Improving Predictability of Wind Power Generation

by

Vivienne Jiao Zhang

B.A., Yale University (2015)

Submitted to the
INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY
and
DEPARTMENT OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

in partial fulfillment of the requirements for the degree of
MASTER OF SCIENCE IN TECHNOLOGY AND POLICY
MASTER OF SCIENCE IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2023

Authored by: Vivienne Jiao Zhang
Institute for Data, Systems, and Society, and
Department of Electrical Engineering and Computer Science
January 14, 2023

Certified by: Christopher R. Knittel
Professor, Applied Economics, MIT Sloan School of Management
Thesis Supervisor

Certified by: Caroline Uhler
Professor, Department of Electrical Engineering and Computer Science
Professor, Professor, Institute for Data, Systems and Society
Thesis Reader

Accepted by: Noelle Selin
Director, Technology and Policy Program
Professor, Institute for Data, Systems, and Society and
Department of Earth, Atmospheric and Planetary Sciences

Accepted by: Leslie A. Kolodziejski
Professor, Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Improving Predictability of Wind Power Generation

by

Vivienne Jiao Zhang

Submitted to the Institute for Data, Systems, and Society
Department of Electrical Engineering and Computer Science
on January 11, 2023, in partial fulfillment of the
requirements for the degrees of
Master of Science in Technology and Policy
and
Master of Science in Electrical Engineering and Computer Science

Abstract

Wind energy plays an important role in decarbonizing the economy and increasingly accounts for a growing share of electricity supply in the United States. However, availability of wind resource is highly dependent on variable factors such as weather and local geographies, making wind power generation forecast a particularly difficult task. This adds to the challenge of grid management, which requires that the supply of electricity equates the demand at all times.

Complicating the effort to improve wind power predicitability is a lack of empirical data, since wind power generation data are proprietary and often considered business secrets. To address this lack of empirical study, this thesis uses actual generation data between 2016 to 2021 from seven anonymized wind farms in Midwestern United States that range from 50MW to 235MW in size. The experiments demonstrate how machine learning methods can be used to forecast wind power generation at different time intervals, and how the accuracy of forecasting can be significantly improved when using a combination of newly extracted weather forecast data and weather measurement data. The economic benefits of more accurate forecasting are then studied using a using a simulation with market data from the Midcontinent Independent System Operator and the Southwest Power Pool. The thesis then explores whether predictability of wind power generation can be improved by placing weather stations closer to the wind forecast sites. Implications of these findings can inform investment decisions regarding weather monitoring stations and forecasting models, which can help electricity market participants adapt to a grid with an increasing share of renewable resources.

Thesis Supervisor: Christopher R. Knittel
Title: Professor, Applied Economics, MIT Sloan School of Management

Thesis Reader: Caroline Uhler

Title: Professor, Electrical Engineering and Computer Science
Professor, Institute for Data, Systems and Society

Acknowledgments

The past two and half years have been some of the most intellectually challenging and yet fulfilling time of my life. I couldn't have wished for a better learning experience, and for that, I have so many wonderful people to thank.

I'd like to first and foremost thank my advisor Professor Christopher Knittel for giving me incredible opportunities to explore my interests and never failing to offer support and guidance whenever I needed it. I also want to thank Professor Caroline Uhler for generously serving as my thesis reader and teaching me many of the data analysis skills necessary for this thesis in her IDS.131 class.

Big thanks also go to Joshua Hodge, Director of the MIT Center for Energy and Environmental Policy Research(CEEPR), where I have been incredibly lucky to work at as an Research Assistant, for connecting me with the wind power forecasting project that ultimately led to this thesis. I am also grateful for the research sponsor, for sharing their proprietary data and spending many hours with me on this analysis – without which this thesis would not exist. Tony Tran at CEEPR has also gone above and beyond to respond to every question or request of mine and helped make working at CEEPR such a joy.

My experience at MIT wouldn't have been complete without the classes and communities I have found myself a part of. To all of my TPP peers, never before have I met such a passionate and talented group of scholars, advocates, and professionals. Thank you for being who you are and for doing what you do. To Barb and Frank, I cannot thank you both enough for helping me navigate the peculiarities brought to our academic life by the pandemic. To my TPP friends who I was fortunate to team up with for class projects: Cathy, Jack, Tony, Joonhee, Disha, Paul, Farri, Lan, thank you for your wonderful company on the journey! To Nancy and Haruna, my undergraduate friends who I was reunited with again at MIT, going to school with you again makes it twice as fun. Thank you also for welcoming me into the Sloan community.

Finally, I want to thank my family, whose support encouraged me to pursue

graduate studies in the first place. And to Tejas, who brought so much light, laughter and love (and our wonderful puppy QiQi!) to the journey. Thank you for this, and so much more!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation and Background

Renewable energy now accounts for about one fifth of electricity generation in the United States, of which the biggest share comes from wind energy. After growing exponentially in the past two decades from 6 billion kilowatt-hours (kWh) in 2000 to 380 billion kWh in 2021, wind electricity generation today accounts for more than 9% of total utility-scale power generated in the US [14].

A challenge with wind power is that it is an intermittent renewable resource, generated only when the wind blows. At the same time, consumers expect electricity supply 24/7, and the power grid is a delicate system that requires supply equal demand at all times, making electricity the only commodity traded with instantaneous balancing due to physical constraints of the system. Storing power is possible, through forms such as pumped hydro-electricity, and more frequently in recent times, large-scale lithium ion batteries. But power storage at the moment is not an economic solution, at least not nearly at the level of the consumption level required at the national scale [17]. Accurate forecasting of the demand and supply of electricity power is therefore essential. As wind power becomes the largest category of variable electricity resource, how to more accurately forecast wind power generation has attracted much research.

Improving predictability of wind power generation is challenging, not just because the results depend on a variety of factors, from models, to weather data quality and local geospatial features, but also because there is a lack of empirical data. Wind power generation data belong to the owner of the wind farm. Because of how electricity is traded, wind farm owners usually hold the generation data confidential to prevent being disadvantaged in the market. So while there exist a multitude of studies on how to build the best machine learning model [32] and ensemble methods [18][21][26] for wind power generation to date using simulated data [19] or data from one individual turbine, few studies are based on empirical data from a cluster of wind power plants.

In this thesis, I use actual generation data between 2016 and 2021 from seven anonymized wind power farms in Midwestern United States that range from 50MW to 235MW in size. In applying machine learning methods to forecast wind power generation for these wind farms, I find that the predictability of wind power generation can be significantly improved when we use wind speed forecasts from National Weather Service, instead of using only past weather measurement data. The economic effect of improvements in wind power forecasting accuracy is then studied using a simulation. Depending on the trading strategy and the market the power plants trade in, I find that the more accurately forecasted energy can generate annual savings for more than $300,000/year. From this analysis, we gain a better understanding of the economic value we can harvest from investing in better quality data through the improved predictability of wind power generation. Understanding this value helps inform electricity trading strategies as well as investment decisions regarding the placements of weather monitoring stations, which in turn contributes to managing a more efficient and reliable grid that accommodates an increasing share of variable resources.

1.2 Thesis Scope and Outline

The outline of the thesis is as follows: Chapter 1 introduces the motivation for this analysis and outlines the rest of the thesis. Chapter 2 explains the underlying structure of the electricity marketplace in Independent System Operators (ISO) and Regional Transmission Organizations (RTO)-controlled areas in the US, including the day-ahead and real-time market, which account for 60% of the electricity consumed in the country [15]. It is against this backdrop that the economic discussion about wind power generation forecasting is organized. Chapter 3 introduces the data and a variety of machine learning methods used in forecasting wind power for different time-intervals. The data include the proprietary generation data of seven wind farms owned by the research sponsor, public weather station data from the National Oceanic and Atmospheric Administration (NOAA), as well as wind speed forecast data I extracted from the National Weather Service (NWS) archive. How wind power generation predictability can be improved using better quality data is then discussed. Note that in this case, we are focused on the data quality instead of the methodology. So we choose a commonly used machine learning model as the control and vary the data inputs to see improvement in predictability. Chapter 4 answers the question of how much economic benefit the improved forecasting ability can bring, using a simulation with market data from the Midcontinent Independent System Operator (MISO) and the Southwest Power Pool (SPP). Chapter 5 explores whether predictability of wind power generation can be improved by placing weather stations closer to the wind forecast sites. Chapter 6 summarizes the findings. In concluding, the chapter also discusses the implications of data inputs and model design for more accurately forecasting wind power generation, and the future work needed to respond to a grid that has a higher share of renewable resources in its mix.

Chapter 2

Electricity Markets

2.1 RTO/ISOs

The US electricity market consists of retail and wholesale electricity markets. The latter is where electric utilities, plant owners, and other market participants trade before the electricity is sold to retail consumers. The wholesale market is further divided into vertically integrated and competitive markets. The competitive markets are managed by RTO/ISOs, as shown in Figure 2-1. The markets account for the majority of electricity generated in U.S., about 60% of the nation's electricity power supply. The focus of this thesis is on the RTO/ISO regions [15].

With the exception of the Electric Reliability Council of Texas (ERCOT)[1], RTOs and ISOs (hereby referred to simply as RTOs) are regulated by the Federal Energy Regulatory Commission(FERC). They first formed in the 1990s in response to FERC's request to incentivize more competition among power generators by making transmission open and accessible to all qualified power generators. The idea that drove this reform was that, unlike electricity transmission which is a natural monopoly, electricity generation can achieve greater efficiency through competition [15].

In the Northeast, the RTOs evolved from power pools, which were multilateral arrangements with members ceding operational control over their generating units and

[1]Because ERCOT is not connected to other states and there is no inter-state commerce involved in this case, ERCOT falls outside out FERC's jurisdiction

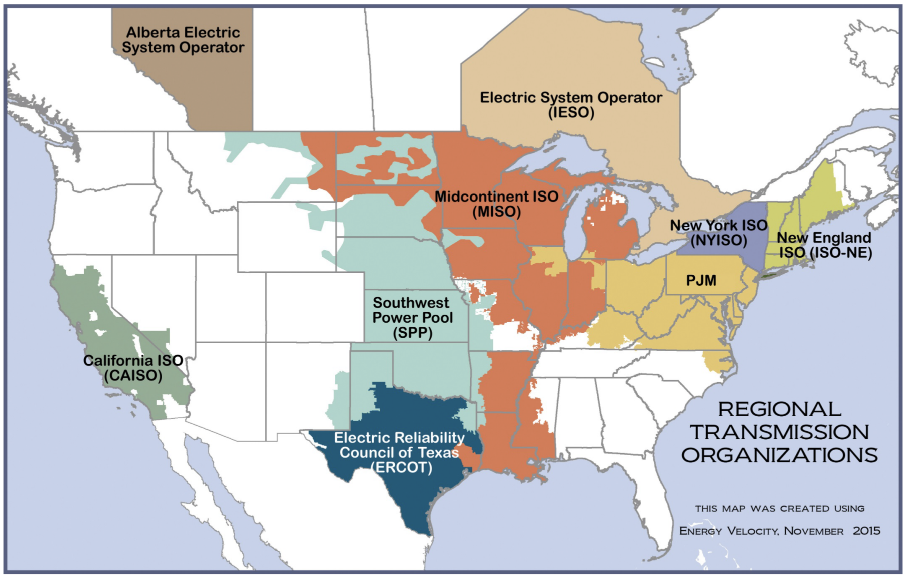


Figure 2-1: Map showing the RTO/ISO territories [17]

transmission facilities to a common operator to optimize for the best dispatch system. The first power pool was Pennsylvania-New Jersey-Maryland Interconnection (PJM), covering all or parts of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia and the District of Columbia. A power pool was in theory more efficient than an individual, standalone, utility because if a utility's load spikes to twice its capacity, instead of having to double its capacity (expensive and often unfeasible), a member connected to a power pool can simply buy power from the pool assumed to be produced by another member with excess capacity. Later, RTOs like California Independent System Operator (CAISO) for California, Southwest Power Pool (SPP) and Midcontinent Independent System Operator (MISO) for the Midwest also emerged to encourage competitive generation and open transmission access [17].

In simplification, there are three primary participants in an RTO marketplace. They include:

- The Seller: Independent generators who own power plants only and make money by selling merchant power or by entering bilateral supply contracts

- "The Neutral Middleman": Transmission companies who earn a fixed rate of return on the transmission infrastructure they invested in subject to FERC and other regulatory bodies' approval

- The Buyer: Load serving entities who buy power in the wholesale marketplace to serve their retail customers. The load serving entities can buy power either by contracting with the independent generators or through the market via different mechanisms that will be explained in the next section.

