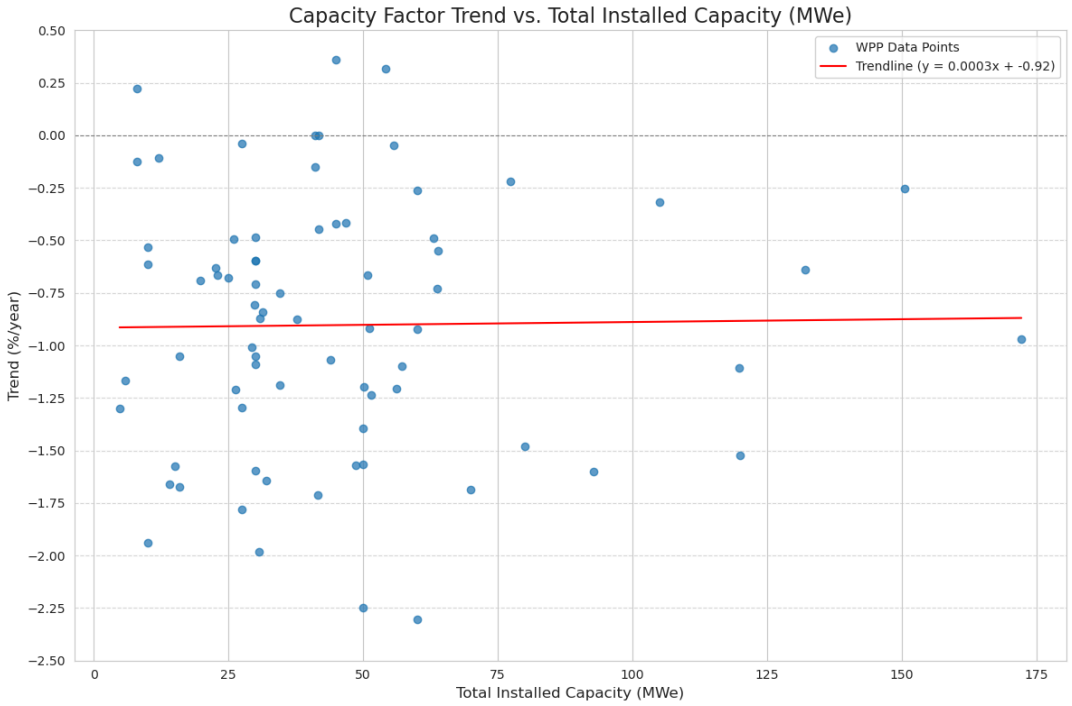

Figure 3.10. Scatter Plot of Trend Values vs WPP's Capacity

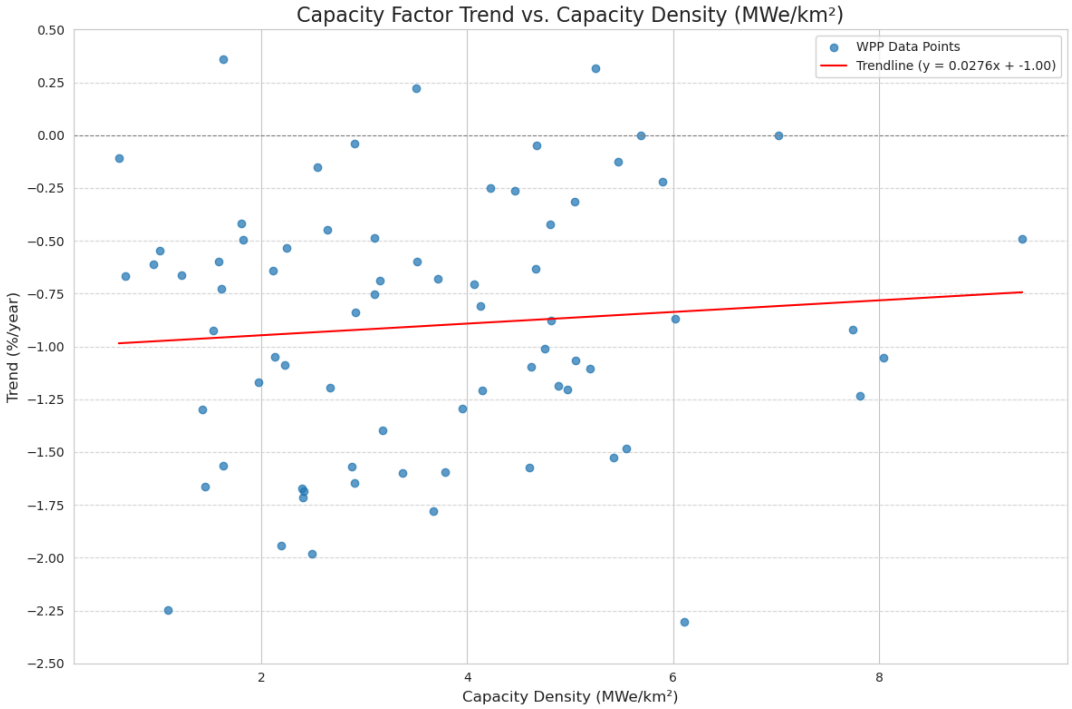

Figure 3.11. Scatter Plot of Trend Values vs WPP's Capacity Density

Visual inspection of the scatter plots for CF trend versus mean site elevation, number of turbines, installed capacity, and capacity density reveals no apparent relationship. To confirm this quantitatively, each characteristic was binned into 'Low' and 'High' groups at its median value, and a two-sample t-test was performed to compare the mean trends of these groups.

The results consistently showed no significant differences. The p-values for all tests were well above the 0.05 significance threshold: Mean Site Elevation ($p = 0.926$), Number of Turbines ($p = 0.291$), Total Installed Capacity ($p = 0.525$), and Capacity Density ($p = 0.537$). These findings quantitatively confirm the lack of correlation observed in the scatter plots.

3.2. LSTM Model Performance

This section presents a detailed evaluation of the Long Short-Term Memory (LSTM) models developed for wind power forecasting. The analysis begins with a high-level comparison of the primary forecasting strategies, followed by an in-depth examination of performance across different time horizons, the influence of key hyperparameters, and variations among individual wind farms.

3.2.1. Justification of the Autoregressive Forecasting Preference

In developing the forecasting model, two primary strategies were considered: the single-shot and the autoregressive methods. As detailed in the Methodology chapter, the single-shot approach requires training a distinct, and often more complex model for each forecast horizon. This leads to a substantial increase in computational demand and training time, particularly when scaling the analysis across multiple wind farms and hyperparameter configurations. Given the available computational resources, pursuing an exhaustive single-shot analysis was deemed impractical for the scope of this project.

However, to provide an empirical basis for selecting the most suitable strategy, a comparative study was conducted on a representative case. For this purpose, two LSTM models with identical architectures, input variables, and hyperparameters were trained. The key distinction lay in the output layer and training procedure:

- **Autoregressive Model:** Utilized a dense layer with a single output (dense=1), trained for 100 epochs, to predict the next hour's generation. This single model was then used iteratively to produce a 24-hour forecast.
- **Single-Shot Model:** Utilized a dense layer with 24 outputs (dense=24) to predict the entire 24-hour horizon at once. Due to the increased complexity, an initial training of 100 epochs yielded poor results; therefore, the training was extended to 500 epochs to achieve a reasonably well-trained model.

The performance of both strategies over the 24-hour forecast horizon is presented in Figure 3.12 and Figure 3.13 below.

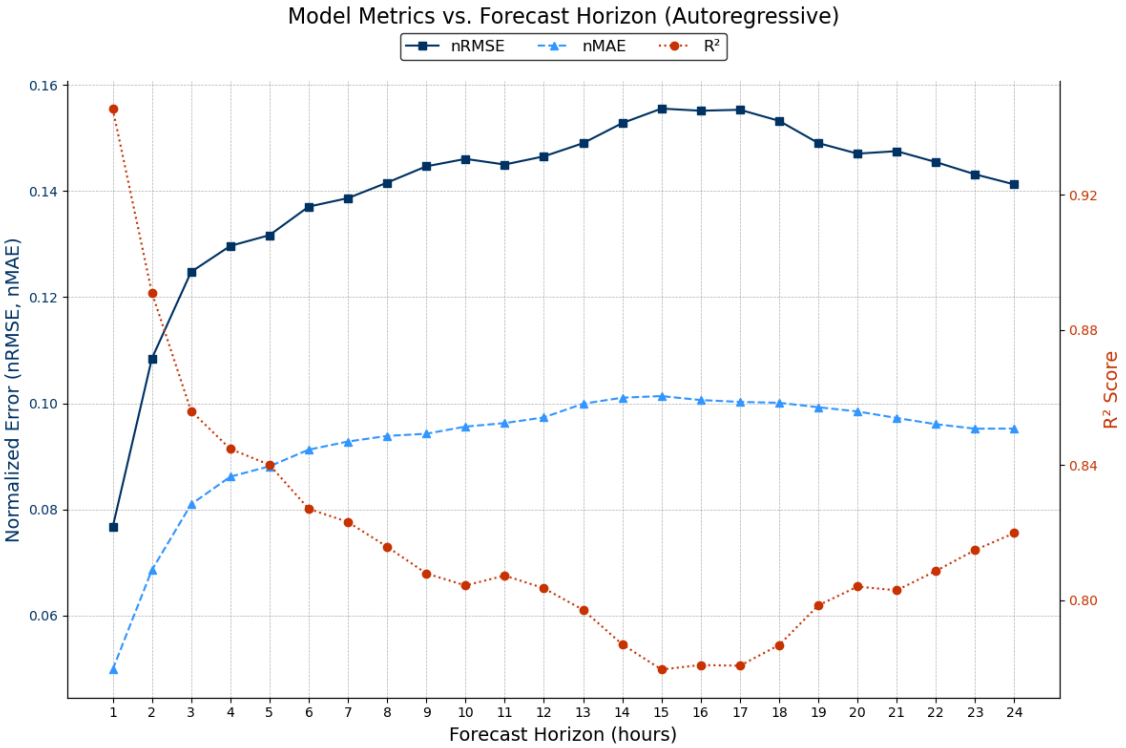

Figure 3.12. Error Metrics for 24 Hour Forecast Horizon (Autoregressive Inference)

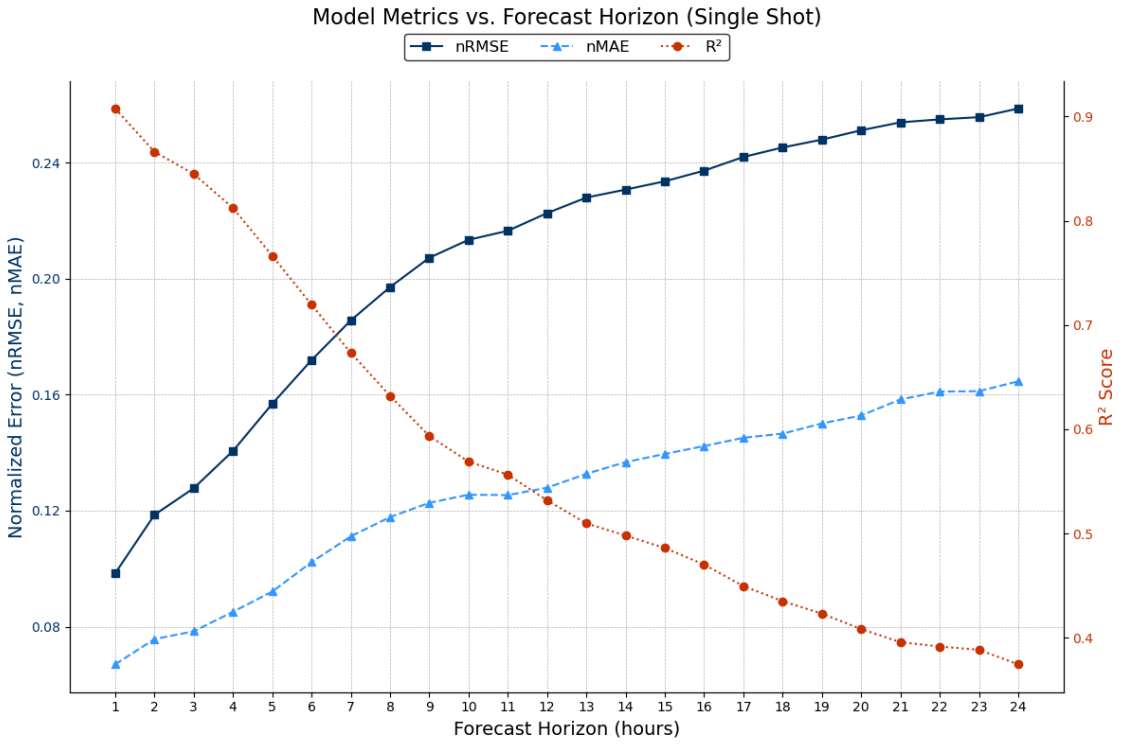

Figure 3.13. Error Metrics for 24 Hour Forecast Horizon (Single-Shot Inference)

A quantitative comparison immediately reveals a critical difference: the single-shot model produces significantly higher errors across all forecast horizons. As seen in Figure 3.13, the nMAE for the single-shot model starts from around ~0.07 and goes up to ~0.16. In stark contrast, the autoregressive model's nMAE is substantially lower, starting at ~0.05 and peaking at ~0.10.

Interestingly, the autoregressive model's error does not increase monotonically with the forecast horizon, a behavior that is counter-intuitive for a typical iterative process (we expect the errors to accumulate). While an initial increase in error is observed, it tends to stabilize and, in some cases, even decline at longer horizons (like as can be seen in the Figure 3.13 above). This stability is not merely a byproduct of using ‘ground-truth’ ERA5 inputs, but rather, it points to a sophisticated, self-regulating mechanism learned by the LSTM model.

A detailed analysis of the model's predictions reveals a sophisticated, self-regulating mechanism, which we term 'Conditional Importance with Inertia', that

explains this counter-intuitive error stabilization. This is explained further, under section 3.2.4. Qualitative Analysis of Prediction Case Studies.

