

IMPROVING WIND POWER FORECASTS: INTEGRATING ADAPTIVE
HISTOGRAM OF ORIENTED GRADIENTS AND MULTIPLE NUMERICAL
WEATHER PREDICTIONS

by

İlayda Çelenk

B.S., Mathematics, Boğaziçi University, 2019

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Industrial Engineering
Boğaziçi University
2023

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to Assoc. Prof. Mustafa Gökçe Baydoğan for his invaluable and unwavering guidance throughout my master's studies. His endless patience and willingness to ensure my understanding have been truly helpful. I will carry the lessons I have learned from him throughout my life and career. I am truly fortunate to have him as my thesis advisor, and I will always cherish the positive impact he had on my education and academic career.

I am also grateful to the members of Department of Industrial Engineering for seeing me as a suitable master's student and a teaching assistant. I also would like to thank Prof. Mustafa Necati Aras, Assoc. Prof. Gönenc Yücel and Assist. Prof. Burak Barutçu for taking part in my thesis committee and making valuable comments.

I would like to thank my parents Yasemin and Ahmet, and my elder sister Tuba. Their steady confidence in me, unlimited support, and sacrifices have been the foundation upon which I built my dreams. I am also very grateful to have Lokum, she brings joy to my life and caring for her needs always gave me the time to freshen my ideas.

I also would like to thank my closest friends: Melis, for supporting me for 23 years; Gözde, for caring my needs; and Esra, for understanding me. I am grateful to all my friends who has always been one phone call away.

ABSTRACT

IMPROVING WIND POWER FORECASTS: INTEGRATING ADAPTIVE HISTOGRAM OF ORIENTED GRADIENTS AND MULTIPLE NUMERICAL WEATHER PREDICTIONS

In this research, alternative representations of the grid based Numerical Weather Prediction (NWP) models are proposed for wind power forecasting purposes. Wind speed is the major indicator for the power generation. However, wind direction has nonlinear effects coming from the wind's dynamic nature. With the traditional representations, the continuous and cyclic behavior of the wind direction is not perceived by the learners. To address these problems, Standard and Supervised transformation methods based on the Histogram of Oriented Gradients (HOG) of the NWP models are proposed. Our experiments on forty-seven wind farms show that the Standard HOG transformation on the naive features of multiple NWP models from distinct locations provides superior performance in linear learners. Additionally, Supervised HOG transformation is proposed with the utilization of tree based clustering methods. According to the results of the experiments with combined representations, linear learning methods outperform the tree based learners in wind power forecasting. This result indicates that the nonlinearity deriving from the wind direction's circular behavior is represented in a suitable form for the linear learners.

ÖZET

RÜZGAR GÜCÜ TAHMİNLERİİNİ İYİLEŞTİRME: YÖNLÜ GRADYANLARIN UYARLANABİLİR HİSTOGRAMI İLE ÇOKLU SAYISAL HAVA TAHMİNLERİNİN ENTEGRASYONU

Bu araştırmada, rüzgar enerjisi tahmini amacıyla grid tabanlı Sayısal Hava Tahmini (NWP) modellerinin alternatif temsilleri önerilmiştir. Rüzgar hızı, enerji üretimi için ana etkendir. Bununla birlikte, rüzgar yönü, rüzgarın dinamik yapısından kaynaklanan doğrusal olmayan etkilere de sahiptir. Geleneksel gösterimlerle rüzgar yönünün sürekli ve döngüsel davranışını öğrenme modelleri tarafından kavranmaz. Bu sorunları ele almak için, NWP modelleri üzerinde Yönlü Gradyanların Histogramı (HOG) tabanlı Standart ve Uyarlanabilir yöntemler önerilmiştir. Kırk yedi rüzgar santrali üzerinde yapılan deneyler, Standart HOG dönüşümünün farklı konumlardan çoklu NWP modellerinin özniteliklerine entegrasyonunun lineer öğrenme modellerinde üstün performans sağladığını göstermektedir. Buna ek olarak, ağaç tabanlı topluluk öğrenme yöntemlerine dayalı Uyarlanabilir HOG dönüşümü önerilmiştir. Birleştirilmiş dönüşümlerdeki deneylerin sonuçlarına göre, lineer öğrenme yöntemleri, özniteliklerle rüzgar gücünü tahmininde ağaç tabanlı öğrenme modellerine göre önemli ölçüde iyi performans göstermektedir. Bu sonuç, rüzgar yönünün döngüselliğinden kaynaklanan doğrusal olmamanın, lineer öğrenme metodlarına uygun bir şekilde temsil edildiğini göstermektedir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
ÖZET	v
LIST OF FIGURES	viii
LIST OF TABLES	xi
LIST OF SYMBOLS	xiii
LIST OF ACRONYMS/ABBREVIATIONS	xiv
1. INTRODUCTION	1
2. LITERATURE REVIEW	10
2.1. Wind Power Forecasting Methods	10
2.2. Feature Extraction Methods for Wind Power Forecasting	12
2.3. Feature Descriptor Histogram of Oriented Gradients	13
3. BACKGROUND	16
3.1. Data Description	16
3.2. Naive Methods	19
4. METHODOLOGY	21
4.1. Standard HOG Representation	25
4.2. Supervised HOG Representations by Tree Based Clusters	27
4.3. Combined Representations and Ensemble Modelling	35
5. EXPERIMENTS ON REAL DATA	37
5.1. Representations and Learners	37
5.2. Performance Metric	41
5.3. Results and Model Comparisons	42
5.3.1. Benefits of Combining Multiple NWPs and Outlier Elimination in Linear Models	42
5.3.2. Comparison of the Naive Models	44
5.3.3. Integrating Continuity and the Cyclic Behaviour of the Wind Direction	45
5.3.4. Location Embedded Representations by Supervised Approaches	47

5.3.5. Random Forest vs. GLM	49
5.3.6. Combined Representations vs. Ensemble Modelling	50
6. CONCLUSION	53
REFERENCES	55
APPENDIX A: WIND FARMS AND NWP INFORMATION	61
APPENDIX B: OVERALL RESULTS OF ALL WIND FARMS	65
APPENDIX C: VALIDATION OF REPRESENTATION PARAMETERS	83

LIST OF FIGURES

Figure 1.1.	GFS Power Curve.	4
Figure 1.2.	GFS Outliers.	5
Figure 1.3.	Prevailing Wind Direction and Wind Speed Relationship in GFS..	6
Figure 1.4.	Wind Direction, Wind Speed and Production Relationship in GFS. .	6
Figure 1.5.	HOG Example for a Picture.	8
Figure 2.1.	Simple HOG Transformation with Unit Length Gradient Vectors..	13
Figure 2.2.	HOG Transformation in Wind Components.	15
Figure 3.1.	NWP Locations of a WF.	16
Figure 3.2.	Components of the Wind.	17
Figure 3.3.	Power Curve at a GFS Location.	18
Figure 3.4.	Power Curve at the Top Correlated GFS Location.	19
Figure 4.1.	Aggregated- u - v vs. Production on GFS.	28
Figure 4.2.	Aggregated- u - v Clusters on GFS.	29
Figure 4.3.	Aggregated- u - v Clusters on GFS with Standard HOG Bins.	29

Figure 4.4. Regression Tree with Depth 3 on the Aggregated- <i>u-v</i> Features of GFS.	30
Figure 4.5. Aggregated- <i>w-d</i> Clusters on GFS.	30
Figure 4.6. Regression Tree with Depth 3 on the Aggregated- <i>w-d</i> Features of GFS.	31
Figure 4.7. Location-Based- <i>u-v</i> Clusters on GFS.	32
Figure 4.8. Regression Tree with Depth 3 on the Location-Based- <i>u-v</i> Features of GFS.	32
Figure 5.1. Wind Farm Locations Used in Experiments.	37
Figure 5.2. Average Ranks of the GLM Models with Single vs. Ensemble of NWPs.	43
Figure 5.3. Average Ranks of the GLM and Random Forest Models on Naive Representations.	44
Figure 5.4. Sensitivity Results for the Number of LHOG Bins.	46
Figure 5.5. Average Ranks of the Selected GLM Models.	46
Figure 5.6. Sensitivity Results for the Number of AHOG Bins.	47
Figure 5.7. Sensitivity Results for Depth and Number of HOG Bins in LuvHOG. .	48
Figure 5.8. Average Ranks of the GLM Models with Supervised Representations. .	49
Figure 5.9. Pairwise WMAPE Comparisons.	50

Figure 5.10. Average Ranks of the Ensemble Approaches. 51

Figure 5.11. Production Histogram for a Wind Farm. 51

Figure 5.12. Nemenyi Tests in Categories. 52

LIST OF TABLES

Table 2.1. Example Binning of HOG Transformation	14
Table 4.1. Before Aggregation	23
Table 4.2. Aggregated Features	23
Table 4.3. Example of HOG Transformation	26
Table 4.4. Example of Aggregated-HOG Bins	27
Table 4.5. Example of Location-Based-HOG Bins	27
Table 5.1. Naive Representations for a Single NWP Model	39
Table 5.2. Additional GLM Representations	40
Table 5.3. Ensemble Techniques	41
Table A.1. Wind Farms and the Number of NWP Coordinates	61
Table A.2. Time Period of Each NWP Model	63
Table B.1. WMAPEs of the GLM Models for Power Generation	65
Table B.2. WMAPEs of the GLM Models for Power Utilization with Aw	67
Table B.3. WMAPEs of the Random Forest Models	69
Table B.4. WMAPEs of the GLM Models	71

Table B.5. WMAPEs of Ensemble Approaches	73
Table B.6. WMAPEs of Ensemble Approaches on High Production	75
Table B.7. WMAPEs of Ensemble Approaches on Moderate Production	77
Table B.8. WMAPEs of Ensemble Approaches on Low Production	79
Table B.9. WMAPEs of Luv.RF on the Production Categories	81
Table C.1. AHOG Sensitivity Results on Training Set	83
Table C.2. LHOG Sensitivity Results on Training Set	85
Table C.3. LuvHOG Sensitivity Results on Training Set	87
Table C.4. LwdHOG Sensitivity Results on Training Set	89

LIST OF SYMBOLS

$d_{i,j,k}(t)$	Direction of the wind speed for i, j, k at time t
$d_{i,j,k^*}(t)$	Direction of the wind speed for i, j, k^* at time t
$\bar{d}_{i,k^*}(t)$	Aggregated d in $ J_i $ locations for i, j, k^* at time t
i	NWP model i
I	Set of NWP models
j	Location j
J_i	Set of locations of i
$ J_i $	Number of locations of i
k	Level k
K_i	Set of levels of i
k^*	Top correlated level
$p(t)$	Amount of power generation in megawatts at t
$p^*(t)$	Corrected amount of power generation at t
$\hat{p}^*(t)$	Forecast for $p^*(t)$
T	Common time interval of all NWP models i
T_o	Outlier eliminated time points
u	u component of wind vector in meter per second
$u_{i,j,k}(t)$	u component for i, j, k at time t
$u_{i,j,k^*}(t)$	u component for i, j, k^* at time t
$\bar{u}_{i,k^*}(t)$	average of u vectors in $ J_i $ locations for i, k^* at time t
$u^*(t)$	Power utilization at time t
v	v component of wind vector in meter per second
$v_{i,j,k}(t)$	v component of wind for i, j, k at time t
$v_{i,j,k^*}(t)$	v component for i, j, k^* at time t
$\bar{v}_{i,k^*}(t)$	Average of v vectors in $ J_i $ locations for i, k^* at time t
$w_{i,j,k}(t)$	Wind speed for i, j, k at time t
$w_{i,j,k^*}(t)$	Wind speed for i, j, k^* at time t
$\bar{w}_{i,k^*}(t)$	Aggregated w in $ J_i $ locations for i, k^* at time t

LIST OF ACRONYMS/ABBREVIATIONS

AHOG	Standard Histogram of Oriented Gradients with Aggregated Features
ARPEGE	Action de Recherche Petite Échelle Grande Échelle (Large-Scale Small-Scale Research Action)
<i>Auv</i> HOG	Supervised Histogram of Oriented Gradients with Aggregated <i>u-v</i> components
<i>Awd</i> HOG	Supervised Histogram of Oriented Gradients with Aggregated <i>w-d</i> components
GFS	Global Forecast System
GLM	Generalized Linear Models
HIRLAM	High Resolution Limited Area Model
HOG	Histogram of Oriented Gradients
LHOG	Standard Histogram of Oriented Gradients with Location Based Features
<i>Luv</i> HOG	Supervised Histogram of Oriented Gradients with Aggregated <i>u-v</i> components
<i>Lwd</i> HOG	Supervised Histogram of Oriented Gradients with Aggregated <i>w-d</i> components
NWP	Numerical Weather Forecast
RF	Random Forest
UGRD	North/South wind component
VGRD	East/West wind component
WF	Wind Farm
WMAPE	Weighted Mean Absolute Percentage Error

1. INTRODUCTION

Increased global energy demand creates bottlenecks in the supply and raises concerns about the energy crisis. At the non-renewable power plants such as coal and gas, production is controlled by the operators according to the consumption forecasts. On the other hand, the rapid changes in weather conditions affect the production at renewable power plants i.e., solar, hydro, and wind power plants. With the concerns about environmental pollution and limited non-renewable resources that are consumed faster than nature's speed of restoration, clean energy production solutions have become highly attractive.

Uncertainties in weather conditions create an imbalance between consumption and weather-dependent energy production. To address this problem, national and regional energy market operations are conducted to balance the generation and consumption in both short and long term, bringing energy suppliers and buyers together to ensure the citizens have access to reliable energy. National energy markets are organized and regulated within a single country, while regional energy markets encompass multiple countries in a particular geographic region. In general, the scope of an energy market is determined by the energy resources being traded, the infrastructure needed to distribute those resources, and the regulations of the market. The electricity grids are made up of a variety of small-scale and dispersed energy generation and consumption devices. To prevent possible infrastructure damages, electricity consumption and production in these grids should be at the same rate. Therefore, deviations in the forecasts of consumption and production lead to severe penalties in the balancing market. For the supplier, acquiring a consistent and accurate picture of the demand of their customer is essential for taking place in the Bilateral Market via long term contracts and trading at a fair price in the Day-Ahead Market in which the hourly price depends on the next day's supply and demand forecasts. To balance the supply and demand, tradings are conducted in the Intraday Market based on bidding. Lastly, to resolve the imbalance between actual supply and demand, tradings are conducted in the Balancing Market with severe penalties so that the aforesaid tradings and more accurate

forecasts are encouraged. As the renewable energy production gets more attractive, accurate forecasting becomes vital for the suppliers in the energy market. The local transparency platforms provide the actual production amounts of the power plants and the consumption levels. For example, Energy Exchange Istanbul(EXIST) [1] operates in Turkey, while the European Network of Transmission System Operators for Electricity (ENTSO-E) [2] operates in the Europe.

In case of renewable energy production from solar and wind, their non-dispatchable nature of the source does not allow adjusting the power output to match the demand as the resources are weather dependent. The weather conditions vary significantly from day to day, and the unstable weather conditions induce fluctuations in the amount of energy produced [3]. Forecasting the amount of renewable energy to be produced at a given time becomes a challenging task. Even though, the wind has become a highly preferred energy source due to its efficiency in clean energy production, considering its non-dispatchable character and the fluctuations in power production, satisfying the demand in the Day-Ahead Energy Market relies on accurate production forecasts.

Inherently, the wind power depends on weather conditions at the power plant location. Since the wind speed has a superior influence on the rotation speed of the wind turbines, it is the primary indicator of the energy generated. Although, measuring the wind speed at a plant site is valuable, the production is severely affected by the air density and the rapid changes in the weather conditions. In the decision making process of building a wind farm on a land, the actual wind speed information for a period of time is not available unless collected in advance. Although, the actual wind information determines the power generated, it is not sufficient for forecasting the production since it does not account for the future weather conditions.

Importance of the future weather conditions brings up the necessity to use Numerical Weather Prediction (NWP) models such as Global Forecast System (GFS), Action de Recherche Petite Échelle Grande Échelle (ARPEGE), High Resolution Limited Area Model (HIRLAM), and etc. GFS and ARPEGE are global models [4]. The former model divides the earth's surface into 20 kilometer grid squares horizontally,

and the atmosphere into 64 temporal layers vertically. It has a horizontal resolution of $0.25^\circ \times 0.25^\circ$ [5]. The latter uses triangular spectral truncation with 5 kilometers of horizontal resolution over France, 11 kilometers over Europe and approximately 7 kilometers over Turkey [6]. It has a horizontal resolution of $0.1^\circ \times 0.1^\circ$ over Turkey. These models are actively used for weather forecasting purposes.

NWP models are classified in two groups based on the prediction types, such as deterministic and probabilistic. While deterministic NWP models provide a point forecast for a particular time or interval, Ensemble Prediction Systems (EPS) create multiple forecasts with probabilities based on varying initial weather conditions, and accounting the uncertainty in a forecast [7–9]. For example, Global Ensemble Forecast System (GEFS) is developed based on multiple simulations of GFS and it provides multiple forecasts. Also, some NWP models perform better at predicting certain types of weather events [4]. Hence, employing multiple NWP models at alternative locations enclosing the wind farm is helpful to express the weather conditions.

In cases that the historical wind speed information is not available, NWP models provide adequate information to describe the relationship between the wind speed and the amount of electrical power that can be generated by a turbine. power curves are widely used. A power curve is a function of the wind speed and it is generally used for modelling this relationship [10]. The power curve of a wind turbine shows the amount of electricity in megawatts or kilowatts generated for different wind speeds and it is possible to predict the energy production based on such curves. In line with this, based on the wind speed forecasts from multiple locations enclosing the wind farm, an estimate of the power curve is shown Figure 1.1. Since the actual wind speed information at the rotor level of the turbine is not available, this power curve is based on the weather forecast data. As seen in Figure 1.1, high production occurs at a wind plant where the wind speed is large and the black curve represents an estimate of the power curve based on the average wind speed of multiple locations. In this estimate the polynomial wind speed terms up to degree three are used.

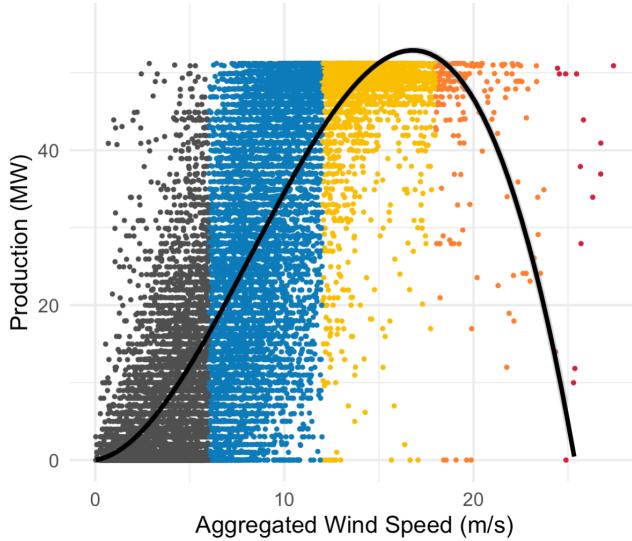


Figure 1.1. GFS Power Curve.

There are several reasons behind the fact that the power curve in Figure 1.1 is not a perfect fit for the actual power curve and it does not represent the exact relationship between the wind speed and the power production. One obvious reason is employing the aggregated wind speed forecasts from several locations instead of the actual wind speed at the wind farm. The second reason is based on the operations on site and the equipment. Scheduled turbine shut downs may occur due to maintenance or to prevent unnecessary power flowing to the grids. When the rotors reach to the maximum speed or on extreme weather conditions, turbines are also shut down to avoid possible damages. Such cases causes the power generated to diminish, even though the wind speed is high.

Site operations and data transmission failures may create outlier points that do not follow the expected relationship between production and wind speed. These points do not represent the actual relationship that is desired to be discovered in wind power prediction. There are approaches detecting the outliers based on the actual wind speed recorded at a turbine's rotor level [11, 12]. In the case of the absence of the actual wind speed, NWP data sets are functional. For a wind farm, Production-Wind Speed scatter plot at the top correlated GFS location in Figure 1.2 illustrates the outliers with A, B and C labels. At the points in Group A, wind speed is very low but the power

generation is not zero. For the actual wind speed information at a turbine's rotor level, Group A would represent data transmission failure. Using information from the top correlated NWP location may create additional source of variation in estimating the power curve. In any case, these outlier points are expected to be detected in Group A. Another outlier type referring to power plant shutdown is labeled as Group B. At these points there is no production recorded despite high wind speed. Other outliers are detected in Group C, representing the wind curtailment and maintenance. It is crucial to detect and treat these outliers to discover the actual relationship between production and wind speed.

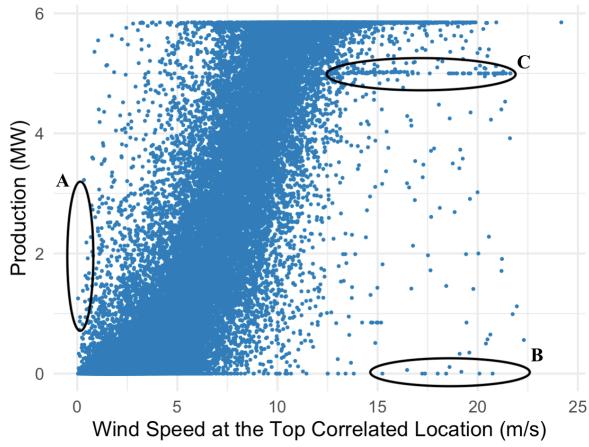


Figure 1.2. GFS Outliers.

Wind direction also has significant influence on the power generated in a wind farm. Prevailing wind direction is the primary factor determining the position of a turbine at the selected location. To integrate the wind direction into the wind power prediction problem, its spatial correlation characteristics together with the wind speed are utilized in the literature [13]. As the wind direction values lie in $[0, 360)$ degree range, some arrangements are needed to shift the degree axis based on the data [14]. To understand the link between the wind speed and direction, the wind rose plot is useful. The wind rose plot of a plant in Figure 1.3(a) shows from which directions the wind is blowing with respect to the aggregated wind speed. It indicates that the wind blows from the southwest to northeast almost 30% of the time, and from the southwest 12% of the time. The strongest winds are along the dominant or prevailing wind direction. This direction is also detected in Figure 1.3(b) where the aggregated

wind speed is high.