2.2 Day-Ahead and Real-Time Energy Markets

An RTO market typically has two components: the Day-Ahead(DA) Energy Market and the Real-Time (RT) Energy Market. Both components constitute a "multi-settlement" system for transactions to take place [2]. As explained in the introduction, due to the physical constraints of the system, there needs to be instantaneous balancing between the demand and supply of electricity. This can make electricity prices extremely volatile when either supply or demand drops suddenly. By offering a Day-Ahead Energy Market where participants can buy or sell wholesale electricity one day before the day of operation, some price volatility can be avoided. This is the first financial settlement.

The prices are location-dependent within an RTO. According to ISO New England, the regional pricing "reflects (1) the operating characteristics of, and (2) the major constraints on, the ... transmission system at each area, as well as (3) the losses resulting from physical limits of the transmission system." They are calculated at intervals between 1 hour to 5-minute [2]. Figure 2-2 is an example of the DA price map valid before 5pm EST on December 22, 2022.

During the day of operation at intervals up to 5-minute, participants will buy and sell electricity in the Real-Time Energy Market. Any differences between the day-ahead commitments and the actual amount of electricity transacted will then be adjusted. Notably, Real-Time Energy Market will have different prices for electricity transacted from the Day-Ahead prices. Figure 2-3 offers an example of the RT price.

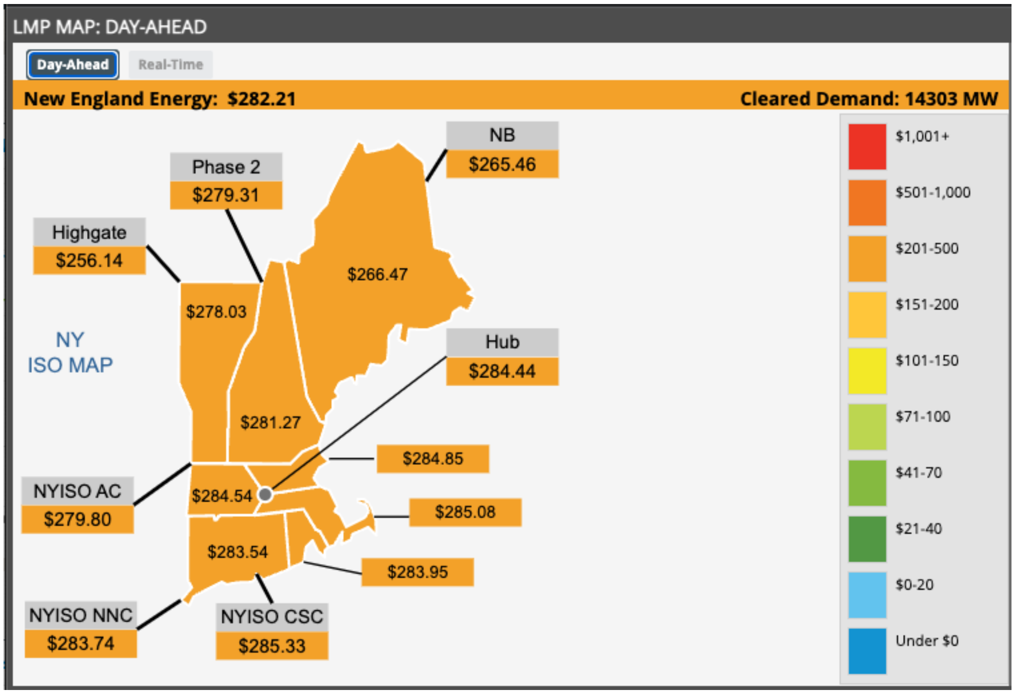


Figure 2-2: Map showing the DA prices in different "nodes" within New England ISO territory on December 23, 2022 [1]

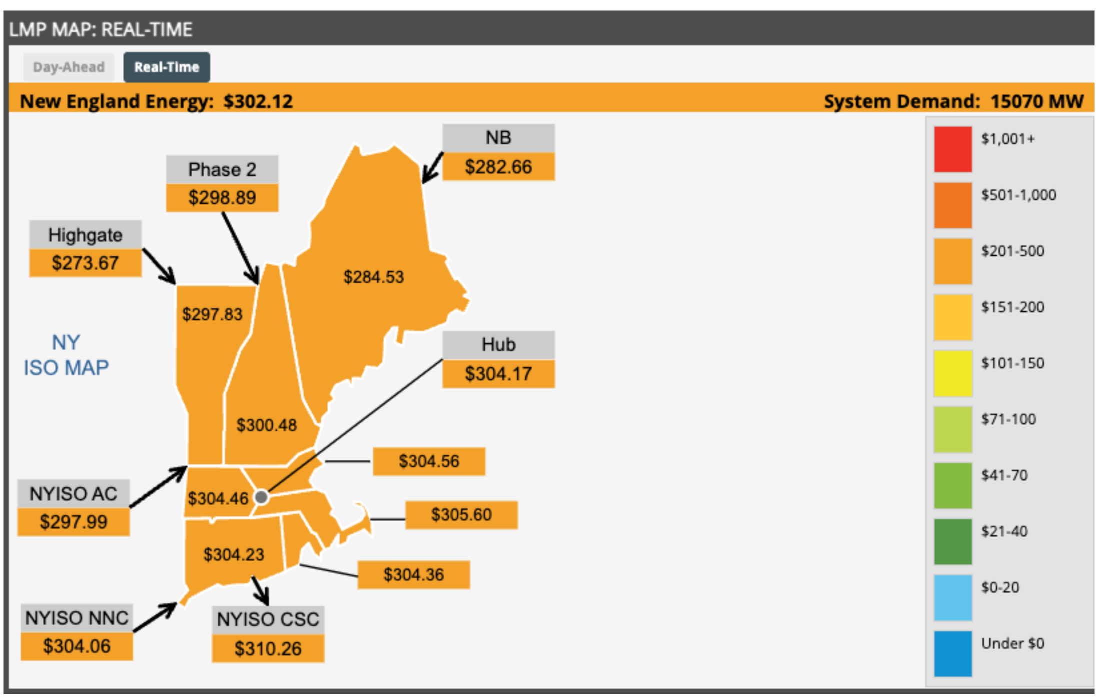


Figure 2-3: Map showing the RT prices in different "nodes" within New England ISO territory. Sampling period is 4:40 pm EST, December 23, 2022 [1]

The Real-Time Energy Market is needed to incentivize generation and transmission of electricity to areas where the electricity demand is high and puts the local load zone under strain. In nodes where the grid is stressed, the regional prices can increase rapidly up to upwards of $1000/MW, as shown in red in the legend in Figure 2-3. The high prices encourages electricity providers to either release stored energy or fire up a natural gas power plant that can be turned up on short notice to meet the high demand. This RT market leads to a second financial settlement, in addition to the first DA settlement. It uses the the established real-time prices to either charge or pay participants in the Day-Ahead market for consumption or generation that deviates from the DA settlements.

There may be some variances between the different RTO markets, but the fundamental principle of the multi-settlement system stays the same for a competitive electricity market. This multi-settlement system is the primary reason for why forecasting of one's wind power plant production can have direct and significant impact on their earnings, in the case of the electricity seller, which will be explained in detail in Chapter 4. Before that, in the next chapter, I will first discuss how to forecast wind power generation using machine learning and how the accuracy of forecasting can be improved using different data inputs.

Chapter 3

Wind Power Generation Forecasting

3.1 Introduction

This chapter first surveys the methodologies currently used to forecast wind power generation, and then focuses on one machine learning method, the Long Short-Term Memory as the main methodology in this analysis to make forecasting comparisons. Data for the analysis are introduced, including wind generation data, weather measurements from various weather stations, as well as wind speed forecast data by the National Weather Service. Models and the analysis are then presented. In the last section of this chapter, I discuss the results and some of the implications of how to better forecast wind power generation.

3.2 Methodology

3.2.1 Literature Review

Wind power generation forecasting generally consists of two methods: a physics-based approach and a statistics-based approach. The former is usually more accurate, particularly for longer-range forecasting, but is computationally intensive as it has to simulate the physical environments. Some methods combine both to form a hybrid approach [20][27][16][28].

This thesis is focused on using machine learning methods, which fall under the statistics-based approach. Machine learning broadly is "the science of getting computers to learn by a practice of using algorithms to analyze data and make predictions. Machine learning can learn from data without being explicitly programmed [24]." The process generally involves extracting training data and fitting the data on a model. At the end of training, the model outputs the parameters that can be generalized to new data to make predictions. Because a machine learning forecasting model does not need for the algorithms to be specifically programmed, it is relatively fast to develop a model [24]. In terms of data inputs, a machine learning forecasting model usually requires numerical weather information, time series data of wind speed and temperature, which are easier to obtain and process than a physics-based approach requires, such as geospatial features. Lastly, apart from the process of developing models, it is also faster for a machine learning model than a physics-based model to make a prediction. These attributes makes machine learning a more attractive approach for individual plant owners to adopt.

For these reasons, and with the rapid development of machine learning in Computer Science in the past decade, there is a plethora of machine learning methods researchers have developed and experimented with to forecast wind power generation. They can be broadly summarized into four categories: Neural Networks, kNN, SVM, and Random Forest. Jørgensen et.al [24] in Figure 3-1 show the rise of Neural Networks as the most popular approach for its flexibility and accuracy across seasons for forecasting wind power generation.

Within Neural Networks, many novel and complex networks have emerged including more recently auto-encoders [32] and ensemble methods [18][30][31][21] that combine a few different machine learning methods into one.

In this thesis, I choose long short-term memory networks (LSTM) as the Neural Network for forecasting wind power generation. An LSTM is an variation of Recurrent Neural Networks (RNN), which have the special property of retaining some memory of data inputs from earlier periods, allowing the networks to capture more temporal relationships. As a result, LSTM is commonly used in time series data, including

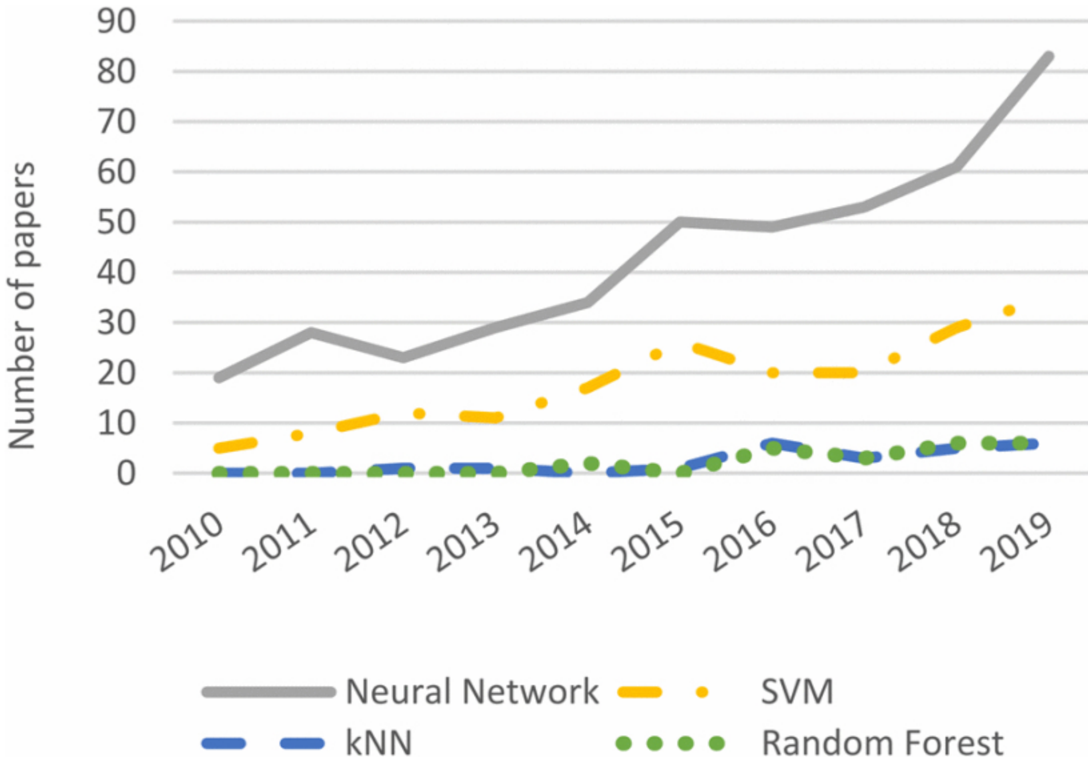


Figure 3-1: A chart showing the development of research interest in machine learning forecasting methods for wind power generation over the last decade [24].

the practice of forecasting wind power generation [26][23]. It also has reasonable performance. I experimented with a few of LSTM's variations such as Convolutional LSTM and Bidirectional LSTM [33], in the hope of capturing some spatial relationships of the data, but no noticeable improvement were observed. LSTM stood out as a reasonably simple and performant method for this analysis. Additionally, LSTMs are fairly fast to implement. This attribute is important as the focus of the thesis is to understand how data inputs influence the accuracy of wind power generation forecasting. So we need to control the method of forecasting and vary the data inputs to seven wind power plants, which results in a total of 14 models and their variations. The LSTM method is discussed in detail in the next section.