Coming back to forecasting method comparison, the performance gap is also evident in a qualitative, head-to-head comparison of their forecasts for the same time periods, as shown in Figure 3.14 and Figure 3.15 below.

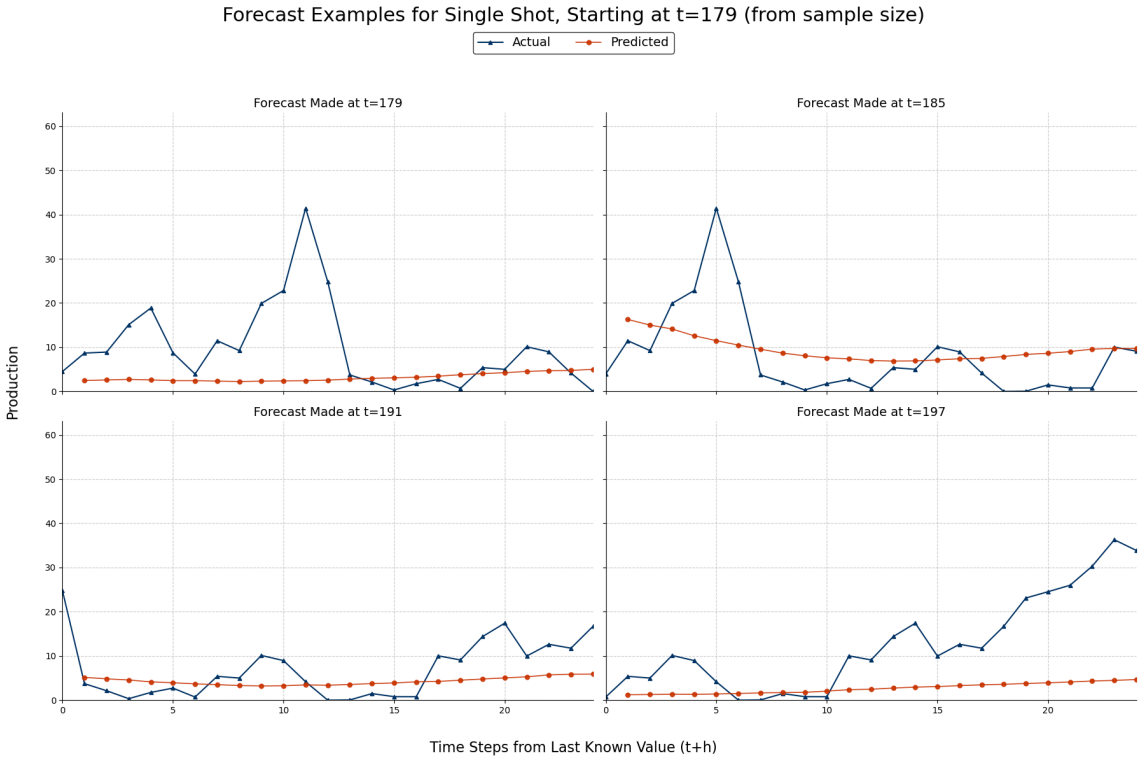

Figure 3.14. Examples of Singled-Out Inferences (for Autoregressive)

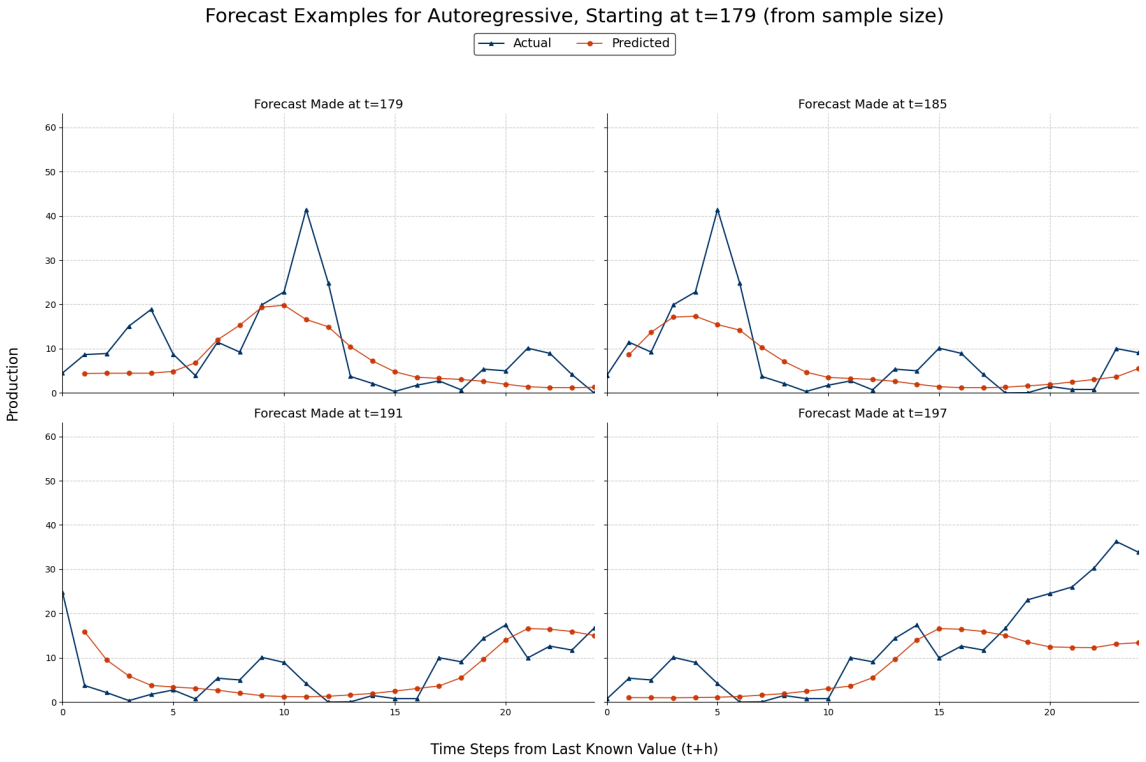

Figure 3.15. Examples of Singled-Out Inferences (for Autoregressive)

This illustrates that the autoregressive forecast tracks the actual production with much greater fidelity. The single-shot model's prediction, despite extensive training, exhibits a noticeable bias and fails to capture the essential dynamics of the wind generation profile.

In conclusion, the autoregressive model is selected as the definitive strategy for this thesis based on two, connected decisive factors:

- **Predictive Accuracy:** It achieves demonstrably lower prediction errors compared to the single-shot model for the same model complexity.
- **Practical Usability and Training Efficiency:** It is vastly more computationally efficient and flexible, requiring only a single model to be trained and can be used to predict variable horizons.

3.2.2. Hyperparameter Optimization and Model Selection

To develop a robust and accurate forecasting model, a systematic hyperparameter optimization process was undertaken. The goal of this process was to evaluate the impact of key architectural and training parameters on the model's predictive performance and to select a final, high-performing configuration for the main analysis.

The primary metrics used for this evaluation were the overall R^2 score, and Normalized Mean Absolute Error (nMAE) calculated **across the entire forecast horizon**. The Normalized Root Mean Square Error (nRMSE) is omitted in the figures shown below to decrease the clutter made by the already quite excessive number of figures. The non normalized results were also omitted as the metrics shown are grouped, and normalization helps with having some roughly comparable results. These aggregated metrics provide a general assessment of a model's predictive capability, rather than its performance at a single forecast step.

A two-tiered approach was adopted to balance computational expense with thoroughness:

- An extensive hyperparameter search was conducted on a single, representative wind farm for an 8-hour forecast horizon. This allowed for a detailed analysis of how each individual parameter influences model performance.
- Based on the insights from the exhaustive search, a refined set of promising hyperparameters and an adaptive training strategy were selected. This configuration was then tested on a diverse set of wind farms with a 24-hour forecast horizon to ensure the model's architecture and training strategy were robust and could generalize effectively.

3.2.2.1. Analysis on the Representative Wind Farm

3.2.2.1.1. Time Steps (Lookback Period)

The length of the input sequence determines the historical window the model uses for its predictions. Figure below shows that performance improves as the number of time steps increases from 1 to 4, and at 8 it starts to worsen slightly. This suggests that approximately 4 hours of preceding data contains the most relevant information for the next 8-hour forecast. It should be noted that a time step of 8 could still be useful for longer horizons.

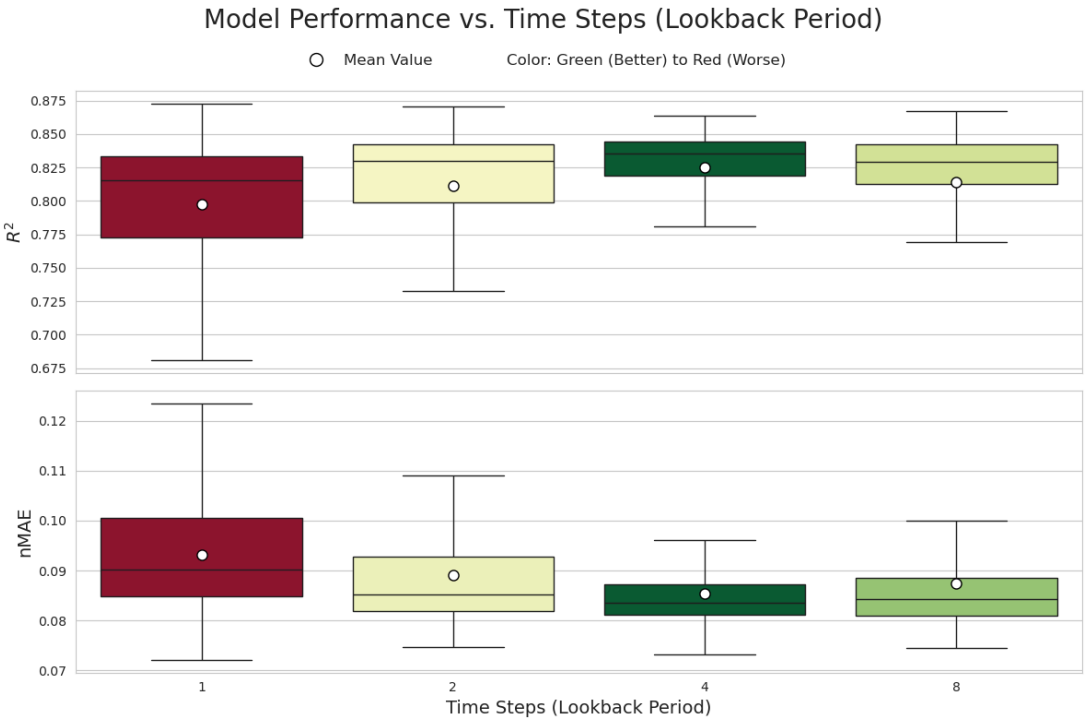

Figure 3.16. Error Metrics for Various Time Steps (single WPP)

3.2.2.1.2. Network Layer Cells

The structure of the LSTM network is a critical determinant of its capacity to learn complex temporal patterns. Figure below illustrates the performance across twelve different layer configurations. The results clearly indicate that deeper networks with 2 or 3 layers consistently outperformed shallower single-layer models. The 128-64-32

configuration achieved a high median R^2 score with less variance than larger models, offering the best trade-off between performance and computational cost.

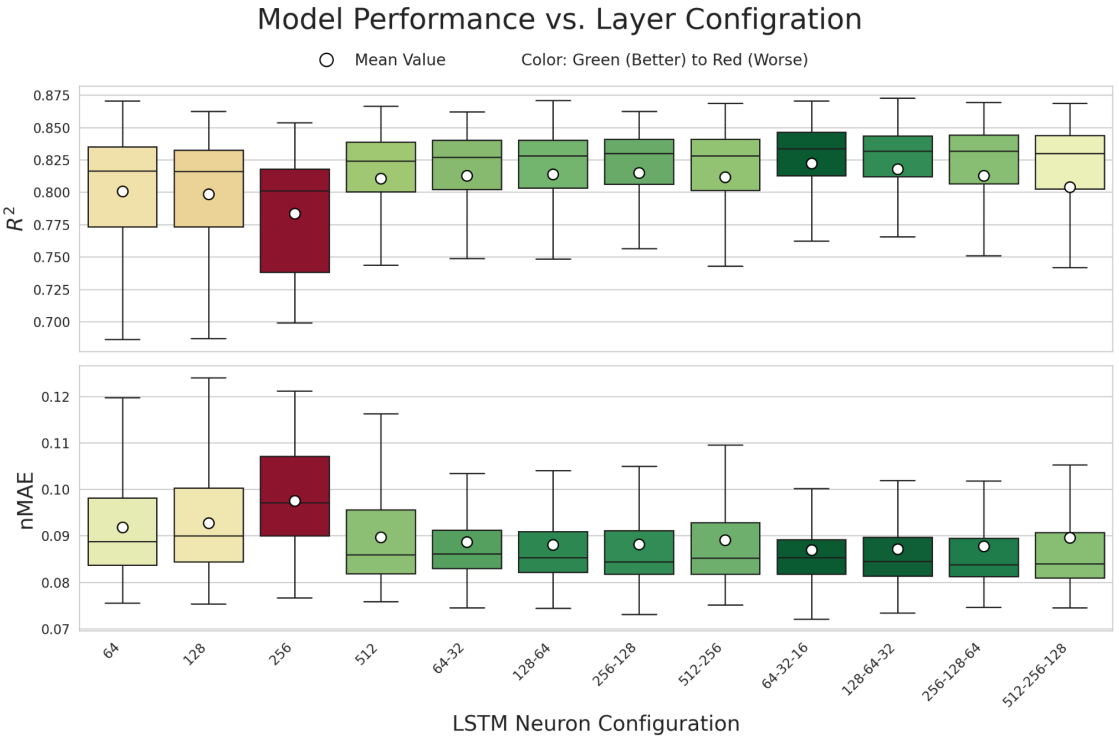

Figure 3.17. Error Metrics for Various Cell Layer Configurations (single WPP)

3.2.2.1.3. Learning Rate

Model convergence is highly sensitive to the learning rate and batch size. Figure below shows that a rate of 0.001 and 0.01 achieved the best results, while a smaller rate failed to converge within the set epochs.

