T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Development of an integrated system for wind power forecasting under complex geographic conditions
著者(和文)	CheYuzhang
Author(English)	Yuzhang Che
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第10616号, 授与年月日:2017年9月20日, 学位の種別:課程博士, 審査員:肖 鋒,奥野 喜裕,岡村 哲至,末包 哲也,長﨑 孝夫
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第10616号, Conferred date:2017/9/20, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Powered by T2R2 (Science Tokyo Research Repository)

Development of an integrated system for wind power forecasting under complex geographic conditions

Yuzhang CHE

Supervisor: **Feng XIAO**

Department of Energy Science, Tokyo Institute of Technology

This dissertation is submitted for the degree of

Doctor of Science

August 2017

Abstract

Due to the clean, renewable and sustainable characteristics, the generation of wind energy around the world grows rapidly. This rapid development of wind energy and the associated high uncertainties and fluctuations in power generation present a big challenge for both wind power generators and electric grids in many countries, especially Japan where the topography condition is very complex. An effective way to solve these problems is providing accurate and reliable wind power predictions in advance to properly adjust the integration of wind power into existing electrical systems. However, to our best knowledge there is very limited report in literature on any practice to establish a comprehensive forecasting system for wind farm sites in Japan.

Therefore, the purpose of this thesis is to develop a wind power forecasting system for a wind farm of interest in Awaji island of Japan as an effort to facilitate the short-term wind power forecast in Japan area. Considering the specific situation in Japan, we have combined meso-scale WRF (Weather Research and Forecasting) model, power curve (approximated by a 10th-order polynomial), Kalman filter, data assimilation and the micro-scale OpenFOAM (Open source Field Operation and Manipulation) model together to build a novel and integrated forecasting system for wind energy prediction under complex terrain conditions.

We firstly evaluated the forecasting ability of the WRF and power curve model separately as the basic components of the integrated forecasting system. It is noted that the WRF model has been tuned to adapt the wind farm in Japan. Compared with the observed data of both wind speed and power, it is found that the two components are able to provide reasonably reliable forecasting results in the target site which has complex geographic environment very

typical in Japan. However, significant errors and uncertainties were also observed in this preliminary system, for example the systematic overestimation of the wind power.

In order to improve the accuracy and reliability of wind power prediction, we have integrated Kalman filter, data assimilation and a micro-scale CFD (computational fluid dynamics) model (OpenFOAM model) as new modules in the system to reduce the errors and uncertainties. The performance of those three modules has been validated with the observed data. With Kalman filter, the raw wind prediction can be substantially improved. The 15-turbine averaged improvements of ME (mean error), RMSE (root mean square error) and CC (correlation coefficient) are 97%, 22% and 10% respectively. Meanwhile, the Kalman filter also demonstrates a promising capability of reducing the uncertainties in the power curve model. More specifically, Kalman filter could significantly improve the raw model prediction of power by 92%, 33% and 15% in ME, RMSE and CC respectively. The validation results of data assimilation also indicate that the WRF model forecasts can be markedly improved after assimilating nacelle wind data, with the relative improvements of 34.3%, 23.9% and 8.8% in ME, RMSE and IA (index of agreement) respectively. It is noted that the data assimilation module can handle part of random errors which cannot be eliminated by Kalman filter module, and integrating both Kalman filter and data assimilation with WRF model can obtain the best performance. The resolution (500-m of horizontal direction) of the current forecasting system is too coarse to capture the detailed flow information caused by the complex terrain in the atmospheric boundary layer. Thus, the micro-scale OpenFOAM model has been coupled with the WRF system to build a multi-scale forecasting system for short-term prediction of hub-height wind under forcing of local geographic conditions. The ability of this multi-scale system for simulating wind flow the complex terrain is firstly validated with an arbitrary case. It is found that this system can capture reasonable distribution of the velocity and turbulent kinetic energy at the atmospheric boundary layer compared with other researchers' work. Moreover, this multi-scale forecasting system shows remarkable advantages against the single meso-scale WRF component through validations with a 8-day series of observed data (192 cases).

In summary, a novel integrated forecasting system has been developed by combining the meso-scale WRF model, power curve model, Kalman filter, data assimilation and the micro-scale OpenFOAM model in this study. Its performance has been validated with the real-case observed data from the Awaji wind farm in Japan. Part of this system has been installed and used as routine tool for operational prediction.

Acknowledgements

I would like to thank all the people who contributed in some way to this thesis. Without their guidance and support, the thesis would not have been possible. First and foremost, I would like to express my deepest appreciate to my academic advisor, Prof. Xiao, for accepting me into his group, giving me intellectual freedom in my work, supporting my attendance at various conferences, engaging me in new ideas and providing invaluable advices ranging from daily life to future development. Those experiences will be great treasures in my whole life.

Then I would like to thank Prof. Peng, who provided me with the computational resources I need to conduct numerical experiments, important instructions on my paper revisions and valuable suggestions to my future career. In addition, many thanks to the *Clean Energy Factory* (CEF) company for providing crucial dataset as well as fundings for my research. I have obtained a lot of experience of building an operational system through working and discussing with CEF members, especially Mr. Kamata, Mr. Ganba, Mr. Shimata and Mr. Zheng. Thanks again for your enthusiasm and patience.

I would like to acknowledge the various members of Xiao laboratory with whom I had the opportunity to work and live: Xie, Abe, Ikebata, Sun, Deng, Jin Peng, Qian, Jin Jonghoon, Inaba, Nakayama, Wang, Tsubone, Kondo, Saito, Hashiguchi, Andres, Siengdy, Shimizu, Endo, Miyakawa and Koseki. They all provided a friendly and cooperative atmosphere at work and also useful feedback and insightful comments on my work.

I would like to thank my family and friends. Foremost, I would like to thank my best friend and beautiful wife, Miao, who is paramount to my happiness throughout writing this thesis and my life in general. In addition, I also would like to thank my Mom, Dad and sister to support me with constant love. Last but not the least, many thanks to my friends overseas, Sun, Chang, Gu, Liu, Chen, Zhou, Jiang, Che, Gao, Guo and Niu for supporting me on many aspects. Thank you all.

I would like to dedicate this thesis to my family.

Contents

List of Figures

List of Tables

Acronym

ABL	Atmosphere Boundary Layer
ANN	Artificial Neural Network
ARIMA	Auto-regressive Integrated Moving Average
ARMA	Auto-regressive Moving Average
BL	Boundary Layer
CFD	Computational Fluid Dynamics
COD	Coefficient Of Determination
DES	Detached-eddy Simulation
DME	Daily Mean Error
FA	Free Atmosphere
FNN	Feed-forward Neural Network
GM	Grey Models
GWEC	Global Wind Energy Council
LES	Large-eddy Simulation
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ME	Mean Error
ML	Mixed Layer
MLP	Multi-layer Perceptron
MM5	Fifth-Generation Penn State/NCAR Mesoscale Model
MSE	Mean Square Error
NCAR	
NCEP	National Center for Environmental Prediction
NN	Neural Network
NWP	Numerical Weather Prediction
PDF	Probability Distribution Function

RAMS	Regional Atmospheric Modeling System
RANS	Reynolds-averaged Navier-Stokes
RBF	Radial Basis Function
RL	Residual Layer
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
SBL	Stable Boundary Layer
SVR	Support Vector Regression
WEP	Weather Ensemble Prediction
WPFS	Wind Power Forecast System
WPPT	Wind Power Prediction Tool
WRF	Weather Research and Forecasting

Chapter 1

Introduction

1.1 Global status of wind power

Wind power is an important form of energy, which is collected from renewable resource of air motion driven by heat energy from the sun. As similar to other forms of renewable energy, e.g., tides and geothermal heat, it has apparent advantages against the traditional generators, including the characteristic of sustainable, environmental friendly and the potential economic benefits. Those are also why the wind power is now established around the world as a mainstream source of energy and even is projected to provide at least 36% global electricity demand in 2050, according to the latest report from the Global Wind Energy Council (GWEC).

The development of wind power is dramatic in the last 16 years. Based on the statistical data from GWEC (Figure 1.1), except for 2013, the global annual installed wind capacity is increasing consistently. From Figure 1.2, it is obvious that the global cumulative installed wind capacity was about 432,900 MW at the end of 2015, growing from 17,400 MW in the year 2000. This growth at some extent directly meets the demand of cutting the emissions of carbon dioxide with no-extra decrease of established economies. In addition, with the development of wind power, a number of new jobs and industries are created and the energy security also has been enhanced at the same time.

A fact which can not be neglected is the current situation of wind energy in different countries or regions has large distinctions. This can be proved with the evidences displayed in the Figure 1.3. It is clear that several countries have already achieved relatively high levels. China has the largest installed capacity (145,104 MW, 33.6% of the global installed wind power), followed by the USA (17.2%) and Germany (10.4%). However, the development of wind power for the rest of counties is not very well due to many factors, for example, lacking of expertise on how to select perfect site in some developing countries or the government policy on electricity generation. In this thesis, we mainly pay attention to the specific situation in Japan.

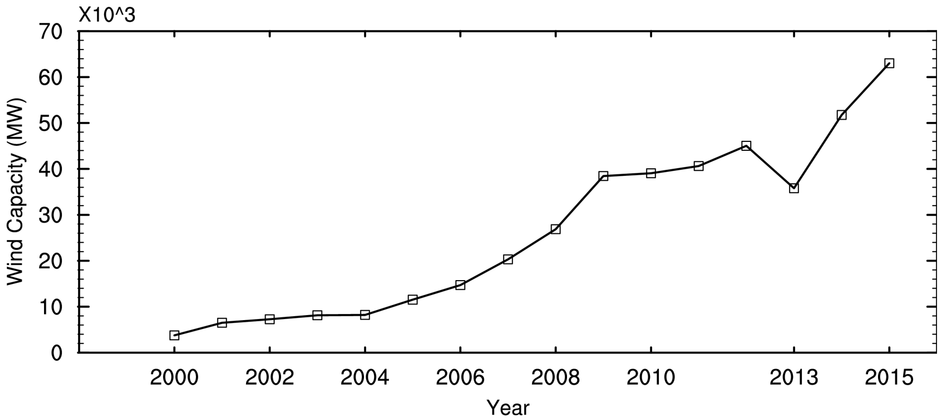

Figure 1.1 The global annual installed wind capacity from 2000 to 2015. Data source: GWEC.

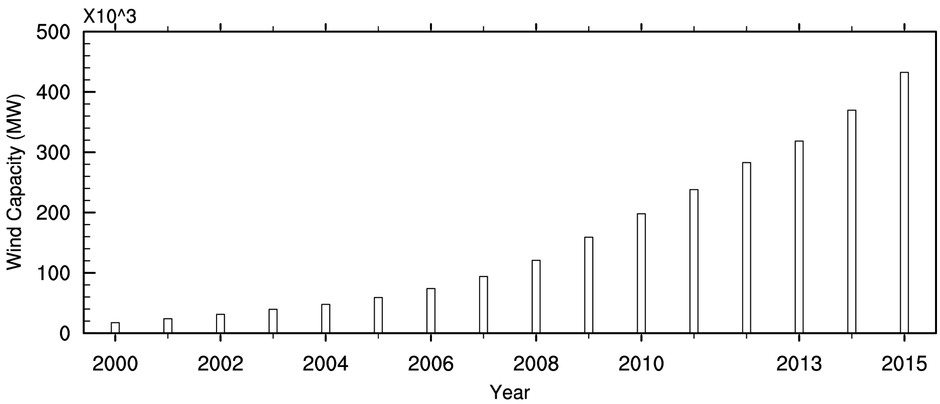

Figure 1.2 The global cumulative installed wind capacity by the end of 2015, source: GWEC.

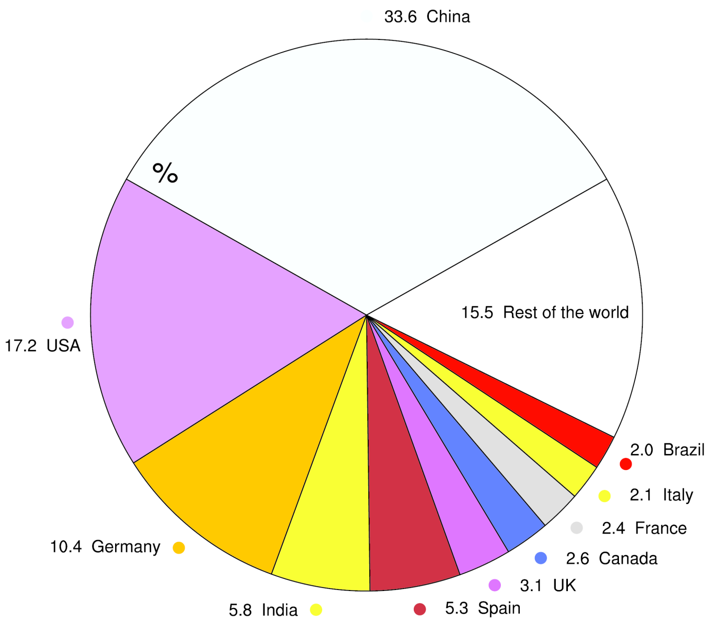

Figure 1.3 The cumulative capacity share of top 10 countries in 2015, source: GWEC.

Table 1.1 Current status of renewable energy in Japan. Estimated renewable power generation, by type and by proportion of total power generated in Japan by the end of 2012 and 2015, respectively. Source: JFS [54]

Type of Renewable Energy	2012		2015	
	Annual Power Generation Capacity (GWh)	Proportion of Total (%)	Annual Power Generation Capacity (GWh)	Proportion of Total (%)
Solar photovoltaic (PV)	7,635	0.69%	34,085	3.31%
Wind	**4,838**	**0.44%**	**5,381**	**0.52%**
Geothermal	2,609	0.24%	3,115	0.30%
Small-scale hydropower	17,401	1.58%	17,777	1.73%
Biomass	12,186	1.11%	16,395	1.59%
Total	44,670	4.1%	76,205	7.41%

1.2 Current situation in Japan

In contrast to many other countries, the situation of wind power in Japan looks a bit different. Japan has not made much progress yet since there are only 245 MW of wind capacity are installed in 2015 and the cumulative capacity just reaches to 3,038 MW [26], which satisfied a mere 0.52% of domestic electricity demand (Table 1.1).

In Table 1.1, we also can find the status of estimated annual power generation capacity by other renewable energy sources and their percentage shares of total power generation in Japan. Compared to the numbers in 2012, all of them have been developed at the end of

2015, indicating positive trends for the expansion of renewable energy. However, except for the largest increase of solar photovoltaic of 2.62%, the developments of other forms of renewable energy are still not evident, especially for the wind power. This may suggest that the transformation of energy system to allow for a more diverse energy mix in Japan is very slow and far behind other developed countries.

To return to our interest of wind energy, the slow development in Japan may owe to following factors:

- *Energy Policy*. Before 2011, energy supply of Japan was dominated by fossil fuels and nuclear energy. The consequence of overusing traditional fossil fuels leads to plenty of emissions of carbon dioxide and other poisonous gases which will generate severe environmental problems, such as global warming and air pollution. Hence, Japan had pledged to increase nuclear power's share in electricity supply from 30% to 50% in 2009, in order to reduce its future greenhouse gas emissions, though this wish had been broken by the 2011 Great East Japan earthquake and the subsequent nuclear accident. Security of the nuclear plant is the key concern. Thus, the government decided to fundamentally rethink its energy policy to balance the security, economic efficiency and environmental protection. And this provides a favorable opportunity for developing wind energy and other renewables. Unfortunately, such development may have been constrained as nuclear power plants are being restarted in Japan.

- *High Population Density*. Japan has high population density of around 336 people per km^2. This directly leads to an unsuitability for constructing a wind farm which generally needs extensive open areas without too many artificial buildings.

- *Social Acceptance Issues*. Social acceptance of wind farm projects is a serious bottleneck, due to construction of wind farm has a couple of disadvantages. In particular, the issues related to large noise and low frequency vibration which are harmful for the local residents' normal lives, are considered most important for local community. Moreover, the manufacturing of wind turbines somehow will cause some pollution, for example destruction of landscape and damage on ecosystem.

- *Complex Terrain Conditions.* Japan is an island nation and nearly 72% of its land is mountainous. Thus, the development of its wind energy potential is very limited compared to US or Europe.

- *Grid Connection Issues.* The distribution of wind resources across Japan is quite uneven. Many of the most suitable wind farm sites are located in rural or mountainous terrain, e.g., Hokkaido and Tohoku, where power demand is low. Besides, the local grid condition is generally inadequate in those areas and the installation costs are high.

Those mentioned concerns could be solved or partly improved by making more appropriate energy policy (especially on energy mix) or enhancing the grid connection and integration with the existing smart grids. However, integrating the available wind power with smart grid is a really tough task if the intrinsically variable and uncontrollable characteristics of wind cannot be understood well.

1.3 The nature of the wind

Wind is the movement of air caused by differences in the atmospheric pressure. Specifically, when a difference in atmospheric pressure exists, air will move from the higher region to the zone of low pressure where air is rising. The strength of wind (i.e, wind speed) is mainly affected by two factors. One is the magnitude of the pressure difference between the high and low region. The greater such differences exit, the stronger winds may happen. The other one is the funneling effects which might be caused by the regional topography, for example canyons, passes, and downtown streets.

Types of the winds

Roughly, the winds may be classified into four types, namely the planetary winds, the monsoon winds, cyclones and anticyclones and local winds.

- *The Planetary Winds.* It is known as the general distribution of winds throughout the lower atmosphere, including the North-east and South-east Trade winds, the Temperate

Westerlies and the Polar Easterlies. These winds blow very regularly throughout the year and are controlled on the whole by the latitudinal pressure belts.

- *The Monsoon Wind* is traditionally defined as a wind system in which there is a complete or almost complete reversal of prevailing direction from season to season. The largest developed monsoonal area in the world is the southeast Asia.

- *The Cyclone and Anticyclone Winds.* A cyclone or anticyclone is a region of low or high atmospheric pressure and associated storm system. Both of them move as compete systems while cyclones moving much faster than the anticyclones.

- *Local Winds,* which are usually caused by local factors are basically confined to a limited area compared to planetary winds. There are some well-known examples of local winds including Fohn, Chinook, Tornado, and Sea and Land Breeze.

Except for some of local winds, the pressure systems of other winds are relatively stationary due to the pressure gradients are balanced by the centripetal and the Coriolis accelerations. However, this kind of balance of forces do not conserved because of the occurring drag forces brought in by the earth's surface roughness, which made the winds within the atmosphere boundary layer (ABL) extremely complicated. Whereas, the atmospheric flows in ABL need to be understood as much as we can, as there is a great increase in the demand for production of wind energy.

Atmosphere boundary layer

The troposphere can be roughly classified to two parts: the free atmosphere (FA) and a boundary layer (BL) near the surface (Figure 1.4), which is known as the atmospheric boundary layer (ABL). ABL is defined originally as the lowest part of the atmosphere which has a direct interaction with the earth's surface and responds to surface forcing with a time scale of about 1 hour or less [105]. It plays an important role in many fields, including aeronautical meteorology, air pollution, agricultural meteorology, hydrology, weather forecasting and climate, and wind energy. The depth of the ABL is variable, typically ranges from 100 m

Table 1.2 Comparison of boundary layer and the free atmosphere characteristics. (from Stull (1988) [105])

Property	Free Atmosphere	Atmosphere Boundary Layer
Turbulence	Mostly Laminar.	Almost continuously turbulent over its whole depth.
Friction	Small viscous dissipation.	Strong drag against the earth's surface. Large energy dissipation.
Dispersion	Small molecular diffusion.	Rapid turbulent mixing in the vertical and horizontal.
Winds	Winds nearly geostrophic.	Near logarithmic wind speed profile in the surface layer. Subgeostrophic, cross-isobaric flow common.
Vertical Transport	Mean wind dominates.	Turbulence dominates.
Thickness	Less variable. 8-18 km. Slow time variations.	Varies between 100 m to 3 km in time and space, Diurnal oscillations over land.

to 2000 m (i.e., occupying the bottom 10% to 20% of the troposphere). There are many differences when we compare ABL with the free atmosphere, part of which have been summarized in Table 1.2.

The ABL structure that evolves with the diurnal cycle is displayed in Figure 1.5. There are three main components of this structure, namely the mixed layer (ML), residual layer (RL) and the stable boundary layer (SBL). Near the surface, there is a thin surface layer at all time, in which the vertical turbulent fluxes are nearly constant.

Flow over the complex terrain

The on-shore wind farms mostly locate in the mountainous areas in Japan, where the terrain conditions are very complex. As known, the terrain effect (e.g., strong turbulence) plays a vital role in modifying wind speed and direction at the site of interest, and indirectly makes the evaluation of wind resources or wind-farm profitability very difficult. For the sake of figuring

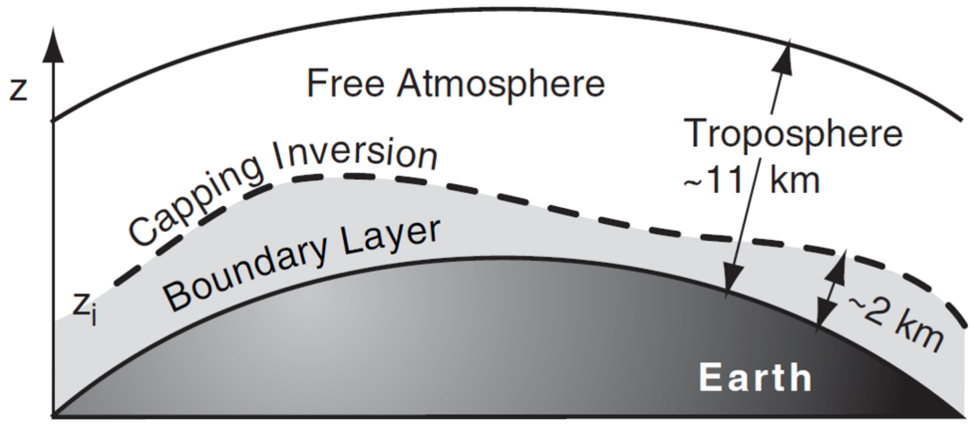

Figure 1.4 Location of the boundary layer, with top at z_i

out how complex terrain affects the wind, a field measurement campaign is conducted, though its impact is limited owing to the sparse measure locations. As a complement, additional methods including experimental and numerical models have been developed. The wind

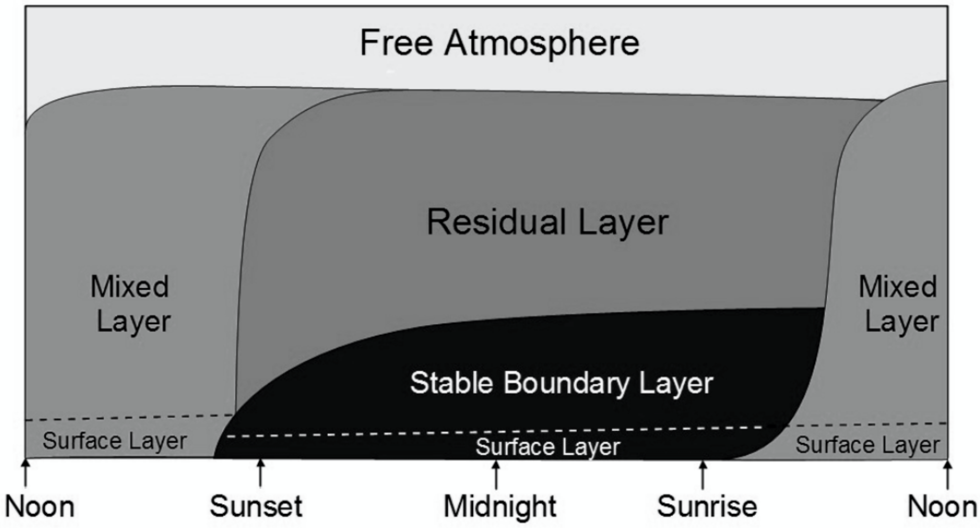

Figure 1.5 The structure of the atmospheric boundary layer (adaptive modified from Stull (1988)[105]).

tunnel test is a common way to study the state of flow over a specific terrain feature. The key step is to reproduce the complex terrain using various model scales (usually smaller than 1/1000). The first wind-tunnel investigations of flow over real topography are implemented by Meroney [75] and Neal and Stevenson [82]. They studied the wind flow over mountains in New Zealand using scales of 1/5000 and 1/4000 respectively. Bowen [9] further suggested taking 1/6000 as a limit in order to properly simulate turbulence. In recent years, there are still several studies [19, 95, 21] related to the wind tunnel experiments with complex terrain features. Nevertheless, to do the wind tunnel experiment needs a significant investment such as experimental instruments and maintenance costs. Meanwhile, it is very difficult to bring the turbulence process in the wind tunnel tests while it always exits in real ABL. Considering those problems, using the method of numerical modeling could be an effective way to obtain a detailed and accurate information of the airflow over the complex region.

In these days, using the computational fluid dynamics (CFD) methods to simulate the airflow over complex terrain is a prevail trend. There are many models have been developed, ranging from the very simple linear models to complex non-linear models, for example, the Reynolds-averaged Navier-Stokes (RANS), Large-eddy Simulation (LES) and the Detached-eddy Simulation (DES). The RANS model is a couple of time-averaged equations of motion for fluid flow. The basic idea behind those equations is Reynolds decomposition (Eq. (1.1)) [90], whereby an instantaneous quantity can be decomposed into its time-averaged and the corresponding fluctuating quantities.

$$u(\mathbf{x},t) = \overline{u}(\mathbf{x}) + u^{'}(\mathbf{x},t) \tag{1.1}$$

where u is an arbitrary and instantaneous variable, $\mathbf{x} = \mathbf{x}(x,y,z)$ and $\overline{u^{'}(\mathbf{x},t)} = 0$. The RANS approach has been studied in the literature for flow over complex terrain, including the Serra das Meadas mountainous [72], the Askervein hill [62, 13, 84] and the Blashaval hill [41]. Most of the results in those studies indicate that numerical models based on the RANS approach could perform reasonably well on simulating mean flow conditions. However, the RANS models somewhat lack the ability of simulating turbulent vortexes of different size and scales owing to the assumption of steady-state solution in the numerical analyses. That's why the LES models has been used increasingly nowadays. Some of the studies [10, 108, 17] on LES have been reported for atmospheric flow simulating over two dimensional hills or ridges at laboratory scales. Unfortunately, it is difficult to apply the LES approach over real topography because of some numerical challenges, including reproducing realistic upstream boundary conditions, inhomogeneous surface roughness, resolution of near-wall turbulence structures, or high-Reynolds-number turbulent flow [20]. In addition, the computational costs required by LES simulation are much greater that RANS. Thus, aiming to build a wind power prediction system for operational use, the RANS-based models which have a good compromise between accuracy and computational costs [88] will be chosen in our study.

1.4 Review of the wind and power forecasting

Though the chaotic property of wind results in large difficulties to predict wind power accurately, numerous efforts have been done to develop and improve the power forecasting methods, in order to schedule the spinning reserve capacity and to manage the grid operations. Basically, a wind power forecasting is an estimate of the expected power production from wind turbines (or wind farms) in the near future. It is usually generated using one or a combination of wind forecast models. Therefore, in this section, the forecasting techniques of wind are firstly reviewed. After that, the relationship between wind and power will be discussed and finally an overview of wind power forecasting during various time periods is reviewed.

1.4.1 Wind forecasting

Classification of wind forecasting

The basic role of wind speed forecasting is to provide useful information for operators in the next few minutes, hours, or days. Based on the forecasting horizon, the wind forecasts can be divided into different categories as following:

1. Very-short-term forecasting: from few minutes to 1 hour ahead.

2. Short-term forecasting: from 1 hour to several hours ahead.

3. Medium-term forecasting: from several hours to 1 week ahead.

4. Long-term forecasting: from 1 week to 1 year or more ahead.

Table 1.3 presents the specific time-scale in view of the operation of electricity systems. The applications of specific time-scale in electricity systems are consequentially different. Very short-term forecasts are usually used for load tracking and regulation actions. Short-term forecasts are mainly utilized for making load reasonable decisions. In terms of medium-term forecasting, it could provide useful information for power system management and energy trading. Long-term forecasting is always used for maintenance scheduling of the wind turbines or designing the wind farm.

Table 1.3 Time-scale classification for wind forecasting. (from Stull (1988)[105])

Time-scale	Range	Applications
Very-short-term	Few minutes to 1 hour ahead	Electricity market clearing Real-time grid operations Regulation actions
Short-term	1 hour to several hours ahead	Economic load dispatch planning Load reasonable decisions Operational security in electricity market
Medium-term	Several hours to 1 week ahead	Unit commitment decisions Reserve requirement decisions Generator online or offline decisions
Long-term	1 week to 1 year or more ahead	Maintenance planning Operation management Optimal operating cost Feasibility study for design of the wind farm

Wind forecasting techniques

A wind forecasting model is actually a couple of computer programs which uses various inputs at current time or before to produce wind outputs for future times. There are many wind forecasting techniques including persistence method, physical approach, statistical technique and hybrid approach.

A. Persistence method

In this model, the prediction for all times ahead is just set to the value which it has now, i.e, $w_{t+\delta t} = w_t$. In other words, it is assumed that the wind speed w at time '$t + \delta t$' will be the same as it was at time 't'. Apparently, smaller δt is, better forecasts will be obtained. It is a quite easy way to forecast wind speed but can fabulously provide more accurate forecasts than most of the physical and statistical methods for very-short or short-term forecasts [87].