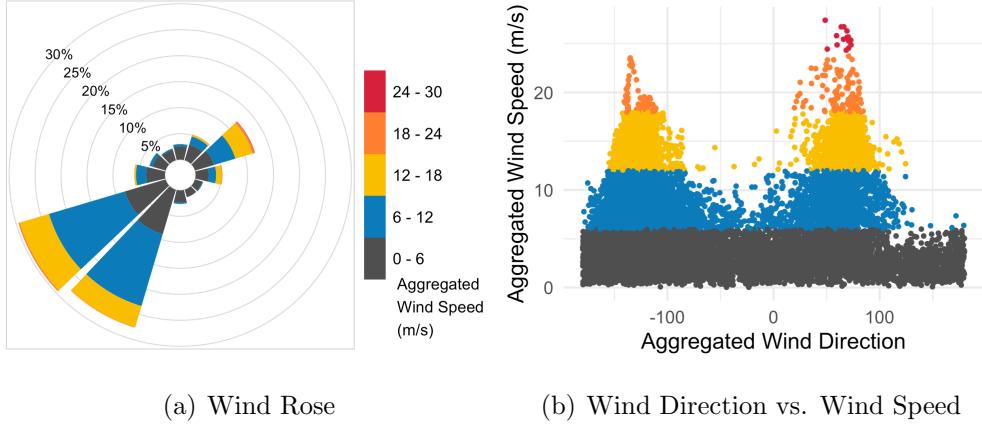


Figure 1.3. Prevailing Wind Direction and Wind Speed Relationship in GFS.

Most of the power generation occurs when the wind blows along the prevailing wind direction at different wind speed values as shown in Figure 1.4(a). Even at the times with lower wind speed, the production may be significantly high. Moreover, analyzing the prevailing wind direction with respect to the the power generation, Figure 1.4(b) indicates that a significant part of the high electricity generation occurs when the wind blows closely to the prevailing direction. Ideally, the dominant wind directions are expected to be similar in each NWP location covering the wind farm. However, there are significant differences in locations due to the physical properties of the land. For example, a nearby mountain may not be affecting the wind farm location but it may affect NWP at a location considered. As a result, the aggregated wind speed and direction may not represent the reality properly.

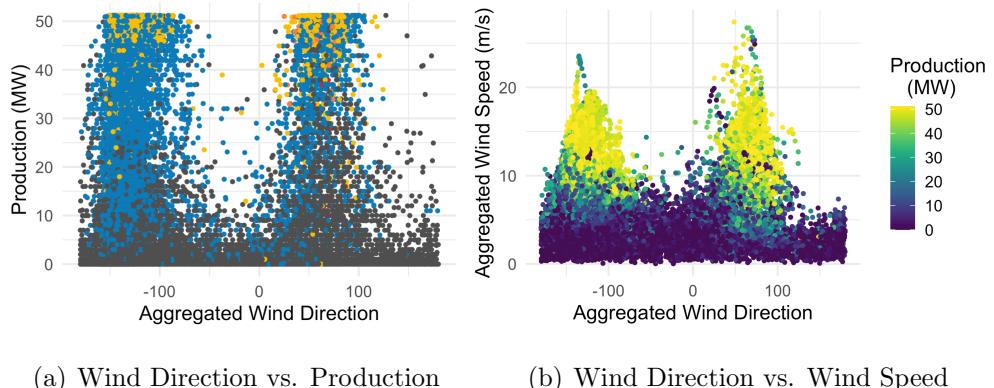


Figure 1.4. Wind Direction, Wind Speed and Production Relationship in GFS.

In wind power prediction, the nonlinear relationship between wind speed and direction is indicated as a vital information. Unlike the wind speed, the wind direction is not continuous. Therefore, integrating the wind direction to the linear models is a challenging aspect and it requires a transformation. As shown in Figure 1.3(b), wind direction has a discontinuity along the semi-infinite line $x < 0$ and $y = 0$, since $-180 + \epsilon$ and 180 degrees are considered as extremely different angles for a very small number ϵ . Therefore, using wind speed and direction as inputs to the wind power forecasting model may lack identifying the cyclic nature of the direction. Partitioning the direction interval into a sufficiently large number of bins and adding the corresponding wind speed values to these bins would create features representing the continuity of the wind direction. However, this approach would increase the complexity of the problem. This brings a need for a new representation addressing the problems with wind direction.

A transformation which maintains the continuity over the features with respect to the cyclic nature of the wind directions is a necessity. The Histogram of Oriented Gradients (HOG) method maintains the cyclic behavior of the direction and reduces the problem regarding discontinuities along coordinates. The main idea behind the HOG method is described in a patent application in 1986, without mentioning the exact terminology [15]. Orientation Histograms method was proposed as a feature descriptor for hand gesture recognition in static images, and its application has proven a significant improvement for gesture classification [16]. After the implementation of HOG descriptors for pedestrian detection [17], the method became widespread for image recognition.

To demonstrate an example of how the HOG features are extracted for human detection purpose, a picture in Figure 1.5(a) is used. In this example, HOG descriptors using a gradient range $(0, 180)$ are extracted. First, the cell size and the number of HOG bins are set to 12 and 9, respectively. The original image shape is $(520, 558, 3)$, and the gray scale image in Figure 1.5(a) shape is $(520, 558)$. To obtain the cells with size 12, the x-axis of the image is divided into $\lfloor \frac{558}{12} \rfloor = 46$ cells. Similarly, the y-axis of the image is divided into $\lfloor \frac{520}{12} \rfloor = 43$ cells. As a result, the HOG descriptor covers the picture pixels with a window of size $(552, 516)$ and the black frame in Figure 1.5(b)

corresponds to the discarded area. For each pixel in a cell, gradients are computed using the nearby cells and then its average gradient is computed. From the average gradients, the vector position and magnitude are extracted. This procedure is conducted for each cell and it provides Figure 1.5(b). A specific cell is shown in Figure 1.5(c) to illustrate the relationship with its adjacent cells. For this cell, the 9 HOG bins covering the $(0, 180]$ interval are shown in Figure 1.5(d). Note that this is a version of HOG and for the wind power prediction problem the vector positions and magnitudes are already available. Therefore, instead of taking the average gradients reflecting the relationship between points, weighted distribution is suitable in this case.

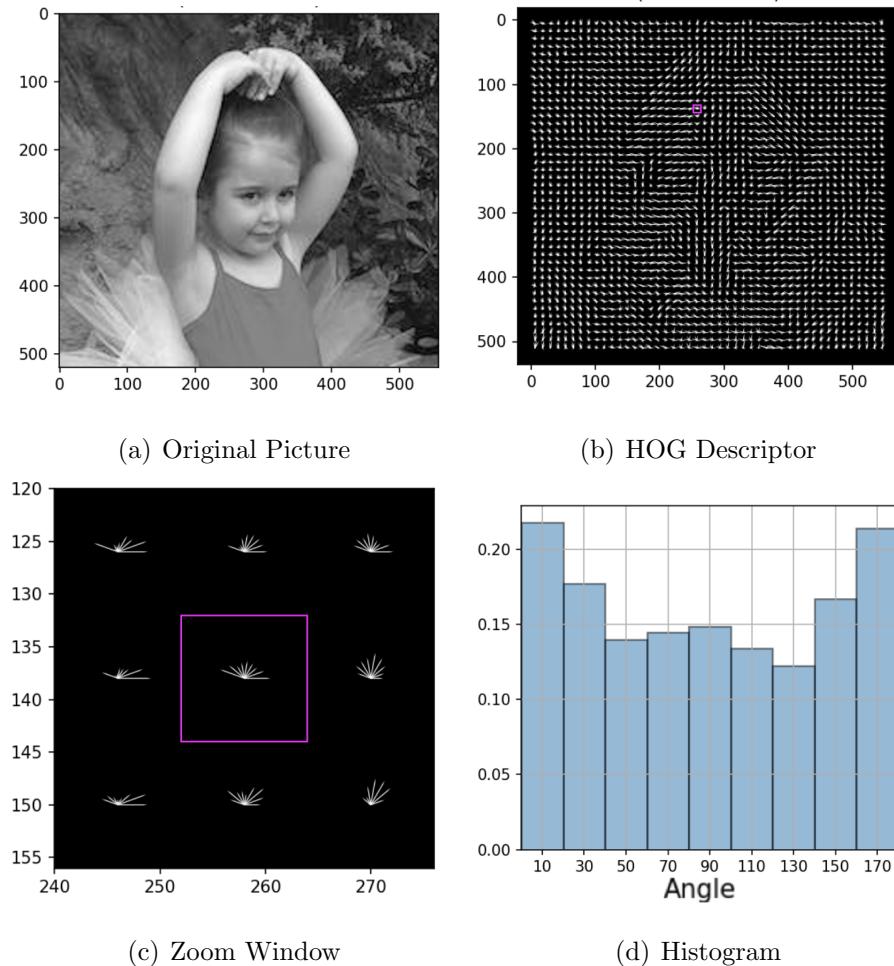


Figure 1.5. HOG Example for a Picture.

For the wind power forecasting, adapting the Standard HOG transformation on the range $(-180, 180]$, the continuity over the extracted features is maintained and the cyclic behavior is addressed on the boundaries. Moreover, Supervised HOG transfor-

mations are conducted after employing tree based clusters of data points, to uncover the embedded information in the wind speed components. Standard and Supervised HOG features over multiple NWP models are claimed to hold valuable information related to the dynamics of the wind. In this thesis, alternative feature extraction methods, Standard and Supervised HOG, are proposed with alternative model ensembling options. To test these claims, outliers are eliminated as a pre-processing step and multiple NWP models are combined. Then, to evaluate their performances linear and nonlinear learners are utilized and comparisons are made with the pre-existing approaches.

This thesis is organized as follows: In Section 2, a detailed literature review on wind power forecasting and feature extraction methods are provided. Section 3 provides background information on the data sets used in thesis, including the numerical weather prediction models, and some naive methods with the base components. In Section 4, a detailed explanation for the suggested representations is presented. In Section 5, the learners used in this thesis and the performance metric are introduced then the overall results with their comparisons are presented, including the parameter selection method. Section 6 provides the conclusion.

2. LITERATURE REVIEW

2.1. Wind Power Forecasting Methods

To forecast wind production, the first step is to analyze the components affecting the power generated. These components depend on the turbine properties and the characteristics of the weather. Size of the turbine rotors is a constant variable affecting the power generated. On the other hand, combining the kinetic energy and power formulations, the theoretical power P that can be harvested is proportional to the cube of the actual wind speed v at a turbine's rotor level [18] and it is formulated as:

$$P = \frac{1}{2}CA\rho v^3, \quad (2.1)$$

where C is the wind turbine factor, A is the area swept out by the rotor and ρ is the air density. The cube of the actual wind speed with the other defined variables in the formula represents the theoretical power curve of the turbine. The theoretical formula indicates that the fluctuations in the wind power mainly depends on the wind speed, but their relationship is not linear. Based on the formulated relationship, polynomial regression models with up to third-order wind speed information are suggested as simple and reliable linear methods [19].

Forecasting methods for the wind power can be divided into two main categories depending on involving the historical wind data as an input or the NWP data [4]. In the first category, statistical approaches are employed to forecast power production directly or via forecasts of hourly wind speed. In the other category, features are extracted from NWP models to forecast N-steps ahead. The Wind Power Prediction Tool (WPPT), developed by the Technical University of Denmark's Informatics and Mathematical Modelling Department, applies a statistical approach [20]. Such methods work well in estimating monthly mean wind speed. Prediktor is considered as a physical model developed by *Landberg* to model the wind power curve based on HIRLAM data [21]. Accounting the dynamics of the atmospheric events, these models provide sufficient results in the short-term horizons such as hourly or daily.

Various techniques are employed in the literature for the wind power forecasting purpose and to reduce the forecasting error [4], [22]. These methods benefit the spatial and temporal relationships of the data points. Forecasts based on the actual weather information and NWP are conducted in the literature. The variability of the weather is a challenging part of the power prediction methods and reliable wind speed information leads accurate wind power forecasts. The weather conditions at the plant site can be represented by a single type of NWP at several distinct nearby locations and some of the wind power forecasting methods utilizes a single NWP model [23]. Since the accuracy of the NWP models used has a major impact on the wind power forecasting, using a combination of different NWP models reduces the uncertainties of the weather and it performs significantly better with extreme weather conditions [22, 24]. An ensemble of NWP models benefits from each model's specialty in diverse weather conditions, therefore integration of NWP models provide more reliable forecasts [25]. Post-processing methods are also utilized in the literature. Kalman filtering algorithm is adopted to eliminate the systematic forecast errors in wind speed forecasts of NWP models provided that the weather conditions form a dynamical system [22], [26].

When the past wind speed information is not available at a wind farm location, a possible approach to estimate the actual wind speed is to employ the NWP information from multiple surrounding locations. NWP models are also useful for integrating the future weather events with the power generated. Statistical downscaling is an accepted approach to create more descriptive features for the precise weather conditions based on the information from multiple NWP locations. These features are often generated based on the physical aspects of the wind and the results show that their utilization can improve the wind power forecasting performance [22]. Spatial interpolation of the wind speed forecasts from multiple grid points is an option for such purposes [27].

To reduce the forecasting error, some methods focus on the reliability of the features, and put emphasis on the outlier detection and treatment as well. In the literature, several approaches are employed to remove the outliers in the wind power prediction setting. K-nearest neighbor based [28,29] and correlation-based classification algorithms [30] are proposed in the literature to detect the outliers. In most of the

methods outliers are removed [28, 29], while in some of them they are recovered [30]. Quantile regression is also proposed to detect outliers caused by shutdowns and wind curtailment [11].

2.2. Feature Extraction Methods for Wind Power Forecasting

Feature extraction methods have been employed on NWP data sets to obtain useful time and weather related features in the wind power forecasting [31, 32]. For example, transforming the powers of wind speed and direction into a single feature and then addressing it in categories of direction intervals and 6-hour time frames lead significant improvement [33]. In this study, the direction interval is partitioned and the wind speed values are distributed over these bins with respect to their corresponding directions. Although this feature representation addresses the wind direction, it is conducted in a discrete manner, which means that the continuity and circularity of the direction is not taken into account. With the addition of temporal relationships, this representation yields accurate results.

For the temporal relationship within a feature, lags and leads are often used in the time series forecasting setting to create auto-regressive models and reflect the changes in time [34, 35]. It is shown that the temporal variance is high in the periods with high power variability, furthermore its utilization reduces the forecasting error [35]. To extract the spatial information among features and reduce dimension, Principal Component Analysis (PCA) is employed with a preset variance threshold [35]. Since PCA is based on linear transformations, its utilization is more meaningful for the linearly correlated features. Spatial standard deviation and spatial smoothing effects are also discussed in the literature [22], [35–37]. Suggested models in the literature often combine spatial and temporal information. In the literature, three-dimensional convolutional neural networks are also utilized to extract the spatio-temporal features from NWP data sets [38].

This section summarizes the basic wind power forecasting and feature extraction methods are explained. Comprehensive studies analyzing more methods are available

in the literature [3, 4], [22]. The majority of these methods neglect continuity, and the extracted features are inadequate to use the linear learners since they do not represent the nonlinear relationship at the feature level. The focus of this thesis is to obtain a meaningful and efficient feature extraction method suitable to the characteristics of the wind vectors in the NWP models.

2.3. Feature Descriptor Histogram of Oriented Gradients

HOG features are often used in computer vision and image processing for the purpose of object detection in classification setting. First, the image is resized so that its pixel matrix is divisible into $s \times s$ blocks. For each pixel, the gradient vectors has two components. Each gradient component is calculated by the differences of the preceding and succeeding pixels with respect to the component. After obtaining the pixel gradients, gradient magnitude and angle matrices are obtained based on the vector magnitude and angle formulas. Then these matrices are divided into cells to form the blocks. Finally, for each block the gradient magnitudes are distributed over a predetermined number of bins representing the gradient direction. For n bins, each $s \times s$ block is represented by a vector of $1 \times n$ by distributing the magnitudes over the bins based on the weighted voting principle [39]. Consider a block of 6×6 pixels with unit gradient vectors in Figure 2.1(a). For this cell there are 7 different orientation bins and their histogram represents the 1×7 features as shown in Figure 2.1(b).

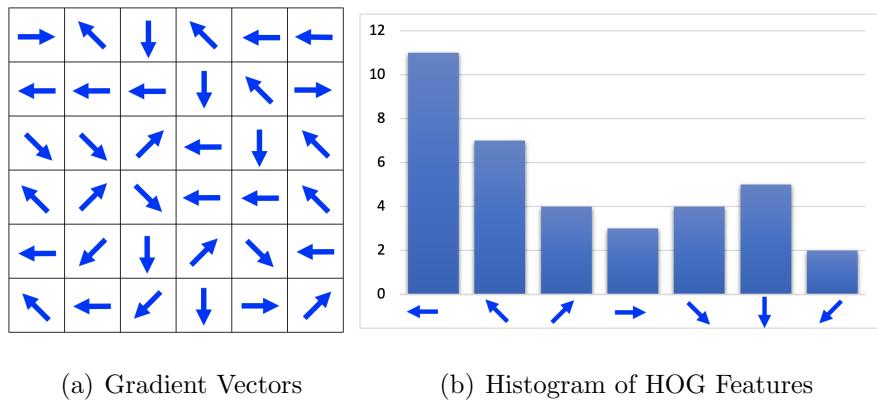


Figure 2.1. Simple HOG Transformation with Unit Length Gradient Vectors.

For the wind power forecasting setting, wind components in different altitudes at multiple locations are available in NWP data sets. These wind components are treated as the gradient components, from which the magnitudes and directions are computed. For a simple demonstration of HOG transformation, suppose the magnitudes and direction angles in Table 2.1 are given and the interval of directions is predefined as $(-180, 180]$ degrees.

Table 2.1. Example Binning of HOG Transformation.

		Angles:	180	-120	-60	0	60	120
ID	Magnitude	Direction	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6
1	15	180	15	0	0	0	0	0
2	24	-160	16	8	0	0	0	0
3	12	-80	0	4	8	0	0	0
4	20	150	10	0	0	0	0	10
5	15	-30	0	0	7.5	7.5	0	0
6	30	100	0	0	0	0	10	20

Suppose the number of bins are set to 6, which means partitioning the space into 6 intervals of 60 degrees as illustrated in Figure 2.2(a). In the 6–HOG transformation, bins represent 6 angles of -120, -60, 0, 60, 120, 180, and for the cases where the direction is between two angle bins the magnitude is proportionally distributed, assuring continuity over the transformed features. The values for the rest of the bins are zero. Treating -180 and 180 as the same, any direction in $(-120, -180)$ is distributed over Bin 6 and Bin 1 as shown in Table 2.1 which assures the cyclic behavior. After this transformation, the HOG features can be aggregated according to the other properties of the data points, which are not available in the table. Aggregating all the given data points results in 6-HOG features as shown in Figure 2.2(b). Note that, this transformation labels the features with vector directions, it does not use the interval partitions directly. And, when the direction of a point is between two labels, its corresponding wind speed is mapped to two adjacent bins.

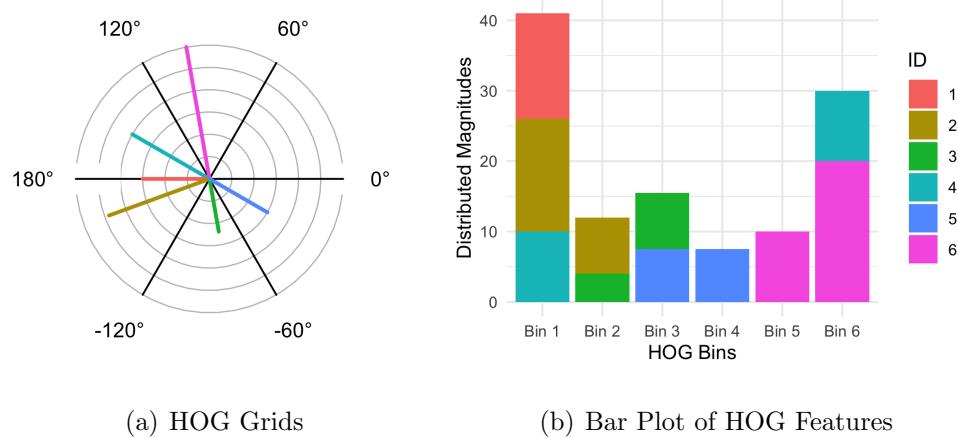


Figure 2.2. HOG Transformation in Wind Components.

As illustrated, the HOG transformation assures the continuity over the features involving magnitudes and represents the cyclic behavior at degree 180 by connecting the boundaries. The number of bins is a representation parameter determining the intervals among bins. When it is small, the intervals get wider and a wind speed value between two bins has partial contributions. These characteristics of the bins are sensitive to the weather conditions at the wind farm location and the sensitivity analysis regarding the optimal number of bins is provided in the following sections.

In this thesis, HOG transformation is implemented on the multiple NWP models in order to forecast the wind power generated. By this representation, the continuity of the wind direction is integrated with the wind speed information at multiple locations. In comparative with partitioning the direction interval, HOG features address the continuity over the features and the cyclic behavior on the boundaries.

3. BACKGROUND

In this section, the data sets used in the suggested wind power prediction models are explained. Then, the advantages and disadvantages of some naive prediction models will be discussed.

3.1. Data Description

In this thesis, instead of the actual wind speed information, the forecasts are obtained from multiple deterministic NWP models. Two NWP models are used, namely GFS and ARPEGE. The former is developed by the United States' National Weather Service (NWS) [5] and the latter is developed by Meteo France [6]. The forecasts from both of the models are available online [40, 41].

The forecasts are available at 16 and 25 locations enclosing the wind farm, respectively. NWP locations for a wind farm is shown in Figure 3.1.

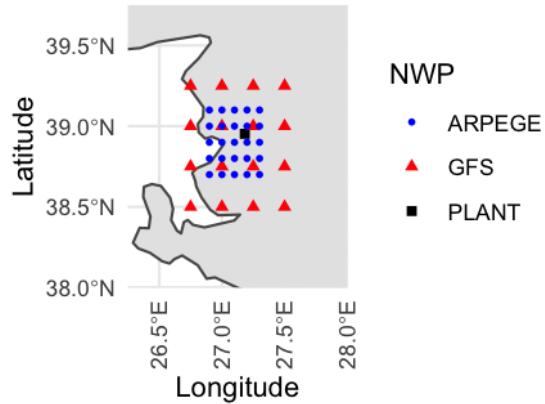


Figure 3.1. NWP Locations of a WF.

Each forecast data set provides wind components UGRD (u) and VGRD (v) in meter per second at several altitudes. The wind component u corresponds to the East-West direction, and being positive means that the wind is coming from the West. The wind component v corresponds to the North-South direction, and positive u means

that the wind is coming from the South. An example of a wind vector is shown in Figure 3.2, where the wind is blowing from South-West and its direction is towards North-East since both of the u and v components are positive.