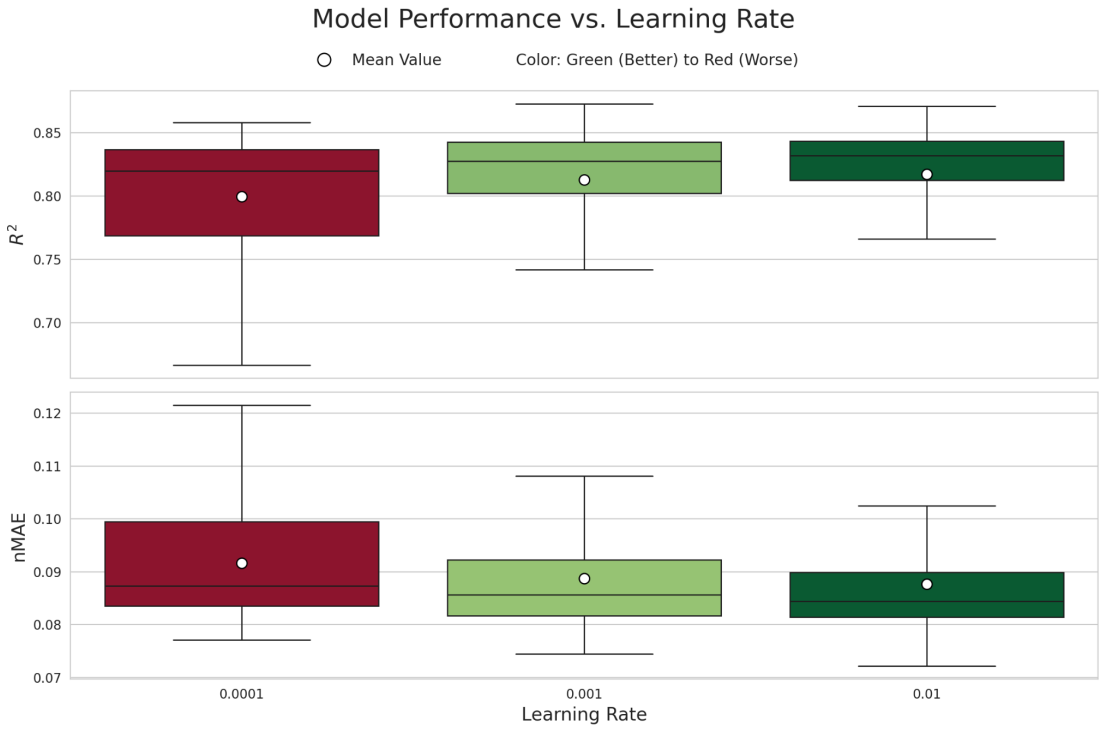

Figure 3.18. Error Metrics for Various Learning Rates (single WPP)

3.2.2.1.4. Batch Size

For batch size, Figure below reveals that the values 32 and 64 produced the most stable and high-performing results. The higher batch sizes often resulted in worse performing models which took the average error rates very high.

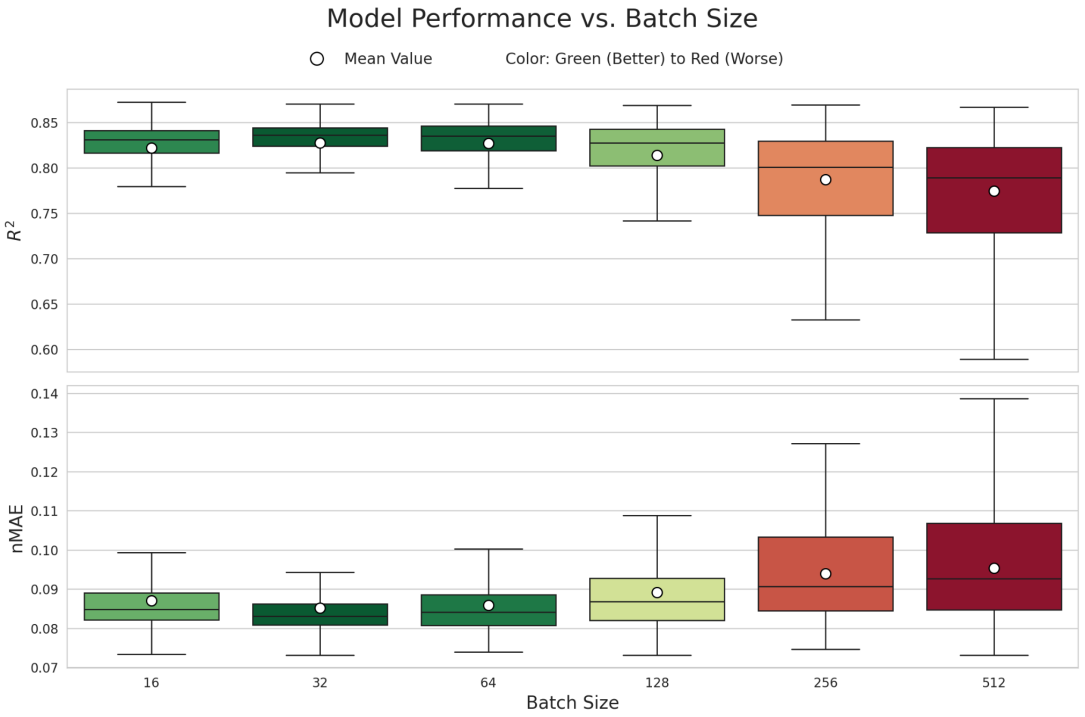

Figure 3.19. Error Metrics for Various Batch Sizes (single WPP)

3.2.2.1.5. Number of Epochs

Based on the results from the Figure below, 10 epochs on average was not a sufficient number. The higher epochs performed consistently better.

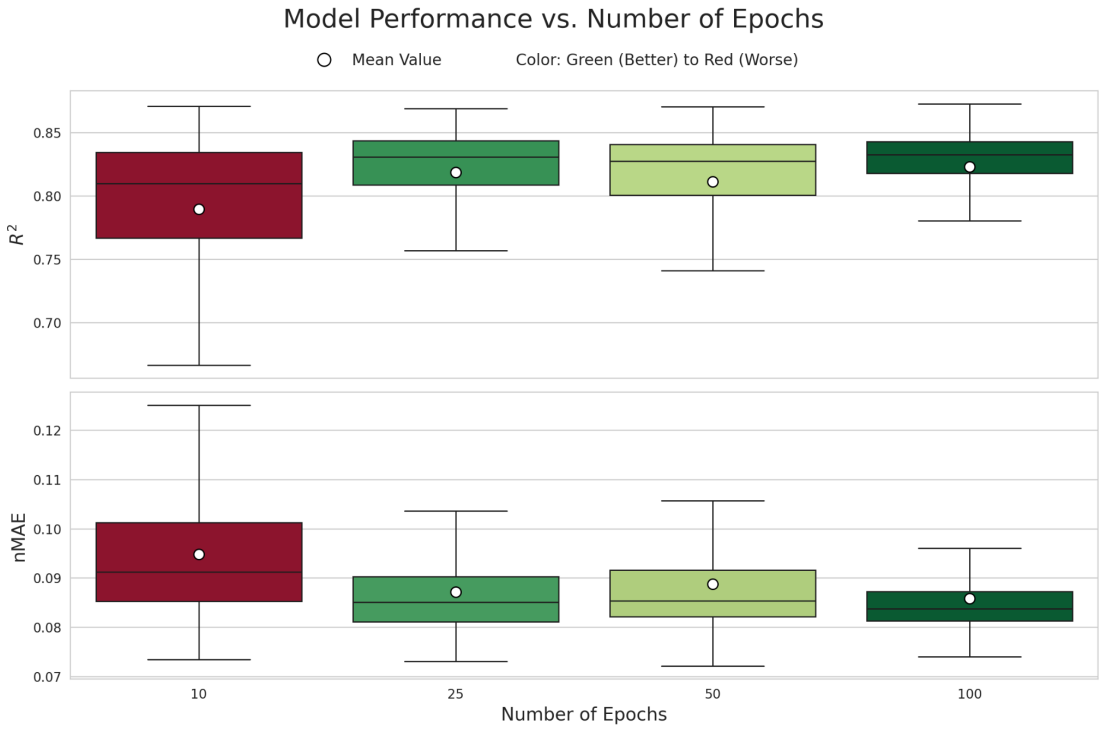

Figure 3.20. Error Metrics for Various Epochs (single WPP)

3.2.2.1.6. Dropout Rate

The effect of the dropout rate was not much. While the figure below shows 0.2 and 0 drop rates are best for 2 different error rates (for nRMSE 0.2 is also the best), their difference is mostly a rounding error. It is known that the dropout rate helps the model with generalization so that the model does not over train on the training data, so it is reasonable to not see the effect directly on the error metrics.

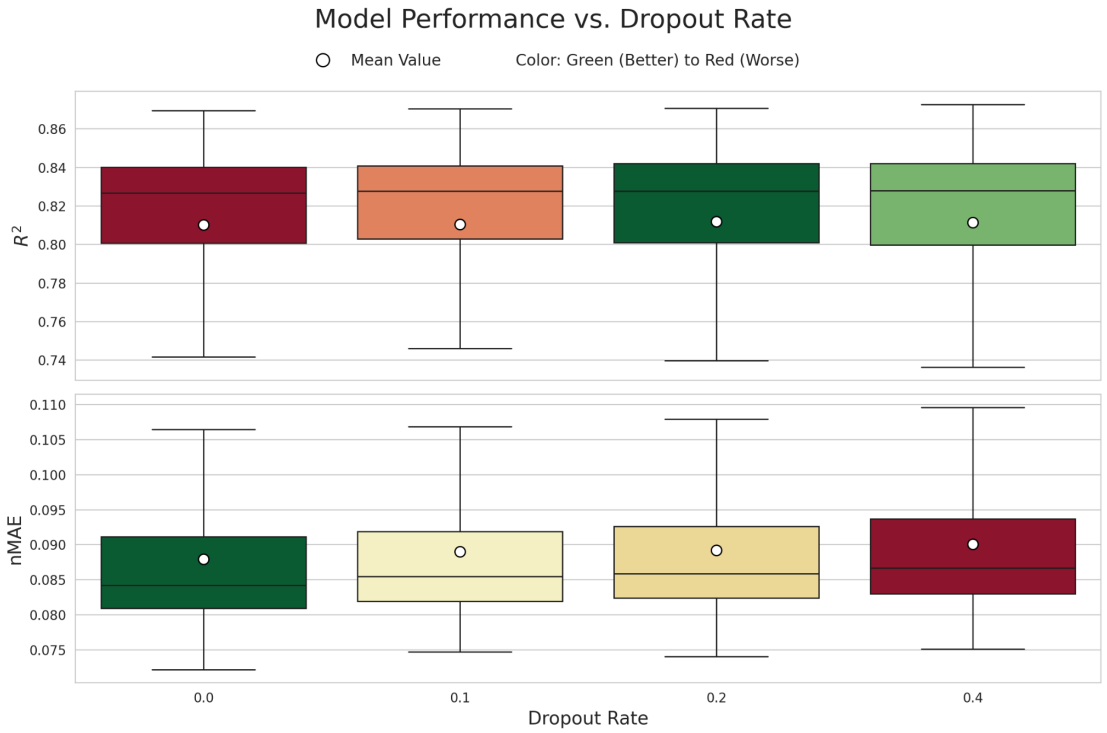

Figure 3.21. Error Metrics for Various Dropout Rates (single WPP)

3.2.2.2. Adaptive Training Strategy

After analyzing the insights from the hyperparameter analysis done on the representative farm, we decided to implement a more robust and adaptive training strategy to ensure optimal convergence. This strategy incorporated two key callback functions during the training of the models: Early stopping of the model and reducing the learning rate automatically based on the loss function.

This callback reduced the learning rate by a factor of 0.2 if the validation loss stagnates, allowing the model to take smaller, finer steps as it approaches a solution. A maximum of 100 epochs was chosen as a sufficiently large window for these adaptive mechanisms to function effectively. The patience is set to 40 epochs and the best model weights are set to restore if the loss does not improve after the patience runs out. Finally, a dropout rate of 0.2 was also chosen to be an effective regularization technique. We used this value because it was already chosen from the literature in the Methodology section of this study.

3.2.2.3. Multi-Farm Validation and Globally Optimal Configuration

While the initial sensitivity analysis on a single wind farm provided valuable insights, it is crucial to ensure that these findings are not specific to a single site's unique conditions. Therefore, a targeted set of the most promising hyperparameters was selected and validated across a diverse portfolio of wind farms for a 24-hour forecast horizon.

3.2.2.3.1. Time Steps

The results seen on the figure below show the time steps of 2 and 4 generally performed better. On the other hand, 8 time steps performed worse again, this time on a longer horizon, which suggests a higher time step parameter might not be useful even for longer horizons.