3.2.2 LSTM

RNNs are a class of neural networks that can utilize previous outputs as inputs using special units called the hidden states [9]. This design allows RNNs to retain some memory of past information in giving a new output. As a result, RNNs usually outperform traditional Feedforward Neural Networks on sequential data.

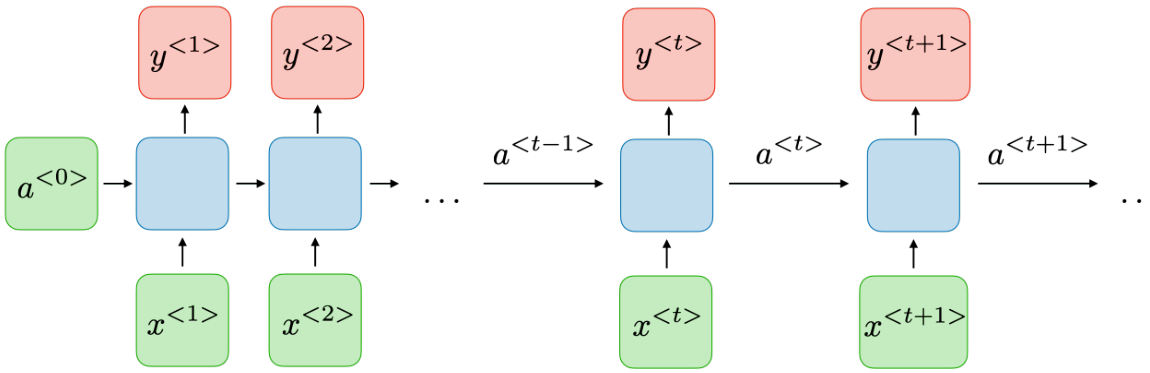


Figure 3-2: An illustration showing how previous outputs are used as inputs in a typical RNN architecture. [9]

As shown in Figure 3-2, t represents each time step. At time t, $x^{<t>}$ is the input, given to $a^{<t>}$, the activation (with g_1 and g_2 being the activation functions), whereas $y^{<t>}$ is the output. W_{aa}, W_{ax}, W_{ya}, b_a, b_y are parameters that are shared across different time steps.

$$a^{<t>} = g_1\left(W_{aa}a^{<t-1>} + W_{ax}x^{<t>} + b_a\right) \tag{3.1}$$

$$y^{<t>} = g_2\left(W_{ya}a^{<t\rangle} + b_y\right) \tag{3.2}$$

The blue units shown in Figure 3-2 and Figure 3-3 in more detail are the hidden units, where equations 3.1 and 3.2 are applied. These are the "inner workings" that help to retain information from many time steps back for the current time step t, by adjusting the weights of W_{aa}, W_{ax}, W_{ya}, b_a, b_y.

When training an RNN, the loss function is the sum of the losses at every time step $\mathcal{L}$, as shown by Equation 3.3. The training is done by backpropogation through time as shown in Equation 3.4. Through gradient descent, the weights $\mathcal{W}$ are adjusted to minimize the loss function.

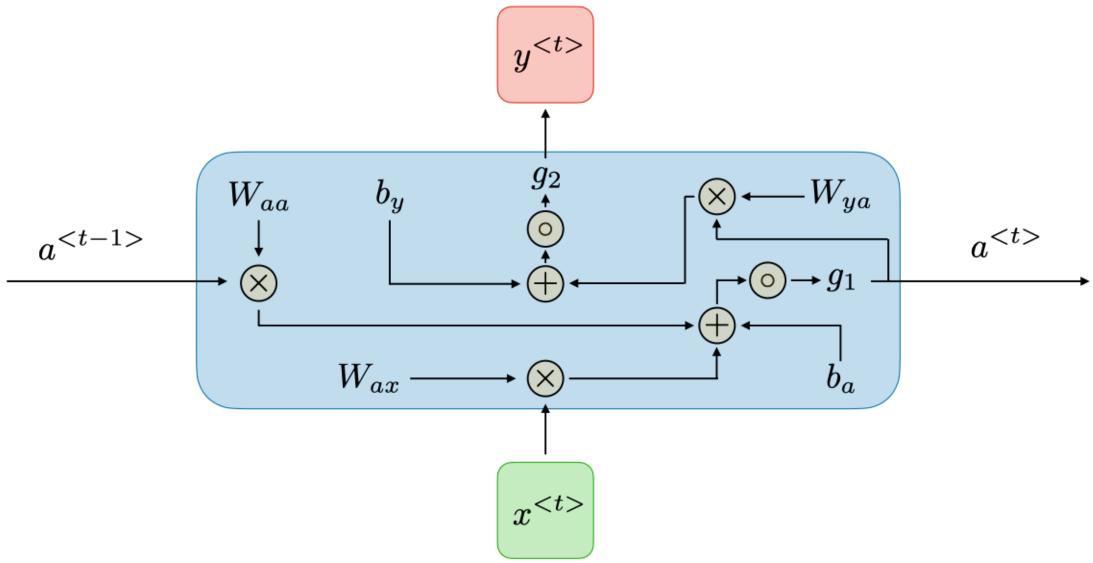


Figure 3-3: The "inner workings" of a hidden unit. [9]

$$\mathcal{L}(\hat{y}, y) = \sum_{t=1}^{T_y} \mathcal{L}\left(\hat{y}^{<t>}, y^{\langle t>}\right) \tag{3.3}$$

$$\frac{\partial \mathcal{L}^{(T)}}{\partial W} = \sum_{t=1}^{T} \left.\frac{\partial \mathcal{L}^{(T)}}{\partial W}\right|_{(t)} \tag{3.4}$$

The challenge with RNNs is vanishing or exploding gradient during backpropagation through time. This happens often as the inputs go through more layers of the networks, the multiplicative gradient can decrease or increase exponentially [9].

LSTM, introduced by in the 1990s [22], helps remedy the vanishing gradient problem by introducing gates to determine whether to retain or discard information when giving a new output. The gates are defined as:

$$\Gamma = \sigma\left(Wx^{\langle t\rangle} + Ua^{\langle t-1>} + b\right), \tag{3.5}$$

σ is the sigmoid activation function. W, U, b are specific coefficients to the gate, depending on the function of the gate. There are four types of gates in LSTM:

- Update gate: How much past information to use?
- Relevance gate: Drop previous information?

- Forget gate: Erase a cell?
- Output gate: How much to reveal of a cell?

The four gates are deployed in one memory cell of an LSTM model as shown in Figure 3-4.

$$\tilde{c}^{<t>} : \tanh\left(W_c\left[\Gamma_r \odot a^{<t-1>}, x^{<t>}\right] + b_c\right) \tag{3.6}$$

$$c^{<t>} : \Gamma_u \odot \tilde{c}^{<t>} + \Gamma_f \odot c^{<t-1>} \tag{3.7}$$

$$a^{<t>} : \Gamma_o \odot c^{<t>} \tag{3.8}$$

The sign $\odot$ denotes the element-wise multiplication between two vectors. At time step t, $x^{<t>}$, the input, is passed through these four gates. If the input helps achieve a lower loss function, the information is likely to be retained through the mechanism of these gates. Otherwise, it can be "forgotten." Because the cells no longer retain information from many layers prior that may no longer have predictive value, the vanishing or exploding gradient challenge is remedied.

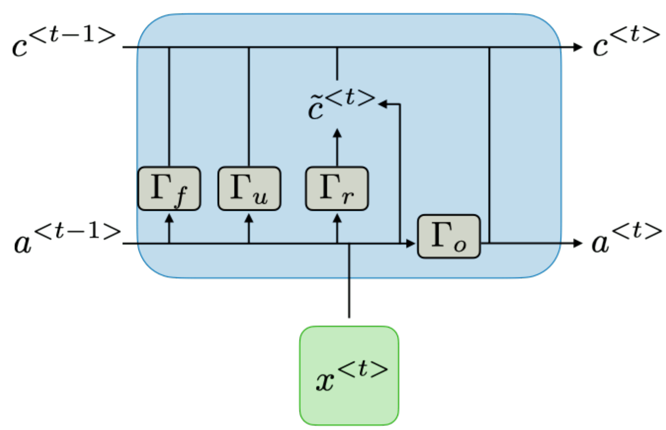


Figure 3-4: The Characterization of LSTM [9]

Note that Figure 3-2 illustrates only one structure of the relationship between inputs and outputs (many-to-many). There exist variations, such as one-to-many, or

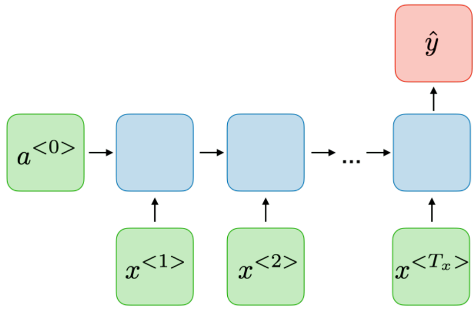


Figure 3-5: Many-to-one RNN architecture [9]

many-to-one (Figure 3-5), which is what I will use in predicting wind power generation. The model will be explained in detail in Section 3.4.

3.3 Data

3.3.1 Wind Power Generation Data

The wind generation data come from a cluster of seven wind farms in about an 80-kilometer-radius area owned by one utility in the Midwest who participate in both the MISO and SPP RTO markets. Because of how electricity is traded, wind farm owner usually holds the generation data confidential to prevent being disadvantaged in the market. The actual names and the exact locations of the power plants are anonymized in this thesis.

The approximate locations of the seven code-named wind farms (in red) are shown below in Figure 3-6. More details about the data are summarized in Table 3.1. Note that with the exception of OS, all the data are in 1 hour intervals. For Project OS, 5-min interval data are also available.