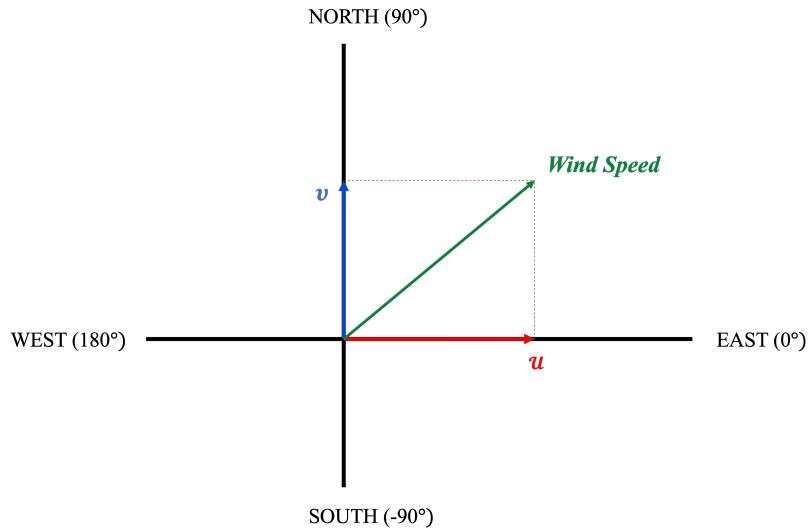


Figure 3.2. Components of the Wind.

Weather forecasting data sets provide information for one or three hour periods depending on the forecast period and NWP model. For example, GFS forecast step varies in 1 hour and 3 hours. Since GFS provide 3-hour forecasts for some time periods, missing hours in such models are filled by linear interpolation in order to prevent information loss in the ensemble of NWP approach where multiple NWP models are combined and remove the concerns related to comparability for single NWP based predictions.

Power production data set of a wind farm provides the actual amount of electricity generated in megawatts hourly. There are distortions in the production that are generally caused by maintenance, manual shutdowns, and the new turbine installations. Production capacity depends on the number of turbines as well as the wind speed. For instance, installation of a new turbine increases the capacity of the wind farm. By these distortions that are independent of the weather conditions, wind power forecasting models are not able to learn the actual possible production. To remove such distortions in high production, wind farm capacity is calculated by 30 days rolling maximum production, then the production above 99% percentile of the respective 30 days

is mapped to the rolling capacity. With this adjustment, it is possible to predict actual possible production hourly. This modification also enables finding the hourly utilization rate where the total capacity of the farm is the maximum production amount. Then, the hourly utilization rate is found by dividing the hourly production by the total farm capacity.

Generally, production amount is correlated with the cube of the wind speed, and production is expected to be low when the wind is weak. To convey an outlier free learning and forecasting, quantile regression curves are utilized. The behavior of these curves differ in each location of an NWP model. At a location the power curve may not represent the true relationship at the plant site as shown in Figure 3.4. In this location, production and wind speed do not follow an explicit relationship. This also indicates that the aggregated wind speed values are not suitable for representing the power curve, since the aggregation is influenced by such locations.

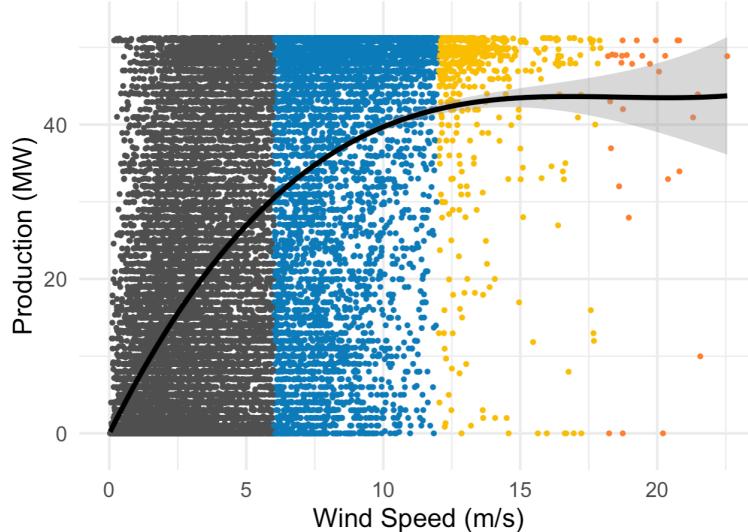


Figure 3.3. Power Curve at a GFS Location.

To represent the power curve based on the NWP models, the location in which the wind speed values are highly correlated with the production is suitable. The thick curve in Figure 3.4 is the power curve at the top correlated GFS location. The two other polynomial regression curves are fit at 1% and 99% percentiles. The outlier points of high production occurrences in low wind speed are above the upper curve. And, the

points of low production with high wind speed are below the lower curve. Removing these points produces data sets that are independent of the operating decisions. Therefore, for each wind farm and NWP model, the two percentile regression curves are fit based on the top correlated location. Data points below the 1% and above the 99% percentile regression curves are detected as outliers and removed from the NWP data set.

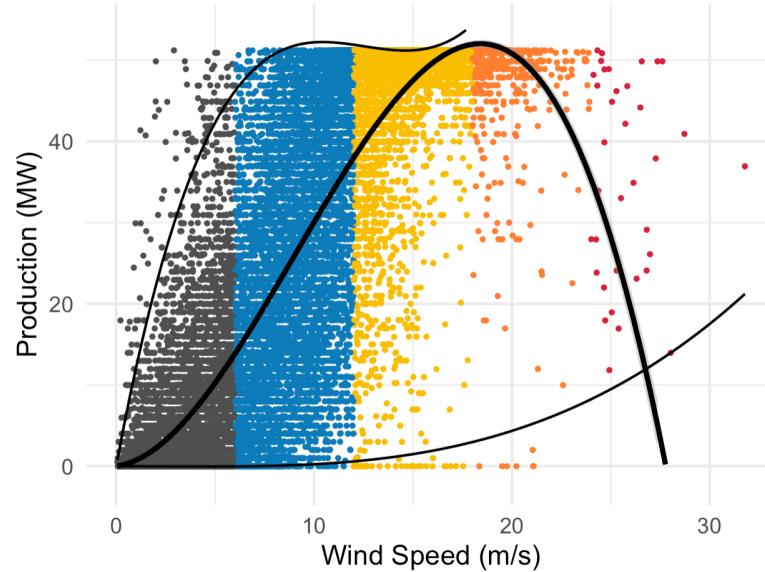


Figure 3.4. Power Curve at the Top Correlated GFS Location.

3.2. Naive Methods

The HOG features are the transformations of the base components u and v . The naive methods considered in this thesis include the base components and their derivatives wind speed w and direction d . Another naive representation is partitioning the $(-180, 180]$ interval into bins and aggregating the wind speed terms in each bin without accounting the continuity over bins. This approach does not connect the -180 and 180 degree points.

Naive methods based on aggregated terms from different locations are in comparative to the aggregated HOG representation. Naive methods with location based terms are suitable to be compared with the representations involving location based HOG features. To forecast the wind power at a wind farm, penalized polynomial re-

gression models and tree based nonlinear learners are employed on different sets of naive features.

Based on the relationship between the cube of the wind speed and the power curve, polynomial wind speed terms up to degree three are taken into account for the linear models. Since the wind direction induces a nonlinear relationship, direct use of these terms is not suitable for the linear models. As opposed to the linear learners, tree based learners are able to detect nonlinear relationships between the features. Therefore, the aggregated and location based components u and v are suitable features for the tree based learners. On the other hand, employing the wind direction as a feature in addition to the wind speed is a suitable feature set for the tree based learners due to the nonlinear relationship between the direction and wind speed features.

Features obtained by partitioning the $(-180, 180]$ interval without addressing continuity are suitable for both linear and tree based learners. However, linear models with polynomial terms are not expected to learn the nonlinear relationship on the boundaries. Moreover, tree based learners are expected to perform better by discovering the nonlinear relationship between the boundary features.

The advantage of tree based learners are their ability to process nonlinear relationships between features. However, in all the naive methods, the boundaries of the wind direction are not connected and the continuity over terms is not established. These models are compared with the suggested HOG representations in the following sections to demonstrate the importance of wind direction properties.

4. METHODOLOGY

In order to describe the conditions at a wind farm, NWP data sets from a number of nearby locations are used. For each NWP model $i \in I$, let $j \in J_i$ be the nearby locations and $k \in K_i$ be the levels to define the characteristics of the weather at the plant site. In this case, levels represent the altitudes such as 250 meters above ground and 925 millibar in terms of atmospheric pressure, etc.

For a WF and an arbitrary NWP model i , $u_{i,j,k}(t)$ and $v_{i,j,k}(t)$ vectors at each location j and level k are available at time t . The magnitude of the wind vector from the origin to $(u_{i,j,k}(t), v_{i,j,k}(t))$ is defined by

$$w_{i,j,k}(t) = \sqrt{u_{i,j,k}^2(t) + v_{i,j,k}^2(t)} \quad (4.1)$$

which represents the wind speed.

The target variable $p(t)$ is the amount of electricity generated in megawatts at time t . Changes in the capacity create distortions in the power production data set $p(t)$. In order to take them into account, the production above 99% percentile for the periods of 30 days is mapped to its maximum production, which gives the corrected power generated namely $p^*(t)$. In addition to the power generated, utilization rate of the farm capacity is used as another target variable. The adjustment on $p(t)$ accounts for capacity changes, which allows defining the utilization rate by the formula

$$u(t) = \frac{p^*(t)}{\max_t p^*(t)}. \quad (4.2)$$

Benefiting from the correlation between wind speed and production, it is efficient to focus on the top correlated level $k^* \in K_i$ of NWP model i , which is found by

$$k^* = \arg \max_k \rho(p^*, w_{i,j,k}), \quad (4.3)$$

for a fixed NWP model i , where ρ is the Pearson's Correlation Coefficient. To aggregate the location based wind vectors at the top correlated level k^* , the average of u vectors

from $|J_i|$ locations is computed as

$$\bar{u}_{i,k^*}(t) = \frac{1}{|J_i|} \sum_{j \in J_i} u_{i,j,k^*}(t). \quad (4.4)$$

One can define \bar{v}_{i,k^*} in a similar manner.

Since $u_{i,j,k^*}(t)$ and $v_{i,j,k^*}(t)$ together define the wind vector of NWP model i in location j at level k^* , the interaction between them can be represented by the wind speed and wind direction. The magnitudes of the corresponding wind vectors are $w_{i,j,k^*}(t)$. For the direction of the wind vector the function

$$\text{atan2}(v, u) = \begin{cases} \arctan\left(\frac{v}{u}\right) & \text{if } u > 0, \\ \arctan\left(\frac{v}{u}\right) + \pi, & \text{if } u < 0 \text{ and } v \geq 0, \\ \arctan\left(\frac{v}{u}\right) - \pi, & \text{if } u < 0 \text{ and } v < 0, \\ +\frac{\pi}{2}, & \text{if } u = 0 \text{ and } v > 0, \\ -\frac{\pi}{2}, & \text{if } u = 0 \text{ and } v < 0, \\ 0, & \text{if } u = 0 \text{ and } v = 0 \end{cases} \quad (4.5)$$

is used, where the standard arctan function has the range $(-\frac{\pi}{2}, \frac{\pi}{2}]$. The angle between the x -axis and the wind vector is defined by

$$d_{i,j,k^*}(t) = \text{atan2}(v_{i,j,k^*}(t), u_{i,j,k^*}(t)) \cdot \frac{180}{\pi}. \quad (4.6)$$

By this definition, the interval for wind direction d_{i,j,k^*} is taken as $(-180, 180]$ degrees. The aggregated wind speed over locations is defined based on the aggregated u and v vectors, by the formula

$$\bar{w}_{i,k^*}(t) = \sqrt{\bar{u}_{i,k^*}^2(t) + \bar{v}_{i,k^*}^2(t)} \quad (4.7)$$

and the aggregated direction is

$$\bar{d}_{i,k^*}(t) = \text{atan2}(\bar{v}_{i,k^*}(t), \bar{u}_{i,k^*}(t)) \cdot \frac{180}{\pi}. \quad (4.8)$$

The explained aggregation is performed based on the wind components u and v , so the aggregated wind speed is not derived from the location based wind speed values. Consider the wind information given in Table 4.1. For the specific date-hour

T1, the weather information is available in 2 locations, namely L1 and L2. In order to represent this time point by a single vector, u and v components are aggregated by taking their means respectively, and then aggregated wind speed and direction values are computed in Table 4.2.

Table 4.1. Before Aggregation

Time	Location	UGRD	VGRD	Wind Speed	Wind Direction
T1	L1	22.825	7.415	24	18
T1	L2	-14.142	14.142	20	135

Table 4.2. Aggregated Features.

Time	UGRD	VGRD	Wind Speed	Wind Direction
T1	4.342	10.779	11.621	68.062

Outlier elimination part requires two quantile regression curves for the wind speed and the production relationship at the NWP location where they are the most correlated. This procedure is not conducted using the aggregated wind speed, since the locations may have very distinct weather predictions based on the physical properties of the land. The top correlated location j_i^* on the top correlated level k^* of an NWP model i is found by the Pearson's Correlation by

$$j_i^* = \arg \max_{J_i} \rho(p^*, w_{i,j_i^*,k^*}). \quad (4.9)$$

Then at the 1% and 99% percentiles two polynomial regression curves up to degree 3 are fit using the relation

$$p^* \sim w_{i,j_i^*,k^*} + w_{i,j_i^*,k^*}^2 + w_{i,j_i^*,k^*}^3. \quad (4.10)$$

The time points below the quantile regression curve at the 1% percentile and above the curve at the 99% percentile are determined as the outliers and removed from the NWP data sets including all locations. This creates an outlier eliminated subset of

p^* . Similarly, the outlier eliminated utilization is calculated by dividing the power generated by the maximum production in this subset. Letting T be the intersection of time points in all NWP models i and T_o be its outlier eliminated subset. After the modifications, there are four targets considered in total; $\{p^*(t)|t \in T\}$, $\{u^*(t)|t \in T\}$, $\{p^*(t)|t \in T_o\}$ and $\{u^*(t)|t \in T_o\}$. The last two of them are the outlier eliminated versions of the first two targets.

Polynomial regression models with power generation as the target variable may make predictions out of the range. Therefore, after obtaining the production forecasts in such models, some adjustments are required before computing the error terms. The highest value of a forecast is expected to be the capacity of the wind farm and the lowest value is zero for the times when there is no production. Evidently, these values generate upper and lower bounds for the forecasts. For a wind farm, the forecast range is determined by $[0, \max_{t \in T_1} p^*(t)]$ where T_1 is the time interval taken for the training data set. Due to the fact that linear learners with Gaussian family are able to extrapolate the linear trend of the production and make forecasts that are out of the boundaries of the actual values, negative forecasts are mapped to zero and forecasts above the capacity are mapped to $\max_{t \in T} p^*(t)$. On the other hand, for tree based learners the explained mapping is redundant since they do not generate a prediction that is out of the forecast range. Eventually, $\hat{p}^*(t)$ is defined as the corrected forecast for the corrected power generated $p^*(t)$ and it is suitable for calculating the error terms properly.

The HOG method is applicable on NWP model data structure for the wind power prediction since the method requires magnitude and direction. The main advantage of using the HOG features is its ability to resolve the discontinuity problem at the boundaries of the direction interval by involving the circular nature of the angles. It also maintains the continuity over the features generated by the wind speed. The implementation of alternative HOG methods will be explained in the following subsections.

4.1. Standard HOG Representation

The wind components, u and v , are the gradients of the wind vectors in horizontal and vertical directions, from which the wind magnitudes and directions are obtained. To obtain the Standard HOG representation, for all data points in all locations, $\{w, d\}$ pairs are mapped to an n -dimensional HOG vector. For n number of HOG features, $(-180, 180]$ interval is divided into n sub-intervals by n points of bins. The corresponding angle for each bin is obtained by a sequence of angles from -180 to 180 in the increasing order by the step size $\frac{360}{n}$. Let α_r be the r^{th} element of this sequence where r is from 1 to $(n + 1)$ by

$$\alpha_r = -180 + (r - 1) \cdot \frac{360}{n} \text{ for } r \in \{1, 2, \dots, n, n + 1\}. \quad (4.11)$$

By this definition, for $r = 1$, α_1 is -180 , and for $r = (n + 1)$, α_{n+1} is 180 . Based on the α_r values, define the sub-intervals A_r as:

$$A_r = (\alpha_r, \alpha_{r+1}]. \quad (4.12)$$

Let b_{α_r} be the r^{th} HOG bin for $r \in \{1, 2, \dots, n\}$, without an $(n + 1)^{th}$ term. Mapping the pair $\{w, d\}$ to the n -HOG features, the selected bins and their values depend on b and w , respectively. Suppose d lies in the interval A_{r^*} for $r^* \neq n$, then the corresponding bins are $b_{\alpha_{r^*}}$ and $b_{\alpha_{r^*+1}}$. Otherwise, $r^* = n$ and d lies in the interval A_n . In this case, to maintain the cyclic behavior the corresponding bins are selected as b_{α_n} and b_{α_1} . Corresponding two bin values are defined in the same manner for both cases of $(b_{\alpha_{r^*}}, b_{\alpha_{r^*+1}})$ and $(b_{\alpha_n}, b_{\alpha_1})$. Based on the weighted voting principle, the two bin values are computed by

$$\left(\frac{\alpha_{r^*+1} - d}{\alpha_{r^*+1} - \alpha_{r^*}} \cdot w, \frac{d - \alpha_{r^*}}{\alpha_{r^*+1} - \alpha_{r^*}} \cdot w \right) \quad (4.13)$$

respectively. The values for the rest of the bins are zero. Transformation T is conducted for all data points in all locations by

$$T : (w, d) \mapsto (b_{\alpha_1}, \dots, b_{\alpha_n}) \quad (4.14)$$

which is defined as

$$T(w, d) = \begin{cases} b_{\alpha_r} = \frac{\alpha_{r+1}-d}{\alpha_{r+1}-\alpha_r} \cdot w, & \text{if } d \in A_r \text{ and } r \neq n, \\ b_{\alpha_{r+1}} = \frac{d-\alpha_r}{\alpha_{r+1}-\alpha_r} \cdot w, & \text{if } d \in A_r \text{ and } r \neq n, \\ b_{\alpha_n} = \frac{\alpha_{r+1}-d}{\alpha_{r+1}-\alpha_r} \cdot w, & \text{if } d \in A_r \text{ and } r = n, \\ b_{\alpha_1} = \frac{d-\alpha_r}{\alpha_{r+1}-\alpha_r} \cdot w, & \text{if } d \in A_r \text{ and } r = n, \\ 0, & \text{otherwise.} \end{cases} \quad (4.15)$$

For the Aggregated-HOG representation (AHOG), HOG features in all locations are aggregated by computing their mean HOG values, resulting in an n -HOG representation for each date and hour in the time interval. This concludes the explanation of how the wind speed and direction pairs are mapped to n -HOG features. Consider the example given in Table 4.1 and their aggregated features given in Table 4.2. To create 4 HOG features, wind speed values are distributed over the $(-180, 180]$ interval for each time and location pair as given in Table 4.3.

Table 4.3. Example of HOG Transformation.

Time	Location	Bin1	Bin2	Bin3	Bin4
T1	L1	0	0	19.2	4.8
T1	L2	10	0	0	10

To form the Aggregated-HOG representation, location based u and v components are aggregated, respectively, by taking the mean of each component at the same time point. This aggregation is on the vector level with a simple averaging operation. Then, the aggregated wind speed and direction values based on the mean u and v features are obtained. The Aggregated-HOG bins with the obtained wind speed and direction are provided in Table 4.4 ensuring that the resulting representation provides a single feature vector for each time point. Based on this transformation, the weather information in different locations contribute to the plant site equally. However, this may not be valid due to the topographical characteristics of the farm location.

Table 4.4. Example of Aggregated-HOG Bins.

Time	Bin1	Bin2	Bin3	Bin4
T1	0	0	2.833	8.788

For the Location-Based-HOG representation (LHOG), instead of aggregating the HOG features in Table 4.3, the transformed data set is reshaped such that the HOG representation of each location forms 4 HOG bins, resulting in 8 HOG features. To form the Location-Based-HOG representation, the HOG features in different locations are concatenated in order to represent each time point by a single feature vector again. The result of the transformation on the same example is provided in Table 4.5. As a result, n HOG bins in each location forms $n \cdot |J_i|$ HOG features where $|J_i|$ is the number of locations for NWP model i in the Location-Based-HOG representation.

Table 4.5. Example of Location-Based-HOG Bins.

Time	L1				L2			
	Bin1	Bin2	Bin3	Bin4	Bin1	Bin2	Bin3	Bin4
T1	0	0	19.2	4.8	10	0	0	10

4.2. Supervised HOG Representations by Tree Based Clusters

Regression Trees are non-parametric methods that are used to partition the data points into clusters [42]. Distinctive properties of clusters hold valuable information to model the collective behavior of data points. In each cluster separate n -HOG transformations are conducted to recover the lying information based on the variables. The trees are formed on training set, with respect to the production, then the clusters are obtained in the whole data set, assuring not to use the production in the test data.

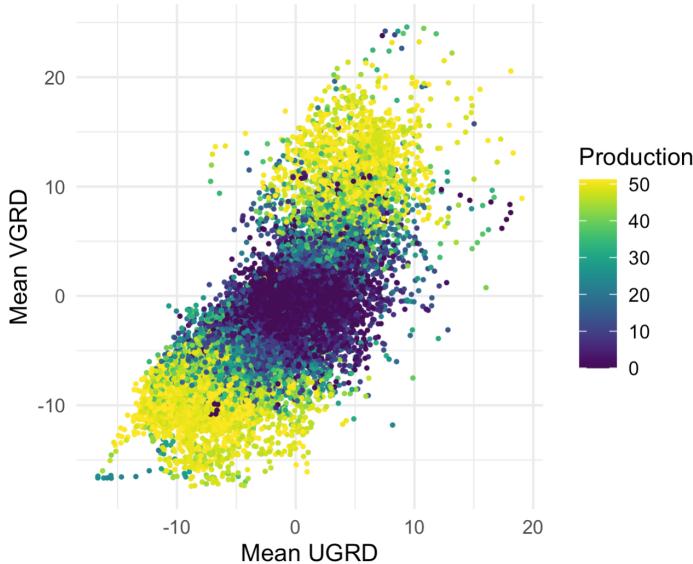


Figure 4.1. Aggregated- u - v vs. Production on GFS.

There are alternative ways to cluster the data points. As illustrated in Figure 4.1, high production is expected to occur at the directions where the wind speed is large. In order to benefit from the relationship between wind direction and production, Aggregated- u - v space is recursively partitioned into m rectangular clusters based on the production in the training data. For each cluster, n -HOG is applied on the arc which covers the points in that cluster. For the regression trees, the minimum number of points required at each node to split further and the minimum number of points required at the leaf node are fixed as 200 and 100, respectively.