B. Physical approach

Physical approaches often based on the numerical weather prediction (NWP) models, which could project the real atmosphere state based on an approximation of known physical laws, and result in accurate and reliable estimates for very long time horizons (from hours to several days). Basically, a NWP model is operated by solving complex mathematical models considering many kinds of conditions such as temperature, wind speed and direction, pressure, surface roughness and so on. It is thus a really comprehensive system and needs a lot of computational costs. Different limited-area NWP models, such as WRF (Weather Research and Forecasting), RAMS (Regional Atmospheric Modeling System) and MM5 (Fifth-Generation Penn State/NCAR Mesoscale Model) have been used for wind energy resource assessment by various researchers [104, 93, 106]. However, to a large extent the wind forecasts derived from NWP models are affected by errors stemming from uncertainties in initial / boundary conditions, simplifications in physics and numerical approximations [2]. Great efforts have been devoted to reduce these uncertainties by improving data quality [36] and developing more accurate numerical models with improved dynamic cores and more sophisticated physical parameterizations [70, 73].

C. Statistical technique

Statistical techniques are easier to conduct and more economical in comparison with the NWP based physical methods. In general, the statistical methods used the previous history of wind data to forecast the state at next few hours. They can provide accurate wind speed forecasts over short time scales with limited computational requirements. Sub-classification of this kind of approach is *time-series based models* and *neural network (NN) based methods*.

The auto-regressive moving average (ARMA) model is a well known type in the time-series based approaches to predict future values of wind speed. Its advantages have been testified in the study of Torres et al. [114]. The results indicate that using an ARMA model will make it possible to get 20% error reduction compared to persistence models for forecasting average hourly wind speed for a 10-h forecast horizon. There are also several derivations of ARMA method including auto regressive integrated moving average (ARIMA), seasonal- and fractional-ARIMA, ARMA with exogenous input (ARMAX or ARX). Few other time-series models are linear predictions, grey predictors and exponential smoothing.

The NNs are trained using historical data taken over a long time period to learn the relationship between input data and output wind speed. The NN-based methods generally do not need any predefined mathematical models. When the same or similar patterns are met, the model will come up with a result with minimum errors. As the ARMA, the accuracy of the forecasts for NNs also drops quickly when the time horizon is extended. A couple of NN models including feed-forward neural networks (FNNs), multi-layer perceptrons (MLP), recurrent neural networks (RNNs), radial basis function (RBF) and Adaline networks [23] have been proposed and investigated.

Generally speaking, NNs methods often show priority over time-series models for almost all time-scales, for example, NNs modeling are able to capture the nonlinear pattern in data while traditional time-series methods may not be able to overcome that problem. However, there are still exceptions which are displayed in two studies [12, 11]. It is found that s-ARIMA better follows the actual pattern when both the s-ARIMA and Adaline NN models are chosen to forecast wind speed in Mexico.

Except for the mentioned methods, there are some new techniques including spacial correlation (good for short-term), fuzzy logic model, wavelet transform, ensemble forecasts and entropy based training. Among them, only a brief description of fuzzy logic will be introduced here. Basically, fuzzy logic is a research field based on the principles of approximate reasoning. The fuzzy models are often employed in cases where a system is very difficult to model exactly or when ambiguity and vagueness is encountered in the problem formulation. In the wind forecasting field, a fuzzy system that can predicts the wind speed and generated electrical power has been developed in [30]. After tuning, the forecast horizon has been considered from some minutes up to several hours ahead.

D. Hybrid approach

Hybrid approach is generally defined as a combination of different approaches mentioned in former subsections. The main objective of hybrid approaches is to benefit from the advantages of each model and then obtain a globally optimal forecasting performance. Many types of hybrid approaches are utilized to predict wind, for example, ANFIS, a combination of ANN and fuzzy logic. ANFIS is very good for providing very-short time wind forecast. Moreover, the NNs can be coupled with physical model (NWP) or spatial correlation methods.

1.4.2 Relationship between wind and power

From the fluid mechanical definition, the power output P of a wind turbine is a function of the wind speed and can be expressed as follows:

$$P = \frac{1}{2} C_p \rho A v^3 \tag{1.2}$$

where C_p is the turbine coefficient of performance, and ρ is the air density which depends on air pressure and temperature. A is the swept area of a turbine blade and v represents the wind speed at the turbine site. From equation (1.2), it can be easily found that the relationship between wind speed and power is nonlinear (basically cubic). Therefore, any errors in wind speed forecasts will directly lead to very large errors in predictions of wind power. An easy way to map wind speed into power is using manufacturer's power curve, though it somehow cannot be directly used to predict wind power in real cases. For example, the results in study [25] indicate that using a power curve model derived based on the observed wind and power can improve the forecast RMSE by nearly 20 percent compared to the application just using the theoretical model as equation (1.2). The possible reasons that make the manufacturer's power curve cannot be used in real application will be discussed in the section 2.4.1.

1.4.3 Wind power forecasting

In this section, the existing forecasting techniques of wind power are briefly introduced in four different timescales as bellow.

A. Very-Short Term Forecasting

Very-short-term forecasting aims to provide useful information from few minutes to 1 hour ahead. It is totally required to enable the real-time scheduling of load-frequency control and electricity production. The accurate very-short-term forecasts are often of particular relevance in deregulated energy markets. Moreover, anticipating very short-term load forecasts also can meet the demands of electricity retailers to make appropriate decisions.

There are very few published works on very short-term forecasting of power. In general, the proposed methods might be divided into two categories. One is classical approaches, such as ARIMA models and exponential smoothing. The other one is artificial intelligence based methods like artificial neural networks (ANN) and Neuro-Fuzzy models. Adoption of these methods, from another aspect indicate that the very-short-time forecasting requires a different approach in comparison with the other forecasting time frames. In other words, instead of modeling relationships between demand, time and weather conditions, the very-short-term forecasting is to pay attention to extrapolating the past or recently observed pattern to the nearest future state [109, 16]. These two types of models can be enhanced through the use of a time series of wind data from the nacelle or meteorological (MET) towers within the wind plant. In particular, the second type of approaches may have an advantage if there are more than one MET tower in a wind farm of interest, due to it may be possible to capture some of the variability in weather conditions within the plant and thus will produce a better power generation forecasting.

B. Short-term forecasting

Most of researches related to wind forecasting have been done in this time scale. The approaches for forecasting very-short-term power forecasting somehow can be applied to short-term (day-ahead) forecasting. The ARMA method with only historical generation data for 6-h ahead forecasts is investigated in the study [78] and it is found that ARMA has a significant improvement over persistence models within six hours. Panteri and Papathanassiou [85] introduced the ARX and NN models and then compared with simple persistence method for the forecasting horizons of 1, 3, 6 and 12 hours. The results indicate that NN model delivers the best performance whereas the statistical model shows inferior to the simple persistence method. The RNN is another technique has been used for scheduling autonomous wind-diesel system for next 2 hours [58]. It outperforms many other methods such as persistence as well as classical methods in the literature. Mori et al. [81] proposed a new method for estimating upper and lower bounds and average of wind speed of next 1-h. This method is based on Gaussian Process (GP) with kernel-machine technique and Bayesian estimation. It has been testified using real data of wind speed in the Muroto Cape in Japan. Compared to the performance of MLP and RBF NN methods, the GP-based method can reduce as much as 27% and 12% of average error, respectively. The spatial correlation method also can be used to make short-term wind power forecasting. For example, developed a model based on local and spatial relations of the wind speed at neighboring sites for the sake of improving the accuracy and efficiency for short and long range forecasts, ranging from a couple of minutes to several hours ahead. By using nearly a whole year's measurements, this method has been verified and the results illustrated that the forecasting efficiency has improved by 28% compared to the persistence method, which indicates that the data at neighboring sites is always very useful. Another similar application of spatial correlation method using TSK fuzzy interface model for 2-hour ahead forecasting is presented in [30]. A novel technique for wind speed forecasting and wind power prediction based on using the Grey model (GM) is [40]. The brief conclusion is that using the GM(1,1) the forecasts of wind speed has an average accuracy of 11.2% better than the persistent model and while the predicted output power has a better average accuracy of around 12.2%. Aiming to absorb the advantages of each methods, a hybrid method is presented in [68] which firstly uses wavelet method to decompose original time-series into a couple of subseries and then an improved ARIMA method is adopted to predict the next future values in each subseries. The performance of this hybrid method is compared with classical time-series model and BP NN using mean absolute error (MAE), mean square error (MSE) and mean absolute percentage error (MAPE) criteria. It is found that the hybrid method gives better results (less error) compared to others for 3-step, 5-step and 10-step ahead prediction respectively. As expected, the bigger the forecasting steps are, the lower is the accuracy. Of course there still many statistical methods for forecasting short-term power output have not mentioned here, more specific review can be found in [101, 46].

Another common way we usually choose to obtain the short-term forecasts of power begins with the grid point output (i.e., atmospheric variable, e.g, temperature, wind speed and direction) from a meso-scale, limited area and physical-based atmospheric model. Then inputing such outputs of atmospheric model as an input into a power forecasting system (e.g, power curve), the expected forecasts of power output will be obtained. The procedure of this kind of methods differ substantially. Some forecasting procedures attempt to start directly from the limited area forecasting results to the local-scale by using either diagnostic physical models, statistical models or a hybrid of both. The Prediktor tool developed by the Risoe National Laboratory adopts this approach, aiming to provide the expected production of wind farms up to 48 hours with a interval of 6-h. However, this approach often misses the detailed information of processes occurring at the sub-regional. An alternative method to resolve that problems is to execute sub-regional scale simulations with a physics-based model, which has already used by TrueWind Solutions in their eWind system. There are lots of other forecasting systems, some of which are called WPPT (wind power prediction tool), Zephyr [47] and WPFS (wind power forecast system), have been implemented by many case studies in Germany, Ireland, Spain, France and Denmark and all have a considerable skill over persistence forecasts for 1 to 2 day periods [119]. Unfortunately, it is difficult to obtain a quantitative assessment of all existing forecasting techniques as the methods, locations and time periods which always vary substantially. An existing system which performed well in other countries may not be suitable in the case of wind farms in Japan. Therefore, one of targets of this thesis is to develop a comprehensive system to provide short-term forecasts for a wind farm considering the specific situation in Japan. Though, in this thesis we only focus on short-term forecasting, the progress on medium-term and long-term is also necessary to known for future plan.

C. Medium-term forecasting

Most of methods developed for power forecasts at this time-scale are mainly based on NN approaches, physical weather models and the hybrid models combining both of these or other new techniques. Some of them are reviewed as below.

The study [14] presents an ANN approach for short-term wind power forecasting in Portugal. The implementation of this approach to do wind power forecasting has been proven to be successful, due to its MAPE has 12% less than persistence model while the average computation time is just 5 seconds. This indicates that the proposed method shows a good compromise between forecasting accuracy and computation time. The method of wavelet transform is used in the study [61] to decompose the signal and cut up original data into different frequency components. This approach is successfully applied on the wind data obtained from Colorado public utility sites and it can provide forecasts up to 24 hours. An advanced statistical method based on NWPs and RBF NN for wind power forecasting (1-48 hours ahead) is developed in [96]. It provides an estimation of the forecasts quality based on errors between forecasted and actual values of power, to subsequently improve the raw NWP outputs using fuzzy rules. Results are compared using indexes of MAE and RMSE against naïve predictor and 'solely' NWP method, which show substantial improvement for both (nearly 46% for persistence). Kavasseri and Seetharaman [59] proposed a new f-ARIMA model, which has been applied to hourly average wind speed records obtained from four potential wind generation sites in North Dakota. The parameters of this model were estimated using 'exact maximum likelihood' and optimized using Akaike's information criterion. The results show that on an average (over the four records considered) the DME (daily mean error) is 79.3% with the persistence model, 117% with the ARIMA models while 47% for the f-ARIMA models. Moreover, the proposed f-ARIMA models yield a much smaller variance (0.24) compared to the persistence (1.07) and ARIMA method (0.89). In addition, the conversion from wind speed to power forecasts also could obtain accurate results for 24 to 48-hour ahead. Lazić et al. [66] planned to apply the regional NWP Eta model and describe its performance in validation of the wind forecasts for wind power plants. Two sets of Eta model forecasts are generated, i.e, one with a very coarse resolution of 22 km and the other one added a nested grid of 3.5 km, centered on the Nasudden power plants at island Gotland, Sweden. The 12 to 36- hour ahead wind speed forecasts are compared with the observed wind at nearest surface station and at 10 m height above the ground at a site in Sweden respectively. The results show a high coefficient of determination (COD) of 0.8, and low MAE and RMSE, which indicates that Eta model is a very useful tool for wind energy modeling. Fan et al. [42] developed a novel forecasting model based on deep investigations of meteorological information. This model adopts a two-stage hybrid network with Bayesian clustering by dynamics and support vector regression (SVR). In this model, the role of “Bayesian clustering” is to classify the input training dataset into several subsets (with similar properties) and then SVR fits training data in each subset in supervised manner. This model is applied and tested on a 74-MW wind farm located in the southwest Oklahoma of the US and results indicate that it can improve over persistent for all predictions by 40% based on RMSE and MAE criteria. Recently, Croonenbroeck and Dahl [27] suggest an accurate wind power forecasting methodology that exploits many of the actual data’s statistical features, in particular both-sided censoring. The proposed model produces turbine-specific forecasts which are significantly more accurate when compared against the established benchmark models (persistence). Later, a new wind power forecasting model that does not focus on providing the most precise forecasts but minimizing the financial loss of forecasting impreciseness is presented in [28].

D. Long-term forecasting

The significance of long-term power forecasting is reflected in power system planning whose purpose is to define the generation capacity trend that sufficiently meets the demand for electricity within a context of technical, economical and environmental constrains. A few studies have been done on long-term wind forecasting approaches. Taylor et al. [110] developed a new type of physical method to predict PDF of wind power generation for 1 to 10-day ahead forecasts using "weather ensemble predictions" (WEP). A calibration method that incorporated kernel smoothing, with parameters optimized using maximum likelihood is used to correct the systematic errors in the location and scale of the distribution of the 51 ensemble members. By comparing with statistical time series methods (i.e., ARMA based models), it is found that WEP could give more accurate results over a period of week. Another method of long-term wind speed and power forecasting using local RNN is developed in [5]. The authors of this paper used three RNNs with NWP results as input from four nearby nodes and the final results of comparing with the persistence model indicate a 50% improvement. However, this kind of technique is very exhaustive and time-consuming. This means that it is required for RNN to forecast task in lesser time and without requiring too much meteorological data or complications. As shown in [92], the problem can be slightly solved by applying RNN for wind forecasts of each month separately. Salcedo-Sanz et al. [91] choose MM5 model using meteorological data from a global NWP as its initial and boundary conditions to study the long-term forecasting. Coupling the MM5 model with NNs could obtain better forecasts compared against persistence model [89].

1.4.4 Meteorological modeling for wind power forecasting

As mentioned earlier, NWP model plays a very important role in wind power forecasting at various time scales. Thus, a brief introduction of NWP models is described here. Generally, NWP is a type of weather forecasting which is based on the output of complex computer programs, known as forecasting model. The model actually is a translation of a set of governing equations which describe the air flows. It usually consists of multiple parts including data collection and assimilation, forecasting, post-processing and distribution to users. The first two parts directly decide the forecasting skills of the whole forecasting procedure. Therefore, many efforts have been done to improve them, for example, developing

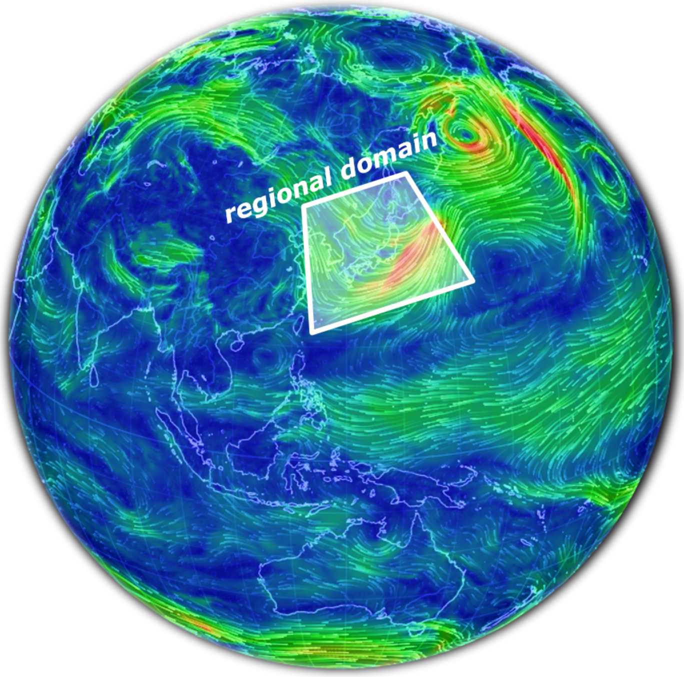

Figure 1.6 An example of a regional model domain over the Japan on a global grid.

optimal data assimilation algorithms to provide more precise initial condition and improving the accuracy of numerical methods to reduce model errors.

According to space scales, the numerous NWP models can be simply divided into two types. One is called global models whose calculation domain covers the entire Earth; the other one is regional models also known as limited-area models which use finer grid resolution to resolve explicitly smaller-scale meteorological phenomena. The relationship between these two is displayed vividly in the Figure 1.6. Regional model is nested within the global model to determine all the details for the region of interest and in turn the global model provides the necessary information for running the regional model. Table 1.4 shows the information of some existing NWP models in the world.

Those models can be used for a wide range of applications. The most important application is to provide precise predictions in both time and space for daily activities such as travel, health and safety. They also can provide timely warning of weather extremes (e.g., floods and gales) which could lead to great financial losses as well as people's lives. To our interest, the NWP models are very useful tools in the field of wind energy. They are widely

Table 1.4 Selected main characteristics of the numerical weather prediction models.

Model name	Hydrostatic or nonhydrostatic	Global or regional model	Owner
ECMWF	Nonhydrostatic	Global	European Union
GEM	Hydrostatic	Global	Canada
GFS	Nonhydrostatic	Global	NCEP
UKMET	Nonhydrostatic	Global/Regional	United Kingdom
NOGAPS	Nonhydrostatic	Global	Navy
T639	Nonhydrostatic	Global	China
GSM	Nonhydrostatic	Global	Japan
WRF	Nonhydrostatic	Regional	NCEP
MM5	Nonhydrostatic	Regional	PSU
NAM	Nonhydrostatic	Regional	NCEP
MSM	Nonhydrostatic	Regional	Japan

used for the wind resource assessment and power forecasting. In addition, they can be easily combined with other models (e.g., micro-scale model) and post-processing methods (e.g., Kalman filter and MOS) to obtain better results. Considering that, in this study, one of them which is named WRF is chosen to reach our research purpose of providing power forecasts as accurate as possible.

1.5 Motivation and Objectives

The recent assessments of the World Energy Council have reported the difficulties that Japan is facing in lieu of the shortage of energy caused by the shutdown of nuclear plants after the Fukushima accident. In order to compensate for the loss of the nuclear power generation, extra fossil resources are currently imported, which has led to the rise of electricity cost and to an increase of the fossil fuel emissions. As an alternative to the traditional fossil energy resources, renewable energy has been the focus of recent developments as a long-term and sustainable solution. In particular, wind energy has by far shown tremendous potential in terms of economic and environmental effects.

However, numerous factors hinder the development of the wind power in Japan. One of the biggest obstacles is the challenge of integrating the wind power with the existing smart girds, which is mainly influenced by the variability and limited predictability of wind. That is, accurate and reliable wind or power predictions are necessary to optimize the integration of wind power into existing electrical systems. Therefore, based on the discussion above, we give the purposes of our study from the following four parts:

- **Building a preliminary forecasting system based on the WRF and a power curve model for the Awaji wind farm.**

  We want to develop a wind power prediction system for the Awaji wind farm in Japan as an effort to facilitate the short-term wind power forecasts in Japan area. The preliminary system is mainly based on the high resolution WRF model and a simple power curve model which is constructed based on a polynomial fit technique using the historical data of the observed wind speed at hub-height and power output. Of course the forecasting ability of the WRF and power curve model needs to be evaluated, to show that they are able to provide reasonably reliable forecasting results in the target site which has complex geographic environment very typical in Japan.

- **Using the Kalman filter algorithm to improve the forecasting ability of the system already built based on the available observations from the wind farm of interest.**

  As known, short-term wind energy prediction relies heavily on the low-level wind forecasts derived from the NWP model. As documented in [63, 24], recent progresses in forecasting skills of NWP models make it possible to provide more reliable predictions of surface wind field which is essential for wind energy management. However, the current NWP models are still far from a mature stage and particularly large errors are found in the prediction of the surface wind forecasts, which has motivated continuous efforts to improve the NWP models themselves [31, 97]. The encouraging results in these works suggest that more reliable forecasts can be made by using more advanced and sophisticated numerical models which are generally believed to possess dynamic cores with less assumptions and accurate numerics, refined parameterization packages for physical processes that directly affecting the phenomena of interests. Unfortunately, other factors beside the model inaccuracy, such as the uncertainties in observations and the chaotic nature of atmosphere, also prevent the outputs of deterministic NWP models being directly usable to many applications. Based on these consideration, the Kalman filter is regarded as a module of the integrated forecasting system we finally want to develop. It mainly aims to post-process the outputs of WRF model (reducing errors) to make the forecasts being suitable for the wind farm of interest.

- **Using data assimilation technique to generate the initial conditions (known as the "analysis") as close as possible to the real atmosphere for the WRF in use to improve its forecasting skills.**

  Apparently the Kalman filter module is aiming to reduce the WRF model errors, however, it somehow cannot deal well with the random errors caused by many factors, for example, the inaccurate initial conditions of the WRF model. The data assimilation technique happens to have an ability of generating better initial conditions (known as the "analysis") [57] for WRF model. That is why we also want to develop the data assimilation technique as another key module of our final integrated forecasting system.

- **Coupling meso-scale WRF model with micro-scale OpenFOAM to investigate how the complex terrain affect the wind flow at a wind farm.**

  Japan has its unique topography which made the wind conditions more complex than other countries (e.g., US and Europe) in the world. Although a forecasting system has been developed by combining the Kalman filter and DA with WRF model and it definitely can be used to do the operational forecasting over a common wind farm, such system still needs to be improved to adapt the special situation in Japan's wind farm. The main starting point is WRF model based system cannot handle the local terrain features because of the relatively coarse resolution. Higher resolution modeling might can fit this shortcoming. With this in mind, Computational Fluid Dynamics (CFD) model is an option for simulating the flow characteristics of smaller scales due to a finer resolution of terrain features. However, it is imperfect when we use those two models solely to predict wind flows over the complex terrain at a real wind farm for operational use. Coupling those two components together is expected to a better way to accurately forecast short-term hub-height wind for any wind farms in Japan.

All in all, we want to develop an integrated and multi-scale forecasting system for operational prediction of wind and power under complex geographic conditions, which is summarized in the Figure 1.7. This integrated system consists of WRF model, a power curve, Kalman filter, data assimilation and micro-scale OpenFOAM model.

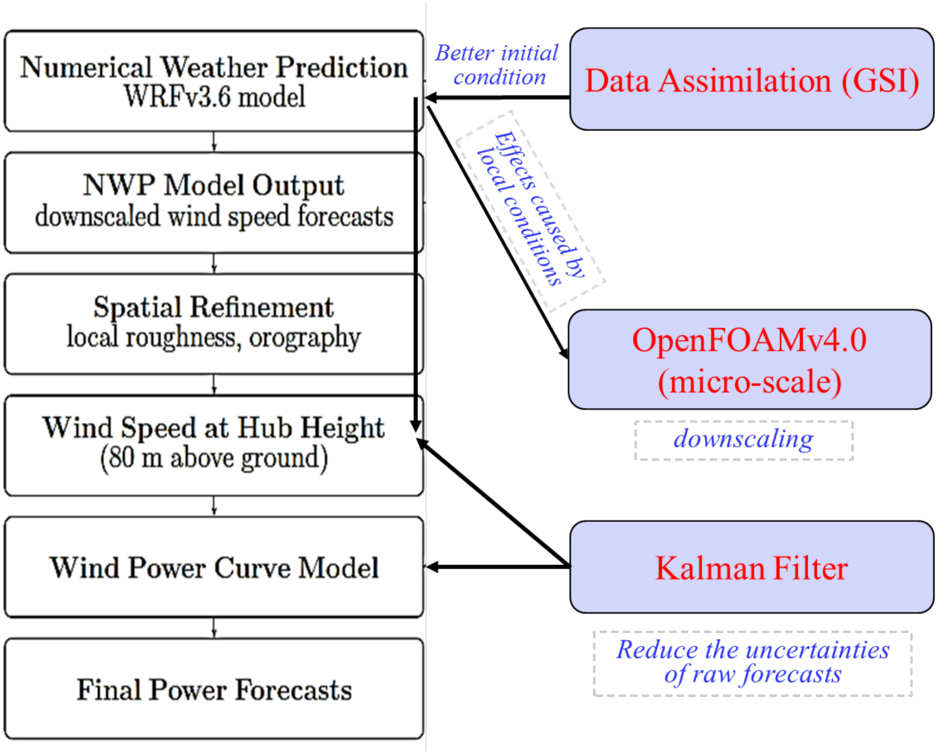

Figure 1.7 A schematic workflow of the thesis.

1.6 Outline of the thesis

This thesis consists of six chapters. Chapter 1 is focused on the global status of wind power development, current situation in Japan, the properties of wind in ABL, the background of wind power forecasting, and finally the motivation and objectives of this thesis.

Chapter 2 gives the information of a preliminary forecasting system we build based on the WRF model and a power curve model. Initially, a brief description of datasets and WRF model used through this thesis is introduced. Then the ability of WRF model for forecasting hub-height wind at the Awaji wind farm is studied. In parallel, a power curve model is constructed based on a polynomial fit technique to predict wind power for each wind turbine of the Awaji wind farm.

In chapter 3, the Kalman filter is adopted as a module of our integrated system, which could improve the forecasting skill of the preliminary system mentioned in chapter 2. In this chapter, firstly, the details of Kalman filter is introduced. Then its impacts on both wind speed and power forecasts are validated.

Chapter 4 considers the possibility of further improving the ability of our system. Thus, the data assimilation technique is also chosen as another key module of the final integrated forecasting system. The benefits obtained by applying data assimilation are clearly shown.

Chapter 5 describes the process of coupling the meso-scale WRF with the micro-scale OpenFOAM model. Initially, the theoretical background based on OpenFOAM is presented. Then the procedure of coupling those two models is stated and the performance of the coupled system for predicting wind flow over the complex terrain conditions is studied eventually.

Finally, chapter 6 summarizes the overall conclusions from the present research works and gives the perspectives for future study.

Chapter 2

The basic system for wind and power forecasts

In this chapter, the two kinds of datasets (i.e., GFS data and nacelle wind and power observations) used in this thesis are introduced firstly in the section 2.1. Then we will give a brief introduction of the WRF modeling system in the section 2.2 and in the next section, its ability of forecasting hub-height wind at the Awaji wind farm is studied. At last, a power curve model is constructed based on a polynomial fit technique to predict wind power for each wind turbine of the Awaji wind farm.

2.1 Datasets

2.1.1 GFS Data

The Global Forecast System (GFS) is a global numerical weather prediction system produced by the National Centers for Environmental Prediction (NCEP). Dozens of atmospheric and land-soil variables, ranging from winds, geopotential height, temperatures and precipitation to soil moisture and atmospheric ozone concentration, are all available through this system. The GFS model covers the entire globe and can provide deterministic and probabilistic guidance out to 16 days. The GFS model is a coupled model, composed of four separate models (i.e., an atmosphere model, an ocean model, a land/soil model, and a sea ice model), which could work together to make an accurate prediction of future weather conditions. The main model used in GFS is a global spectral model (GSM). The current operational dynamical core of the GFS/GSM is based on a two time-level semi-implicit semi-Lagrangian discretization with three dimensional Hermite interpolation [15].

With the increased computing resources and changing computer architecture at NCEP, the GFS has already evolved to higher resolutions, on both horizontally and vertically. The current operational horizontal resolution is 13 *km* for the first 10-day and about 34 *km* from 240 to 384 hours (days 10-16). In the vertical direction, there are 64 sigma-pressure hybrid layers with the most top layer centered around 0.27 hPa (nearly 55 *km*). There is no doubt that many persistent changes are regularly made to the GFS model in order to improve its performance and forecast accuracy.

The gridded products with four horizontal resolutions (i.e., 0.25, 0.5, 1.0 and 2.5 degree) are available for downloading through the NOAA National Operational Model Archive and Distribution System (NOMADS). The global forecasts are made four times daily at 0000, 0600, 1200 and 1800 UTC. The reason we choose those data products is that they can provide initial and/or boundary conditions for other models, for example WRF model of interest in this thesis. Therefore, nearly seven-month (2013/08/01-2014/01/31; 2016/01/01-2016/02/01) GFS data are used as the initial and boundary conditions for the WRF model with a 6-hour interval. The horizontal resolution of all variables we adopted is 1.0×1.0 degree, except for the data in the year 2016 (using 0.5×0.5 degree). In the vertical, 27 pressure levels ranging from 1000 to 10 hPa are chosen.

2.1.2 Observations and data quality control

The main target region is a wind farm located in south Awaji island, Japan, where 15 wind turbines have been installed. All wind turbines (General Electric GE2.5) have a rated capacity of 2.5 MW and the power curve is displayed in Figure 2.1. The rotor diameter of the turbines is 84 *m* and the tower height is 80 *m*. The nacelle wind for each turbine is measured by the anemometers placed on the top of the nacelle behind the rotor. Thus, in this study, the nacelle-based wind speed from 15 turbines at hub-height (80 m above ground) are utilized to build as well as evaluate the performance of our integrated forecasting system. In addition, observations of power output are also used to evaluate the reliability of the power prediction proposed in this thesis. Both wind speed and power observational data are available every 10 minutes for the six-month period from 1 August 2013 to 31 January 2016. In the evaluation process, the forecasts from the forecasting system are firstly interpolated (3D interpolation) to those 15 turbines and then compared with the corresponding observations respectively.