3.3.2 Weather Data

Weather data from 2015 to 2021 were obtained from local weather stations through NOAA [7]. Eight weather stations were selected due to their proximity to the wind farms. Figure 3-6 shows the weather stations (in orange) in relations to the seven wind farms on the map. In this analysis, only the data from the station closest to the

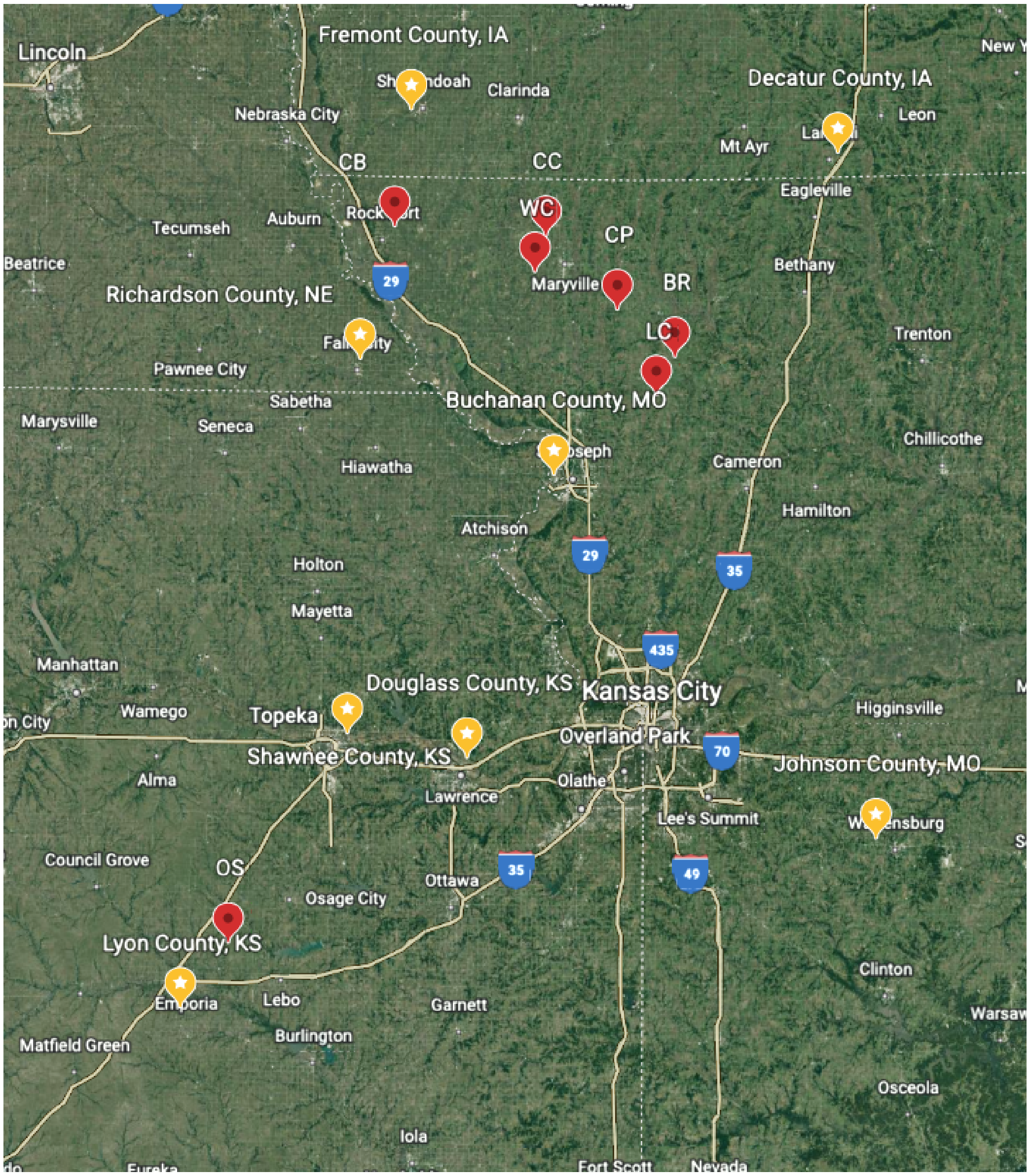


Figure 3-6: Approximate locations of the seven wind power plants. Image made with Google Earth.

Table 3.1: Summary of Wind Generation Data

Project	**Capacity/MW**	**Years Availalble**
OS	150	2015 - 2021
LC	150	2016 - 2021
CB	50	2020 - 2021
CP	50	2020 - 2021
WC	238	2020 - 2021
BR	57	2019 - 2021
CC	235	2019 - 2021

wind farm is used in making predictions. For instance, to forecast the wind power generation of Project OS, only the wind measurement data from Lyon County are used. The data contain temperature, wind direction, humidity level, wind speed, and other measurement data at hourly intervals. Some sources may have occasional missing data due to equipment failure, but more than 99% of data are complete.

3.3.3 National Weather Service Forecast

The NWS as a part of NOAA makes detailed weather forecasts up to seven days in advance for use by emergency managers and the public. Live streaming as the forecast becomes available is easy to access at the NWS website [5]. However, accessing past forecasts by NWS is much more difficult, and to my knowledge, has not been done in current literature for forecasting wind power generation. With the help from the staff at NOAA, I was able to download archives of forecasts from the National Digital Forecast Database (NDFD) [4].

The data are available in GRIBv2 format, which stands for General Regularly-distributed Information in Binary form. GRIB is commonly used in meteorology as an efficient format to store historical and forecast weather data and is standardized by the World Meteorological Organization. The National Centers for Environmental Prediction maintains the documentation of GRIB [6].

One challenge with the GRIBv2 data is that they are presented in grids only, as shown in Figure 3-7. To translate these data into numerical wind speed forecasts of reasonable spatial granularity, I used a GRIBv2 decoder, DeGRIB, developed by the Meterological Development Laboratory. DeGRIB has a point-probe capability: by querying a location that ranges about a 5-km square in longitude and latitude, it can output the requested forecast data at a specific time.

I downloaded the NDFD Central Plains grids from 2015 - 2021 (Data Header: YBK and YCK [4]), to match the wind power generation data, and then used a Python script that queried the GRIB files with DeGrib, to get point-wise numerical forecasts of wind speeds at specified locations within a 5km radius to where the wind power plants are located. Note that wind speed forecasts are only made in 3-hour intervals, meaning that between hour 00z to 03z, 03z to 06z..., only one forecast wind speed is given for each 3-hour period.

How the forecast data in 3-hour intervals are added to weather station data in hourly intervals as inputs for machine learning models will be explained in the next section on Models.

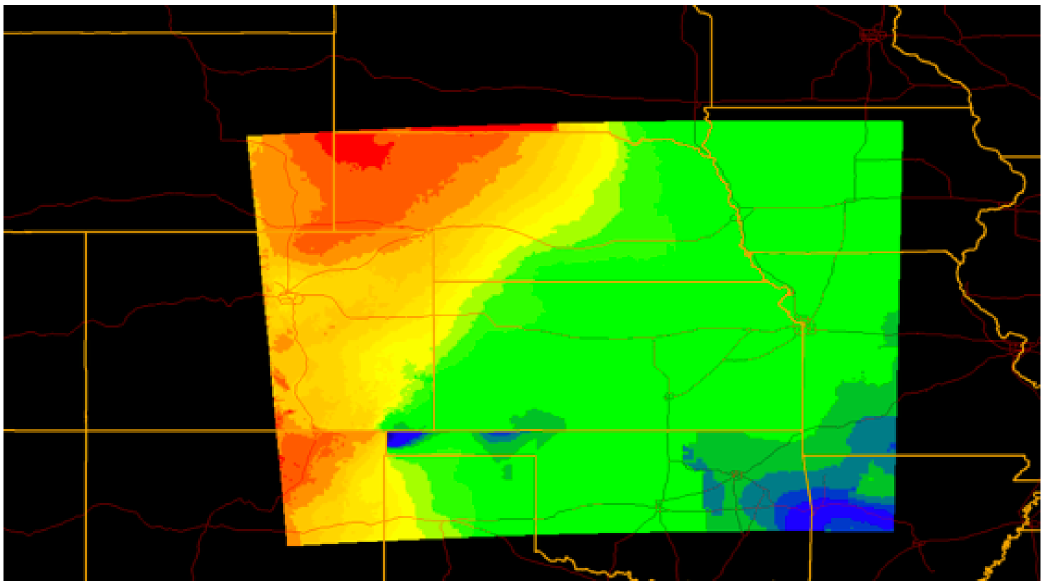


Figure 3-7: GRIBv2 data accessed using NOAA software.

3.4 Models

The LSTM model has a many-to-one structure. Three different kinds of predictions are done using this structure. The first kind, illustrated by Figure 3-8, is a short-range model that uses past generation data to predict the generation of the immediate next time step. This is only done with the data from Wind Project OS, which has both 5-min interval data as well as hourly interval data. The analysis is done simply to contrast performance of the machine learning model on short-range (1 hour-ahead) vs longer-range prediction(24 hour-ahead).

The second kind is a 24-hour-ahead forecast using generation data and wind speed data in the nearest weather stations for the past 24 hours to predict the wind power generation at the 48th hour, as shown in Figure 3-9. The 24-hour period in advance

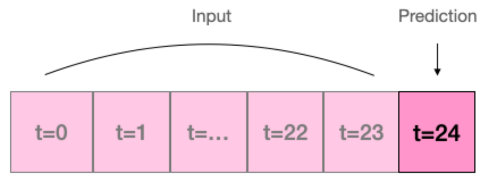


Figure 3-8: Data input for next hour prediction

is a minimum requirement for the forecasts to be useful for day-ahead trading, as explained in Chapter 2. This serves as the baseline model.

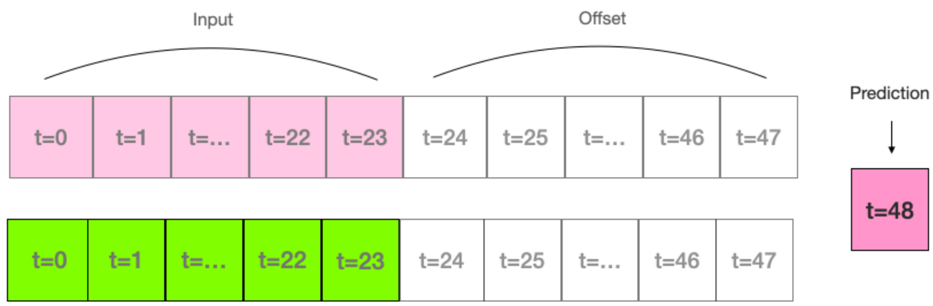


Figure 3-9: Data input for 24-hour ahead prediction using past generation and past wind speed measurements. The pink tiles stand for generation data while the green tiles are for the wind speed data.

The third experiment is to contrast with the baseline model. It also makes a 24-hour-ahead forecast, but in addition to using the generation data and wind speed data in the nearest weather stations, it also utilizes the weather forecast data as explained in Section 3.3.3. Figure 3-8 shows the three different kinds of data inputs and predictions diagrammatically.

For the second and third experiments, the data used for training each LSTM for each individual wind farm are summarized in Table 3.2. As an example, for the Project OS, in the base case, the data used are the past generation data from OS and the weather data from Lyon county. For the NWS case, both the generation data from OS and the weather data from Lyon county, in addition to the NWS forecast data for the location with lat/long (38.53, -96.03) are used for the prediction.

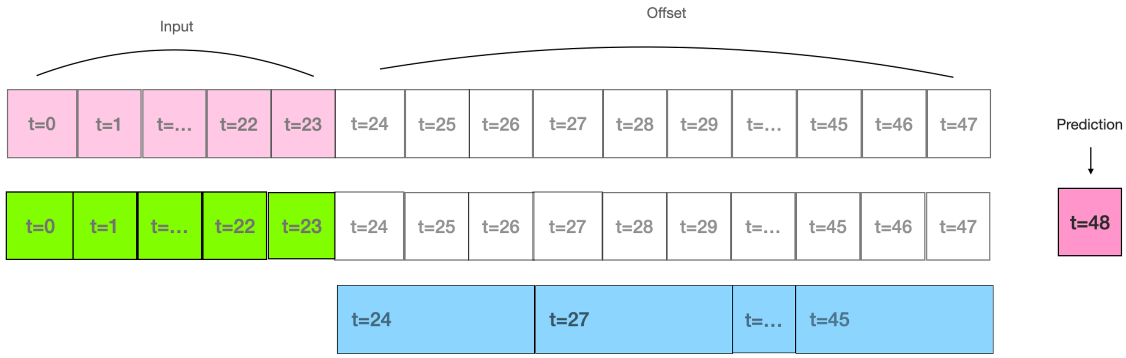


Figure 3-10: Data input for 24-hour ahead prediction using past generation and past wind speed, as well as weather forecasts. The blue tiles stand for wind speed forecast data from NWS. Note that because wind forecasts are available only in 3-hour intervals, for a 24-hour span, we have 8 evenly-spaced data points.

Project	**Base Case Data Inputs**	**NWS Source (Lat/Long)**
OS	OS Generation, Lyon County	38.53, -96.03
LC	LC Generation, Buchanan County	39.98, -94.54
CB	CB Generation, Fremont County	40.41, -93.93
CP	CP Generation, Buchanan County	40.20, -94.69
WC	WC Generation, Richardson County	40.30, -94.98
BR	BR Generation, Buchanan County	40.09, -94.49
CC	CC Generation, Fremont County	40.44, -95.48

Data cleaning was done through Pandas [8]. The three data panels were joined by timestamp into one large data-frame, which are then coded to form tensors as shown above as data inputs for the machine learning model. The LSTM model is written in TensorFlow and some of the code follows the TensorFlow tutorial on time series forecasting [12]. The data are split into training, validation, and test data-sets by the ration of 7:2:1. The optimizer used for training is ADAM [25]. The loss function used for optimization is Mean Squared Error, defined as:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2, \tag{3.9}$$

The performance metrics I used was Mean Absolute Error, which is defined as

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n}. \tag{3.10}$$

We find that the model performs best when training with the loss function of MSE. However, MAE is easier for interpretation, since it translates directly to the divergence between the forecast and the actual generation, and is therefore used as the validation and test metric. Note that the data are not normalized. I experimented with data normalization and did not see an improvement in performance. The hyper-parameters for the first LSTM model used for Project OS were finely tuned, and then kept the same and used subsequently for all other models used to predict other projects. This was done because we want to control the model and training itself while we investigate how performance of prediction changes as we vary data inputs. The hyperparameters selected are listed below [11]:

- Hidden units: 32
- Batch size: 32
- Hidden layers: 2
- Activation: relu
- Number of training epochs: 50

The model that has the lowest MAE during the 50 epochs of training is saved and then used to make predictions on the test data sets. The results are reported in the next section.