The feature representation with separate HOG features in each cluster based on the aggregated u and v terms is called AuvHOG. Again, the number of clusters m and the number of bins n are parameters to be tuned. For simplicity, the pair with the lowest error on the training set is chosen. An example of 8 rectangular clusters with Regression Tree of depth 3 for GFS data set is illustrated in Figure 4.2(a), where the arc covering each cluster is a continuous subset of $(-180, 180]$ which allows overlapping. There is a cluster covered by $(-180, 180]$ and the others are covered by acute or wide angles. This clustering method allows focusing on shorter intervals to be binned. This approach creates splits by lines in Figure 4.2(a). However, they correspond to nonlinear splits with respect to wind speed and direction as illustrated in Figure 4.5(b). It is

seen that the magnitude of the wind speed is captured by this clustering method and the different directions with high wind speed are also grouped in separate clusters.

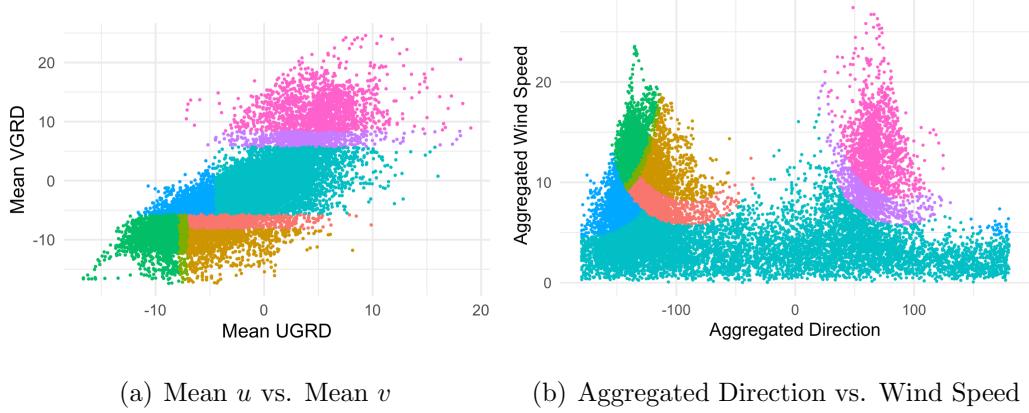


Figure 4.2. Aggregated- u - v Clusters on GFS.

The Standard HOG bins do not use the relationship between the data points and they partition the whole space by the wind direction. The corresponding Standard 6-HOG bins are illustrated in Figures 4.3(a) and 4.3(b). Tree based clusters enable the bins to shift and change its width.

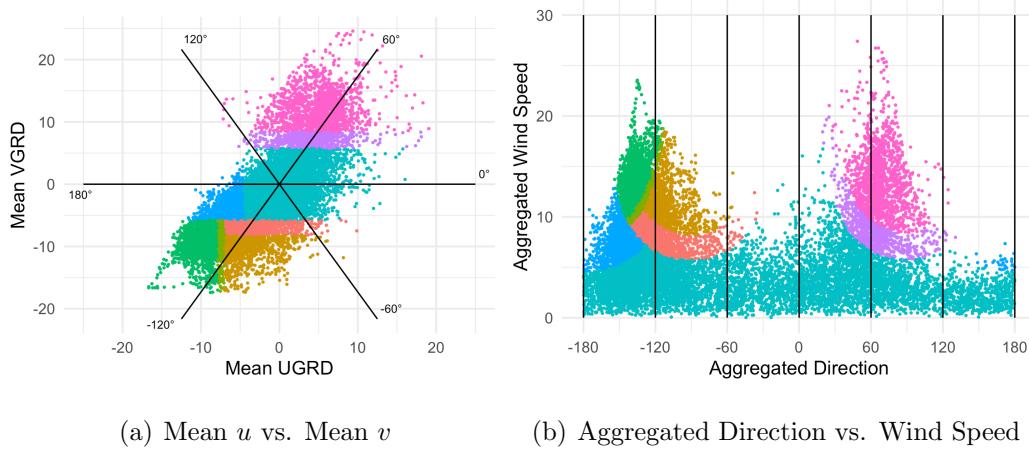


Figure 4.3. Aggregated- u - v Clusters on GFS with Standard HOG Bins.

Figure 4.4 shows the corresponding regression tree built for Figure 4.2(a) and 4.2(b). For this wind farm, aggregated v component is the most important feature since it is the determiner in the first and second split levels.

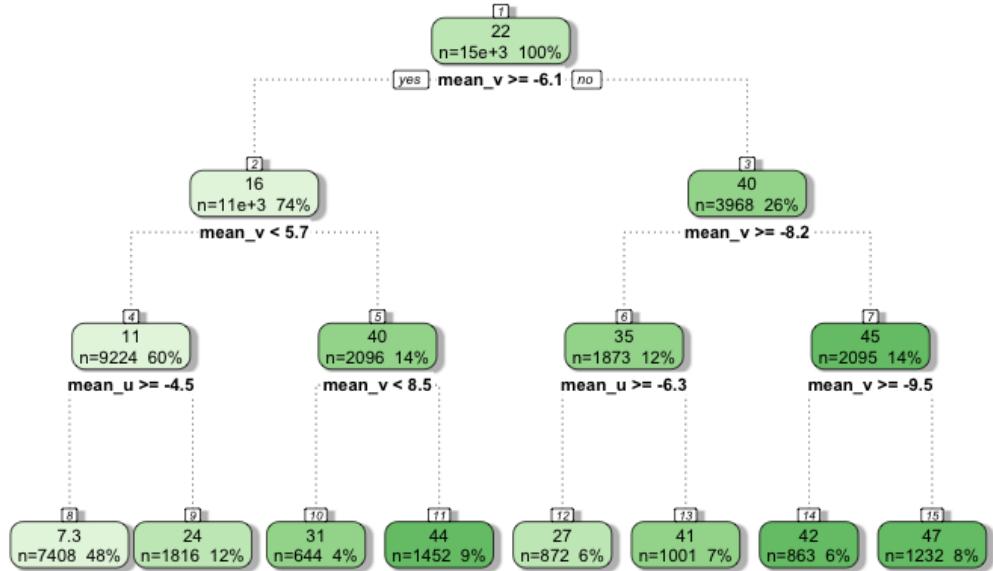


Figure 4.4. Regression Tree with Depth 3 on the Aggregated- u - v Features of GFS.

Another approach is partitioning the Aggregated- w - d set. An example of 8 clusters on GFS data set is shown in Figure 4.5(a). Mean wind speed is more valuable for the Regression Tree clusters since only a few splits are made at the lower split levels based on mean wind direction. Illustrating same clusters on the corresponding Aggregated- u - v data set in Figure 4.5(b) supports this conclusion since the clusters are torus shaped. Also, there are no clear splits visible based on the wind direction. The feature representation with separate HOG features based on these clusters is called *AwdHOG*.

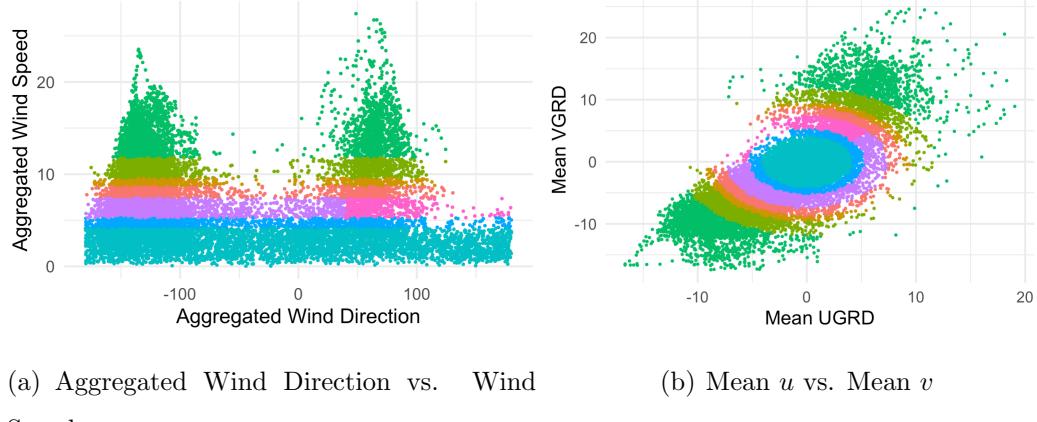


Figure 4.5. Aggregated- w - d Clusters on GFS.

The corresponding regression tree built for Figure 4.5(a) and 4.5(b) is shown in Figure 4.6. For this wind farm, aggregated wind speed component is the most important feature. Only one split is made at the last level with respect to the aggregated direction. Based on this regression tree approach, the contribution of the wind direction is not important as much as the wind speed.

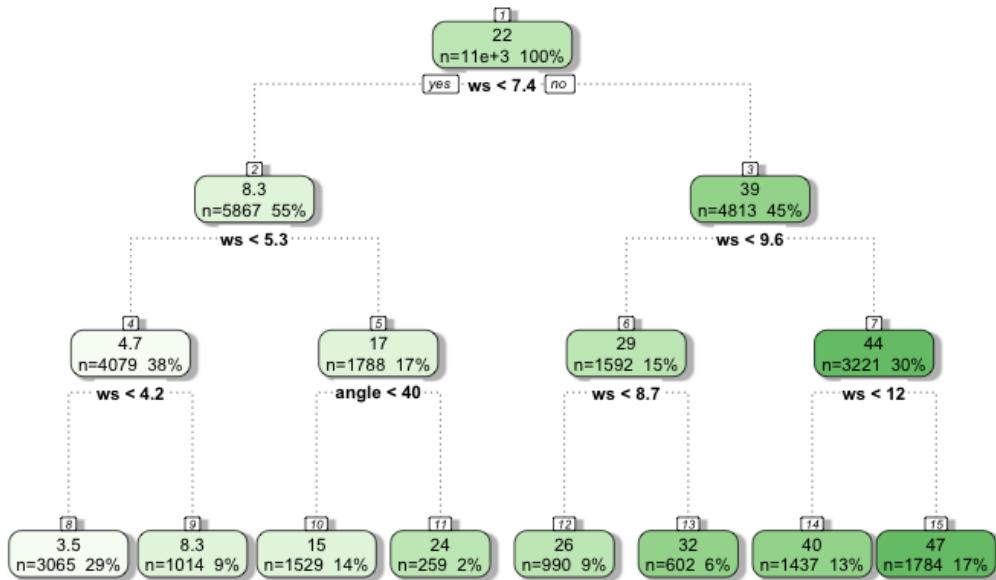
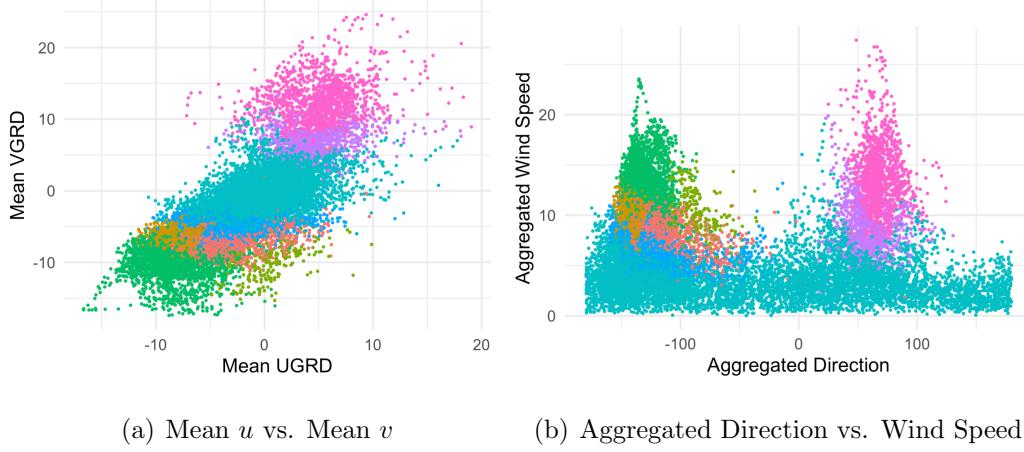
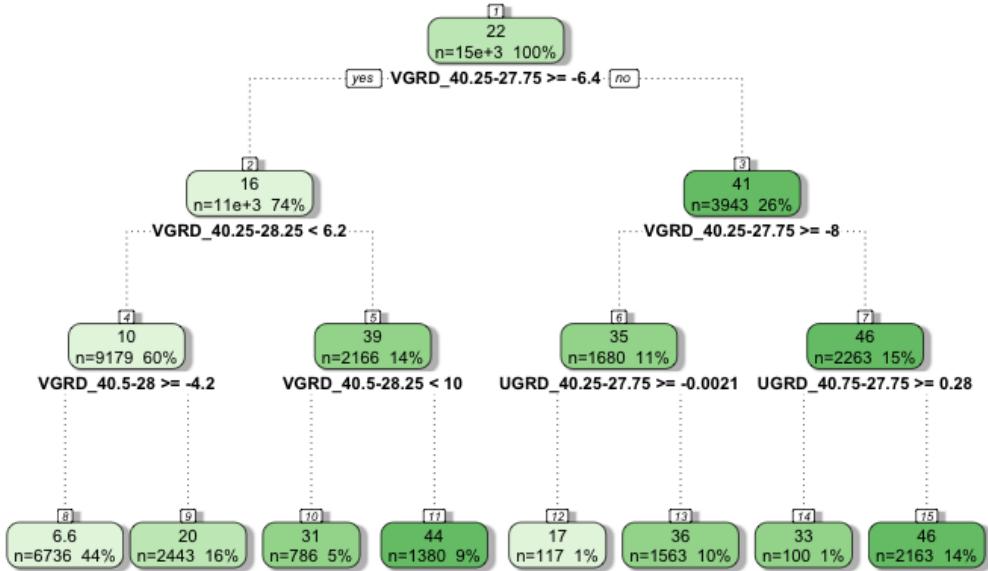


Figure 4.6. Regression Tree with Depth 3 on the Aggregated-*w-d* Features of GFS.

Clustering the data set with respect to aggregated components ignores the spatial relationships of the location based components. To attach importance to the location based components, the Location-Based-*u-v* feature space is clustered. The clusters obtained from a tree with depth 3 in aggregated space are illustrated in Figure 4.7(a). Comparatively to the Aggregated-*u-v* clusters, the location based feature space is linearly partitioned, and this corresponds to non-linearly separated clusters in the aggregated space. Two points with similar aggregated *u* and *v* components may be in different clusters based on their location based components.

Figure 4.7. Location-Based- u - v Clusters on GFS.

The corresponding regression tree preformed based on the u and v components at each location of an NWP model is shown in Figure 4.8 and one can see that the v component is the most important distinguishing variable. The supervised representation of the aggregated features is performed in each Location-Based- u - v cluster, and it is called LuvHOG.

Figure 4.8. Regression Tree with Depth 3 on the Location-Based- u - v Features of GFS.

Similarly, Location-Based- w - d feature space is partitioned into clusters and the corresponding representation is named as LwdHOG. Since the Aggregated- w - d tree

creates torus shaped clusters in Figure 4.5(b), Location-Based- $w\text{-}d$ tree clusters have distorted torus shapes on the aggregated $u\text{-}v$ feature space. This is due to the fact that the wind direction is considered at the further splits.

It is necessary to define a more generalized version of n -HOG transformation within a given interval instead of $(-180, 180]$. The continuity over the boundaries is disregarded, since they are not connected. Let S denote a cluster of points obtained by a regression tree. The convex hull of S , $\text{conv}(S)$ is the smallest convex set enclosing all its points [43]. Call this subset S_c , it is clear that $S_c \subset S$ is satisfied. By definition, S_c is minimal and unique, and the inner angles of its elements are at most 180 degrees. Connecting the outermost points in S_c yields a polygon containing all the points in S . The elements of S_c are of the form $(\bar{u}_{i,k^*}, \bar{v}_{i,k^*})$ and two of its data points are on the boundaries of the arc covering S . In order to find the arc covering all data points in S , the pairwise angles between the points in S_c are calculated. The points with the largest angle between them, create the interval of direction to be partitioned by n -HOG and the step size between bins is calculated as

$$\frac{1}{n-1} \cdot \max_{S_c} \angle(\overrightarrow{\bar{u}_{i,k^*}}, \overrightarrow{\bar{v}_{i,k^*}}). \quad (4.16)$$

The two data points of interest are found by

$$\arg \max_{S_c} \angle(\overrightarrow{\bar{u}_{i,k^*}}, \overrightarrow{\bar{v}_{i,k^*}}) \quad (4.17)$$

with the corresponding directions d_1 and d_2 . Intuitively, the interval to be binned is $(d_1, d_2]$ for $d_1 < d_2$ when both of the directions are of the same sign and their acute angle encloses all the data points in the cluster. After conducting a case by case analysis over the possible quadrants where these two data points are located, necessary manipulations are carried out on the data points in the cluster and the boundaries of the interval to be binned. For example, a third point is considered to determine whether d_1 and d_2 form an acute or wide angle. This changes the interval and steps size. And, in case the arc covers a point with direction of -180 degrees, to obtain the correct partitions, all the directions in that cluster are shifted by $|d_1|$ where $|\cdot|$ denotes the absolute value. After these manipulations $(d_1, d_2]$ is updated accordingly. For simplicity, call the updated interval $(\theta, \beta]$.

For each cluster, the corresponding $(\theta, \beta]$ interval is used for representing its data points. The interval covered by a cluster of points may be too narrow when the aggregated wind speeds for the time points are very similar. Due to this reason, the clusters with intervals narrower than 60 degrees are not partitioned into bins. Instead, for the points in these clusters, the aggregated wind speed is used directly. For the clusters with intervals wider than 60 degrees, $(\theta, \beta]$ interval is divided into n sub-intervals by n points of bins in order to obtain n number of HOG features. In case the arc covers 360 degrees, the Standard HOG representation is implemented to maintain the cyclic behavior at the boundaries. Otherwise, the corresponding angle for each bin is a sequence of angles from θ to β in the increasing order by the step size $\frac{\theta-\beta}{n-1}$. With this partitioning there is no cyclic behavior among the first and the last bins since they do not correspond to the same direction.

The interval is found by the following procedure. Let α_r be the r^{th} element of this sequence where r is from 1 to n by

$$\alpha_r = \theta + (r - 1) \cdot \frac{\beta - \theta}{n - 1} \text{ for } r \in \{1, 2, \dots, n\}. \quad (4.18)$$

Based on the α_r values, define the sub-intervals $A_r = (\alpha_r, \alpha_{r+1}]$. One can observe that $\alpha_1 = \theta$ and $\alpha_n = \beta$ based on this definition. Then, to transform the features into two consecutive bins the bin values need to be determined. Suppose d lies in the interval A_{r^*} for a given pair of w, d . Then the corresponding two bin values are defined as

$$(b_{\alpha_{r^*}}, b_{\alpha_{r^*+1}}) = \left(\frac{\alpha_{r^*+1} - d}{\alpha_{r^*+1} - \alpha_{r^*}} \cdot w, \frac{d - \alpha_{r^*}}{\alpha_{r^*+1} - \alpha_{r^*}} \cdot w \right). \quad (4.19)$$

The values for the rest of the bins are zero again. Since there is no continuity among the bins corresponding to the boundaries of the arc, they do not have to be connected. This transformation is conducted for all data points in the particular cluster when $(\theta, \beta] \neq (-180, 180]$.

In the Supervised-HOG representation, for each cluster, aggregated wind speeds are distributed over n bins with respect to their directions creating a comparative representation to the aggregated naive features. This transformation assures continuity over the features in a cluster except for the bins corresponding to the boundaries of the arc when they are not connected. For the clusters spanning $(-180, 180]$ interval,

weighted distribution is conducted among the boundary bins to maintain the circularity and continuity. In this case, the transformation described for the Standard HOG representation is used.

4.3. Combined Representations and Ensemble Modelling

In addition to representations with a single NWP, all available NWP models for a WF are combined to obtain an ensemble of NWP representation to create a better understanding on the actual weather conditions. Ensemble of NWP representations benefit from each NWP model's advancements. The improvements in wind power forecasting performances of the HOG representations are expected due to the Naive results.

In comparative to the naive representations, Standard HOG and Supervised HOG features are concatenated with respect to date and hour. With the combined representations, the aim is to increase the information gained from the transformations. In parallel, the transformed and the naive features are combined in order to identify the additional impact of the transformations. Recall that, the arc intervals binned in the supervised tree models are more narrow, while AHOG bins are spread over the whole space. This means that the supervised models are expected to capture valuable information when combined with LHOG, since the intervals binned are alike.

Predictions based on alternative representations may be combined to improve the prediction accuracy. A weighted prediction is found by solving a minimization problem for the least squared error. This method finds the coefficients corresponding multiple predictions subject to the weight constraints. The first constraint is that the weights are in the range $(0, 1)$, and the second one is their summation is equal to 1. Explained problem is solved based on the predictions of the training set to prevent overfitting and then the obtained weights are used to find the ensemble prediction for the test set.

After obtaining the predictions for the wind power production and utilization from the single representation models, their ensemble is used to find the weight of each

model's prediction. Standard HOG and Supervised HOG predictions for the train sets are employed as the ensemble members to determine the prediction weights then they are used to calculate the error for the test sets.

5. EXPERIMENTS ON REAL DATA

In this thesis, the experiments are conducted on 47 wind farms in Turkey. Most of the wind farms are located on the west coast of Turkey and the locations are shown in Figure 5.1. All the experiments for a wind farm are performed on its top correlated level k^* of NWP model i . Since the time interval differs for each NWP model i and wind farm, the common interval T of all NWP models is considered for each wind farm. The available time periods of each NWP model and information on the wind farms are included in Appendix A.

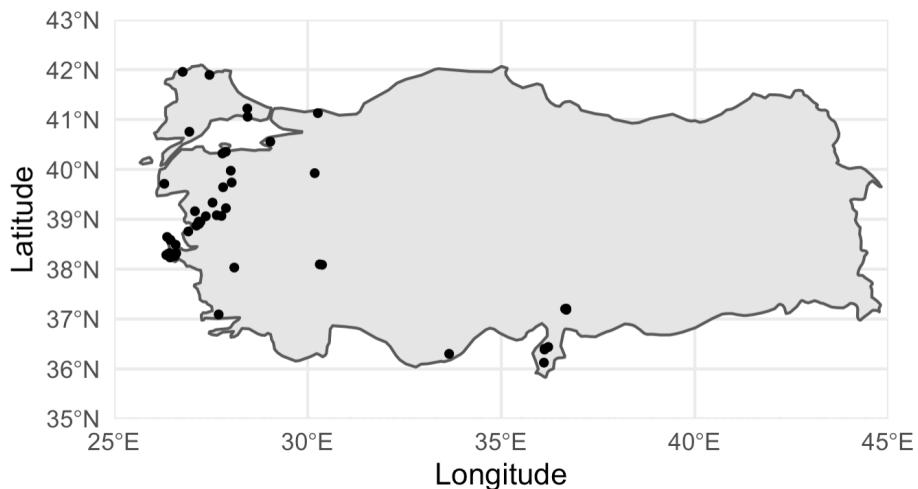


Figure 5.1. Wind Farm Locations Used in Experiments.