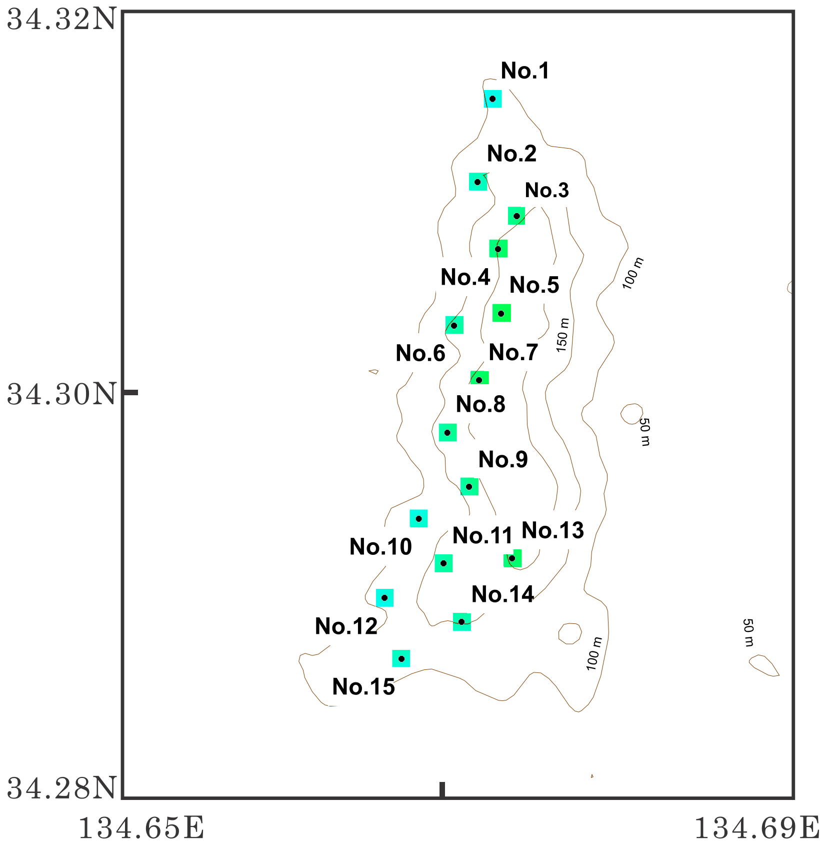

Figure 2.1 The wind farm site at Awaji island, Japan. Contours stand for the terrain elevation.

In general, the nacelle-based wind data are always used by wind farm operators directly for the turbine control (e.g. to determine the cut-in/cut-out speeds). However, some studies has pointed out that this kind of data cannot be used for research or application directly without considering the existing uncertainties, for example, the effects caused by the rotating blades and nacelle. It seems that the observed data from the upwind meteorological tower are more reliable. The effects of the rotating blades and nacelle on the observed wind can be taken into account by adjusting the relationship between the nacelle-based observations and the measurements from the upwind meteorological tower [100]. Unfortunately, we cannot expected to build a lot of meteorological towers close to turbines, which would increase the cost and is always not adopted in practice, especially for small-scale wind farms in Japan. Therefore, to some extent, as suggested in the study [29], the nacelle-based wind speed observation might be a more reliable estimation to the target turbine site than the measurement obtained from a meteorological tower with a large distance away. Thus, the nacelle wind data still are chosen in this thesis. Apparently, it is very necessary that the quality of the nacelle wind observations needs to be evaluated before using them, especially for data assimilation. When the wind data are assimilated into the NWP models, its quality must be checked so as to avoid the degradation of forecasting skill due to the assimilation of a few bad data points, which might even outweigh the benefits of assimilating many other good data points. Thus, we, in this study, implement the standards addressed in the technical report of NOAA Earth Systems Research Laboratory (ESRL) [44] to flag out the unreasonable data points for each turbine separately.

2.2 Introduction of WRF modeling system

The Weather Research and Forecasting (WRF) model is numerical weather prediction and atmospheric simulation system which is designed for operational forecasting as well as research. Its development has been a multi-agency efforts to build a next-generation NWP model and data assimilation system to advance the understanding and prediction of mesoscale weather system. Those efforts began in the latter part of the 1990's and was a collaborative partnership principally among the National Center for Atmospheric Research (NCAR), the National Oceanic and Atmospheric Administration (represented by the National Centers for Environmental Prediction (NCEP) and the Forecast Systems Laboratory (FSL)), the Air Force Weather Agency (AFWA), the Naval Research Laboratory, the University of Oklahoma, and the Federal Aviation Administration (FAA) [115].

2.2.1 Structure of WRF modeling system

In general, the WRF modeling system consists of three major parts, including WPS (WRF pre-processing system), the WRF model and post-processing and visualization component (e.g., GrADS and NCL). The inter-relationships among those three parts are displayed in the Figure 2.2. It is easily found that there are two kinds of dynamical cores. One is Advanced Research WRF (ARW) and the other one is Nonhydrostatic Mesoscale Model (NMM). In this thesis, the former one is chosen which can be used to do both idealized and real simulation. In the subsequent subsections, the detail of the each components of WRF modeling system will be introduced briefly.

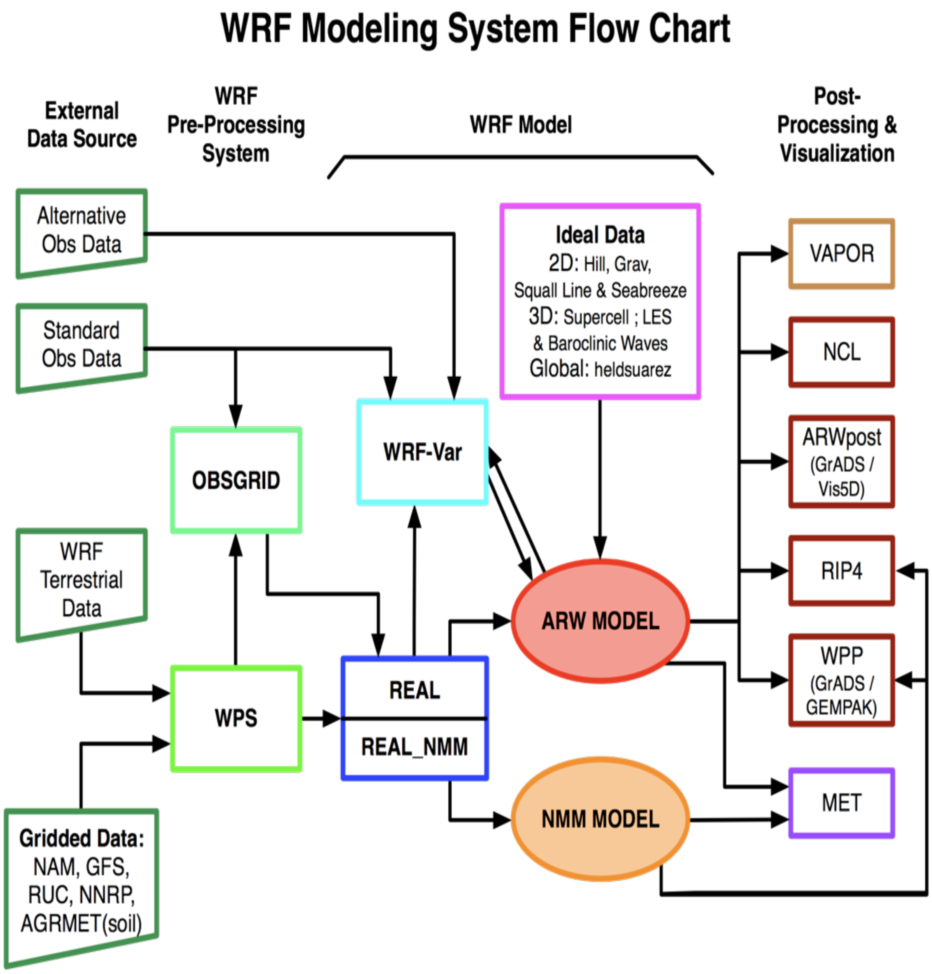

Figure 2.2 The structure of the general WRF modelling system.

2.2.2 WRF pre-processing system

This part actually is to generate the necessary information for the sake of running WRF model properly. Its functions can be roughly divided into three classes, including defining simulation domains, interpolating topography data (e.g., terrain, landuse, and soil types) to the simulation domain and degribbing and interpolating background data from another model (global or regional) to this simulation domain. The detail of each is briefly described below.

Map projections and domain configuration. In the current ARW WRF system, there are four projection methods including Lambert conformal, Polar stereographic, Mercator and latitude-longitude projection. Different projections are suitable for different areas on the globe, for example, the Mercator projection is the best choice for low latitudes regions. In the current study, we focus on the wind farms in Japan and thus the Lambert conformal projection is chosen.

The choice of the resolutions (both horizontal and vertical) for WRF model is important. In general, higher resolution lead to better forecasts over a limited region, though this will increase the computational burden. Therefore, the nested-domain is always adopted for doing real simulations. A common nested-domain configuration consists of one parent domain and one or more children with higher horizontal resolution than the parent domain. The nested domains receive information from their parent domain (the adjacent outer domain), for example, obtaining boundary conditions. This kind of configuration is very important for real prediction. It somehow can overcome the shortage of computational resource while could obtain the same precision of forecasts for a target region (limited) as a single-domain simulation with uniformly high resolution does. The configuration of the nested-domain are various, some examples of nested-domains are displayed in the Figure 2.3.

There are two kinds of nesting options, i.e., one-way and two way. It is defined as one-way if information exchange between the parent and the nested domain is strictly down-scale. In other words, the nested solution does not feedback to the coarser or parent solution. On the other hand, in two-way mode, information exchange between the parent and the nest is bi-directional. The feedback from sub-domain usually impacts the coarse-grid domain's solution. The one-way nesting is chosen in this thesis.

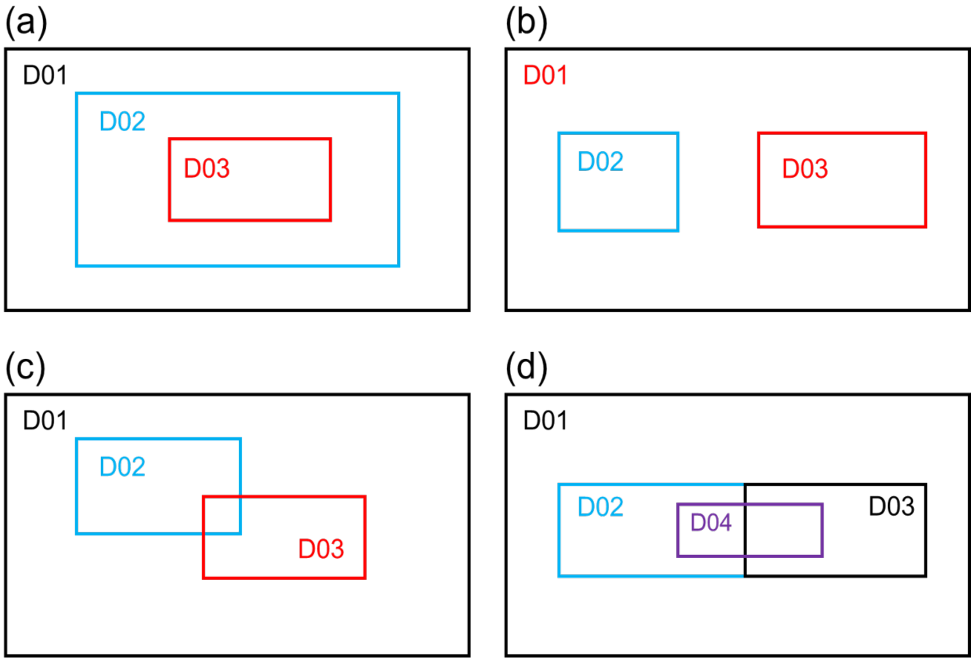

Figure 2.3 A couple of nested domain configurations for WRF model. (a) Telescoping nested grid. (b) Nesting at the same level with respect to one parent domain. (c) Overlapping nesting. (d) Most inner grid has more than one parent grid. It should be noted that the ways of (c) and (d) are not allowed.

Topography data and interpolation. Using WRF modeling system to do the real prediction, the terrestrial data are necessary. Various terrestrial datasets such as terrain height, landuse, soil type, annual deep soil temperature, monthly vegetation fraction, maximum snow albedo, monthly albedo and slope data are available. A few of the datasets are available in only one resolution, but others are made available in resolutions of 30”, 2’, 5’, and 10’. Those data are to be interpolated using an inside program named GEOGRID to the model grids which have been decided at the domain configuration step. Except above data, new and additional datasets also can be interpolated to the simulation domain through modifying a table file named GEOGRID.TBL, which defines each of the fields that will be produced by GEOGRID. Outputs from GEOGRID are written in the WRF I/O API format. Generally, selecting the NetCDF I/O format, GEOGRID can be made to write its output in NetCDF for easy visualization using some external software packages, for example ncview.

Degribbing and interpolating meteorological data. Here the meteorological data refers to the various atmospheric fields such as wind, temperature and pressure which are derived from other models, for example NCEP GFS or ECMWF interim (European Centre for Medium-Range Weather Forecasts). Those models usually can provide global and relatively large-scale forecasts for nearly two weeks in advance. Therefore, they always are chosen as the initial and boundary conditions for WRF model (over limited regions). However, the format of these data usually is GRIB, which cannot be recognized by WRF model directly. Fortunately, another program which is called UNGRIB could extract meteorological fields from GRIB formatted files. It also uses a specific table file named VTABLE to control what kind of information is needed. Various types of VTABLE are available for degribbing different GRIB data taken from different models.

After degribbing the GRIB data using UNGRIB, the immediate outputs should be interpolated to model grids, in order to use them as the initial/boundary conditions appropriately. This procedure could be done by using METGRID program. As GEOGRID, there is also a similar table file, which is called METGRID.TBL. This file is possible to specify options such as the interpolation methods to be used for the fields, the field that acts as the mask to be used for masked interpolations, and the staggering (e.g., U, V in ARW) to which a field is to be interpolated.

2.2.3 WRF model

WRF model initialization

The WRF model has ability to do two large classes of simulations and thus there are two kinds of WRF initializations: an ideal initialization and utilizing real data. The WRF model itself is not altered whichever is chosen, nevertheless the WRF pre-processors are to be specifically built based upon a specific selection. The real data initialization is selected in this thesis because of our interest which is to build a developed forecasting system for a real wind farm.

In order to explain how to do the real case initialization of WRF model, we recall the description of WPS in section 2.2.2. In the first step, the model grids should be prepared, which are controlled by some parameters in the file of *namelist.wps*, which can be generated more easier by using a external package named WRF Domain Wizard. It is a graphical user interface (GUI) for the new WRF Preprocessing System and it enables users to easily define and localize domains (limited) by selecting a region of the Earth as well as choosing a map projection. When the model grids are prepared, the program of GEOGRID will interpolate various terrestrial data to this model grids and some files (i.e., geo_em.d0$_*$.nc) are generated. Similarly, the METGRID program also can interpolate the meteorological fields generated by program of UNGRIB to the model grids, and finally the files of met_em.d0$_*$.YYYY-MM-DD_HH:MM:SS.nc are obtained. The generated data from METGRID consists of meteorological fields and terrestrial information, however, it cannot directly recognized by WRF model. Therefore, there is another program named REAL which is in charge of merging the data provided by the WRF preprocessing system and finally outputting the initial and boundary conditions for running WRF model. A brief summary [115] of other additional roles of real data initialization are listed as follows,

- Read various data generated from the WRF preprocessing system
- Compute dry surface pressure, model levels, and vertically interpolate data
- Compute reference temperature profile (differently than with ideal cases, to allow for seasonal norms)
- Prepare soil fields for use in model (usually, vertical interpolation to the requested levels)
- Checks to verify soil categories, land use, land mask, soil temperature, sea surface temperature are all consistent with each other
- Multiple input time periods are processed to generate the lateral boundary conditions
- Three dimensional boundary data (u, v, t, q, ph) are coupled with map factors (on the correct staggering) and total mu

WRF model solver

The WRF model is a fully compressible, Euler non-hydrostatic model (with a hydrostatic option). The model uses the Runge-Kutta 2nd and 3rd order time integration schemes. The spatial discretization may be selected from 2nd to 6th order advection schemes in both horizontal and vertical directions.

The horizontal grid staggering is the Arakawa C-grid (as shown in Figure 2.4a). It is easy to find that the components of horizontal wind (u, v) are normal to the respective faces of the grid cell while the mass/thermodynamic/scalar variables (e.g., T, ρ, q_v) are located in the center of the cell. Similarly, the staggering grid is also adopted in the vertical direction which can be found in Figure 2.4b.

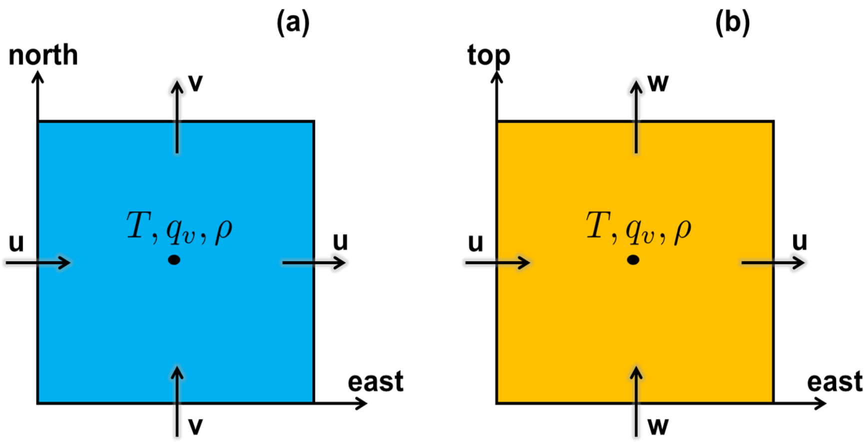

Figure 2.4 Horizontal (a) and vertical (b) grids of the WRF model (Arakawa C staggering).

As displayed in Figure 2.5, in the vertical, WRF model choose a terrain-following hydrostatic pressure coordinate denoted by η, which is defined as:

$$\eta = \frac{P_h - P_{ht}}{\mu} \tag{2.1}$$

where p_h is the hydrostatic component of th pressure and P_{ht} stands for the pressure at the top boundary. The parameter μ is expressed as,

$$\mu = P_{hs} - P_{ht} \tag{2.2}$$

where P_{hs} refers to the value of pressure along the surface. This kind of coordinate [64] which is also called a mass vertical coordinate is similar to the traditional σ coordinate used in many hydrostatic atmospheric models. Apparently, η varies from a value of 1 at the surface to 0 at the top boundary of the model domain (Figure 2.5).

Due to the μ in equation 2.1 stands for the mass per unit area within the column in the model domain, the appropriate flux form variables are defined as,

$$\mathbf{V} = \mu \mathbf{v} = (U, V, W) \tag{2.3}$$

$$\Omega = \mu \dot{\eta} \tag{2.4}$$

$$\Theta = \mu \theta \tag{2.5}$$

where, $\mathbf{v} = (u, v, \omega)$ are the covariant velocities in the two horizontal and vertical directions respectively, and $\omega = \dot{\eta}$. θ refers to the potential temperature.

Based on the above definitions, the flux-form Euler governing equations can be expressed as following:

$$\partial_t U + (\nabla \cdot \mathbf{V} u) - \partial_x (p \phi_\eta) + \partial_x (p \phi_x) = F_U \tag{2.6}$$

$$\partial_t V + (\nabla \cdot \mathbf{V} v) - \partial_y (p \phi_\eta) + \partial_y (p \phi_y) = F_V \tag{2.7}$$

$$\partial_t W + (\nabla \cdot \mathbf{V} \omega) - g(\partial_\eta p - \mu) = F_W \tag{2.8}$$

$$\partial_t \Theta + (\nabla \cdot \mathbf{V} \theta) = F_\Theta \tag{2.9}$$

$$\partial_t \mu + (\nabla \cdot \mathbf{V}) = 0 \tag{2.10}$$

$$\partial_t \phi + \mu^{-1} [(\mathbf{V} \cdot \nabla \phi) - gW] = 0 \tag{2.11}$$

along with the diagnostic relation for the inverse density and the state equation:

$$\partial_\eta \phi = -\alpha \mu \tag{2.12}$$

$$p = p_0 (R_d \theta / p_0 \alpha)^\gamma \tag{2.13}$$

where F_U, F_V, F_W and F_Θ indicate forcing terms arising from model physics, turbulent mixing, spherical projections, and the earth's rotation. R_d is the gas constant for dry air and p_0 is a reference pressure (generally 10^5 Pascals). Also appearing in the governing equations are the non-conserved variables $\phi = gz$ (the geopotential), p (pressure), and $\alpha = 1/\rho$ (the inverse density). γ stands for the ratio of the heat capacities for dry air

$$\gamma = c_p/c_v \tag{2.14}$$

where c_p and c_v are the the heat capacity at constant pressure and at constant volume, respectively.

Except for the equation 2.11, the governing equations 2.6 – 2.10 are all cast in conservative form. In fact, the equation 2.11 could be written in flux form, however we cannot find any advantages in doing so due to $\mu\phi$ is not a conserved quantity. Additionally, it should be noted that the relation for the hydrostatic balance (equation 2.12) does not represent a constraint on the solution, rather it is a diagnostic relation that formally is part of the coordinate definition.

Physics options

In WRF model, the physics processes are insulated from the rest of the dynamics solver by the use of physics drivers.This section will briefly outline the physics options available in the WRF model, which fall into several groups: microphysics, cumulus parameterization, planetary boundary layer (PBL), land surface model (LSM) and radiation. Each category also contains several options.

Microphysics. It includes water vapor, cloud, and precipitation processes. In the current WRF model, microphysics is carried out at the end of the time-step as an adjustment process, and so does not provide tendencies. The start point of doing this is that condensation adjustment should be at the end of the time-step in order to guarantee that the final saturation balance is enough accurate for the updated temperature and moisture. There are many options for these processes in the current WRF model. *Kessler scheme* [60] is one of them, which is take from the COMMAS model and is a simple warm cloud scheme that contains water vapor,

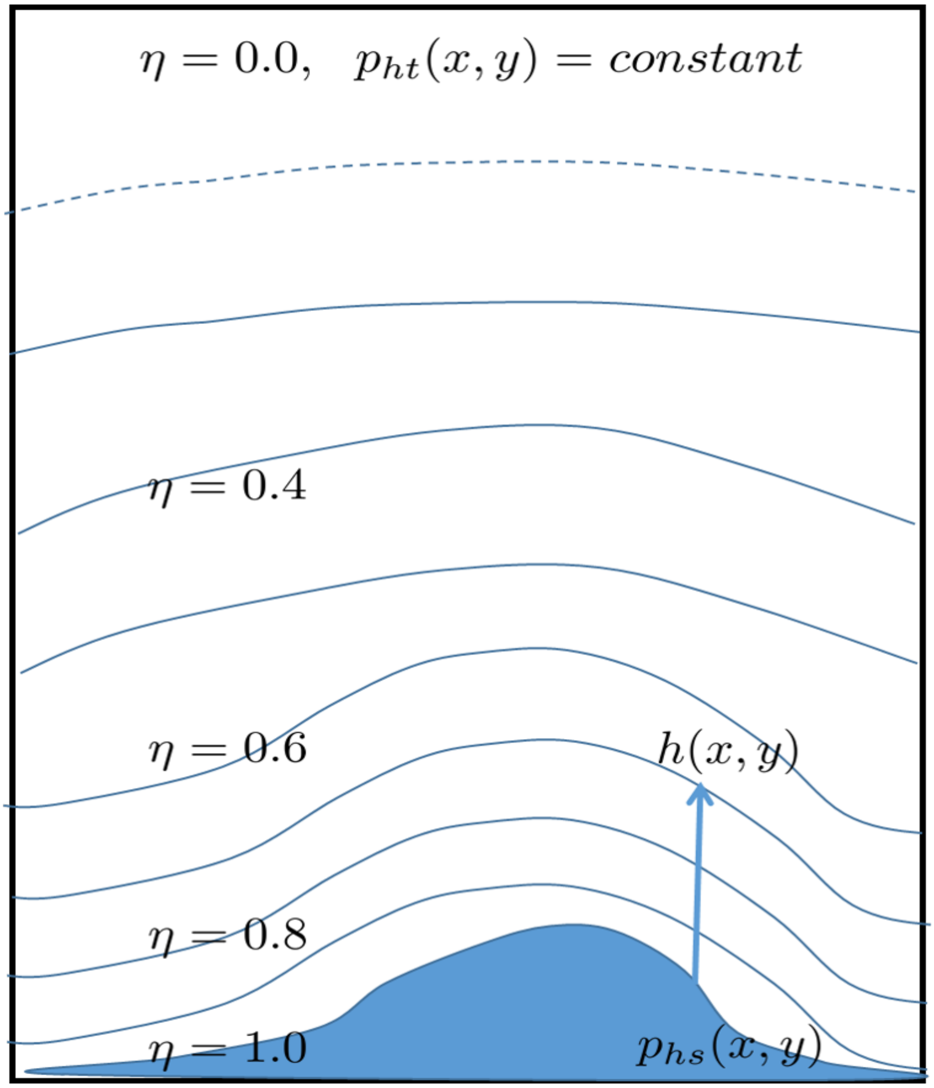

Figure 2.5 The vertical-coordinate of the WRF model.

rain and cloud water. *Purdue Lin scheme* which is based on the study of Lin et al. [67] is a very famous scheme to resolve cloud water, rain, water vapor, cloud ice and snow. Another scheme named *WRF Single-Moment 3-class (WSM3) scheme* [50] includes ice sedimentation and other new ice-phase parameterizations revised from the older NCEP3 scheme. Similar to the WSM3, there are other two schemes, i.e., *WSM5* and *WSM6* [51] are added in the WRF model. In addition, there are still several schemes are available such as *Thompson Scheme*, *Stony–Brook University Scheme* and *NSSL 1–moment 6–class Scheme* and the more information can be found in the reference [98].

Cumulus parameterization. These kinds of parameterization schemes are used to resolve the sub-grid-scale effects of convective and/or shallow clouds. They are adopted only on individual columns where the scheme is triggered and provide vertical heating and moistening profiles. Moreover, some schemes could additionally provide cloud and precipitation field tendencies in the column. Some of schemes and its basic characteristics are summarized in the Table 2.1.

Table 2.1 Available cumulus parameterization options in the WRF model [98]

Scheme	Cloud Detrainment	Type of Scheme	Closure
Kain-Fritsch	Y	Mass flux	CAPE removal
Betts-Miller-Janjic	N	Adjustment	Sounding adjustment
Grell-Devenyi	Y	Mass flux	Various

Surface layer. The role of surface layer schemes is to calculate friction velocities and exchange coefficients which will enable the calculation of surface heat and moisture fluxes by the land-surface models and surface stress in the PBL schemes. As same as most of the schemes of microphysics, it also cannot provide tendencies but the stability-dependent information on the surface layer for the land-surface and PBL schemes. That is, these schemes may be tied to particular boundary-layer options in the current WRF model. *MM5 Similarity Schemes*, *Eta Similarity Schemes* and *MYNN Scheme* are all available.

Land-surface model. The purpose of the land-surface models is to provide heat and moisture fluxes over land points and sea-ice points. In order to realize that, those land-surface models need to take atmospheric information from the surface layer scheme, precipitation forcing from the microphysics and convective schemes, radiative forcing from the radiation scheme. In the current version of WRF model, it includes *5-layer thermal diffusion* [39], *Noah LSM* [112] and *Rapid Update Cycle (RUC) Model LSM* [7].

Planetary Boundary Layer (PBL). The PBL is responsible for vertical sub-grid-scale fluxes due to eddy transports not just in the boundary layer but in the whole atmospheric column. It determines the flux profiles within the well-mixed BL and the SBL, and thus provide atmospheric tendencies of temperature, horizontal momentum and moisture in the whole atmospheric column. Similarly, there also many options (e.g., *MRF, YSU, MYJ, QNSE or MYNN*) can be chosen for a particular prediction, however, it should be balanced with other physics schemes. In fact, the choice is vital to do the wind prediction at low level (e.g. at hub-height) and that is the reason we intend to do some sensitivity tests before running long period forecasts. The specific information will discussed in the section 2.3.1.3.

Atmospheric Radiation. The atmospheric radiation schemes are in charge of providing atmospheric heating because of radiative flux divergence and surface downward longwave and shortwave radiation for the ground heat budget. The longwave radiation contains infrared or thermal radiation absorbed and emitted by gases and surfaces, while shortwave radiation includes visible and surrounding wavelengths that make up the solar spectrum. For longwave radiation, the upward radiative flux from the ground is determined by the surface emissivity, which in turn depends on land-use types or the skin temperature. The upward radiative flux for shortwave radiation is the reflection of surface albedo. All of the radiation schemes in current WRF model are column (one-dimensional) scheme. Several of them are listed in the Table 2.2 with the basic features of the radiation schemes.

Table 2.2 Available Radiation Options in the WRF model [98]

Scheme	Longwave or Shortwave	Spectral Bands	CO2, O3, clouds
RRTM	LW	16	CO2, O3, clouds
GFDL LW	LW	14	CO2, O3, clouds
GFDL SW	SW	12	CO2, O3, clouds
MM5 SW	SW	1	clouds
Goddard	SW	11	CO2, O3, clouds

2.2.4 WRF post-processing

In general, the format of outputs of WRF model is NETCDF. Many post-processing utilities in current model are supported. One of them is NCL (NCAR Command Language), which is a free interpreted language designed specifically for scientific data processing and visualization. It has robust file input and output and has ability of reading the file data in the format of

NETCDF, HDF4, HDF4-EOS, GRIB, binary and ASCII data. RIP4 (Read/Interpolate/Plot 4) is a fortran program that invokes NCAR Graphics routines for the purpose of visualizing output from gridded meteorological datasets. It was originally designed for sigma-coordinate-level output from meteorological models, for example WRF model. Another package usually used is the ARWpost which can read WRF model data and then create outputs to either GrADS or Vis5D format.