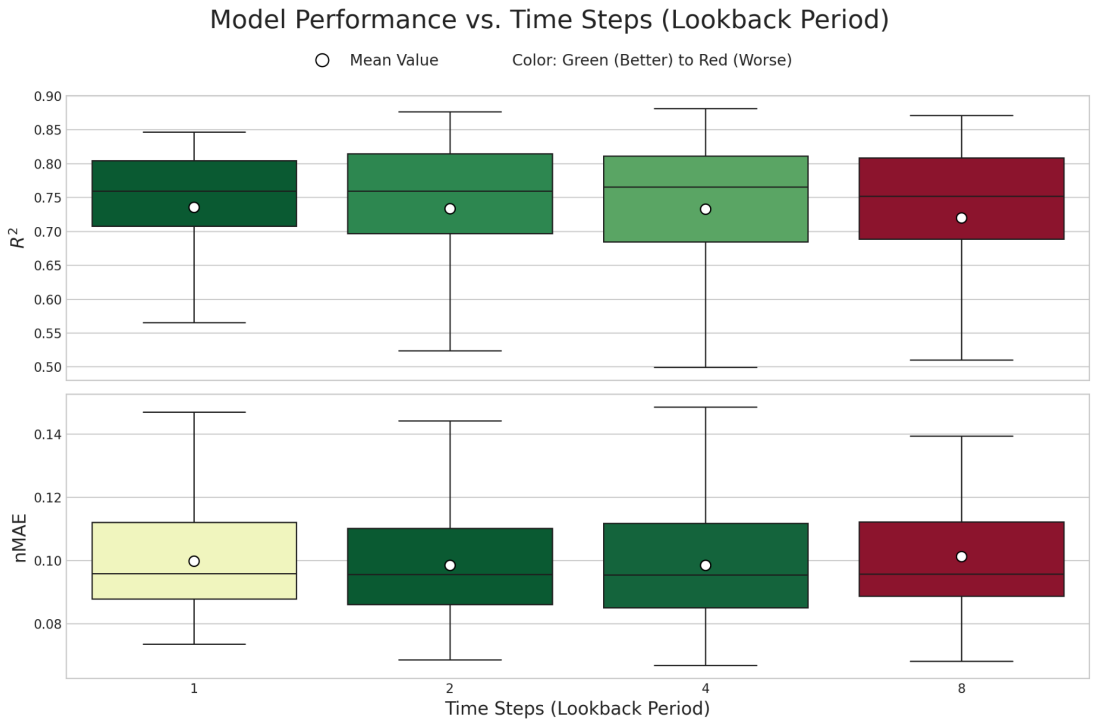

Figure 3.22. Error Metrics for Various Time Steps (All WPPs)

3.2.2.3.2. Network Layer Cells

From the initial step of finding the best hyperparameters, the 3 layered configuration is selected. Looking at the figure below, while the most complex layer with the highest number of neurons performed slightly worse, the others generally perform at the same range.

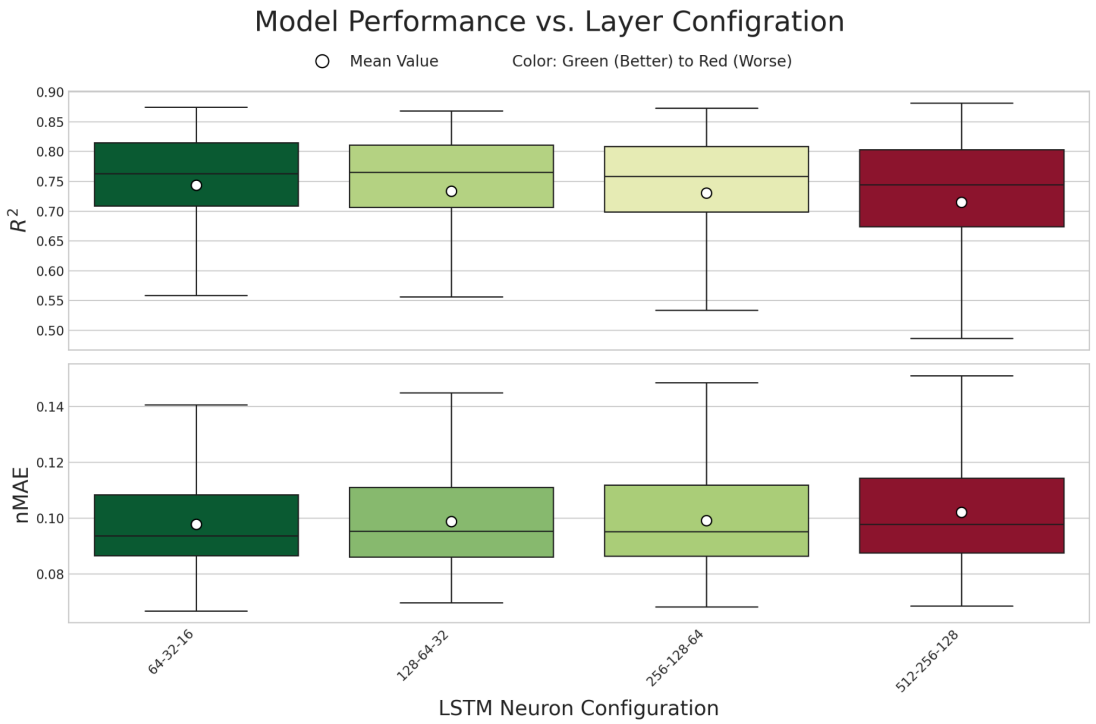

Figure 3.23. Error Metrics for Various Cell Layer Configurations (All WPPs)

3.2.2.3.3. Learning Rate

After implementing the adaptive training strategy mentioned above, it would have made sense to start with a faster learning rate, but we wanted to put this on a test and did these tests with both 0.01 and 0.001 learning rates. Contrary to our initial thoughts, the slower learning rate was still performing considerably better, suggesting that a faster learning, even with adaptive learning strategies such as early stopping and reducing the learning rate, tends to overfit to the training data.

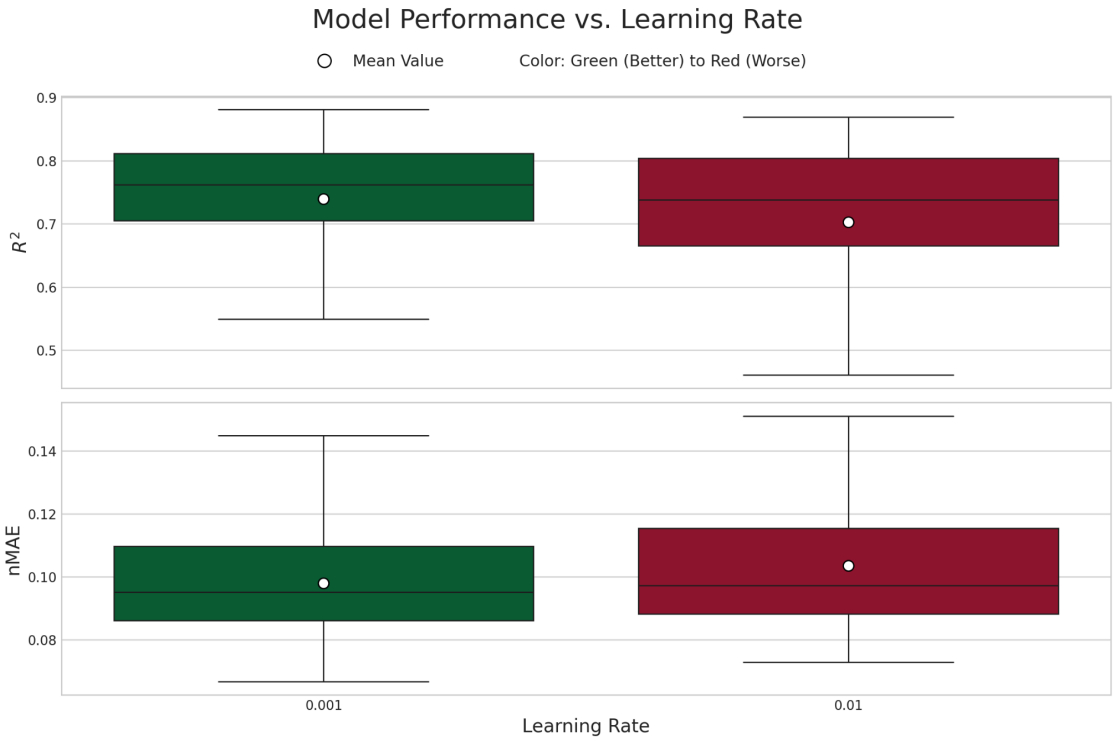

Figure 3.24. Error Metrics for Various Learning Rates (All WPPs)

3.2.2.3.4. Batch Size

Finally, for the batch size we saw the 64 and 128 were giving better results on average. A batch size of 32 was having worse resulting models compared to the others.

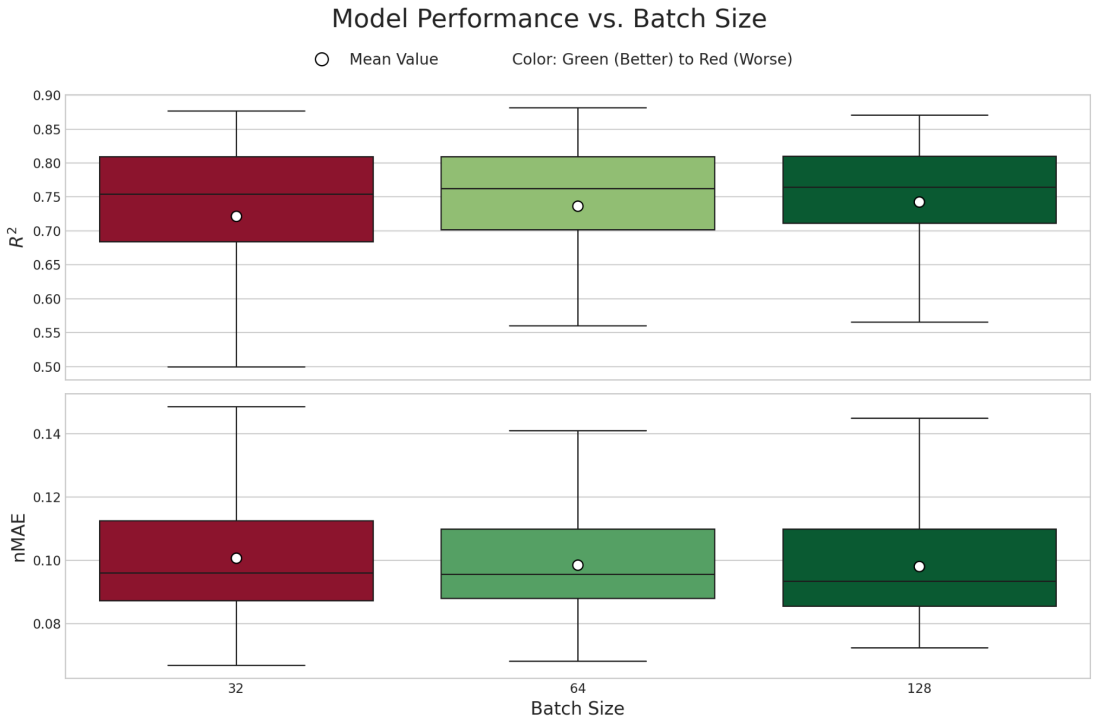

Figure 3.25. Error Metrics for Various Batch Sizes (All WPPs)

At the end of this hyperparameter analysis, to select a single, overall good performing model parameters that we can apply to all wind farms, we have filtered the top 10 performing models for each wind farm, and selected the combination that is used the most across these 10 top performing models list. The model with the following parameters were found to be the best performing among the most farms:

Time Steps	4
Learning Rate	0.001
Epochs	100
Layer Configuration	256-128-64
Batch Size	128
Dropout Rate	0.2

Table 3.2. Selected Hyperparameters Based on The Experimental Hyperparameter Optimization

3.2.3. Quantitative Performance of the Optimized Model

Following the hyperparameter optimization process, the best-performing autoregressive model configuration was selected for a performance evaluation. This section presents the quantitative results of this model, to establish its overall effectiveness and analyze its behavior over the 24-hour forecast horizon.

3.2.3.1. Overall Predictive Accuracy and Error Dynamics

The primary measure of the model's performance is shown in Figure 3.4, which plots the average nMAE, nRMSE, and R^2 score as a function of the forecast horizon. These metrics represent the average performance across the entire test set for all wind farms.

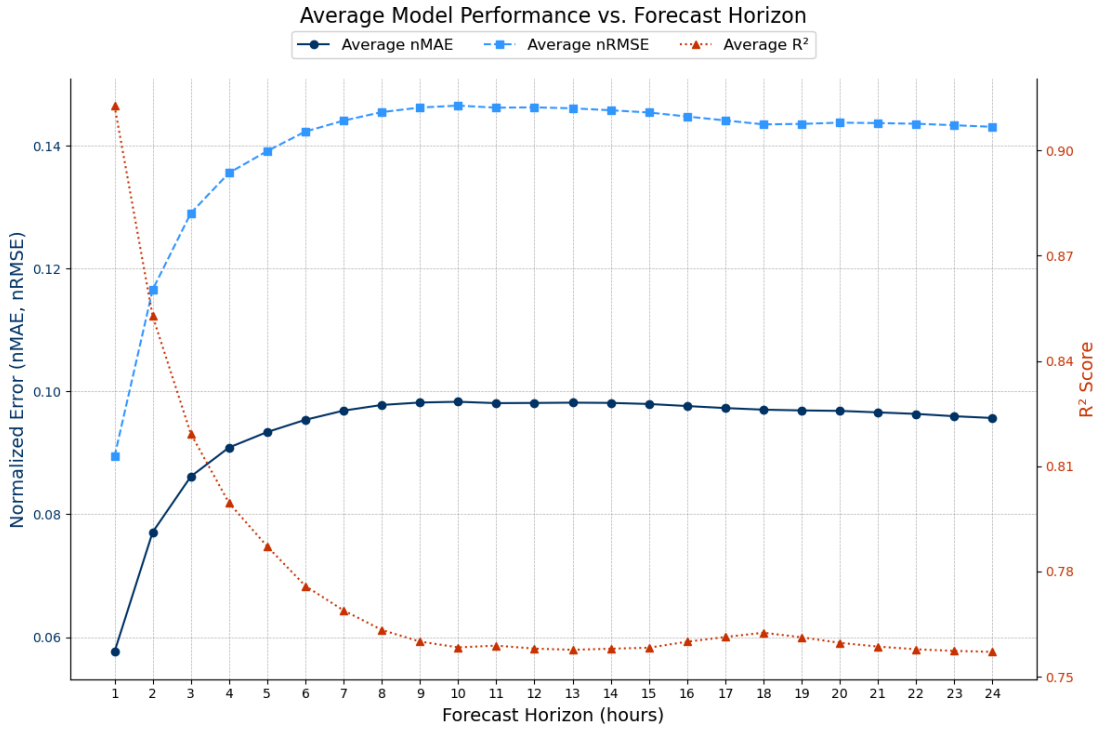

Figure 3.26. Averaged Error Metrics for All WPPs in Study

The model demonstrates very high accuracy for short-term forecasts. At the 1-hour horizon, the average nMAE is approximately 6%, and the R^2 score is 0.91.