In the following sections, the learners are introduced and the representations are specified. After introducing the performance metric, comparisons of multiple regressors over 47 wind farms are conducted. The approach in selecting the representation parameters is also provided for Standard and Supervised HOG representations.

5.1. Representations and Learners

The data sets are splitted into train and test sets where test set size is set to the last 365 days. Model parameters are tuned by five-fold cross-validation on the training set. The reason for not employing time series cross validation with sliding windows is

that the lagged features are not included in any of the proposed representations and in such cases the usage of K-fold cross validation is proven to be valid [44, 45].

Two learners are selected for wind power forecasting; Generalized Linear Models (GLM) [46] with Lasso Regularization and Ranger [47] which is a fast implementation of Random Forests (RF). For GLM learners with Gaussian and Quasibinomial family, LASSO (L1) regularization term is tuned for 100 values, and for RF learners the number of features to possibly split at each node and the splitting rule are tuned on the training sets. Depending on the representation, two values for the number of features to possibly split at each node are considered. These values are $\frac{1}{6}$ and $\frac{1}{3}$ of the number of features. For the naive representations with less than 6 features, the number of features to possibly split at each node is tuned over the set $\{1, 2\}$. The splitting procedure is performed for two criteria. The first one makes splits based on the variances of possible splits with randomly selected features. In the second splitting strategy, the nodes are created by choosing cut-points randomly [48]. This results in a total parameter set of size 4 to be tuned in RF models. And the number of trees is fixed as 500.

First, the models with naive representations are established. For linear learners, two representations are considered. Aggregated-Wind-Speed (Aw) is based on the features \bar{w}_{i,k^*} and for the Location-Based-Wind-Speed (Lw) w_{i,j,k^*} are used. Regarding the linear relationship between production and the cube of the wind speed, the features along with their polynomial terms up to degree 3 are used in Aw and Lw representations. Random Forest learners are able to detect nonlinear relationships between the features and the target variable [47] whereas the GLM is limited to the linear relationships. Lw without the polynomial terms is also considered in RF modelling. In addition to Lw features, location based wind directions d_{i,j,k^*} are also used as input variables in the Lwd representation, employing their nonlinear interactions that GLM disregards.

Another naive representation is Auv where the aggregated features are \bar{u}_{i,k^*} and \bar{v}_{i,k^*} . Luv representation is also evaluated with u_{i,j,k^*} and v_{i,j,k^*} in each $|J_i|$ location,

aiming to conceive the importance of locations. Even though, the nonlinear interaction of wind components is detected with RF, the continuity along the semi-infinite line $x < 0$ and $y = 0$, where -180 degrees and 180 degrees represent the same direction is still disregarded. All the naive representations and their corresponding features are shown in Table 5.1. In the first column representation names for RF models are indicated with the suffix .RF, and the GLM representations with polynomial terms do not have any suffix. In the second and third columns, the naive features and the total number of them are indicated for each naive representation. The last column is added to indicate the feature types.

Table 5.1. Naive Representations for a Single NWP Model i .

Representation	Features	N. of Features	Version
Aw	$\bar{w}_{i,k^*}, \bar{w}_{i,k^*}^2, \bar{w}_{i,k^*}^3$	3	Aggregated
Lw	$w_{i,j,k^*}, w_{i,j,k^*}^2, w_{i,j,k^*}^3$	$3 \cdot J_i $	Location Based
$Aw.RF$	\bar{w}_{i,k^*}	1	Aggregated
$Auv.RF$	$\bar{u}_{i,k^*}, \bar{v}_{i,k^*}$	2	Aggregated
$Lw.RF$	w_{i,j,k^*}	$1 \cdot J_i $	Location Based
$Lwd.RF$	w_{i,j,k^*}, d_{i,j,k^*}	$2 \cdot J_i $	Location Based
$Luv.RF$	u_{i,j,k^*}, v_{i,j,k^*}	$2 \cdot J_i $	Location Based

The last naive representation is obtained by partitioning the $(-180, 180]$ interval into bins and taking the sum of location based wind speed values corresponding to each bin with respect to their directions. The number of bins considered in this representation is determined by the sensitivity results of the Standard HOG modellings which is explained in the following sections. This representation is called LwB . Instead of using the location based wind speed and direction, their aggregated values are used to form AwB . These representations do not address the continuous and cyclic behaviour of the wind direction.

The standard HOG representations AHOG, LHOG, and the supervised HOG representations LuvHOG, and LwdHOG are suitable to be employed in both of the learners, but for GLM the polynomial features need to be considered as well. In these representations, the number of features depends on the sensitivity results. AHOG and LHOG are the more complex versions of AwB and LwB, in which the continuity and cyclic behaviour of the wind direction is addressed. The additional representations used for GLM are listed in Table 5.2. Second column is added to specify how the representations are obtained and the third column is an indicator of comparative representations for comparison of the results. Remark that for the GLM learner, the polynomial terms up to degree 3 are employed in the representations.

Table 5.2. Additional GLM Representations.

Representation Name	Type	Version
AwB	Naive	Aggregated
LwB	Naive	Location Based
AHOG	Standard HOG	Aggregated
LHOG	Standard HOG	Location Based
LuvHOG	Supervised HOG	Aggregated
LwdHOG	Supervised HOG	Aggregated

Additionally, combined representations and ensemble modeling are employed. For the combined representations, $Lw.LHOG.LwdHOG$ and $LHOG.LwdHOG$ regressors are formed by concatenating the features and GLM is performed. For RF models, $Luv.LHOG$ without the polynomial terms is considered. This type of ensembling is on the feature level. On the other hand, taking the weighted averages of the predictions is considered for ensemble modelling, and they are listed in Table 5.3 with the prefix E. This type of ensembling is on the model prediction level.

Table 5.3. Ensemble Techniques.

Representation Name	Ensemble Type	Version
LHOG.LwdHOG	Feature	Combined
Lw.LHOG.LwdHOG	Feature	Combined
Luv.LHOG.RF	Feature	Combined
E.LHOG.LwdHOG	Prediction	Combined
E.Lw.LHOG.LwdHOG	Prediction	Combined
E.Luv.LHOG	Prediction	Combined

In power utilization forecasting, only the GLM models are used. Since the utilization is between $[0, 1]$, the quasi-binomial family is suitable for the GLM learner. It also provides predictions that lie within the same interval with the target variable. By employing this family, overdispersion in the response variable is accounted.

5.2. Performance Metric

The metric used to evaluate the performance of the forecast models is Weighted Mean Absolute Percentage Error (WMAPE) given by

$$\frac{1}{\sum_{t \in T_2} p^*(t)} \sum_{t \in T_2} |p^*(t) - \hat{p}^*(t)| \times 100, \quad (5.1)$$

where T_2 is the time interval taken for the test set, $p^*(t)$ is the actual corrected power generated, and $\hat{p}^*(t)$ is its corrected forecast. The selected performance metric weights the error terms by the total power generated and shows the general accuracy of forecasts instead of specific hours with distortions. WMAPE is used in every model learning process, including tuning and selection of representation parameters. The overall model performances are also compared based on their WMAPE results.

5.3. Results and Model Comparisons

In the model comparison process, the Friedman [49] and the post-hoc Nemenyi tests [50] are performed to detect the performance differences between multiple methods. Nemenyi test is a non-parametric version of the Analysis of Variance allowing to make pairwise comparisons on the average ranks of the WMAPE values obtained for 47 wind farms. For a given confidence level, the critical distance shows whether there is a significant difference in the performances of pairwise representations [51]. There is a significant difference in the performances of two representations if the difference of their average ranks is at least the critical distance.

5.3.1. Benefits of Combining Multiple NWPs and Outlier Elimination in Linear Models

The NWP models, GFS and ARPEGE, provide information at distinct locations, and each model specializes in different weather conditions. For the purpose of improving the prediction accuracy by integrating advancements of different NWP models, their ensemble is used. On the other hand, outlier points in each NWP model are removed as a pre-processing step. Recall that to remove the outliers in each NWP model, the top correlated location with the production is determined and then the points out of the polynomial regression curves at 1 and 99 percentiles are detected. The corresponding ensemble of NWP is obtained by the outlier eliminated NWP models. To show the effects of using ensemble of NWP and removing outliers, separate NWP models are compared with their ensemble. Separate Nemenyi Tests for power generation and power utilization are conducted on the Aw representation with GLM learner. In this process, the forecasts for the same test points are compared.

Nemenyi Test results in Figure 5.2(a), provides the average ranks of the WMAPE values for power prediction, with the best average rank obtained using the ensemble of NWPs. In this figure, the suffix .O stands for the outlier removed data sets. In the same figure, the effects of removing outliers is also available. The interpretation of the Nemenyi test depends on the pairwise mean ranks and the critical distance.

Based on the Nemenyi test in 95% confidence level, there is a significant difference in the performances of single and ensemble of NWP models, since the difference of their average ranks is at least the critical distance, 1.1. As stated in the literature and shown in the Nemenyi Test, ensemble of NWPs provide better forecasts for the wind power. Moreover, the performances with single NWP models are not significantly different in this confidence level. On the other hand, based on the average mean ranks, removing outliers improves the prediction accuracy in all NWP models.

In Figure 5.2(b), another Nemenyi Test is conducted for power utilization, indicating highly similar inferences. In this figure, the suffix .U stands for the utilization prediction. Again, removing outliers and combining NWP models is beneficial. As a result, for both of the power generation and utilization problems, using an ensemble of NWPs and removing outliers are the two important steps for improving the prediction accuracy. Therefore, the ensemble of outlier removed NWP models is used in the following representation comparisons and the corresponding suffixes are dropped.

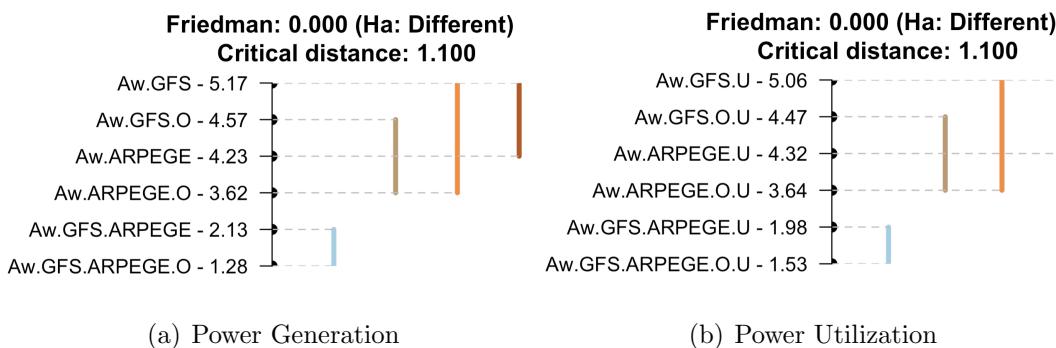


Figure 5.2. Average Ranks of the GLM Models with Single vs. Ensemble of NWPs.

To compare the performances of the models for power generation and utilization, a paired T-test is conducted on the WMAPE values of outlier removed ensemble of NWP representations. These two models correspond to the best models in Figure 5.2(a) and Figure 5.2(b). Based on the T-test results, these two models are significantly different in 99% level. Moreover, the mean WMAPE values of the model for power utilization is significantly less than the model for power generation with a p-value $4.381e - 07$ which is less than 0.01. Therefore, the linear models with power generation are eliminated. And, in the following comparisons, power utilization is the response variable for the

linear models. Therefore, the corresponding suffixes .O and .R are dropped in the following representations.

5.3.2. Comparison of the Naive Models

To compare the naive representations with linear models and RF models, the Nemenyi test in Figure 5.3 is conducted for two GLM and five RF models with the suffixes .GLM and .RF, respectively. The response variable for the GLM models is the power utilization, and RF models are fit for the power generation, both on the outlier eliminated ensemble of NWP representations. The WMAPE performances are comparable since the responses are proportional.

Based on the average ranks on 47 wind farms, *Lw* representation performs significantly better than *Aw*, both in linear and tree based models. This result indicates the importance of modeling the spatial dependencies by the location based wind components. On the other hand, *Luv* representation with RF outperforms the others, indicating the importance of the nonlinear relationship between the location based wind components. RF with *Lwd* does not perform as well as *Luv* due to the disconnection on the wind direction boundaries. Representing this nonlinear relationship with the circularity on the direction boundaries would be useful for the linear models.

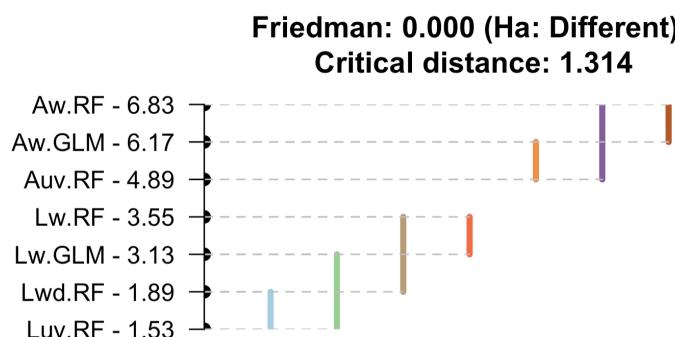


Figure 5.3. Average Ranks of the GLM and Random Forest Models on Naive Representations.

5.3.3. Integrating Continuity and the Cyclic Behaviour of the Wind Direction

Partitioning the wind direction interval for each location, naive LwB representation integrates the direction information with the wind speed. However, it does not maintain the continuity between the features, and lacks the cyclic information over the boundaries. Increasing the number of bins would be an option to represent the continuity. However, this approach would lead high dimensional feature representations. To address these problems with less complex representations, the standard HOG is integrated.

The selected number of bins in LHOG representation depends on the wind conditions at the WF. Although, it is a parameter that needs to be tuned, for each NWP model, the parameters are selected among {6, 9, 12, 18, 24, 36} based on their GLM performances on their corresponding outlier eliminated ensemble of NWP training set by five-fold cross-validation. One approach to maintain continuity is to increase the number of bins. Instead of using high number of bins, HOG transformation is expected to provide continuity with less features. For this reason, it is necessary to analyze how the train error values change as the number of bins increases. The normalized cross-validation WMAPE results on the LHOG representation for the first 5 farms shown in Figure 5.4 imply that the selected number of LHOG bins is highly dependent on the wind farm location and the performance is not linearly dependent on the number of HOG bins. For this reason, the first number of bins where the train error values stop decreasing is selected as the parameter for the Standard HOG representations. After selecting the selected representation parameters on the training data, the test results are reported. Comparatively, the selected number of bins are also used in LwB representations.

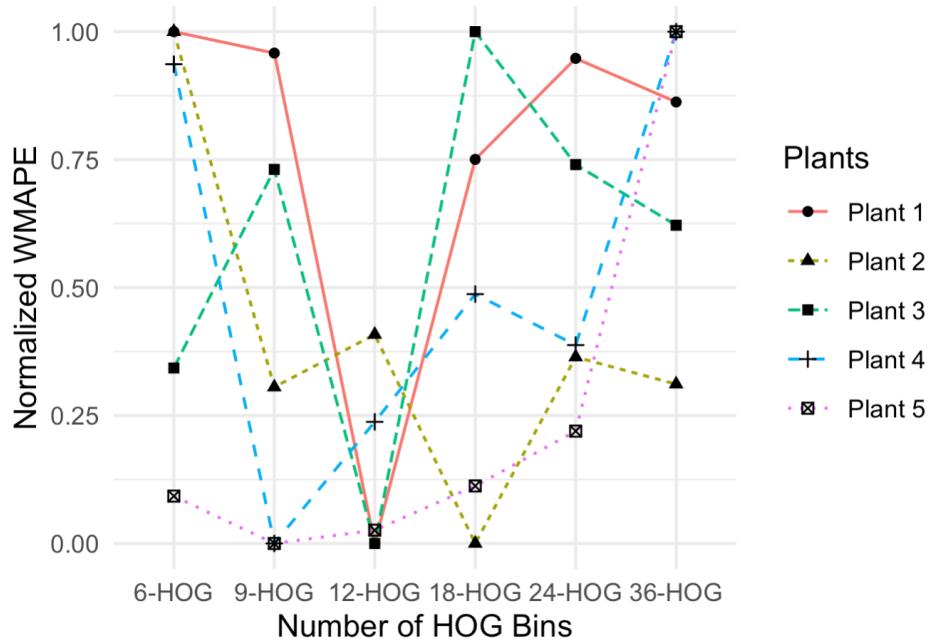


Figure 5.4. Sensitivity Results for the Number of LHOG Bins.

To compare the approaches for integrating location based information, Nemenyi test in Figure 5.5 is conducted. Since the mean rank of Aw is 4, aggregating the location based features increases the error for all wind farms. However, integrating continuity and cyclic behavior of the wind direction with LHOG improves the performance significantly compared to Lw and LwB representations.

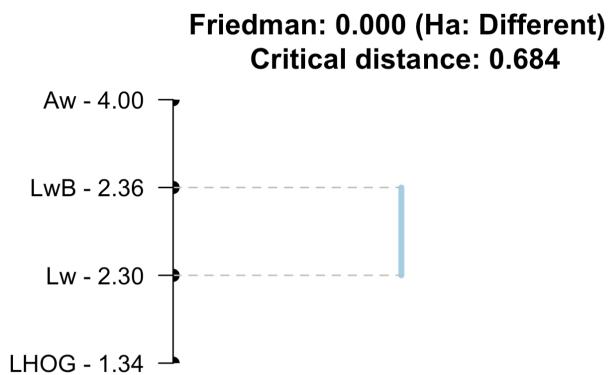


Figure 5.5. Average Ranks of the Selected GLM Models.

5.3.4. Location Embedded Representations by Supervised Approaches

With the additional information gained from the tree based clustering representations, Supervised HOG representations transform the aggregated feature space. Therefore, *LuvHOG* and *LwdHOG* are in comparative to the standard AHOG representation and its naive alternative AwB. For the AHOG representation, the parameter selection approach in LHOG is conducted similarly. In Figure 5.6, the normalized WMAPE values of the first five wind farms are provided. It is observed that the improvement is not correlated with the number of bins. As an example, for the first wind farm, 9-AHOG performance is the best and 24-AHOG performance is the worst and the performances do not get better with the increasing number of bins.

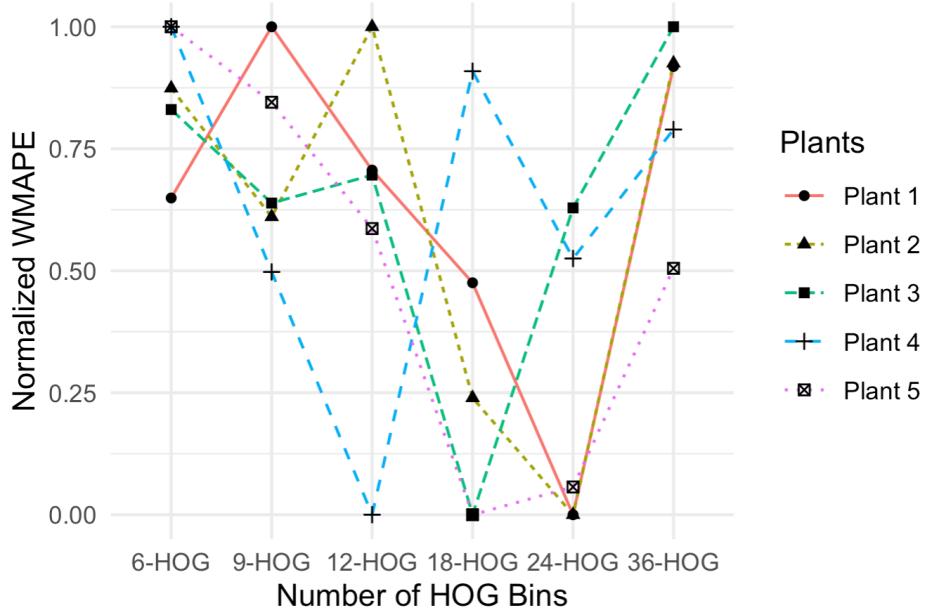


Figure 5.6. Sensitivity Results for the Number of AHOG Bins.

For the Supervised HOG representation setting, Regression Trees are fit on the outlier eliminated training data sets and then the clusters are obtained for test sets as well. Then the HOG is performed on the aggregated features. On each cluster of a Regression Tree with depth τ , n -HOG is applied separately. The parameters, τ and the number of bins n are selected among $\{3, 4, 5, 6\}$ and $\{3, 6, 9\}$, respectively by GLM employing five-fold cross-validation on the training set. The WMAPE values on the training set of *LuvHOG* representation for the first farm is shown in Figure 5.7.

Although, the performance is again not linearly dependent on either of the parameters, the performances with 3 HOG bins are better with all depth options. The results on the train sets are analyzed based on the tree depth and the number of bins. Since a regression tree with depth τ creates at most 2^τ clusters and then each cluster is represented in n bins, the maximum number of features is $2^\tau \cdot n$ for the pair $\{\tau, n\}$. To address the complexity and the number of bins of the resulting representation, the training results are analyzed based on the maximum complexity and the number of bins in an increasing manner. Starting from the lowest complexity and number of bins, the training error values are compared. In this order, $\{\tau, n\}$ is selected when the training error stops decreasing. The parameters are selected separately for each farm by this procedure. Based on these results, to obtain the LuvHOG representation for the first farm, the previously formed clusters are used with the selected τ . The regression tree is formed only on the training data sets, and then they are used to find the clusters of the whole data set. Then the Supervised HOG representations are obtained with the selected n bins. The same approach is conducted for LwdHOG representation as well.

