A NEW DATA MINING BASED UPSCALING APPROACH FOR REGIONAL WIND POWER FORECASTING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MEHMET BARIŞ ÖZKAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
COMPUTER ENGINEERING

FEBRUARY 2021

Approval of the thesis:

A NEW DATA MINING BASED UPSCALING APPROACH FOR REGIONAL WIND POWER FORECASTING

submitted by **MEHMET BARIŞ ÖZKAN** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Computer Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** ________________

Prof. Dr. Halit Oğuztüzün
Head of Department, **Computer Engineering** ________________

Prof. Dr. Pınar Karagöz
Supervisor, **Computer Engineering, METU** ________________

Examining Committee Members:

Prof. Dr. Suat Özdemir
Computer Engineering, Hacettepe University ________________

Prof. Dr. Pınar Karagöz
Computer Engineering, METU ________________

Prof. Dr. İbrahim Körpeoğlu
Computer Engineering, Bilkent University ________________

Prof. Dr. Cem İyigün
Industrial Engineering, METU ________________

Assoc. Prof. Dr. Şeyda Ertekin
Computer Engineering, METU ________________

Date: 08.02.2021

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Mehmet Barış Özkan

Signature :

ABSTRACT

A NEW DATA MINING BASED UPSCALING APPROACH FOR REGIONAL WIND POWER FORECASTING

Özkan, Mehmet Barış
Ph.D., Department of Computer Engineering
Supervisor: Prof. Dr. Pınar Karagöz

February 2021, 87 pages

Together with the increasing need for energy, the importance of renewable energy sources has been increasing day by day. Although wind is one of the most important alternative energy sources due to its high potential, it is not a stable source since it depends on the weather conditions. So, in order to include the power produced by the wind into electricity grid with planned manner, it must be predicted accurately beforehand. To produce a reliable wind power forecast, getting a Wind Power Plant's (WPP) power data in real time and constructing the model with past production values is an expected and optimal situation. However, this situation cannot be applicable for all WPPs in the country due to the difficulties on getting the power data of WPP in real time. Hence there is a need for accurate upscaling algorithm for generating the power forecast of such WPPs and producing a regional power forecast for a given region. These forecasts are especially so crucial for the management and planning of the electricity grid. It is very important that the system operators who manage the energy flow in the country use their energy resources correctly by using these power forecasts day in advance.

In this thesis, six different statistical based models are developed to solve the regional wind power estimation problem and the results obtained are compared with some well known statistical based machine learning models in the literature such as ANN and SVM. The most important contribution of the proposed method is that it produces regional power forecasts while also generating power estimates for wind power plants in the system for which we do not have their historical power data. Another advantage of the method is that the wind potential of a candidate point where a wind farm is planned to be established can be determined by this method. In this thesis, this feature of the model is tested for 16 different candidate points from four different cities of country. In addition, regional forecasts results are tested for 9 different load distribution regions and performance results of the models are analyzed.

Keywords: Regional Wind Power Forecasting, Data Mining, Numerical Weather Prediction, Upscaling

ÖZ

YENİ BİR VERİ MADENCİLİĞİ TABANLI BÖLGESEL RÜZGAR GÜCÜ TAHMİNİ İÇİN YUKARI ÖLÇEKLEME ALGORİTMASI

Özkan, Mehmet Barış
Doktora, Bilgisayar Mühendisliği Bölümü
Tez Yöneticisi: Prof. Dr. Pınar Karagöz

Şubat 2021 , 87 sayfa

Enerjiye olan ihtiyacın artmasıyla birlikte yenilenebilir enerji kaynaklarının önemi gün geçtikçe artmaya başlamıştır. Rüzgâr, yüksek potansiyeli ve verimliliği ile önemli bir alternatif enerji kaynağı olmasına rağmen hava olaylarına bağlı olması sebebiyle istikrarlı değildir. Bu özelliğinden dolayı rüzgar gücünün elektrik gridine planlı bir şekilde dahil olması için iyi bir şekilde önceden tahmin edilmesi gereklidir. İyi bir rüzgâr gücü tahmini için anlık olarak santrale ait güç verisinin alınıp tahmin modelinin geçmiş güç verisinden direkt oluşturulması istenilen durumdur. Fakat bu durum ülkedeki tüm rüzgar santraller için güç verisini anlık olarak almanın getirdiği bazı zorluklardan ve maliyetlerden dolayı uygulanabilir olmayabilir ve bu sebepten bu santraller ve sistemin tümü için iyi bir yukarı ölçekleme algoritmasına ihtiyaç duyulmaktadır. Oluşturulacak model ile ülke geneli ya da bölgesel toplam tahminler üretebilecektir. Bu tahminler özellikle elektrik hattının yönetimi ve planlaması açısından çok önemlidir. Ülkedeki enerji akışını yöneten sistem operatörlerinin bu güç tahminleri gün öncesinden kullanarak enerji kaynaklarını doğru bir şekilde kullanmaları oldukça önemlidir.

Bu tezde, bölgesel rüzgar gücü tahmin problemini çözmek için altı farklı istatistiksel temelli model uygulanmış ve elde edilen sonuçlar literatürde iyi bilinen ANN ve SVM gibi bazı istatistiksel tabanlı modellerle karşılaştırılmıştır. Uygulanan yöntemin en büyük katkısı, bölgesel güç tahminleri üretirken, aynı zamanda tarihsel güç verilerine sahip olmadığımız sistemdeki rüzgar santralleri için de güç tahminleri üretmesidir. Uygulanan modelin bir diğer avantajı da rüzgar santrali kurulması planlanan bir aday noktanın rüzgar potansiyelinin bu yöntemle belirlenebilmesidir. Bu tezde modelin bu özelliği, ülkenin dört farklı şehrinden 16 farklı aday noktası için test edilmiştir. Ayrıca 9 farklı yük tevzi bölgesi için bölgesel tahmin sonuçları test edilerek modellerin performans sonuçları analiz edilmektedir.

Anahtar Kelimeler: Bölgesel Rüzgar Enerjisi Tahmini, Veri Madenciliği, Sayısal Hava Tahminleri, Yukarı Ölçekleme

To my family

ACKNOWLEDGMENTS

First of all, I would like to thank Prof. Dr. Pınar Karagöz, the advisor of the thesis, for the ways she showed me at every stage of the thesis work. It was a great honour for me to work with her on such a novel subject.

I would like to thank Prof. Dr. İbrahim Körpeoğlu and Assoc. Prof. Dr. Şeyda Ertekin, members of the thesis committee, for their valuable comments that they made during my thesis work. With their contribution, this work has become more valuable.

I would also like to thank the jury members Prof. Dr. Suat Özdemir and Prof. Dr. Cem İyigün for their valuable comments during my thesis defense.

I would like to thank my executives, Dr. Cem Şahin, Assoc. Prof. Dr. Dilek Küçük and Dr. Turan Demirci for their support during the thesis work.

A special thank to Dr. Serkan Buhan who guided me with his knowledge in many places when I grapple with the subject.

Thanks to my great friends Umut Güvengir and Aydın Başkaya for their great contributions to my work and help throughout this study.

I would like to thank all my colleagues in TÜBİTAK MAM especially Cevdet Ünlüer, Sinan Eren, Süleyman Yeşilyurt, Gökhan Cantürk, Mustafa Çalışkan, Ali Ünver Seçen, Okan Kılıçoğlu and Uğur Şahbaz who contributed indirectly to this study.

I would like to thank all of the members of RITM project.

I would like to thank my precious family who stood by me in my most difficult and beautiful moments all my life.

Finally, I would like to thank my dear wife Güneş for her support in this process. She was always with me.

TABLE OF CONTENTS

ABSTRACT v

ÖZ vii

ACKNOWLEDGMENTS x

TABLE OF CONTENTS xi

LIST OF TABLES xiv

LIST OF FIGURES xv

LIST OF ABBREVIATIONS xvii

CHAPTERS

1 INTRODUCTION 1

1.1 General Context 1

1.2 Essentials of the Wind Power Forecasting 3

1.3 Regional Wind Power Forecasting 6

1.4 Contributions of the Thesis 8

1.5 Organization of the Thesis 9

2 RELATED WORK 11

2.1 Models for Regional Wind Power Forecasting 11

2.2 Models for Wind Resource Assessment 14

2.2.1 Models based on Site Measurements 17

2.2.2 Models based on NWP 18

2.2.3 Hybrid Models 19

2.2.4 Turkey Wind Potential Atlas 19

2.3 Renewable Energy Monitoring and Forecast Centers 21

3 BACKGROUND 23

3.1 ANN 23

3.2 SVM 26

3.3 Similarity Analysis Methods 28

3.3.1 Pearson's Product Moment-Correlation 28

3.3.2 RV Correlation 30

4 PROPOSED TECHNIQUE: REGIONALSHWIP 31

5 LARGE SCALE ENERGY MONITORING AND FORECAST 43

5.1 Data Collection Facilities 43

5.2 Forecast Facilities 45

5.3 Data Presentation Facilities 48

5.4 Wind Power Monitoring and Forecast System: A Case Study 49

6 EXPERIMENTAL RESULTS 53

6.1 Experiments for Wind Power Forecasting 53

6.1.1 Data Used In Experiments 53

6.1.2 Experiment 1: Analysis on Power Forecast for *Offline* Plants . 54

6.1.3 Experiment 2: Analysis on Regional Power Forecasting 59

6.2 Experiments for Wind Power Assessment 61

6.2.1 Experiment 1: Analyzing the Model on Candidate WPPs . . . 61

6.2.1.1 Balikesir Wind Map Results 65

6.2.1.2 Manisa Wind Map Results 65

6.2.1.3 Hatay Wind Map Results 66

6.2.1.4 Tokat Wind Map Results 68

6.2.2 Experiment 2: Analyzing the Model on *Online* WPPs 68

7 CONCLUSION AND FUTURE WORK 71

REFERENCES 73

CURRICULUM VITAE 85

LIST OF TABLES

TABLES

Table 2.1 Summary of various wind resource assessment models 16

Table 4.1 NWP Sources 31

Table 4.2 r value correlation table 36

Table 6.1 Results of all Models (In terms of NMAE %) 55

Table 6.2 Results of all Models (In terms of NRMSE %) 55

Table 6.3 p-value and Standard Deviation Test Results between Models 56

Table 6.4 Results of all Regions (In terms of NMAE %) 62

Table 6.5 Results of all Regions (In terms of NRMSE %) 63

Table 6.6 Forecast Result for all Cities 67

Table 6.7 Hourly Efficiencies of Real WPPs in Hatay. WPP1 and WPP2 are located in CP-1 area, WPP3 and WPP4 are in CP-2 plant area. 67

Table 6.8 Forecast Result for *online* WPPs 69

LIST OF FIGURES

FIGURES

Figure 1.1 Wind Power Capacity of Top 10 Countries in 2017 [1] 2

Figure 1.2 Classification of Wind Power Forecasting 4

Figure 1.3 A sample region from the Marmara Sea 6

Figure 1.4 Regional Wind Power Forecast Approaches 7

Figure 2.1 Turkey Wind Energy Potential Atlas 20

Figure 3.1 Typical Neuron Structure 24

Figure 3.2 ANN Structure 25

Figure 3.3 SVM Classification 27

Figure 4.1 An overview of the power forecast models for *online* and *offline* WPPs in RITM Center 32

Figure 4.2 General Architecture of the Proposed RegionalSHWIP Method. . 33

Figure 4.3 Power Curve of a Typical 1 MW Wind Turbine 35

Figure 4.4 The flowchart of the Model-5 38

Figure 4.5 The flowchart of the Model-6 41

Figure 5.1 The architecture of the renewable energy monitoring and forecast system 44

Figure 5.2 Plant Based and Regional Based Forecasting 46

Figure 5.3 The general architecture of the RITM system. 50

Figure 5.4 Numerical weather forecasts in RITM system 51

Figure 6.1 Model-6 and Model-1 Comparison in Determination of the Peak Values 58

Figure 6.2 Seasonal Performance of the all Models 59

Figure 6.3 Model-6's Results for West Anatolia YTM 60

Figure 6.4 Model-6's Results for National Anatolia YTM 60

Figure 6.5 **(a)** Wind Map of Balikesir, **(b)** Wind Map of Manisa, **(c)** Wind Map of Hatay, **(d)** Wind Map of Tokat 64

Figure 6.6 **(a)** WPPs in Balikesir, **(b)** WPPs in Manisa, **(c)** WPPs in Hatay, **(d)** WPPs in Tokat 66

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
AR	Auto Regressive Model
ARX	Auto Regressive Model with Exogenous Variables
CECRE	The Special Regimes Control Center
CFD	Computational Fluid Dynamics
CNN	Convolutional Neural Network
DCS	Data Collection Service
ECMWF	European Centre for Medium-Range Weather Forecasts
GFS	Global Forecast System
GIC	Geomagnetic Originated Current
GPS	Global Positioning System
GUI	Graphical User Interface
GW	Gigawatt
CP	Candidate Plant
IEC	International Electrotechnical Commission
IP	Internet Protocol
IPC	Inter-Process Communication
LIDAR	Light Detection and Distance
LightGBM	Light Gradient Boosted Machine
MGM	Turkish Met Office
MLP	Multilayer Perceptron
MOS	Model Output Statistics
MSE	Mean Square Error
MW	Megawatt

NCAR	National Center for Atmospheric Research
NMAE	Normalized Mean Absolute Error
NOAA	National Oceanic and Atmospheric Administration
NRMSE	Normalized Root Mean Square Error
NWP	Numerical Weather Prediction
OT	Orthagonal Test
PCA	Principal Component Analysis
QuickSCAT	Quick Scatterometer
RBF	Radial Base Function
RBFNN	Radial Base Function Neural Network
REMFS	Renewable Energy Monitoring and Forecast System
RITM	Wind Power Monitoring and Forecast Center
RV	Random Variable
RegionalSHWIP	Regional Statistical Hybrid Wind Power Forecast Technique
SCADA	Supervisory Control and Data Acquisition
SFTP	Secure File Transfer Protocol
SHWIP	Statistical Hybrid Wind Power Forecast Technique
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SVR	Support Vector Regression
TEIAS	Turkish Electricity Transmission Company
TSO	Transmission System Operator
TUBITAK	The Scientific and Technological Research Council of Turkey
VDRAS	Variational Doppler Radar Analysis System
WAsP	Wind Atlas Analysis and Application Program
WPP	Wind Power Plant
YEGM	General Directorate of Renewable Energy of Turkey
YTM	Load Distribution Center

CHAPTER 1

INTRODUCTION

1.1 General Context

Energy is one of the basic needs of people in order to continue their daily lives. Energy sources exist in nature in various forms and we can categorize them into three basic categories: fossil origin, nuclear or renewable (alternative). Technological developments, which accelerated especially with the industrial revolution, increased the energy need of human beings day by day and these needs were met mostly from fossil sources. However, the overuse and rapid depletion of fossil energy resources have brought some problems in recent years.

Climate change is at the top of these problems and this is one of the most important problems experienced in the world in recent years [2, 3]. The main reason for the increasing climate change is the high usage of fossil-based energy sources such as coal, oil and natural gas, which release high carbon dioxide into the atmosphere [2, 3]. People need to lower their energy consumption and greenhouse gas emissions for a safer and healthier future. In the last century, this situation was taken into consideration by world leaders and in 1997, the Kyoto Protocol was put into effect. The key goal of this agreement is to reduce the rate of atmospheric emissions of methane and carbon dioxide through the use of renewable energy sources by reducing the consumption of fossil fuels [4].

Since fossil fuels are exhausted and dangerous for the environment, reducing dependence on these fuels and increasing the use of alternative clean energy sources is vital for a sustainable environment. Compared to fossil fuel types and nuclear energy, clean energy types such as solar, wind and hydro derive their source from the sun and are always in natural flow [5, 6]. Therefore, the use of these resources in nature is very important not only for our climate and health, but also for our economy.

In recent years, with developments in turbine technology and solutions, wind energy has become an important source of renewable energy, rising rapidly in the energy industry. Because the energy produced from this source does not pollute the atmosphere, it both decreases dependency on fossil fuels and reduces greenhouse gas emissions that are environmentally harmful [7]. In addition, this energy source is available all over the world and not only contributes to the increase of energy diversity but also creates new opportunities for people and contributes to the development of national economies [8].

According to the Renewables Global Status Report, the wind power capacity in the world reached approximately 550 GW as of the end of 2017 and within this scope, China (35 %) and the USA (17 %) will account for about half of the total installed power as shown in Figure 1.1 [1]. According to the same report, approximately 26.5 % of the world's energy production is obtained from renewable energy sources such as hydro, solar and wind, while wind energy is the second largest renewable energy source among all energy sources with a production share of 5.6 % [1]. Despite its high potential, the most important feature that distinguishes wind energy from other clean energy sources is that it is extremely variable and therefore it is important to

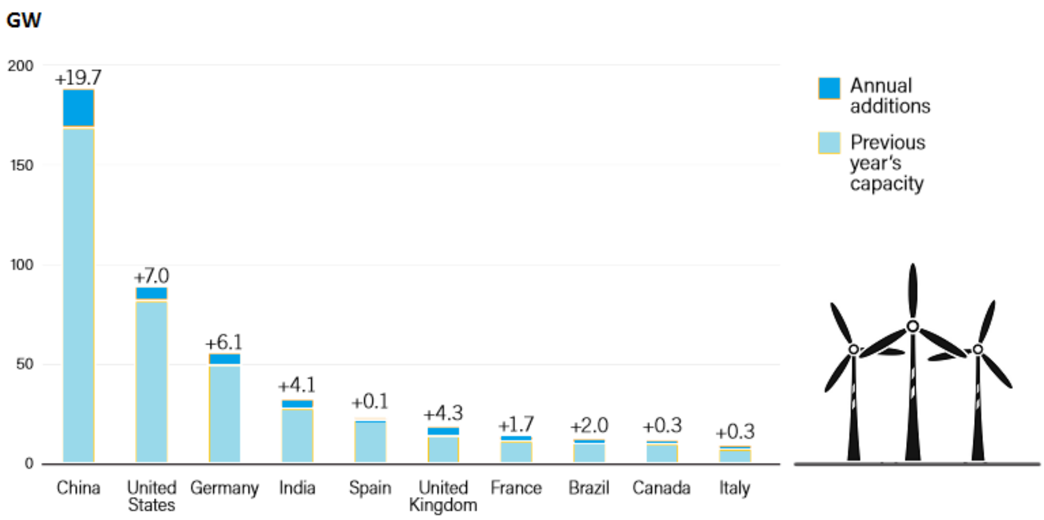

Figure 1.1: Wind Power Capacity of Top 10 Countries in 2017 [1]

include this energy source in the network management in a controlled manner. Due to this requirement, various wind energy estimation methods have been proposed by researchers in the literature and it is aimed to safely include the power obtained from the wind into the energy flow with these methods [9, 10, 11].

Wind energy investments in Turkey as in other developed countries has been increasing in recent years. By the end of 2020, there are approximately 180 WPPs with a total installed power of 8.1 GW across the country. Installed wind power capacity in the country corresponds to about 8 % of Turkey's total installed capacity. It is aimed to increase this ratio to 20 % with the investments made in new wind power plants [12]. In order to control the energy generated from the wind in the country, the RITM project was initiated by TUBITAK for YEGM in 2010 and the first wind energy estimates in this system were produced in 2011 [13, 14, 15]. The two main objectives of this project are to monitor wind power in the country in real time and to generate reliable wind power estimates for advanced energy management. At the center, various power prediction models are working daily to solve different problems such as short term wind energy forecast (up to 48 hours), very short term wind energy forecast (up to 6 hours), probabilistic wind energy forecast and ramp forecast. For regional wind energy estimation, the method described in this thesis has been integrated into the system as the regional power estimation module of the RITM system.

1.2 Essentials of the Wind Power Forecasting

There are inherent difficulties in obtaining energy from alternative energy sources. Wind energy has an important place in alternative energy sources with its high potential. However, due to its stochastic nature, it has a very unstable and volatile unlike other renewable energy sources. Therefore, a reliable forecast system is inevitable for the system operator that manages the energy flow within the country [9, 10].

The power output of a typical wind turbine depends on the parameters as presented in Equation 11 [16]. In this formula ρ represents air density, C is for the maximum theoretical coefficient of performance, A is the frontal area of the wind turbine and V is the velocity of the wind speed. According to this formula, the most important input

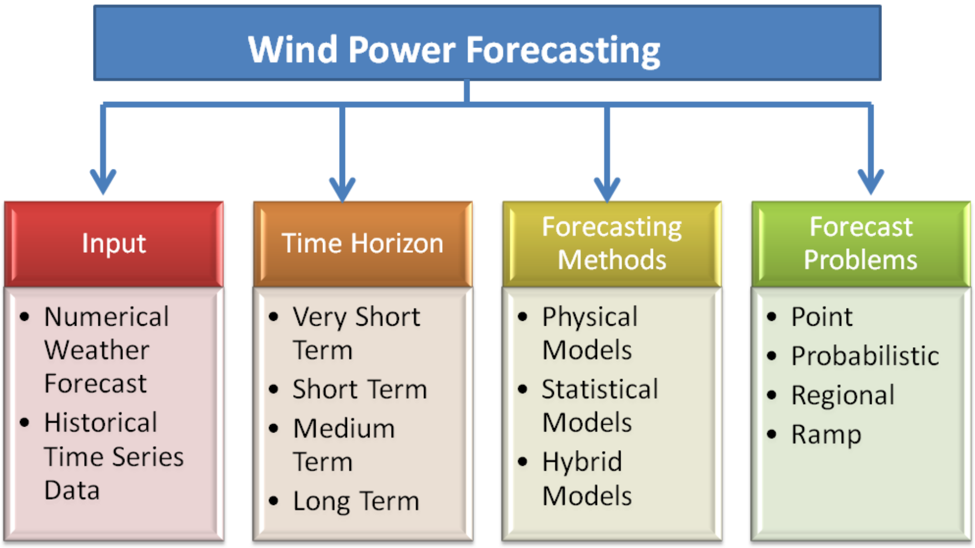

Figure 1.2: Classification of Wind Power Forecasting

for the wind power is the wind speed and the power output of the wind turbine has a cubic relation between it.

$$Power = \frac{1}{2} C \rho A V^3 \tag{11}$$

This equation shows that even a small increase in wind speed causes a large increase in power, and achieving maximum power is more associated with achieving maximum speed in the right place. Therefore, wind energy prediction models are often concerned with the prediction of accurate wind speed values.