As the forecast horizon extends, the error metrics increase as the model predicts further into the future. The nMAE rises to a peak of almost 10% around the 9-hour mark. Mirroring this, the R^2 score declines to its lowest point of 0.76 at the same horizon.

While the aggregated results in Figure 3.26 confirm the model's overall robustness, it is crucial to recognize that these averages mask performance variations among individual wind farms. The model's accuracy is not uniform across all sites due to unique factors such as local topography, turbine characteristics, and data quality. To illustrate this variance, Figure 3.27 below presents box plots of the nMAE, nRMSE and R^2 score distribution across all wind farms.

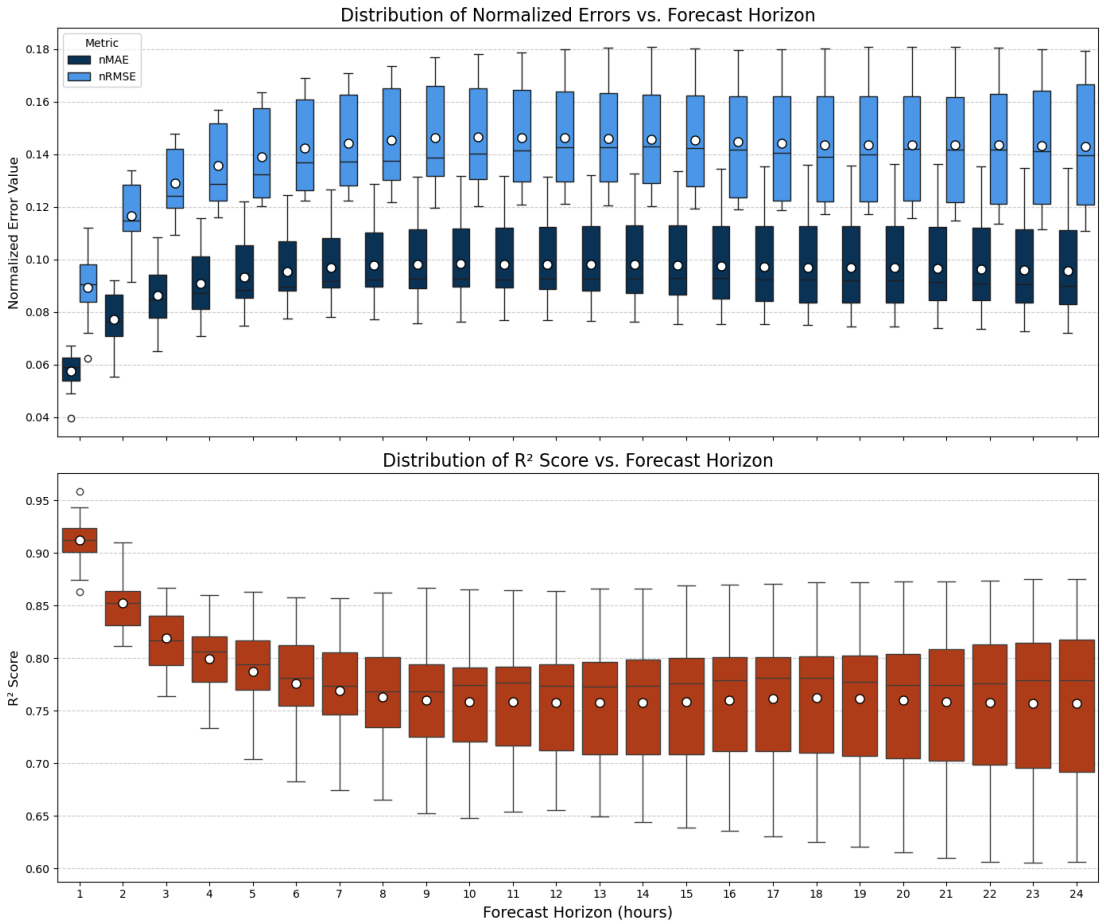

Figure 3.27. Box Plot of the Average Error Metrics for All WPPs is Study

The plot clearly shows that the spread of errors across farms widens as the forecast horizon extends. The height of the boxes and the reach of the whiskers grow larger, indicating that while the model is consistently accurate for most farms at short horizons, its performance becomes more variable with longer horizon forecasts. One simple possible explanation for this observed disparity can be attributed to the diverse characteristics of the wind farms in Türkiye. Factors such as the complexity of the local wind regime, the specific turbine technology in use, and the quality of the available historical data all contribute to the site-specific accuracy of the forecasts. The other possible reason is the ability of the LSTM model in capturing the relation of wind production with the weather parameters.

3.2.4. Qualitative Analysis of Prediction Case Studies

While quantitative metrics provide a high-level summary of accuracy, a qualitative analysis is essential for a deeper understanding of the model's behavior. This section visually inspects the model's predictions in two distinct ways. First, we assess the model's performance at fixed, short-term horizons across the entire dataset. Second, we analyze complete 24-hour forecast case studies to observe the iterative prediction process during specific events.

3.2.4.1. Performance at Fixed Forecast Horizons

To evaluate the model's baseline accuracy at different lookaheads, Figure 3.28 presents the predictions for the first four forecast horizons, each stitched together over the entire test set. This is not an iterative forecast, but rather a collection of all t+1 predictions, all t+2 predictions, and so on. The figures onwards show the rest of the fixed forecast horizons until the 24 hour mark.

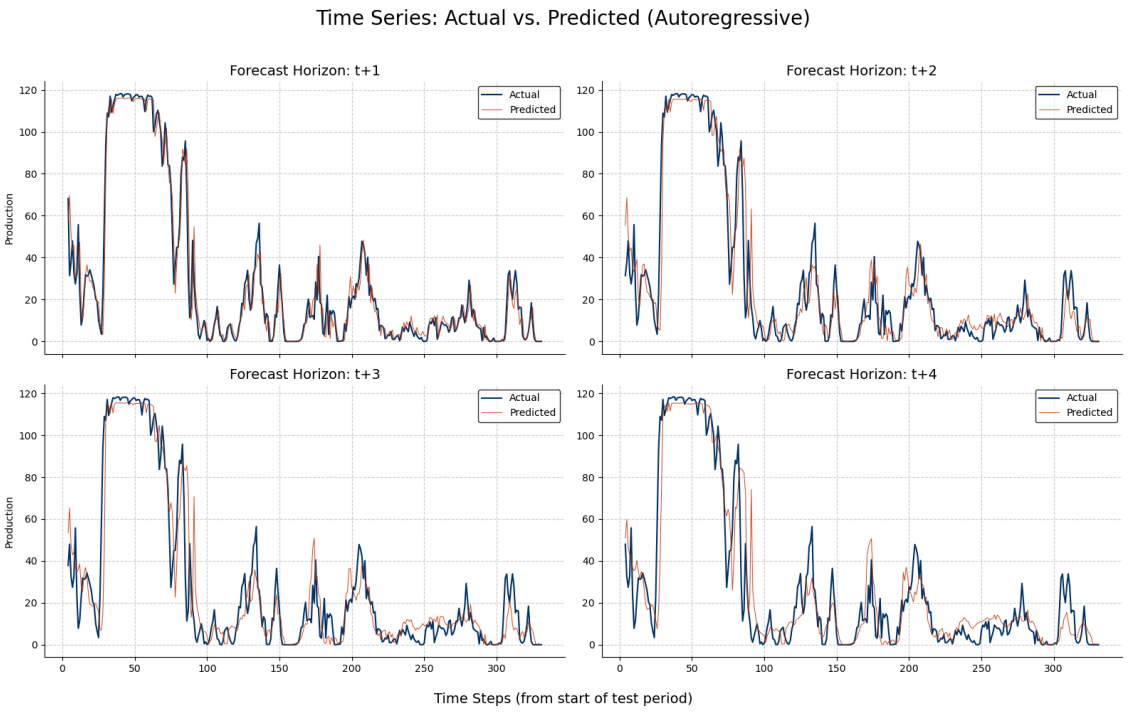

Figure 3.28. Time Series View of the Forecast Horizons (t+1 to t+4)

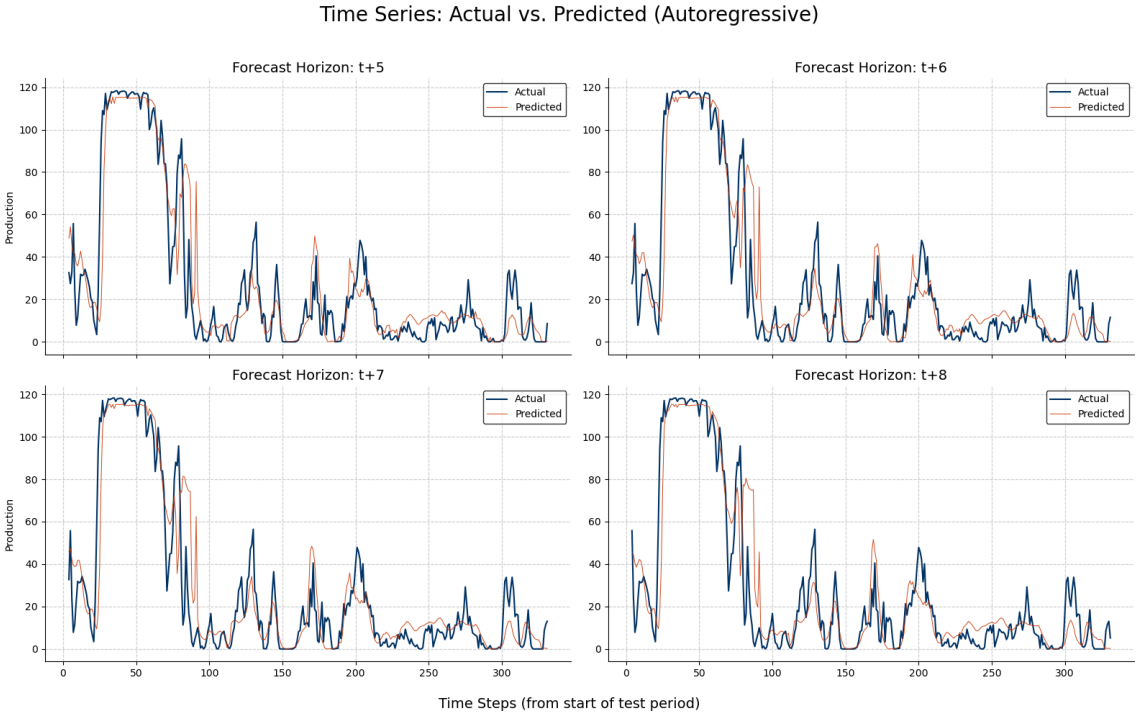

Figure 3.29. Time Series View of the Forecast Horizons (t+5 to t+8)

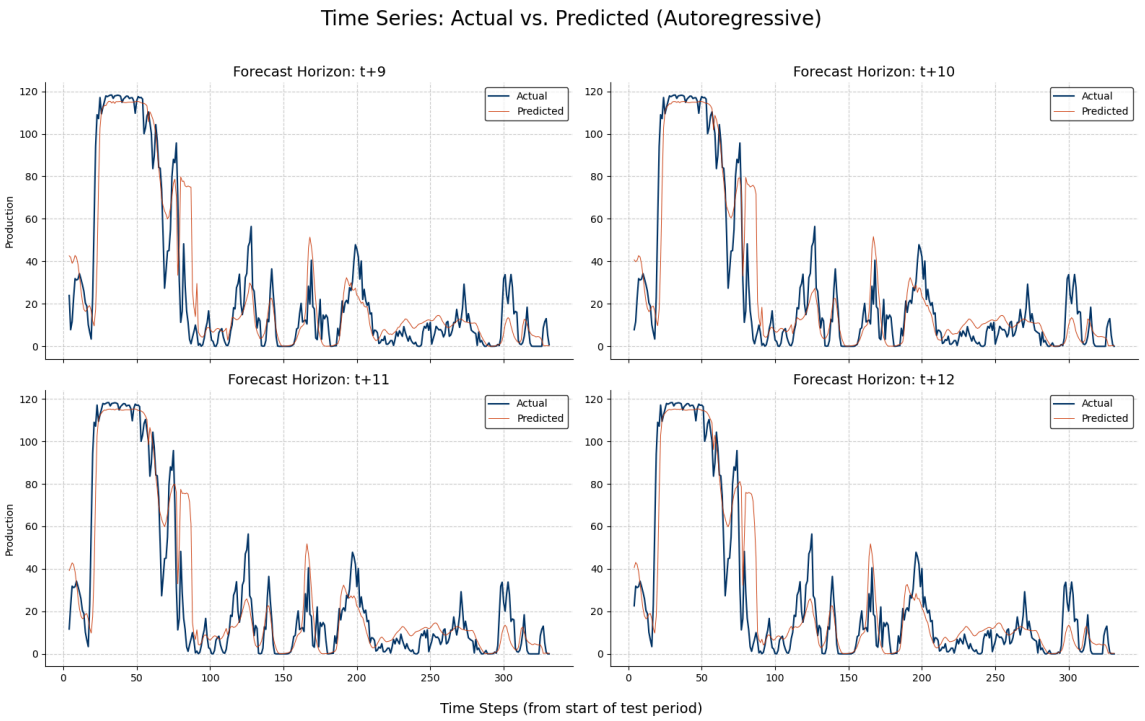

Figure 3.30. Time Series View of the Forecast Horizons (t+9 to t+12)