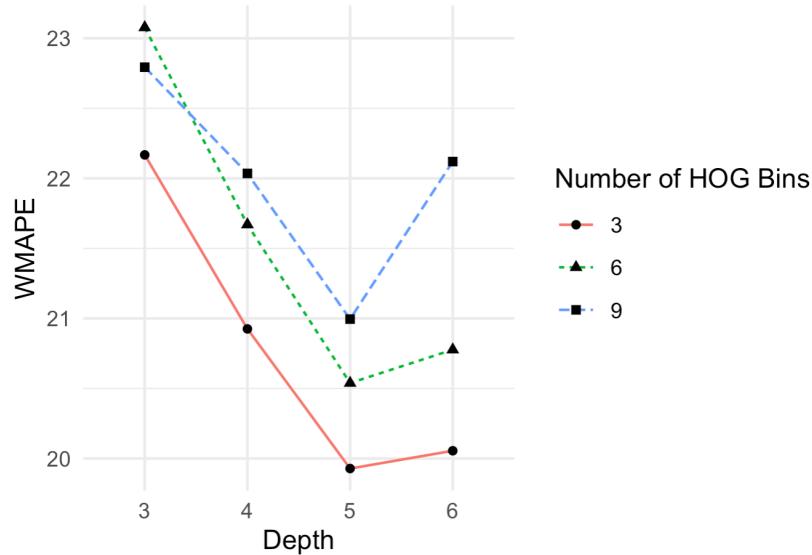


Figure 5.7. Sensitivity Results for Depth and Number of HOG Bins in LuvHOG.

To analyze the additional information gained by the discretization of the feature space, a Nemenyi test is utilized for the HOG representations transforming aggregated features, AwB and LHOG. The results in Figure 5.8 indicate that obtaining a supervised HOG representation with location based clustering methods improves the

prediction accuracy for the aggregated setting. On the other hand, LHOGL still performs better than the other compared representations, since it transforms the location based features.

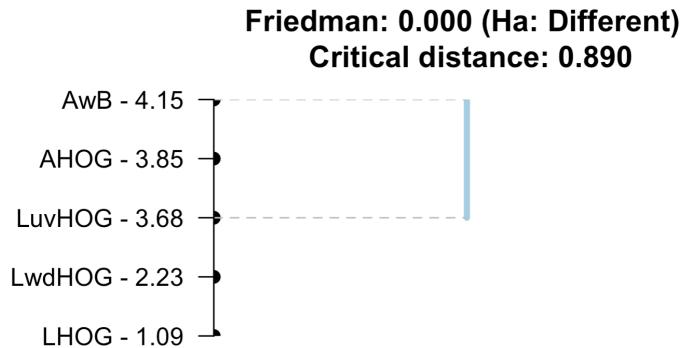


Figure 5.8. Average Ranks of the GLM Models with Supervised Representations.

Recall from the Figure 5.3, that *Luv.RF* performed better than *Lwd.RF*. However, Figure 5.8 suggests that GLM performance of *LwdHOG* is significantly better than *LuvHOG*. The regression tree using location based wind speed and direction features creates distorted torus shaped clusters. Moreover, the wind speed values are the most important features partitioning the space. The splits based on the directions are in the deeper nodes. As a result, in each cluster, the wind speed values are similar, leading a wind speed sensitive representation. Since the power generation is highly dependent on the wind speed, the supervised HOG with this clustering method is more convenient than *LuvHOG*. As a result, supervised transformations are beneficial in the aggregated setting.

5.3.5. Random Forest vs. GLM

Integration of the nonlinear relationship of the features by standard and supervised HOG methods, improves the predictive performance. For the representations LHOGL, *LwdHOG* and *Luv.RF*, their pairwise comparisons of the 47 WMAPE values are illustrated in Figure 5.9. WMAPE values of the *Luv.RF* model are slightly lower than LHOGL's as shown in Figure 5.9(a). On the other hand, the performance of *Luv.RF* surpasses *LwdHOG*'s, as seen in Figure 5.9(b). Lastly, in Figure 5.9(c), LHOGL

representation leads lower errors compared to $LwdHOG$. This is expected since in the supervised representations the aggregated features are transformed.

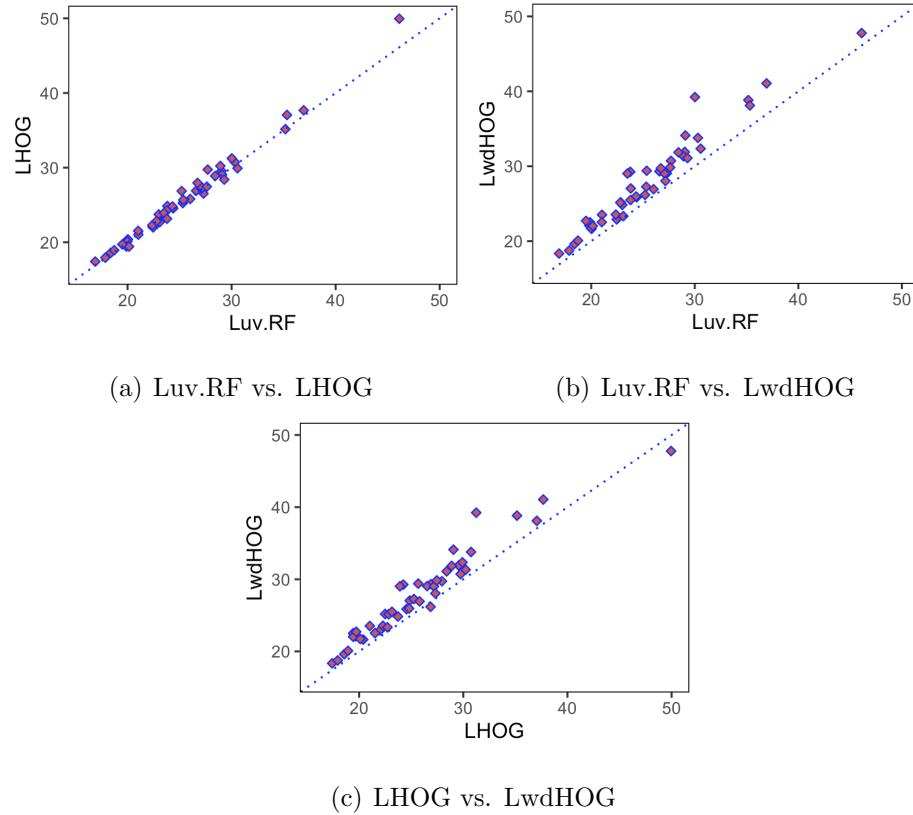


Figure 5.9. Pairwise WMAPE Comparisons.

5.3.6. Combined Representations vs. Ensemble Modelling

Finally, ensembling methods are compared with $Luv.RF$ representation. Combining representations and using GLM or RF is an ensembling on the feature level. For the ensemble modelling, a weighed average of the predictions from multiple models is considered and these methods are indicated with the prefix .E. The corresponding Nemenyi test results in Figure 5.10 indicates that combining the predictions of Luv and $LHOG$ representations decreases the mean rank. However, RF model on their combined features is not preferable. On the other hand, combining features of Lw , $LHOG$ and $LwdHOG$ leads the minimum rank in the Nemenyi test. So as a result, the Standard and Supervised HOG techniques hold valuable information regarding the nonnegativity originating the continuous and cyclic behaviour of the wind direction. With the naive location based wind speed values, the application of GLM on the

combined representation *Lw.LHOG.LwdHOG* outperforms the RF models.

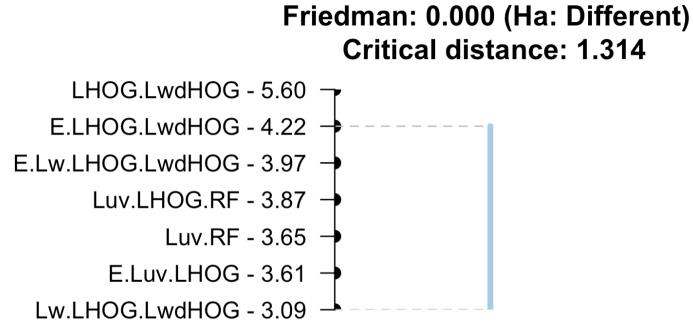


Figure 5.10. Average Ranks of the Ensemble Approaches.

The production amount in the test set can be divided into three categories; low, moderate and high. Depending on the wind farm, below 5% of its capacity is determined as the low production set and above its 95% is selected as the high production set. And, the remaining points form the moderate production set. Production histogram of a wind farm with capacity 119 MW is illustrated in Figure 5.11, the vertical dashed lines correspond to the splits for low, moderate and high production.

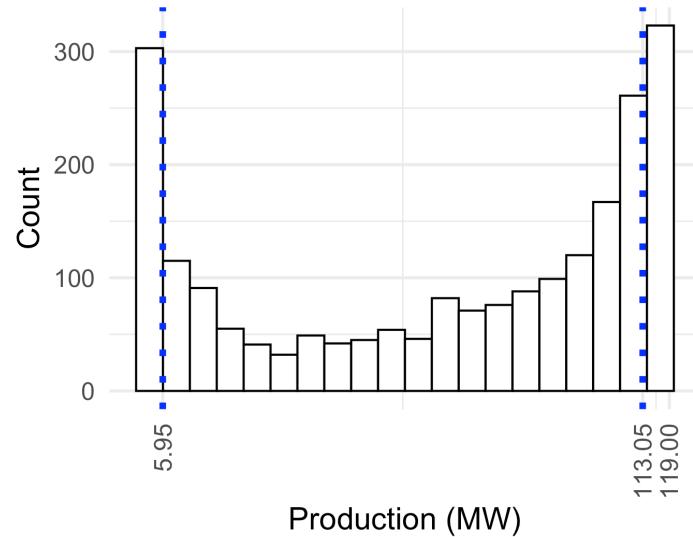


Figure 5.11. Production Histogram for a Wind Farm.

To see the performances of ensemble models in each of these categories, separate Nemenyi Tests are employed. For the low and high production, the corresponding

results in Figure 5.12(a) and Figure 5.12(c) indicate that the *Lw.LHOG.LwdHOG* representation with GLM is preferable to *Luv.RF*. Moreover, *Lw.LHOG.LwdHOG* outperforms the others in the high production category, and its mean rank is significantly lower than RF models. For the moderate production, RF models are preferable to the GLM models based on the results in Figure 5.12(b).

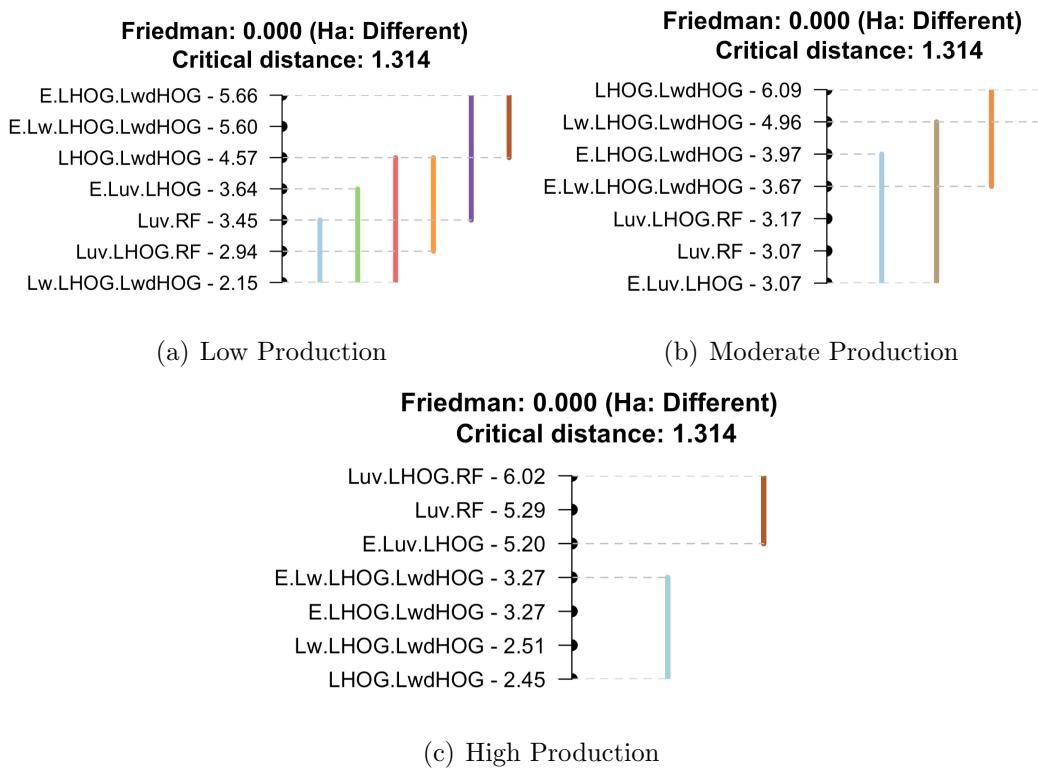


Figure 5.12. Nemenyi Tests in Categories.

6. CONCLUSION

Wind power forecasting based on the NWP models is a challenging task due to the nonlinear relationships between components and the complexity of the feature sets. This nonlinear relationship derives from the wind direction and the spatial relations of the wind components. Existing approaches do not address the continuity and the cyclic behavior of the wind direction. On the other hand, it is observed that using outlier eliminated Ensemble of NWP models for power utilization prediction enhances the performance. Moreover, models with location based representations outperform the aggregated models.

In this thesis, alternative feature extraction techniques for the wind power prediction are proposed. The main purpose of the proposed representations is to extract features expressing the spatial relationships of the wind components such as the continuity and the cyclic behavior of the wind direction. For this aim, Standard and Supervised HOG methods on the outlier eliminated ensemble of NWP models are proposed. It is observed that, the Standard HOG methods represent the continuity over sufficiently small number of features. On the other hand, the Supervised HOG representations transforming the aggregated wind features built on the location based components improves the performance in the aggregated setting by holding spatial information. Combined modeling approaches with the proposed methods improve the performance of the GLM and give sufficiently better results compared to the combined modelling in RF. As a result, it is shown that the proposed representations hold the nonlinear information in a suitable form for the linear learners.

An extension to the proposed approach can be using a well constructed method for aggregating the location based wind components. Moreover, after conducting a tree based clustering method, the location based components can be transformed by a supervised HOG method. However, the drawback would be working with a extremely sparse high dimensional feature representation.

The NWP model data sets and power generation information are publicly available. For reproducibility, all the WMAPE results, selected values in the tuning process and the related scripts are also available [52]. Example data sets for a single wind farm, including production amounts and NWP models, are also provided.

REFERENCES

1. “EXIST-EPIAS Transparency Platform”, <https://seffaflik.epias.com.tr/transparency/index.xhtml>, accessed on October 9, 2022.
2. “ENTSO-E Transparency Platform”, <https://transparency.entsoe.eu/>, accessed on October 9, 2022.
3. Hanifi, S., X. Liu, Z. Lin and S. Lotfian, “A Critical Review of Wind Power Forecasting Methods—Past, Present and Future”, *Energies*, Vol. 13, No. 15, 2020.
4. Foley, A. M., P. G. Leahy, A. Marvuglia and E. J. McKeogh, “Current Methods and Advances in Forecasting of Wind Power Generation”, *Renewable Energy*, Vol. 37, No. 1, pp. 1–8, 2012.
5. “GFS (Global Forecasting System)”, https://www.emc.ncep.noaa.gov/emc/pages/numerical_forecast_systems/gfs.php, accessed on December 9, 2022.
6. “ARPEGE (Action de Recherche Petite Échelle Grande Échelle)”, <http://www.umr-cnrm.fr/spip.php?article121>, accessed on December 9, 2022.
7. Wu, Y.-K., P.-E. Su and J.-S. Hong, “An Overview of Wind Power Probabilistic Forecasts”, *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, pp. 429–433, 2016.
8. Mylne, K., J. Chen, A. Erfani, T. Hamill, D. Richardson, S. Vannitsem, Y. Wang and Y. Zhu, *Guidelines on Ensemble Prediction System Postprocessing*, World Meteorological Organization, 2021.
9. Deppe, A., E. Takle and W. Gallus, “Creation of a WRF Ensemble for Improved Wind Forecasts at Turbine Height”, *Weather and Forecasting*, Vol. 28, 2013.

10. Manwell, J., J. McGowan and A. Rogers, *Introduction: Modern Wind Energy and its Origins, Second Edition*, chap. 1-2, pp. 1–89, John Wiley & Sons, Ltd, 2009.
11. Jing, B., Z. Qian, H. Zareipour, Y. Pei and A. Wang, “Wind Turbine Power Curve Modelling with Logistic Functions Based on Quantile Regression”, *Applied Sciences*, Vol. 11, No. 7, 2021.
12. Zou, M. and S. Z. Djokic, “A Review of Approaches for the Detection and Treatment of Outliers in Processing Wind Turbine and Wind Farm Measurements”, *Energies*, Vol. 13, No. 16, 2020.
13. Wu, Q., F. Guan, C. Lv and Y. Huang, “Ultra-short-term Multi-step Wind Power Forecasting Based on CNN-LSTM”, *IET Renewable Power Generation*, Vol. 15, No. 5, pp. 1019–1029, 2021.
14. Barbounis, T., J. Theocharis, M. Alexiadis and P. Dokopoulos, “Long-term Wind Speed and Power Forecasting Using Local Recurrent Neural Network Models”, *IEEE Transactions on Energy Conversion*, Vol. 21, No. 1, pp. 273–284, 2006.
15. McConnell, R. K., “Method of and Apparatus for Pattern Recognition”, U.S. Patent 4567610 A, January 28, 1986, accessed on October 1, 2022.
16. Freeman, W. T. and M. Roth, “Orientation Histograms for Hand Gesture Recognition”, *International Workshop on Automatic Face and Gesture Recognition*, Vol. 12, pp. 296–301, 1995.
17. Dalal, N. and B. Triggs, “Histograms of oriented gradients for human detection”, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 1, pp. 886–893, 2005.
18. Breeze, P., “Chapter 2 - The Wind Energy Resource”, P. Breeze (Editor), *Wind Power Generation*, pp. 9–17, Academic Press, 2016.
19. Niu, W., J. Huang, H. Yang and X. Wang, “Wind turbine power Prediction Based

- on Wind Energy Utilization Coefficient and Multivariate Polynomial Regression”, *Journal of Renewable and Sustainable Energy*, Vol. 14, p. 013306, 2022.
20. Madsen, H., H. A. Nielsen and T. S. Nielsen, “A tool for predicting the wind power production of off-shore wind plants”, *Proceedings of the Copenhagen Offshore Wind Conference & Exhibition*, 2005.
 21. Landberg, L., “Short-term Prediction of Local Wind Conditions”, *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 89, No. 3, pp. 235–245, 2001, 10th International Conference on Wind Engineering.
 22. Jung, J. and R. P. Broadwater, “Current Status and Future Advances for Wind Speed and Power Forecasting”, *Renewable and Sustainable Energy Reviews*, Vol. 31, pp. 762–777, 2014.
 23. Sideratos, G. and N. Hatziaargyriou, “An Advanced Statistical Method for Wind Power Forecasting”, *IEEE Transactions on Power Systems*, Vol. 22, No. 1, pp. 258–265, 2007.
 24. Nielsen, H., T. Nielsen, H. Madsen, M. Pindado and I. Marti, “Optimal Combination of Wind Power Forecasts”, *Wind Energy*, Vol. 10, pp. 471–482, 2007.
 25. Al-Yahyai, S., Y. Charabi, A. Al-Badi and A. Gastli, “Nested Ensemble NWP Approach for Wind Energy Assessment”, *Renewable Energy*, Vol. 37, No. 1, pp. 150–160, 2012.
 26. Louka, P., G. Galanis, N. Siebert, G. Kariniotakis, P. Katsafados, I. Pytharoulis and G. Kallos, “Improvements in Wind Speed Forecasts for Wind Power Prediction Purposes Using Kalman Filtering”, *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 96, No. 12, pp. 2348–2362, 2008.
 27. Nielsen, T. S., H. Madsen, H. A. Nielsen, P. Pinson, G. Kariniotakis, N. Siebert, I. T. Martí, M. Lange, U. Focken, L. von Bremen, G. A. Louka, G. Kallos and

- G. N. Galanis, “Short-term Wind Power Forecasting Using Advanced Statistical Methods”, *Proceedings of the European Wind Energy Conference and Exhibition*, 2006.
28. Kusiak, A., H. Zheng and Z. Song, “On-line Monitoring of Power Curves”, *Renewable Energy*, Vol. 34, No. 6, pp. 1487–1493, 2009.
29. Mangalova, E. and E. Agafonov, “Wind Power Forecasting Using the K-nearest Neighbors Algorithm”, *International Journal of Forecasting*, Vol. 30, No. 2, pp. 402–406, 2014.
30. Ye, X., Z. Lu, Y. Qiao, Y. Min and M. O’Malley, “Identification and Correction of Outliers in Wind Farm Time Series Power Data”, *IEEE Transactions on Power Systems*, Vol. 31, No. 6, pp. 4197–4205, 2016.
31. Qu, G., J. Mei and D. He, “Short-term Wind Power Forecasting Based on Numerical Weather Prediction Adjustment”, *11th IEEE International Conference on Industrial Informatics (INDIN)*, pp. 453–457, 2013.
32. Xu, Q., D. He, N. Zhang, C. Kang, Q. Xia, J. Bai and J. Huang, “A Short-Term Wind Power Forecasting Approach With Adjustment of Numerical Weather Prediction Input by Data Mining”, *IEEE Transactions on Sustainable Energy*, Vol. 6, pp. 1283–1291, 2015.
33. Silva, L., “A Feature Engineering Approach to Wind Power Forecasting”, *International Journal of Forecasting*, Vol. 30, No. 2, pp. 395–401, 2014.
34. Wei, H., W. sheng Wang and X. xuan Kao, “A Novel Approach to Ultra-short-term Wind Power Prediction Based on Feature Engineering and Informer”, *Energy Reports*, Vol. 9, pp. 1236–1250, 2023.
35. Andrade, R. J., Jose R.; Bessa, “Improving Renewable Energy Forecasting With a Grid of Numerical Weather Predictions”, *IEEE Transactions on Sustainable En-*

- ergy*, Vol. 8, No. 4, pp. 1571–1580, 2017.
36. Focken, U., M. Lange, K. Mönnich, H.-P. Waldl, H. G. Beyer and A. Luig, “Short-term Prediction of the Aggregated Power Output of Wind Farms—a statistical Analysis of the Reduction of the Prediction Error by Spatial Smoothing Effects”, *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 90, No. 3, pp. 231–246, 2002.
 37. Han, Y. and L. Chang, “A Study of the Reduction of the Regional Aggregated Wind Power Forecast Error by Spatial Smoothing Effects in the Maritimes Canada”, *IEEE Electrical Power & Energy Conference*, pp. 1–6, 2010.
 38. Higashiyama, K., Y. Fujimoto and Y. Hayashi, “Feature Extraction of NWP Data for Wind Power Forecasting Using 3D-Convolutional Neural Networks”, *Energy Procedia*, Vol. 155, pp. 350–358, 2018.
 39. Carcagnì, P., M. Del Coco, M. Leo and C. Distante, “Facial Expression Recognition and Histograms of Oriented Gradients: A Comprehensive Study”, *SpringerPlus*, Vol. 4, pp. 1–25, 2015.
 40. National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, “NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive”, .
 41. “Atmospheric Models from Météo-France”, <https://registry.opendata.aws/meteo-france-models>, accessed on December 9, 2022.
 42. Loh, W.-Y., “Classification and Regression Trees”, *WIREs Data Mining and Knowledge Discovery*, Vol. 1, No. 1, pp. 14–23, 2011.
 43. Rockafellar, R., *Convex Analysis*, Princeton Mathematical Series, Princeton University Press, 1970.
 44. Bergmeir, C. and J. M. Benítez, “On the Use of Cross-validation for Time Series

Predictor Evaluation”, *Information Sciences*, Vol. 191, pp. 192–213, 2012, data Mining for Software Trustworthiness.