3.5 Results and Discussion

3.5.1 Short-range (Next Timestep) Prediction

For the 5 min-interval data, the MAE for this prediction is 5.48, which translates to 3.65% MAPE (mean absolute percentage error) for Project OS at size 150MW.

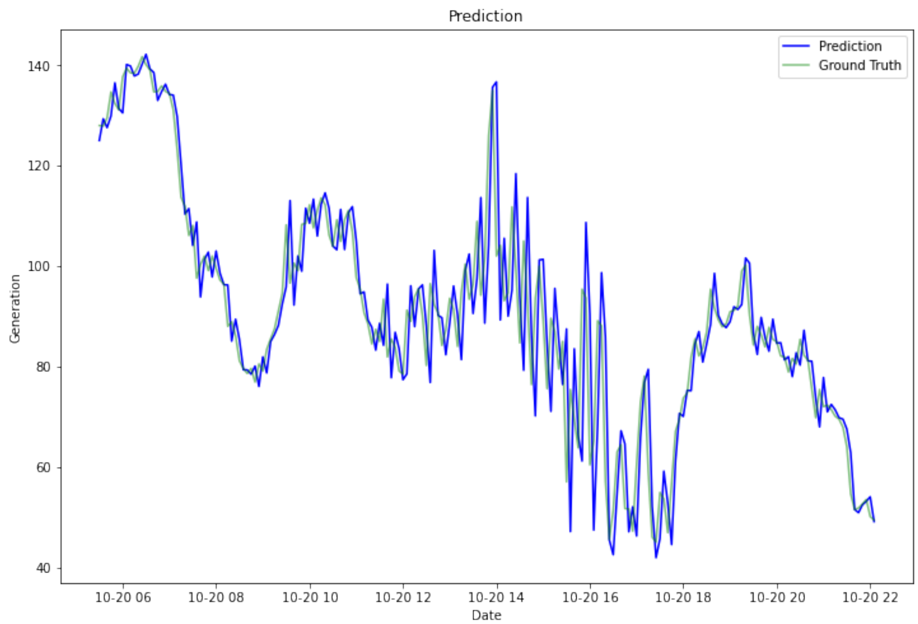
The prediction and the ground truth are plotted in Figure 3-11 which shows that the prediction follows the actual generation very closely. This MAPE is much lower than the typical industry-quoted 10% MAPE goal for wind power forecasting [26]. LSTM is able to outperform a persistence model or physics-model in this case because the immediate next timestep (the label for the machine learning forecast) has a strong statistical relationship with the past data (inputs) and the machine learning model is able to accurately capture this relationship.

Figure 3-12 shows the performance of the same LSTM model structure trained on hourly data. The performance has deteriorated somewhat from the previous example using 5-min interval data, since wind speed is much more likely to change suddenly in an hour than in 5 minutes. However, the approximately 10% error rate performs still very well in comparison to the persistence methods and can be used for the RT electricity market as described in Chapter 2 where price settlements are determined down to 5-min intervals.

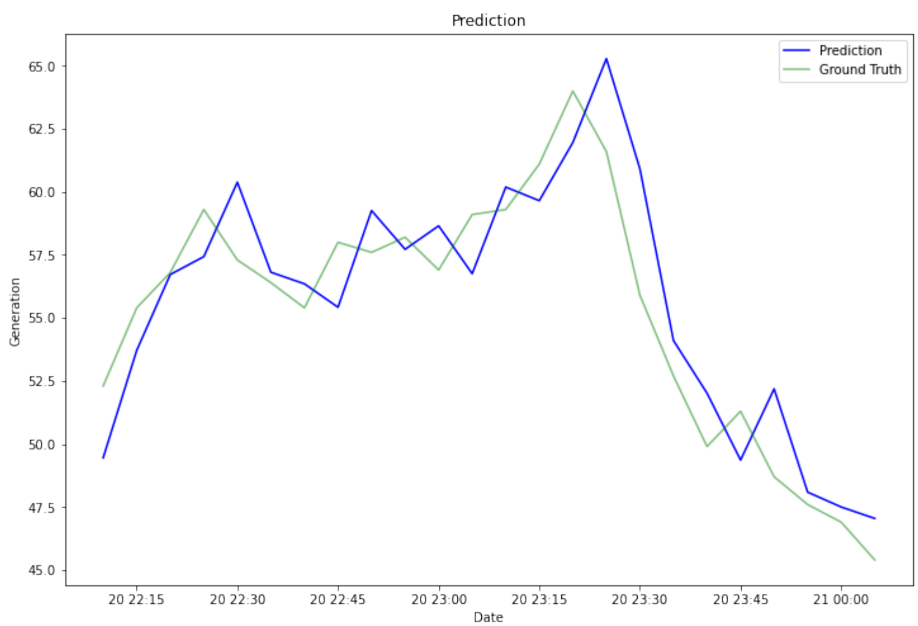
3.5.2 Long-range Prediction - 24 hour-ahead

Graphs 3-13 shows the performance of the forecasting models on the seven wind power plants. The green line is the ground truth. The red line is the base case, while the blue line is the NWS case. As discussed in the previous section (Table 3.2), each wind farm has its own model trained with its unique data. Data inputs for the base case are past wind power generation and wind speed from the the power plant's nearest weather station in the past 24 hours to predict the generation at the 48th hour (Figure 3-9). This 24-hour offset is necessary to make a forecast useful for the day-ahead market. Data inputs for the NWS case, which we are evaluating against the base case, are the NWS wind forecast data collected in the 5km radius of which the power plant is situated, in addition to the past wind power generation and wind speed data used in the base case (Figure 3-10). The performance metrics of the base case vs. the NWS case are shown in Table 3.3.

The first observation is that the performance of both the base case and the NWS case is worse than predictions made for the immediate next timestep, declining from



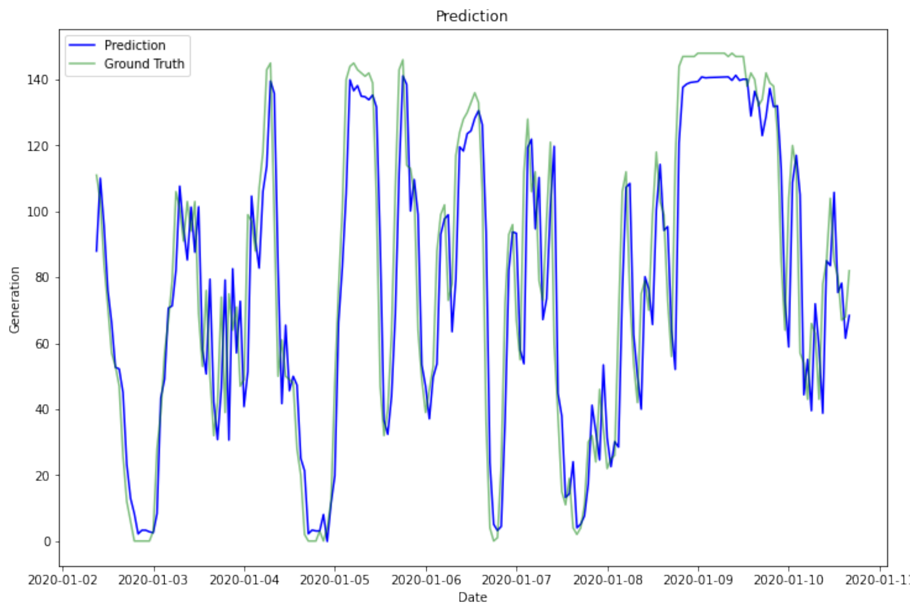
(a) Graph shows near perfect next timestep prediction using LSTM for 5-min intervals, with an MAE of 5.48.



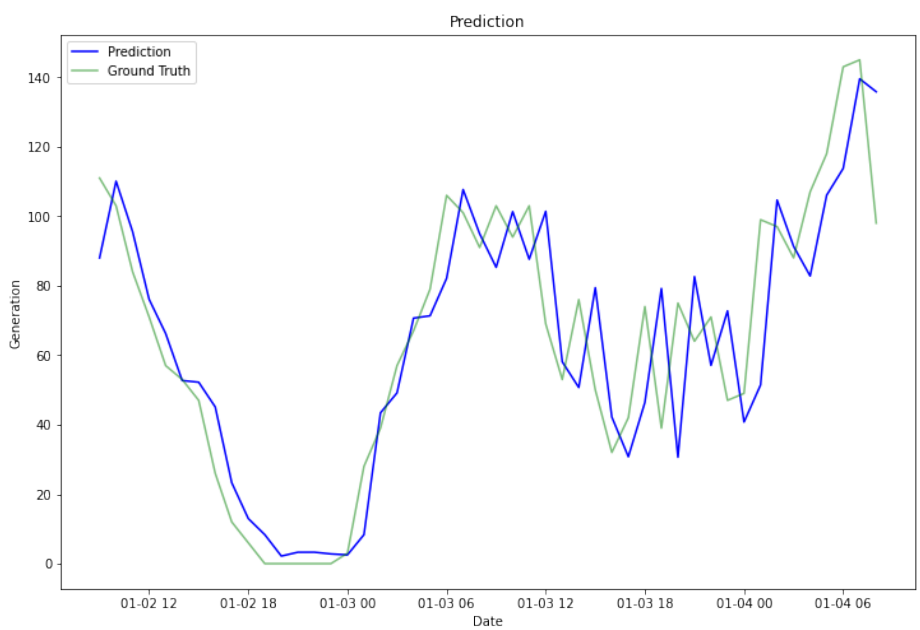
(b) This is Graph(a) zoomed in. Note that even though the predictions closely follow the ground truth, the learned model is not "cheating" by simply copying the value of the previous timestep, which would make this essentially persistence forecasting.

Figure 3-11: Prediction vs Ground Truth using 5-min interval generation data.

about 10% to on average 23.46% for the base case and 17% for the NWS case. This decline in accuracy is expected as performance of statistics-based methods typically



(a) Next timestep prediction using LSTM for 1 hour intervals with generation and wind data. The MAE is 16.15

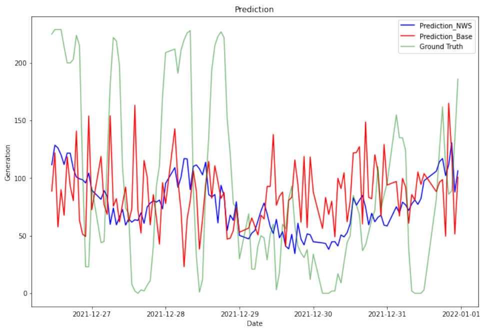


(b) Graph (a) zoomed in

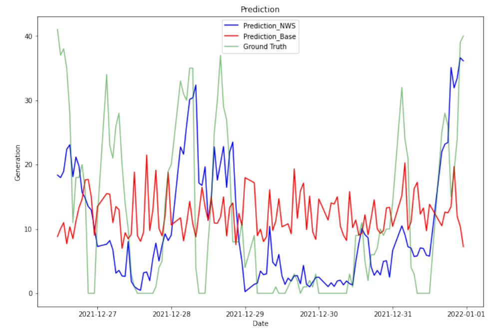
Figure 3-12: Prediction vs Ground Truth using 5-min interval generation data.

deteriorates as forecasting horizon increases.