Wind energy prediction methods usually use weather forecast and SCADA data to estimate the power to be produced on a wind farm. It is possible to classify the power estimation methods developed for this purpose as statistical, physical and hybrid [9, 10] as shown in Figure 1.2. Statistical methods use historical production values and weather forecasts to find the relationship between them when creating the model, while physical methods use features such as the roughness, height, and slope of the area as input to the model [10, 17]. In hybrid methods, the combination of statistical and physical methods is often preferred. We can summarize the type of estimation problems targeted by all these methods as follows:

- *Very Short Term Wind Power Forecasting*: These are generally high resolution (at least 1 minute) power estimates ranging from 0-6 hours. These estimates are used in real-time network transactions [18].

- *Short Term Wind Power Forecasting*: This type of forecast is usually 0-72 hours per hour. These estimates are used in the energy market. Therefore, accuracy is important in the decision-making processes of the plant owners [19, 20, 21].

- *Regional Wind Power Forecasting*: This type of power estimate targets a specific region using data from some reference facilities. These estimates are used by Transmission System Operator (TSO) in maintenance planning [22].

- *Probabilistic Wind Power Forecasting*: In this type of prediction, the uncertainty in wind energy estimation for WPPs is presented within certain confidence intervals, also called percentages. They are used in operation management and load decision [23, 24].

- *Ramp Wind Power Forecasting*: The main purpose of this type of prediction is to predict sudden drops and rises in power. These estimates are often used in the facility's risk analysis [25].

We can summarize the main objectives of all these forecast types as follows:

- Setting energy reserves according to accurate wind forecasts.

- Optimizing the planning of power plants economically and increasing the profitability of power plants in the market.

- Ensuring system optimization for TSOs and to create a healthy energy flow.

- Producing maintenance planning for WPP components such as wind turbines.

- Determining whether a place is sufficient for wind potential using actual measurements and weather forecasts.

1.3 Regional Wind Power Forecasting

In an ideal wind forecast center system, it is desirable to store power and weather forecast data for all power plants and to create a wind power forecast model from these data. However, in systems with a large number of wind power plants, it is a costly process to obtain this data instantly for all power plants, and therefore a different approach is required for the total wind power estimation of the whole system or a particular region. As a result of this problem, regional power prediction algorithms have emerged.

Regional wind energy prediction models are about estimating the wind power generation of the region or the whole country using measurements of some reference *online* wind power plants [26]. A sample region in the country is shown in Figure 1.3. In this demonstration, green circles represent *online* WPPs in the system, while blacks show *offline* WPPs where actual production data are not available at the RITM center. In the center the NWP weather forecasts are available both type of plants and real production values are collected by the *online* WPPs in real time with 3-second resolution. The purpose of such estimates is to model the region in terms of *online* plants and to produce a regional wind energy forecast. Such power estimates are crucial for the management of the electricity grid and for more efficient planning of energy resources.

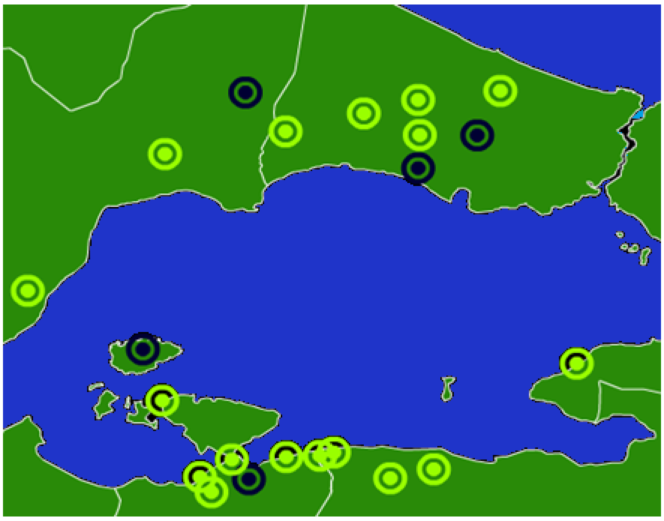

Figure 1.3: A sample region from the Marmara Sea

Although regional wind power forecast has become a very popular research topic in recent years, there is no standard definition for the *region*. While some researchers define it as the area for several hundred kilometers, for some researchers, regions are defined according to the density of the wind farms or the neighbor relationship. Regions in Turkey are determined by TEIAS, who is responsible for electricity transmission in the country, according to their proximity to power plants and electricity transmission lines. They named these regions as YTM [27], and when testing the accuracy of the studies proposed in this thesis regionally, these regions were taken into consideration.

General approaches for the regional wind power forecast problem can be summarized as shown Figure 1.4 [26]. In general, we can classify the approaches to this problem as three groups namely direct upscaling, cascaded approach and cluster approach. In direct upscaling approach, the regional wind power forecasts are produced directly by using the power forecasts of the *online* WPPs. In this modelling, the direct upscaling function must be updated after the addition of the new *online* WPPs in to system. In the cascaded approach firstly power forecasts are generated to reference *online* WPPs and then the sum is extrapolated to the total region. Finally, in cluster or sub-regions approach the WPPs are aggregated into clusters that contain neighbouring relation. For all of these models, selection of the correct reference wind power plants is important and all of these approaches are examined in detail in Related Work chapter.

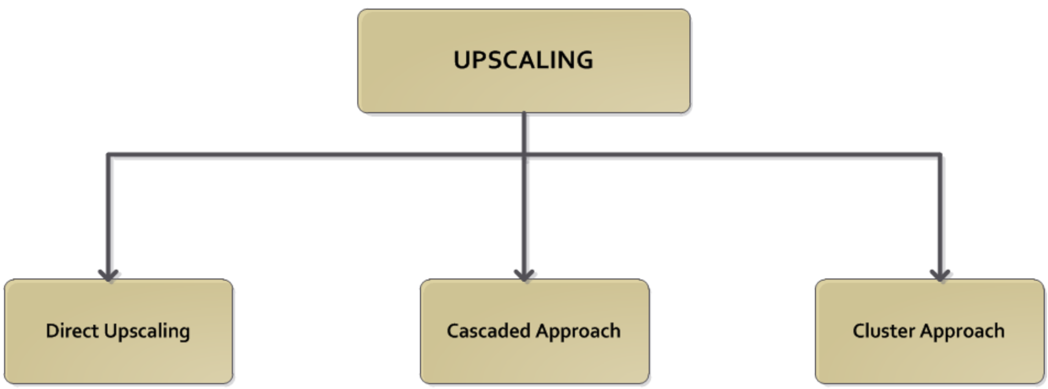

Figure 1.4: Regional Wind Power Forecast Approaches

1.4 Contributions of the Thesis

In this thesis study, it is focused on regional wind power forecasting problem, and a new method is proposed for forecasting. *Regional Wind Power Forecasting* (*i.e., upscaling*) problem deals with the prediction of the aggregated power output of wind farms located within a defined region [9, 10, 26, 28, 29]. The main problem in the regional wind power forecasting is that dynamic information of all wind power plants such as weather forecasts and production data is not usually available. Currently, in RITM project, a hybrid forecasting method namely *Statistical Hybrid Wind Power Forecast Technique (SHWIP)* [19, 30] is in use as the point forecasting model. However, this method needs past power data of the wind power plant for forecast model construction. So this model is applicable for *online* plants whose production data is monitored by the system in real time. Therefore, for power prediction of *offline* plants, and hence, of a given region, there is a need for an alternative model.

In regional wind power forecasting, many models in the literature propose techniques for the estimation of the entire zone rather than an *offline* plant. *Offline* plants are the plants such that their production data is not available in the system. In this work, a new method is proposed for generating power forecasting for *offline* plants, and for a region at the same time. The technique is based on firstly power forecasting for *offline* plants, and then upscaling to region by using forecasts for both *online* and *offline* wind power plants.

The contributions of this thesis can be summarized as follows:

- With this method, power estimation can be done for each wind power plant in the country even if the power generation history is not available for some of them. The model produces power forecasts for the *offline* wind power plants by historical weather forecast similarities with reference *online* plants.
- Proposed model uses a bottom-up approach (different from the methods in the literature) by first producing power forecasts of all WPPs. Therefore, regional forecast results are more accurate and stable compared to the other direct scaling models and it considers all of the WPPs in the country while calculating the power forecasts for the regions.

- This model can be used on a candidate plant area to estimate the wind power potential of the location. This problem is named as *Wind Resource Assessment* and by applying this method the wind power potential of a wind farm location under consideration can be estimated. Wind maps of the country is a useful resource to determine high-potential locations. However, such maps often present an overview to provide the general picture, and there is a need for a more detailed analysis, specifically on the potential plant area. By applying the model on a potential plant site area, an investor can get the power prediction, and hence the wind potential of the area, without any time and financial cost

- Finally, in this thesis, the system components and system structure required for a generic renewable energy monitoring center are presented. The details of the system components such as data collection facilities, forecast facilities and data presentation facilities are presented in the thesis.

1.5 Organization of the Thesis

The remainder of this work is structured as follows.

The models for the regional wind power forecast and related studies are discussed in the literature in Chapter 2. This section also gives general frameworks of these models with real world examples.

In Chapter 3 of the thesis, background information about the methods used in the work are described. Some statistical based forecast algorithms in the literature, such as ANN and SVM, are used in the implementation of the proposed work to compare the model with the existing models in the literature. Also some similarity index methods such as Pearson Moment Correlation and RV correlation are used for selection of the reference *online* WPPs. In this chapter general information and computational details of these methods and how they are used in the RegionalSHWIP are presented.

In Chapter 4, the proposed regional wind power forecast model, namely RegionalSHWIP, is presented with computational details. This model produces power forecasts both for the entire region and *offline* plants at the same time.

The general structure of a dynamic renewable energy monitoring and forecast center is presented in Chapter 5. In this section, the system components of such system are defined and as a case study RITM system is presented.

The evaluation results of the RegionalSHWIP is discussed in Chapter 6 in detail. The performance of the model is investigated both for *offline* plant and regional basis in terms of NMAE and NRMSE error rates. Additionally, another contribution of the RegionalSHWIP is that it can be applied to *Wind Resource Assessment* problem and the performance of this feature of the model is validated in different cities and their results are discussed in this section.

Finally, in Chapter 7, the study is concluded with further reflections and potential future work.

CHAPTER 2

RELATED WORK

In this section, models for regional wind power forecast, models developed for wind resource assessment problem and renewable energy monitoring centers designed in the literature are discussed separately.

2.1 Models for Regional Wind Power Forecasting

Regional wind power forecast models are about estimating wind power generation for a specific region or whole country using measurements and forecasts of some reference *online* wind power plants. It is costly to obtain the power measurement data of all power plants in the country instantly and it is very difficult in cases where there are many wind power plants. Therefore, there is a need for a model that can produce consistent wind power estimates for the entire region using available data in the system [26]. Such power estimates are crucial for the successful management of the electricity grid and more efficient planning of energy resources.

In [31], Siebert and Kariniotakis focus on the correct selection of reference wind power plants in the regional prediction model. They propose to aggregate k-means for the selection of the reference wind farms and investigate the effect of the number of reference wind power plant selections on regional prediction accuracy. According to their results, increasing the number of reference wind farms to the optimum number reduces the error rate.

Bremen et al. [32] suggest a hybrid physical-statistical scaling approach based on PCA in their work. It shows that the spatial decomposition of wind power generation can be done with PCA to extract the variability pattern. With this approach, they prepare forecast maps for different regions of Germany.

In their study, Daniel et al. [33] investigate the role of regional wind energy prediction to determine ramps in the energy grid. They discuss three different high-resolution weather forecast models to understand their impact on power forecasts. They conclude that an accurate and reliable regional wind energy forecast model minimizes costs for energy grid management.

Lobo and Sanchez propose a regional wind energy prediction model based on flattening techniques [34]. They apply their models to the Spanish peninsula system. Their approach focuses on looking for similarities between current wind speed estimates in the region and past wind speed estimates. Energy production is derived from historical power data based on estimated wind speed similarity.

In [35, 36], the regional wind power forecast model integrated in the Previento forecast tool is described. This forecast tool uses physical characteristics of the WPPs such as turbine hub height, roughness in the plant area etc. Finally, based on the physical properties, several reference wind power plants are selected and the power estimates of the entire region are calculated from these power plants.

Ernst et al. [37] propose a regional wind power prediction approach based on the neural network. They use two different networks in the model. The first network is used to calculate the power output of some selected reference wind farms. These estimates are then given as input to the second network to calculate wind power forecasts for the entire region.

Jan et al. [38] develop a multi-scale regional wind energy forecast model to control the high penetration of wind energy into the electricity grid. Their model is built on multi-to-multi mapping networks and the use of stacked deionized auto-codecs. They generate wind power forecasts for the region, with a time interval of 24 to 72 hours. According to the reported results, their models perform better than other regional power prediction models.

A post-processing technique with PCA is demonstrated by Davo et al. in [22]. They apply the model not only for regional wind energy prediction, but also for solar radiation prediction. They use ECMWF and GFS as sources for numerical weather forecasts, and forecast horizons are 0-72 hours. According to the experimental results, combining PCA with these post-processing techniques provides better performance than the results of the technique without PCA reduction.

RegioPred is developed by Marti et al. [39] by using the point forecasts of reference *online* wind plants, which are generated by their statistical point forecasting model, namely LocaPred. The model uses Computational Fluid Dynamics (CFD) and power curve modeling together to generate power estimates. The selection of reference wind farms is carried out by clustering.

In [40], Barbero et al. compare MLP and SVR performance for reference wind farm selection in the regional wind power prediction model. It is reported that the SVR based model offers better performance since the training period is much less than the MLP.

A hybrid approach for regional wind power forecasting is proposed by Yamaguchi et al. in [41]. They suggest the transfer coefficient method for the online implementation of the mesoscale model and multiple time scale mode. According to the results, the error rate of the model based on the mesoscale model is lower than the model based on NWP data at 20km resolution.

The model proposed in this thesis is based on statistical models such as most of the methods described above. In the proposed model, reference *online* power plants were selected as in the models described in [31, 34, 35, 36] based on geographic relationships, neighborhood, historical wind speed and wind direction similarities. However, unlike these models, in the proposed model, the reference *online* WPPs are determined separately for each *offline* plant instead of the region. Therefore, unlike models in the literature, the proposed model generates the power forecast for each *offline* power WPP in the system before estimating regional power forecasts. As a result, the main advantage of the proposed method is that it is possible to generate power forecasts for each power plant in the system while generating power forecasts to the region even if some WPPs have not historical power data.

2.2 Models for Wind Resource Assessment

Another important problem in the field of wind energy is to determine where to install the wind farm as well as obtaining the correct power forecasts. This problem is called *Wind Source Assessment* and it should be determined whether the facility area has sufficient potential for wind energy by performing these analyzes [42, 43]. Advance analysis of the wind potential of the power plant area is crucial to address investors' financial concerns. With the analysis methods developed today, not only the location of the turbines but also the placement of the blades in the right direction is of great importance for the plant owners in terms of economy.

Wind atlases are created based on actual wind speed measurements made on the field to reveal the wind potential of a region or country. Wind maps are produced by real wind measurement values (especially wind speed) obtained at certain measurement points and meteorological values obtained by satellites. Generally, these measurements are given as input to modeling software such as WAsP [44], windPRO [45] and WindSim [46] to create the field's wind atlas. These wind atlases are used for energy management on country basis and wind power investments of power plant owners.

The RegionalSHWIP forecast model is created to generate energy forecasts for wind farms without historical production data [47]. The model uses meteorological forecasts in the power plant area and wind energy forecasts of the designated reference *online* wind power plants as input when creating the model. A distinctive feature of the model compared to similar solutions in the literature is that this method can also be applied to a candidate wind farm to determine the wind energy potential of the area. In this thesis, this aspect of the method has been analyzed in detail and a verification is made through wind atlas, and the predictability of the wind potential for a given area without time and financial cost is shown. The model has been tested in different candidate wind power plants in different cities from different parts of the country to demonstrate the performance of the model in the selection of the premise area.

The problem of determining the most suitable place for producing maximum wind power can be examined under three headings as follows [42, 43]:

- *Initial Site Selection*: At this stage, the geographical conditions, topographic structure and the distance from the transmission lines of the area planned to be established, are determined whether the area is suitable for the wind farm installation.
- *Wind Resource Assessment*: After determining the suitability of the site in terms of physical conditions, it is determined whether the wind potential is sufficient to invest in this place by real and meteorological measurements.
- *Micro-siting*: At this stage, in which coordinates and in which direction the turbines should be located, is determined in the power plant area to get the highest gain from the wind power plant.

In previous days, wind resource potential was determined based on biological and physical indicators such as flora, tree species and leaf shapes (tree deformation caused by wind) [48]. Strong winds permanently deform some tree species, and this deformity is called as *flagging*. Then, when looking at the deformity levels of the trees, the wind potential and the dominant direction of the wind can be investigated. These deformity levels are determined by the Griggs-Putnam index [49]. With technological developments, the potential of the plant area can be determined both by real measurements and meteorological parameters.

The approaches for this problem can be grouped in three main headings as follows and summary of the related works of these models are presented in Table 2.1[48]:

- *Models based on Site Measurements*: In these models, real wind speed and direction values within a certain period of the plant area are modeled using statistical models. Weibull Distribution based methods are generally used for modeling [50].
- *Models based on NWP*: In these models, meteorological weather forecasts are reduced to the plant area at the turbine level using simulation tools such as WindSim, WAsP and WindPRO, and the modeling of the plant is done in this way.
- *Hybrid Models* : In these methods, both real field measurements and weather forecasts are combined in modelling process.

Table 2.1: Summary of various wind resource assessment models

Classification	Methods	Forecasted Region	Data Feature	Main Observation
Based on Site Measurement	Wind roses,Weibull Distrubion and Roughness Lengths [51]	Singapore	two years data	- Southern part of the country is more ideal for wind power investments
	Weibull Parameters [52]	Kiribati Island	one year data	- Moments method is the optimal solution
	Measure–Correlate–Predict method [53]	South Korea	one year data	- Produces probability models for the wind resources
	Weibull and Rayleigh Distribution [54]	Iran	ten years data	- With a 97% likelihood, the Weibull and Rayleigh distribution functions forecast the observed results.
Based on NWP	Analog Ensemble Methodology [55]	United States	two years data	- Analog ensemble based model shows better performance compared to other two methods
	Satellite Winds [56]	Great Lakes	three to eleven years data	- A robust wind atlas offshore can result from the combination of several observational data sets
	WAsP, OpenWind and WindSim [57]	Brazil	not stated	- WindSim provides more reliable results
	WAsP, MM5 [58]	German Bight	one year data	- MM5 model produces reliable results for the offshore areas
Hybrid	WAsP together with Weibull [59]	China	twenty years data	- For higher points Weibull shows better performance. For the lower parts they combine two models.
	Hybrid-Wind Systems with Battery Storage [60]	Five different city from different countries	eight years data	- With very high precision, the annual Wind Fraction can be predicted.

2.2.1 Models based on Site Measurements

The proposed models in this section generally use to wind speed and direction measurements for different time domains and different height levels.

In [51], Kerthikaya et al. investigate the wind energy potential in Singapore using fixed rooftop wind mast and mobile Light Detection and Distance (LIDAR) measurements located in various parts of Singapore. They make statistical analysis on wind rose and Weibull distribution on measurement data. According to the results, the southern part of the country is more ideal for wind energy investments.

Aukitino et al compare different methods to find the optimum Weibull parameters to determine the wind resource potential of the island of Kiribati [52]. In experiments, they use wind speed, direction, ambient temperature and pressure measurements as input parameters. Optimum Weibull parameters and wind energy density were obtained with the Moments method, which provides higher performance than the other seven models discussed in the study. According to the findings, the wind resources of the Pacific region can be used to reduce dependence on fossil-based energy generation.

In [61], the authors investigate the wind potential for an offshore region in Atlantic Iberian Cost. They divide wind speed and wind direction measurements into different sectors by taking into consideration different wind speed bins allows identifying the intervals associated to poorer results.

Kwon provides a framework for wind resource assessment using probability uncertainty analysis in his work [53]. It configures probability models based on air density, average wind speed and associated Weibull parameters. According to the results of the experiment, the proposed framework can effectively estimate the expected annual energy production with certain confidence intervals.

In [54], Pishgar-Komleh et al. examine the wind characteristics of Iran's Firouzkooh region. In their calculations, they use the 10-year wind speed data at a 3-hour resolution. In addition to the Weibull distribution, they use the Rayleigh distribution functions to determine the most suitable model. According to experimental evaluations, both of these models can fit well with coefficients similar to wind speed actual values. They also determine the dominant wind direction in the region by wind rose diagram analysis, and according to the results, winds blowing from the north are more efficient at the plant site.

2.2.2 Models based on NWP

In that approach, maps are generally created by using the historical NWP data obtained by the satellites.

In their research, Zhang et al. compare three wind-resource assessment approaches based on NWP [55]. They intend to investigate the impact of NWP data sets with high and low resolution in the study of wind potential. High-resolution NWP with an analog ensemble-based methodology model shows significantly better performance in the coarse downscale estimates, according to their research.

Satellite wind observations are used by Doubrawa et al to create an offshore wind atlas for the Great Lakes [56]. They use QuickSCAT wind measurements from the National Aviation and Space Administration and radar wind measurements from the European Space Agency to make the wind speed forecast map. Using WAsP software, they modeled all inputs together and came together in a single wind speed forecast for the Great Lakes region.

Ramos et al. discuss the differences of the WAsP, OpenWind and WindSim modeling tools for the wind resource assessment problem [57]. They use the same turbine design for modeling for all models. Based on its results, WindSim shows more robust results, but its computational complexity is higher than the other two tools. It gives better results for OpenWind software at an altitude of 80-100 m compared to 120 m. Similarly, WAsP shows better performance at 100m and 24 % wake loss at 120m.

An assessment analysis of offshore wind sources is conducted by Jiménez et al. as a case study for German Bight [58]. In the experiments, they compare two approaches, the MM5 mesoscale air model and the WAsP wind resource assessment program. In general, both models perform similarly, but the biggest difference between the models is measured at a distance of 5-50 km from the beach.