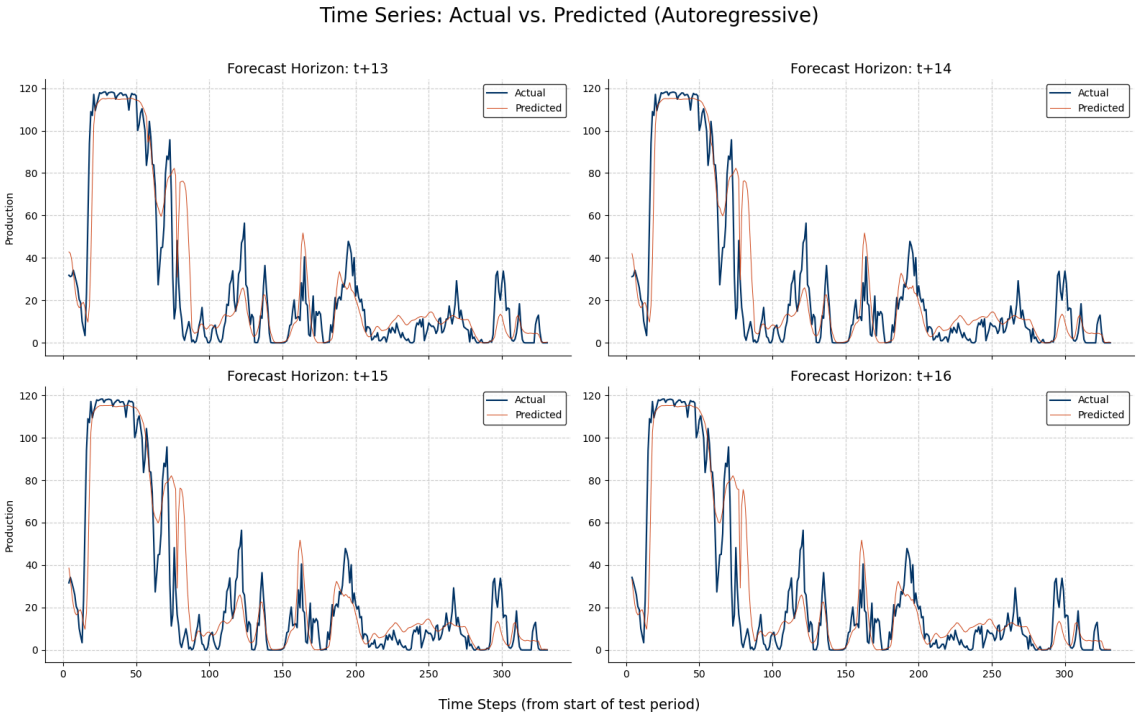

Figure 3.31. Time Series View of the Forecast Horizons (t+13 to t+16)

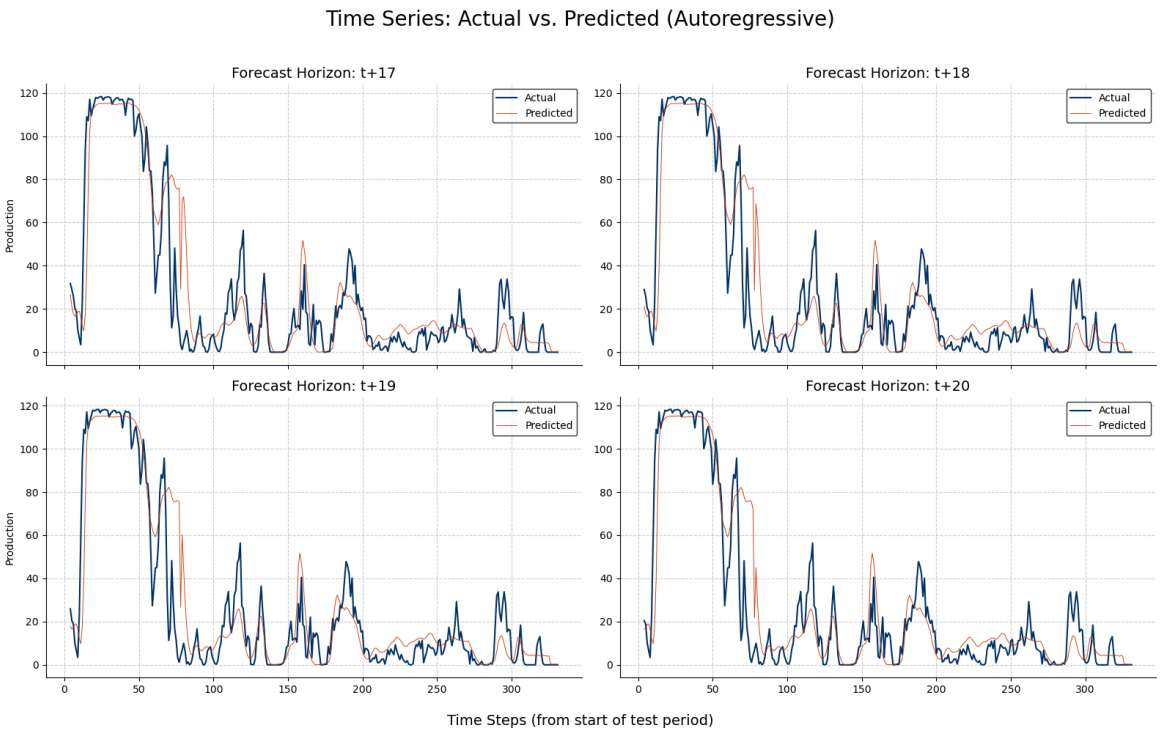

Figure 3.32. Time Series View of the Forecast Horizons (t+17 to t+20)

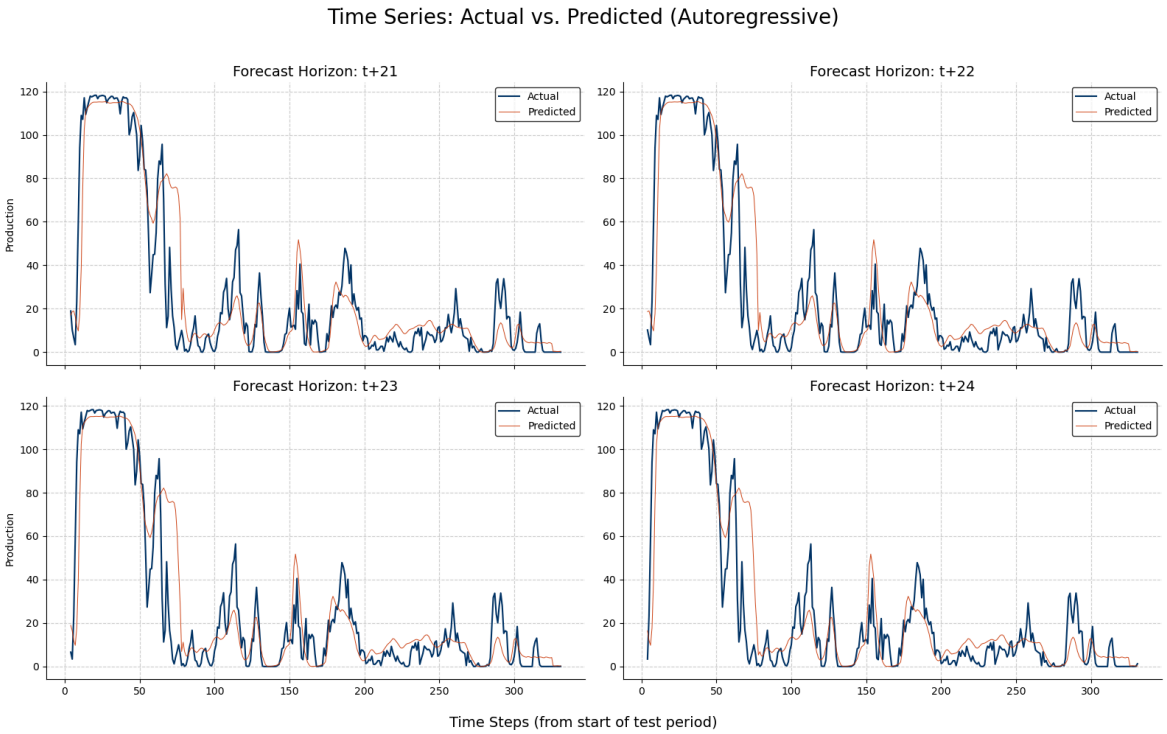

Figure 3.33. Time Series View of the Forecast Horizons (t+21 to t+24)

These visualizations confirm the findings from the quantitative analysis. At the t+1 horizon (top left), the forecast is remarkably accurate. The predicted values align very closely with the actual generation, capturing both the timing and magnitude of fluctuations. This corresponds to the low nMAE and high R^2 value observed for the one-hour-ahead forecast. This is possible thanks to using the last production in the input. From t+2 and onwards, a gradual and subtle degradation in performance is observable, and at a certain point the degradation comes to a stop. While the model continues to track the primary trends, the predictions become less precise, with an increase in divergence from the actual values, particularly at the sharpest peaks.

3.2.4.2. Analysis of an Individual Inference

To understand how the model performs in a real-world, iterative scenario, we now analyze complete 24-hour forecasts for specific events. Unlike the previous figure, each plot in this section represents a single, continuous forecast generated step-by-step over 24 hours from a given starting point.

3.2.4.2.1. Case Study A: Forecasting a Major Ramp-Up Event

Figure 3.34 showcases the model's performance when tasked with predicting a slight dip in wind generation.

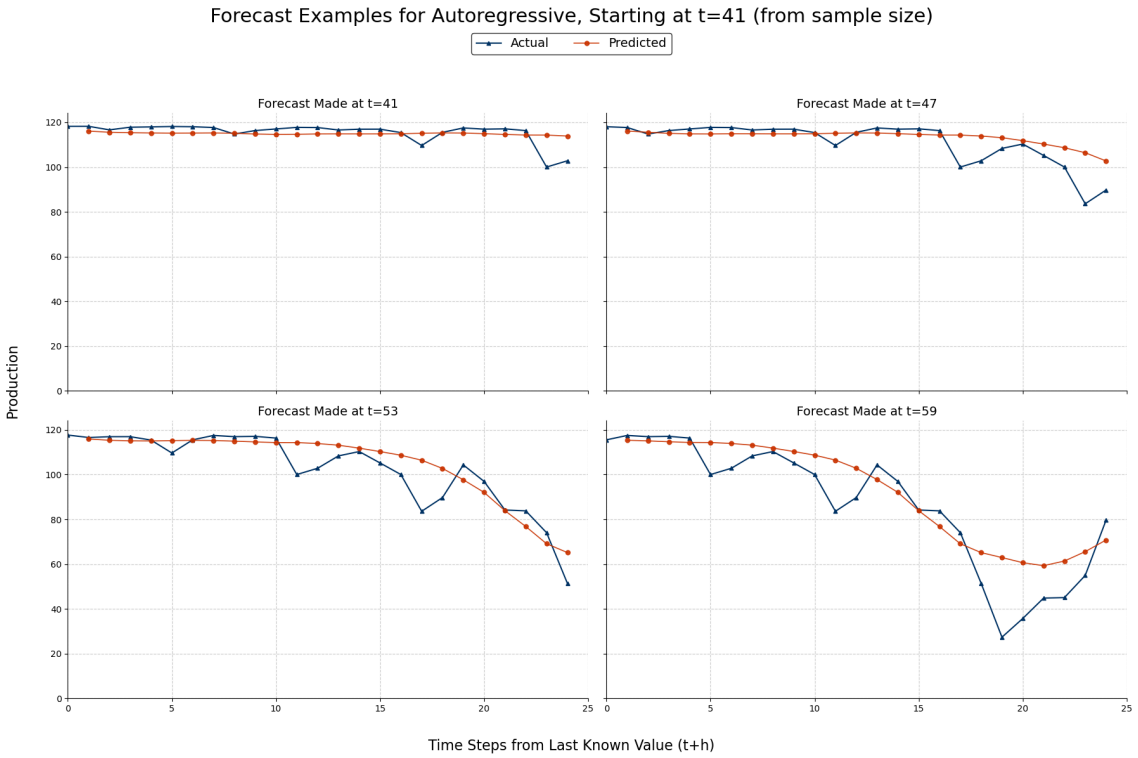

Figure 3.34. Examples of Singled-Out Inferences (Autoregressive)

In this scenario, the model demonstrates a key strength, its ability to capture the clear trends. The forecast correctly follows the downward trend and predicts its trajectory over the 24-hour period with impressive accuracy.

3.2.4.2.2. Case Study B: Forecasting Through Volatility

Figure 3.35 presents a more challenging case where the generation profile is less stable and has a sudden ramp up.

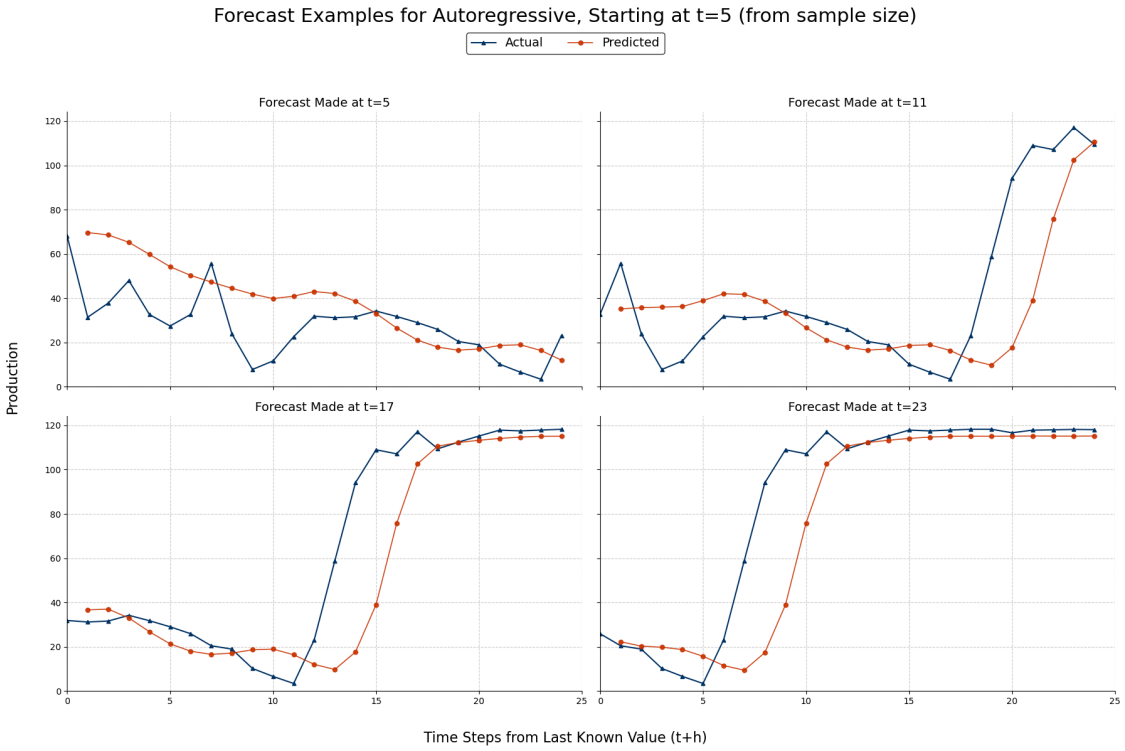

Figure 3.35. Examples of Singled-Out Inferences (Autoregressive)

This case study highlights the nature of the autoregressive process. The initial predictions track the actual generation closely. However the moment the production ramps up, the autoregressive model struggles with breaking away from the lagged production input and has a clear ‘latency’ when reacting to change. This is tied closely to the behavior described earlier as the Conditional Importance with Inertia.