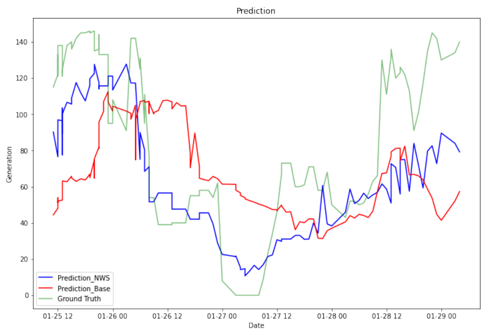
Another interesting finding is that the MAPE can vary quite considerably even for projects with essentially the same data inputs. For instance, the MAPE for OS is 19.86% while the MAPE for LC is 13.81% for both the NWS cases. Because the sizes



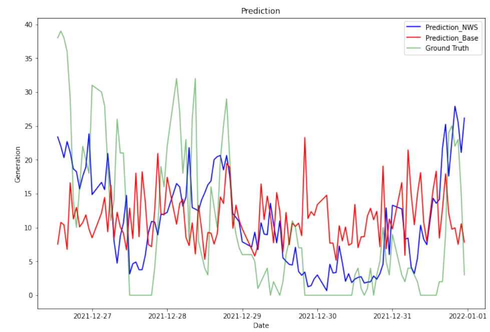
(a) CC



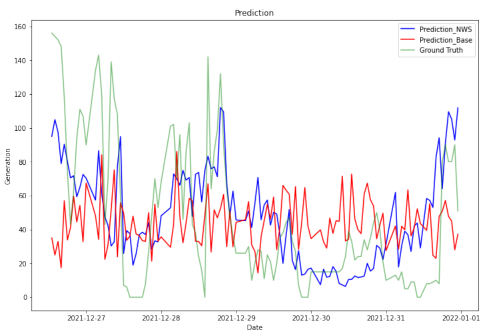
(b) CB



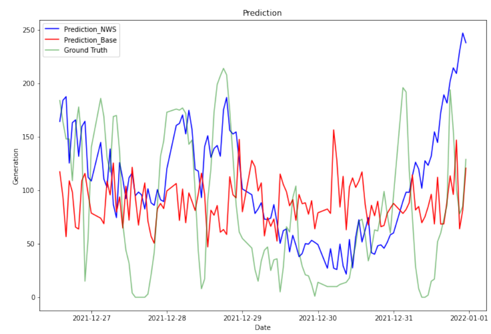
(c) OS



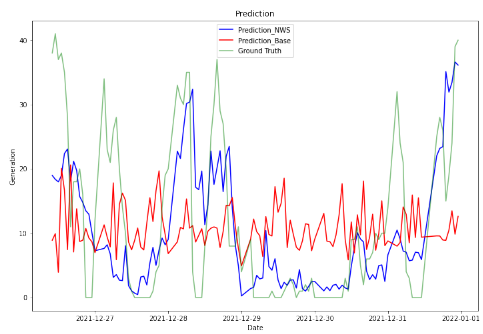
(d) BR



(e) LC



(f) WC



(g) CP

Figure 3-13: LSTM Performance - 24 hour-ahead: Baseline (red) vs NWS (blue) vs Ground Truth (green) for all Seven Projects

Table 3.2: Summary of LSTM Performance - 24 hour-ahead

Project	**Capacity (MW)**	**MAE Base**	**MAE NWS**	**MAPE Base**	**MAPE NWS**	**Improvement (%)**
OS	150	40.82	29.79	27.21%	19.86%	7.35%
LC	150	32.92	20.71	21.95%	13.81%	8.14%
CB	50	10.67	6.65	21.34%	13.30%	8.04%
CP	50	9.4	7.00	18.80%	14.00%	4.80%
WC	238	67.67	51.96	28.43%	21.83%	6.60%
BR	57	9.96	7.04	17.47%	12.35%	5.12%
CC	235	68.20	56.39	29.02%	24.00%	5.03%

of the projects are the same at 150MW, and their data inputs and model structures are also identical, it is rather surprising that the MAPE can differ so significantly, as much as more than 6%. One possibility is that wind direction, which is not an input in the model, can have a large predictive value on wind power. A simple example is that if two equally sized wind farms are equal distance from a weather station, but one is situated to the east while the other is to the west of the weather station. If the wind blows towards the west, the prediction for the wind power plant to the west of the weather station will be much better, even if we control for every other factor. We may have encountered such a case for Project OS vs Project LC. Incorporating wind direction may be an extension of our model in the future.

Based on the reading of the graphs, compared to the base case, the prediction made with NWS wind forecast data is able to much more closely trace the ground truth, especially when there are sudden drops or peaks. For instance, in the graphs of Project CC and OS, the increasing or falling trends of wind generation is followed accurately by the NWS prediction, even though there is still a difference in absolute values. On the contrary, the baseline case shows an opposing trend to the ground truth: decreasing when the ground truth is increasing, and vice versa. This may suggest that the stochastic nature of wind speed can make it very difficult for a purely statistics-based method to predict wind power generation over a longer horizon. The information may have been better captured by an NWS weather forecast model that is physics-based. The improvement in performance as measured by MAPE averages around 6.44% for seven projects, with the biggest increase more than 8%. This improvement is significant, especially because the observed improvement from tuning the initial LSTM or trying variations of the LSTM was no bigger than 1%.

This experiment demonstrates that using solely wind data from the past is insufficient for predicting wind power generation, potentially due to LSTM's inability to forecast sudden wind speed changes that are not statistically correlated with the past and may only be captured by a physics- or simulation- based model such as those used by NWS to forecast weather data. It also shows that good data inputs play a significant role in accurate wind power forecasting and should probably receive more attention from both researchers and industry participants in the effort to improve predictability of wind power generation.

Chapter 4

Economic Benefits of Accurate Wind Power Forecasting

Chapter 3 answers the question of how wind power forecasting can be improved with better data using machine learning. This chapter aims to gain an understanding of how much economic benefit can be attributed to better wind power forecasting. Based on real market price data from SPP and MISO, I use a simple simulation to compare earnings under the base case forecasting versus the NWS forecasting, before extending to a broader discussion of the benefits of more accurate forecasting.

4.1 Analysis

4.1.1 Model and Assumptions

Recall that under the multi-settlement system discussed in Chapter 2, the earnings of an electricity seller would be equal to the following:

$$R = Q_{DA} * P_{DA} - (Q_{DA} - Q_{RT}) * (P_{RT}), \tag{4.1}$$

where Q stands for quantity of electricity sold at each hour and P is the price, measured as \$ per MW at different hours of the day. DA is for day-ahead and RT is for real-time. Note that Q_{RT} by definition is the actual power generation. If we make the important assumption that the plant owner will always bid a day ahead exactly the amount of power as the model predicts (in other words $Q_{DA} = Q'_{DA}$, where Q'_{DA} is the predicted generation), if P_{DA} and P_{RT} are known, we can calculate the earnings under a forecasting model and make comparisons of the earnings across different models.

Specifically, we can use Equation (4.1) to define earnings under the baseline forecast model and the improved NWS model:

$$R_{Base} = Q'_{Base_{DA}} * P_{DA} - (Q'_{Base_{DA}} - Q_{RT}) * (P_{RT}). \tag{4.2}$$

$$R_{NWS} = Q'_{NWS_{DA}} * P_{DA} - (Q'_{NWS_{DA}} - Q_{RT}) * (P_{RT}). \tag{4.3}$$

4.1.2 Price Data

Archives of P_{DA} and P_{RT} in different RTO territories is public information and can be downloaded from their respective websites. Since the seven power plants participate in SPP and MISO markets, we utilize the MISO and SPP market data [3][10]. Nodal prices within one RTO region can still differ significantly, so I rely on the SPP and MISO DA and RT data provided by the research sponsor/plant owner to make the calculation.

The price data are available in hourly intervals from January 2016 to January 2022. As discussed, many RT settlements are done in 5-min intervals, instead of hourly. However, for simplicity, in this simulation, we are going to assume that RT settlements are captured by the hourly prices.

4.2 Results and Discussion

While there are significant savings for more accurate forecasting in MISO, particularly for the larger projects, it is surprising to see that there are actual losses and some sizeable in SPP. However, it would be premature to conclude that more accu-

Table 4.1: The table summarizes R_{NWS}, R_{Base}, and $(R_{NWS} - R_{Base})$ for each project in the two respective markets.

Project	**Capacity /MW**	**MISO NWS/$**	**SPP NWS/$**	**MISO Base/$**	**SPP Base/$**	**Diff MISO/$**	**Diff SPP/$**
OS	150	4,697,130	5,789,517	4,457,849	6,195,229	239,281	-405,712
LC	150	5,809,732	6,563,803	5,608,228	6,661,708	201,504	-97,905
CB	50	631,187	678,829	599,253	720,412	31,934	-41,583
CP	50	630,279	677,671	595,837	716,224	34,442	-38,553
WC	238	4,104,923	4,138,045	3,779,358	4,328,679	325,565	-190,634
BR	57	910,785	1,075,534	887,102	1,079,992	23,683	-4,458
CC	235	5,005,334	5,782,456	4,880,094	5,801,721	125,240	-19,265

rate forecasting may not generate savings. This likely happens because of the price differences between P_{DA} and P_{RT} in different markets. One way to conceptualize this is that suppose the revenue under perfect forecasting is simply:

$$R_{perfect} = Q_{actual} * P_{DA}, \tag{4.4}$$

given that $Q_{actual} = Q'_{DA} = Q_{RT}$. Revenue under a forecast with errors is:

$$R_{forecast} = Q_{forecast} * P_{DA} - (Q_{forecast} - Q_{actual}) * (P_{RT}). \tag{4.5}$$

Difference in revenue or the cost of forecast errors is Equation (4.5) - Equation (4.4), which gives us:

$$R_{forecast} - R_{perfect} = (Q_{forecast} - Q_{actual}) * (P_{DA} - P_{RT}). \tag{4.6}$$

If we define forecast error as:

$$FE = (Q_{forecast} - Q_{actual}), \tag{4.7}$$

Then:

$$R_{NWS} - R_{Base} = FE * (P_{DA} - P_{RT}) \tag{4.8}$$

In the case of $R_{NWS} - R_{Base}$ in Table 4.1, even if FE stays constant, because P_{DA} and P_{RT} are very different in MISO and SPP, the results can be completely different, which may explain what we are seeing here. It would be interesting to explore the patterns of $(P_{DA} - P_{RT})$ in MISO and SPP respectively that may explain how the same forecasts yielded such different outcomes.

This observation also goes to show how a forecast number from a machine learning model is not sufficient for decision making. Our metrics from such forecasts are measured in MSE, MAE, or MAPE, which does not take into account of whether the prediction is over or under the actual generation, whereas the revenue from this trading process is precisely dependent on whether $Q_{forecast} - Q_{actual}$ is larger or smaller than zero. This might identify an area of research in the future for a better loss function and performance metrics that can help data users make better trading decisions. For instance, an industry market participant once remarked that a confidence interval may be more helpful than the number of MAE or MAPE. An output that takes into account of the probabilistic context will be helpful in this case.

I should also note that some of our assumptions are not realistic here. We assume that the plant owner will always bid a day ahead exactly the amount of power as the model predicts ($Q_{DA} = Q'_{DA}$). This is not realistic since a plant owner knows that there exists uncertainty in this forecasts and will likely adjust their trading strategies accordingly given their existing knowledge of the uncertainty and the DA and RT prices.

Additionally, we assume that DA and RT prices are known in this simulation. In reality, that is not how it works. In MISO, participants would have to put in their DA bids (with quantity and price) by 10:30 am Eastern Time in order to be considered for the DA market for the next day [3]. DA prices are then determined based on the bids. RT prices by definition are not known the day before, and are even more volatile compared with DA. This dynamic nature of the multi-settlement process for electricity trading further goes to highlight the importance of more accurate wind power forecasting – there are already significant uncertainties involved in the DA and RT prices, not knowing the quantity of generation further complicates the trading process and risk making the marketplace much less efficient.

Last but not the least, the methodology to assess the economic benefit of more accurate forecasting is not complete in just looking at the revenue under each forecast model. A more comprehensive assessment would also look at the risks. Because P_{RT} is much more volatile than P_{DA}, being exposed to the RT market can be very risky. In February 2021 in Texas, extreme weather caused extensive blackouts and caused some nodal prices to shoot to $9000/MW [13]. If a plant owner over-forecasts and oversells, in this case, they would have to make it up by buying RT generation. Depending on how much they have to buy in the RT market, this exposure can bankrupt them. Large risks like this are not factored in this study, but are very important to market participants and should be further studied.