At the end of this section, a simplified WRF flow chart is displayed in the Figure 2.6.

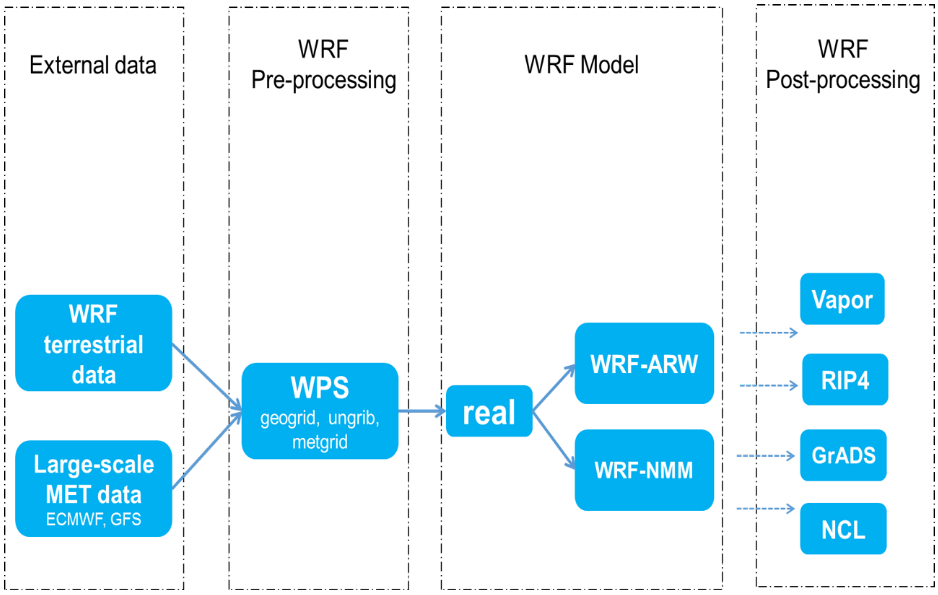

Figure 2.6 A simplified work flow of running WRF modeling system.

2.3 WRF forecasts for Awaji wind farm

2.3.1 Configuration of the WRF model

The meteorological model adopted in this work is the Advanced Research WRF (ARW) model version 3.6 (WRFv3.6 hereafter), which is based on a fully compressible and non-hydrostatic dynamic core [115]. The WRFv3.6 is a limited-area mesoscale model, with a terrain-following hydrostatic-pressure vertical coordinate, designed for operational forecasting as well as research. Its ability of forecasting hub-height wind at wind farms of interest in Japan should be validate firstly for the sake of building an integrated and developed forecasting system for operational use.

2.3.1.1 Domain configuration

In this chapter, the domain configuration of WRFv3.6 which follows the steps recommended by Warner et al [117, 116] includes a parent domain (D01) and three nested domains (D02, D03 and D04) (Figure 2.7) with one-way interaction. The D01 is centered at 34.65 °N and 134.635 °E with a 75×73 mesh of 48 *km* resolution. The horizontal resolution of D02, D03 and D04 are 12 *km* (97×97 grid points), 3 *km* (101×109 grid points) and 1 *km* (103×109 grid points) respectively. The model top is located at 50 hPa and there are 35 vertical stretched eta levels, 10 of which are within the lowest 1 *km*. Initial and boundary conditions are all given by the GFS dataset and no data assimilation or grid nudging was used in this part. The geographical data for the land use and topography are obtained from the U.S. Geological Survey datasets and have resolutions of 5 arc minutes for the parent domain, 2 arc minutes for D02 and 30 arc seconds (about 925 m × 925 m) for the nested D03 and D04. It should be noted that the aforementioned model configuration allows the system to run operationally on workstations for routine real-case use.

The main physical options adopted include the WRF Single-Moment 6-class (WSM6) microphysics parameterization [51], the Rapid Radiative Transfer Model (RRTM) scheme [79] for long-wave radiation with Dudhia's scheme [38] for shortwave radiation, the Kain-Fritsch convective parameterization [55] and the Noah land surface model (LSM) [18]. The specific information of the planetary boundary layer (PBL) scheme we selected is discussed in section 2.3.1.3.

The model prediction period is from 1 August to 31 January 2014 and the numerical results are output at a one-hour interval. we re-initialize WRFv3.6 as a "cold-start" at 18:00 UTC each day and each re-initialization runs for 30 hours. Due to the cold start, there is typically a spin-up time period of 6-hour as recommended by Wang et al.[115] before the model turns to a stable state. Therefore, the forecasts during the initial 6 hours of each run are excluded from the forecasting data series used to compute the performance metrics.

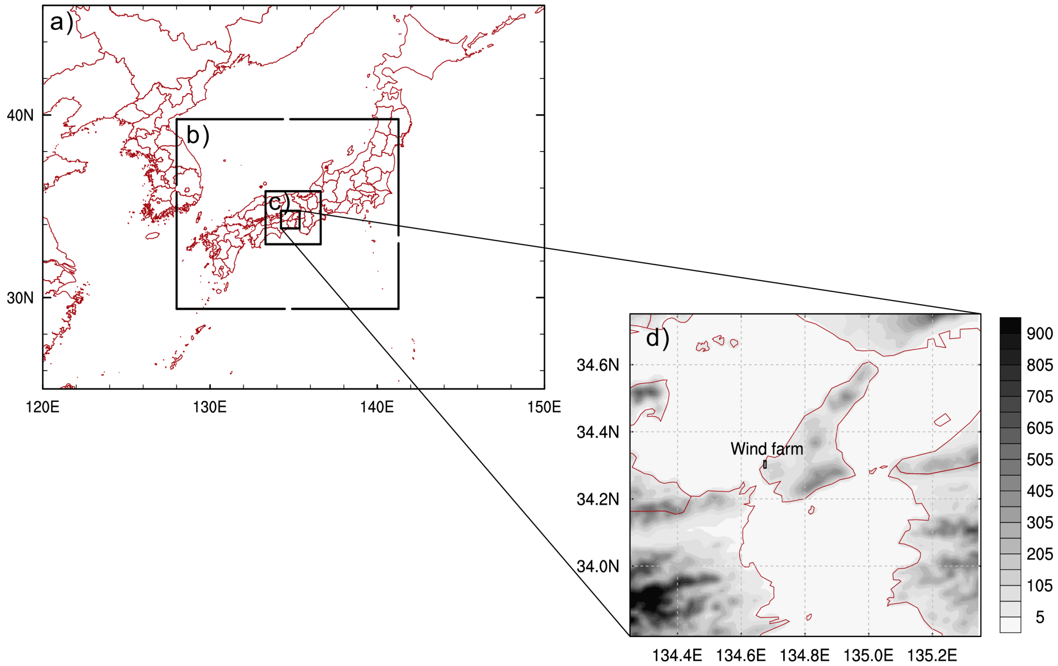

Figure 2.7 WRF domains and model topography. a) D01, b) D02, c) D03 and d) D04 are indicated by black frames. The detailed terrain height (shaded with the gray bar in meter) of the D04 is shown in the panel d).

2.3.1.2 Evaluation metrics

To evaluate the performance of the integrated forecasting system we build quantitatively, the following set of statistical metrics is used.

Mean error (ME):

$$ME = \frac{1}{N}\sum_{i=1}^{N}(fore_i - obs_i) \tag{2.15}$$

where i is the time point and N is the total number of verification time points. $fore$ and obs represent the predicted and observed values, respectively.

Root mean square error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(fore_i - obs_i)^2}{N}} \tag{2.16}$$

Index of agreement (IA):

$$IA = 1 - \frac{\sum_{i=1}^{N} (fore_i - obs_i)^2}{\sum_{i=1}^{N} (\mid fore_i - \overline{obs} \mid + \mid obs_i - \overline{obs} \mid)^2} \tag{2.17}$$

where $\overline{obs}$ denotes the mean of observations. The value of IA which indicates the agreement between the observations and forecasts, ranges from 0 to 1. A larger IA value means better agreement.

Pearson product-moment correlation coefficient (CC):

$$CC = \frac{\sum_{i=1}^{N} (fore_i - \overline{fore})(obs_i - \overline{obs})}{\left[\sum_{i=1}^{N} (fore_i - \overline{fore})^2 \sum_{i=1}^{N} (obs_i - \overline{obs})^2\right]^{1/2}} \tag{2.18}$$

where $\overline{fore}$ indicates the average of forecasts.

2.3.1.3 Sensitivity test of the PBL schemes

The PBL is the lowest part of the atmosphere in which turbulent motions dominate the atmospheric flow. In atmospheric models, the turbulent effects are taken into account by PBL parameterizations. The current WRFv3.6 model has 12 PBL schemes that might exhibit different performances even for the same simulation region [94, 37]. Therefore, prior to applying the WRFv3.6 model to the target wind farm, it is worthwhile to examine the prediction skills of different PBL schemes for the low-level wind field. To this end, the sensitivity of five PBL schemes, i.e., the Quasi-normal Scale Elimination (QNSE) [107], the Asymmetric Convective Model version 2 (ACM2) [86], the Mellor-Yamada-Janjic (MYJ) [74], the Mellor-Yamada-Nakanishi-Niino (MYNN) [53] and the Yonsei University Scheme (YSU) [52], are tested over 15-day (from 1 October to 15 October 2013) for predictions of wind speed. The setup of the numerical experiments for this inter-comparison of PBL schemes is summarized in Table 2.3.

Figure 2.8 (a) illustrates the predicted diurnal variations of the predicted wind speed of No.3 turbine at 80 m from the five sets of experiments with different PBL parameterization schemes, referred to as QNSE, ACM2, MYJ, MYNN and YSU, and the corresponding observations. All the experimental runs capture the wind speed variations well and the sensitivity of PBL schemes is more apparent in period 09:00-23:00 UTC than in period 00:00-08:00 UTC. The forecasts from all experiments overestimate the wind speed during the whole period. The largest bias (0.94 ms^{-1}, shown in the third row of Table 2.4) at No.3 turbine site is observed in the QNSE experiment, while the smallest bias (0.63 ms^{-1}) is seen in the ACM2 prediction. The RMSEs also indicate that the performances of the ACM2, MYJ, MYNN and YSU schemes are better than that of the QNSE scheme, and ACM2 has the smallest RMSE of 0.87 ms^{-1}.

Table 2.3 The five sets of numerical experiments for the inter-comparison of PBL schemes.

Experiment	PBL scheme	Land surface model	Surface-layer scheme
QNSE	Quasi-normal Scale Elimination	Unified Noah LSM	QNSE
ACM2	Asymmetric Convective Model	Pleim-Xu	Pleim-Xu
MYJ	Mellor-Yamada-Janjic	Unified Noah LSM	Eta similarity
MYNN	Mellor-Yamada-Nakanishi-Niino	Unified Noah LSM	MYNN
YSU	Yonsei University Scheme	Unified Noah LSM	Monin Obukhov

Table 2.4 Error statistics of diurnal variation between WRFv3.6 forecasts and observations of wind speed at hub-height for different experiments. The verification metrics are computed over the 1-15 October 2013 period at No.3, No.7 and No.14 wind turbine sites.

Experiment	QNSE		ACM2		MYJ		MYNN		YSU	
Turbine \ Errors	ME	RMSE	ME	RMSE	ME	RMSE	ME	RMSE	ME	RMSE
No.3	0.94	1.16	**0.63**	**0.87**	0.75	0.99	0.90	1.13	0.88	1.09
No.7	0.58	0.98	**0.20**	**0.79**	0.39	0.88	0.53	0.96	0.55	0.93
No.14	1.42	1.68	**1.08**	**1.30**	1.21	1.49	1.33	1.62	1.38	1.60

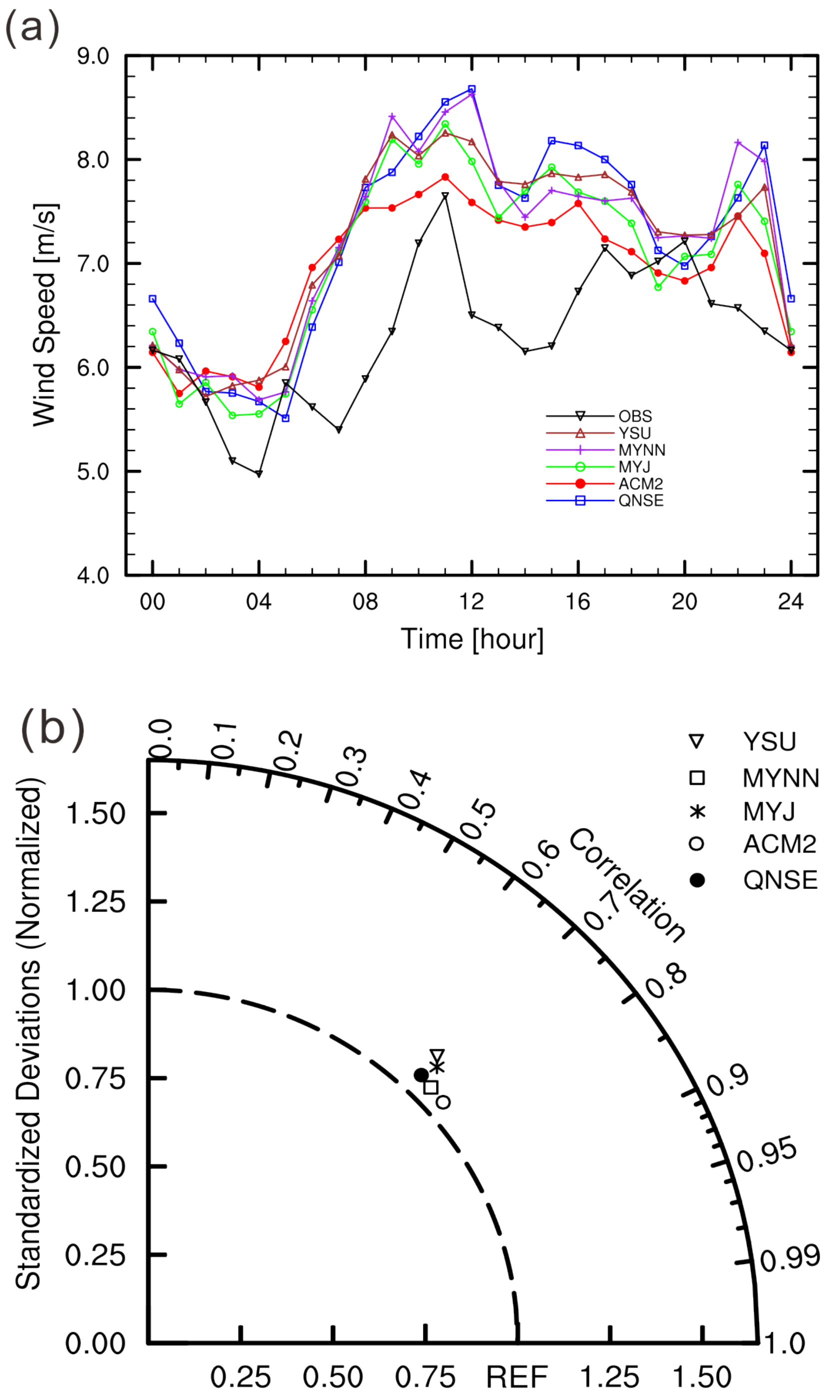

Figure 2.8 Comparisons of observations and forecasts of five sensitivity experiments shown in Table I at No.3 turbine site. (a) The diurnal variation of 15-day (1-15 October 2013) averaged wind speed at hub-height. The black line depicts the series corresponding to the observations, whereas the colored lines correspond to the forecasts with different experimental setups. (b) Taylor diagram shows normalized standard derivation and correlation of wind speed at hub-height for five experiments referred to observation ("REF"). The number of samples is 360 (one hour interval from 1 October to 15 October 2013).

The correlation and normalized standard deviation (NSD) of each experiment are calculated and summarized in a Taylor diagram [111], which provides a synthetically visual comparison in terms of centered RMSE, correlation and NSD. From Figure 2.8 (b), although the NSD difference among five experiments is probably not significant, advantage of ACM2 is still observed. Moreover, regarding both NSD and correlation, the ACM2 prediction is the one closest to the observations. Thus, the ACM2 scheme is chosen as the optimum PBL scheme for the prediction of wind speed at the wind farm site of interest. We further substantiated this conclusion by examining the values of ME and RMSE for other two turbines, i.e., No.7 and No.14 in Table 2.4, which show that the ACM2 scheme gives the best forecast for the local wind field of the target area.

2.3.2 Validation of wind forecasts

Although the performance of WRFv3.6 does not appear highly sensitive to the PBL schemes tested, the results of the experiments above partly demonstrate that the configuration including the ACM2 PBL scheme with Pleim-Xu land surface model and Pleim-Xu surface-layer scheme works best for the wind speed prediction at the wind farm site analyzed in this study. With this configuration of the WRFv3.6 model, a six-month time series (i.e., from 00:00 UTC 2 August 2013 to 23:00 UTC 31 January 2014) of the low-level wind speed prediction were generated following the procedure described in section 2.3.1. In this section, No.3 turbine is firstly chosen as an example to show that the WRFv3.6 model is able to predict the wind speed at hub-height for the target area with reasonably good accuracy. The conclusion is then confirmed by the consistent results obtained from other 14 turbines.

2.3.2.1 wind forecasts of the No.3 and No.7 turbine

Figure 2.9 shows the comparison between predicted (black), observed (red) and bias (green) of wind speed at No.3 turbine site during the period from 00:00 UTC 2 August 2013 to 23:00 UTC 31 January 2014. The green dotted curve which lies close to the zero reference line reveals that the predicted raw wind speed reproduces the observation with good accuracy. Although there are occasional large errors, in general the predicted wind speed coincides

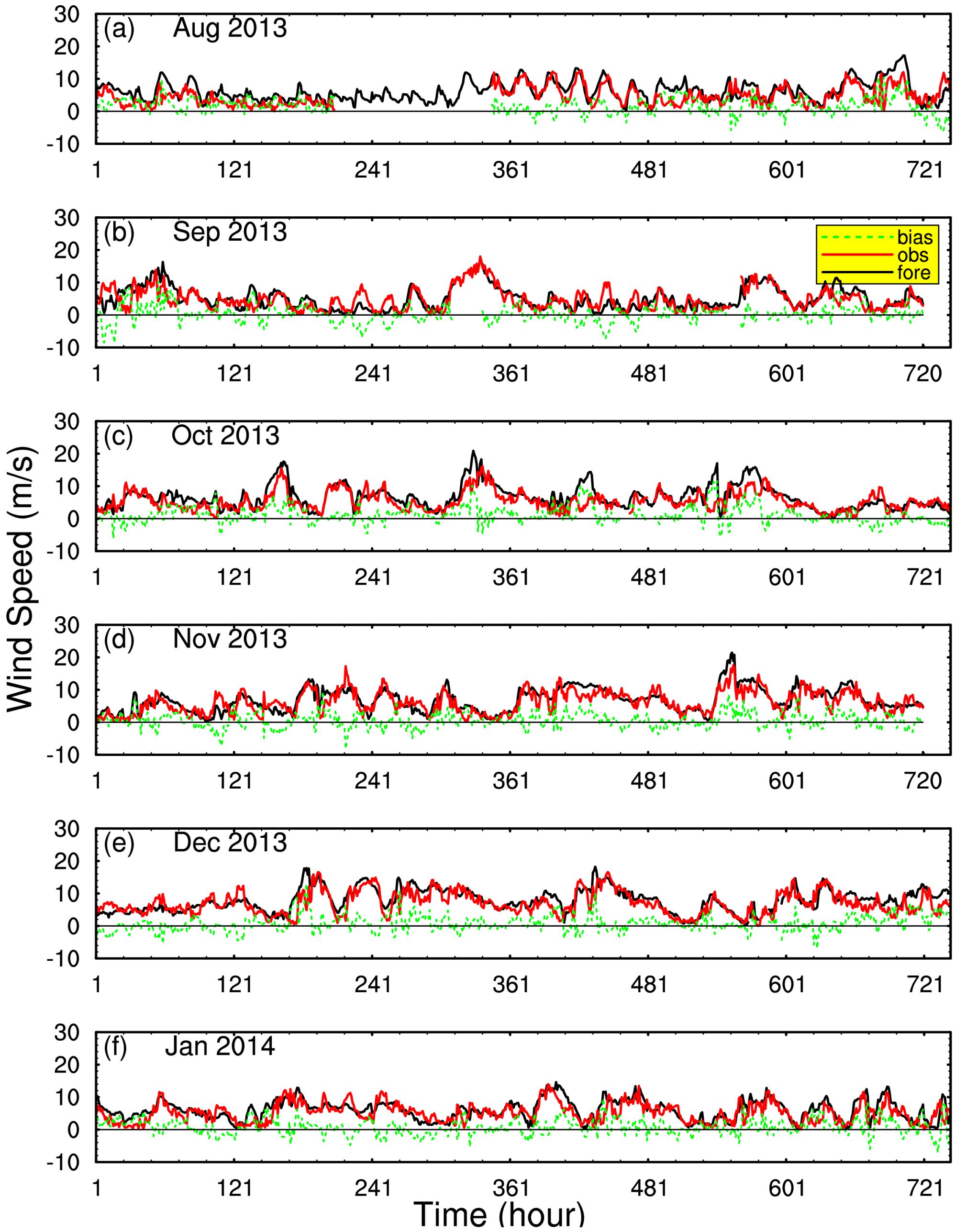

Figure 2.9 Six-month series of the predicted raw (black), observed (red) wind speed (ms^{-1}) at hub-height and bias (green) of No.3 turbine. Panels from (a) to (f) stand for August, September, October, November, December 2013 and January 2014, respectively.

well with the observation through all six months. Similar conclusions characterized by small value of ME ($\leq$1.23) and RMSE ($\leq$2.84), as well as relatively large value of CC ($\geq$0.62), can be drawn from the statistics listed in Table 2.5 (column 5). Those results show that WRF based forecasting system has a relatively high ability of forecasting hub-height wind of No.3 turbine. In addition, the results displayed in the Figure 2.10 indicates the success of No.3 turbine is not by chance due to very similar conclusions can be obtained.

2.3.2.2 Overall results for 15 turbines

Table 2.5 also exhibits error statistics for other 14 turbines to ensure that the reasonable prediction for No.3 turbine is not a success by chance. This table shows that only 3 out of 90 CC are smaller than 0.60, which indicates that the trend of predictions is in good accordance with observations. For most of turbines, the ME varies from -1.45 ms^{-1} to 2.00 ms^{-1}, and the smallest ME (0.03 ms^{-1}) is found in December for No.1 turbine. All predictions overestimate the wind speed. Except for a few large values (bold in Table 2.5), the RMSE retains a relative small value and does not change much through the six months for all turbines. All these results substantiate that the WRFv3.6 model has reasonably good forecasting skill in predicting low-level wind speed for the Awaji-island wind farm. However, the relatively large variation in MEs and RMSEs still shows the possibility to further improve the prediction of wind speed by using the Kalman filter as a post-processing approach or data assimilation to provide better initial conditions, which will be discussed in the next two chapters. Before doing that, the wind power forecasts need to be explored since the final interests of wind companies is not wind speed forecasts but the near future power output.

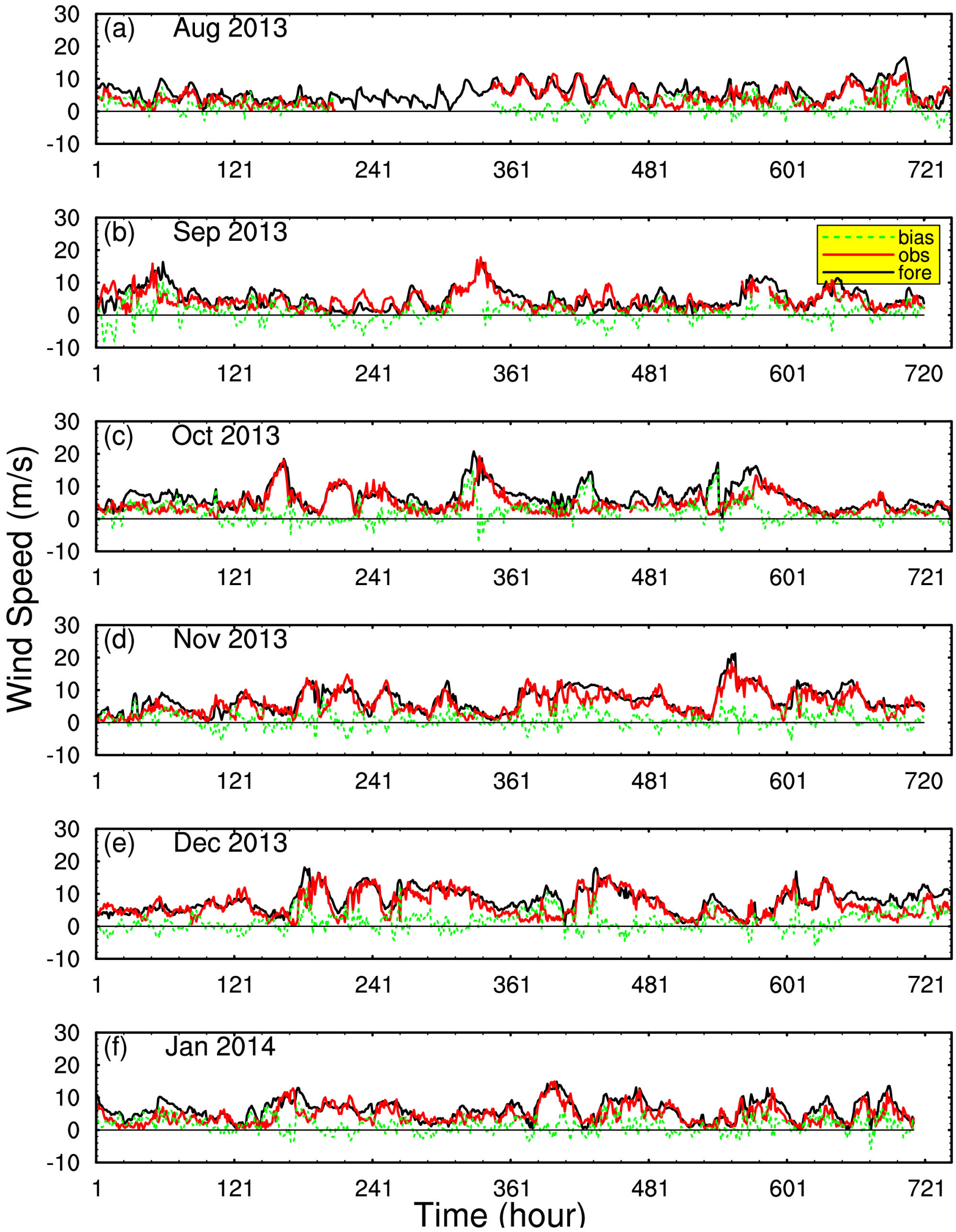

Figure 2.10 Same as Figure 2.9, but for No.14 turbine.

Table 2.5 The monthly ME, RMSE and CC calculated with the predictions and corresponding observations of wind speed at hub-height of 15 turbines. The values of ME≥2.00, RMSE≥3.50 and CC<0.60 are bold.

Month	Statistics	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10	No.11	No.12	No.13	No.14	No.15
August	ME	0.89	1.39	0.51	1.06	1.05	-1.45	1.00	0.94	1.89	1.07	1.15	1.24	1.28	1.31	0.54
	RMSE	2.43	2.51	2.35	2.37	2.76	3.20	2.78	2.83	3.06	2.90	2.54	2.84	2.45	2.60	2.53
	CC	0.70	0.62	0.71	0.68	0.68	**0.54**	0.66	0.68	0.64	0.67	0.61	0.65	0.67	0.70	0.73
September	ME	0.68	0.82	0.25	0.66	0.24	1.19	0.46	0.80	1.05	1.11	1.04	0.72	0.70	0.76	0.49
	RMSE	2.70	2.76	2.63	2.91	2.90	2.72	2.93	2.70	2.77	2.84	2.68	2.71	2.70	2.77	2.48
	CC	0.71	0.72	0.73	**0.58**	0.68	0.67	0.68	0.70	0.69	0.72	0.73	0.73	0.70	0.68	**0.52**
October	ME	1.10	0.91	1.23	1.43	0.93	1.71	1.37	1.68	1.97	**2.36**	**2.17**	1.35	1.34	**2.01**	1.38
	RMSE	2.83	2.91	2.84	3.14	2.83	3.12	3.11	3.33	2.77	**3.72**	**3.90**	**3.45**	3.02	**3.59**	3.28
	CC	0.72	0.71	0.74	0.67	0.74	0.68	0.67	0.65	0.74	0.63	0.60	0.75	0.72	0.63	0.66
November	ME	0.54	0.68	0.62	0.79	1.01	1.27	0.35	0.99	1.14	1.28	1.12	0.92	0.89	1.04	0.45
	RMSE	2.75	2.67	2.74	2.42	2.64	2.65	2.22	2.51	2.52	2.60	2.60	2.34	2.54	2.47	2.41
	CC	0.70	0.74	0.68	0.78	0.75	0.78	0.81	0.78	0.79	0.77	0.85	0.80	0.78	0.80	0.78
December	ME	0.03	0.47	0.87	0.79	1.28	1.78	0.77	1.05	1.42	1.24	1.16	1.06	1.15	1.45	0.88
	RMSE	2.59	2.59	2.67	2.51	2.65	**3.54**	2.57	2.86	3.05	3.02	2.98	2.80	2.79	3.17	3.03
	CC	0.72	0.73	0.72	0.75	0.75	0.66	0.75	0.70	0.68	0.70	0.70	0.70	0.72	0.71	0.74
January	ME	0.34	0.61	0.84	0.80	0.89	1.63	1.23	1.14	1.37	1.60	1.47	1.05	1.14	1.53	0.71
	RMSE	2.49	2.45	2.58	2.34	2.44	2.77	2.58	2.57	2.75	2.74	2.74	2.54	2.58	2.82	2.52
	CC	0.72	0.71	0.66	0.74	0.70	0.70	0.69	0.72	0.69	0.70	0.72	0.76	0.70	0.70	0.71

2.4 Power forecasts

2.4.1 A power-curve model

Recall the fluid mechanical definition of the power output:

$$P = \frac{1}{2} C_p \rho A v^3$$

where C_p is the turbine coefficient of performance, and ρ is the air density which depends on air pressure and temperature. A is the swept area of a turbine blade and v represents wind speed over the wind farm site. The theoretical power curve of turbine No.3, which is a 2.5 MW horizontal-axis wind turbine (HAWT) with three blades, is shown in Figure 2.11 (black line and points). In the Figure 2.11, the blue points represent the observed electrical power output obtained for turbine No.3 measured from routine operation. It is easily found that although the overall trend shows an agreement with the theoretical manufacture power curve, there are remarkable uncertainties and deviations from it. This is not surprising considering that there are still other factors that affect the power output, such as the unresolvable sub-scale fluctuations, wake effects, wind direction, as well as operational control [49]. Consequently, the manufacturer power curve cannot be directly used to predict wind power in real cases, and the approach followed in this study is to build the empirical power model for each turbine based on 4-month (from 1 August 2013 to 30 November 2013 with a 10-min interval) observed wind speed and power data using a polynomial fit technique. The empirical power curve is expressed by

$$P = a_{10} \cdot v^{10} + a_9 \cdot v^9 + \cdots + a_1 \cdot v + a_0 \tag{2.19}$$

where v is the wind speed, P the prediction of wind power, and $a_0, a_1, \cdots, a_{10}$ are the coefficients separately generated for different turbines. Table 2.6 displays the values of coefficients of power curve models for the No.3 wind turbine and the corresponding power curve is plotted in Figure 2.11 (red circles), which looks noticeably different from the manufacturer power curve. It is also seen that there is an uncertainty between the observed wind speed and power output, which might be attributed to other factors, such as variations

Table 2.6 The coefficients of power curve models for the No.3 wind turbine.