First, the model exhibits a form of "inertia," where it gives significant weight to its previous prediction by default. This is evidenced by a noticeable delay in the model's response to sharp, rapid changes in actual production and a tendency to smooth out high-frequency fluctuations. The model has learned that wind production generally has a smooth profile and is therefore hesitant to predict drastic, moment-to-moment changes.

However, this inertia is not absolute. The model has learned to apply "conditional importance" to its inputs. When the autoregressive input, the model's own prior prediction, drifts into a state that is physically implausible given the corresponding weather data, the model has learned to distrust and discount this anomalous input.

This self-regulating behavior explains the observed error curves where the error starts to decrease with the longer forecasting horizons. The error increases in the initial hours as minor inaccuracies propagate through the model's "inertial" phase. Subsequently, the error stabilizes because the conditional, self-correcting mechanism acts as a powerful feedback loop, preventing the cumulative error from spiraling out of control by constantly pulling the forecast back to a physically plausible trajectory dictated by the exogenous weather data.

It is crucial, however, to contextualize these findings within the limitations of the input data used. This behavior and stable long-term error profiles observed are predicated on the use of ERA5 reanalysis data. As a historical reanalysis product, ERA5 serves as a near-perfect, error-free representation of past weather conditions, or a ‘ground-truth’ data and not a true forecast. In a real-world operational setting, this type of forecasting model must rely on weather predictions from services like the ECMWF’s HRES or GFS, which contain their own inherent errors and an uncertainty that typically grows with the forecast horizon. To support this hypothesis we also experimented where we introduced artificial noise into the weather inputs, and found that the error metrics do not come to a plateau and instead continue to rise slowly. Therefore, while this analysis reveals the model's impressive ability to learn and enforce physical plausibility, the stability observed should be viewed as a best-case scenario. The deployment of this model in a live environment would inevitably lead to more pronounced error accumulation, highlighting the critical performance gap between academic studies using reanalysis data and the challenges of real-world, operational forecasting.

CHAPTER 4.
DISCUSSION

The primary finding from the aging analysis is the statistically significant mean degradation rate of -0.90% per year across the Turkish wind farm fleet for the 2020-2023 period. This figure provides a crucial, data-driven benchmark for a region where such studies have been previously lacking. It offers a quantitative basis for investors, operators, and policymakers to factor age-related performance decline into economic models and long-term energy planning. Interestingly, the analysis did not reveal a significant correlation between this degradation trend and key wind power plant characteristics such as turbine brand, installation year, or site elevation. This data suggests that, within the analyzed timeframe, these factors may be less influential than broader operational and maintenance strategies or atmospheric conditions that are common across the fleet. It is possible that the four-year analysis window is not sufficient to capture the more pronounced age-related effects that might emerge over a decade or more.

On the forecasting front, the decision to proceed with an autoregressive LSTM model was vindicated. The comparative analysis demonstrated that this approach is not only more computationally efficient but also yields a better predictive accuracy compared to a single-shot model. The single-shot method's requirement to train a separate, complex model for each forecast horizon makes it impractical for real-world, scalable deployment, whereas the autoregressive model’s flexibility is a significant operational advantage.

A key insight from the forecasting results was the model's self-regulating behavior, described as "Conditional Importance with Inertia". This learned mechanism, where the model discounts its own anomalous predictions when they conflict with plausible weather inputs, explains the stability of the forecast error at longer horizons. However, it is critical to contextualize this finding within the study's primary limitation: its reliance on historical ERA5 reanalysis data as a proxy for perfect weather forecasts. This creates an extremely idealized scenario. In a real-world operational setting, the

model would ingest weather forecasts that have their own inherent and increasing uncertainty, which would inevitably lead to more pronounced error accumulation.

This highlights the primary challenge of autoregressive forecasting: the propagation of errors at each iterative step. While our model showed resilience, its performance in a live environment would be lower. Future work should focus on bridging this gap. One promising technique to mitigate error accumulation is scheduled sampling (Bengio et al., 2015). This training strategy involves systematically replacing some of the ground-truth inputs with the model's own predictions during the training phase. By exposing the model to its own errors during training, it can learn to become more robust and less sensitive to the imperfections of its own iterative forecasts, which could significantly improve its operational performance. Furthermore, future research should validate the model's performance using actual weather prediction data from sources to quantify the real-world performance gap and refine the model accordingly.

CHAPTER 5.

CONCLUSION

This thesis set out to perform a comprehensive time series analysis of wind energy production in Türkiye, with two primary objectives: to quantify the effects of aging on wind turbine performance and to develop an accurate and practical short-term power forecasting model. Both objectives were successfully met through the creation of a detailed national database and the application of rigorous analytical and machine learning techniques.

The major findings of this research can be summarized as follows:

- **Performance Degradation:** The Turkish wind farm fleet exhibits a statistically significant average performance degradation, measured by capacity factor, of -0.90% per year.
- **Influencing Factors on Aging:** For the 2020-2023 period, no significant statistical link was found between the rate of degradation and the wind farm's turbine manufacturer, installation year, or physical characteristics such as elevation and capacity density.
- **Forecasting Model Efficacy:** An autoregressive LSTM model was developed and proven to be a highly effective tool for short-term forecasting. It demonstrated high predictive accuracy and practical flexibility compared to a single-shot alternative.
- **Model Behavior and Limitations:** The autoregressive model exhibits a self-correction mechanism that stabilizes long-horizon errors when using "perfect" reanalysis data. However, this underscores a critical performance gap between academic studies and real-world applications, which must contend with the uncertainty of actual weather forecasts.

This study contributes to the literature as a quantitative benchmark for wind turbine aging in Türkiye, providing invaluable data for the nation's energy sector. Furthermore, it presents a robust, validated, and computationally efficient LSTM

architecture for wind power forecasting. The insights into the model's behavior and the clear path outlined for future improvements, particularly the implementation of scheduled sampling and testing with real-world forecast data, lay the groundwork for the development of even more reliable and accurate forecasting tools. As Türkiye continues to expand its portfolio with more renewable energy sources, such data-driven analyses will be more essential for ensuring the efficiency, stability, and economic viability of its wind power infrastructure.

REFERENCES

Astolfi, D., Byrne, R., & Castellani, F. (2020, 10 27). Analysis of wind turbine aging through Operation Curves. *Energies*, *13*(21), 5623. https://doi.org/10.3390/en13215623

Astolfi, D., Byrne, R., & Castellani, F. (2021, 02 09). Estimation of the performance aging of the Vestas V52 wind turbine through comparative test case analysis. *Energies*, *14*(4), 915. https://doi.org/10.3390/en14040915

Astolfi, D., Castellani, F., Lombardi, A., & Terzi, L. (2021, 12). Data-driven wind turbine aging models. *Electric Power Systems Research*, *201*. https://doi.org/10.1016/j.epsr.2021.107495

Bengio, S., Vinyals, O., Jaitly, N., & Shazeer, N. (2015, 09 23). Scheduled sampling for sequence prediction with recurrent neural networks. https://doi.org/10.48550/arXiv.1506.03099

Byrne, R., Astolfi, D., Castellani, F., & Hewitt, N. J. (2020, 04 21). A Study of Wind Turbine Performance Decline with Age through Operation Data Analysis. *Energies*, *13*(8), 2086. https://doi.org/10.3390/en13082086

Demirtop, A., & Sevli, O. (2024). Wind speed prediction using LSTM and ARIMA time series analysis models: A case study of Gelibolu. *Turkish Journal of Engineering*, *8*, 524–536. 10.31127/tuje.1431629

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *kdd*, *96*(34), 226-231. https://cdn.aaai.org/KDD/1996/KDD96-037.pdf

Foley, A. M., Leahy, P., & Mckeogh, E. (2010). Wind Power Forecasting & Prediction Methods. *9th International Conference on Environment and Electrical Engineering*. https://doi.org/10.1109/EEEIC.2010.5490016

Foley, A. M., Leahy, P. G., Marvuglia, A., & McKeogh, E. J. (2012, 01). Current methods and advances in forecasting of wind power generation. *Renewable Energy*, *37*(1), 1-8. https://doi.org/10.1016/j.renene.2011.05.033

Geerts, H. (1984). Short range prediction of wind speeds: a system-theoretic approach. In *Proceedings of the European Wind Energy Conference, Hamburg (DE)* (594--599).

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Dinand, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*. https://doi.org/10.1002/qj.3803

International Renewable Energy Agency. (2024, July 11). *Renewable energy highlights*. International Renewable Energy Agency. https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2024/Jul/Renewable_energy_highlights_FINAL_July_2024.pdf

Kisvari, A., Lin, Z., & Liu, X. (2021, 01). Wind power forecasting – a data-driven method along with gated recurrent neural network. *Renewable Energy*, *163*, 1895-1909. https://doi.org/10.1016/j.renene.2020.10.119

Memarzadeh, G., & Keynia, F. (2020, April 21). A new short-term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets. *Energy Conversion and Management*, *213*. https://doi.org/10.1016/j.enconman.2020.112824

Memarzadeh, G., & Keynia, F. (2020, 06). A new short-term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets. *Energy Conversion and Management*, *213*. https://doi.org/10.1016/j.enconman.2020.112824

Oh, J., Park, J., Ok, C., Ha, C., & Jun, H.-B. (2022, 12 10). A study on the wind power forecasting model using transfer learning approach. *Electronics*, *11*(24). https://doi.org/10.3390/electronics11244125

scikit-learn developers. (n.d.). *DBSCAN — scikit-learn 1.6.0 documentation*. Scikit-learn. https://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html

Shahid, F., Zameer, A., & Muneeb, M. (2021). A novel genetic LSTM model for wind power forecast. *Energy*, *223*. https://doi.org/10.1016/j.energy.2021.120069

Shahid, F., Zameer, A., & Muneeb, M. (2021, 05). A novel genetic LSTM model for wind power forecast. *Energy*, *223*. https://doi.org/10.1016/j.energy.2021.120069

Shao, B., Song, D., Bian, G., & Zhao, Y. (2021). Wind Speed Forecast Based on the LSTM Neural Network Optimized by the Firework Algorithm. *Advances in Materials Science and Engineering*, *2021*. https://doi.org/10.1155/2021/4874757

Staffell, I., & Green, R. (2014). How does wind farm performance decline with age? *Renewable Energy*, *66*, 775-786. https://doi.org/10.1016/j.renene.2013.10.041

Taslimi Renani, E., Elias, M. F., & Rahim, N. A. (2016, 06). Using data-driven approach for wind power prediction: A comparative study. *Energy Conversion and Management*, *118*. https://doi.org/10.1016/j.enconman.2016.03.078

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley Publishing Company.

Turkish Ministry of Energy and Natural Resources. (2024, May). *TÜRKİYE ELEKTRİK YATIRIMLARI 2024 YILI MAYIS AYI ÖZET RAPORU*. Enerji İşleri Genel Müdürlüğü Yatırımlar Dairesi Başkanlığı. https://enerji.gov.tr//Media/Dizin/EIGM/tr/Raporlar/PRP/EY%C3%96RA/2024/May%C4%B1s.pdf

Wang, J., Zhou, Q., & Zhang, X. (2018, 11 19). Wind power forecasting based on time series Arma model. *IOP Conference Series: Earth and Environmental Science*, *199*. https://doi.org/10.1088/1755-1315/199/2/022015

Xu, J., Zou, J., Ziegler, A., Wu, J., & Zeng, Z. (2023, 05 12). What drives the change of capacity factor of wind turbine in the United States? *Environmental Research Letters*, *18*(6). https://doi.org/10.1088/1748-9326/acce4f

Yoldaş, İ. S. (2022). Short-term Wind Speed and Power Forecasting: A Comprehensive Case Study for Three Operational Wind Farms. *Master's thesis, Izmir Institute of Technology (Turkey)*.