2.2.3 Hybrid Models

The works in this part generally combine two approaches for better investigation of the wind potential of the area.

In [59], a mixed approach is suggested by Liu et al. For the analysis of wind potential in Beijing. They use wind measurements from 325 m high meteorological towers from 1991 to 2011. According to the reported results, the Weibull function can accurately simulate the actual wind speed at higher levels. On the other hand, in other cases, performance should be increased. They use WAsP with the Weibull function to produce the Beijing wind atlas for these levels.

Celik [60] offers a new algorithm to estimate the annual wind rate in a hybrid wind power system. He uses 8-year-old wind speed measurement data, collected in five different positions at the hourly resolution. With the proposed approach, battery voltage status can be simulated to estimate monthly and annual wind speed fraction. The proposed algorithm relates the annual wind fraction to the parameters of the Weibull distribution function.

2.2.4 Turkey Wind Potential Atlas

A country's wind atlas is important to investigate the wind energy potential of cities and areas in the region. In this work, the wind potential analysis generated by the proposed regional wind power forecast algorithm, RegionalSHWIP, is compared against the Wind Potential Atlas of Turkey, which is published by YEGM [62]. As shown in Figure 2.1, atlas is produced hybridly using a mid-size digital weather forecast model and a micro-scale wind flow model. Actual measurements are collected by 60 wind observation stations located in different parts of the country at an altitude of 200 m. In the country there are 81 cities and these are shown in the map with their borders and names. In the figure additionally seasonal maps for winter (December, January, February), spring (March, April, May), summer (June, July, August) and autumn (September, October, November) are presented. According to atlas, the western part of the country has more wind potential and most of the WPPs are built especially in this region. Likewise, the wind energy potential in winter is higher than in other

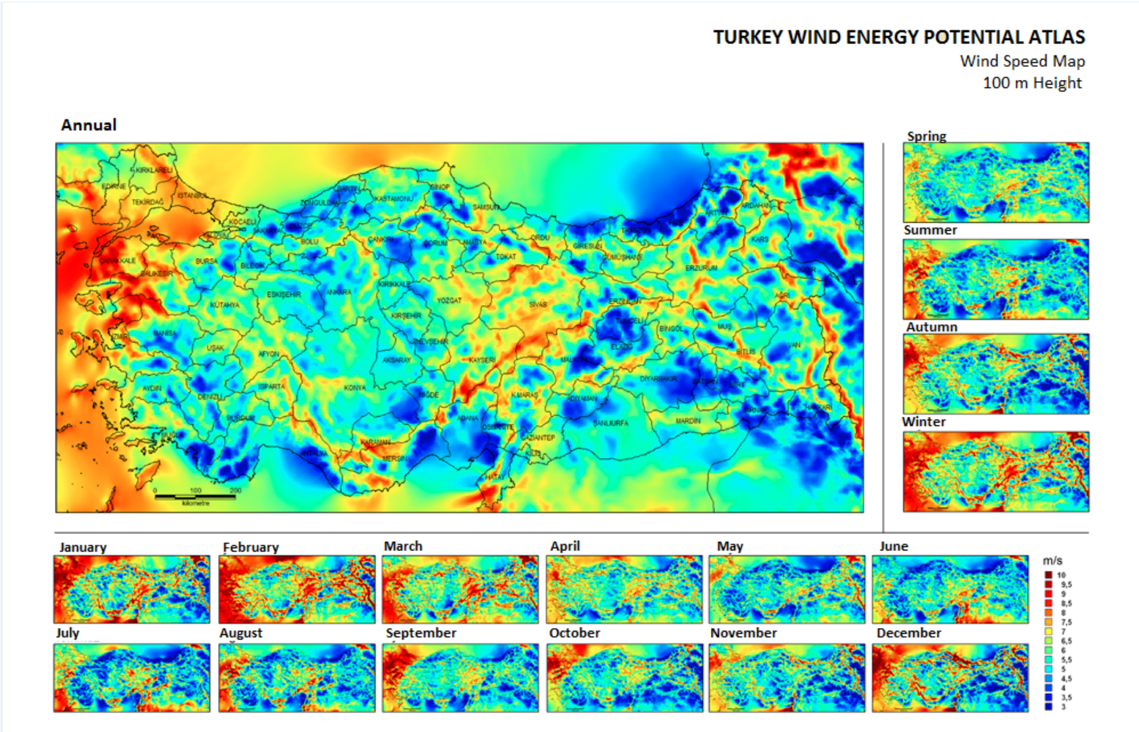

Figure 2.1: Turkey Wind Energy Potential Atlas

seasons. The country's monthly wind potential is also shown on the map. According to the results, Turkey has nearly 48 GW wind power potential with 38 GW terrestrial and 10 GW offshore capacity. With the help of this atlas, the following information is available at 200m x 200m resolution:

- Average annual, seasonal, monthly and daily wind velocity in 30, 50, 70 and 100 m heights
- Annual, seasonal and monthly wind power densities at 50 and 100 m altitudes
- Annual capacity factor at a height of 50 m for a reference wind turbine
- Annual wind classes at a height of 50 m
- Monthly temperature values at heights of 2 and 50 m
- Monthly pressure values at sea level and 50 m altitude

2.3 Renewable Energy Monitoring and Forecast Centers

Instant monitoring and forecasting of energy from energy sources are the two most important issues for the institutions that manage the energy flow in the country. For good planning, data collected in the field can be safely stored on central servers, and system operators can instantly monitor system reliability using the user software. In addition to instant monitoring, the second most important pillar of a reliable system is reliable forecasting. A good forecasting is crucial for the system operator especially for better planing of the electric grid. Depending on the climatic conditions, the integration of energy sources such as wind, sun and water into the system should be carried out with reliable estimates for the next day. In this section, the general structures of energy prediction and monitoring centers and algorithmic approaches used for prediction applications are examined in the literature.

In [63], the authors provide details of the geomagnetic originated current (GIC) monitoring and forecasting system used by The National Grid Company, a high voltage electrical power systems operator in England and Wales. The main purpose of this system is to determine oscillations in the power output, voltage drop, and transformer failures in the transmission system, using global solar weather forecasts. The estimated time horizon in the system is determined as long-term (up to one year), medium-term (up to one month), short-term (up to one week) and real-time (1 hour ahead). According to the performance results of the system, the most useful forecasts are obtained 72, 24 and 1 hour ago, and the system accurately predicts two main solar storm events and alarms the system operator.

The Special Regimes Control Center (CECRE) is located in Spain to integrate renewable energy sources into the transmission line [64, 65]. This center is used in Spain to monitor and control wind power generation. The system uses active and reactive power, voltage, connection, temperature and wind speed values as input. When using these inputs, CECRE automatically calculates wind energy, which must be integrated into the energy flow at any time. In addition, voltage drops and production triggers are estimated at the center and this information is used for congestion management. According to the regulation, generators with an installed capacity of 10 MW should be directly monitored and integrated into the CECRE system.

Technologies and communication standards of the smart grid system have been discussed by the authors in [66]. It is an expanded version of traditional electricity grid infrastructure integrated with renewable and alternative energy sources with smart grid systems, automatic control and modern communication technologies. In such systems it is important to determine the security of the data and the standard of the data format. In the article, the authors discuss the basic features of communication standards for various parts of the system, such as power line network, measurement and communication control of residential network devices, communication from inter-control centers and interoperability.

In [67], the authors propose a smart referral structure for regional transmission organizations and transmission system operators to control large power networks. Unlike traditional methods, using this structure, the system operator is intended to make real-time decisions based on changes in system conditions such as load, production, changes in the electricity network. Based on the experimental results, the proposed shipping structure provides a more holistic and forward-looking view of the system conditions.

Researchers from [68] propose a wind energy forecast framework used by The National Center for Atmospheric Research (NCAR) to optimize network integration. In the system, researchers present the predictive estimates of mesoscal wind energy and probability estimates of wind energy calibrated by Kalman analog filter and quantile regression. The estimates produced are used not only for real-time integration of the network, but also for the next day trading in the energy market. The Variable Doppler Radar Assimilation System (VDRAS) developed by NCAR is used for data entry and the system obtains data with lidars, Doppler radars and surface networks.

CHAPTER 3

BACKGROUND

In recent years, artificial intelligence-based statistical methods have been used effectively in important research topics especially in the fields of banking, medicine, energy and social media. Different versions of statistical based ANN and SVM methods are used in many different studies, especially in wind power prediction problems also. The performance of the statistical-based RegionalSHWIP prediction model developed in this thesis is compared with two methods based on ANN and SVM. In this section, the definitions of these two methods and their usage in wind power prediction in the literature are explained with case studies.

In addition, two different methods are used in the RegionalSHWIP algorithm to select similar reference *online* plants. These methods are Pearson Moment Correlation and RV Correlation techniques and their definitions are also explained in this section.

3.1 ANN

Artificial neural networks are structures that are based on the working principle of the human brain and they work with the logic of the nervous system [69]. In these structures, as in the nervous system, every past information (similar situations, similar patterns) is learned and reflected in the guesswork. The characteristic behaviors of the data are taught with the training sets and predictions are produced in the test set according to these learning. These structures have been highly preferred in recent years for the solution of different problems in many fields such as economy, medicine, energy, communication and social media [70].

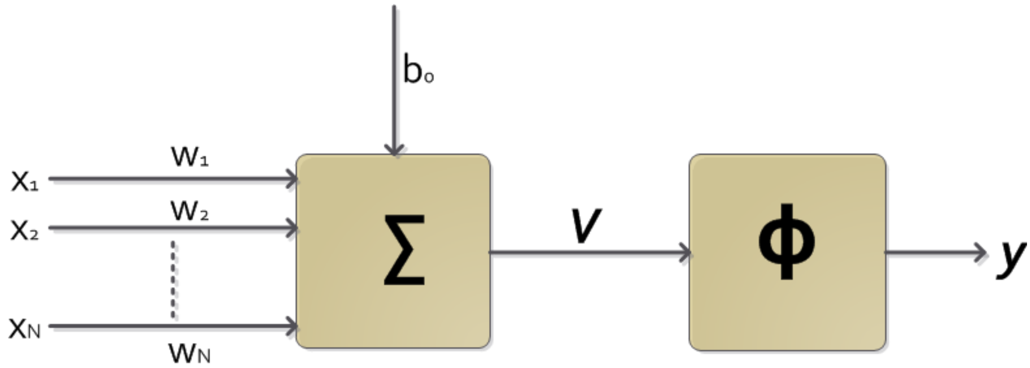

Figure 3.1: Typical Neuron Structure

The basic building block of ANN is called the neuron as in the brain, and a typical artificial neuron structure is shown in Figure 3.1. In such a structure, the total output value of the neuron is first calculated by the formula in Equation 31. In this equation, *x* symbolizes the input data, w_i weight coefficient, the total number of data in the training set *N* and *b* bias value. The calculated total value is then passed through the activation function as in Equation 32, and the final output value of the neuron is calculated. In this equation, *y* represents the final value of the neuron and Φ is for the activation function [71].

$$v = b_0 + x_1 w_1 + x_2 w_2 + ... + x_N w_N = \sum_{j=1}^{N} x_j w_j + b_0 \tag{31}$$

$$y = \Phi(v) \tag{32}$$

Depending on the type of problem studied, linear, threshold, sigmoid and tangent activation functions are used, but in wind prediction problems, sigmoid function is generally preferred as the activation function due to the wind characteristics.

Although neurons are effective structures, a neuron alone is not enough to solve complex problems. For this reason, a large number of neurons connect to each other to form the ANN model shown in Figure 3.2.

There is mainly a multi-layered feedback structure in the ANN models. There are input, hidden and output layers in this structure. The input layer varies depending

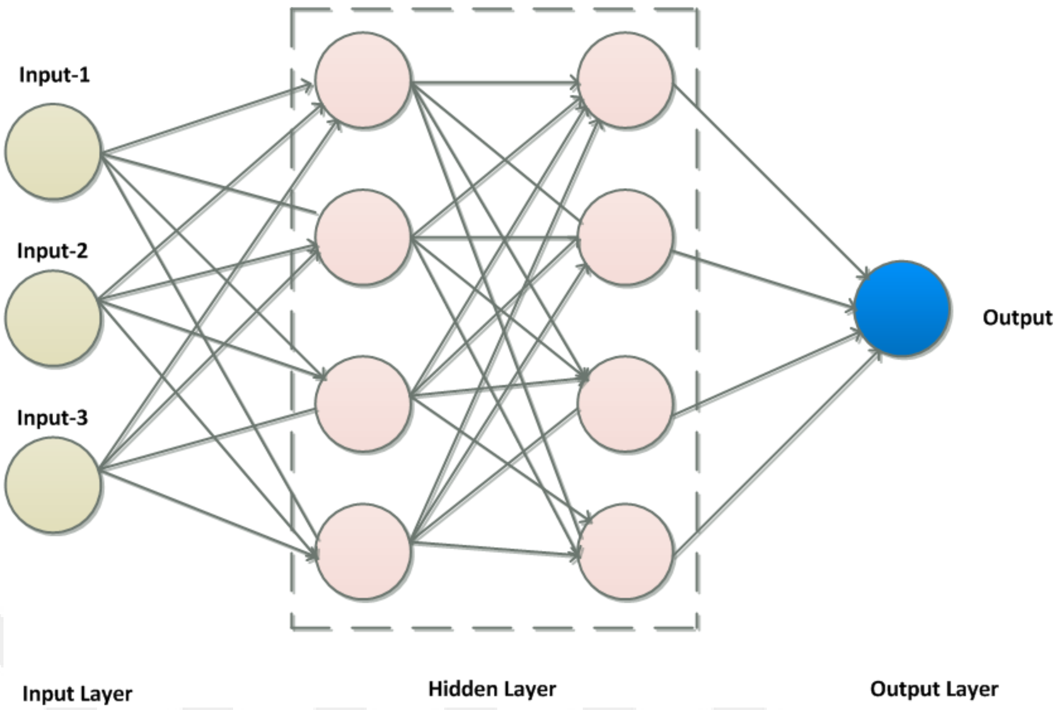

Figure 3.2: ANN Structure

on the number of inputs given to the model. The number of hidden layers can be designed specific to the problem used in a different number. The number of neurons is determined in the output layer according to the response produced as a result of the problem. In the training process, in each iteration, the outcome value is compared with the goal value and the error is determined. Then the error is tried to be minimized by using weights and bias with feedback in each iteration.

There are many different types of ANN, developed for different problem types, as follows:

- Feed-forward Neural Network – Artificial Neuron
- Radial basis function Neural Network
- Kohonen Self Organizing Neural Network
- Recurrent Neural Network(RNN) – Long Short Term Memory
- Convolutional neural network (CNN)
- Modular Neural Network

ANN structures are also frequently preferred in many different models in wind power prediction. Hong and Rioflorido [72] propose a hybrid ANN model for short term wind power forecasting by combining the Radial Basis Function Neural Network (RBFNN) with Convolutional Neural Network (CNN). In their models, first they determine best number of layers, length of kernels and the dropout value and they compare their model with traditional models also in seasonal. According to their results, proposed method shows better performance compared other traditional models.

In [73], Ju et al. use ANN for ultra-short-term wind power forecasting by combining Convolutional Neural Network and LightGBM algorithm. In their model, they first construct a new feature sets on the wind raw data then CNN is used to extract information from this data set. In the final stage, LightGBM algorithm is used in the classification of the power forecasts. According to their results, proposed approach shows better performance compared to existing support vector machines for the very short-term time scale.

In [74], the authors discuss the effect of other meteorological parameters such as temperature and pressure on wind power forecasting by applying ANN based model. They propose three different ANN structures in which the first is given only wind speed, in the second wind speed and temperature, in the third, wind speed, temperature and pressure are given together as input. According to their results, the lowest error rates are obtained in the third structure which also uses temperature and pressure value in the construction of the ANN.

3.2 SVM

Support Vector Machines are one of the important machine learning algorithms developed for the classification problem, which is especially important for data mining [75]. Like ANN, the SVM model also references neural networks as its working principle. The aim of these structures is to carry the problem to a hyper flat that can be solved in different dimensions. As shown in Figure 3.3, a linear SVM model allows different clusters to be separated from each other, considering the maximum margin in the hyper-plane.

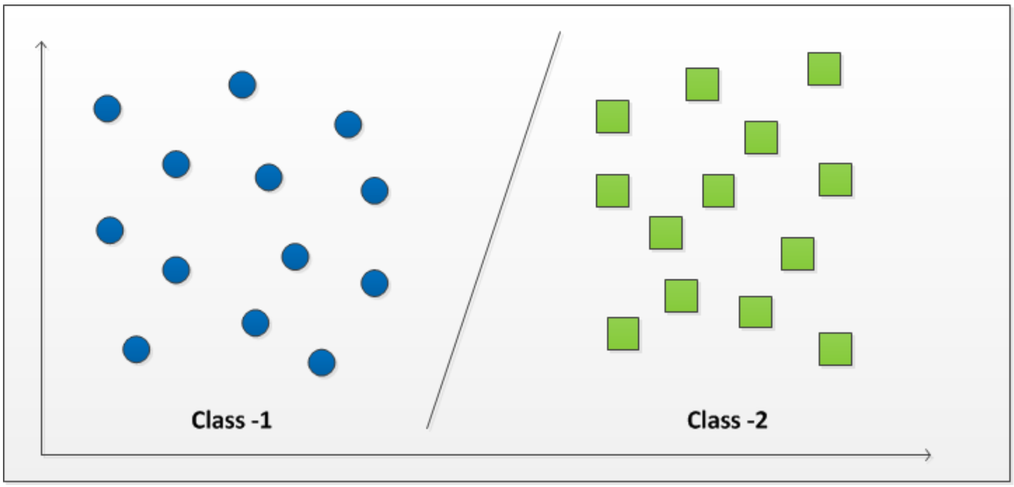

Figure 3.3: SVM Classification

A linear SVM model can be formulated as in Equation 33 [76]. In this notation, y represents the output vector, w is the hyperplane normal, x is the input vector, and b is the scroll window. The purpose of these structures is to maximize the margin interval with the training set used. Lagrange multipliers and Kernel functions are often used to separate data sets as represented in Equation 34.

$$y = wx - b \tag{33}$$

$$y = \sum y_j a_j K(x_j, x) - b \tag{34}$$

In this notation a represents the Lagrange multiplier and K represents the kernel function used in distance measurement. Radial Based Kernel Function (RBF) is generally preferred in wind power forecasting [77].

In the literature, as with ANN, many wind power prediction models are proposed based on SVM. For determining the ramps on the wind power, Yiu et al. propose a forecasting model which is based on Orthagonal Test and Support Vector Machine [78]. They named their model as OT-SVM and the proposed model is tested on three different WPPs in China by compared with the pure SVM models. According to their findings, proposed OT-SVM model shows better performance compared to other models.

Barman and Choudhury propose a hybrid SVM based approach for short-term load forecasting [79]. Proposed model combines firefly algorithm with SVM by looking at the seasonal similarities of the load forecasting data. Their model is tested on a city of India with four different seasons and according to experimental results the model that uses seasonal characteristics as input gives better results than other standard models.

Li et al. propose a SVM based model which is supported with dragonfly algorithm for short-term wind power forecasting [80]. In the proposed approach, the improved dragonfly algorithm is used to choose the optimal parameters of SVM and the model is tested on wind plant located in France. They compare their models with a back propagation neural network and Gaussian process regression based model and according to their findings SVM based model shows better performance compared to former approaches.

3.3 Similarity Analysis Methods

Determining the similarity on two different time series data is an important problem occurring in many different areas and many different similarity indexes have been proposed for the solution of this problem. In the model proposed in this thesis, Pearson Moment Correlation technique is used to determine one-dimensional wind speed similarities, and RV Correlation technique is used to determine the similarity in the data consisting of two-dimensional wind speed and direction.

3.3.1 Pearson's Product Moment-Correlation

Pearson's correlation coefficient is a method that measures the statistical relationship between two continuous variables. Since it is based on the covariance method, it is known as the best method to measure the relationship between the relevant variables. It provides information about the magnitude or correlation of the relationship as well as the direction of the relationship as defined with positive or negative related of two independent data sets and it attempts to draw a line of best fit through the data of two variables [81].

The mathematical definition of the Pearson's Product Moment-Correlation is presented in Equation 35, Equation 36 and Equation 37.

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_x \sigma_y} \tag{35}$$

$$\rho_{X,Y} = \frac{\varepsilon[(X-\mu_x)(Y-\mu_y)]}{\sigma_x \sigma_y} \tag{36}$$

$$\rho_{X,Y} = \frac{\varepsilon[XY] - \varepsilon[X]\varepsilon[Y]}{\sqrt{\varepsilon[X^2] - (\varepsilon[X])^2}\sqrt{\varepsilon[Y^2] - (\varepsilon[Y])^2}} \tag{37}$$

In these notations symbol representation are as follows:

- ρ is the Pearson's correlation coefficient
- *X* and *Y* are the independent data sets
- *cov* is the covariance
- σ_x is the standard deviation of x
- σ_y is the standard deviation of y
- μ_x is the mean of x
- μ_y is the mean of y
- ε is the expectation

It has a value between +1 and -1. A value of +1 symbolizes total positive linear correlation, 0 is for no linear correlation, and -1 is for totaly negative linear correlation. In this study, Pearson Correlation method is applied on wind speed forecast data of the WPPs in order to select the reference *online* WPPs which are used in the power forecasting. In this approach, the power plants that have a similarity above a certain threshold value are selected as reference and the power forecasts are obtained by establishing a relationship based on the power estimates of these plants. Also, these similarity coefficients are later used as weighting coefficients in the combination phase.