Turbine	a_{10}	a_9	a_8	a_7	a_6	a_5	a_4	a_3	a_2	a_1	a_0
No.3	2.23E-06	-0.0002	0.0061	-0.1125	1.2089	-7.8032	29.3869	-55.9710	48.3195	-16.7577	2.0053

of wind direction, air temperature, as well as the effects of mechanical and operation control systems. Our main interest in this thesis is to generate a power output prediction system as a function of wind speed.

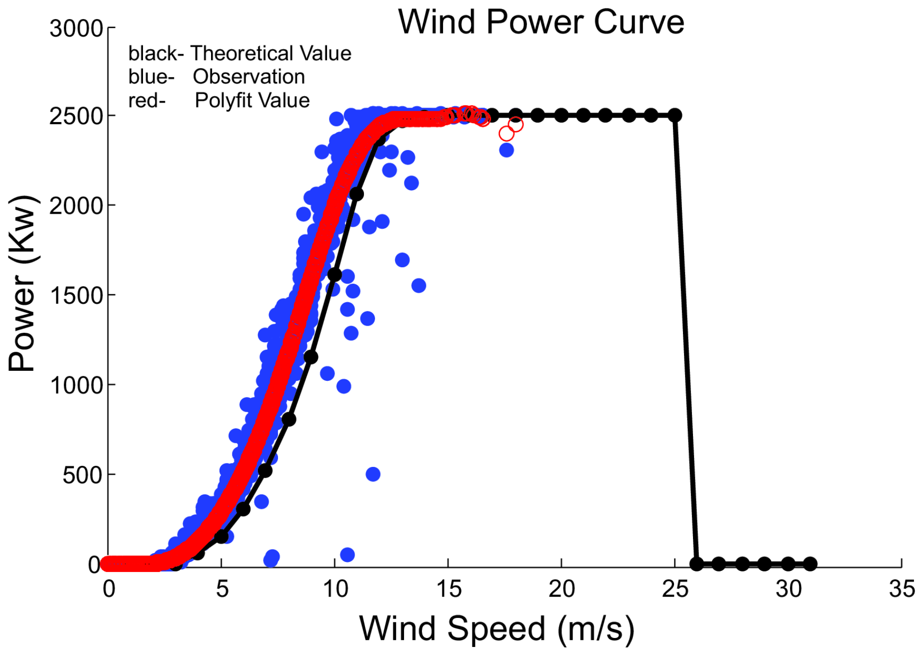

Figure 2.11 The theoretical (black line and points), observed (blue point) and the tenth-order polynomial (red point) wind power curve of No.3 turbine.

2.4.2 Validation of the No.3 turbine

Figure 2.12 displays the comparison of predicted and observed power for No.3 turbine during November , December 2013 and January 2014, respectively. It is clear that the forecasting power reproduces the observation with a relatively good accuracy. Similar to the wind speed forecasts, although there are occasional large errors, the predicted wind power generally coincides well with the observed power output through all three months. This conclusion can be further demonstrated with the relatively small value of ME (121.60 kW) and RMSE

(644.34 kW), as well as large value of CC (72.43%), can be drawn from the statistics shown in Table 2.7 (row 5). All of those results indicate that the power curve we build using the polynomial fit technique and historical data has a relative high forecasting skill for a single wind turbine.

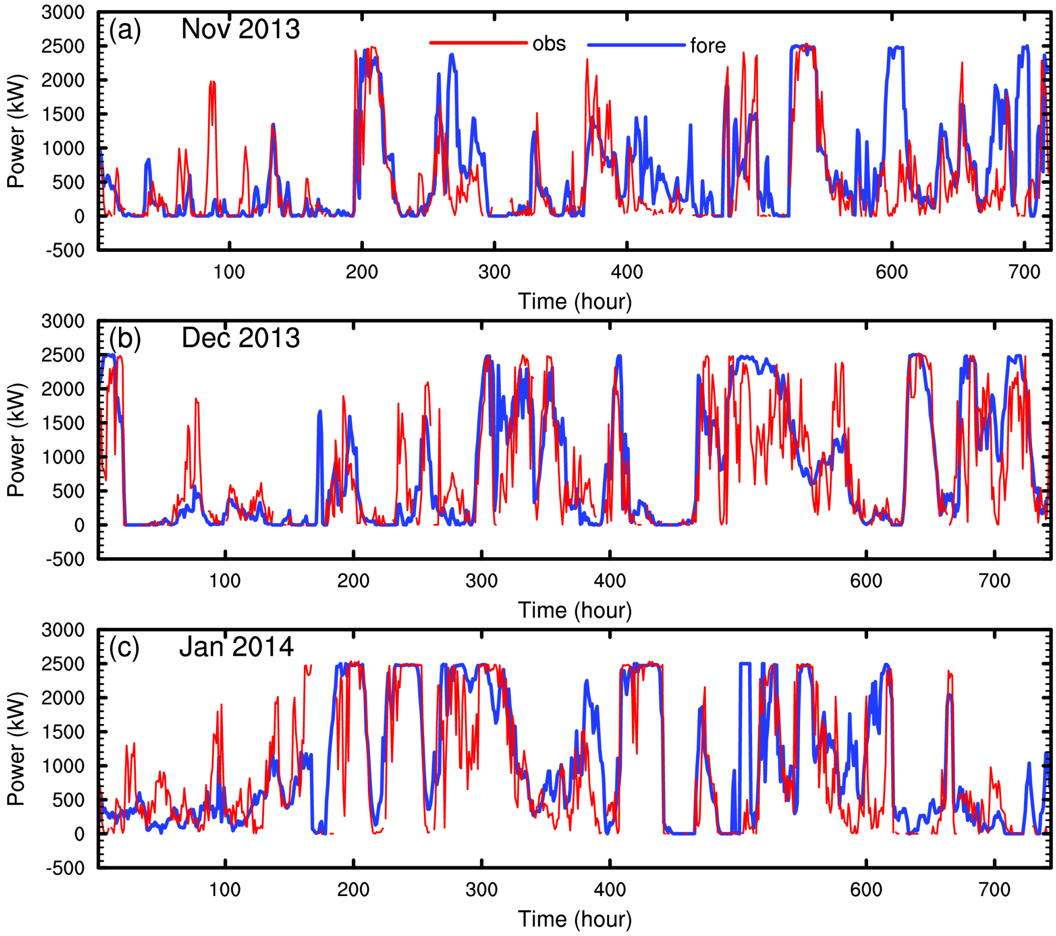

Figure 2.12 Three-month series of the predicted (blue) and observed (red) power (kW) of No.3 turbine. Panels from (a) to (c) stand for November, December 2013 and January 2014, respectively.

2.4.3 Overall results for 15 turbines

In addition, Figure 2.13 also shows a 10-day series of predicted and observed power for other two wind turbines (No.7 and No.14) except for the No.3 turbine, in order to validate whether the procedure of constructing power model is suitable for other turbines where the

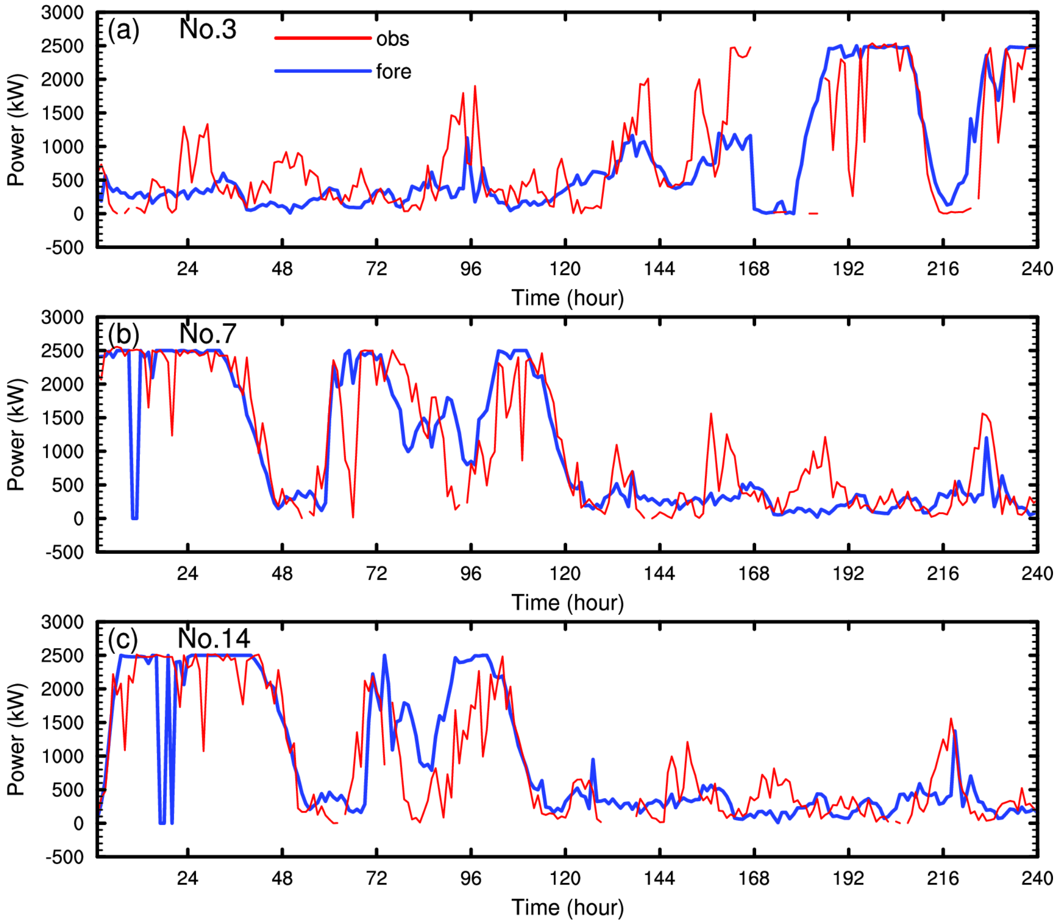

Figure 2.13 A comparison of the raw wind power forecast (blue) against the observed power output (red) for No.3 (a), No.7 (b) and No.14 (c) turbine over the same time period, which is from 00:00 UTC 01 to 23:00 UTC 10 January 2014.

terrain features vary largely. Fortunately, the results from the Figure 2.13 indicate that the predicted raw wind power generally coincides well with the observations at all three turbine sites. Moreover, Table 2.7exhibits error statistics for all 15 turbines over a period of 3-month. From this table, we can find that all predictions overestimate the wind power due to the values of ME are all positive and the values of RMSE are smaller than 737.643 kW. The values of CC are around 70% which indicates that the trend of predictions is in good accordance with observations. All these results substantiate that the power curve model has a reasonably good forecasting skill in predicting power output for a single turbine, as well as over the whole wind farm in Japan.

Table 2.7 The ME, RMSE and CC of power prediction for total 15 turbines over 3-month period.

Turbine	ME (kW)	RMSE (kW)	CC (%)
No.1	70.42	684.82	70.22
No.2	108.36	688.87	70.55
No.3	121.60	644.34	72.43
No.4	143.33	656.70	74.13
No.5	150.27	639.31	69.47
No.6	235.11	687.05	72.85
No.7	88.86	623.33	74.99
No.8	193.48	684.77	71.19
No.9	212.75	668.41	70.11
No.10	258.17	661.89	73.55
No.11	289.83	737.63	69.26
No.12	158.49	561.91	71.32
No.13	243.46	684.67	68.23
No.14	252.83	677.32	73.71
No.15	253.49	676.36	72.29

2.5 Summary

In this chapter, we have established a preliminary forecasting system for wind power prediction, based on the meso-scale meteorological model WRFv3.6 and a power curve. The system has been validated for the targeted wind farm in Awaji-island, Japan, which is characterized by complex topographic features.

The global-scale GFS dataset is adopted as both initial and boundary conditions for the regional-scale and high resolution WRFv3.6 model through a 4-level nesting refining the horizontal grid resolution down to 1 $km \times 1$ km for the target region. The model has been tuned, and the ACM2 PBL and the corresponding parameterization schemes were chosen for predicting the wind speed at hub height (80 m above ground) in the wind farm site. Compared to the observed wind speed of 15 turbines in the target wind farm, from 1 August 2013 to

31 January 2014, the WRFv3.6 model shows good performance in forecasting the surface wind field. The power curve model used in this thesis is constructed using polynomial fit technique. Its forecasting skills for 15 turbines of the Awaji wind farm have been validated and the results indicate that given reliable wind prediction the power curve model constructed from the historic wind speed and power data provides reasonable projection for wind power.

Therefore, combining the WRF model and the power curve model together, a preliminary forecasting system has been built. In fact, the system has been installed by a wind energy company to provide operational prediction twice each day. However, there is no doubt that many errors and uncertainties exist in this system, which could affect the forecasting skill of the system, for example the consistently overestimating the power in this chapter. These problems lead to the possibility to further improve the prediction of wind speed and power using Kalman filter and data assimilation which will be described in the next two chapters.

Chapter 3

Kalman filter module

As mentioned at the end of last chapter, there are many errors or uncertainties exist in both WRF model and power forecasts. Specifically, to a large extent the wind forecasts derived from NWP model (WRF) are affected by errors stemming from uncertainties in initial/boundary conditions, simplifications in physics and numerical approximations [2]. Great efforts have been devoted to reduce these uncertainties by improving data quality [36] and developing more accurate numerical models with improved dynamic cores and more sophisticated physical parameterizations [70, 73].

Although efforts to improve NWP models have led to substantial progress in the accuracy of deterministic predictions, it cannot be expected to eliminate all uncertainties in real-case applications, which results in deviations between the NWP output and the real atmospheric state. An effective way to reduce the uncertainties of the NWP models is by implementing post-processing methods to revise or correct the NWP model outputs based on their past performances. Typical and widely-used post-processing methods include Model Output Statistics (MOS) [48] and Kalman Filter [56, 45, 32]. The Kalman filter is a popular algorithm due to the simplicity of the algorithm, the moderate computational costs, and the short training period required. It has been applied successfully for wind energy modeling to produce more accurate predictions. The works by Louka et al [69] and Al-Hamadi et al [1] clearly demonstrated that the forecasting errors of both wind power and electric load can be effectively reduced with the Kalman filter. They have shown that combining NWP models and statistical post-processing into a tuned prediction system, can further improve wind speed and power forecasts. Unfortunately, to the best knowledge of the authors, there is no report in literature on any practice to establish such a prediction system for wind farm sites in Japan. In this chapter, the Kalman algorithm used throughout this thesis is presented firstly. Then its impact of improving both wind and power is investigated.

3.1 The Kalman filter

The Kalman filter is an estimation algorithm named after Rudolf E Kálmán, which operates recursively on streams of input data (containing random variations) to produce a statistically optimal estimate of the underlying system state. It is over 50 years old but is still one of the most important and common data fusion algorithms in use nowadays. The great success of the Kalman filter is due to its adaptive, recursive, optimal characteristics and small computational requirement. Therefore, it is widely used in various fields from radar and computer vision to meteorological purposes. The specific set of mathematical equations can be found in [56, 118]. In this study, we focus on estimating and removing the bias of WRF model and power model using the Kalman filter method presented in [32].

The implementation of the Kalman filter can be divided into two main steps: one is "time update", aiming to project forward the bias of the current state to estimate the forecasting bias at the next time step; the other part is a "measurement update", namely incorporating a new observation into the previous estimation to obtain a corrected estimate of the forecasting bias.

In general, the forecasting bias between the forecasts and measurements of a variable at time t is related to the state at previous time $t-\delta t$:

$$x_{t|t-\delta t} = x_{t-\delta t|t-2\delta t} + \eta_{t-\delta t} \tag{3.1}$$

where x_t is the true forecasting bias at time t, δt is a time lag, $x_{t|t-\delta t}$ is the priori state estimate at time t, η is the white noise that has zero-mean, and the variance (σ_η^2) is uncorrected in time. Although the real forecasting bias is unknown, it has certain relationships with the forecasting errors (also called measurement bias) y_t. That is, the forecasting errors equal the forecasting bias plus a random error ε_t:

$$y_t = x_t + \varepsilon_t = x_{t|t-\delta t} + \eta_t + \varepsilon_t \tag{3.2}$$

where ε_t is normally distributed with zero-mean and variance σ_ε^2. The source of random errors ε_t comes mainly from uncertainty or errors in numerical models, as well as inaccuracy in initial and boundary conditions.

The Kalman filter gives the recursive estimation of the unknown forecasting bias x_t based on the bias estimation at previous time and the historical forecasting errors y:

$$\hat{x}_{t+\delta t|t} = \hat{x}_{t|t-\delta t} + K_{t|t-\delta t}(y_t - \hat{x}_{t|t-\delta t}) \tag{3.3}$$

where the hat ($\hat{}$) notation indicates the estimation of the variable. $K_{t|t-\delta t}$ is the Kalman gain, which is recursively calculated as follows:

$$K_{t|t-\delta t} = \frac{p_{t-\delta t|t-2\delta t} + \sigma_\eta^2}{p_{t-\delta t|t-2\delta t} + \sigma_\eta^2 + \sigma_\varepsilon^2} \tag{3.4}$$

where p is the expected mean square error:

$$p_{t|t-\delta t} = (p_{t-\delta t|t-2\delta t} + \sigma_\eta^2)(1 - K_{t|t-\delta t}) \tag{3.5}$$

Given a reasonable initial guess of p_0 and K_0, as well as the model forecast M_t and observation time series, the Kalman Filter can recursively generate an estimate of forecast bias x at $t + \delta t$ through equations 3.3-3.5. Then, the model forecast can be corrected as follows:

$$M_{kf\,t+\delta t} = M_{t+\delta t} - \hat{x}_{t+\delta t|t}. \tag{3.6}$$

It is worthwhile to note that the calculation of white noise σ_η^2 and σ_ε^2 is crucial to the implementation of Kalman filter procedure, though their priori are not usually known. We first define a new variable z_t as following,

$$z_t = y_{t+\delta t} - y_t = \varepsilon_{t+\delta t} - \varepsilon_t + \eta_t \tag{3.7}$$

which has variance [35],

$$\sigma_z^2 = 2\sigma_\varepsilon^2 + \sigma_\eta^2 \tag{3.8}$$

Assuming the estimation of σ_η^2 is derived from the estimation of σ_ε^2 with a ratio r:

$$\sigma_\eta^2 = r\sigma_\varepsilon^2 \tag{3.9}$$

The ratio r is a parameter reflecting the relative weighting of observation and forecasts. As r is somewhat sensitive to different models and predicted variables, several tests have to be carried out to find the best values of r for in specific situation. Thus, the equation 3.8 can be rewritten as,

$$\sigma_z^2 = (2+r)\sigma_\varepsilon^2 \tag{3.10}$$

According to [32, 34], σ_ε^2 is a time-varying quantity which can be calculated with the Kalman algorithm itself (using equations 3.3-3.5). Specifically, σ_ε^2 is firstly estimated by applying equation 3.5:

$$p_{t|t-\delta t}^{\sigma_\varepsilon^2} = (p_{t-\delta t|t-2\delta t}^{\sigma_\varepsilon^2} + \sigma_{\sigma_\eta^2}^2)(1 - K_{t|t-\delta t}^{\sigma_\varepsilon^2}) \tag{3.11}$$

where, $p^{\sigma_\varepsilon^2}$ is the expected mean square error of the σ_ε^2, $\sigma_{\sigma_\eta^2}^2$ is the variance in the σ_η^2 estimate and $K^{\sigma_\varepsilon^2}$ is the Kalman gain for estimating σ_ε^2. Similarly to equation 3.4, this Kalman gain can be written as,

$$K_{t|t-\delta t}^{\sigma_\varepsilon^2} = \frac{p_{t-\delta t|t-2\delta t}^{\sigma_\varepsilon^2} + \sigma_{\sigma_\eta^2}^2}{p_{t-\delta t|t-2\delta t}^{\sigma_\varepsilon^2} + \sigma_{\sigma_\eta^2}^2 + \sigma_{\sigma_\varepsilon^2}^2} \tag{3.12}$$

where $\sigma^2_{\sigma^2_\varepsilon}$ is the variance of σ^2_ε. Constant values of 1 and 0.0005 are assigned to $\sigma^2_{\sigma^2_\varepsilon}$ and $\sigma^2_{\sigma^2_\eta}$ respectively. Considering the equation 3.3, 3.10 and 3.12, the σ^2_ε can be estimated as follows:

$$\sigma^2_{\varepsilon,t+\delta t} = \sigma^2_{\varepsilon,t-\delta t} + K^{\sigma^2_\varepsilon}_{t|t-\delta t}\left[\frac{(y_t - y_{t-\delta t})^2}{2+r} - \sigma^2_{\varepsilon,t-\delta t}\right] \tag{3.13}$$

7-day running mean method

Stensrud and Skindlov [103] showed that a simple bias correction method using the previous 7-day mean bias correction can improve the direct model forecasts of maximum temperature. This method is easy to implement meanwhile has ability of improving the raw predictions. Therefore, it is chosen as a reference to validate the performance of Kalman filter algorithm for correcting the raw prediction of wind speed in this thesis.

3.2 Results of using Kalman filter

The Kalman filter procedure is applied independently to every prediction lead time. For instance, WRFv3.6 raw predictions at 00:00 UTC are revised by the Kalman filter that is updated by using the predictions and observations at the same time on the previous days. The first 60 days (August and September) are chosen as a training period for implementing the Kalman filter. The following discussions are all based on the statistic metrics computed over the 4-month prediction period (from October 2013 to January 2014).

The following two subsections present and discuss the improvement of the Kalman filter predictions for both wind speed and power in terms of the error quantifications, such as ME, RMSE and CC. Additionally, the results of the 7-day running mean (7-day hereafter) method are also included for comparison.

3.2.1 wind

A comparison of the 7-day method and the Kalman filter to correct the WRF prediction is depicted in Figure 3.1. It presents hourly model raw forecasts (black line) and observed wind speed at 80-m of No.3 turbine (red line) as well as the corrected predictions using the Kalman filter (blue line) and 7-day method (green line) for the 10-day period, from 00:00 UTC 14 to 23:00 UTC 23 October 2013. Again, the WRFv3.6 model demonstrates the capability of predicting the local wind speed. Moreover, both the Kalman filter and 7-day method are able to significantly improve the raw prediction of the WRF model, particularly the systematic bias has been largely reduced.

When comparing the correction results of Kalman filter and the 7-day method, we see the remarkable advantage of the Kalman filter in reducing the forecasting errors. This advantage is further illustrated by the statistic parameters in Table 3.1. Compared to the 7-day method, the Kalman filter shows much smaller RMSE and larger CC. This may be due in part to the fact that the current Kalman filter can not only correct the systematical error but also part of stochastic uncertainties, while the 7-day method has an effect barely on the systematic bias.

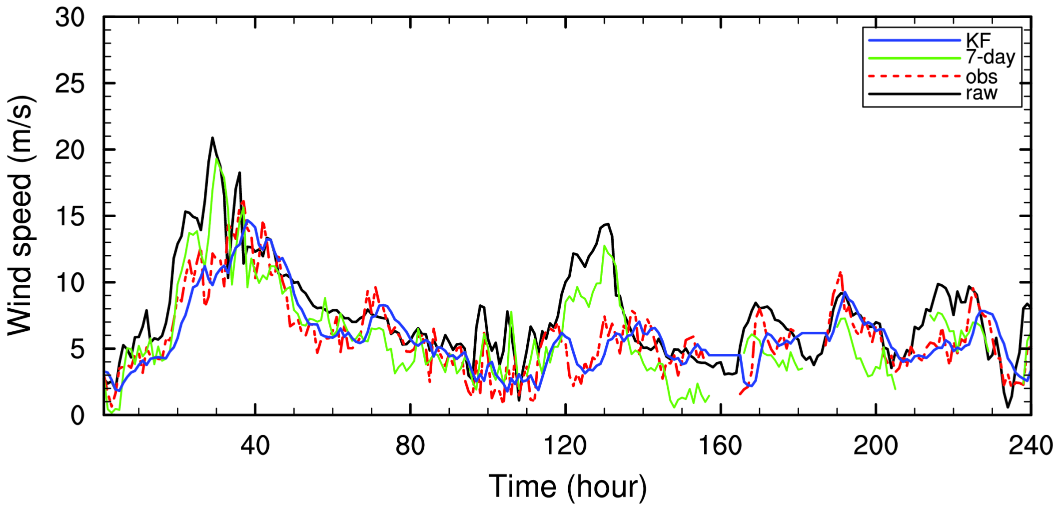

Figure 3.1 Hourly WRFv3.6 model raw forecasts (black) and corresponding observations (red) of wind speed at hub-height of No.3 turbine for the 10-day period from 00:00 UTC 14 October to 23:00 UTC 23 October 2013. The blue and green line present the predictions corrected by the Kalman filter and 7-day method respectively.

The ME, RMSE and CC of the Kalman filter and 7-day method predictions with respect to the raw WRFv3.6 prediction are shown in Figure 3.2 for the total 15 turbines in the

Awaji-island wind farm. From Figure 3.2a, it can be seen that most of the ME of raw forecast range from 0.28 ms^{-1} to 1.54 ms^{-1}. Although the ME looks different for each turbine, both the Kalman filter and 7-day method can largely alleviate the systematic error tendency in the raw forecast of the wind speed. The reduction of ME from the Kalman filter correction ranges from 92% to 99%, while that of the 7-day method ranges from 74% to 96%. The comparison of RMSE is given in Figure 3.2b. As expected, the Kalman filter reduces the 15-turbine mean RMSE by 22% which is much more significant compared to that of 7-day method (4%). Consistent to RMSE, the CC displayed in Figure 3.2c further demonstrates that the Kalman filter algorithm is superior than the 7-day method in improving wind speed forecast .

From the validations discussed above, we may conclude that 1) the raw forecasts of WRFv3.6 model with a tuned PBL package is able to produce reasonably good prediction for the wind speed at the hub height in Awaji-island wind farm site which is characterized by complex topography, and 2) The Kalman filter, as a better post-processing method against 7-day method, can significantly improve the forecasting skill for the surface wind at the target turbine site considered in this study.

Table 3.1 Averaged ME, RMSE and CC of No.3 turbine over a 10-day period of 00:00 UTC 14-23:00 UTC 23 October, 2013.

Statistic quantity	ME			RMSE			CC(%)		
Predicted variable \ Approach	Raw	KF	7-day	Raw	KF	7-day	Raw	KF	7-day
wind speed (ms^{-1})	1.79	-0.02	0.05	2.97	1.58	2.62	69.21	85.42	71.11
wind power (kW)	371.19	-62.43	–	762.94	443.27	–	64.26	75.53	–

3.2.2 wind power

In this section, the power curve model described in section 2.4.1 is firstly tested with the dataset of No.3 turbine over a 10-day period from 00:00 UTC 14 to 23:00 UTC 23, 2013.