Chapter 5

Predictability and Data Quality

The previous two chapters discuss how better data lead to more accurate forecasts and how more accurate forecasts may be beneficial economically. This chapter continues the discussion by exploring whether more accurate forecasts can be achieved by using data from weather stations closer to the wind power plants, which helps to inform the question of whether installing weather stations or meteorological towers closer to wind power plants may be a worthwhile investment.

Intuitively, a weather station closest to a wind power plant would have the highest predictive power on wind speed and other meteorological factors that would affect the power plant's generation. This is why many wind power plant developers or owners install meteorological towers, often temporary, right on the site of the wind power plants. The measurement data can verify the productivity of the wind power plant, which plays a significant role in forecasting the plant's lifelong earning. A meteorological tower's one- or two-year measurements can often help a wind power plant developer obtain a lower cost of financing and much better returns on the power plant [29]. However, maintaining a meteorological tower is expensive, and many plant owners would like to find out whether it is a worthwhile ongoing expense, particularly after the project is built and financing has already been obtained.

5.1 Data and Methodology

To study the question of whether a meteorological tower can meaningfully improve the predictability of a wind power plant, in the absence of actual onsite weather data, I rely on those from the weather stations of the nearest counties. The power plants relative to their nearby county weather stations maintained by NOAA are shown in Figure 5-1.

For each of the seven wind power plants, I measure the distance between the power plant and the weather station. This distance will be the independent variable. The values are recorded in Table 5.1. I then use the weather station data and fit them on the same LSTM baseline model discussed in Chapter 3 to make power generation forecasts. The generation and weather data are explained in Chapter 3. This performance will be the dependent variable. A total of 56 models will be trained and evaluated. I will then discuss whether there is a discernible pattern to the relationship between the distance (which serves as a proxy for data quality) and the predictability of wind power generation.

Table 5.1: Summary of the direct distance between a project and a county weather station in kilometers, rounded to the 2nd decimal place.

Project \Station	**Lyon /km**	**Buchanan /km**	**Johnson /km**	**Richardson /km**	**Fremont /km**	**Decatur /km**	**Shawnee /km**	**Douglass /km**
OS	24.60	170.55	197.41	179.86	255.13	299.96	73.69	90.61
LC	232.31	39.03	147.87	89.36	112.96	91.19	136.98	122.13
CB	243.15	90.08	231.35	43.56	36.99	135.33	152.19	160.96
CP	248.12	54.68	176.76	78.17	85.67	81.51	150.5	141.53
WC	245.23	61.65	198.07	58.86	60.98	97.65	148.85	147.00
BR	245.50	50.85	156.52	94.54	108.55	78.44	149.24	135.05
CC	255.63	71.39	205.50	66.83	55.33	91.20	160.59	157.97

5.2 Results

Table 5.2 summarizes the performance in MAE of the 56 models trained with the combination of the corresponding weather station data and the generation data for each power plant. To take away the effect of individual plant size on performance,

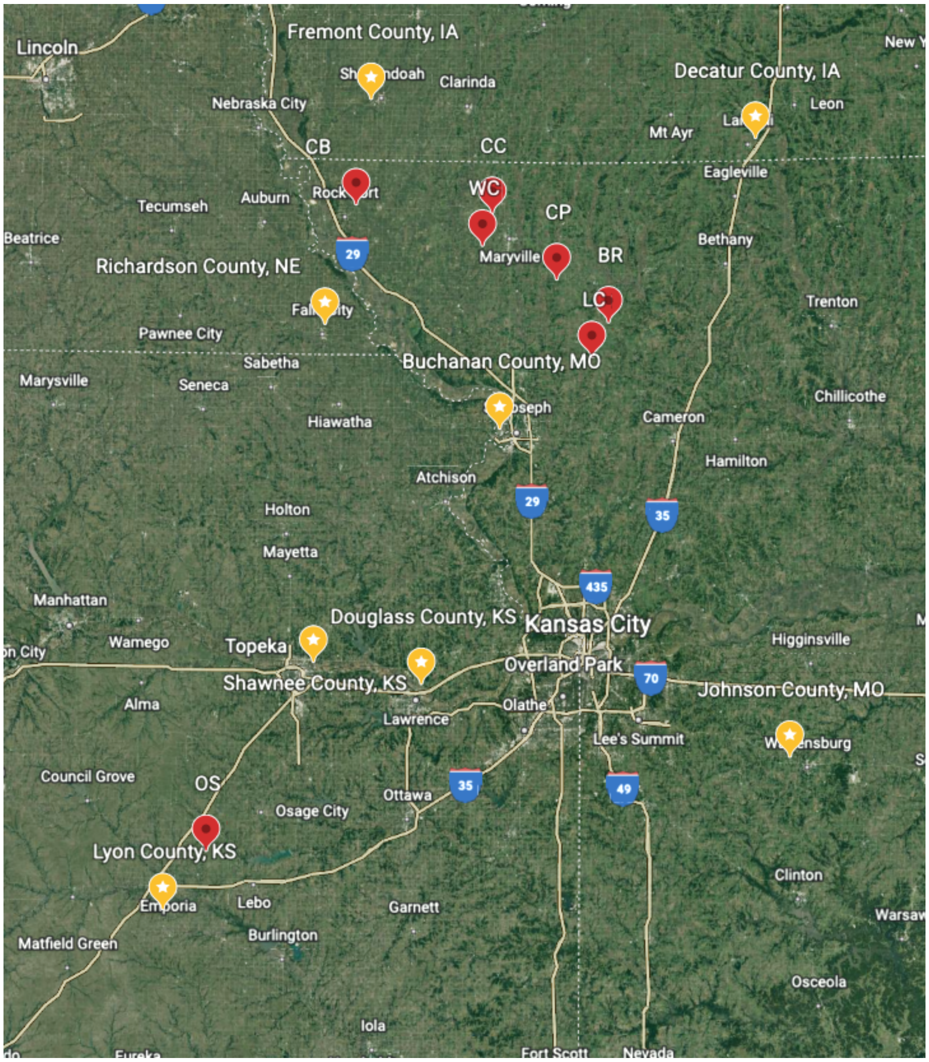


Figure 5-1: Approximate locations of the seven wind power plants(red) in relation to the eight nearest weather stations(orange).

Table 5.3 shows the same results but in MAPE. Figure 5-2 plots the results in Table 5.3 against the distance shown in Table 5.1 for each wind power plant with a line of best fit.

5.3 Discussion

Figure 5.2 shows that for six out of the seven wind projects, there is a positive relationship between MAPE and the distance. This follows our expectation that as distance increases, the performance of the predictive model deteriorates. However,

Table 5.2: Summary of the performance in MAE of the baseline LSTM model with data from respective weather station. The performance in bold indicates that the corresponding weather station is the closest to the wind power plant and mirrors the results in Table 3.2.

Project \ Station	**Lyon**	**Buchanan**	**Johnson**	**Richardson**	**Fremont**	**Decatur**	**Shawnee**	**Douglass**
OS	**40.82**	42.14	41.38	43.22	43.35	41.94	42.23	42.26
LC	34.37	**32.92**	34.10	34.74	35.44	35.07	33.60	33.61
CB	10.86	11.03	12.03	11.59	**10.67**	11.52	11.39	11.54
CP	10.13	**9.4**	10.01	10.37	10.36	10.56	10.09	10.16
WC	81.86	67.20	84.41	**67.67**	77.27	70.00	76.71	70.99
BR	9.85	**9.96**	11.11	10.32	10.44	10.77	10.73	10.58
CC	50.08	**60.42**	65.17	59.08	**68.20**	60.17	59.68	59.47

Table 5.3: Performance in MAPE

Project \Station	**Lyon**	**Buchanan**	**Johnson**	**Richardson**	**Fremont**	**Decatur**	**Shawnee**	**Douglass**
OS	27.21%	28.09%	27.59%	28.81%	28.90%	27.96%	28.15%	28.17%
LC	22.91%	21.95%	22.73%	23.16%	23.63%	23.38%	22.40%	22.41%
CB	21.72%	22.06%	24.06%	23.18%	21.34%	23.04%	22.78%	23.08%
CP	20.26%	18.80%	20.02%	20.74%	20.72%	21.12%	20.18%	20.32%
WC	34.39%	28.24%	35.47%	28.43%	32.47%	29.41%	32.23%	29.83%
BR	17.28%	17.47%	19.49%	18.11%	18.32%	18.89%	18.82%	18.56%
CC	21.31%	25.71%	27.73%	25.14%	29.02%	25.60%	25.40%	25.31%

with the exception of the project WC, the effects appear to be very small and we may not be able to rule out that these effects were not due to randomness.

In the case of Project CC, there seems to be a small negative correlation between distance and MAPE, which is a surprising. This reverse trend could be a result of an outlier: Lyon County, despite being very far away at more than 250km away, yielded the best performance. Further examination seems to confirm that the performance on Lyon County may be an outlier: the MAE on the validation data set is much higher than the test dataset, at more than 60, instead of 50. A trained model usually performs worse on the test data than the validation data, so this large drop in MAE may be the result of a fluke. It is also worth noting that data were only available for year 2020 to 2021 for Project CC, which leaves us about 1000 data points in the test data set. In machine learning, that may not be sufficient.

In either case, a more rigorous statistical model is needed for us to draw conclusions regarding the relationship between the distance of the power plant and data source

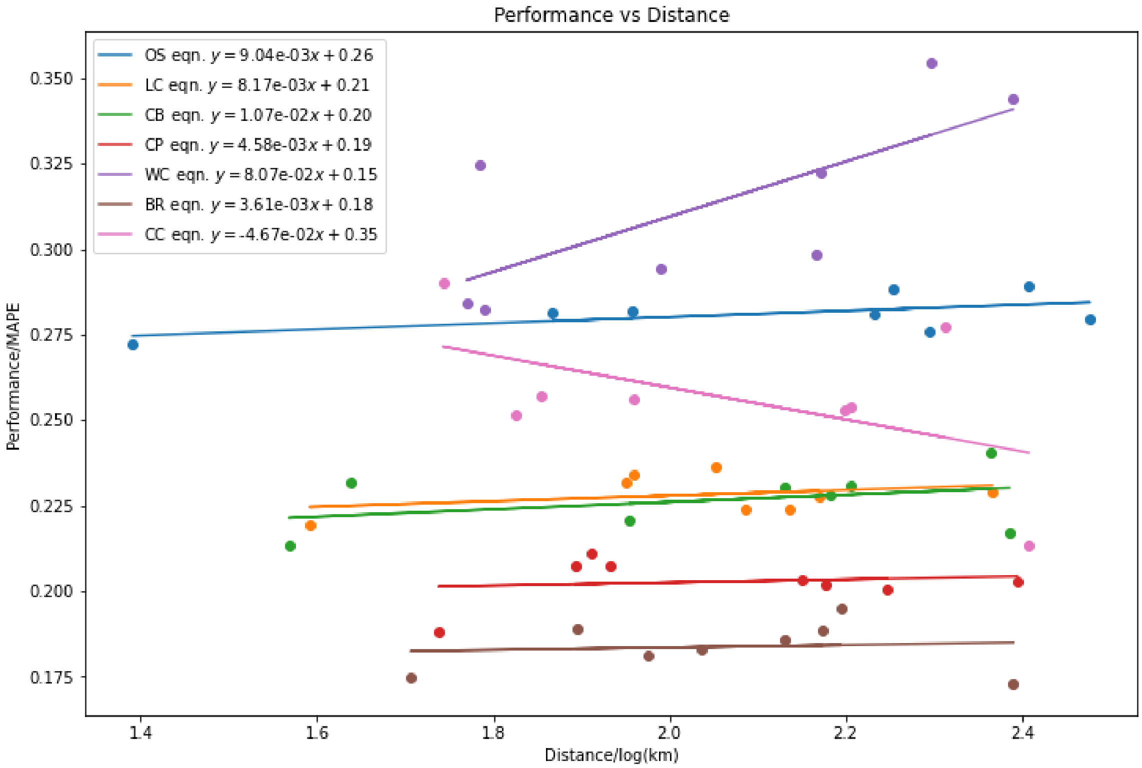


Figure 5-2: Distance in log vs Performance in MAPE

and the performance of the forecasting model.

One factor that may complicate the pattern here is the fact that only wind speed is used as an input to the models and wind direction is not accounted for here. We understand that if the wind blows from the East to the West and the power plant is situated to the East of a weather station, the data from the station may not have predictive power even no matter how close it is located to the power plant.