3.3.2 RV Correlation

Although Pearson Moment Correlation is an important method in determining the similarities of one-dimensional data, it is not suitable in multi-dimensional data. The RV Coefficient was first introduced by Robert an Escoufier for measuring the similarity between two matrices [82]. This method is also sometimes called as vector or matrix correlation.

$$R_V = \frac{trace\{S^T T\}}{\sqrt{(trace\{S^T S\})x(trace\{T^T T\})}} \tag{38}$$

$$R_V = \frac{vec\{S\}^T vec\{T\}}{\sqrt{(vec\{S\}^T vec\{S\})(vec\{T\}^T vec\{T\})}} \tag{39}$$

$$R_V = \frac{\sum_i^I \sum_j^I s_{i,j} t_{i,j}}{\sqrt{(\sum_i^I \sum_j^I s_{i,j}^2)(\sum_i^I \sum_j^I t_{i,j}^2)}} \tag{310}$$

RV is calculated as presented in the Equation 38, Equation 39 and Equation 310. In this formulas *S* and *T* are two matrices which are have the same dimension. Similar to Pearson Moment Correlation, RV coefficient varies between -1 and 1 where 1 is for the highest similarity. In this thesis, RV coefficient is used on two dimensional data which consists of historical wind speed and wind direction weather forecast. According to this similarity index reference *online* WPPs are selected in order to be used in the power forecast calculations.

CHAPTER 4

PROPOSED TECHNIQUE: REGIONALSHWIP

Regional wind power forecasting is important especially for system planning. Most of the models in the literature make predictions of the wind power of the region by scaling up the power estimates of several reference *online* wind farms in the region. Unlike this approach, in the proposed method, first of all, the wind power forecast is produced for the *offline* power plants in the region and then the power forecast of the region is obtained from the sum of all the *online* and *offline* power plants in the region. With this method, for all the power plants in the country, it is possible to produce power estimates even if they are not monitored in the system, and these forecasts can be used for improving regional forecasts. Unlike previous similar systems, the most important factor in the implementation of this method is the availability of weather forecast data for all the coordinates in the country, not just *online* plants in the system.

Power production values are collected from the wind power analysis tools located in the plant area and these values are sent to RITM center every 3 seconds for the *online* WPPs. As stated in Figure 4.1 by using the NWP and power data inputs SHWIP and RegionalSHWIP models are run daily for every *online* and *offline* WPPs in the system in order to produce power forecasts for these plants [19, 47].

Table 4.1: NWP Sources

NWP Source	NWP Initial Boundary Condition	High Resolution Working Model	Horizontal Grid Spacing (km)
MGM	ECMWF	ALADIN	4 km
ECMWF	ECMWF	WRF	6 km
NCAR	GFS	WRF	6 km

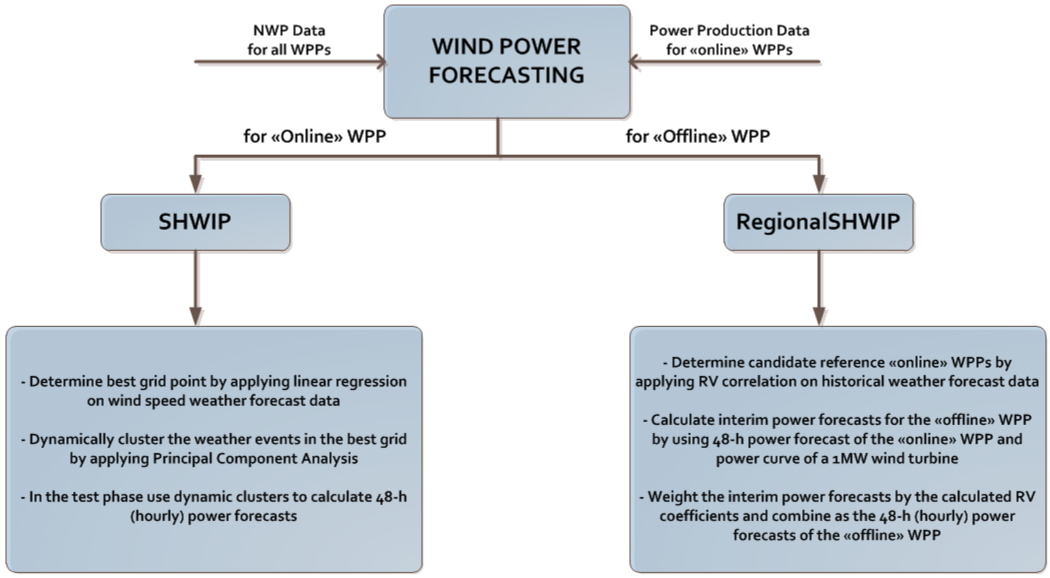

Figure 4.1: An overview of the power forecast models for *online* and *offline* WPPs in RITM Center

Within the scope of the RITM project, weather forecasts are obtained from three different sources as shown in Table 4.1. In RegionalSHWIP, each of the forecasts contribute to the model with equal weight. For the weather forecasts obtained from MGM [83], ALADIN [84] model is run on the initial boundary data obtained from ECMWF and forecast resolution is 4 km. For the GFS [85] ,initial data obtained by NCAR [86], and ECMWF [87], WRF [88] model is run daily and the distance between two grid points is 6 km in these two forecasts. For all these weather sources, 48-hour forecasts (in hourly resolution) are calculated daily.

The general architecture of the proposed work is presented in Figure 4.2. The inputs of Stage-1 are 48-hour power forecasts of the *online* plants obtained by SHWIP method [30, 19], 48-hour NWP data for the *online* plants and 48 hour NWP data for the *offline* plants. Individual offlineSHWIP power forecasts are produced in Stage-1 for the *offline* plants and these forecasts are combined to obtain the final 48-hour power forecasts for *offline* WPPs. Six alternative models are proposed that can be used for this phase, and their performance have been analyzed in the next chapter. The forecasts of the whole region are determined in Stage-3 by summing up the SHWIP power forecasts of the *online* plants and offlineSHWIP power forecasts of the *offline* plants.

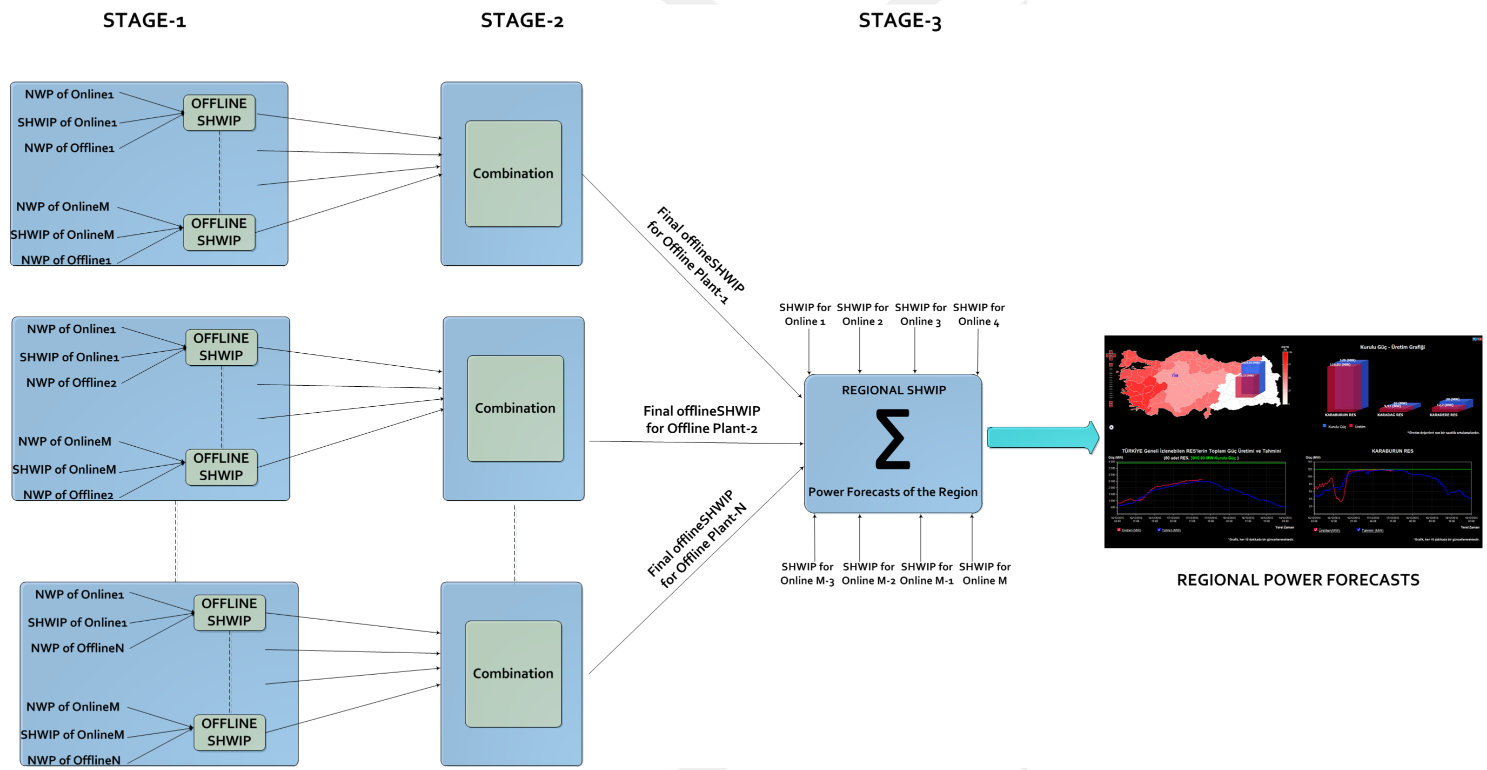

Figure 4.2: General Architecture of the Proposed RegionalSHWIP Method.

Stage 1. In this stage, firstly, 48-hour weather forecast data is retrieved for the candidate reference *online* plants and *offline* plants. In the constructed model, wind speed values are used in the generation of the power forecasts.

Let $v_{offline}$ and v_{online} are the 48-hour wind speed forecast vector for the time instance *t*. Equation 41 and Equation 42 represent the 48-hour wind speed input vector for the *offline* and *online* plant, respectively. From these values, capacity factor of each hour is calculated as given in Equation 43, by using a typical power curve of 1 MW wind turbine given in Figure 4.3. According to power curve in Figure 4.3, according to the wind speed forecast, expected capacity is determined by using the ranges in Equation 43.

$$v_{offline}[t, t+1, t+2, ..., t+47] \tag{41}$$

$$v_{online}[t, t+1, t+2, ..., t+47] \tag{42}$$

$$C_f(v) = \begin{cases} 0 & v < 5\,m/s \\ 0.014 & v = 5\,m/s \\ 0.07 & v = 6\,m/s \\ 0.1408 & v = 7\,m/s \\ 0.232 & v = 8\,m/s \\ 0.35 & v = 9\,m/s \\ 0.5044 & v = 10\,m/s \\ 0.6744 & v = 11\,m/s \\ 0.8348 & v = 12\,m/s \\ 0.9596 & v = 13\,m/s \\ 1 & v > 13\,m/s \end{cases} \tag{43}$$

$$offlineSHWIP(t) = \frac{C_f(v_{offline(t)})}{C_f(v_{online(t)})} \times \frac{C_{offline}}{C_{online}} \times SHWIP(t) \tag{44}$$

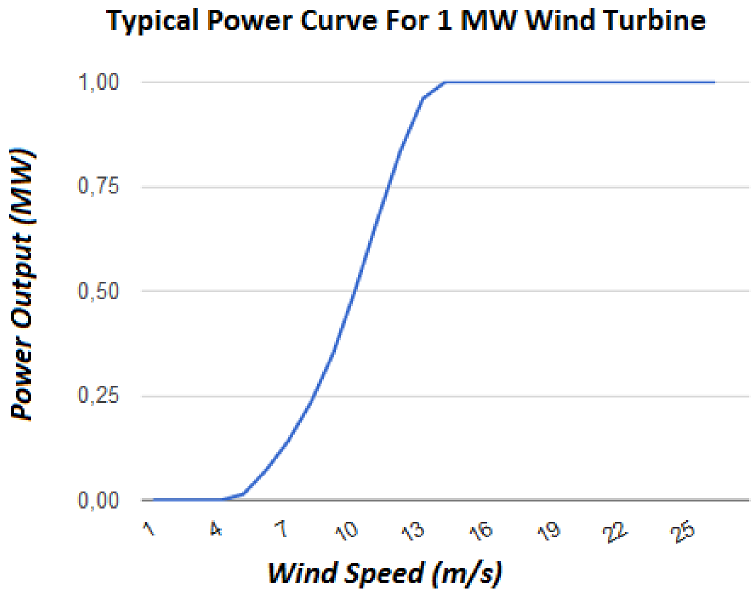

Figure 4.3: Power Curve of a Typical 1 MW Wind Turbine

From the capacity factor ratio, capacity of the plant and 48-hour SHWIP power forecasts of the *online* plant, 48-hour offlineSHWIP power forecast of the *offline* plant is calculated as given in Equation 44.

$$distance = \sqrt{(POnline_{lat} - POffline_{lat})^2 + (POnline_{lon} - POffline_{lon})^2} \tag{45}$$

$$weight = e^{\frac{-distance}{100}} \tag{46}$$

Stage 2. The main job of the second stage is selection of the reference *online* WPPs and combination of the individual offlineSHWIP forecasts that were calculated in the previous stage in order to generate the final power forecasts of the *offline* plant. For the selection of the *online* plants, geographical region, neighborhood relation and historical weather data similarities are considered. Hourly power values and NWP values for the wind speed and direction are used as the historical inputs while constructing the models.

To this aim, following six alternative models are defined:

Table 4.2: r value correlation table

r value	Relation
+.70 or higher	Very Strong positive relationship
+.40 to +.69	Strong positive relationship
+.30 to +.39	Moderate positive relationship
+.20 to +.29	Weak positive relationship
+.01 to +.19	No or negligible relationship
0	No relationship
-.01 to -.19	No or negligible relationship
-.20 to -.29	Weak negative relationship
-.30 to -.39	Moderate negative relationship
-.40 to -.69	Strong negative relationship
-.70 or higher	Very Strong negative relationship

- **Model 1 : Region Based Exponential Distance Weighting.** Each *online* plant within the *offline* plant's geographical region is selected as reference and individual offlineSHWIP forecasts are weighted in inverse proportion to the distance as shown in Equations 45 and 46.

- **Model 2 : Region Based Average Weighting.** Each *online* plant within the *offline* plant's geographical region is selected as reference and individual offlineSHWIP forecasts are equally weighted for each reference *online* WPP.

- **Model 3 : Neighbourhood Based Exponential Distance Weighting.** The nearest 15 *online* plants with respect to the *offline* power plant's central coordinates are selected reference and individual offlineSHWIP forecasts are weighted in inverse proportion to the distance given in Equations 45 and 46.

- **Model 4 : Neighbourhood Based Average Weighting.** The nearest 15 *online* plants with respect to the *offline* power plant's central coordinates are selected as reference. Individual offlineSHWIP forecasts are equally weighted for each reference *online* WPP.

- **Model 5 : Historical Wind Speed Similarity k Nearest Neighborhood based Weighting.**

  For this model initially, for the data retrieval step, the following operations are performed and a flowchart for this model is presented in Figure 4.4:

  - For every plant, 48-hour daily weather (wind speed) forecast is obtained.
  - For every plant, 30-day historical weather forecast data is obtained. Weather forecasts are obtained as the average forecasts of the 4 (Experimentally shown that 4 is optimal as the nearest neighbor count) nearest grid points to the WPP's central coordinate.
  - For the *online* plants, 48-hour SHWIP power forecasts are obtained.

  After the data retrieval for the selection of the reference *online* WPPs, the following procedure is followed:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}} \tag{47}$$

  1. For an *offline* plant, from the historical wind speed vectors, a correlation coefficient is calculated for each *online* plant in the system, independent of the regional neighbourhood. For the correlation coefficient *The Pearson Product-Moment Correlation* equation whose formula is given in Equation 47 is used [89].

     In this equation, *x* value is the 30-day historical wind speed vector of the *offline* wind power plant and y value is the corresponding historical wind speed vector of the *online* WPP.
  2. According to r value correlation in Table 4.2 [1], k *online* WPPs whose correlation coefficient is greater than 0.4 are selected as the reference *online* plants.
  3. Finally from the selected k *online* WPPs final power forecasts of the *offline* plant are calculated as correlation coefficients are assigned as the combination weights.

[1] r value correlation table shows the borders of strong, moderate and weak relation between two data sets

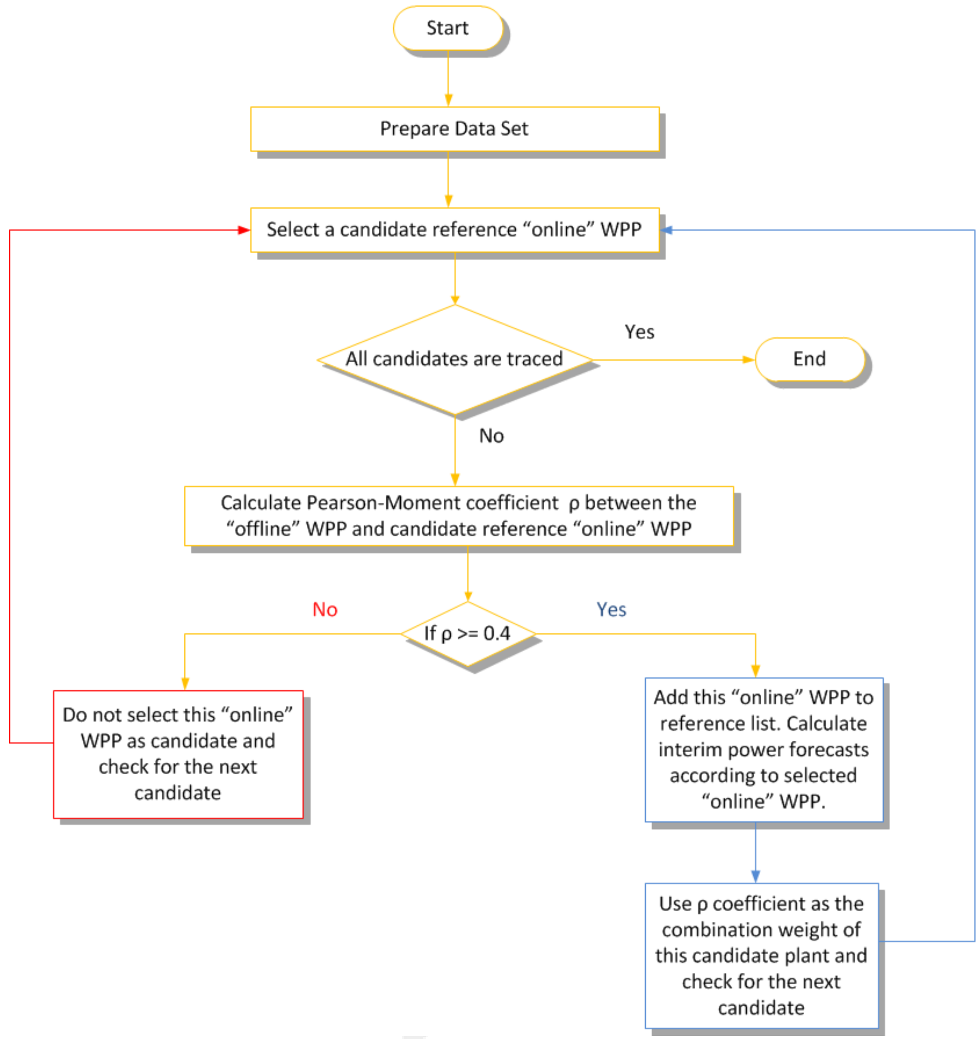

Figure 4.4: The flowchart of the Model-5

- **Model 6 : Historical 2D Wind Data Similarity Nearest Neighborhood based Weighting.** In Model-5, data similarity is investigated with respect to the most recent 30-day wind speed data of the *offline* plant, and reference *online* plants. While finding the similarity, *the Pearson Product-Moment Correlation* is used. In this model, not only wind speed forecast data but also wind direction forecast data are considered. Additionally, for extracting the seasonal pattern, the training day size is extended to 90 days.

  In this model, the Pearson Product-Moment Correlation is not suitable since the correlation calculation involves 2D matrices that include wind speed and wind direction data. Therefore, instead, we use *RV Coefficient* technique [90]. The RV coefficient is a generalization of the squared Pearson correlation coefficient. As in Pearson correlation, in RV coefficient values range in [0,1], where 1 denotes the highest similarity. As in the Model-5, the *online* plants whose RV correlation coefficients are higher than 0.4 are selected as reference *online* plants and these coefficients are used as the combination weights. Detail algorithm for Model 6 is presented in Algorithm 1 and a flowchart of the model is presented in Figure 4.5.

$$\rho = \frac{traceS^T P}{\sqrt{(traceS^T S)x(traceP^T P)}} \tag{48}$$

$$\rho = \frac{vec\{S\}^T vec\{P\}}{\sqrt{(vec\{S\}^T vec\{S\})(vec\{P\}^T vec\{P\})}} \tag{49}$$

$$\rho = \frac{\sum_i^I \sum_j^I s_{i,j} p_{i,j}}{\sqrt{(\sum_i^I \sum_j^I s_{i,j}^2)(\sum_i^I \sum_j^I p_{i,j}^2))}} \tag{410}$$

  While determining the most similar reference *online* WPPs, RV Correlation Technique is applied on historical weather forecast data [82]. RV coefficient is a similarity index between two matrices. It is calculated as in the Equation 48, Equation 49 and Equation 410 [82]. In this formulas, ρ is the RV coefficient, *S* represents historical NWP data matrices (contains wind speed and wind direction measurements) of the *offline* plant and *P* is for historical NWP data matrices for the reference *online* WPP. *I* is the total number of data is used in historical data which is 90x24.