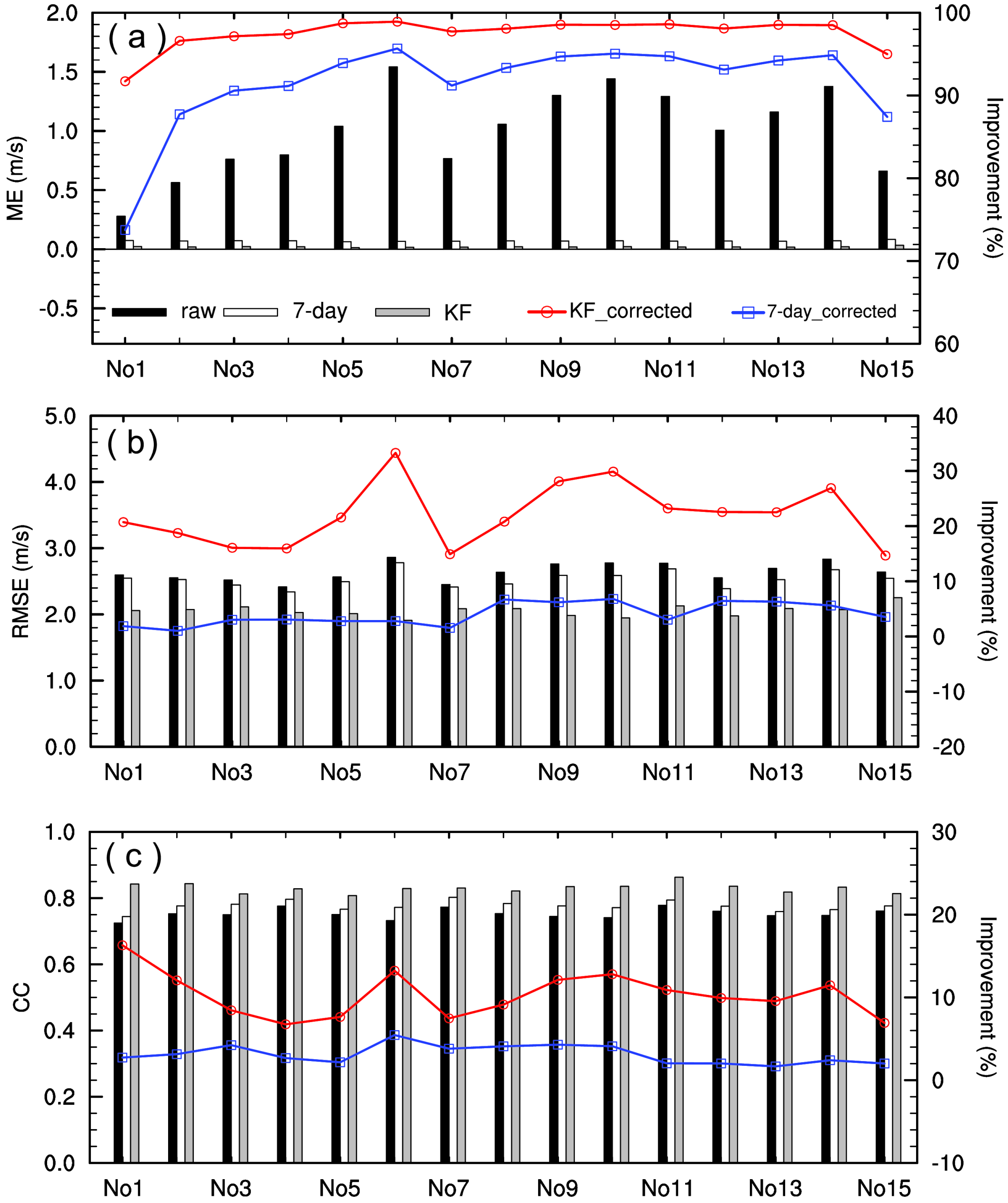

Figure 3.2 A comparison of the ME (a), RMSE (b) and CC (c) for wind speed at hub-height of the Kalman filter (solid gray bar) and 7-day method (solid white bar) predictions with respect to the raw WRFv3.6 prediction (solid black bar) for the 15 turbines of the Awaji-island wind farm. The marked lines stand for the relative improvement of the Kalman filter (red) and 7-day method (blue) against the raw forecasts of WRFv3.6 model.

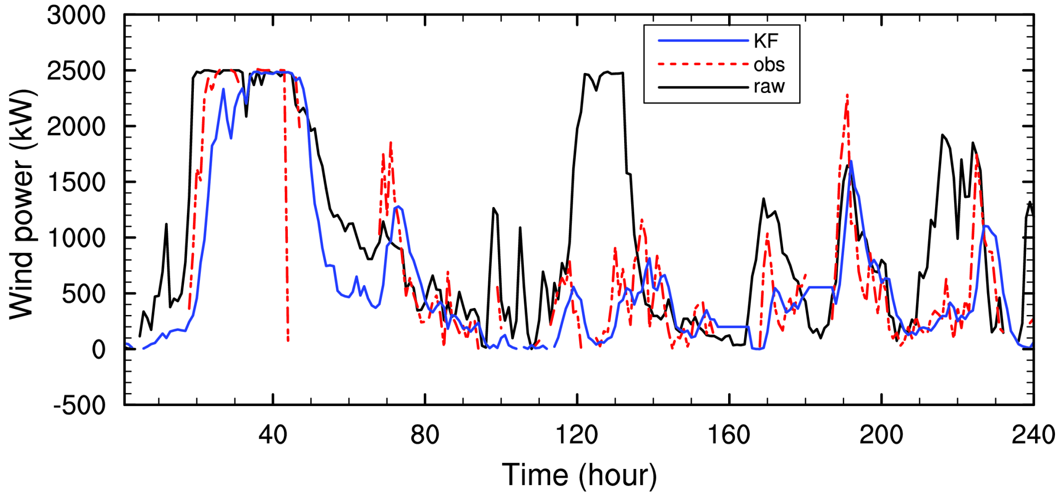

Figure 3.3 A comparison of the raw wind power forecast (black) and the Kalman filter corrected prediction (blue) against the observed power output (red) for No.3 turbine over the period from 00:00 UTC 14 to 23:00 UTC 23 October 2013.

Then the Kalman filtered wind speed is used as an input of the power curve model to investigate whether improvement can be seen in the power output by using the corrected wind speed. The results are concluded in Table 3.1 and Figure 3.3.

Figure 3.3 shows a 10-day example for comparing the performances of raw forecast and the Kalman filter prediction of wind power. The Kalman filter significantly improves the power output prediction. The systematic overestimation of the raw wind power forecasts are consistently reduced during the whole period. From Table 3.1, it can be seen that the power output from the raw wind forecast is overestimated with an ME of 371.19 kW, which has been effectively reduced down to -62.43 kW by using the Kalman filtered wind speed. Other two statistic quantities, RMSE and CC, show consistent results revealing that the Kalman filtered wind field effectively improves the power prediction. These results also indicate that given reliable wind prediction the power curve model constructed from the historic data provides reasonable projection for wind power.

Furthermore, three-month datasets (November, December 2013 and January 2014) of the total 15 turbines were used to validate the performance of Kalman filter in predicting the power output for the whole farm site. Figure 3.4 shows ME, RMSE and CC of the power predictions with both raw wind speed and the Kalman filter corrected wind speed as the input of the power curve model for each turbine. It is found that the Kalman filter predictions (Figure 3.4a) improved the power output predictions for all turbines. We also show the relative improvement for each case (the right vertical axis in percentage), which reveals that the improvement brought by the Kalman filter to each turbine is different from one another. The relative improvement in ME varies from 44% to 97% with an averaged of 83%. As shown in Figure 3.4b, the RMSE of the power forecasts for all 15 turbines are largely reduced by Kalman filter, with No.13 being the best, having the value of relative improvement over 36%. Regarding to the CC parameter displayed in Figure 3.4c, the Kalman filtered wind speed leads to an averaged improvement of 15% for all 15 turbines.

From all results shown above, it can be concluded that the accuracy of power output prediction can be significantly improved when the Kalman filtered wind speed is used as the input of power curve model equation (2.19).

In the power predictions discussed above, the Kalman filter is implemented to reduce the systematical and random errors in the wind prediction of the WRFv3.6 model, which exhibit significant improvement in power prediction. However, uncertainties still remain in the power curve model as mentioned before. It motivates us to implement the Kalman filter further to reduce the uncertainties in the power curve. To this end, we carried out four experiments to evaluate the impact of Kalman filter for both wind speed prediction and power curve model as follows.

- **Baseline**: We use it as a controlled case, where the raw forecasts of wind speed from the WRF model is directly used to calculate the power output using (2.19). The Kalman filter is not used to correct either wind speed or power curve model (2.19).

- **KF-speed**: The Kalman filter is used to correct the wind speed from the WRF model, and the corrected wind speed is used to calculate the wind power by using (2.19).

- **KF-power**: The raw forecast of wind speed of the WRF model is used as the input for the power curve model (2.19) and the Kalman filter is only applied to the power output.

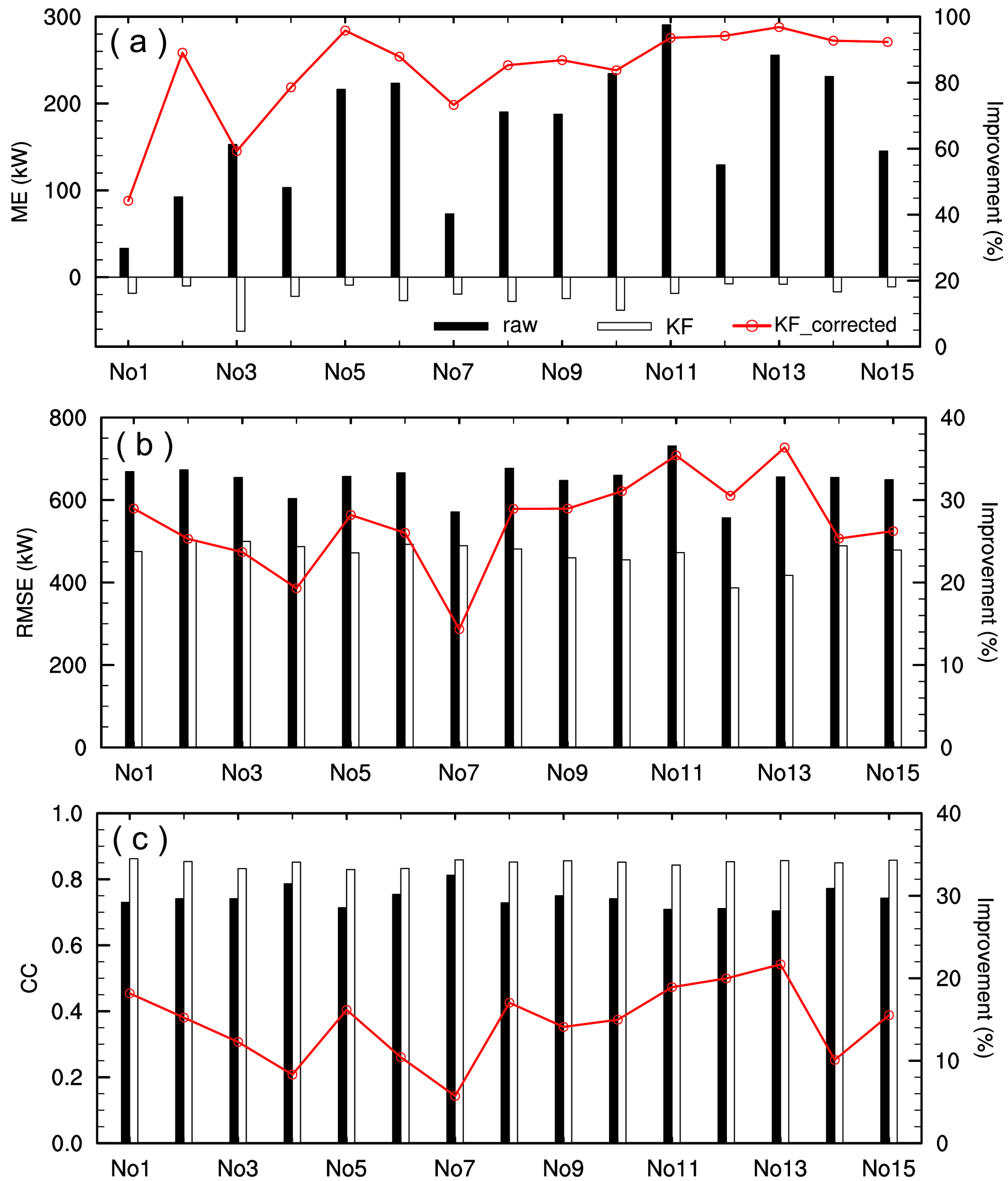

Figure 3.4 A comparison of the ME (a), RMSE (b) and CC (c) of the Kalman filter (solid white bar) and the raw power forecasts (solid black bar) for the 15 turbines of the Awaji-island wind farm. The red line stands for the relative improvement of the Kalman filter against the raw forecasts of the power curve model.

- **KF-speed & power**: The Kalman filter is applied to the predictions of both wind speed and the power curve model.

The results of the four cases are shown in Table 3.2. The larger positive value of ME for controlled case (Baseline) indicates the overestimation of wind power. Having implemented the Kalman filter, the ME is largely reduced, especially for cases KF-speed and KF-speed & power showing 85% and 92% reductions in bias respectively. Furthermore, case KF-speed & power has got the most significant reduction in RMSE and improvement in CC. Compared with the baseline case, the KF-power case demonstrates that the Kalman filter indeed makes significant positive impact on error correction of wind power curve model. This conclusion is further validated by the differences between the case KF-speed and KF-speed & power presented in Table 3.2.

Table 3.2 The ME, RMSE and CC of inter-comparison among four experiments with different configurations of implementing the Kalman filter. Shown are the average results for total 15 turbines.

Case	Baseline	KF-power	Improvement (%)	KF-speed	Improvement (%)	KF-speed & power	Improvement (%)
ME(kW)	142.79	-65.52	54%	-21.42	85%	-11.27	92%
RMSE (kW)	648.43	500.64	23%	470.57	27%	431.32	33%
CC(%)	74.24	76.37	3%	84.94	14%	85.46	15%

3.3 Summary

In this chapter, a hybrid forecasting system of wind power generation has been developed by integrating the Kalman filter module with the high resolution Weather Research and Forecasting (WRF) model as well as an empirical formula of wind power output (power curve). The system has been validated with observations including wind speed and power output over a six-month period for 15 turbine sites at a wind farm in Awaji-island, Japan. The results show that the tuned WRF model is able to provide hub-height wind speed prediction for the target area with reliability to some extent. The predicted wind field can be substantially improved by the Kalman filter as a post-processing procedure.

The Kalman filter presented in this thesis is a linear and adaptive algorithm which can minimize both the systematical and random errors by recursively combining direct model outputs with the most updated observations. It demonstrates the ability to improve the ME, RMSE and CC in both wind speed and power predictions based on the WRFv3.6 NWP model and the empirical power curve model. As shown in Table 3.2, Kalman filter significantly improves the raw model prediction of power by 92%, 33% and 15% in ME, RMSE and CC respectively. Compared with other post-processing methods, such as MOS and 7-day method, Kalman filter is able to provide more reliable prediction with a short training period, and is more flexible to adapt to any target prediction with available observation and forecasting model. However, as other statistical correction methods, Kalman filter has limited ability to predict the sudden changes in forecasting error [33]. It should be also noted that extra initial tests are always needed to successfully implement the Kalman filter presented in this study, since the correction effect depends on the ratio r ($\sigma^2_{\eta,t}/\sigma^2_{\varepsilon,t}$) which is somewhat sensitive to different models and predicted variables.

In spite of these, the wind power forecasting system presented in the chapter can be expected as an effective tool for short-term operational control for both single turbine and whole wind farm in a target site. Having validated the system as a hybrid wind power forecasting system of practical significance for the Awaji-island wind farm site, we are planning to adopt it to other wind farm sites in Japan.

The present research has shown the promising performance of the proposed integrated model for wind power prediction under complex topographic conditions which feature almost all land-based wind farms in Japan. It also indicates some new directions worthy of further investigations, for example, a more precise initial condition for WRF model to obtain more reasonable input for power curve model; a computational fluid dynamic model with finer grid resolution coupled with the WRF model to directly resolve the topographic effects on the surface wind field; and more reliable power curve models that include more factors and are thus able to remove the uncertainties due to the processes not reflected in the current model.

Chapter 4

Data assimilation module

As mentioned earlier, the data assimilation is chosen as another module of our integrated forecasting system. It is worth to note that the observations are vital to the data assimilation for the purpose of improving the forecasting skills of NWP models. For operational wind farms, the conventional observations include upwind meteorological (MET) tower measurement and nacelle wind data. In general, the measurements from an MET tower at a wind farm site cannot accurately reflect the real wind field around the turbines which are located at different locations away from the MET tower, particularly when the terrain of a wind farm is complex. Instead, the nacelle-mounted anemometer which is placed on the top of nacelle behind the rotor can provide the routine data of wind speed and direction for each turbine. Although the wind observation of the nacelle-mounted anemometer is always affected by the design/shape of the wind turbine and nacelle, as well as the operation condition of the turbine [120, 99], some studies still show that nacelle-based wind speed observation, after proper calibration and data quality control, is more representative to the wind behavior (e.g., wind disturbance) experienced by the wind turbines in a wind farm than that from an upwind MET tower [100, 29]. Thus, the nacelle-based wind data will be chosen for the data assimilation module to further improve our system mentioned in the last chapter.

Before implementing the data assimilation procedure, several basic theories including some basic elements in probability theory, the ingredients of data assimilation, mathematical description and classification of data assimilation will be firstly stated. Then the impact of assimilating nacelle wind data on the wind forecasts derived from WRF model will be checked. Lastly, the role of Kalman filter and data assimilation is compared.

4.1 Data assimilation and GSI system

Data assimilation is the process of combining all available information from models, background and observations, in order to obtain an optimal estimate of the state of a physical system. It can be used for data reanalysis (using past observations), calibration and validation, observing system design, better understanding of model errors, parameters and physical process interactions, and for providing more accurate initial conditions for physical models. In the subsequent section, we firstly state the mathematical description of data assimilation and then several classical data assimilation algorithms are briefly introduced.

4.1.1 Basic elements in probability theory

4.1.1.1 Random experiment

A random experiment is a process for which the outcome cannot be predicted with certainty. In general, it is described as follows [22],

- define a set Ω as all possible outcomes from an experiment and a subset of Ω is called an event;
- the corresponding probability function, P: a numerical expression of a state of knowledge. Assuming any two disjoint events A and B, P has following properties,

$$
\begin{aligned}
&0 \le P \le 1 \\
&P(A \cup B) = P(A) + P(B) \\
&P(\Omega) = 1
\end{aligned}
$$

If A and B are not independent, knowing the state of knowledge of A has occurred changes on B, known as conditional probability:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

- the outcomes of a random experiment are called random variables.

4.1.1.2 Probability density function

The probability density function is commonly referred as pdf (p), which is defined as:

$$P(a < X \leq b) = \int_a^b p(x)dx$$

where X is a set and x is a random variable ($x \in X$).

4.1.1.3 Joint and conditional pdf

Assuming x and y are two random variables, their joint pdf would be $p(x,y)$:

$$p(x,y) = p(x|y)p(y) \tag{4.1}$$

4.1.1.4 Expectation and variance

Generally, a pdf is rarely known completely, but some properties such as expectation and variance, the definitions of which are written as follows:

$$E(x) = \int_{-\infty}^{+\infty} xp(x)dx$$
$$Var(x) = E([x - E(x)]^2) = \int_{-\infty}^{+\infty} [x - E(x)]^2 p(x)dx$$

The standard deviation is the square root of the $Var(x)$.

When the random variable is a vector (x), its expectation also is a vector and the covariance matrix is defined by

$$E([\mathrm{x}-E(\mathrm{x})][\mathrm{x}-E(\mathrm{x})]^T)$$

4.1.1.5 Gaussian distribution

The Gaussian (normal) distribution is a very common continuous probability distribution and often used in natural sciences. A Gaussian random variable x is characterized by:

- Mean: μ
- Covariance: σ^2

We usually write the Gaussian random variable as $x \sim N(\mu, \sigma^2)$. If $\sigma^2 > 0$, its pdf is given by:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

This distribution has some nice properties, in particular:

- It is a natural distribution for many signal noises;
- If there are two independent variables $x_1 \sim N(\mu_1, \sigma_1^2)$ and $x_2 \sim N(\mu_2, \sigma_2^2)$, then the combination $x_1 + x_2$ is also a Gaussian random variable and $(x_1 + x_2) \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$;
- In addition, if m and n are real numbers, then $mx+n$ also obey the Gaussian distribution, i.e., $mx+n \sim N(m\mu+n, m^2\sigma^2)$.

When a Gaussian variable (x) is a vector with n size, its expectation and covariance are usually noted as μ and P which is a $n \times n$ matrix, therefor the pdf is given by:

$$p(\mathrm{x}) = \frac{1}{(2\pi)^{n/2}|\mathrm{P}|^{1/2}}\exp\left[-\frac{1}{2}(\mathrm{x}-\mu)^T\mathrm{P}^{-1}(\mathrm{x}-\mu)\right] \tag{4.2}$$

where $|\mathrm{P}|$ stands for the determinant of P. It should be noted that any linear combinations of Gaussian vectors are still Gaussian. For example, given $\mathrm{x} \sim N(\mu, \mathrm{P})$, then $\mathrm{Lx} \sim N(\mathrm{L}\mu, \mathrm{LPL}^{\mathrm{T}})$ where L is a linear transformation matrix.

4.1.1.6 Bayes' and marginalization rules

Those two are fundamental rules in estimation theory. Bayes' rule:

$$p(\mathrm{x}|\mathrm{y}) = \frac{1}{p(\mathrm{y})} p(\mathrm{y}|\mathrm{x}) p(\mathrm{x}) \tag{4.3}$$

where the x and y are all random variables. The marginalization rule is shown as below:

$$p(\mathrm{y}) = \int p(\mathrm{x}, \mathrm{y}) dx = \int p(\mathrm{y}|\mathrm{x}) p(\mathrm{x}) dx$$

In the Bayes' formula, the $p(\mathrm{x}), p(\mathrm{y}|\mathrm{x})$ and $p(\mathrm{x}|\mathrm{y})$ are always called **Prior**, **Likelihood** and **Posterior**, respectively.

4.1.2 Ingredients of data assimilation

Generally, the concept of data assimilation is based on three main ingredients, including a dynamical system (stochastic or deterministic) based on some prior information, a sequence of observations and a numerical model.

4.1.2.1 Prior information

We assume throughout this section that $\mathbf{x}$ is the true state. Usually, the prior knowledge of $\mathbf{x}$ is available under the form of the prior pdf ($p(\mathbf{x})$). If this $\mathbf{x}$ obeys Gaussian distribution, its expectation could be noted as $\mathbf{x}^b$, which is also called the *background* state. Thus the background error e^b and corresponding covariance matrix B can be defined as:

$$e^b = \mathbf{x}^b - \mathbf{x} \tag{4.4}$$

$$\mathrm{B} = e^b (e^b)^T \tag{4.5}$$

4.1.2.2 Observation

The true field $\mathbf{x}$ results in a signal $\mathbf{y}$ in the observation space. This causality relation involves a function (possibly nonlinear) H known as *observation operator*, as well as an additive noise:

$$\mathbf{y} = H(\mathbf{x}) + \eta \tag{4.6}$$

where η is the observation error and often is assumed as a Gaussian variable in practice. In the data assimilation procedure, the measurement of pdf (or likelihood) $p(\mathbf{y} \mid \mathbf{x})$ is needed. In fact, this can be simply obtained with a translation of $p(\eta)$ by $H(\mathbf{x})$.

4.1.2.3 Numerical model

Here are considered the dynamical models, that is , the models that compute the time evolution of the simulated state. Assuming our model is M and considering the Markov chain $\mathbf{x} = \{\mathbf{x}_t\}_{t \in Z^+}$,

$$\mathbf{x}_{t+1} = M(\mathbf{x}_t) + \xi_t \tag{4.7}$$

where ξ_t refers to model error which accounts for the errors in the numerical model (e.g., misrepresentation of physical processes) and for the errors due to the discretization.

4.1.3 General formulation of data assimilation

Together equation 4.7 and 4.6 provides a probabilistic model for the jointly varying random variable $(\mathbf{x}, \mathbf{y})$. The aim of data assimilation is to find out information on the signal $\mathbf{x}$, given data of observation $\mathbf{y}$. There are two key problems of data assimilation, namely smoothing and filtering. In the subsequent section those two problems of which the key concept is Bayes' formula are briefly described respectively.

4.1.3.1 Smoothing problem

For this problem, an instance of observe data $\mathbf{y}$ is needed, in order to determine information about the signal $\mathbf{x}$. To be more precise we wish to obtain the probability measure describing the signal $\mathbf{x}$ on a time interval $T_0 = (0, 1, \cdots, T)$, given observation on the interval of $T_1 = (1, \cdots, T)$, denoted as $\mathbf{x}|\mathbf{y}$. Here T_0 is often called data assimilation window. In the next contexts, we define $\mathbf{x} = \{x_t\}_{t\in T_0}$, $\xi = \{\xi_t\}_{t\in T_0}$, $\mathbf{y} = \{y_t\}_{t\in T_1}$ and $\eta = \{\eta_t\}_{t\in T_1}$ and assume that ξ and η are two independent Gaussian random variable. The Bayes' theory (4.3) is used to find the $p(\mathbf{x}|\mathbf{y})$ in what follows.

Prior ($p(\mathbf{x})$. Using the condition pdf (4.1) and the structure of ξ in turn,

$$\begin{aligned}
p(\mathbf{x}) &= p(x_T, x_{T-1}, \cdots, x_0) \\
&= p(x_T|x_{T-1}, x_{T-1}, \cdots, x_0)p(x_{T-1}, \cdots, x_0) \\
&= p(x_T|x_{T-1})p(x_{T-1}, \cdots, x_0) \\
&\cdots \\
&= \prod_{t=0}^{T-1} p(x_{t+1}|x_t)p(x_0)
\end{aligned} \tag{4.8}$$

where

$$\begin{aligned}
x_0 &\sim N(\mu_0, P_0) \\
x_{t+1}|x_t &\sim N\big(M(x_t), P\big)
\end{aligned}$$

therefore using (4.2), we can get :

$$p(x_0) \propto \exp\left[-\frac{1}{2}(x_0-\mu_0)^T P_0^{-1}(x_0-\mu_0)\right] \tag{4.9}$$

$$p(x_{t+1}|x_t) \propto \exp\left[-\frac{1}{2}(x_{t+1}-M(x_t)^T P^{-1}(x_{t+1}-M(x_t)\right] \tag{4.10}$$

Combining (4.9) and (4.10), the prior $p(\mathbf{x})$ (4.8) could be written as:

$$p(\mathbf{x}) \propto \exp\big(-J_1(\mathbf{x})\big) \tag{4.11}$$

where $J_1(\mathbf{x})$ is defined by:

$$J_1(\mathbf{x}) = \frac{1}{2}(x_0 - \mu_0)^T P_0^{-1}(x_0 - \mu_0) + \sum_{t=0}^{T-1} \frac{1}{2}(x_{t+1} - M(x_t))^T P^{-1}(x_{t+1} - M(x_t)) \tag{4.12}$$

Likelihood ($p(\mathbf{y}|\mathbf{x})$. It is usually a Gaussian probability distribution with pdf proportional to $J_2(\mathbf{x};\mathbf{y})$, which could be similarly obtained as described in the prior part:

$$p(\mathbf{y}|\mathbf{x}) \propto \exp\big(-J_2(\mathbf{x};\mathbf{y})\big) \tag{4.13}$$

where

$$J_2(\mathbf{x};\ \mathbf{y}) = \sum_{t=0}^{T-1} \frac{1}{2}(y_{t+1} - H(x_{t+1}))^T R^{-1}(y_{t+1} - H(x_{t+1})) \tag{4.14}$$

In the above equations, μ_0 and P_0 are usually referred to as the background expectation and background covariance, respectively. P and R are the covariance of model errors and observation errors respectively. The J_2 are named as the model-observation misfit functional.

Posterior ($p(\mathbf{x}|\mathbf{y})$. This is the purpose of data assimilation, which is looking for information about the signal $\mathbf{x}$ given the observed data $\mathbf{y}$. Recalling the Bayes' formula in equation (4.3), the posterior pdf can be defined by:

$$\begin{aligned} p(\mathbf{x}|\mathbf{y}) &= \frac{1}{p(\mathbf{y})} p(\mathbf{y}|\mathbf{x}) p(\mathbf{x}) \\ &\propto p(\mathbf{y}|\mathbf{x}) p(\mathbf{x}) \end{aligned} \tag{4.15}$$

Combining (4.11) and (4.13) together,

$$p(\mathbf{x}|\mathbf{y}) \propto \exp\big(-J(\mathbf{x};\mathbf{y})\big) \tag{4.16}$$

where $J(\mathbf{x};\ \mathbf{y})$is defined by

$$
\begin{aligned}
J = J_1 + J_2 = &\frac{1}{2}(x_0 - \mu_0)^T P_0^{-1}(x_0 - \mu_0) \\
&+ \sum_{t=0}^{T-1} \frac{1}{2}(x_{t+1} - M(x_t))^T P^{-1}(x_{t+1} - M(x_t)) \\
&+ \sum_{t=0}^{T-1} \frac{1}{2}(y_{t+1} - H(x_{t+1}))^T R^{-1}(y_{t+1} - H(x_{t+1}))
\end{aligned}
\tag{4.17}
$$

Here J is often referred to a cost function. Apparently, the smoothing problem can only be performed off-line (e.g., reanalysis process in atmospheric area), because this kind of algorithm is aiming to find information about signal $\mathbf{x}_{t_0}$ based on the observation$\mathbf{y}_{t_1}$, however, where t_0 is often smaller than t_1. On the contrary, the filtering problem can overcome this drawback and can determine $p(\mathbf{x}_{t_1}|\mathbf{y}_{t_1})$ by using Bayes' theory, which will be described in the following section.

4.1.3.2 Filtering problem

A general filtering problem consists of two main steps: prediction and analysis. In order to describe the theory conveniently, let $\mathbf{Y}_t = \{y_j\}_{j=1}^t$, where y_j stands for the observation data at time j. The signal is still referred to $\mathbf{x}$. Thus, the aim of the filtering is to obtain the $p(x_{t+1}|\mathbf{Y}_t + 1)$ from $p(x_t|\mathbf{Y}_t)$. The details of each step are shown as below.