45. Bergmeir, C., R. J. Hyndman and B. Koo, “A Note on the Validity of Cross-validation for Evaluating Autoregressive Time Series Prediction”, *Computational Statistics & Data Analysis*, Vol. 120, pp. 70–83, 2018.
46. Myers, R. H. and D. C. Montgomery, “A Tutorial on Generalized Linear Models”, *Journal of Quality Technology*, Vol. 29, No. 3, pp. 274–291, 1997.
47. Wright, M. and A. Ziegler, “ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R”, *Journal of Statistical Software*, Vol. 77, 2015.
48. Geurts, P., D. Ernst and L. Wehenkel, “Extremely Randomized Trees”, *Machine Learning*, Vol. 63, No. 1, pp. 3–42, 2006.
49. Friedman, M., “A Comparison of Alternative Tests of Significance for the Problem of m Rankings”, *The Annals of Mathematical Statistics*, Vol. 11, No. 1, pp. 86–92, 1940.
50. Nemenyi, P., *Distribution-free Multiple Comparisons*, Princeton University, 1963.
51. Demšar, J., “Statistical Comparisons of Classifiers over Multiple Data Sets”, *Journal of Machine Learning Research*, Vol. 7, No. 1, pp. 1–30, 2006.
52. Celenk, I., “Master’s Thesis”, 2023, <https://ilaydacelenk.github.io/MSThesis.html>, accessed on August 9, 2023.

APPENDIX A: WIND FARMS AND NWP INFORMATION

Table A.1. Wind Farms and the Number of NWP Coordinates.

WF	Code	Plant Name	Number of Coordinates	
			GFS	ARPEGE
1	40W000000000573X	BARES	16	25
2	40W000000000587M	YUNTDAG RES	16	25
3	40W000000000663W	TEPE RES	16	25
4	40W000000000694L	GÖKÇEDAĞ RES	16	25
5	40W000000000726Y	DUZOVA RES	16	25
6	40W000000000748O	MAZI RES	14	25
7	40W000000001581T	KOCADAG RES	14	25
8	40W000000002141F	ALİAĞA BERGAMA RES	16	25
9	40W000000003182X	ZİYARET RES	16	25
10	40W000000003302C	SOMA1-2 RES	16	25
11	40W0000000041237	SARES	16	25
12	40W000000004889N	SEYİTALİ RES	16	25
13	40W0000000051208	METRISTEPE RES	16	25
14	40W000000005540N	KARADAG RES	16	25
15	40W000000005541L	GERES	16	25
16	40W000000005611Q	POYRAZ RES	16	25
17	40W000000005874V	GUNAYDIN RES	16	25
18	40W000000005895N	KARABURUN RES	15	25
19	40W0000000065377	ZEYTİNELİ RES	14	25
20	40W000000006616B	EDINCIK RES	16	25
21	40W0000000066802	SENBUK RES	16	25
22	40W0000000070982	SALMAN RES	15	25
23	40W000000007158A	KARADERE RES	16	25

Table A.1. Wind Farms and the Number of NWP Coordinates. (cont.)

24	40W0000000071598	SADILLI RES	16	25
25	40W000000007879B	GOK2 RES	16	25
26	40W0000000082581	INCESU RES	16	25
27	40W0000000083189	CERCIKAYA RES	16	25
28	40W000000008459S	ORTAMANDIRA RES	16	25
29	40W0000000094059	SULOGLU RES	16	25
30	40W000000009921S	ULUBORLU RES	16	25
31	40W000000009922Q	YAYLAKOY RES	15	25
32	40W000000010013S	GERMIYAN RES	14	25
33	40W000000010041N	CATALTEPE RES	16	25
34	40W000000010062F	SARITEPE RES	16	25
35	40W000000010064B	ALACATI RES	14	25
36	40W000000010081B	URLA RES	14	25
37	40W0000000100829	DEMIRCILER RES	16	25
38	40W0000000101639	MORDOGAN RES	14	25
39	40W000000010403F	OVA RES	16	25
40	40W0000000104611	ELMALI RES	16	25
41	40W000000010501F	KIRKAGAC RES	16	25
42	40W0000000106215	KINIK RES	16	25
43	40W000000010903W	BERGRES RES	16	25
44	40W000000011525X	KUREKDAGI RES	16	25
45	40W000000011549J	KAROVA RES	16	25
46	40W000000011703Z	GOKDAG RES	16	25
47	40W000000015986E	ÖZBEK RES	16	25

Table A.2. Time Period of Each NWP Model.

WF	GFS	ARPEGE
1	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
2	01/01/2017 - 30/12/2022	09/08/2020 - 30/12/2022
3	01/10/2017 - 05/11/2022	09/08/2020 - 05/11/2022
4	01/01/2020 - 05/11/2022	09/08/2020 - 05/11/2022
5	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
6	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
7	01/01/2017 - 30/12/2022	09/08/2020 - 30/12/2022
8	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
9	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
10	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
11	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
12	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
13	01/01/2017 - 30/12/2022	09/08/2020 - 30/12/2022
14	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
15	01/04/2017 - 05/11/2022	09/08/2020 - 05/11/2022
16	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
17	20/09/2018 - 05/11/2022	09/08/2020 - 05/11/2022
18	01/01/2020 - 25/02/2023	09/08/2020 - 25/02/2023
19	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
20	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
21	21/01/2017 - 10/03/2023	09/08/2020 - 10/03/2023
22	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
23	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
24	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
25	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
26	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
27	01/01/2017 - 09/03/2023	09/08/2020 - 09/03/2023
28	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
29	01/01/2017 - 07/11/2022	09/08/2020 - 07/11/2022

Table A.2. Time Period of Each NWP Model. (cont.)

30	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
31	01/01/2017 - 25/02/2023	09/08/2020 - 25/02/2023
32	01/01/2017 - 25/02/2023	09/08/2020 - 25/02/2023
33	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
34	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
35	01/01/2017 - 25/02/2023	09/08/2020 - 25/02/2023
36	01/01/2017 - 25/02/2023	09/08/2020 - 25/02/2023
37	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
38	01/01/2017 - 25/02/2023	09/08/2020 - 25/02/2023
39	01/07/2017 - 05/11/2022	09/08/2020 - 05/11/2022
40	01/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
41	01/03/2017 - 05/11/2022	09/08/2020 - 05/11/2022
42	20/01/2017 - 05/11/2022	09/08/2020 - 05/11/2022
43	06/07/2017 - 05/11/2022	09/08/2020 - 05/11/2022
44	01/01/2020 - 25/02/2023	09/08/2020 - 25/02/2023
45	01/12/2017 - 05/11/2022	09/08/2020 - 05/11/2022
46	01/12/2017 - 05/11/2022	09/08/2020 - 05/11/2022
47	01/02/2021 - 05/11/2022	01/02/2021 - 05/11/2022

APPENDIX B: OVERALL RESULTS OF ALL WIND FARMS

Table B.1. WMAPEs of the GLM Models for Power Generation.

WF	Aw GFS .O	Aw ARPEGE .O	Aw GFS ARPEGE .O	Aw GFS	Aw ARPEGE	Aw GFS ARPEGE
1	32.428	29.341	29.032	32.582	29.407	29.047
2	41.351	34.244	34.329	41.561	34.485	34.343
3	29.318	26.059	24.654	29.375	26.355	24.568
4	55.877	39.948	38.988	56.237	39.552	39.200
5	49.671	39.888	39.534	49.922	39.699	39.877
6	30.567	32.850	29.982	30.639	33.004	30.179
7	28.820	28.670	27.068	29.145	28.532	27.053
8	37.307	32.168	31.494	37.299	32.411	31.930
9	42.892	33.889	33.959	42.697	34.102	34.134
10	23.159	22.304	20.879	23.176	22.494	20.942
11	28.562	31.776	28.218	28.585	31.703	28.555
12	46.212	38.424	37.885	46.718	38.842	38.689
13	52.831	44.626	44.558	53.310	45.056	44.921
14	34.802	43.214	34.840	35.055	43.142	34.949
15	35.195	24.661	24.541	35.510	24.786	24.745
16	30.537	31.050	28.666	30.570	30.356	28.733
17	39.727	34.024	33.825	40.537	34.747	34.280
18	33.976	30.477	30.302	34.367	31.023	31.180
19	27.985	28.639	26.704	27.981	28.935	26.894
20	35.118	31.091	30.416	35.683	31.452	30.552
21	69.449	69.875	68.525	68.566	70.510	68.321
22	41.122	36.258	35.593	40.983	35.627	35.924

Table B.1. WMAPEs of the GLM Models for Power Generation. (cont.)

23	29.824	28.415	27.425	30.014	28.274	27.483
24	26.647	28.109	25.806	26.816	28.513	26.171
25	37.536	29.820	29.429	38.045	30.125	29.725
26	42.625	44.419	40.527	42.266	45.399	41.459
27	54.814	44.872	43.067	54.808	44.937	43.971
28	30.252	27.854	26.814	30.414	27.859	26.982
29	32.519	32.292	30.740	32.822	32.472	31.593
30	37.332	41.772	37.154	37.428	42.048	36.930
31	32.801	31.472	31.038	32.892	31.976	31.317
32	33.425	36.092	32.442	33.526	36.212	32.939
33	29.817	30.127	28.067	29.952	30.379	28.579
34	60.927	55.958	52.381	60.522	55.947	52.518
35	34.201	38.037	34.414	34.508	38.185	34.184
36	29.188	25.000	24.430	29.648	25.263	24.770
37	50.203	46.762	45.286	50.802	47.034	45.565
38	32.422	24.924	24.637	32.272	25.110	24.989
39	46.814	49.521	46.943	46.923	49.385	46.588
40	50.593	43.124	40.811	51.362	44.016	41.998
41	21.978	20.828	19.578	22.145	20.880	19.750
42	32.047	25.868	25.400	32.113	26.029	25.588
43	25.622	28.956	24.154	25.537	28.913	24.367
44	39.566	38.838	37.565	39.866	39.462	37.632
45	40.877	39.304	37.366	40.817	39.386	37.265
46	54.368	53.999	51.849	54.860	54.296	52.381
47	54.656	50.359	49.105	54.897	50.499	49.512

Table B.2. WMAPEs of the GLM Models for Power Utilization with Aw.

WF	Aw GFS .O.U	Aw ARPEGE .O.U	Aw GFS ARPEGE .O.U	Aw GFS .U	Aw ARPEGE .U	Aw GFS ARPEGE .U
1	31.463	28.539	27.889	31.602	28.164	27.636
2	39.693	33.855	33.471	39.843	33.974	33.971
3	28.604	25.840	24.256	29.445	26.382	25.201
4	54.690	39.455	38.473	54.895	39.349	38.707
5	48.883	40.116	40.487	48.804	40.271	39.537
6	30.415	32.438	29.782	30.688	33.411	29.876
7	28.163	28.105	26.748	28.384	28.218	26.655
8	35.734	31.541	30.888	36.220	31.902	31.270
9	42.886	33.964	34.033	42.926	33.943	33.833
10	22.806	22.286	20.642	22.918	22.363	20.692
11	27.265	31.127	27.191	27.684	31.201	27.401
12	44.871	37.560	36.940	45.096	37.922	37.409
13	52.039	44.017	44.029	52.564	44.738	44.320
14	34.059	42.393	34.107	34.392	43.456	34.277
15	35.017	24.342	24.325	34.946	24.653	24.612
16	31.507	30.893	29.211	31.598	31.163	29.533
17	39.322	33.654	33.859	38.807	33.654	33.482
18	33.620	29.237	29.191	33.265	28.984	29.255
19	27.346	28.380	26.280	27.666	29.239	26.334
20	34.561	31.207	29.735	34.930	31.495	30.615
21	69.349	69.974	68.632	68.957	70.469	68.516
22	41.177	35.399	35.088	40.476	34.432	34.656
23	29.783	28.461	27.440	29.723	28.445	27.523
24	26.028	27.255	24.853	25.798	28.191	25.105
25	36.622	29.436	29.294	37.104	30.136	29.677
26	40.686	44.654	40.045	40.763	45.282	40.008

Table B.2. WMAPEs of the GLM Models for Power Utilization with Aw. (cont.)

27	54.144	43.207	41.883	54.109	43.316	41.975
28	30.051	27.771	26.672	30.251	27.993	26.806
29	31.539	31.147	29.615	32.100	33.057	30.305
30	35.570	41.136	35.287	36.020	40.956	35.817
31	32.597	30.994	30.438	32.500	30.604	30.464
32	32.466	36.497	32.244	32.730	36.920	32.587
33	29.945	30.075	27.791	30.078	30.909	28.623
34	57.306	55.153	50.867	57.676	55.374	50.764
35	35.205	38.934	35.111	35.772	39.352	35.170
36	29.038	25.351	24.740	29.107	25.406	24.775
37	48.658	46.349	44.438	48.520	46.390	44.195
38	31.906	25.213	24.334	31.390	24.253	24.148
39	45.867	49.391	46.067	45.532	48.159	45.428
40	49.760	41.509	40.185	50.465	42.070	41.343
41	20.595	20.569	19.091	20.992	20.755	19.274
42	31.385	26.076	25.582	32.056	26.300	25.843
43	24.855	28.900	23.921	25.093	29.091	24.201
44	37.459	37.888	35.617	37.773	38.284	36.081
45	40.260	39.433	36.759	40.670	39.759	37.498
46	53.721	54.538	51.788	54.757	54.954	52.095
47	53.549	49.795	47.881	53.480	49.901	48.044

Table B.3. WMAPEs of the Random Forest Models.

WF	Aw.RF	Lw.RF	Auv.RF	Luv.RF	Lwd.RF
1	27.944	25.569	25.804	22.998	23.753
2	33.806	20.131	25.109	19.810	19.983
3	24.873	23.026	23.499	22.461	22.414
4	37.957	24.489	27.078	23.821	24.113
5	40.078	27.380	32.300	26.587	26.731
6	30.177	25.567	29.177	25.310	25.257
7	26.953	24.715	25.975	24.388	24.469
8	31.130	20.377	23.658	20.031	19.911
9	33.962	24.607	28.993	23.769	24.263
10	21.204	17.876	17.957	16.893	16.946
11	27.407	20.343	24.367	19.984	20.116
12	37.720	27.350	31.244	27.301	26.947
13	45.005	37.194	40.429	35.172	36.045
14	34.219	30.038	32.583	29.030	29.589
15	24.474	18.306	20.603	18.379	18.294
16	29.298	23.590	26.719	22.841	23.095
17	34.274	27.454	31.439	27.087	26.834
18	29.556	20.408	24.369	19.890	19.918
19	26.669	22.327	25.267	22.377	22.213
20	30.286	27.640	28.902	27.169	27.120
21	65.749	28.209	28.944	23.497	24.231
22	35.160	27.423	29.028	26.724	26.792
23	27.482	24.976	26.911	25.200	24.818
24	25.262	23.419	23.877	23.101	23.167
25	29.438	21.562	24.889	21.037	21.359
26	40.848	31.324	31.128	28.918	29.100
27	41.779	29.676	33.920	29.090	29.238
28	27.267	24.483	26.411	24.317	24.283
29	30.455	26.482	28.197	26.021	26.081

Table B.3. WMAPEs of the Random Forest Models. (cont.)

30	36.087	29.233	30.738	27.703	27.665
31	31.045	23.439	26.650	22.801	23.020
32	32.372	28.338	30.377	27.620	27.800
33	27.899	24.457	26.256	23.792	23.698
34	51.069	29.803	29.593	28.425	28.982
35	34.270	30.471	31.753	29.296	29.695
36	24.896	21.266	22.130	21.020	21.048
37	44.879	36.351	38.519	35.324	35.937
38	25.131	20.791	21.863	20.143	20.283
39	45.893	37.978	40.952	36.930	37.023
40	39.200	27.007	28.982	25.360	25.692
41	19.244	18.368	18.307	17.865	17.886
42	26.104	19.974	23.708	19.497	19.653
43	24.150	18.902	22.660	18.720	18.672
44	36.637	31.598	32.837	30.567	30.729
45	37.250	31.615	33.995	30.304	31.011
46	52.730	48.616	50.670	46.115	46.698
47	48.412	30.954	33.710	30.016	30.535

Table B.4. WMAPEs of the GLM Models.

WF	AwB	AHOG	Lw	LwB	LHOG
1	26.415	26.244	25.119	24.194	23.727
2	25.349	25.719	20.007	19.836	19.657
3	23.388	23.607	22.273	22.353	21.990
4	27.816	27.703	24.220	25.022	24.855
5	33.582	33.084	27.671	27.441	26.894
6	28.741	29.101	25.902	25.781	25.256
7	26.198	26.076	24.742	24.906	24.546
8	24.137	24.002	20.533	20.429	20.392
9	30.610	30.015	24.378	24.251	24.235
10	19.175	19.049	18.145	17.908	17.408
11	25.313	25.242	20.358	20.286	20.112
12	30.935	31.010	26.879	26.832	26.527
13	39.835	41.092	37.369	35.780	35.154
14	32.468	32.949	30.891	30.054	29.560
15	21.852	21.390	18.051	19.138	18.573
16	27.139	26.770	23.563	22.812	22.494
17	31.729	30.806	26.936	27.615	27.177
18	24.832	24.644	19.944	19.975	19.426
19	25.735	25.328	22.117	22.580	22.287
20	28.678	28.974	27.476	27.955	27.317
21	33.335	29.141	28.888	24.458	23.913
22	29.847	28.941	27.128	27.360	27.948
23	26.903	26.542	25.193	27.161	26.858
24	23.775	24.120	22.602	23.176	22.716
25	27.385	26.092	21.201	21.714	21.022
26	31.415	31.235	31.542	31.085	30.224
27	37.727	33.469	29.914	29.706	29.052
28	26.274	26.451	24.382	24.832	24.799
29	28.893	28.225	25.758	26.045	25.800

Table B.4. WMAPEs of the GLM Models. (cont.)

30	30.652	31.875	30.134	30.074	29.737
31	27.316	27.091	23.418	23.240	22.836
32	31.430	30.783	28.760	28.012	27.453
33	26.813	26.257	24.107	23.500	23.149
34	30.658	30.064	29.548	29.018	28.885
35	31.763	31.557	31.374	29.842	28.399
36	23.519	22.788	21.584	22.076	21.527
37	38.180	39.277	36.833	37.535	37.058
38	22.984	21.826	20.035	19.940	19.423
39	43.222	41.721	37.829	38.480	37.678
40	29.848	30.355	28.598	27.207	25.671
41	18.478	19.375	18.104	18.359	17.935
42	24.809	23.909	19.818	19.513	19.714
43	22.803	23.059	19.021	19.327	18.930
44	33.300	33.290	31.216	30.615	29.916
45	34.065	34.993	31.638	31.267	30.745
46	48.450	50.145	48.248	49.883	49.952
47	36.069	33.108	31.047	31.179	31.238

Table B.5. WMAPEs of Ensemble Approaches.

WF	E.Luv LHOG	E LHOG LwdHOG	E.Lw LHOG LwdHOG	LHOG LwdHOG	Lw.LHOG LwdHOG	Luv LHOG RF
1	22.998	23.588	23.588	23.558	23.285	23.174
2	19.810	19.599	19.599	19.608	19.468	19.818
3	22.461	21.902	21.902	21.998	21.871	22.555
4	23.821	24.800	24.800	24.717	24.357	23.803
5	26.587	26.836	26.812	26.848	26.830	26.471
6	25.310	25.117	25.117	25.474	25.098	25.434
7	24.388	24.465	24.465	24.708	24.484	24.272
8	20.031	20.106	20.106	20.306	19.769	20.023
9	23.769	24.169	24.169	24.047	24.059	23.749
10	16.893	17.056	17.056	17.185	16.954	16.880
11	19.984	20.020	19.999	20.075	19.896	19.998
12	27.301	26.380	26.380	26.604	26.295	27.169
13	35.172	35.042	35.042	35.361	35.555	35.164
14	29.030	29.570	29.570	29.779	29.742	29.001
15	18.379	18.337	18.273	18.623	18.291	18.472
16	22.841	22.490	22.490	22.510	22.398	22.836
17	27.087	26.995	26.995	27.190	26.749	27.099
18	19.890	19.440	19.440	19.507	19.295	19.892
19	22.377	22.333	22.333	22.355	22.079	22.428
20	27.169	27.148	27.148	27.374	27.003	27.276
21	23.497	23.991	23.998	24.022	24.093	23.612
22	26.724	27.782	27.581	27.986	27.096	26.795
23	25.200	26.417	26.417	26.329	25.341	25.285
24	23.101	22.716	22.716	22.805	22.479	23.164
25	21.037	21.020	20.895	20.951	20.854	20.971
26	28.918	30.224	30.224	30.175	29.435	29.082
27	29.090	29.052	29.052	29.271	28.735	29.007

Table B.5. WMAPEs of Ensemble Approaches. (cont.)

28	24.317	24.674	24.674	24.936	24.261	24.404
29	26.021	25.686	25.623	26.034	25.357	26.076
30	27.703	29.737	29.737	29.605	28.749	27.936
31	22.801	22.732	22.732	23.042	22.855	22.721
32	27.620	27.463	27.463	27.415	27.292	27.569
33	23.792	23.515	23.515	23.509	23.305	23.856
34	28.425	28.741	28.692	28.777	28.542	28.351
35	29.296	28.444	28.444	28.748	28.747	29.072
36	21.020	21.446	21.446	21.508	21.415	21.060
37	35.324	37.024	37.024	36.607	36.049	35.424
38	20.143	19.485	19.476	19.664	19.642	20.149
39	36.930	37.536	37.366	37.152	37.022	37.000
40	25.360	25.622	25.622	27.672	27.148	25.397
41	17.865	17.794	17.794	18.028	17.832	17.841
42	19.497	19.512	19.512	20.324	20.084	19.571
43	18.720	18.785	18.785	18.912	18.729	18.837
44	30.567	29.880	29.880	30.062	29.981	30.513
45	30.304	30.630	30.630	30.843	30.423	30.292
46	46.115	49.471	49.471	48.893	46.881	46.315
47	30.016	31.216	31.216	31.265	30.663	30.007

Table B.6. WMAPEs of Ensemble Approaches on High Production.