Algorithm 1 Algorithm for RegionalSHWIP

Input: 90 days historical NWP data of the *online* and *offline* WPP, 48-h NWP test data of the *online* and *offline* WPP, 48-h SHWIP power forecast for the *online* WPP, 1 MW wind turbine power curve

Output: Hourly 48-h RegionalSHWIP Power Forecast for the *offline* WPP

```
for i = 1 to # of WPP in the System do
    Prepare the hourly historical NWP data for the most recent 90 day for all online
    and offline plant
    Prepare 48-hour NWP test data for all online and offline plant
end for
for i = 1 to # of offline WPP in the System do
    for j = 1 to # of online WPP in the System do
        Calculate 2D (wind speed and wind direction) similarity matrices for each
        online WPP by applying RV Coefficient Technique on the historical NWP
        data
        Select the online WPPs whose RV correlation coefficient is greater 0.4 and
        add these plants to reference plant list
    end for
    for j = 1 to # of selected reference online WPP do
        Calculate interim power forecasts for the offline plant by using 48-hour NWP
        test data, 48-hour SHWIP power forecast for the online plant and power
        curve of a 1MW typical wind turbine
        Normalize the interim power forecasts according to installed capacities of
        the plants
        Weight the interim forecasts with the calculated RV coefficients
    end for
    Sum the all interim forecasts for the offline plant and divide to total weight
end for
```

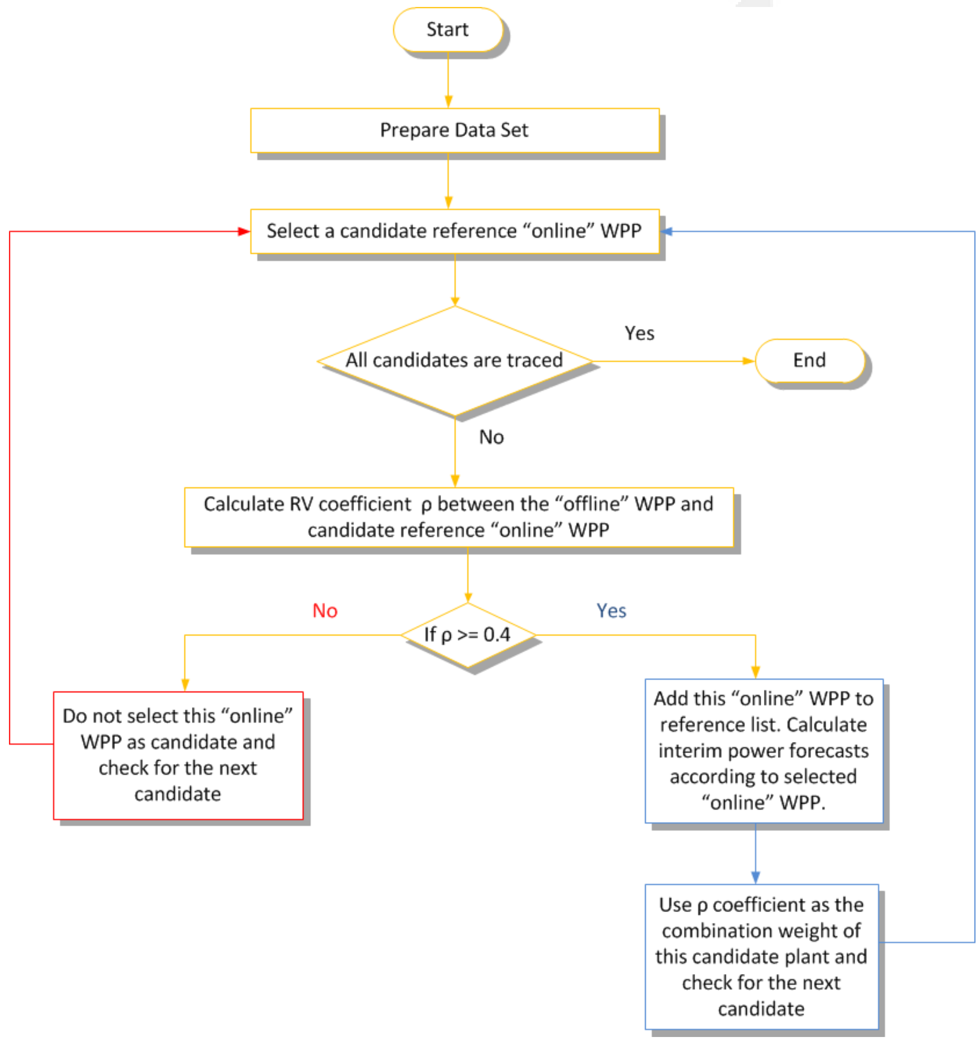

Figure 4.5: The flowchart of the Model-6

For all 6 models, the computational complexity for each of 6 models is *O(k x n)*, where k is the number of *offline* plants in the system and *n* is the total number of reference *online* plants used in the calculation of the individual offlineSHWIP forecasts. In the first four models *n* is determined from a specified region with the average number of reference plants are 20-25 for these models. On the other hand, for Model-5 and Model-6, *n* is the number of all *online* plants in the system (nearly 150 WPPs) and reference plants are selected among them. Since the value for *k* is expected to be limited and much lower than n, it brings negligible time cost and the growth functions of the models can be formalized as *O(n)*.

CHAPTER 5

LARGE SCALE ENERGY MONITORING AND FORECAST

In this section, the details of a generic and large-scale renewable energy monitoring and forecast system (henceforth, REMFS) and as the case study RITM system are presented. The architecture of this system proposal is presented in Figure 5.1. In the central monitoring and forecasting systems, besides the plant-specific forecasts, regional forecasts have a critical importance. With these forecasts, system operators can easily make day-ahead planning on regional and country basis, and energy resources are used efficiently. Therefore, the integration of a reliable regional estimation method into such a central system is very important for system reliability. The main capabilities of the generic large scale energy monitoring and forecast system are listed below and they are described in the following subsections.

- Data collection facilities
- Forecast facilities
- Data presentation facilities

5.1 Data Collection Facilities

The main data sources of REMFS are (1) systems, sensors and devices of renewable energy plants including (but not limited to) wind, solar, hydro-power, and wave energy plants, and (2) national and international weather services. As depicted in Figure 5.1, there are mainly three types of data that are retrieved from the plant sites in an automatic and periodic manner:

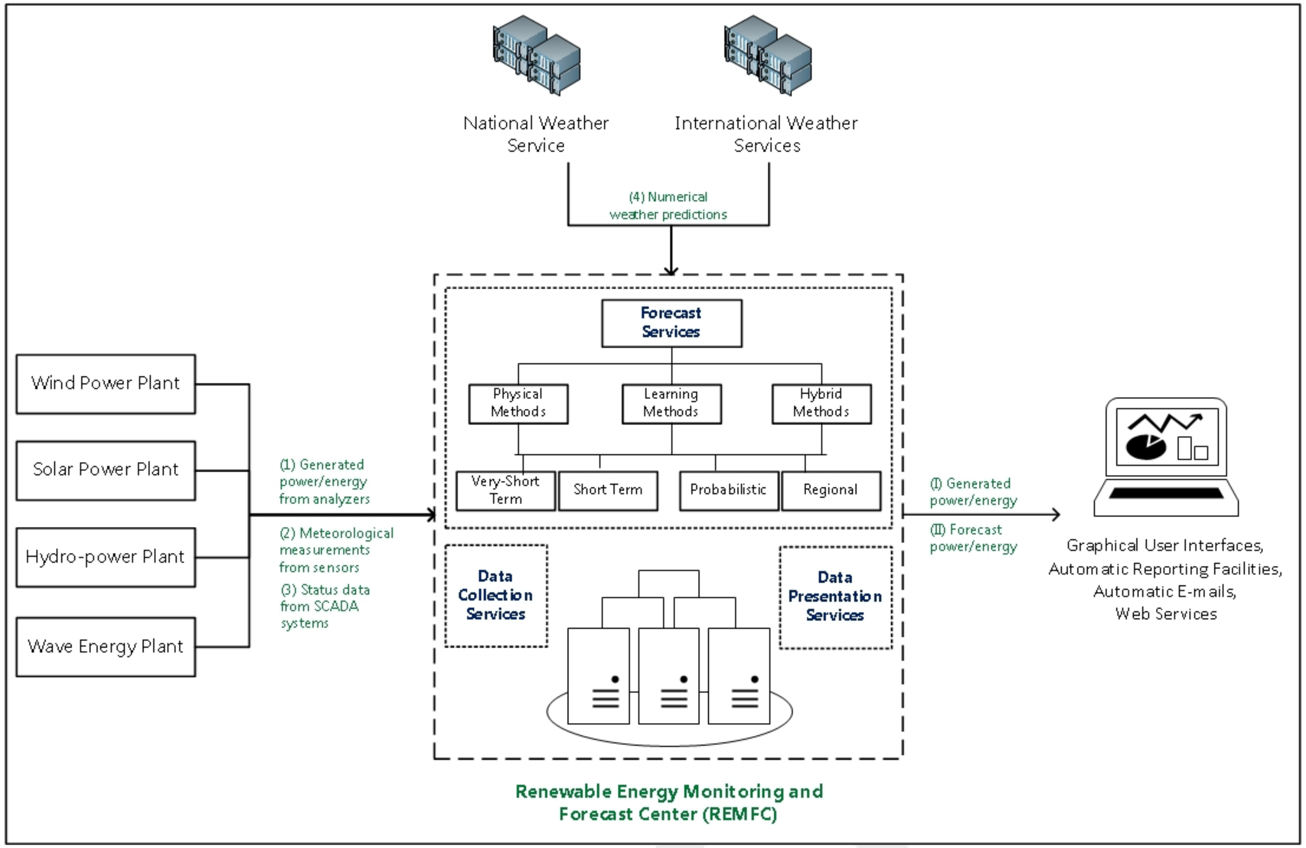

Figure 5.1: The architecture of the renewable energy monitoring and forecast system

- Measurements of energy and/or power analyzers deployed at the plants which correspond to actual generations of the plants.
- Measurements of meteorological sensors such as wind speed and direction for wind plants, solar radiation for solar plants, and precipitation and temperature for hydro-power plants, and from the corresponding meteorological sensors for other types of plants.
- On/off status of the generating units of the plants from their SCADA systems, whenever applicable.

The above-listed types of data are retrieved by the DCSs of the overall system and stored in a database at REMFS.

Apart from the data obtained directly from the plants, NWPs produced by weather services are known to be significant features for learning-based energy and power forecasting for renewable energy plants. Within the proposed architecture, NWPs at different ranges (such as very short-range, short-range, medium-range, etc.) can be automatically obtained and stored in the REMFS's database, again with the corresponding services (DCSs). It should be noted that, national weather services usually provide high-resolution NWPs for the country or region, with a large number of grid points for the particular country/region, while international weather services such as ECMWF and NOAA usually provide NWPs with low resolution with less number of grid points for the country/region under consideration. As a final note on data collection, there may be several means of data collection: some data may be presented by the related providers through file sharing, while others may be provided through Web services or IPC methods.

5.2 Forecast Facilities

As mentioned earlier in this chapter, power and energy production forecasts of renewable energy resources are particularly important for two main reasons:

- The organization within the country that is responsible for the operation of the electrical grid (which is usually called TSO or grid operator) requires energy forecasts from the renewable energy plants to balance electricity demand and supply, and to perform this task in an optimized way in order not to waste the potential of natural resources.

- The corresponding power/energy forecasts are also useful for the plant owners themselves to optimize their daily and maintenance operations and to optimize their gain in the energy market.

Hence, considerable research effort has been devoted to renewable energy forecast in a high-performance fashion, as reviewed in the previous section.

As demonstrated in Figure 5.1, Forecast Services of the overall system, REMFS, can be considered in three classes first: physical methods, learning methods, and hybrid methods. Physical methods are usually based on the geographical properties of the plant site and on related methods such as CFD to transform NWPs to power/energy generation forecasts. The learning methods will further be elaborated below and hybrid methods (in the system proposal) correspond to those methods that produce combined methods using physical and learning methods.

In a good designed montioring and forecast center, all of these approaches must serve for different forecast problems such as very-short term, short term, probabilistic and regional forecasting

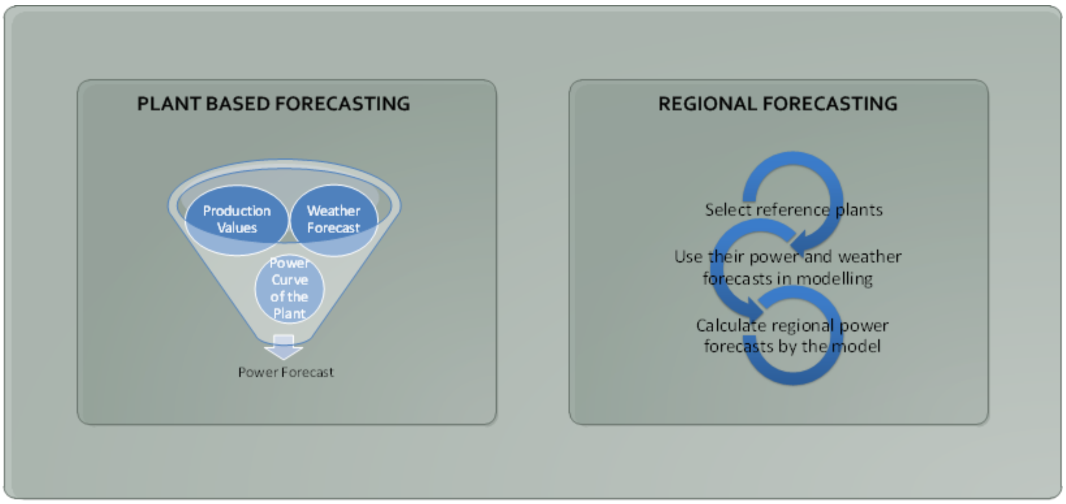

Figure 5.2: Plant Based and Regional Based Forecasting

- **Very-Short Term Forecasting:** The very short-term prediction models aim to predict the production amount of the power plant in the next few hours by using time series approaches such as AR and ARX, especially by using recent actual power production of the power plants. [91, 92]. These forecasts are important for power plants especially in the intra-day energy market operations.

- **Short Term Forecasting:**In this type of prediction problem, the aim is to accurately predict the power estimates of the power plants for the next few days using statistical, physical and hybrid approaches [93, 94]. Such estimates are one of the cornerstones of a good monitoring system. Because these forecasts have an important economic dimension as they are used in day-ahead market and planning.

- **Probabilistic Forecasting:** In a good monitoring and forecasting center, probabilistic forecasts within some confidence intervals should be presented as well as point power estimates [95]. In such estimation methods, it is very important to present the production estimates of the power plants within certain intervals in terms of planning and efficient use of energy resources.

- **Regional Forecasting:** Although it is desirable to monitor the power values of all power plants in a monitoring and forecast system and to create prediction models according to these values, it is not possible in terms of economic costs. Therefore, in a good monitoring and forecast system, an effective regional estimation method is needed for the whole system or for certain regions [96, 97]. Regional prediction methods generate power estimates for the entire region by using the data of some reference power plants in the region as indicated in the Figure 5.2. As indicated in the Figure 5.2, power values and weather forecasts of the plant are directly used in the production of power plant based power estimation, while in the regional estimation approach, the total power estimation of the region is made by selecting some reference plants in the region.

All of the above-mentioned intelligent data analysis methods are included in REMFS, and both the individual outputs of singleton methods and/or the combined outputs of hybrid methods and ensemble learning methods can be presented to the system users.

5.3 Data Presentation Facilities

REMFS has a number of GUI applications, automatic and periodic reports, automatic e-mails, and Web services as its data presentation modules. These modules enable interested parties (system users) to access both the actual generations and generation forecasts stored in the system center. The system center hosts a suite of software called Data Presentation Services to interface with and supply data to these modules. In a central system, these softwares have an important place in terms of direct access of the user to information and results.

Details of these modules are provided below:

- **Desktop Applications:** REMFS includes a suite of desktop, Web-based, and mobile GUI applications to help its users access the generation and forecast data. Desktop applications usually facilitate detailed analysis of the underlying data, including detailed long-range reports, long-range queries for historical data (similar to a historian application), and graph, table, and map-based visualizations.

- **Web-Based Applications:** These applications will facilitate access to the generation and forecast data with some limitations on the time range of queries and reports. Yet, these applications are easier to maintain due to centralized installation at the Web server, compared to desktop applications, which require updates to the computers of the users.

- **Mobile Applications:** These applications will enable ubiquitous access to the data stored at REMFS. However, similar to Web-based applications, there may be some limitations on the time range of queries and reports within mobile applications.

- **Automatic Report Generation and Periodic E-mails:** The system also has facilities to produce general and plant-based automatic reports which include the error rates of previous forecasts and new forecasts. These reports can be periodically sent to system operators and plant owners through automatic e-mails.

- **Web Services:** The data collected and produced within REMFS can also be supplied to related stakeholders like the grid operator and plant owners through Web services which also include an authentication mechanism.

5.4 Wind Power Monitoring and Forecast System: A Case Study

RITM project was started in 2010 by TUBITAK [13, 98]. Aim of this project is to forecast the electrical power to be generated from the wind, using the meteorological data received from NWP model outputs and the data received from the power quality analyzers. Project is launched for YEGM. Project is made up of two phases and first phase is conducted 14 pilots WPPs selected from the different regions of the country [99]. In the second phase of the project, all power plants in the country are started to be included in the system after the forecast performance criteria were met for the pilot plants used in the first phase. By the end of 2020, approximately 180 wind farms with a total installed capacity of 8100 MW can be monitored instantly and daily power estimates are made for these power plants in the center. Power forecasts produced in the system are used by the TSOs in grid management and planning while they are used by the power plant owners in the energy market. For a better planning they need reliable power forecasts for all kind of energy sources. Also a reliable wind power forecast system is inevitable for the plant owners. WPP owners attend the energy markets with their day-ahead and very short-term power forecasts and according to the accuracy of these forecasts they get profits or pay penalty. This system solves these problems for the wind energy sources.

General architecture of RITM system is shown in Figure 5.3. In this structure, inputs required for power forecasts are provided from the plant site and the meteorological stations and these values are stored in the central servers.While actual production values are obtained instantly from power quality devices, weather forecasts are obtained from different meteorological sources as hourly average values on a daily basis. Then, the power prediction algorithms developed by using these inputs are run on the servers and the estimation-power values can be displayed by the user software.

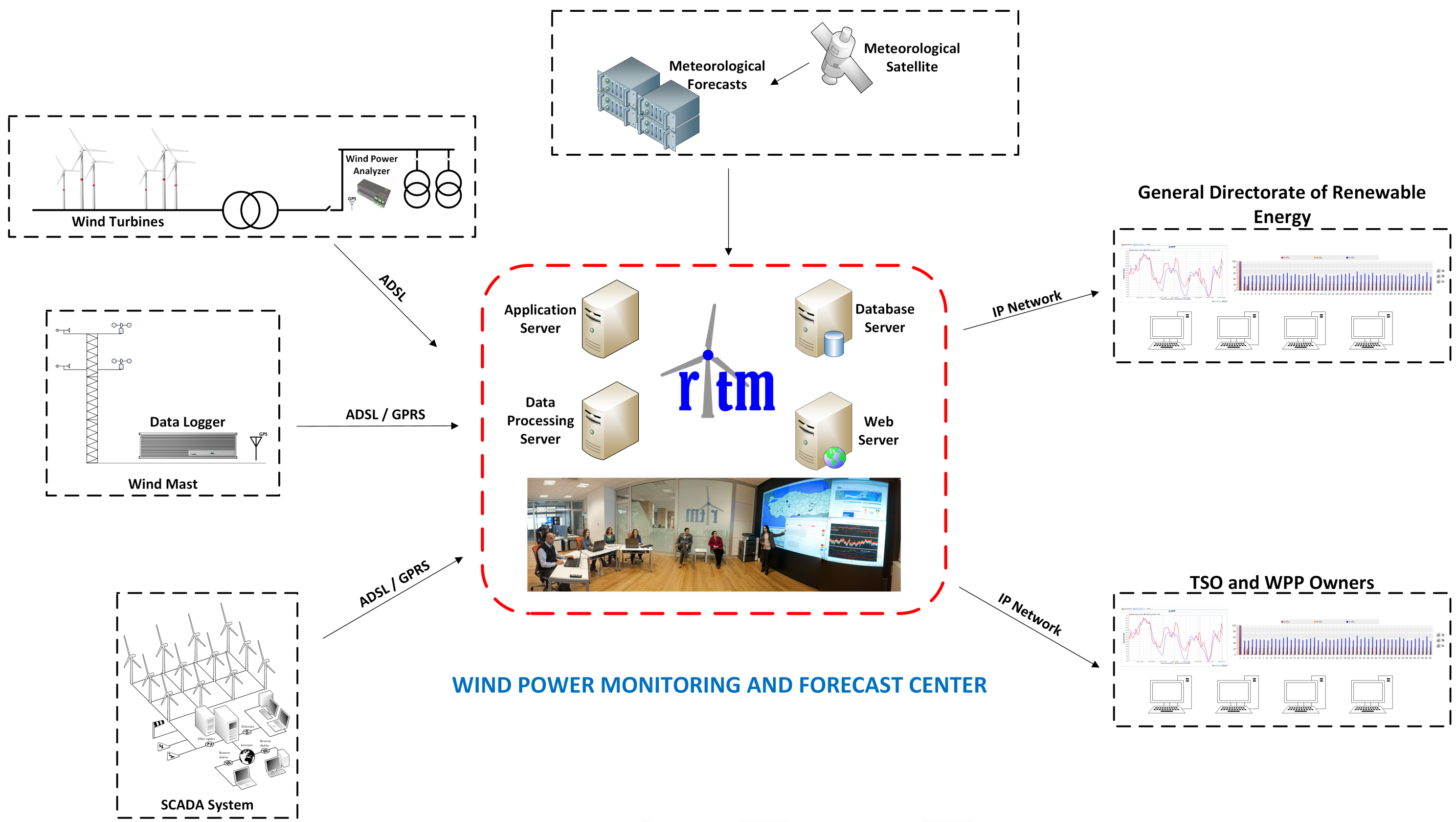

Figure 5.3: The general architecture of the RITM system.