Prediction. Firstly, it is worth to note that $p(x_{t+1}|\mathbf{Y}_t,\ x_t)$ actually equals $p(x_{t+1}|x_t)$ since $\mathbf{Y}_t$ includes information about x_t and cannot be improved based on perfect knowledge of x_t. Using the theory of conditional pdf, we can get

$$
\begin{aligned}
p(x_{t+1}|\mathbf{Y}_t) &= \int p(x_{t+1}|\mathbf{Y}_t,\ x_t)p(x_t|\mathbf{Y}_t)dx_t \\
&= \int p(x_{t+1}|x_t)p(x_t|\mathbf{Y}_t)dx_t
\end{aligned}
\tag{4.18}
$$

In (4.18), the $p(x_{t+1}|x_t)$ is determined by the forward model ($M(x_t)$). Therefore, the prediction step can be regarded as a transition from $p(x_t|\mathbf{Y}_t)$ to $p(x_{t+1}|\mathbf{Y}_t)$.

Analysis. In this step, we want to deduce $p(x_{t+1}|\mathbf{Y}_{t+1})$ from $p(x_{t+1}|\mathbf{Y}_t)$. It is easy to find that $p(x_{t+1}|\mathbf{Y}_{t+1}) = p(x_{t+1}|\mathbf{Y}_t, y_{t+1})$. Then using Bayes' formula, the $p(x_{t+1}|\mathbf{Y}_{t+1})$ can be written as:

$$\begin{aligned} p(x_{t+1}|\mathbf{Y}_{t+1}) &= p(x_{t+1}|\mathbf{Y}_t, y_{t+1}) \\ &= \frac{p(y_{t+1}|x_{t+1}, \mathbf{Y_t})p(x_{t+1}|\mathbf{Y}_t)}{p(y_{t+1}|\mathbf{Y_t})} \end{aligned} \tag{4.19}$$

In the above equation, $p(y_{t+1}|x_{t+1}, \mathbf{Y_t}) = p(y_{t+1}|x_{t+1})$ because of Y_t only contains noisy and indirect information about the signal $\mathbf{x}$ and also cannot be improved upon the knowledge of $\mathbf{x}_{t+1}$, thus (4.19) can be rewritten as:

$$p(x_{t+1}|\mathbf{Y}_{t+1}) = \frac{p(y_{t+1}|x_{t+1})p(x_{t+1}|\mathbf{Y}_t)}{p(y_{t+1}|\mathbf{Y_t})} \tag{4.20}$$

From (4.20), it is clear that the analysis step finally provides a map from $p(x_{t+1}|\mathbf{Y}_t)$ to $p(x_{t+1}|\mathbf{Y}_{t+1})$.

4.1.4 Data assimilation algorithms

Many assimilation algorithms on smoothing problems have been developed. One of them is Kalman smoother, which is aiming to find explicit expressions for the pdf $p(\mathbf{x}|\mathbf{y})$ in the linear and Gaussian scenario. If the setting is non-Gaussian, an algorithm named Metropolis-Hastings is proposed. However, sampling the posterior pdf using this method can be prohibitively expensive due to the procedure of sampling involves generating a lot of points in the state space of Markov chain. Thus, it will be more efficient if it is possible to find a limited number of points which could represent the salient features of the pdf. The variational method happens to meet this consideration, which aims to minimize the so-called cost function to indirectly maximize the posterior probability. The typical one is four dimensional variational method (4DVAR). As mentioned before, those methods of smoothing problems usually can not be used on-line rather than applied to do some reanalysis works. There are also various algorithms for the filtering problem, which could use the observations to sequentially update the probability distribution on the state. Several commonly used algorithms are Kalman filter for linear system, Extend- or Ensemble Kalman filter and three dimensional variational (3DVAR) method for non-linear dynamical system. In this thesis, we choose 3DVAR to produce a better initial condition for WRF model, given the nacelle observed wind data (as described in the chapter 2). Thus in the following a brief description of 3DVAR which has been widely used in atmospheric field is given. The 3DVAR method

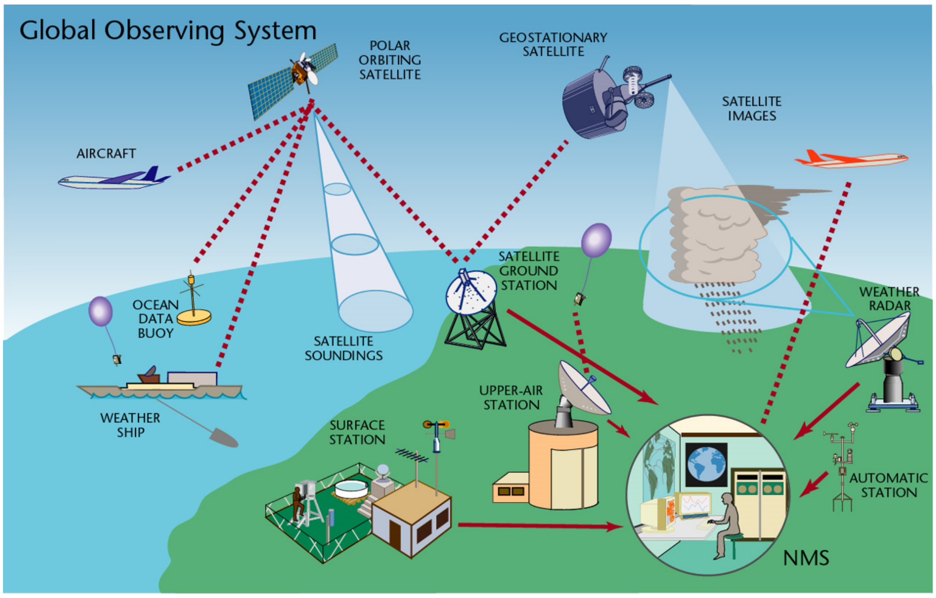

Figure 4.1 Global Observing system (http://www.wmo.ch/web/www/OSY/GOS.html).

uses a background (a priori state), which could provide a realistic reference state needed to generate the nonlinear observation operators which is used to assimilate many of the indirect observations (e.g., satellite-radiances) [4]. It is clear that the final estimate (i.e., analysis) is determined by a good background as well as a number of observations with relatively high resolution in both time and space (as shown in Figure 4.1). This also can be seen from the so-called cost function:

$$J(x) = \frac{1}{2}(x - x_b)^T B_f^{-1}(x - x_b) + \frac{1}{2}[y - H(x)]^T R^{-1}[y - H(x)] \qquad (4.21)$$

where x is the state variable, x_b is the state of model background, B_f and R are the static background and observation error covariance matrices respectively, y stands for the observation, and H is the observation operator. The aim of 3DVAR is to find a optimal x to minimize this cost function (4.21) in turn to maximum the posterior pdf. Generally, the evaluation of the error between the model solution and the observations is based on a cost function which assumes that background and observation error covariances are all described using Gaussian pdf with zero mean error [6]. It should be noted that the model error is assumed to be not a source of error (i.e., assumption of "perfect-model") during the assimilation process.

One of the biggest limitations of WRF wind forecasts at hub-height is the difficulty of obtaining accurate information on the current state of the atmosphere, which can be partly solved by using data assimilation technique based on the available observations. Here the Gridpoint Statistical Interpolation (GSI) analysis system which is capable of assimilating a diverse set of observations, is integrated with the WRF-ARW mesoscale system. More specifically, this study will implement the GSI 3DVAR system using the nacelle wind data to improve the hub-height wind forecasts. More detailed description and information of the GSI system can be found on the GSI website (http://www.dtcenter.org/com-GSI/users).

4.2 Model configuration

It should be noted that the GFS data and the domain configuration of WRF model in this section are slightly different from what we adopted in the last chapter. Specifically, the horizontal resolution of GFS data used in this section is 0.5×0.5 degree and the model domain configuration is displayed in the Figure 4.2. There are one parent domain (D01) and three nested domains (D02, D03 and D04) with horizontal resolution of 24.0 *km*, 6.0 *km*, 1.5 *km* and 0.5 *km* respectively. The resolution of the most inner domain is half of the former one ($1km$). We do this change based on some previous studies which show that better forecasting performance will be obtained when the high resolution is chosen for WRF model as well as the GFS data which are usually used as the initial and boundary conditions for WRF model. Thus, such change might improve the forecasting skills of our integrated forecasting system at some extent. As same as former configuration, 35 vertical stretched eta levels, 10 of which are within the lowest 1 *km* is used for all domains and the top level is located at 50 hPa. The topography height with a 50-meter resolution data obtained from the Geospatial Information Authority of Japan is used for D04, to furnish the local real observation information during the GSI data assimilation processing.

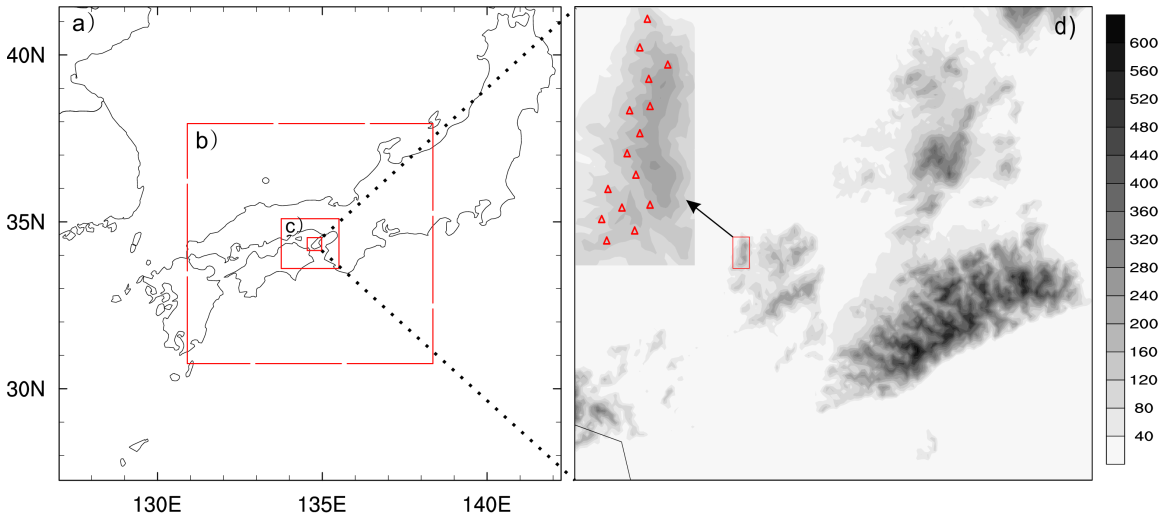

Figure 4.2 Four nested domains (a) D01, b) D02, c) D03 and d) D04) and model topography. The detailed terrain height (shaded with the gray bar in meter) of the D04 is shown in panel d). The red triangles stand for the observational sites for wind speed and direction in a wind farm, in south Awaji island, Japan.

4.3 Experiment design

Four experiments (shown in Table 1) were carried out to investigate how assimilating nacelle wind data and using Kalman filter algorithm influence the performance of WRF wind forecasts at hub-height, and to understand the priority of those two procedures. In the first experiment (Case1), only GFS data were used to obtain the raw wind speed forecasts by re-initializing WRF model as a “cold-start" at 12:00 UTC each day. In each re-initialization runs for 30 hours, the initial 6 hours (spin-up time) were excluded from the forecasting data series. The second experiment (Case2) was conducted to evaluate the impact of assimilating the nacelle wind data with cyclic mode, in comparison with the results of Case1. As displayed schematically in the Figure 4.3, the final analysis field at 18:00 UTC each day was cyclically assimilated three times with a 6-hour interval. Using this analysis field as initial condition, the wind speed forecasts of continued 24 hours were obtained. We designed the experiment of Case3 to evaluate the contribution of the Kalman filter algorithm for improving the raw forecasts of Case1 based on the available nacelle wind speed observations. Finally, the experiment of Case4 was carried out to compare the contributions of the data assimilation technique and the Kalman filter algorithm in improving the wind speed forecasts.

Table 4.1 The four sets of experiments for evaluating the impact of data assimilation and Kalman filer. The "WRF", "GSI_DA" and "KF" with "+" represent the use of WRF model, GSI analysis system and Kalman filter, respectively.

Experiment	Case1	Case2	Case3	Case4
WRF	+	+	+	+
GSI_DA		+		+
KF			+	+

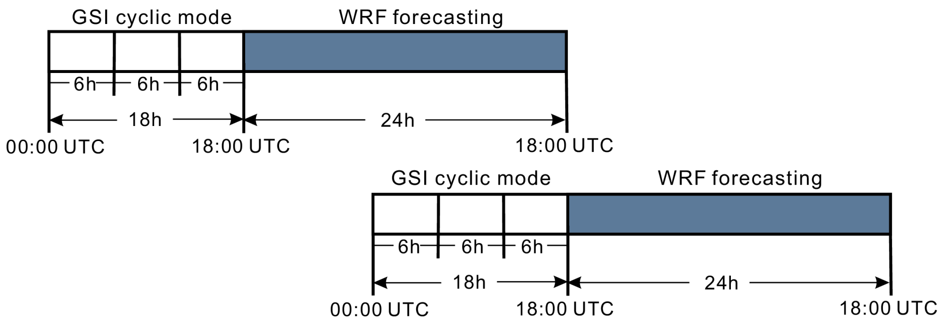

Figure 4.3 The schematic of implementing GSI system in cyclic mode. The white boxes stand for the total assimilation time of 18-h with an interval of 6-h, while the gray boxes represent the forecasting length (24-h) of WRF model after assimilating the nacelle wind data.

4.4 Impact of assimilating nacelle wind observations

In this section, firstly the overall result of comparison between the raw wind speed forecasts of WRF model and the forecasts with data assimilation is presented based on the statistical parameters introduced in section 2.3.1.2. Then the role of the data assimilation technique and the Kalman filter algorithm in improving the raw wind speed forecasts of WRF model will be investigated. It is noted that all of the statistics and discussions are based on the 15-turbine averaged data unless there is a special instruction.

4.4.1 Impact of assimilation on the hub-height wind forecasts for the experiment period

Figure 4.4 (a) illustrates the comparison between raw forecasts (black), observations (red) and the forecasts with data assimilation (blue) of wind speed, which are symboled with "Case1", "Obs" and "Case2" respectively, during the experiment period from 18:00 UTC 2 January to 23:00 UTC 31 January 2016. It can be seen that the raw wind speed forecasts (Case1) reproduces the observation with relatively good accuracy, though there are occasional large errors, especially when the observed wind speed is larger than 15 ms^{-1}. It is also easy to find that the blue line lies closer to the red line than the black one during almost the whole period, which implies that the forecasts of wind speed with data assimilation are remarkably improved in comparison with the forecasts without data assimilation (Case1). Obviously, the evidence from Figure 4.4 (a) also shows that the forecasting skill of wind ridge is largely increased after assimilating nacelle wind data, though the contribution of data assimilation for other periods is relatively slighter or even not clear in some cases. This might be attributed to the WRF model itself whose forecasting ability of large wind speed at hub-height is inferior to that of regular wind speed (e.g. ranging from 4 ms^{-1} to 15 ms^{-1}). It reveals that implementing data assimilation with the nacelle wind data can significantly improve the forecasting skill of WRF model in extreme weather conditions. The values of ME, RMSE, IA and CC are listed in Table 4.2 to quantify the effects of assimilating nacelle wind data into the WRF model, in comparison with the raw forecasts. Examining the second column, the positive values of ME indicate that both cases overestimate the wind speed in the whole period (nearly one month). Compared to the raw forecasts (Case1), the ME in case 2 is reduced by 34.3% where the nacelle wind data is assimilated. Regarding RMSE, the value of Case2 is much smaller compared to Case1, with a relative error reduction of 23.9%. Similarly, both IA and CC are increased when assimilation is implemented.

The conclusions obtained from the 15-turbine averaged results needs to be further verified on the individual turbine. Figure 4.4 (b) illustrates the comparison of the RMSE between the raw wind speed forecasts and the forecasts with data assimilation for 15 turbines separately. It is found that forecasting skills of wind speed are improved after assimilating nacelle wind data for all 15 turbines, though small differeces are existing. The relative improvement in RMSE varies from 19.5% to 25.9% with an average of 21.5%.

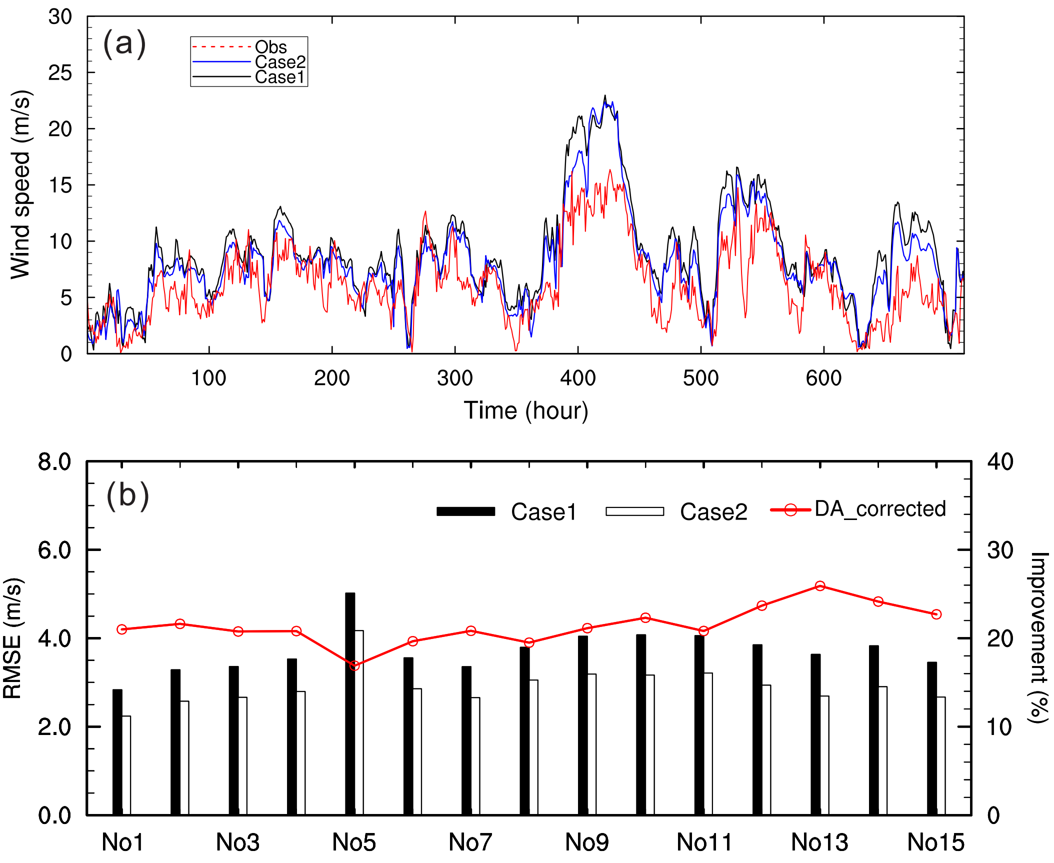

Figure 4.4 (a) One month series of the raw wind speed forecasts (black), the forecasts with assimilation (blue) and the corresponding observations (red). (b) The comparison of the RMSE of wind speed using data assimilation (solid white bar) with respect to the raw forecasts (solid black bar) for the 15 turbines. The marked line stands for the relative improvement after using data assimilation. The period is from 18:00 UTC 1 January to 23:00 UTC 31 January 2016.

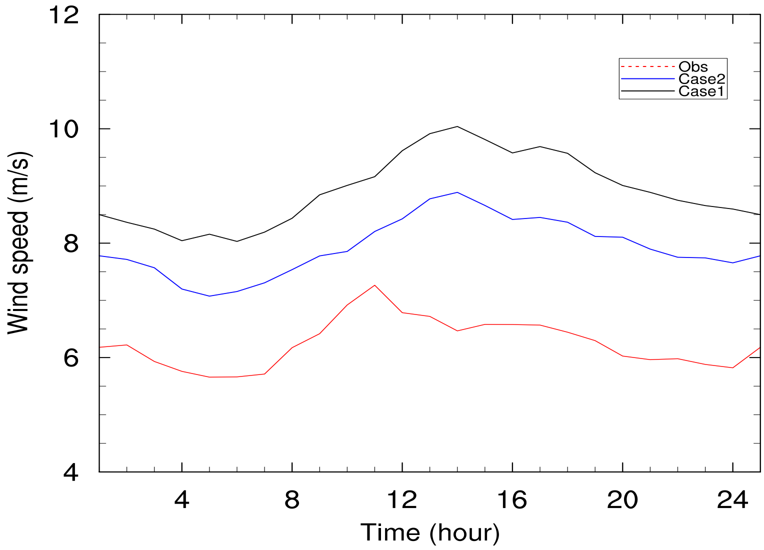

Figure 4.5 Comparison between the hub-height wind speed forecasts with (blue) or without (black) assimilating the nacelle wind data based on the corresponding observations (red), during the 24-h forecasting period (30-day averaged).

Table 4.2 The monthly mean ME, RMSE and CC calculated with the forecasts (with or without assimilating the nacelle wind data) and the corresponding observations of wind speed at hub-height.

Experiment	ME (ms^{-1})	RMSE (ms^{-1})	IA	CC
Case1	2.54	3.51	0.80	0.83
Case2	1.67	2.67	0.87	0.84

Figure 4.5 displays the 30-case mean (30 days) forecasts of Case1, Case2 and observations during the 24-h forecasting length. Apparently, the overestimation of wind speed is found no matter whether the nacelle wind data is assimilated or not. However, this systematic discrepancy has been significantly corrected by using data assimilation technique. The relative decrease of RMSE (36.4%) further demonstrates the large impact of assimilating nacelle wind data in reduction of the systematic bias in WRF model.

From the above discussions, we may conclude that assimilating nacelle wind data can substantially improve the accuracy of WRF model in forecasting the hub-height wind field in the target wind farm site of interest.

4.4.2 Comparison of the role of the Kalman filter and data assimilation

Having confirmed the effect of assimilating nacelle wind data on improving the raw wind forecasts at hub-height, we further evaluated the integrated forecasting system which uses the Kalman filter as another key technique to improve the prediction. To this end, we conducted other two experiments, i.e. Case3 and Case4, to include the Kalman filter as another module. In order to implement Kalman filter properly, the first 15 days are chosen as a training period and thus the following discussions are all based on the forecasts and corresponding observations of the second half 15 days. Figure 4.6 displays the statistical parameters that quantify the performance of the integrated prediction system and the contributions of its different components in forecasting the hub-height wind under configurations of the four test cases.

As observed above, the ME and RMSE of Case2 are largely reduced compared to Case1, while the values of IA and CC are increased, which shows the large improvement due to assimilating nacelle wind data.

The effects of implementing Kalman filter to the raw forecasts of WRF model are also examined by comparing the results of Case1 and Case3 in Figure 4.6. It seems that the bias (Figure 4.6 (a)) in the raw forecasts can be largely revised and meanwhile the random errors (Figure 4.6 (b)) can be partly reduced as well. Furthermore, the values of IA and CC of Case3 become larger after using Kalman filter compared to the Case1. All of these results demonstrate that the Kalman filter as a post-processing method, can significantly improve the forecasting skill of hub-height wind speed.

Figure 4.6 also illustrates the difference between data assimilation and Kalman filter when one compares among the results of Case2, Case3 and Case4. For Case4, in which the data assimilation is used to improve the initial condition and then Kalman filter is adopted to post-process the forecasts, the RMSE is further reduced and the values of IA and CC are larger than both Case 2 and Case3, while the value of ME is nearly same. This implies that combining the Kalman filter and the nacelle wind data assimilation can provide the best forecasts and the role of Kalman filter is more important in calibrating the systematic bias. On contrary, comparing the RMSE and IA of Case3 and Case4 suggests that assimilation of nacelle wind data shows better performance against Kalman filter in revising random uncertainties. However, if we consider the differences represented by all four statistic parameters of Case2 and Case3 synthetically, the Kalman filer shows the priority over data assimilation for wind speed forecasts at hub-height.

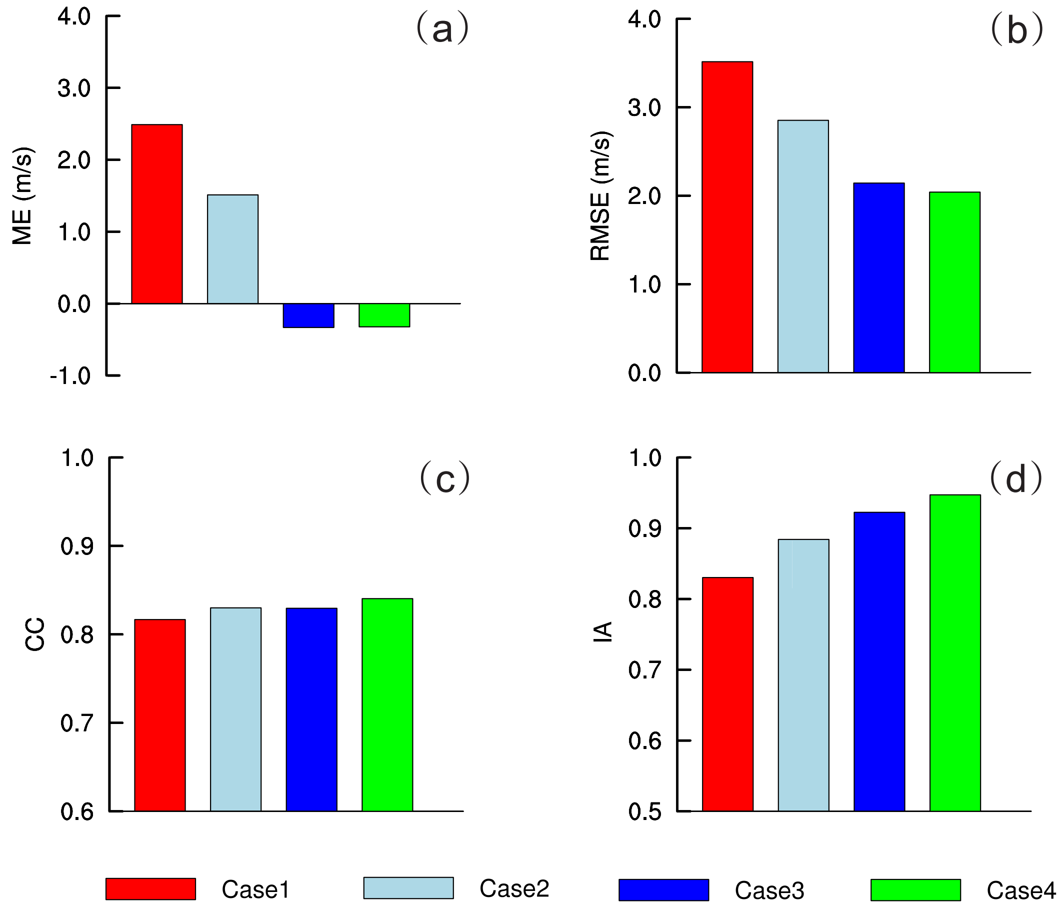

Figure 4.6 The ME (a), RMSE (b), CC (c) and IA (d) of four cases described in Table 2.3. All results are based on data series of the second half 15 days (16-31 January 2016 with 1-h interval).

To further evaluate the improvements of assimilating nacelle wind data and Kalman filter, we show in Figure 4.7 (a) to Figure 4.7 (c) the ME, RMSE and IA at different forecasting periods (i.e., 0–12-h and 12–24-h). In regard to the raw forecasts (Case1), the forecasting skill at period of 0–12-h is slightly higher than the period of 12–24-h. The same conclusion can be drawn for the forecasts with data assimilation (Case2) where the errors in the raw forecasts have been largely reduced after assimilating nacelle wind data. As same as shown in the Figure 4.6 (a, b and d), Kalman filter (Case3 and Case4) can significantly improve the model forecasts under the situations with or without data assimilation for different forecasting periods. Comparing the values of ME, RMSE and IA of Case3 and Case4 during different periods, we observe that the impact of data assimilation is more apparent compared to Kalman filter in the period of 12–24-h, due to the difference between Case3 and Case4 during 12–24-h is larger than that in the period of 0–12-h.

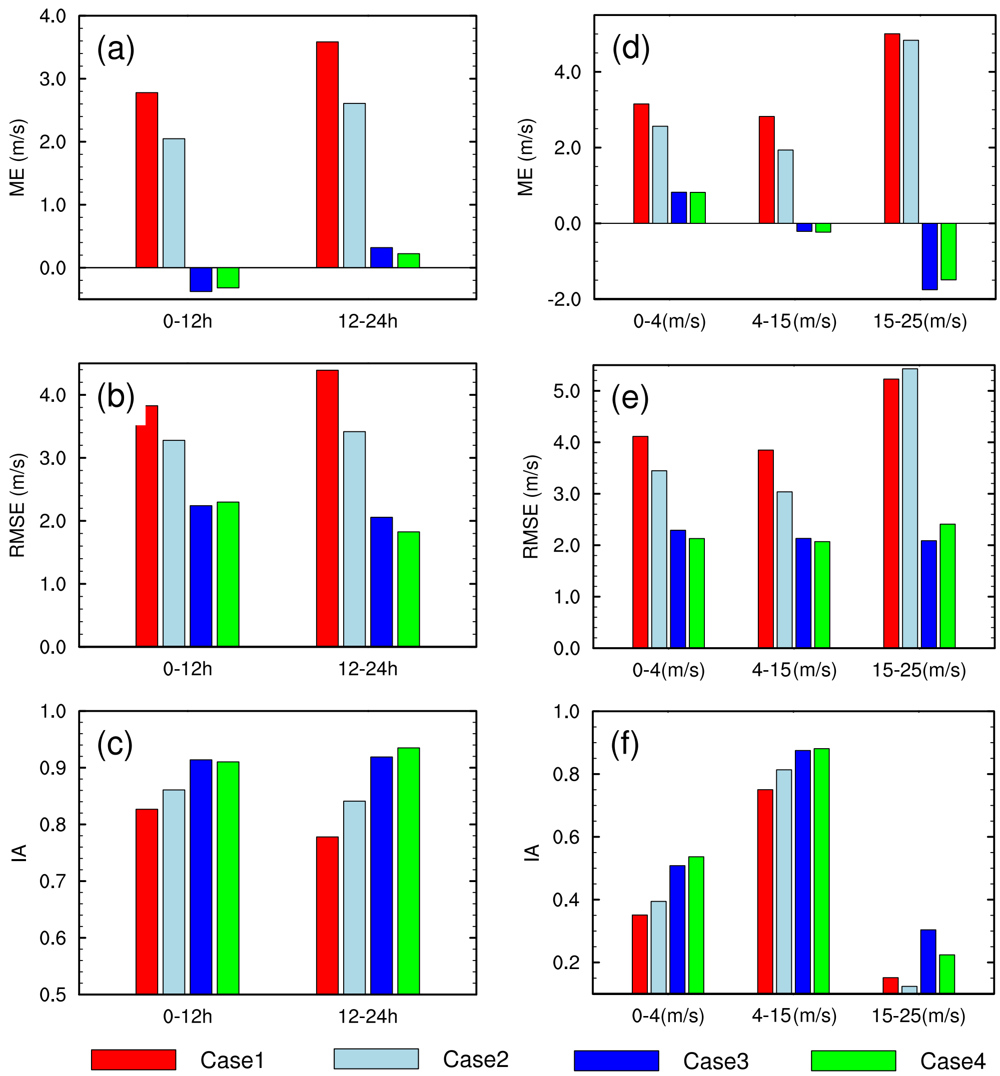

Figure 4.7 The statistical parameters of the hub-height wind speed forecasts for four cases in different forecasting periods ((a)-(c)) and different wind speed bands ((d)-(f)). Same as in Figure 4.6, the evaluation period is 15 days from 16 to 31 January 2016.