Another factor may be individual geospatial differences that are peculiar to each project. For instance, while the direct distance between a wind power plant and the weather station may not be large, but a mountain ridge may separate the project and the data source and the wind speeds at either location can be completely unrelated because of this local feature. It would be difficult for our model to capture such a factor.

In summary, while there seems to be a general positive correlation between distance and MAPE, the magnitude is small. More data points may be needed for a statistical framework to apply before we can draw conclusive conclusions. Two ways to improve this experiment in the future: First, incorporate wind direction as an input in the forecasting machine learning model; Second, use data from meteorological towers directly, if they can be obtained, instead of weather data from nearby stations as proxies for onsite data. That way, individual geospatial differences can be accounted for and wind direction may no longer be needed as an input.

Chapter 6

Conclusion

This thesis discusses three key questions:

1. How can wind power generation forecasting be improved with better data?
2. What are the economic benefits of more accurate forecasting of wind power generation?
3. Can prediction for wind power generation be improved with data sources closer to the site of generation?

We have clear answers for Question 1, and our understanding of Question 2 is now better illuminated through the analysis with empirical data. A preliminary answer exists for Question 3, which does show a generally positively correlated trend between MAPE and distance between power plants and weather stations, although the significance of this result needs further research. These findings are summarized in this chapter. I will also discuss the implications of the findings as well as the recommendations and directions for future work in the last section.

6.1 Summary of Key Findings

6.1.1 Wind Power Generation Forecasting

The first result that stands out is that machine learning models are the best solutions for short-term predictions. The next-step prediction for both the 5-min interval and the hourly predictions outperform traditional persistence or time series methods.

The second observation is that machine learning models perform a lot better, on average with a 6% lower MAPE, when NWS wind speed forecasts are added to the past measurement data of power generation and wind speed. This improvement is significant, as we rarely see a similar leap of performance through modifying model architectures alone. The finding is also important for industry participants who may not be able to build elaborate machine learning models for forecasting themselves but would have access to the publicly available NWS forecast data.

6.1.2 Economic Benefits of Accurate Wind Power Forecasting

It is generally understood that more accurate forecasts would be beneficial economically. However, through the analysis with empirical data, I find that the economic benefit would actually largely depend on the market itself, namely the differences between DA and RT prices, assuming trading strategies are simplified to always bidding what the forecast says. There can be significant savings in the MISO territory with up to $300,000 higher revenue per year, but there were also consistent losses in SPP territories with the same forecasts. The dynamic nature of the market with DA and RT prices unknown prior to bidding makes it difficult to pinpoint the exact economic benefit from the perspective of potential revenues. A helpful framework may be to conceptualize how more accurate forecasting simplifies the trading process: Given the complexity of the multi-settlement process for trading electricity, a cascade of uncertainties in both quantity (wind generation, electricity demand) and price (DA and RT prices) can risk making the trading process impossible to manage. One other way to evaluate the economic benefit of more accurate forecasting is through the idea of risks, as poor forecasting can expose a market participant to significant price risks of the RT market.

6.1.3 Predictability and Data Quality

The analysis shows that there exists a generally positively correlated trend between MAPE and distance between power plants and weather stations. As expected, the performance of machine learning models deteriorate as distance between the data source and the project site under evaluation increases. However, the magnitude of this trend is small and varies among the seven wind power plants, which suggests that there may be idiosyncratic factors that are not captured by the regression between distance and model performance or LSTM, the predictive model used itself.

6.2 Recommendations and Future Work

Wind farm owners should expect to have more accurate forecasting of wind power generation by switching to an RNN based machine learning model for near-term, next-step prediction for the RT market. If they currently have machine learning models to predict generation for the DA market with a longer forecasting horizon, they will benefit from accessing and incorporating NWS' publicly available wind speed forecasts that are updated hourly for forecasting horizons under three days.

At the same time, while NWS' wind speed forecasts are shown to be very useful, the performance of forecasting models may be further improved if NWS makes predictions for wind speeds that are in hourly intervals, instead of 3-hour intervals, as it is done currently. From the analysis on wind power generation, we see that generation can vary wildly from hour to hour: it can go from zero to full capacity in under 5 minutes. However, for meteorologists, weather forecasts are only made every 3 hours. This could be due to the long compute time needed for physics-based forecasting models, but if the forecasts can speed up to hourly, the performance of predictive model for wind power generation may further improve.

Electricity market participants may not be able to measure the economic value of more accurate forecasting by using the equation $R = Q_{DA}*P_{DA} - (Q_{DA} - Q_{RT})*(P_{RT})$, because whether the revenue is positive or negative will be entirely dependent on the differences between P_{DA} and P_{RT}. Instead, another metric to assess the benefits of more accurate forecasting is the reduced uncertainty in decision making in the trading process and the lowered risks due to more limited exposure in the more volatile RT market. A different economic model may be needed to comprehensively assess these benefits of accurate forecasting.

From the perspective of building predictive models that are more user-friendly to industry practitioners, it may be helpful for the model to output a confidence interval for each prediction, since a mean absolute error is insufficient. It matters for the bottom line for a trader to know whether a model tends to over- or under-forecast. An improvement on the model architecture may be to build a model that integrates both the forecasting of wind power plants and DA and RT electricity prices with a loss function that maximizes the revenue in Equation 4.1, although such efforts should be mindful of the complexity involved and the dynamic nature of the market, which may require the model to be frequently retrained, if not redeveloped.

Regarding the question of whether to invest in meteorological data for more accurate forecasting, while it is not untrue that the data from nearby sources perform better, it is not clear whether a wind farm owner can benefit significantly by buying proprietary data from, for instance, a weather station 50km away from a wind farm, especially if such data already exist from nearby, public weather stations. To better answer this question, wind direction should be included in the forecasting models in the future, given that it would help determine whether wind speeds at a weather station situated in a certain direction would have large predictive value for a wind farm.

It may be worthwhile to maintain data from a meteorological tower right onsite and compare them with those from a nearby weather station. The former, the onsite meteorological tower, will avoid issues such as wind direction and local geospatial features that may reduce the predictive power of the data from a public weather station. As we have seen in estimating the economic value of more accurate forecasting, particularly in the case of larger wind power plants, the annual savings can be more than \$300,000, which is still likely an underestimate, since we do not factor in the benefit of the reduced risk due to more limited exposure to the volatile RT market. If the annual cost of maintaining a meteorological tower is less than 10% of the potential annual savings, the effort is likely to be worthwhile. A study contrasting data from actual onsite meteorological tower with public weather station data from a distance will be a better experiment to more precisely answer the question of whether such investments are good decisions.

Bibliography

[1] ISO-New England. https://www.iso-ne.com/markets-operations/markets/da-rt-energy-markets/. Accessed: 2022-12-23.

[2] ISO-New England. https://www.iso-ne.com/isoexpress/web/charts. Accessed: 2022-12-22.

[3] MISO Market Data. https://www.misoenergy.org/markets-and-operations/#t=10&p=0&s=&sd=. Accessed: 2022-12-31.

[4] National Digital Forecast Database. https://www.ncei.noaa.gov/has/HAS.FileAppRouter?datasetname=9959_02&subqueryby=STATION&applname=&outdest=FILE. Accessed: 2022-12-28.

[5] National Weather Service. https://www.weather.gov/mob/forecast#. Accessed: 2022-12-28.

[6] NCEP WMO GRIB2 Documentation. https://www.nco.ncep.noaa.gov/pmb/docs/grib2/grib2_doc/. Accessed: 2022-12-28.

[7] NOAA, National Center for Environmental Information, Local Climatological Data. https://www.ncdc.noaa.gov/cdo-web/datatools/lcd. Accessed: 2022-12-28.

[8] PANDAS Documentation. https://pandas.pydata.org/docs/. Accessed: 2022-12-29.

[9] Recurrent Neural Networks cheatsheet. https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks. Accessed: 2022-12-27.

[10] SPP Market Data. https://marketplace.spp.org/groups/real-time-market-data. Accessed: 2022-12-31.

[11] TensorFlow LSTM Documentation. https://www.tensorflow.org/api_docs/python/tf/keras/layers/LSTM. Accessed: 2022-12-29.

[12] TensorFlow Time Series Forecasting. https://www.tensorflow.org/tutorials/structured_data/time_series. Accessed: 2022-12-29.

[13] Texas regulators keep prices near $9,000/MWh cap during rotating outages. https://www.spglobal.com/commodityinsights/en/market-insights/latest-news/natural-gas/021621-texas-regulators-keep-prices-near-9000mwh-cap-during-rotating-outages. Accessed: 2022-12-31.

[14] U.S. Energy Information Administration. https://www.eia.gov/energyexplained/electricity/electricity-in-the-us.php. Accessed: 2022-12-22.

[15] U.S. Energy Information Administration. https://www.eia.gov/todayinenergy/detail.php?id=790. Accessed: 2022-12-22.

[16] Wen-Yeau Chang et al. A literature review of wind forecasting methods. *Journal of Power and Energy Engineering*, 2(04):161, 2014.

[17] Federal Energy Regulatory Commission et al. Energy primer: A handbook of energy market basics. *Federal Energy Regulatory Commission: Washington, DC, USA*, 2015.

[18] Ramon Gomes da Silva, Matheus Henrique Dal Molin Ribeiro, Sinvaldo Rodrigues Moreno, Viviana Cocco Mariani, and Leandro dos Santos Coelho. A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting. *Energy*, 216:119174, 2021.

[19] Caroline Draxl, Andrew Clifton, Bri-Mathias Hodge, and Jim McCaa. The wind integration national dataset (wind) toolkit. *Applied Energy*, 151:355–366, 2015.

[20] Aoife M Foley, Paul G Leahy, Antonino Marvuglia, and Eamon J McKeogh. Current methods and advances in forecasting of wind power generation. *Renewable energy*, 37(1):1–8, 2012.

[21] Justin Heinermann and Oliver Kramer. Machine learning ensembles for wind power prediction. *Renewable Energy*, 89:671–679, 2016.

[22] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[23] Md Alamgir Hossain, Ripon K Chakrabortty, Sondoss Elsawah, Evan MacA Gray, and Michael J Ryan. Predicting wind power generation using hybrid deep learning with optimization. *IEEE Transactions on Applied Superconductivity*, 31(8):1–5, 2021.

[24] Kathrine Lau Jørgensen and Hamid Reza Shaker. Wind power forecasting using machine learning: State of the art, trends and challenges. In *2020 IEEE 8th International Conference on Smart Energy Grid Engineering (SEGE)*, pages 44–50. IEEE, 2020.

[25] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[26] Wen-Hui Lin, Ping Wang, Kuo-Ming Chao, Hsiao-Chung Lin, Zong-Yu Yang, and Yu-Huang Lai. Wind power forecasting with deep learning networks: Time-series forecasting. *Applied Sciences*, 11(21):10335, 2021.

[27] Hui Liu, Hong-Qi Tian, Chao Chen, and Yan-fei Li. A hybrid statistical method to predict wind speed and wind power. *Renewable energy*, 35(8):1857–1861, 2010.

[28] William P Mahoney, Keith Parks, Gerry Wiener, Yubao Liu, William L Myers, Juanzhen Sun, Luca Delle Monache, Thomas Hopson, David Johnson, and Sue Ellen Haupt. A wind power forecasting system to optimize grid integration. *IEEE Transactions on Sustainable Energy*, 3(4):670–682, 2012.

[29] Paul D Schwabe, David J Feldman, Donald E Settle, and Jason Fields. Wind energy finance in the united states: Current practice and opportunities. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2017.

[30] Heng-Yi Su and Chun-Rong Huang. Enhanced wind generation forecast using robust ensemble learning. *IEEE Transactions on Smart Grid*, 12(1):912–915, 2020.

[31] Huai-zhi Wang, Gang-qiang Li, Gui-bin Wang, Jian-chun Peng, Hui Jiang, and Yi-tao Liu. Deep learning based ensemble approach for probabilistic wind power forecasting. *Applied energy*, 188:56–70, 2017.

[32] Lin Wang, Rui Tao, Huanling Hu, and Yu-Rong Zeng. Effective wind power prediction using novel deep learning network: Stacked independently recurrent autoencoder. *Renewable Energy*, 164:642–655, 2021.

[33] Henning Wilms, Marco Cupelli, Antonello Monti, and Thorsten Gross. Exploiting spatio-temporal dependencies for rnn-based wind power forecasts. In *2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia)*, pages 921–926. IEEE, 2019.