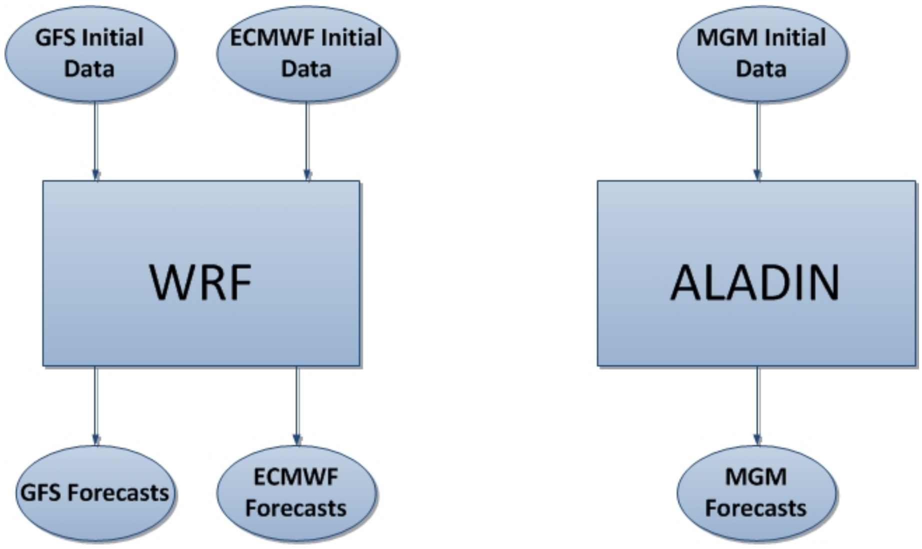

Figure 5.4: Numerical weather forecasts in RITM system

In order to produce power estimations, the most important component of the system is to collect the parameters related to the wind in the center. For this purpose, estimation inputs from the plant site and meteorological sources are collected in the system with many different methods. The inputs collected from the plant site consist of atmospheric information, such as wind speed and direction in the field, real production power value and SCADA data on turbine basis [100].

NWPs are the main input of the power forecast algorithms. Due to their core functionality, these inputs are included in the system from three different sources, namely MGM, GFS and ECMWF. As shown in Figure 5.4, GFS and ECMWF are served as the initial input of the WRF model, while MGM is served as the initial input of ALADIN model. For the GFS and ECMWF sources WRF model run at RITM center after initial data are download from the weather sources. ALADIN [101] model is run at the servers located in the MGM and outputs are taken to RITM center. These forecasts cover all country and the grid resolution is 6 km for GFS and ECMWF and 4 km for MGM forecasts [13].

Power forecast algorithms running at the center generate power estimates for various problems by using inputs from the plant site area and meteorological stations. In the center, algorithms for very short-term [102], short term [19, 15, 28], probabilistic [103] and regional [47] wind power forecasting problems are in operation.

RegionalSHWIP algorithm which is developed within the scope of this thesis has been integrated to RITM system by at end of 2019. With this developed model, in addition to producing regional estimates in the country, daily power estimates can be created for all wind power plants in the country even if their production values are absence in the system. Furthermore, with the integration of this model into the RITM system, geographic region-based and load distribution region-based daily power estimates can be automatically generated from the system and these predictions can be used effectively in regional day-ahead planning.

Finally, in order for the system users to view power forecast results obtained from the models, user interface software is required. In order to fulfill these needs, three main graphical user interface modules namely Real-Time Monitoring and Querying Module, Map-Based Monitoring and Forecast Module and Web-Based Monitoring and Forecast Module serve for system users.

CHAPTER 6

EXPERIMENTAL RESULTS

In this section, detailed results of the experiments conducted for the wind power forecast and wind resource assessment problems of the ReginonalSHWIP model are presented.

6.1 Experiments for Wind Power Forecasting

In this section, experimental evaluation results are presented for the wind power forecasting problem. Firstly, data used in the experiments is introduced. Experiments on power forecasts for *offline* plants are presented. Following this, the experiments on regional wind power forecasting are presented in the next section.

6.1.1 Data Used In Experiments

In the experiments, we used 10 *offline* WPPs in order to evaluate the forecasting accuracy of the proposed method. Additionally, 9 regions, determined by TSO, are used for evaluating the performance of the regional power forecasts. As the ground truth, we used the power values for these *offline* WPPs and the regions for the period between 1 January 2017 to 31 December 2017, which are obtained from the study in [27].

Performance results are obtained for the hourly power forecast and power data in this time period. Error results are reported as average NMAE and NRMSE, by using Equation 61 and Equation 62.

In Equation 61 and Equation 62 , x_i is the real power, y_i is the power forecast at i^{th} hour, *C* is the installed capacity of the WPP and *N* is the total number of hours processed in the test phase which correspond 8760 (365x24) hour data in one year test region.

$$NMAE = \frac{\frac{\sum_{i=1}^{N} |y_i - x_i|}{N}}{C} \times 100\,\% \tag{61}$$

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{N} (y_i - x_i)^2}{N}}}{C} \times 100\,\% \tag{62}$$

6.1.2 Experiment 1: Analysis on Power Forecast for *Offline* Plants

In the first experiment, performance of the final offlineSHWIP power forecasts of the 10 *offline* plants for all of the 6 models are investigated in terms of NMAE and NRMSE error rates. These plants are selected from different geographical regions of country. Additionally, these models are compared with two other forecast models whose details are as follows:

- **Base Model:** In this method, the weather forecasts of the *offline* plant are directly used in the power curve in Equation 43 without checking a relation between reference *online* plants.
- **Persistent Model:** In this approach, mean power data of the previous day is assigned as the power forecasts of the next day for each 24-hour tuple.

According to error rates presented in Table 6.1 and Table 6.2, Model-6 error rates are lower than Model-5 in 7 plants. Also, for 7 of 10 WPPs the lowest error rates are obtained by Model-6. For three other plants, the error rates of all models are close to each other and these plants have high number of geographical neighbors. Also according to the experimental which are stated in Table 6.1 and Table 6.2 results Model-6 shows better performance compared to Base Model and Persistent Model in all of the plants.

Table 6.1: Results of all Models (In terms of NMAE %)

WPP	Model-1	Model-2	Model-3	Model-4	Model-5	Model-6	Base Model	Persistent Model
WPP1	13.27 %	13.36 %	12.87 %	13.47 %	12.90 %	**12.71** %	13.43 %	19.32 %
WPP2	16.14 %	16.42 %	16.66 %	16.70 %	15.42 %	**15.09** %	17.65 %	18.95 %
WPP3	**13.92** %	14.07 %	14.34 %	14.65 %	15.10 %	14.79 %	14.40 %	16.40 %
WPP4	13.75 %	13.89 %	13.73 %	13.78 %	**13.11** %	13.17 %	16.37 %	17.15 %
WPP5	17.97 %	18.17 %	18.22 %	19.56 %	17.93 %	**17.55** %	21.79 %	21.01 %
WPP6	15.54 %	16.39 %	15.58 %	16.17 %	15.72 %	**15.43** %	16.52 %	21.70 %
WPP7	15.49 %	15.09 %	16.83 %	16.26 %	14.76 %	**14.00** %	14.99 %	18.14 %
WPP8	16.19 %	16.40 %	17.22 %	18.75 %	16.02 %	**15.84** %	20.32 %	21.12 %
WPP9	17.14 %	17.33 %	17.04 %	17.27 %	**16.59** %	16.70 %	19.43 %	19.21 %
WPP10	17.03 %	17.25 %	16.67 %	17.07 %	14.70 %	**14.38** %	18.83 %	20.78 %

Table 6.2: Results of all Models (In terms of NRMSE %)

WPP	Model-1	Model-2	Model-3	Model-4	Model-5	Model-6	Base Model	Persistent Model
WPP1	19.24 %	20.46 %	18.57 %	19.70 %	17.91 %	**17.72** %	19.87 %	27.07 %
WPP2	24.48 %	24.91 %	24.90 %	25.01 %	23.82 %	**23.35** %	26.89 %	25.93 %
WPP3	**20.74** %	20.97 %	21.41 %	21.76 %	22.43 %	22.05 %	22.82 %	24.24 %
WPP4	19.56 %	19.75 %	19.48 %	19.54 %	**18.48** %	18.55 %	23.68 %	24.08 %
WPP5	25.83 %	26.90 %	26.67 %	28.94 %	26.20 %	**26.06** %	31.50 %	30.12 %
WPP6	22.40 %	23.35 %	22.35 %	22.88 %	22.85 %	**22.22** %	24.66 %	29.98 %
WPP7	21.73 %	21.21 %	23.29 %	22.48 %	21.07 %	**20.21** %	22.11 %	26.77 %
WPP8	23.95 %	23.98 %	24.48 %	26.39 %	**23.07** %	23.16 %	28.97 %	31.51 %
WPP9	24.71 %	25.19 %	**24.00** %	24.27 %	24.13 %	24.28 %	28.59 %	26.52 %
WPP10	25.14 %	25.51 %	24.66 %	25.31 %	22.11 %	**21.60** %	28.37 %	29.62 %

Table 6.3: p-value and Standard Deviation Test Results between Models

WPP	Model-1 vs Model-6	Model-2 vs Model6	Model-3 vs Model-6	Model-4 vs Model6	Model-5 vs Model-6
WPP1	p = 0.01 σ = 2.17	p = 0.01 σ = 3.02	p = 0.03 σ = 1.07	p = 0.01 σ = 1.65	p = 0.59 σ = 0.60
WPP2	p = 0.02 σ = 0.58	p = 0.02 σ = 0.61	p = 0.03 σ = 0.71	p = 0.02 σ = 0.68	p = 0.27 σ = 0.12
WPP3	p = 0.01 σ = 0.97	p = 0.01 σ = 1.00	p = 0.05 σ = 0.36	p = 0.2 σ = 0.40	p = 0.35 σ = 0.25
WPP4	p = 0.14 σ = 0.60	p = 0.25 σ = 0.45	p = 0.01 σ = 0.92	p = 0.01 σ = 0.94	p = 0.43 σ = 0.21
WPP5	p = 0.01 σ = 0.48	p = 0.01 σ = 0.31	p = 0.01 σ = 0.43	p = 0.01 σ = 0.52	p = 0.01 σ = 0.39
WPP6	p = 0.01 σ = 2.24	p = 0.01 σ = 2.62	p = 0.01 σ = 1.98	p = 0.01 σ = 1.59	p = 0.01 σ = 1.34
WPP7	p = 0.01 σ = 1.34	p = 0.01 σ = 1.25	p = 0.01 σ = 1.52	p = 0.01 σ = 1.35	p = 0.14 σ = 0.20
WPP8	p = 0.04 σ = 1.24	p = 0.01 σ = 1.34	p = 0.01 σ = 1.85	p = 0.01 σ = 2.14	p = 0.03 σ = 1.31
WPP9	p = 0.01 σ = 1.94	p = 0.01 σ = 1.74	p = 0.01 σ = 2.11	p = 0.01 σ = 2.24	p = 0.07 σ = 0.64
WPP10	p = 0.01 σ = 1.40	p = 0.01 σ = 1.35	p = 0.01 σ = 1.25	p = 0.01 σ = 1.44	p = 0.04 σ = 0.74

One of the tests that reveals how the two groups of data sets are statistically different from each other is the t-test method [104]. This statistical tool is used with continuous data and tests whether any difference between the means of the two groups is significant. This test result is calculated according to the mean and standard deviation values in two data structures. If the *p* value obtained as a result of the test is 0.05 and below, it is evaluated that difference in the given two data groups is statistically significant and if it exceeds this value, it means that there is much similarity between the data structures. In Table 6.3, the statistical difference index (p value) of Model-6 with five other models is presented for all power plants. According to the results obtained, it is seen that the difference for the results of Model-6 from the first four models is statistically significant for almost all of the power plants. According to the results, it is determined the performances of Model-6 and Model-5 are similar. In these test results, it is observed that there is a statistically significant difference between Model-6 and Model-5 in 5 plants, while the estimation data for the other 5 plants are found to be statistically close to each other for these two models. At this point, it can be evaluated that adding wind direction similarity to the model for these plants brings only a slight change in the reference *online* plant list selected compared to the other test plants.

$$\sigma = \sqrt{\frac{1}{n}\sum_{1}^{n}(x_i - \overline{x})^2} \tag{63}$$

Another criterion that gives information about the relationship between data structures is standard deviation. Standard deviation describes how the data spread according to the arithmetic mean. The standard deviation value is calculated from the values in the data as indicated in Equation 63. In this formula x_i represents the data element, $\overline{x}$ is used for the mean of the forecasts of selected model and n is the total number of hourly forecast data. In Table 6.3, the absolute value standard deviation difference values of the estimation data sets obtained between Model-6 and other Models are also presented. According to the values obtained here, it is determined that the absolute standard deviation differences are more in the models with statistically significant differences are high. Also for the wind power plants which have more installed capacity, the absolute standard deviation differences are higher compared to small WPPs.

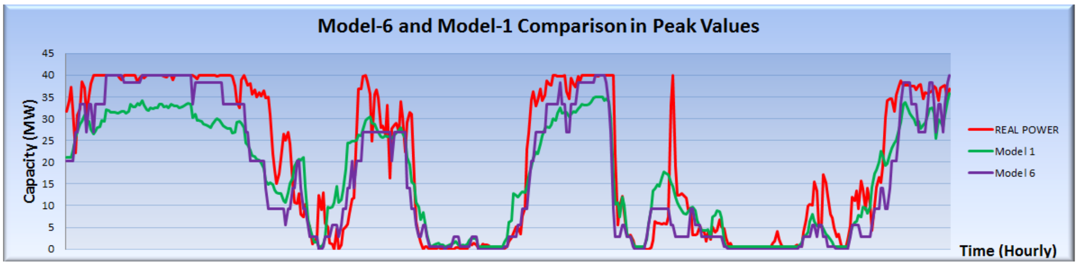

Figure 6.1: Model-6 and Model-1 Comparison in Determination of the Peak Values

Compared to other models, Model-6, which considers wind direction while finding similar *online* plants, appears to be more stable especially in the determination of the peak values. We can define the peak points for a power plant as the points where the production value is highest (between 90 % and 100 % of the plant capacity) and the lowest (between 0 % and 10 %). According to this definition, there are 250 peak points falling within these ranges within the test period given in Figure 6.1. In this period, while Model-1 predicts 106 of these points correctly, it is seen that this number is 218 in Model-6. Although WPP1 has 40 MW installed capacity, maximum power forecast value obtained by the Model-1 is at the level of 35 MW, while Model-6 reaches the plant installed capacity as the power forecast values for the same period. These results also show that Model-6 provides about 40 % more improvement in accurately predicting peaks and troughs than Model-1.

All of the models are also compared in terms of seasonal performances during the one-year test period as shown in Figure 6.2. Seasonal error rates are calculated as the average NMAE rates of the 10 *offline* test plants during the seasons. According to the seasonal performances, for all of the seasons, Model-6 shows better performance compared to other models on the average. Similarly, for all of the models, the lowest error rates are obtained in the autumn season while the highest error rates are obtained in the winter season. This situation is merely related with the weather forecasts. Generally, the weather conditions in the winter season are more volatile and the accuracy of the weather forecasts are low compared to other seasons. On the other hand, in summer or autumn, weather conditions are more stable and this issue makes weather forecasts more accurate. Since the weather forecasts are the main and indispensable input to the power forecasts, the seasonal performances show differences.

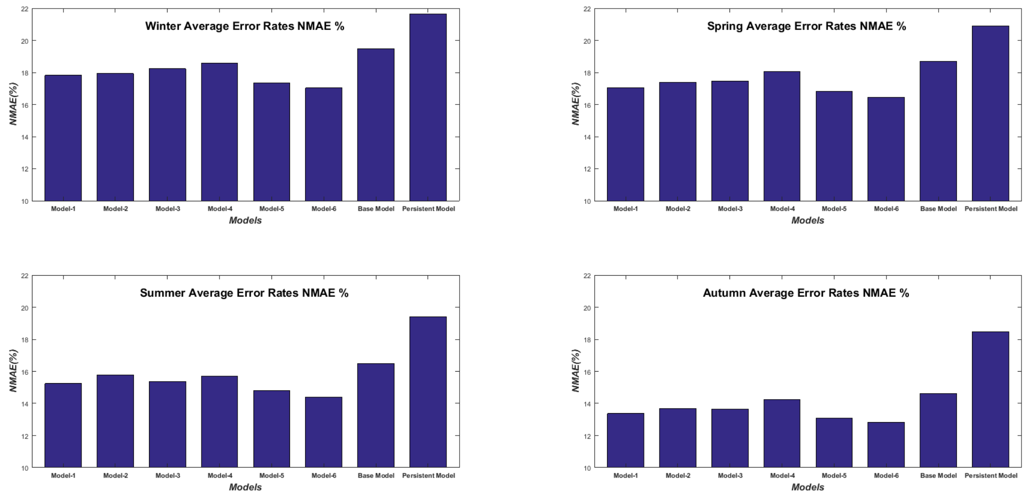

Figure 6.2: Seasonal Performance of the all Models

6.1.3 Experiment 2: Analysis on Regional Power Forecasting

In this experiment, performance of the RegionalSHWIP for 9 regions, determined by TSO of Turkey is evaluated. These regions are determined by the transmission system operator according to distribution of the electricity transmission lines in the country. All of the regions have different number of plant density and total installed capacity. In this experiment, proposed models are compared (in terms of regional bases) with three different approaches described in the literature, whose details are as follows:

- **Base Model:** In this method, for every region, 3 largest (in terms of installed capacity) *online* plants are selected. Then their total of power forecasts is upscaled as the forecast of the whole region with the *upscaling coefficient*. This coefficient is the ratio of total installed capacity of region to total installed capacity of the 3 largest plants in the region.

- **ANN based Model:** In this approach, 1-year of training power forecasts data obtained by SHWIP model (10 months for training, 2 months for validation) of 3 largest *online* plants in the region are given as input to the Feed Forward Back Propagation based ANN [15] with the real production data of the whole region in this 1-year period, which covers the whole 2016 year values. While constructing the network, 3 layers (Input, Hidden, Output) are used with 5 to 10 neurons. The model is implemented in Matlab [105].

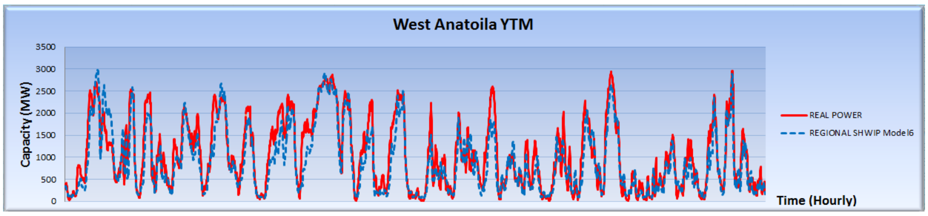

Figure 6.3: Model-6's Results for West Anatolia YTM

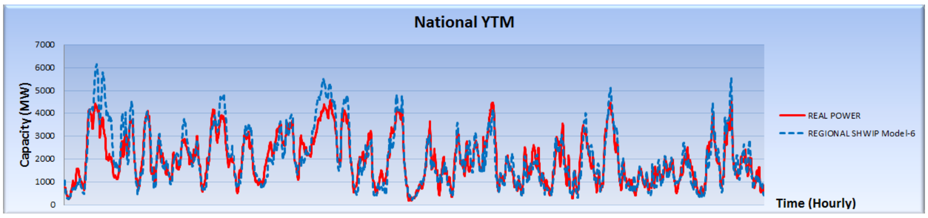

Figure 6.4: Model-6's Results for National Anatolia YTM

- **SVM based Model:** Similar to the ANN based model, for the same reference *online* plants' 1-year of training data is used for constructing the SVM model with the same inputs in ANN model. The model is implemented by using *libsvm* library with the interfaces for Matlab [106, 105]. For the kernel, Radial Bases Function (RBF) is used and optimum values are selected for γ (gamma) and c (penalty) parameters [15, 28]. (γ=5, c=2)

Unlike geographical regions, TEIAS identified 9 regions according to the distribution of electricity lines in the country which they named these regions as load distribution center. The results of the all models for these regions are presented in Table 6.4 and Table 6.5. Also, total number of plants in the region and total installed capacities of the regions are presented in the result tables. In general, it is seen that wind farms are concentrated in the western part of the country, as this region has higher wind potential compared to other parts of the country. According to the results, the lowest error rate in the regional bases is obtained for the West Anatolia YTM and maximum error rates are obtained for the Middle Black Sea YTM. According to the regional estimation results, it is observed that there is an inverse correlation between the total installed power of the region and the average error rate, and it is observed that the error rate is lower in regions with high installed capacity. For the total of all WPPs, namely National YTM, NMAE error rate is calculated as 5.25 % for Model-6.

There is a correlation between the error rates and total installed capacities of the region and generally lowest error rates are obtained in the largest regions. The reason behind this result is that, if the number of wind power plants in a region is high, than a forecast error in one plant can be compensated with better performance of another plant in the same region. Hence the error rate drops on the total. In addition to this, SVM and ANN show different performances for the different regions. The reason behind this factor is that, for the regions where these two models show better performance, the 3 largest WPPs in the region may represent the whole region better in comparison to the other regions where the models show lower performance. For the West Anatolia YTM and National YTM, the forecast graphics of Model-6 are presented in Figure 6.3 and Figure 6.4, respectively. In these graphics, red line represents the actual measured power data in the test region and blue scatter line represents the corresponding power forecasts for the same period. As observed in the figures, the overlap between the forecast and real production value is very high.