YÜREK, Ö., BİRANT, D., & YÜREK, İ. (2021, 01 15). Makine öğrenmesi algoritmalarını Kullanarak Rüzgar Enerjisi üretimi tahmini. https://doi.org/10.21205/deufmd.2021236709

Zhang, J., Jiang, X., Chen, X., Li, X., Guo, D., & Cui, L. (2019, 04 12). Wind Power Generation Prediction based on LSTM. *Proceedings of the 2019 4th International Conference on Mathematics and Artificial Intelligence*, 85-89. https://doi.org/10.1145/3325730.3325735

Zhang, L., Wang, K., Lin, W., Geng, T., Lei, Z., & Wang, Z. (2019, 07 09). Wind power prediction based on improved genetic algorithm and support vector machine. *IOP Conference Series: Earth and Environmental Science*, *252*. https://doi.org/10.1088/1755-1315/252/3/032052

APPENDIX A

CALCULATED TRENDS FOR THE WIND PARKS

Name of the Wind Power Plant	wf_id	Trend
ALİBEY RES	1	-0.45
İNTEPE RES	3	-0.05
Gökçedağ RES	5	-0.25
SAYALAR RES	7	-1.10
SEBENOBA RES	8	-0.26
YUNTDAĞ RES	9	-2.30
MAZI I RES	11	-1.20
MAZI 3 RES	13	-1.05
KORES KOCADAĞ RÜZGAR ENERJİ SANTRALİ	14	-0.68
SEYİT ONBAŞI RES	18	-1.94
SARIKAYA RES	25	-1.09

Name of the Wind Power Plant	wf_id	Trend
KELTEPE RES	26	-0.81
ÇATALTEPE RES	27	-1.78
ÇAMSEKİ RES	28	-0.49
DÜZOVA RES	29	-1.23
BOREAS I ENEZ RES	30	-0.87
DATÇA RES	31	-1.71
TURGUTTEPE RES	32	-0.88
KUYUCAK RES	34	-1.20
AKSU RES	36	-1.48
SAMURLU RES	37	-1.07
KOZBEYLİ RES	39	-1.19
BANDIRMA-3 RES	42	0.00

Name of the Wind Power Plant	wf_id	Trend
ŞENBÜK RES	43	-0.42
ŞAH RES	44	-0.32
BOZYAKA RES	45	-0.69
GEYCEK RES	46	-0.97
SARPINCIK RES	47	-1.64
KORKMAZ RES	48	-1.60
MORDOĞAN RES	49	-1.98
ALAÇATI RES	55	-1.05
YAYLAKÖY RES	56	-1.57
SARES RES	57	-1.30
ALİAĞA RES	58	-1.52
TEPE RES	62	-1.17

Name of the Wind Power Plant	wf_id	Trend
MADRANBABA RES	65	-1.01
ULUBORLU RES	68	-0.92
DİLEK RES	71	-0.04
SAKARBAYIR RES	76	-0.13
ZİNCİRLİ RES	87	1.56
ÖDEMİŞ RES	89	-0.42
AKBÜK RES	95	-1.21
EDİNCİK RES	102	-0.22
BAĞARASI RES	103	-0.66
KAROVA RES	107	-0.60
GÖKDAĞ RES	109	-0.63
YAHYALI RES	111	-1.60
KURTKAYASI RES	112	-1.57
İÇDAŞ BİGA RES	117	-0.73

Name of the Wind Power Plant	wf_id	Trend
KARADERE RES	128	-0.84
YAHYALI RES	131	-0.55
YAMAÇTEPE-2 RES	135	-0.71
KUYULUKOYAK RES	144	-1.67
ÇATALTEPE RES	146	-0.53
AKDAĞ RES	147	-0.67
AKBÜK II RES	148	-0.49
PİTANE RES	151	-1.30
BERGAMA RES	154	-0.75
BERGRES RES	155	-1.69
GERES RES	156	-0.60
GÖKRES 2 RES	159	-0.15
MAHMUT ŞEVKET PAŞA-1 RES	161	0.22
KINIK RES	169	-2.25

Name of the Wind Power Plant	wf_id	Trend
KIYIKÖY RES	174	0.36
AİRRES-4 RES	175	-0.64
BARBAROS RES	184	-0.11
YENİKÖY RES	186	0.32
KURTİNİ RES	187	-1.66
ŞİLE RES	190	-1.40
POYRAZ RES	191	-0.48
KALFAKÖY RES	194	-0.61
TİRE RES	196	-1.57
KIRKAĞAÇ RES	200	-0.92
MANASTIR-ESENKÖY RES	203	0.00
ATAKALE RES	228	-1.11

APPENDIX B

REASONS FOR WIND FARM REMOVAL FROM THE STUDY

Reason	ID of the wind farm in the WFD
Data does not cover an important amount of the study period	83, 92, 122, 132, 136, 226, 236, 264, 268, 287
Plant has too many availability events	12, 15, 19, 20, 22, 35, 41, 52, 53, 54, 59, 60. 61, 63, 64, 70, 72, 75, 79, 80, 82, 84, 86, 88, 93, 96, 104, 106, 108, 113, 114, 115, 116, 119, 127, 137, 138, 139, 141, 143, 145, 152, 157, 158, 162, 165, 167, 170, 173, 178, 195, 201, 202, 204, 211, 212, 221, 232, 233, 235, 242, 252, 258
New turbine installation/capacity increase during the study period	2, 4, 6, 16, 17, 33, 50, 51, 66, 67, 69, 73, 74, 77, 78, 91, 94, 98, 99, 100, 101, 105, 110, 121, 123, 124, 129, 130, 140, 149, 150, 153, 160, 166, 168, 171, 172, 176, 179, 180, 181, 181, 183, 185, 192, 197, 205, 209, 213, 217, 218, 219, 220, 223, 224, 227, 229, 231, 234, 237, 240, 243, 244, 246, 247, 248, 250, 251, 254, 262, 265, 267, 270
Other factors affecting data quality	38, 97, 126, 133, 230, 260

APPENDIX C

WFD ID-PLANT NAME MATCHES

Plant Name	wf_id
ALİBEY RES	1
KEMERBURGAZ RES	2
İNTEPE RES	3
KARAKURT RES	4
GÖKÇEDAĞ RES	5
ŞAMLI RES	6
SAYALAR RES	7
SEBENOBA RES	8
YUNTDAĞ RES	9
MAZI I RES	11
ÇEŞME RES	12
MAZI 3 RES	13
KOCADAĞ-2 RES	14
AYYILDIZ RES	15
BELEN RES	16
KAPIDAĞ RES	17
SEYİT ONBAŞI RES	18

Plant Name	wf_id
AKBÜK RES	19
SEYİTALİ RES	20
SOMA RES	22
BALIKESİR RES	24
SARIKAYA RES	25
KELTEPE RES	26
ÇATALTEPE RES	27
ÇAMSEKİ RES	28
DÜZOVA RES	29
BOREAS 1 ENEZ RES	30
DATÇA RES	31
TURGUTTEPE RES	32
MERSİN RES	33
KUYUCAK RES	34
BANDIRMA RES	35
AKSU RES	36
SAMURLU RES	37

Plant Name	wf_id
SÖKE-ÇATALBÜK RES	38
KOZBEYLİ RES	39
ÇANAKKALE RES	41
BANDIRMA III RES	42
ŞENBÜK RES	43
ŞAH RES	44
BOZYAKA RES	45
GEYCEK RES	46
SARPINCIK RES	47
KORKMAZ RES	48
MORDOĞAN RES	49
URLA RES	50
ŞENKÖY RES	51
MORDOĞAN RES	52
GERMİYAN RES	53
URLA RES	54
ALAÇATI RES	55

Plant Name	wf_id
YAYLAKÖY RES	56
SARES RES	57
ALİAĞA RES	58
DAĞPAZARI RES	59
KİLLİK RES	60
ARDIÇLI RES	61
TEPE RES	62
SEFERİHİSAR RES	63
ÇEŞME RES	64
MADRANBABA RES	65
MERYEM RES	66
ZONGULDAK RES	67
ULUBORLU RES	68
CERİT RES	69
AMASYA RES	70
DİLEK RES	71
KANİJE RES	72
HAMZABEYLİ RES	73
MUTLU RES 5 RES	74
SARAY RES	75
SAKARBAYIR RES	76

Plant Name	wf_id
KANDIRA RES	77
KOCATEPE RES	78
GAZİ RES	79
BEREKETLİ RES	80
ALBAY ÇİĞİLTEPE RES	82
ALİBEYHÜYÜĞÜ RES	83
İNCESU RES	84
KANGAL RES	86
ZİNCİRLİ RES	87
UMURLAR RES	88
ÖDEMİŞ RES	89
ADARES RES	91
EMRES RES	92
KARAÇAYIR RES	93
DİKİLİ RES	94
AKBÜK RES	95
ELMALI RES	96
ALARES RES	97
KİRAZLI RES	98
KARATEPE RES	99
TAYAKADIN RES	100

Plant Name	wf_id
KUŞADASI RES	101
EDİNCİK RES	102
BAĞARASI RES	103
POYRAZGÖLÜ RES	104
ÇANTA RES	105
DENİZLİ RES	106
KAROVA RES	107
YENİHİSAR RES	108
GÖKDAĞ RES	109
BAFA RES	110
YAHYALI RES	111
KURTKAYASI RES	112
KONAKPINAR RES	113
ÜÇPINAR RES	114
FATMA RES	115
SALMAN RES	116
İçdaş Biga RES	117
AYVALIK-I RES	119
KARTAL RES	121
GERMİYAN RES	122
SELİMİYE RES	123

Plant Name	wf_id
ADA 2 RES	124
MARMARA RES	126
PAŞALİMANI RES	127
KARADERE RES	128
HAVZA RES	129
YILMAZ RES	130
YAHYALI RES	131
YENİ RES	132
SİBELRES RES	133
ALAPINAR RES	134
YAMAÇTEPE-2 RES	135
ERTAN RES	136
ÇAMINBAŞI RES	137
EGE RES	138
PAMUKOVA RES	139
ULU RES	140
KOCALAR RES	141
AKYURT RES	143
KUYULUKOYAK RES	144
ZELİHA RES	145
ÇATALTEPE RES	146

Plant Name	wf_id
AKDAĞ RES	147
AKBÜK II RES	148
ÖMERLİ RES	149
HACI BEY RES	150
PİTANE RES	151
ŞENBÜK RES	152
ŞAPDAĞI RES	153
BERGAMA RES	154
BERGRES RES	155
GERES RES	156
KAVAKLI RES	157
G RES	158
GÖKRES-2 RES	159
EVRENCİK RES	160
MAHMUT ŞEVKET PAŞA-1 RES	161
SADILLI RES	162
KARABEL RES	164
ÇERÇİKAYA RES	165
HASANOBA RES	166
GÜNDOĞDU RES	167
ÖZBEK RES	168

Plant Name	wf_id
KINIK RES	169
ELMALI RES	170
AYDOS RES	171
MASLAKTEPE RES	172
KIZILCATERZİ RES	173
KIYIKÖY RES	174
AİRRES-4 RES	175
KÜPTEPE RES	176
GÜLLÜK RES	177
ORTAMANDIRA RES	178
YENİKÖY RES	179
ORHANLI RES	180
TATLIPINAR RES	181
BAĞLAR RES	182
AKKUŞ RES	183
BARBAROS RES	184
MERSİNLİ RES	185
YENİKÖY RES	186
KURTİNİ RES	187
ŞİLE RES	190
POYRAZ RES	191

Plant Name	wf_id
KORUDAĞI RES	192
KALFAKÖY RES	194
OVARES RES	195
TİRE RES	196
PETKİM RES	197
KARADAĞ RES	198
SİNCİK RES	199
KIRKAĞAÇ RES	200
KARACABEY RES	201
DATÇA RES	202
MANASTIR-ESENKÖY RES	203
KÜREK DAĞI RES	204
ARZU RES	205
ATİK RES	206
DEMİRCİLİ RES	207
GREİF RES	208
KARABURUN RES	209
ALAÇATI RES	210
ÇAKIL RES	211
YAHŞELLİ RES	212
GÜLPINAR RES	213

Plant Name	wf_id
SOMA RES	214
ZEYTİNELİ RES	215
KARADAĞ RES	216
SOĞANLI RES	217
BAĞLAMA RES	218
GAZİ-9 RES	219
AKYEL-1 RES	220
SÖKE RES	221
YALOVA RES	222
UŞAK RES	223
HİLAL-2 RES	224
ÇATALCA RES	225
VİZE-2 RES	226
GÖKZİRVE RES	227
ATAKALE RES	228
PAZARKÖY RES	229
ILGARDERE RES	230
AKYEL-2 RES	231
KARLITEPE RES	232
ÇAYPINAR RES	233
BOZÜYÜK RES	234

Plant Name	wf_id
GELİBOLU RES	235
SERTAVUL RES	236
GÖKTEPE RES	237
SARITEPE RES	238
DEMİRCİLER RES	239
ANDOZ RES	240
BOZCAADA RÜZGAR SANTRALİ RES	241
TAŞPINAR RES	242
İSTANBUL RES	243
GÜNEY-1 RES	244
ADAPAZARI RES	246
ERİMEZ RES	247
METAFOR RES	248
ARAPKİR RES	249
YAKAAĞZI RES	250
İPEKTEPE RES	251
SEFERİHİSAR RES	252
KAYADÜZÜ RES	253
GELİBOLU RES	254
GÜLÜMUŞAĞI RES	255
BEYYURDU RES	256

Plant Name	wf_id
ÖRLEMİŞ RES	257
EBER RES	258
ÇANDIR RES	260
ÇANDARLI RES	262
MİSKEVANK RES	264
ERCİYES RES	265
GAZİOSMANPAŞA RES	266
ALARES 2 RES	267
KARAMAN RES	268
ATARES-1 RES	270
GÖNEN RES	271
AKÇA RES	272
ALİAĞA RES	273
BURGAZ RES	274
METRİSTEPE RES	275
SİLİVRİ RES	277
OVACIK RES	278
SUSURLUK RES	279
HASANBEYLİ RES	280
FENER RES	281
KARAMÜRSEL RES	282

Plant Name	wf_id
HAMSİ RES	283
ÇERKEŞ RES	284
AKHİSAR RES	285
GERİŞ RES	286
GEYVE RES	287
BANDIRMA RES	288
SAROS RES	289
HARMANLIK RES	290
KARTALDAĞI RES	291
MENEKŞE RES	292
FUATRES RES	293
KORU RES	294
MUT RES	295
ÖRENLİK RES	296
ATARES-2 RES	297
GÜMÜŞTEPE RES	298
ŞULE RES	299
YELDEĞİRMENİ RES	300
KIYIKÖY RES	301
PELİT RES	302
D1 RES	303

Plant Name	wf_id
POYRAZ RES	304
ARIKONAK RES	305
BALABANLI RES	306
IŞIKLAR RES	307
SELİN4 RES	308
HARMANCIK RES	309
SÜLOĞLU RES	310
OVACIK RES	311
GÜNAYDIN RES	312
ZİYARET RES	313
AKKÖY RES	314
D3 RES	315
ÇEPNİ RES	316
ATASA RES	317
BAYIR RES	318