WF	E.Luv LHOG	E LHOG LwdHOG	E.Lw LHOG LwdHOG	LHOG LwdHOG	Lw LHOG LwdHOG	Luv LHOG RF
1	8.733	8.631	8.631	8.598	8.480	9.004
2	8.167	6.817	6.817	6.468	6.419	8.217
3	10.374	9.502	9.502	9.006	9.049	10.411
4	7.015	6.157	6.157	6.397	6.191	7.071
5	14.739	13.379	13.404	13.286	13.197	14.581
6	9.955	8.560	8.560	8.329	8.381	10.241
7	7.739	7.649	7.649	7.633	7.570	7.740
8	7.278	7.505	7.505	7.049	6.872	7.421
9	11.481	10.086	10.086	9.996	9.993	11.448
10	7.261	6.892	6.892	6.310	6.730	7.139
11	9.220	7.851	7.862	7.929	8.087	9.024
12	11.484	8.660	8.660	8.216	8.422	11.293
13	18.294	15.825	15.825	14.217	14.292	18.155
14	13.853	12.408	12.408	12.291	12.512	14.067
15	9.446	8.526	8.528	8.391	8.347	9.578
16	8.922	6.874	6.874	6.587	6.492	9.062
17	8.893	8.593	8.593	8.140	8.189	8.959
18	6.459	5.729	5.729	5.374	5.508	6.489
19	9.366	8.008	8.008	7.696	7.518	9.347
20	13.034	10.879	10.879	11.279	11.288	13.207
21	9.521	9.118	9.162	8.865	8.821	9.610
22	12.926	12.077	11.982	11.938	11.486	13.194
23	12.456	11.261	11.261	11.216	11.042	12.598
24	8.288	7.048	7.048	7.235	7.170	8.473
25	12.545	11.187	11.298	10.948	11.478	12.343
26	15.367	13.108	13.108	13.280	13.860	16.003
27	13.241	11.962	11.962	11.876	11.887	13.451

Table B.6. WMAPEs of Ensemble Approaches on High Production. (cont.)

28	11.169	9.749	9.749	10.005	9.941	11.431
29	11.885	10.509	10.535	10.618	10.364	12.091
30	17.127	15.311	15.311	15.913	16.080	17.690
31	11.044	9.814	9.814	10.048	9.865	11.204
32	8.454	7.065	7.065	6.552	6.491	8.471
33	8.794	7.560	7.560	7.211	7.188	8.953
34	7.882	7.169	7.171	7.069	7.005	8.042
35	11.447	10.326	10.326	10.300	10.370	11.576
36	6.238	6.181	6.181	5.983	5.872	6.286
37	20.425	16.504	16.504	17.057	16.382	21.171
38	8.012	7.067	7.067	6.945	6.849	8.099
39	18.341	17.683	17.670	16.763	17.296	18.473
40	11.972	10.087	10.087	11.576	11.173	12.190
41	8.836	8.334	8.334	8.514	8.240	8.904
42	9.120	7.788	7.788	8.227	7.699	9.484
43	8.489	8.154	8.154	7.737	7.846	8.693
44	11.344	9.167	9.167	9.667	9.365	11.427
45	15.352	14.831	14.831	13.958	13.650	15.269
46	21.979	17.805	17.805	20.392	17.974	23.430
47	13.308	11.899	11.899	11.731	12.226	13.757

Table B.7. WMAPEs of Ensemble Approaches on Moderate Production.

WF	E.Luv LHOG	E LHOG LwdHOG	E.Lw LHOG LwdHOG	LHOG LwdHOG	Lw LHOG LwdHOG	Luv LHOG RF
1	26.744	27.563	27.563	27.577	27.331	26.839
2	21.885	21.832	21.832	22.108	22.037	21.899
3	25.935	25.141	25.141	25.676	25.692	26.037
4	22.013	23.005	23.005	22.929	22.801	22.059
5	23.867	24.452	24.420	24.579	24.519	23.787
6	27.298	27.241	27.241	27.881	27.450	27.374
7	25.499	25.502	25.502	25.977	25.815	25.370
8	23.042	22.912	22.912	23.323	22.832	22.899
9	21.418	21.846	21.846	21.746	21.873	21.453
10	17.321	17.317	17.317	17.698	17.428	17.299
11	20.620	20.594	20.577	20.871	20.722	20.759
12	31.141	31.071	31.071	31.617	31.208	31.102
13	32.700	32.348	32.348	33.459	33.795	32.707
14	31.900	32.573	32.573	33.078	33.013	31.764
15	21.542	21.415	21.344	21.890	21.612	21.626
16	24.514	24.665	24.665	24.733	24.733	24.457
17	29.132	28.993	28.993	29.458	28.943	29.151
18	21.219	20.691	20.691	20.968	20.833	21.229
19	24.923	24.994	24.994	25.403	25.227	24.970
20	27.481	27.628	27.628	28.081	27.682	27.588
21	28.707	29.742	29.585	30.030	30.056	28.913
22	28.713	30.229	30.037	30.545	29.934	28.755
23	25.578	26.910	26.910	26.885	26.023	25.707
24	24.659	24.485	24.485	24.636	24.469	24.669
25	20.231	20.207	20.109	20.229	20.187	20.186
26	28.967	30.431	30.431	30.491	29.938	29.103
27	26.712	27.226	27.226	27.264	27.071	26.669

Table B.7. WMAPEs of Ensemble Approaches on Moderate Production. (cont.)

28	27.603	28.066	28.066	28.437	27.923	27.589
29	27.588	27.287	27.216	27.705	27.231	27.631
30	26.014	27.654	27.654	27.819	27.034	26.197
31	23.190	23.201	23.201	23.590	23.505	23.098
32	32.737	32.950	32.950	33.456	33.423	32.660
33	29.158	29.241	29.241	29.512	29.506	29.151
34	27.598	28.139	28.090	28.242	28.018	27.513
35	31.507	30.886	30.886	31.151	31.248	31.249
36	22.084	22.400	22.400	22.585	22.745	22.148
37	29.600	31.209	31.209	30.937	30.741	29.760
38	21.469	21.107	21.100	21.407	21.447	21.412
39	36.562	37.497	37.332	37.490	37.145	36.562
40	30.024	30.964	30.964	31.529	31.535	30.059
41	19.876	19.810	19.810	19.997	20.003	19.819
42	20.085	20.300	20.300	20.975	21.143	20.117
43	19.427	19.403	19.403	19.742	19.577	19.444
44	29.432	29.180	29.180	29.345	29.543	29.402
45	28.671	28.831	28.831	29.380	29.082	28.664
46	38.251	39.452	39.452	39.433	38.705	38.553
47	28.934	30.967	30.967	31.023	30.218	28.913

Table B.8. WMAPEs of Ensemble Approaches on Low Production.

WF	E.Luv LHOG	E LHOG LwdHOG	E.Lw LHOG LwdHOG	LHOG LwdHOG	Lw LHOG LwdHOG	Luv LHOG RF
1	464.632	480.067	480.067	461.491	463.592	467.656
2	281.875	314.314	314.314	275.327	281.875	278.765
3	262.007	299.016	299.016	262.922	262.007	264.014
4	348.692	364.480	364.480	336.465	348.692	341.782
5	522.247	506.560	506.578	500.495	521.150	517.483
6	391.116	422.027	422.027	400.318	391.116	390.809
7	282.351	293.374	293.374	266.718	282.351	281.022
8	206.656	223.823	223.823	206.348	206.656	215.694
9	710.306	722.447	722.447	689.885	710.306	697.079
10	204.360	255.657	255.657	211.646	204.360	209.495
11	222.767	271.285	269.755	222.848	222.767	209.614
12	191.809	191.654	191.654	184.169	191.809	188.409
13	404.920	445.051	445.051	377.232	404.252	404.031
14	613.925	728.731	728.731	691.758	613.925	613.925
15	290.935	381.820	377.040	354.030	290.935	289.060
16	206.333	199.062	199.062	190.448	206.333	207.896
17	300.885	310.161	310.161	294.810	290.521	288.635
18	191.010	203.123	203.123	180.114	191.010	189.757
19	324.866	375.869	375.869	317.000	324.866	329.094
20	293.526	322.919	322.919	293.561	293.526	292.131
21	504.197	516.844	529.128	524.992	504.197	497.786
22	470.571	510.344	503.644	456.615	470.571	464.433
23	457.380	531.391	531.391	455.778	457.380	446.068
24	476.642	483.730	483.730	432.081	476.642	479.324
25	304.229	329.507	320.750	297.441	304.229	303.337
26	317.163	382.440	382.440	324.183	317.163	309.475
27	715.274	658.990	658.990	629.381	715.274	703.115

Table B.8. WMAPEs of Ensemble Approaches on Low Production. (cont.)

28	480.762	609.633	609.633	504.729	480.762	486.060
29	391.656	428.347	425.813	386.631	391.656	387.169
30	299.616	373.646	373.646	316.869	299.616	303.767
31	163.237	173.247	173.247	157.829	163.237	160.614
32	359.997	378.211	378.211	343.565	356.125	355.619
33	428.059	451.983	451.983	402.228	425.527	427.617
34	306.426	301.280	300.735	295.146	306.426	304.654
35	418.345	404.840	404.840	408.215	418.345	409.601
36	294.956	323.861	323.861	280.721	294.956	291.460
37	503.927	535.693	535.693	495.129	503.927	497.131
38	298.437	266.272	265.776	254.953	298.437	304.058
39	493.950	499.420	495.187	483.932	493.950	497.637
40	320.544	372.403	372.403	449.025	320.544	313.057
41	294.527	325.415	325.415	305.331	294.527	294.193
42	245.610	255.225	255.225	232.750	245.610	243.347
43	207.835	232.568	232.568	205.635	207.835	220.016
44	492.527	474.697	474.697	446.582	492.527	487.604
45	488.476	524.953	524.953	476.799	488.476	489.068
46	642.690	814.189	814.189	684.866	636.051	624.374
47	505.810	478.044	478.044	478.557	505.810	496.947

Table B.9. WMAPEs of Luv.RF on the Production Categories.

WF	High Production	Moderate Production	Low Production
1	8.733	26.744	474.708
2	8.167	21.885	290.711
3	10.374	25.935	284.864
4	7.015	22.013	360.869
5	14.739	23.867	494.408
6	9.955	27.298	410.751
7	7.739	25.499	276.618
8	7.278	23.042	227.206
9	11.481	21.418	716.817
10	7.261	17.321	222.452
11	9.220	20.620	236.849
12	11.484	31.141	191.014
13	18.294	32.700	389.031
14	13.853	31.900	704.944
15	9.446	21.542	388.342
16	8.922	24.514	202.105
17	8.893	29.132	301.149
18	6.459	21.219	194.242
19	9.366	24.923	336.952
20	13.034	27.481	299.377
21	9.521	28.707	513.024
22	12.926	28.713	513.759
23	12.456	25.578	516.304
24	8.288	24.659	469.385
25	12.545	20.231	317.601
26	15.367	28.967	368.641
27	13.241	26.712	692.548
28	11.169	27.603	588.718
29	11.885	27.588	429.580

Table B.9. WMAPEs of Luv.RF on the Production Categories. (cont.)

30	17.127	26.014	332.804
31	11.044	23.190	167.384
32	8.454	32.737	353.034
33	8.794	29.158	441.070
34	7.882	27.598	298.001
35	11.447	31.507	419.910
36	6.238	22.084	314.685
37	20.425	29.600	520.928
38	8.012	21.469	260.846
39	18.341	36.562	483.461
40	11.972	30.024	485.949
41	8.836	19.876	338.463
42	9.120	20.085	289.509
43	8.489	19.427	216.820
44	11.344	29.432	470.853
45	15.352	28.671	495.852
46	21.979	38.251	764.229
47	13.308	28.934	482.172

APPENDIX C: VALIDATION OF REPRESENTATION PARAMETERS

Table C.1. AHOG Sensitivity Results on Training Set.

WF	6-Bins	9-Bins	12-Bins	18-Bins	24-Bins	36-Bins	Selected Bin
1	27.969	28.083	27.988	27.913	27.759	28.056	6
2	25.881	25.712	25.961	25.475	25.322	25.914	9
3	27.657	27.506	27.552	27.003	27.499	27.791	9
4	26.812	26.467	26.125	26.750	26.486	26.667	12
5	36.605	36.498	36.318	35.910	35.950	36.261	18
6	30.427	30.086	29.819	29.769	30.374	30.417	18
7	26.689	25.702	25.975	26.350	26.081	26.404	9
8	25.133	25.623	25.481	25.808	25.410	25.256	6
9	26.943	26.512	26.446	26.700	26.666	26.765	12
10	20.195	19.609	20.280	20.620	20.494	20.305	9
11	26.560	26.622	26.694	26.547	26.771	26.681	6
12	31.830	31.116	31.728	31.008	30.728	31.572	9
13	39.418	39.042	38.986	38.389	38.791	38.408	18
14	34.347	34.368	34.079	33.947	35.135	34.416	6
15	23.505	23.942	23.939	23.735	23.449	24.324	6
16	26.968	27.445	27.029	27.547	27.245	27.901	6
17	32.544	33.170	32.238	32.346	32.914	32.470	6
18	27.000	26.597	25.819	26.536	26.182	26.446	12
19	27.225	26.026	25.544	25.821	26.228	26.649	12
20	29.468	29.179	28.821	29.164	29.138	29.259	12
21	33.496	34.042	34.849	34.508	34.491	34.456	6
22	32.457	32.172	31.571	31.846	31.861	32.646	12
23	28.919	29.266	28.979	28.917	28.990	28.609	6
24	27.724	27.369	27.268	27.349	27.656	27.428	12
25	28.765	28.866	28.965	29.098	29.286	29.133	6

Table C.1. AHOG Sensitivity Results on Training Set. (cont.)

26	35.732	35.177	35.921	35.341	35.257	35.616	9
27	32.537	32.780	32.585	32.626	32.504	32.607	6
28	28.478	28.406	28.358	28.214	28.442	28.221	18
29	32.471	32.503	32.270	32.810	32.644	31.981	6
30	35.243	34.082	34.058	33.702	33.995	33.388	18
31	30.856	30.418	30.217	30.686	30.785	30.738	12
32	30.769	31.142	31.187	30.461	30.548	31.111	6
33	27.494	27.784	27.732	27.880	27.184	27.701	6
34	27.407	26.746	26.376	26.228	26.815	26.220	18
35	32.229	31.959	31.924	31.815	31.972	31.767	18
36	22.792	22.886	22.824	22.545	22.797	22.634	6
37	38.025	37.508	37.119	37.506	37.176	37.099	12
38	25.752	25.995	26.007	25.930	26.115	26.224	6
39	47.017	47.030	47.856	47.375	45.893	44.962	6
40	33.353	32.699	32.219	32.526	32.132	32.952	12
41	21.488	21.213	21.045	21.484	21.406	21.024	12
42	28.219	27.874	28.199	27.932	28.009	28.709	9
43	24.541	24.733	24.534	25.030	25.095	25.474	6
44	31.288	31.753	31.369	31.440	31.289	31.566	6
45	36.287	35.326	35.172	35.063	35.196	34.802	18
46	52.027	51.568	52.137	51.405	52.305	51.445	9
47	31.400	30.895	30.260	30.326	30.977	30.481	12

Table C.2. LHOG Sensitivity Results on Training Set.

WF	6-Bins	9-Bins	12-Bins	18-Bins	24-Bins	36-Bins	Selected Bin
1	21.882	21.860	21.349	21.749	21.854	21.809	12
2	17.730	17.292	17.356	17.099	17.329	17.295	9
3	23.577	23.796	23.384	23.948	23.801	23.734	6
4	20.456	20.088	20.181	20.279	20.240	20.481	9
5	26.855	26.478	26.584	26.936	27.371	30.546	9
6	23.969	23.379	22.964	23.017	23.134	22.847	12
7	21.464	21.530	21.464	21.127	21.528	21.477	6
8	18.092	17.919	18.460	18.434	18.541	18.405	9
9	18.877	18.974	18.700	18.872	19.011	20.088	6
10	16.402	15.821	15.876	15.476	15.873	15.518	9
11	19.585	19.048	19.460	18.515	18.793	18.659	9
12	22.348	21.971	22.218	22.378	22.512	23.310	9
13	31.064	30.767	30.997	30.655	30.739	30.428	9
14	28.596	28.247	28.991	27.736	27.324	29.254	9
15	17.888	17.883	17.874	17.573	18.112	17.431	18
16	20.160	20.932	20.346	20.486	20.948	20.502	6
17	23.544	23.267	22.880	22.479	22.960	23.536	18
18	18.421	18.289	18.361	18.128	18.060	17.663	9
19	20.959	20.996	21.558	20.621	20.957	20.800	6
20	23.926	23.484	23.269	23.069	23.494	23.569	18
21	22.858	22.221	22.613	22.942	23.284	23.915	9
22	25.474	24.994	24.528	24.495	25.245	24.882	18
23	23.979	23.625	23.562	23.356	22.859	22.535	36
24	22.857	22.783	22.756	23.009	22.672	22.526	12
25	20.635	20.398	20.543	20.388	20.229	21.069	9
26	28.813	28.596	28.282	27.609	27.314	27.958	24
27	25.571	24.786	24.873	24.854	24.802	26.067	9
28	23.266	23.192	23.078	22.730	22.737	23.072	18

Table C.2. LHOGL Sensitivity Results on Training Set. (cont.)

29	26.462	26.052	25.922	25.305	25.628	25.429	18
30	26.553	26.205	26.141	24.911	24.987	25.635	18
31	22.609	22.120	22.129	21.619	21.530	21.914	9
32	25.056	25.193	24.530	24.341	24.712	24.257	6
33	22.559	23.038	21.896	22.290	22.050	21.572	6
34	22.029	22.134	21.920	22.084	21.867	21.725	6
35	26.813	27.092	27.012	26.828	27.070	26.933	6
36	19.210	19.131	18.622	18.737	18.405	18.437	12
37	29.342	29.017	29.006	28.549	29.282	28.764	18
38	19.916	19.929	19.474	19.791	19.547	20.150	6
39	32.630	32.524	32.526	32.150	31.170	33.331	9
40	24.419	23.643	23.733	23.828	23.705	24.549	9
41	17.173	17.576	17.795	17.001	17.023	16.324	6
42	18.573	19.182	18.796	19.818	19.906	20.036	6
43	17.913	18.631	18.006	17.970	18.017	17.535	6
44	24.939	26.004	25.271	25.000	25.108	24.681	6
45	28.179	28.393	27.911	28.511	28.148	28.130	6
46	44.247	44.069	44.019	43.928	43.641	43.562	36
47	25.424	25.282	25.413	25.620	25.089	25.419	9

Table C.3. LuvHOG Sensitivity Results on Training Set.

WF	Depth	Number of Bins	Train WMAPE	Test WMAPE
1	4	3	24.848	25.411
2	4	3	20.925	22.749
3	3	3	25.305	23.437
4	4	3	34.140	39.177
5	3	9	31.412	30.582
6	4	3	26.583	27.880
7	4	3	23.803	26.258
8	4	3	21.589	23.174
9	4	3	24.935	30.750
10	4	3	17.807	18.720
11	4	3	21.494	22.892
12	4	3	25.946	29.827
13	4	3	35.923	40.080
14	4	3	30.896	31.897
15	4	3	20.345	20.678
16	4	3	23.984	26.096
17	4	3	28.038	31.393
18	4	3	21.666	23.396
19	4	3	22.659	24.553
20	4	3	25.684	28.260
21	3	3	26.457	27.563
22	4	3	27.563	30.133
23	3	3	26.228	26.990
24	3	3	25.322	23.607
25	4	3	25.049	25.220
26	4	3	33.924	32.190
27	4	3	34.820	39.204
28	4	3	26.215	26.893

Table C.3. LuvHOG Sensitivity Results on Training Set. (cont.)

29	3	3	29.030	27.619
30	4	3	31.980	31.647
31	4	3	25.797	26.135
32	4	3	27.983	30.662
33	3	3	24.854	25.797
34	3	3	34.458	39.485
35	3	3	30.449	31.856
36	4	3	20.868	22.741
37	4	3	43.146	48.197
38	4	3	21.929	21.903
39	4	3	39.378	41.509
40	4	3	27.964	29.435
41	3	3	19.452	18.941
42	4	3	23.806	23.493
43	3	3	21.557	21.768
44	3	3	30.075	33.683
45	3	6	32.088	34.392
46	3	3	49.030	50.242
47	4	3	36.409	42.176

Table C.4. LwdHOG Sensitivity Results on Training Set.

WF	Depth	Number of Bins	Train WMAPE	Test WMAPE
1	4	3	24.233	24.838
2	4	3	20.195	22.092
3	3	3	24.801	22.902
4	3	3	23.742	27.041
5	4	3	29.453	29.319
6	4	3	25.538	27.264
7	4	3	23.022	25.863
8	4	3	20.671	21.639
9	4	3	24.281	29.259
10	4	3	17.089	18.347
11	4	3	20.981	21.725
12	4	3	25.024	29.054
13	4	3	34.789	38.827
14	4	3	30.736	31.913
15	4	3	19.464	19.609
16	3	3	23.595	25.184
17	4	3	25.352	29.016
18	4	3	20.986	22.513
19	4	3	21.864	23.551
20	4	3	25.367	28.041
21	4	3	26.979	29.029
22	4	3	27.763	29.726
23	4	3	24.890	26.178
24	3	3	24.926	23.341
25	4	3	22.590	23.523
26	3	6	31.590	31.327
27	3	3	31.512	34.110
28	4	3	25.011	25.958

Table C.4. LwdHOG Sensitivity Results on Training Set. (cont.)

29	4	3	27.642	26.937
30	4	3	30.072	30.723
31	4	3	24.473	25.150
32	4	3	27.323	29.837
33	4	3	23.451	25.486
34	4	3	25.286	31.856
35	3	6	29.415	31.095
36	4	3	19.980	22.538
37	3	3	34.154	38.105
38	4	3	21.968	22.032
39	4	3	39.060	41.057
40	4	3	27.994	29.397
41	3	6	18.244	18.746
42	4	3	22.036	22.715
43	4	3	19.884	20.079
44	4	3	28.244	32.349
45	4	3	31.050	33.783
46	3	3	47.343	47.774
47	3	3	32.875	39.232