6.2 Experiments for Wind Power Assessment

6.2.1 Experiment 1: Analyzing the Model on Candidate WPPs

Another important advantage of the RegionalSHWIP compared to other approaches in the literature is that it can determine the wind power potential of a candidate plant with a given lat/lon coordinate. In this experiment this feature of the RegionalSHWIP is investigated in four different city from different regions of the country. For the analysis, RegionalSHWIP is applied on 100 MW candidate WPPs from four different locations within the city as shown in Figure 6.5. In these maps, the red areas indicate the locations with high wind potential, and the blue areas denote those with low potential. These locations are selected according to different wind speed densities in the wind maps. The results of the four different candidate WPPs are compared with the wind potential of the corresponding locations in the wind atlas in terms of average efficiency percentages as defined in Equation 64.

$$\varphi = \frac{\sum_i^N \hat{P(i)}}{N} \tag{64}$$

Table 6.4: Results of all Regions (In terms of NMAE %)

Region Name	Capacity (MW)	Model-1 Error	Model-2 Error	Model-3 Error	Model-4 Error	Model-5 Error	Model-6 Error	Base Model	ANN	SVM
Trakya YTM	680.35	9.61 %	9.62 %	9.53 %	9.54 %	9.46 %	9.42 %	11.63 %	9.23 %	**9.16** %
North West Anatolia YTM	305	9.79 %	9.82 %	10.17 %	10.04 %	9.92 %	**9.74** %	11.83 %	13.19 %	10.55 %
West Anatolia YTM	3287	8.10 %	8.12 %	7.59 %	7.61 %	7.35 %	**7.11** %	10.10 %	8.30 %	8.40 %
Middle Anatolia YTM	442.5	9.65 %	9.62 %	9.95 %	10.13 %	9.33 %	**9.28** %	12.97 %	11.32 %	11.64 %
West Mediterranean YTM	341.5	10.74 %	11.48 %	11.08 %	10.90 %	10.31 %	**9.54** %	11.00 %	10.77 %	10.22 %
Middle Black Sea YTM	163.8	14.44 %	14.94 %	15.22 %	15.18 %	14.64 %	**14.28** %	16.49 %	15.33 %	15.07 %
Southeastern Anatolia YTM	188.2	9.68 %	9.78 %	9.17 %	9.36 %	9.06 %	8.54 %	10.60 %	**8.15** %	8.34 %
East Mediterranean YTM	804.1	7.55 %	7.61 %	8.10 %	7.39 %	7.16 %	**7.13** %	10.82 %	8.98 %	8.71 %
National YTM (All WPPs)	6212	6.95 %	6.99 %	6.93 %	6.68 %	5.58 %	5.25 %	8.06 %	6.41 %	**5.12** %

Table 6.5: Results of all Regions (In terms of NRMSE %)

Region Name	Capacity (MW)	Model-1 Error	Model-2 Error	Model-3 Error	Model-4 Error	Model-5 Error	Model-6 Error	Base Model	ANN	SVM
Trakya YTM	680.35	12.82 %	12.83 %	12.77 %	12.78 %	12.12 %	12.08 %	14.30 %	12.33 %	**12.00** %
North West Anatolia YTM	305	14.12 %	14.14 %	14.47 %	14.34 %	13.98 %	**13.65** %	16.55 %	17.59 %	14.62 %
West Anatolia YTM	3287	10.50 %	10.51 %	9.95 %	9.96 %	9.27 %	**8.93** %	13.35 %	10.94 %	11.17 %
Middle Anatolia YTM	442.5	13.37 %	13.31 %	13.75 %	13.95 %	12.49 %	**12.39** %	17.74 %	15.33 %	15.94 %
West Mediterranean YTM	341.5	15.41 %	16.39 %	15.87 %	15.63 %	14.83 %	**12.74** %	15.77 %	15.31 %	14.67 %
Middle Black Sea YTM	163.8	22.51 %	23.44 %	24.05 %	24.57 %	**22.90** %	23.35 %	25.28 %	24.72 %	24.71 %
Southeastern Anatolia YTM	188.2	13.76 %	13.76 %	12.88 %	13.22 %	12.49 %	11.08 %	14.99 %	**10.89** %	11.70 %
East Mediterranean YTM	804.1	9.94 %	10.02 %	10.61 %	9.76 %	9.35 %	**9.15** %	13.90 %	11.33 %	11.18 %
National YTM (All WPPs)	6212	8.65 %	8.69 %	9.61 %	8.35 %	7.33 %	6.85 %	10.54 %	8.44 %	**6.79** %

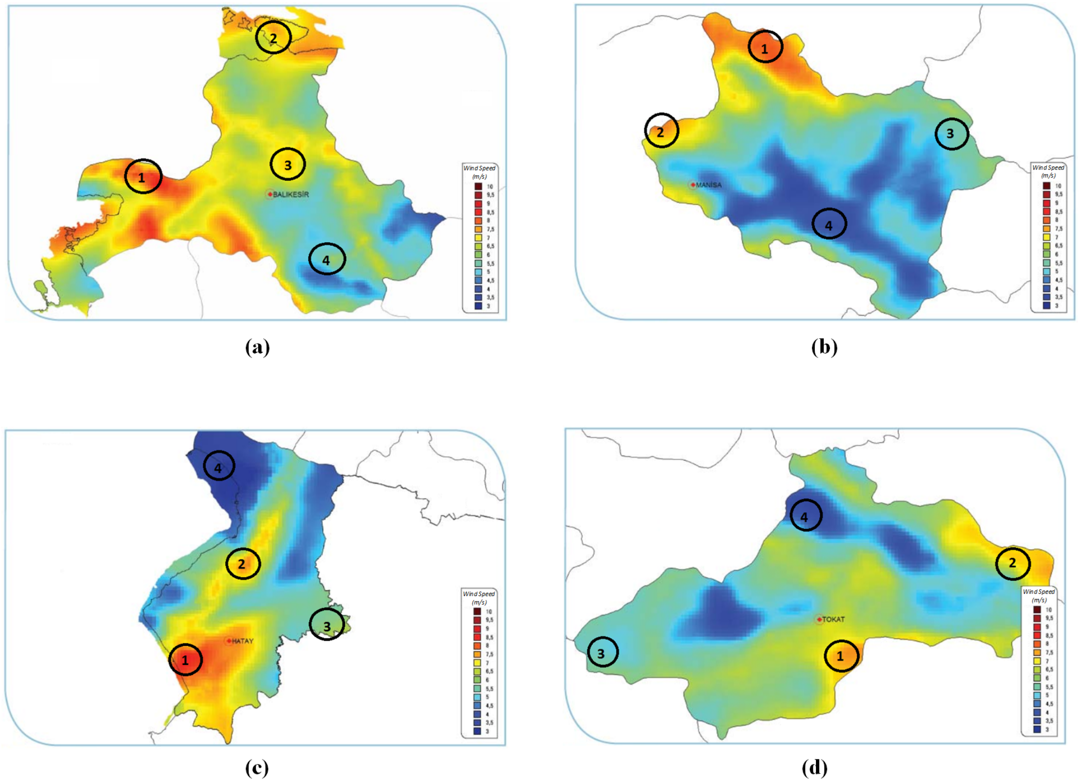

Figure 6.5: **(a)** Wind Map of Balikesir, **(b)** Wind Map of Manisa, **(c)** Wind Map of Hatay, **(d)** Wind Map of Tokat

In this formula, φ represents average efficiency during the period, $\hat{P(i)}$ is the power forecast at i^{th} hour and N is the total number of hourly data in the forecast period. The results are obtained for the time interval between 1 January 2018 and 1 July 2018. The forecasts are generated for each day of the forecast period, which covers the months from January to July. These daily forecasts are combined to generate the monthly power production forecasts. Then the monthly forecasts are compared against the monthly expected productions given in the wind atlas.

6.2.1.1 Balikesir Wind Map Results

Balikesir is located in the western part of Turkey, has high wind power potential, thus 20 of 180 real WPPs of the country are constructed in this city as shown in Figure 6.6 (a). In these city maps, *red* areas indicate the reference city, *brown* areas show the neighbor cities, *blue* areas indicate water and *black* areas show neighboring countries. Wind energy potential of Balikesir is shown in Figure 6.5 (a). In order to analyze the performance of the RegionalSHWIP, four 100 MW candidate plants (named as CP in the results) are selected in different parts of the city according to their wind potential shown in Figure 6.5 (a). For these four candidate plants, the results shown in Table 6.6 are obtained. In this table, Expected Production (MWh) is calculated by using annual wind speed values on a 1MW typical turbine wind power curve on the same forecast period [47]. For Balikesir, the results are consistent with the wind map of Balikesir and higher efficiencies are obtained in CP-1, CP-2 and CP-3 plant area in correlation with the average annual wind speed values of the candidate plant areas shown in Table 6.6 .

6.2.1.2 Manisa Wind Map Results

Manisa, a city in the western part of Turkey, has the largest WPP in the country (with 240 MW installed capacity. The wind potential map of the city is shown in Figure 6.5 (b). The results obtained for the four candidate plants are presented in Table 6.6. The results show that the production forecast and the expected production values are similar. The highest potential is calculated for the CP-1 WPP, which also has the

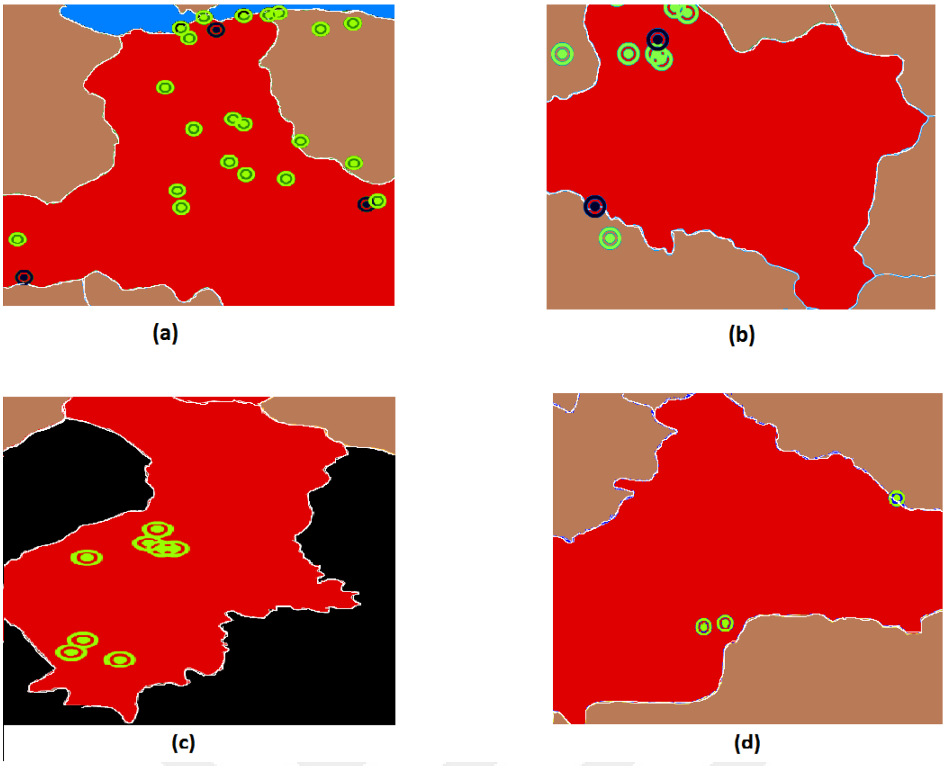

Figure 6.6: **(a)** WPPs in Balikesir, **(b)** WPPs in Manisa, **(c)** WPPs in Hatay, **(d)** WPPs in Tokat

highest expected production. Additionally, when we examine the distribution of the real wind power plants in Manisa as shown in Figure 6.6 (b), nearly all of the WPPs are established in the area of the CP-1 WPP.

6.2.1.3 Hatay Wind Map Results

Hatay is located in the southern part of the country, and according to wind potential atlas of Turkey, it is one of the city with high wind power potential. The wind potential map of Hatay is shown in Figure 6.5 (c). In order to analyze the assessment performance of the RegionalSHWIP, 4 different candidate wind power plants are selected in the map. According to the wind potential map, it is expected to have high wind power potential for the WPPs CP-1 and CP-2 compared to the plants CP-3 and CP-4. For these 4 candidate plants, total wind power production forecasts and efficiency percentages for the forecast period are given in Table 6.6.

Table 6.6: Forecast Result for all Cities

City	WPP	Total Production Forecast (MWh)	Hourly Efficiency Percentages (%)	Annual Wind Speed (m/s)	Expected Production (MWh)
Balikesir	CP-1	141853 MWh	33.58 %	9.0 m/s	147851 MWh
	CP-2	113277 MWh	26.81 %	8.5 m/s	124558 MWh
	CP-3	137292 MWh	32.50 %	8.0 m/s	98005 MWh
	CP-4	73335 MWh	17.36 %	6.5 m/s	47608 MWh
Manisa	CP-1	136107 MWh	32.22 %	9.0 m/s	147850 MWh
	CP-2	119534 MWh	28.29 %	8.5 m/s	124562 MWh
	CP-3	62962 MWh	14.9 %	6.5 m/s	47622 MWh
	CP-4	1950 MWh	0.46 %	4.0 m/s	2119 MWh
Hatay	CP-1	84866 MWh	20.09 %	9.0 m/s	147850 MWh
	CP-2	127537 MWh	30.19 %	8.5 m/s	124537 MWh
	CP-3	46346 MWh	10.97 %	6.5 m/s	47613 MWh
	CP-4	550 MWh	0.13 %	5.0 m/s	6346 MWh
Tokat	CP-1	149725 MWh	35.44 %	9.0 m/s	147866 MWh
	CP-2	133803 MWh	31.67 %	8.0 m/s	98017 MWh
	CP-3	38786 MWh	9.18 %	6.5 m/s	47616 MWh
	CP-4	20368 MWh	4.82 %	5.5 m/s	22776 MWh

Table 6.7: Hourly Efficiencies of Real WPPs in Hatay. WPP1 and WPP2 are located in CP-1 area, WPP3 and WPP4 are in CP-2 plant area.

WPP	Hourly Efficiency Percentages (%)
WPP1	22.30 %
WPP2	25.92 %
WPP3	24.89 %
WPP4	30.22 %

According to the results, for the CP-1 and CP-2 plants, the efficiency is higher than that of CP-3 and CP-4 plants, which is consistent with Hatay wind map shown in Figure 6.5 (c). However according to the results of the RegionalSHWIP, the efficiency in the area of CP-2 is higher than that of CP-1, which is in conflict with the wind potential map. As shown in Figure 6.6 (c), showing the locations of real plants in Hatay, the number of real plants is more in the area of CP-2 in comparison to CP-1. This supports the results obtained by the RegionalSHWIP. Additionally, the production efficiency of the real plants in CP-2 plant area is higher than the efficiencies of the plants in the area of CP-1, as shown in Table 6.7. This also supports the production prediction results of RegionalSHWIP. These values are calculated by the real production values of the plants at the same forecast period. WPP1 and WPP2 are located in CP-1 area whereas WPP3 and WPP4 are in CP-2 plant area.

6.2.1.4 Tokat Wind Map Results

Tokat is a city located in the northern part of Turkey. Among 18 city in this region, 3 of the 5 WPPs are established here. The wind potential map of the city is shown in Figure 6.5 (d). Production efficiencies predicted by RegionalSHWIP for the 4 test areas are given in Table 6.6. As seen in the table and the wind atlas map, the obtained results are completely consistent with the city's wind potential map. The highest values are predicted for the areas of CP-1 and CP-2. When we examine the distribution of the real WPPs as shown in Figure 6.6 (d), they are also located in these areas.

6.2.2 Experiment 2: Analyzing the Model on *Online* WPPs

In the second experiment we have analyzed the assessment performance of the model on real *online* WPPs by considering them as if they are *offline*. We have selected 4 different *online* plants from the test cities, and applied the model for the same forecast period for these plants. Due to confidentiality issues, we name the plants as WPP-x. The results for all these plants are presented in Table 6.8. According to the obtained results, total production forecasts are inline with real obtained power data for these plants and these results also support that the use of RegionalSHWIP is feasible for the preliminary site selection of a WPP.

Table 6.8: Forecast Result for *online* WPPs

City	WPP	Capacity (MW)	Total Production Forecast (MWh)	Real Production (MWh)	Hourly Efficiency Percentages Forecast (%)	Hourly Efficiency Percentages Real (%)
Balikesir	WPP-1	142.5 MW	203821 MWh	21317 MWh	32.91 %	34.77 %
	WPP-2	87 MW	124432 MWh	131960 MWh	32.90 %	34.90 %
	WPP-3	16 MW	19035 MWh	18398 MWh	27.37 %	26.43 %
	WPP-4	52.8 MW	70978 MWh	75291 MWh	30.92 %	32.85 %
Manisa	WPP-1	250 MW	279342 MWh	304184 MWh	25.68 %	28.04 %
	WPP-2	120 MW	171331 MWh	184360 MWh	32.85 %	35.00 %
	WPP-3	57.2 MW	52535 MWh	63841 MWh	21.13 %	25.71 %
	WPP-4	35 MW	36826 MWh	45745 MWh	24.20 %	30.22 %
Hatay	WPP-1	30 MW	29869 MWh	36279 MWh	22.90 %	27.86 %
	WPP-2	60 MW	58601 MWh	69747 MWh	22.46 %	26.91 %
	WPP-3	45 MW	53602 MWh	51763 MWh	27.40 %	26.44 %
	WPP-4	18 MW	20365 MWh	19841 MWh	26.01 %	25.38 %
Tokat	WPP-1	12.8 MW	23694 MWh	23113 MWh	42.57 %	41.56 %
	WPP-2	42 MW	55068 MWh	58483 MWh	30.16 %	32.07 %
	WPP-3	85 MW	98432 MWh	105075 MWh	26.64 %	28.48 %
	WPP-4	30 MW	44664 MWh	40038 MWh	34.23 %	30.70 %

According to the test results obtained in these plants, the maximum average efficiency value is estimated as 42.57 % for the WPP1 plant in Tokat, while the actual efficiency value obtained in this plant is calculated as 41.56 %, also close to the forecasted efficiency. Similarly, in the WPP3 plant (21.13 %) in Manisa city where the lowest efficiency value is determined, it is seen that the actual efficiency value in the test area is around 25.71 % which is second lowest value in the plants.

CHAPTER 7

CONCLUSION AND FUTURE WORK

Wind energy is an important alternative and clean energy source with high potential. In order to benefit from this renewable energy source effectively there is a need for good planning. The aim of this work is constructing an upscaling wind power forecast model to be integrated to the RITM system. The proposed model is based on statistical methods. For the selection of the reference *online* plants and determination of the combination weights, 6 different models are proposed and evaluated. According the experimental results, performance of Model-6, which is based on historical wind speed and wind direction data similarity performs better than the other models both in plant basis and regional basis. With this approach, by using the weather forecasts and output of the *online* plants point forecasts, 48-hour power forecasts are produced for all *offline* plants in the country. As the final step, all of these power forecasts are accumulated as the forecasts of the whole region. Therefore, by using this method it is possible to generate power forecasts for all the plants in the system and for all the regions at the same time.

In this thesis also we investigate the performance of a recent power forecasting technique, RegionalSHWIP, for wind assessment in order to facilitate WPP site selection. Existing wind atlas is an important resource, however keeping wind atlas up to date requires a significant effort and cost. Hence, our motivation is based on the need for light-weight methods to facilitate the process and economize in terms of time and financial cost.

The results of the analysis conducted for different cities having varying characteristics in comparison to an existing wind potential atlas show the compatibility of the predicted wind potential with real cases. The results also reveal that the method can be used for predicting the wind power potential. Furthermore, the analysis demonstrates that, at certain locations, the forecasts obtained by the model reflect the wind potential more accurately than the wind atlas.

For future work, the proposed model can be extended to other renewable energy sources such as solar and hydro by using the related NWP parameters for these sources. Additionally, by using this model, preliminary investigation of the possible plant area can be conducted and after selection of the area, real wind power measurements can be obtained in the area for creating a power curve of the plant. This power curve can also be integrated into the RegionalSHWIP and be utilized in the micro-siting.

REFERENCES

[1] A. Growth, "Ren21 renewables 2018 global status report,"

[2] J. H. Seinfeld and S. N. Pandis, *Atmospheric chemistry and physics: from air pollution to climate change*. John Wiley & Sons, 2016.

[3] C. B. Field, *Climate change 2014–Impacts, adaptation and vulnerability: Regional aspects*. Cambridge University Press, 2014.

[4] H. Iwata and K. Okada, "Greenhouse gas emissions and the role of the kyoto protocol," *Environmental Economics and Policy Studies*, vol. 16, no. 4, pp. 325–342, 2014.

[5] B. V. Solanki, K. Bhattacharya, and C. A. Cañizares, "A sustainable energy management system for isolated microgrids," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1507–1517, 2017.

[6] J. R. Andrade and R. J. Bessa, "Improving renewable energy forecasting with a grid of numerical weather predictions," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1571–1580, 2017.

[7] P. Enevoldsen and B. K. Sovacool, "Examining the social acceptance of wind energy: Practical guidelines for onshore wind project development in france," *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 178–184, 2016.

[8] T. Bauwens, B. Gotchev, and L. Holstenkamp, "What drives the development of community energy in europe? the case of wind power cooperatives," *Energy Research & Social Science*, vol. 13, pp. 136–147, 2016.

[9] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, "The state-of-the-art in short-term prediction of wind power: A literature overview," *ANEMOS. plus*, 2011.

[10] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, G. Conzelmann, *et al.*, "Wind power forecasting: State-of-the-art 2009.," tech. rep., Argonne National Lab.(ANL), Argonne, IL (United States), 2009.

[11] Y. Lin, M. Yang, C. Wan, J. Wang, and Y. Song, "A multi-model combination approach for probabilistic wind power forecasting," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 226–237, 2018.

[12] C. Ilkiliç and H. Aydin, "Wind power potential and usage in the coastal regions of turkey," *Renewable and Sustainable Energy Reviews*, vol. 44, pp. 78–86, 2015.

[13] E. Terciyanlı, T. Demirci, D. Küçük, M. Sarac, I. Çadırcı, and M. Ermiş, "Enhanced nationwide wind-electric power monitoring and forecast system," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1171–1184, 2013.

[14] "Ritm web page." `http://www.ritm.gov.tr/root/index_eng.php`. Last accessed date: 2019-12-14.

[15] S. Buhan and I. Çadırcı, "Multistage wind-electric power forecast by using a combination of advanced statistical methods," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 5, pp. 1231–1242, 2015.

[16] M. Ragheb and A. M. Ragheb, "Wind turbines theory-the betz equation and optimal rotor tip speed ratio," *Fundamental and advanced topics in wind power*, vol. 1, no. 1, pp. 19–38, 2011.

[17] L. Zjavka and S. Mišák, "Direct wind power forecasting using a polynomial decomposition of the general differential equation," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 4, pp. 1529–1539, 2018.

[18] A. Carpinone, M. Giorgio, R. Langella, and A. Testa, "Markov chain modeling for very-short-term wind power forecasting," *Electric Power Systems Research*, vol. 122, pp. 152–158, 2015.

[19] M. B. Ozkan and P. Karagoz, "A novel wind power forecast model: Statistical hybrid wind power forecast technique (shwip)," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 2, pp. 375–387, 2015.

[20] S. Alessandrini, L. Delle Monache, S. Sperati, and J. Nissen, "A novel application of an analog ensemble for short-term wind power forecasting," *Renewable Energy*, vol. 76, pp. 768–781, 2015.

[21] A. Dolatabadi, M. Jadidbonab, and B. Mohammadi-ivatloo, "Short-term scheduling strategy for wind-based energy hub: a hybrid stochastic/igdt approach," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 438–448, 2018.

[22] F. Davò, S. Alessandrini, S. Sperati, L. Delle Monache, D. Airoldi, and M. T. Vespucci, "Post-processing techniques and principal component analysis for regional wind power and solar irradiance forecasting," *Solar Energy*, vol. 134, pp. 327–338, 2016.

[23] J. Wang, A. Botterud, R. Bessa, H. Keko, L. Carvalho, D. Issicaba, J. Sumaili, and V. Miranda, "Wind power forecasting uncertainty and unit commitment," *Applied Energy*, vol. 88, no. 11, pp. 4014–4023, 2011.

[24] G. Sideratos and N. D. Hatziargyriou, "Probabilistic wind power forecasting using radial basis function neural networks," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1788–1796, 2012.

[25] C. Gallego-Castillo, A. Cuerva-Tejero, and O. Lopez-Garcia, "A review on the recent history of wind power ramp forecasting," *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 1148–1157, 2015.

[26] N. Siebert, *Development of methods for regional wind power forecasting*. PhD thesis, 2008.

[27] S. Eren, D. Küçük, C. Ünlüer, M. Demircioğlu, Y. Yanık, Y. Arslan, B. Özsoy, A. H. Güverçinci, İ. Elma, Ö. Tanıdır, *et al.*, "A ubiquitous web-based dispatcher information system for effective monitoring and analysis of the electricity transmission grid," *International Journal of Electrical Power & Energy Systems*, vol. 86, pp. 93–103, 2017.

[28] S. Buhan, Y. Özkazanç, and I. Çadırcı, "Wind pattern recognition and reference wind mast data correlations with nwp for improved wind-electric power fore-

casts," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 991–1004, 2016.

[29] M. Tan and Z. Zhang, "Wind turbine modeling with data-driven methods and radially uniform designs," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 1261–1269, 2016.

[30] M. B. Özkan, "'data mining-based power generation forecast at wind power plants," *Diss. Middle Eeast Technical University*, 2014.

[31] N. Siebert and G. Kariniotakis, "Reference wind farm selection for regional wind power prediction models," 2006.

[32] L. von Bremen, N. Saleck, and D. Heinemann, "Enhanced regional forecasting considering single wind farm distribution for upscaling," in *Journal of Physics: Conference Series*, vol. 75, p. 012040, IOP Publishing, 2007.

[33] D. R. Drew, D. J. Cannon, J. F. Barlow, P. J. Coker, and T. H. Frame, "The importance of forecasting regional wind power ramping: A case study for the uk," *Renewable energy*, vol. 114, pp. 1201–1208, 2017.

[34] M. G. Lobo and I. Sanchez, "Regional wind power forecasting based on smoothing techniques, with application to the spanish peninsular system," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1990–1997, 2012.

[35] U. Focken, M. Lange, and H.-P. Waldl, "Previento-a wind power prediction system with an innovative upscaling algorithm," in *Proceedings of the European Wind Energy Conference, Copenhagen, Denmark*, vol. 276, Citeseer, 2001.

[36] U. Focken, M. Lange, D. Heinemann, and H. P. Waldl, "Regional wind power prediction with risk control," in *Proceedings of the Global Windpower Conference*, pp. 2–5, 2002.

[37] B. Ernst, K. Rohrig, H. Regber, and P. Schorn, "Managing 3000 mw wind power in a transmission system operation center. paper," 2002.

[38] J. Yan, H. Zhang, Y. Liu, S. Han, L. Li, and Z. Lu, "Forecasting the high

penetration of wind power on multiple scales using multi-to-multi mapping," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3276–3284, 2018.

[39] I. Martí, D. Cabezón, J. Villanueva, M. J. Sanisidro, Y. Loureiro, E. Cantero, and J. Sanz, "Localpred and regiopred. advanced tools for wind energy prediction in complex terrain," in *Proceedings of the European Wind Energy Conference EWEC'03*, pp. 16–19, 2003.

[40] Á. Barbero, J. López, and J. Dorronsoro, "Kernel methods for wide area wind power forecasting," in *Proceedings of the European Wind Energy Conference and Exhibition (EWEC 2008), Brussels, Belgium (April 2008)*, 2008.

[41] A. Yamaguchi, T. Ishihara, K. Sakai, T. Ogawa, and Y. Fujino, "A physical-statistical approach for the regional wind power forecasting," in *Proceedings of the European Wind Energy Conference EWEC, Milan (IT)*, vol. 31, 2007.

[42] M. Brower, *Wind resource assessment: a practical guide to developing a wind project*. John Wiley & Sons, 2012.

[43] M. H. Zhang, *Wind resource assessment and micro-siting: science and engineering*. John Wiley & Sons, 2015.

[44] "Wasp web page." `WAsPwebpage`. Last accessed date: 2019-12-14.

[45] "windpro web page." `https://www.emd.dk/windpro/`. Last accessed date: 2019-12-14.

[46] "Windsim web page." `http://windsim.com/`. Last accessed date: 2019-12-14.

[47] M. B. Ozkan and P. Karagoz, "Data mining based upscaling approach for regional wind power forecasting: Regional statistical hybrid wind power forecast technique (regionalshwip)," *IEEE Access*, 2019.

[48] K. Murthy and O. Rahi, "A comprehensive review of wind resource assessment," *Renewable and Sustainable Energy Reviews*, vol. 72, pp. 1320–1342, 2017.

[49] F. W. Telewski, "Is windswept tree growth negative thigmotropism?," *Plant science*, vol. 184, pp. 20–28, 2012.

[50] W. Barreto-Souza, A. L. de Morais, and G. M. Cordeiro, "The weibull-geometric distribution," *Journal of Statistical Computation and Simulation*, vol. 81, no. 5, pp. 645–657, 2011.

[51] B. Karthikeya, P. S. Negi, and N. Srikanth, "Wind resource assessment for urban renewable energy application in singapore," *Renewable Energy*, vol. 87, pp. 403–414, 2016.

[52] T. Aukitino, M. G. Khan, and M. R. Ahmed, "Wind energy resource assessment for kiribati with a comparison of different methods of determining weibull parameters," *Energy Conversion and Management*, vol. 151, pp. 641–660, 2017.

[53] S.-D. Kwon, "Uncertainty analysis of wind energy potential assessment," *Applied Energy*, vol. 87, no. 3, pp. 856–865, 2010.

[54] S. Pishgar-Komleh, A. Keyhani, and P. Sefeedpari, "Wind speed and power density analysis based on weibull and rayleigh distributions (a case study: Firouzkooh county of iran)," *renewable and sustainable energy reviews*, vol. 42, pp. 313–322, 2015.

[55] J. Zhang, C. Draxl, T. Hopson, L. Delle Monache, E. Vanvyve, and B.-M. Hodge, "Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods," *Applied Energy*, vol. 156, pp. 528–541, 2015.

[56] P. Doubrawa, R. J. Barthelmie, S. C. Pryor, C. B. Hasager, M. Badger, and I. Karagali, "Satellite winds as a tool for offshore wind resource assessment: The great lakes wind atlas," *Remote Sensing of Environment*, vol. 168, pp. 349–359, 2015.

[57] D. A. Ramos, V. G. Guedes, R. R. Pereira, T. A. Valentim, and W. A. Netto, "Further considerations on wasp, openwind and windsim comparison study: Atmospheric flow modelling over complex terrain and energy production estimate.," *Brazil WindPower*, 2017.

[58] B. Jimenez, F. Durante, B. Lange, T. Kreutzer, and J. Tambke, "Offshore wind resource assessment with wasp and mm5: comparative study for the german

bight," *Wind Energy: An International Journal for Progress and Applications in Wind Power Conversion Technology*, vol. 10, no. 2, pp. 121–134, 2007.

[59] J. Liu, C. Y. Gao, J. Ren, Z. Gao, H. Liang, and L. Wang, "Wind resource potential assessment using a long term tower measurement approach: A case study of beijing in china," *Journal of cleaner production*, vol. 174, pp. 917–926, 2018.

[60] A. Celik, "A simplified model for estimating yearly wind fraction in hybrid-wind energy systems," *Renewable energy*, vol. 31, no. 1, pp. 105–118, 2006.

[61] N. Salvação and C. G. Soares, "Wind resource assessment offshore the atlantic iberian coast with the wrf model," *Energy*, vol. 145, pp. 276–287, 2018.

[62] "Repa web page." `http://www.yegm.gov.tr/YEKrepa/REPAduyuru01.html`. Last accessed date: 2019-12-14.

[63] I. Erinmez, S. Majithia, C. Rogers, T. Yasuhiro, S. Ogawa, H. Swahn, and J. Kappenman, "Application of modelling techniques to assess geomagnetically induced current risks on the ngc transmission system," *CIGRE Paper*, no. 39–304, 2002.

[64] M. de la Torre, G. Juberias, T. Dominguez, and R. Rivas, "The cecre: Supervision and control of wind and solar photovoltaic generation in spain," in *2012 IEEE Power and Energy Society General Meeting*, pp. 1–6, IEEE, 2012.

[65] J. M. Rodriguez, O. Alonso, M. Duvison, and T. Domingez, "The integration of renewable energy and the system operation: The special regime control centre (cecre) in spain," in *2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century*, pp. 1–6, IEEE, 2008.

[66] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "Smart grid technologies: Communication technologies and standards," *IEEE transactions on Industrial informatics*, vol. 7, no. 4, pp. 529–539, 2011.

[67] K. Cheung, X. Wang, B.-C. Chiu, Y. Xiao, and R. Rios-Zalapa, "Generation dispatch in a smart grid environment," in *2010 Innovative Smart Grid Technologies (ISGT)*, pp. 1–6, IEEE, 2010.

[68] W. P. Mahoney, K. Parks, G. Wiener, Y. Liu, W. L. Myers, J. Sun, L. Delle Monache, T. Hopson, D. Johnson, and S. E. Haupt, "A wind power forecasting system to optimize grid integration," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 4, pp. 670–682, 2012.

[69] A. Afram, F. Janabi-Sharifi, A. S. Fung, and K. Raahemifar, "Artificial neural network (ann) based model predictive control (mpc) and optimization of hvac systems: A state of the art review and case study of a residential hvac system," *Energy and Buildings*, vol. 141, pp. 96–113, 2017.

[70] J. Wang, X. Liao, P. Zheng, S. Xue, and R. Peng, "Classification of chinese herbal medicine by laser-induced breakdown spectroscopy with principal component analysis and artificial neural network," *Analytical letters*, vol. 51, no. 4, pp. 575–586, 2018.

[71] G. Raut, S. Rai, S. K. Vishvakarma, and A. Kumar, "A cordic based configurable activation function for ann applications," in *2020 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, pp. 78–83, IEEE, 2020.

[72] Y.-Y. Hong and C. L. P. P. Rioflorido, "A hybrid deep learning-based neural network for 24-h ahead wind power forecasting," *Applied Energy*, vol. 250, pp. 530–539, 2019.

[73] Y. Ju, G. Sun, Q. Chen, M. Zhang, H. Zhu, and M. U. Rehman, "A model combining convolutional neural network and lightgbm algorithm for ultra-short-term wind power forecasting," *IEEE Access*, vol. 7, pp. 28309–28318, 2019.

[74] D. Zheng, M. Shi, Y. Wang, A. T. Eseye, and J. Zhang, "Day-ahead wind power forecasting using a two-stage hybrid modeling approach based on scada and meteorological information, and evaluating the impact of input-data dependency on forecasting accuracy," *Energies*, vol. 10, no. 12, p. 1988, 2017.

[75] S. Huang, N. Cai, P. P. Pacheco, S. Narrandes, Y. Wang, and W. Xu, "Appli-

cations of support vector machine (svm) learning in cancer genomics," *Cancer Genomics-Proteomics*, vol. 15, no. 1, pp. 41–51, 2018.

[76] A. M. Abd and S. M. Abd, "Modelling the strength of lightweight foamed concrete using support vector machine (svm)," *Case studies in construction materials*, vol. 6, pp. 8–15, 2017.

[77] J. Heinermann and O. Kramer, "Precise wind power prediction with svm ensemble regression," in *International Conference on Artificial Neural Networks*, pp. 797–804, Springer, 2014.

[78] Y. Liu, Y. Sun, D. Infield, Y. Zhao, S. Han, and J. Yan, "A hybrid forecasting method for wind power ramp based on orthogonal test and support vector machine (ot-svm)," *IEEE Transactions on Sustainable energy*, vol. 8, no. 2, pp. 451–457, 2016.

[79] M. Barman, N. D. Choudhury, and S. Sutradhar, "A regional hybrid goa-svm model based on similar day approach for short-term load forecasting in assam, india," *Energy*, vol. 145, pp. 710–720, 2018.

[80] L.-L. Li, X. Zhao, M.-L. Tseng, and R. R. Tan, "Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm," *Journal of Cleaner Production*, vol. 242, p. 118447, 2020.

[81] J. Benesty, J. Chen, and Y. Huang, "On the importance of the pearson correlation coefficient in noise reduction," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 4, pp. 757–765, 2008.

[82] P. Robert and Y. Escoufier, "A unifying tool for linear multivariate statistical methods: the rv-coefficient," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 25, no. 3, pp. 257–265, 1976.

[83] "Mgm web page." `https://mgm.gov.tr/eng/forecast-cities.aspx`. Last accessed date: 2019-12-14.

[84] P. Termonia, C. Fischer, E. Bazile, F. Bouyssel, R. Brožková, P. Bénard, B. Bochenek, D. Degrauwe, M. Derková, R. El Khatib, *et al.*, "The aladin system and its canonical model configurations arome cy41t1 and alaro cy40t1," *Geoscientific Model Development*, vol. 11, pp. 257–281, 2018.

[85] "Gfs web page." https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs. Last accessed date: 2019-12-14.

[86] "Ncar web page." https://ncar.ucar.edu/. Last accessed date: 2019-12-14.

[87] "Ecmwf web page." https://www.ecmwf.int/. Last accessed date: 2019-12-14.

[88] X.-M. Hu, J. W. Nielsen-Gammon, and F. Zhang, "Evaluation of three planetary boundary layer schemes in the wrf model," *Journal of Applied Meteorology and Climatology*, vol. 49, no. 9, pp. 1831–1844, 2010.

[89] R. K. Humphreys, M.-T. Puth, M. Neuhäuser, and G. D. Ruxton, "Underestimation of pearson's product moment correlation statistic," *Oecologia*, vol. 189, no. 1, pp. 1–7, 2019.

[90] F. D'Ascenzi, C. Pisicchio, S. Caselli, F. M. Di Paolo, A. Spataro, and A. Pelliccia, "Rv remodeling in olympic athletes," *JACC: Cardiovascular Imaging*, vol. 10, no. 4, pp. 385–393, 2017.

[91] F. Rodríguez, A. M. Florez-Tapia, L. Fontán, and A. Galarza, "Very short-term wind power density forecasting through artificial neural networks for microgrid control," *Renewable Energy*, vol. 145, pp. 1517–1527, 2020.

[92] B. Dietrich, J. Walther, M. Weigold, and E. Abele, "Machine learning based very short term load forecasting of machine tools," *Applied Energy*, vol. 276, p. 115440, 2020.

[93] T. Ahmad and H. Chen, "Utility companies strategy for short-term energy demand forecasting using machine learning based models," *Sustainable cities and society*, vol. 39, pp. 401–417, 2018.

[94] P.-H. Kuo and C.-J. Huang, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies*, vol. 11, no. 1, p. 213, 2018.

[95] J. Á. G. Ordiano, L. Gröll, R. Mikut, and V. Hagenmeyer, "Probabilistic energy

forecasting using the nearest neighbors quantile filter and quantile regression," *International Journal of Forecasting*, vol. 36, no. 2, pp. 310–323, 2020.

[96] X.-C. Yuan, X. Sun, W. Zhao, Z. Mi, B. Wang, and Y.-M. Wei, "Forecasting china's regional energy demand by 2030: A bayesian approach," *Resources, Conservation and Recycling*, vol. 127, pp. 85–95, 2017.

[97] X. Zhang, Y. Li, S. Lu, H. F. Hamann, B.-M. Hodge, and B. Lehman, "A solar time based analog ensemble method for regional solar power forecasting," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 268–279, 2018.

[98] E. Terciyanli, T. Demirci, D. Kücük, S. Buhan, M. B. Özkan, C. E. Köksoy, E. Koç, C. Kahraman, A. B. Haliloğlu, T. Demir, *et al.*, "The architecture of a large-scale wind power monitoring and forecast system," in *4th International Conference on Power Engineering, Energy and Electrical Drives*, pp. 1162–1167, IEEE, 2013.

[99] M. Özkan, E. Terciyanlı, D. Küçük, S. Buhan, T. Demirci, C. Yıldız, and M. Günindi, "Verification of a real-time wind power monitoring and forecast system for turkey," 2013.

[100] T. Demirci, A. Kalaycıoglu, D. Küçük, Ö. Salor, M. Güder, S. Pakhuylu, T. Atalık, T. Inan, I. Çadırcı, Y. Akkaya, *et al.*, "Nationwide real-time monitoring system for electrical quantities and power quality of the electricity transmission system," *IET Generation, Transmission & Distribution*, vol. 5, no. 5, pp. 540–550, 2011.

[101] "Aladin web page." `http://www.umr-cnrm.fr/aladin/`. Last accessed date: 2019-12-14.

[102] S. Buhan, E. Terciyanli, M. B. Özkan, D. Kücük, A. B. Haliloğlu, T. Demirci, H. Tuna, and H. Ünsal, "Verification of a very short term wind power forecasting algorithm for turkish transmission grid," in *4th International Conference on Power Engineering, Energy and Electrical Drives*, pp. 1217–1221, IEEE, 2013.

[103] M. B. Ozkan, U. Guvengir, D. Küçük, A. U. Secen, S. Buhan, T. Demirci, A. Bestil, C. Er, and P. Karagoz, "Probabilistic wind power forecasting by us-

ing quantile regression analysis," in *International Workshop on Data Analytics for Renewable Energy Integration*, pp. 72–82, Springer, 2017.

[104] M. Delacre, D. Lakens, and C. Leys, "Why psychologists should by default use welch's t-test instead of student's t-test," *International Review of Social Psychology*, vol. 30, no. 1, 2017.

[105] "Matlab web page." `https://www.mathworks.com/products/matlab.html`. Last accessed date: 2019-12-14.

[106] "Libsvm web page." `https://www.csie.ntu.edu.tw/cjlin/libsvm/`. Last accessed date: 2019-12-14.

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Özkan, Mehmet Barış
Nationality: Turkish (TC)

EDUCATION

Degree	Institution	Year of Graduation
M.S.	METU Computer Engineering	2014
B.S.	METU Computer Engineering	2011
High School	Erzurum Anatolian High School	2006

PROFESSIONAL EXPERIENCE

Year	Place	Enrollment
2019- Present	TUBITAK Energy Institute	Chief Researcher
2015- 2019	TUBITAK Energy Institute	Senior Researcher
2011- 2015	TUBITAK Energy Institute	Researcher

PUBLICATIONS

Book Chapters

Ozkan, M. B., Küçük, D., Buhan, S., Demirci, T. and Karagoz, P. (2020). Large-Scale Renewable Energy Monitoring and Forecast Based on Intelligent Data Analysis. In Handbook of Research on Smart Computing for Renewable Energy and Agro-Engineering (pp. 53-77). IGI Global.

International Journal Publications

Ozkan M.B. and Karagoz P. (2021), Reducing the Cost of Wind Resource Assessment: Using a Regional Wind Power Forecasting Method for Assessment, International Journal of Energy Research, John Wiley and Sons (Major Revision)

Ozkan, M. B. and Karagoz, P. (2019). Data mining-based upscaling approach for regional wind power forecasting: Regional statistical hybrid wind power forecast technique (RegionalSHWIP). IEEE Access, 7, 171790-171800.

Ozkan, M. B. and Karagoz, P. (2015). A novel wind power forecast model: Statistical hybrid wind power forecast technique (SHWIP). IEEE Transactions on Industrial Informatics, 11(2), 375-387.

Guder, M., Salor, O., Çadirci, I., **Ozkan, M. B.** and Altintas, E. (2015). Data mining framework for power quality event characterization of iron and steel plants. IEEE Transactions on Industry Applications, 51(4), 3521-3531.

International Conference Publications

Ozkan, M. B. et al. (2017). Probabilistic Wind Power Forecasting by Using Quantile Regression Analysis. In International Workshop on Data Analytics for Renewable Energy Integration (pp. 72-82). Springer, Cham.

Köksoy, C. E., **Ozkan, M. B.**, Buhan, S., Demirci, T., Arslan, Y., Birtürk, A. and

Karagöz, P. (2015, December). Improved wind power forecasting using combination methods. In 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA) (pp. 1142-1147). IEEE

Ozkan, M. B., Küçük, D., Terciyanlı, E., Buhan, S., Demirci, T. and Karagoz, P. (2013). A data mining-based wind power forecasting method: results for wind power plants in Turkey. In International Conference on Data Warehousing and Knowledge Discovery (pp. 268-276). Springer, Berlin, Heidelberg.

Ozkan, M. B., Terciyanli, E., Kucuk, D., Buhan, S., Demirci, T., Yildiz, C., Gunindi, M. (2013). Verification of a Real Time Wind Power Monitoring and Forecast System for Turkey. In Proceedings of the IET Renewable Power Generation Conference. Beijing, China

Buhan, S., Terciyanli, E., **Ozkan, M. B.**, Kucuk, D., Haliloglu, A. B., Demirci, T., Tuna, H., Unsal, H. (2013). Verification of a Very Short Term Wind Power Forecasting Algorithm for Turkish Transmission Grid. In Proceedings of the 4th IEEE International Conference on Power Engineering, Energy and Electrical Drives. Istanbul, Turkey.

Terciyanli, E., Demirci, T., Kucuk, D., Buhan, S., **Ozkan, M. B.**, Er Koksoy, C., Koc, E., Kahraman, C., Haliloglu, A. B., Demir, T., Gunindi, M., Gukmen, M., Karayilanoglu, Z. (2013). The Architecture of a Large-Scale Wind Power Monitoring and Forecast System. In Proceedings of the 4th IEEE International Conference on Power Engineering, Energy and Electrical Drives. Istanbul, Turkey.

Awards

2017 TUBITAK Marmara Research Center Best Research Team Award

IEEE Transactions on Industrial Informatics 2016 Best Paper Award