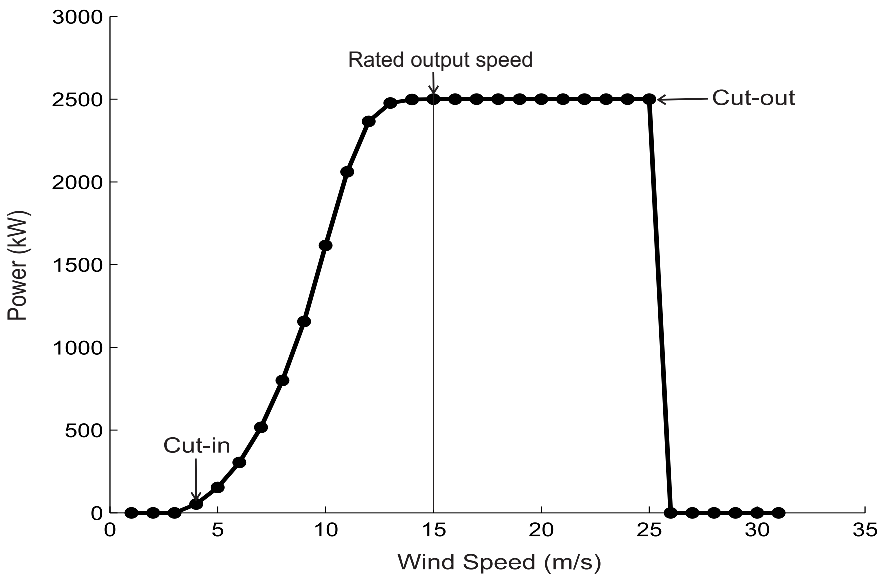

Figure 4.8 The theoretical wind power curve for a 2.5 MW turbine used in this study and the corresponding cut-in, rated output and cut-out speed.

In practice, the simplest way to obtain the wind energy forecasts of a specific turbine is to use the designed (or theoretical) power curve provided by the turbine manufacturer, which is usually a function of the mean hub-height wind speed. In this study, the target wind farm consists of 15 2.5-MW horizontal-axis turbines and the corresponding power curve is shown in Figure 4.8. As displayed, the value of cut-in (4 ms^{-1}), rated output (15 ms^{-1}) and cut-out speed (25 ms^{-1}) is crucial to power management in routine operations. Therefore, the forecasts of 0-4, 4-15 and 15-25 ms^{-1} wind speed bands are further validated. The results are displayed in Figure 4.7 (d) to Figure 4.7 (f). Apparently, almost all of the ME and RMSE of the raw forecasts (Case1) are reduced by applying both data assimilation and the Kalman filter, meanwhile the value of IA is increased. In addition, the smallest value of ME and RMSE and the largest IA in the band of 4-15 ms^{-1} indicate that it is easier to obtain relatively accurate hub-height wind speed forecasts in the interval of 4-15 ms^{-1} compared to other wind speed bands. When the wind speed is larger than 15 ms^{-1} which is the rare case in the experiment period, the performance of the system becomes worse (with small IA), and data assimilation shows more significant effect in comparison with the Kalman filter. It suggests that the data assimilation technique can be more effective in correcting the forecast under rare or extreme weather conditions.

Overall, the forecasts are remarkably improved after assimilating nacelle wind (Case2) or using the Kalman filer (Case3). The largest improvement is found when the two techniques are combined (Case4). It seems that the role of the Kalman filter is more dominated as the difference between the Case3 and Case4 is much smaller than that between Case1 and Case2, while data assimilation becomes more important in rare or extreme weather conditions.

4.5 Summary

We have developed a practical forecasting system for surface wind and power output by integrating data assimilation and Kalman filter into the WRF model. Both data assimilation and Kalman filter modules make use the nacelle wind data which is routinely available, so the system can be easily adopted in different wind farm sites for operational use. Due to the complex topographic features, the surface wind field in Japan region is significantly fluctuating and more difficult to predict, the present system employs data assimilation and Kalman filter to eliminate the uncertainties from two aspects, i.e. the data assimilation improves the accuracy in initial conditions and Kalman filter provides a posterior correction to the raw model output, and thus can be expected as a promising tool in real-case operations.

The system has been validated using the data of a wind farm in Awaji island, Japan. The wind speed forecasts at hub-height have been substantially improved by data assimilation to refine the initial wind field for WRF model, i.e. the ME and RMSE errors in WRF prediction were reduced by 34.3% and 23.9% respectively, while IA has been improved by 8.8% due to the data assimilation technique. On the other hand, the Kalman filter, as a post-processing method, is able to provide more reliable wind forecasts with a short training period (15-day in this study). By using both Kalman filter and nacelle wind data assimilation, the raw forecasts can be further improved. Detailed evaluation indicates that the role of the Kalman filter is more dominant for the wind band of rated out speeds, while data assimilation becomes more important in rare or extreme weather conditions.

Chapter 5

Coupling WRF and a CFD model

So far, an integrated forecasting system based on WRF model, Kalman filter and data assimilation has been built. Comparing the forecasts from this system with the observations indicates that it can provide reasonable predictions for the Awaji wind farm in Japan, where the terrain conditions are usually very complex. In fact, the complex terrain features always limit the forecasting ability of meso-scale WRF model for predicting wind at hub-height, due to the relatively coarse grid resolution of WRF model. Thus it may not be appropriate to rely solely on WRF based system to handle the locale-specific flow in the complex terrain regions. As Zajaczkowski et al. [121] mentioned, the mesoscale model cannot accurately capture the wind flow features (e.g., turbulence) finer than 1 *km* caused by the local terrain conditions.

With this in mind, Computational Fluid Dynamics (CFD) model may be an option for simulating and predicting the flow characteristics of smaller scales due to a finer resolution of terrain features. There are several studies [83, 8, 80] which have used CFD models to investigate how the complex terrain affect the wind flow and several encouraging results have been obtained. However, almost of those studies based on the ideal wind profile or fixed wind direction which is not very suitable for the real applications, especially for short-term forecasting. Actually, a couple of studies which have used CFD models with real boundary conditions given by mesoscale models to study the flow and dispersion in built-up areas [76, 77], though there are very limited application in prediction of wind energy.

Inspired by those aforementioned studies, in this chapter, a multi-scale modeling system which consists of a meso-scale NWP model and a micro-scale CFD model, to the knowledge of the authors, is firstly developed to accurately forecast short-term hub-height wind for a wind farm of interest in Japan. The CFD model adopted is the free and open-source package which is called OpenFOAM, while the WRF model is chosen as the mesoscale component.

The remainder of this chapter is arranged as follows. Initially, the fundamental equations of the mathematical models and a brief introduction of OpenFOAM used in this work are presented. Then the specific procedure of conducting OpenFOAM model is described and some simple simulation results are displayed to show the ability of OpenFOAM of simulating wind flow over the complex terrain. Sequentially, the results from the multi-scale forecasting system are validated with the nacelle wind observations at each wind turbine site, and finally some conclusions are given.

5.1 Theoretical background

The Reynolds Averaged Navier-Stokes (RANS) equations combined with a turbulence model is the most common way in wind engineering when simulating wind flow over complex terrain. As mentioned in the chapter 1, this kind of approach shows a fair compromise between computational burden and modeling accuracy.

5.1.1 Governing equations

This section presents the governing equations of fluid flow. The most common set of equations employed to simulate the characteristics of turbulence flow is the Navier-Stokes equations. RANS equations are often used to govern the so-called incompressible, isothermal flow of a newtonian fluid. The RANS incompressible equation is adopted in this thesis to simulate and solve the turbulent wind flow over the complex terrain. Applying the Reynolds decomposition (a time-averaged part and a fluctuating part), the mass and momentum equations are displayed as follows:

$$\frac{\partial \overline{u_k}}{\partial x_k} = 0 \tag{5.1}$$

$$\rho(\frac{\partial \overline{u_i}}{\partial t} + \overline{u_j}\frac{\partial \overline{u_i}}{\partial x_j}) = -\frac{\partial \overline{p}}{\partial x_i} + \rho\overline{g_i} + \frac{\partial}{\partial x_j}\left[\mu(\frac{\partial \overline{u_i}}{\partial x_j} + \frac{\partial \overline{u_j}}{\partial x_i})\right] - \frac{\partial}{\partial x_j}\rho\overline{u_i^{'}u_j^{'}} \tag{5.2}$$

where μ is the dynamic viscosity, g_i is the gravity acceleration, ρ is the density of dry air and $\overline{p}$ is the averaged pressure. The most common method to capture the time and length scale effectively is to associate the Reynolds stresses with the mean strain rate tensor using the Boussinesq approximation:

$$\tau_{ij} = -\rho\overline{u_i^{'}u_j^{'}} = \mu_t(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}) + \frac{2}{3}\rho k\delta_{ij} \tag{5.3}$$

where k is the so-called turbulent kinetic energy (TKE) and μ_t is the turbulent viscosity. The time averaged bars have been removed for convenience. Due to the flow is incompressible, thus the pressure variations can be referred to the hydrostatic pressure, which implies that the gravity term in the equation 5.2 can be added in the pressure term given by:

$$-\frac{\partial \overline{p^{\dagger}}}{\partial x_i} = -\frac{\partial \overline{p}}{\partial x_i} + \rho\overline{g_i} \tag{5.4}$$

According to the study [102], the term $\frac{2}{3}\rho k\delta_{ij}$ is also included in the above pressure term. Therefore, combining equations from 5.1 to 5.4, omitting the Coriolis force and neglecting superscripts of averaged bars and † for convenience, the governing equations can be rewritten as:

$$\frac{\partial u_k}{\partial x_k} = 0 \tag{5.5}$$

$$\rho u_j\frac{\partial u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}\left[\mu_t(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i})\right] \tag{5.6}$$

5.1.2 Turbulence model

The so-called turbulence model involves the averaging of the Navier-Stokes equations known as RANS. In this thesis, the turbulence modeling is limited to two-equation RANS models where the Reynolds stresses are calculated following the concept of turbulent viscosity.

$k-\varepsilon$ model

In this model, the turbulent viscosity μ_t is calculated as [65]:

$$\mu_t = C_\mu \frac{k^2}{\varepsilon} \tag{5.7}$$

where C_μ is a model constant, $k = \frac{1}{2}(\overline{u'^2} + \overline{v'^2} + \overline{w'^2})$ stands for the turbulent kinetic energy and ε is the dispassion rate. The value of k and ε can be obtained using the following two equations:

$$\rho \frac{\partial(u_j k)}{\partial x_j} = \mu_t \frac{\partial u_i}{\partial x_j}(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}) + \frac{\partial}{\partial x_j}\left[(\frac{\mu_t}{\sigma_k})\frac{\partial k}{\partial x_j})\right] - \rho\varepsilon \tag{5.8}$$

$$\rho \frac{\partial(u_j \varepsilon)}{\partial x_j} = C_{\varepsilon 1}\frac{\varepsilon}{k}\mu_t \frac{\partial u_i}{\partial x_j}(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}) + \frac{\partial}{\partial x_j}\left[(\frac{\mu_t}{\sigma_k})\frac{\partial \varepsilon}{\partial x_j})\right] - \rho C_{\varepsilon 2}\frac{\varepsilon^2}{k} \tag{5.9}$$

where $C_{\varepsilon 1}$ and $C_{\varepsilon 2}$ are two additional model constants, which follows the following metric:

$$C_{\varepsilon 1} = C_{\varepsilon 2} - \frac{\kappa^2}{\sqrt{C_\mu}\sigma_\varepsilon} \tag{5.10}$$

where $C_\mu = (\frac{u_*^2}{k})^2$, u_* is the friction velocity and σ_ε, σ_k are two other model constants. The values of those model constants are usually obtained empirically to tune the model. Several sets can be found in the Table 5.1.

Table 5.1 A set of typical coefficients of the $k-\varepsilon$ model. [71]

Coefficient	κ	C_μ	σ_k	σ_ε	$C_{\varepsilon 1}$	$C_{\varepsilon 2}$
Standard	0.4	0.09	1.00	1.30	1.42	1.92

5.1.3 Micro-scale model of OpenFOAM

The Open Source Field Operation and Manipulation (OpenFOAM) software package was developed at Imperial College of London during 1990-1999. Its core library is written using C++ and designed to effectively solve complex physic problems using finite volume method (FVM) and structured or unstructured discretization grid. The OpenFOAM software package integrates many multi-physics numerical solvers for incompressible and compressible fluid in steady state, transient and turbulent flows as well as laminar, and thus variety range of real fluid dynamic problems can be solved. It also has the ability for parallel calculation which is an advantage for large and complex simulations.

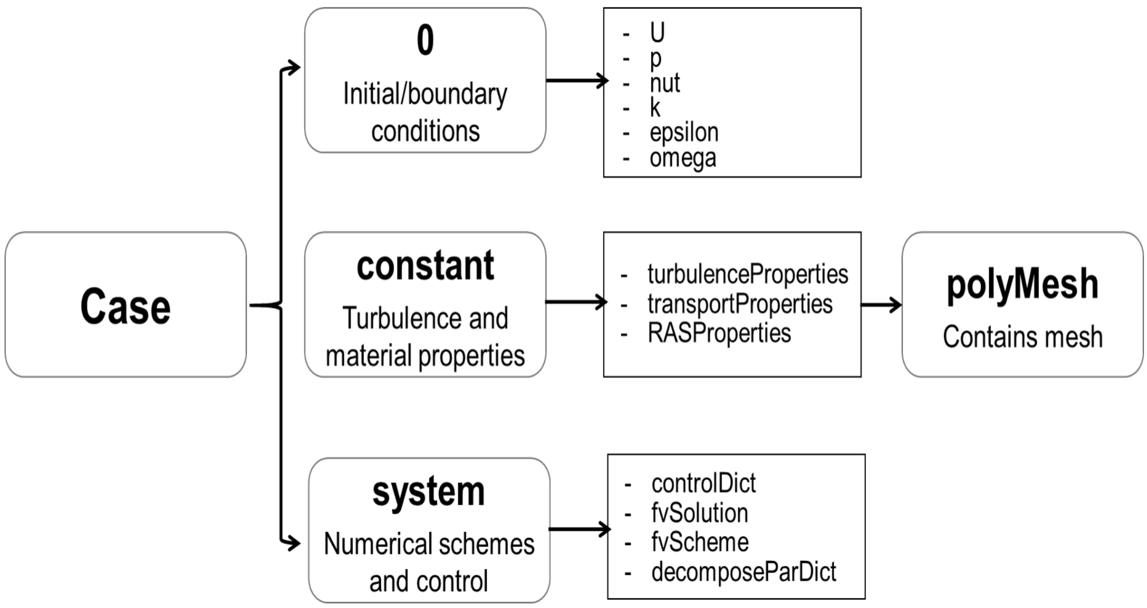

Figure 5.1 Example of an OpenFOAM case folder structure and its content.

Before using the OpenFOAM properly, the users should be familiar with its file structure. As displayed in the Figure 5.1, every OpenFOAM case contains three main folders: *0,* *constant and system*. *0 folder* contains files in which the initial field information for different variables and surface patches can be specified. *Constant folder*, contains Poly mesh folder, turbulence property and material property. The file of *RASproperties* includes the option to choose a RANS turbulence model or change to laminar flow. *transportProperties* contains the possibility to specify the value of the kinematic viscosity ν. By modifying this file, the coefficients for different turbulence models can be decided. *system folder* includes solution settings, time step and solver settings for each field which are controlled by editing the *controlDict* file. Different discretization schemes for time and space are specified in the file of fvSchemes. In the file of fvSolution, users are able to decide which solver is to be used to solve the problems of interest.

5.2 Coupling WRF and OpenFOAM

5.2.1 Coupling Procedure

Two models mentioned above have been demonstrated that they have strong ability of solving scientific problems in their own fields. However, it is imperfect when we use those two models solely to simulate or forecast wind flow over the complex terrain at a real wind farm for operational use. That is just the main reason why we intend to couple those two components together. Typically, the coupling is achieved by using the mesoscale WRF forecasts as the initial and boundary conditions for the micro-scale OpenFOAM. Thus, the OpenFOAM will be provided instantaneous boundary values representing the real atmosphere state. Figure 5.2 displays the relative position of WRF and OpenFOAM domains. Obviously, the grid spacing and domain size of the WRF and OpenFOAM are quite different. The WRF domain shown here is the most inner one among the four-nested domains as configured in the chapter 4, of which the horizontal resolution is 500 $m \times 500$ m. Whilst, the resolution of the OpenFOAM mesh is chosen as 50 m in this study, which is much higher than WRF model. The size of the OpenFOAM is 10 $km \times 10$ km which means that several WRF grids are contained in this region. This kind of configuration is beneficial for extracting more useful information from WRF model for providing initial/boundary conditions to OpenFOAM.

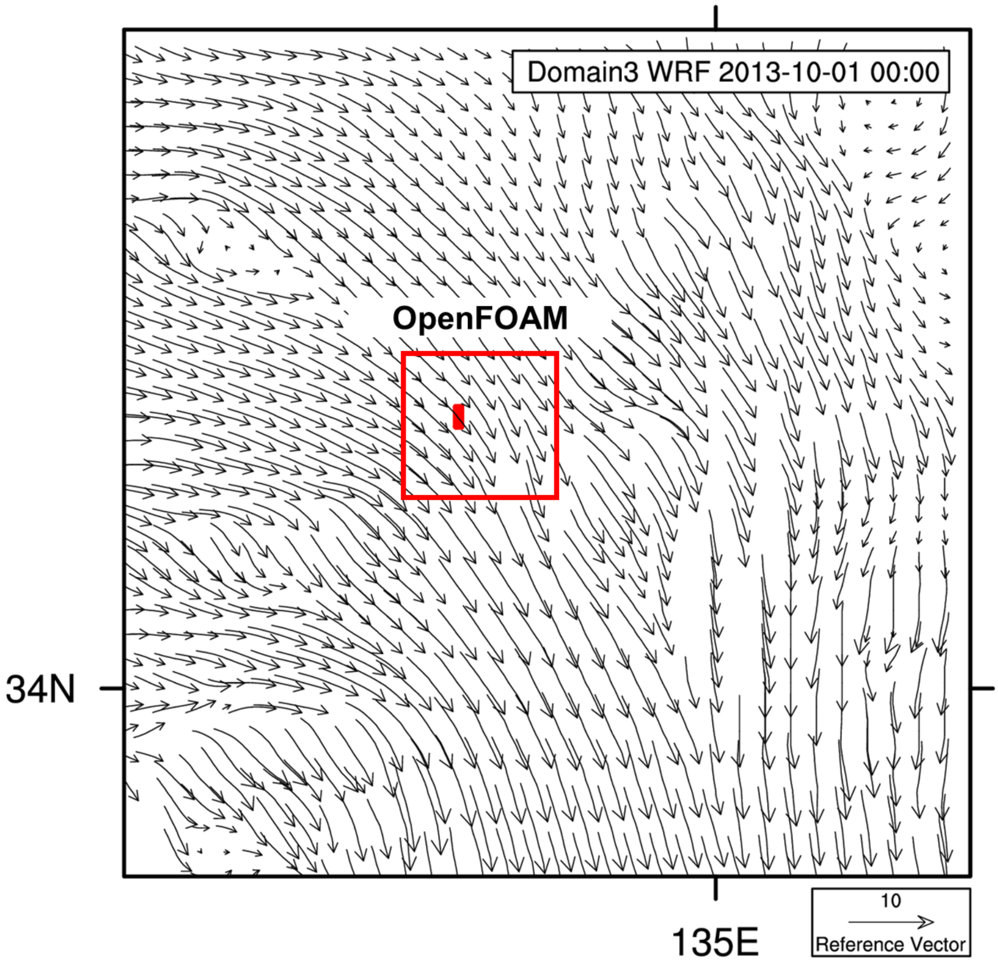

Figure 5.2 Spatial comparison between WRF and the OpenFOAM domain.

The general steps for conducting the multi-scale system are shown as following:

- Meso-scale WRF component
  - Domain configuration, Global data pre-processing and choosing the suitable schemes for the wind farm of interest;
  - Running the main WRF program with Parallel mode (40 cpus are used);
  - Outputting the forecasted variables (u, v) with NetCDF format;
  - Converting the pressure coordinate system to Cartesian coordinate system aiming to OpenFOAM model use conveniently.
- Micro-scale OpenFOAM component

- Processing surface terrain data from a topographical map into a surface mesh;
- Mesh generation using an external Fortran program and the *blockMesh* tool in OpenFOAM;
- Mesh generation and discretization
- calculating the prevail wind direction to decide the patch attribute;
- interpolating the well prepared WRF outputs to the collocated mesh of OpenFOAM, specifying appropriate boundary conditions, defining the fluid properties and selecting the physical/chemical phenomena that need to model;
- Choosing reasonable solver (simpleFoam in this thesis) and turbulence model ($k-\varepsilon$) to solve the problem;
- Using the tool of *sample* in the OpenFOAM to obtain the wind forecast at each wind turbine sites.

The details of running WRF model have been stated in the former chapters. Thus, we mainly describe the micro-scale component (namely OpenFOAM) in the following sections.

5.2.2 Terrain and mesh generation

The terrain data from Shuttle Radar Topography Mission (SRTM) (Jarvis et al., 2008) has been downloaded for the Awaji wind farm region. This data is available in TIFF format on the USGS website with a resolution of approximately 90 meters (3 arc-seconds) with the user providing the latitude and longitude of the area. Since OpenFOAM uses the Cartesian coordinate system, the coordinates are changed from Geographic to the UTM (Universal Transverse Mercator) coordinate. The datum selected is WGS 84 and the zone selected is 53 N. Figure 5.3 shows the final terrain surface used in this study, of which initial size is $10\ km \times 10\ km$.

There are a couple of tools can be used directly to generate the mesh if the terrain data can be converted to *stl* format, for example, the tool of *snappyHexMesh* or *extrudeMesh* in the OpenFOAM is a common choice. In addition, some external meshing softwares such as

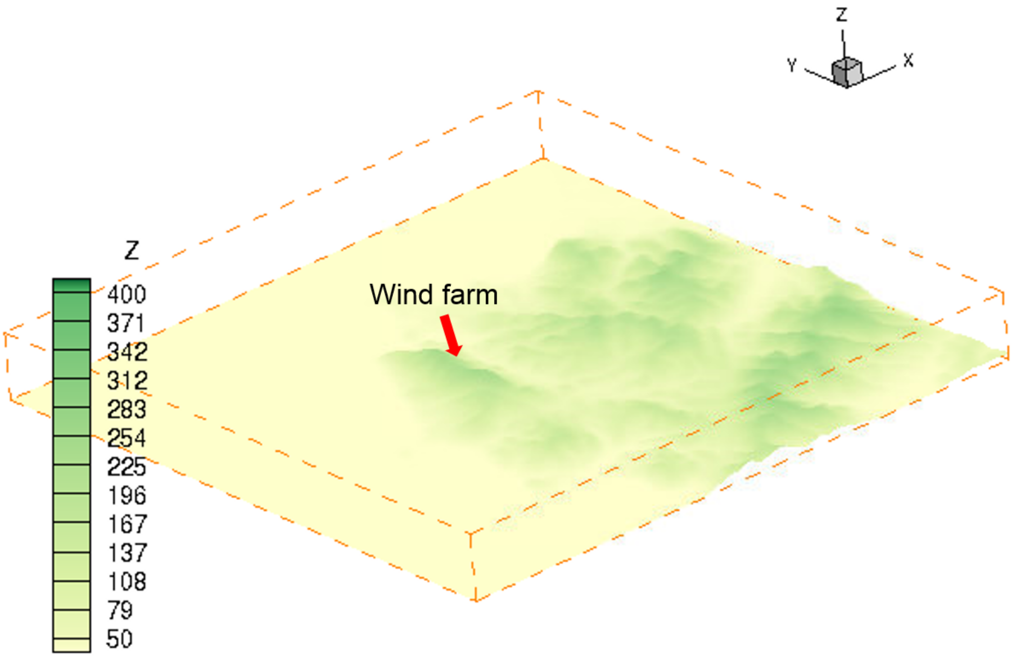

Figure 5.3 Surface model for Awaji wind farm region.

pointwise and *icem* also can be chosen to generate the mesh of the computational domain. Those mentioned tools at some extent have some drawbacks. For example, as pointed out in the study [71], the total amount of cell in the final mesh generated by *snappyHexMesh* tool is difficult to control and the quality of the mesh is highly dependent on the available computational resources as well as the users' skills. Though the whole mesh quality e.g., (*skewness)* can be improved by fining the mesh near terrain surface, the large number of the unstructured mesh will lead to heavy calculation burden and very long computational time which is not favorable for operational forecasting use.

Therefore, we consider to develop an external program coded using Fortran language, which could control the total amount of cells and the adequate aspect ratio near the terrain surface to fine the local region. In fact, this program cannot generate the mesh directly instead output a file named *blockMeshDict* which can be recognized by the tool of *blockMesh* in the OpenFOAM. Thus, combining the Fortran program and the *blockMesh* tool, a structured mesh has been generated as displayed in the Figure 5.4. The terrain surface grid spacing is 50

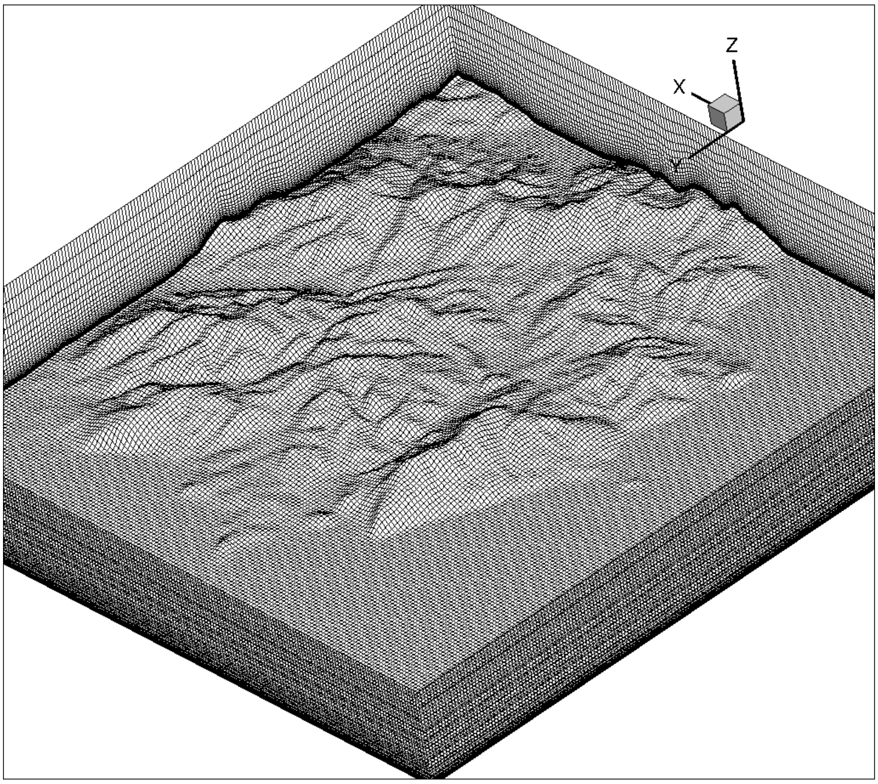

Figure 5.4 The structured mesh of the computational domain.

m and the total number of the mesh is 200×200. In the vertical, the expansion ratio is set as 1.94, indicating the grid spacing in the vertical direction is increasing with the height above the ground, which can be clearly found in the Figure 5.5. The current domain boundaries are decomposed in six patches: ground to which was assigned the *wall* attribute, and the other five named as *top, north, south, west* and *east* are all defined as generic *patch* attribute. It is worth to note that we name the patches with *north, south, west* and *east* rather than *inlet, outlet, back* or *front* because of the consideration of operational use. Usually, the real wind direction changes hour by hour which will require us to change the patch attributes frequently in order to obtain reasonable simulation results. Thus, the way of naming the patches will avoid some misunderstandings. In this study, 30 levels in the vertical are adopted, indicating that the total number of the mesh of the whole computational domain is 1,200,000.

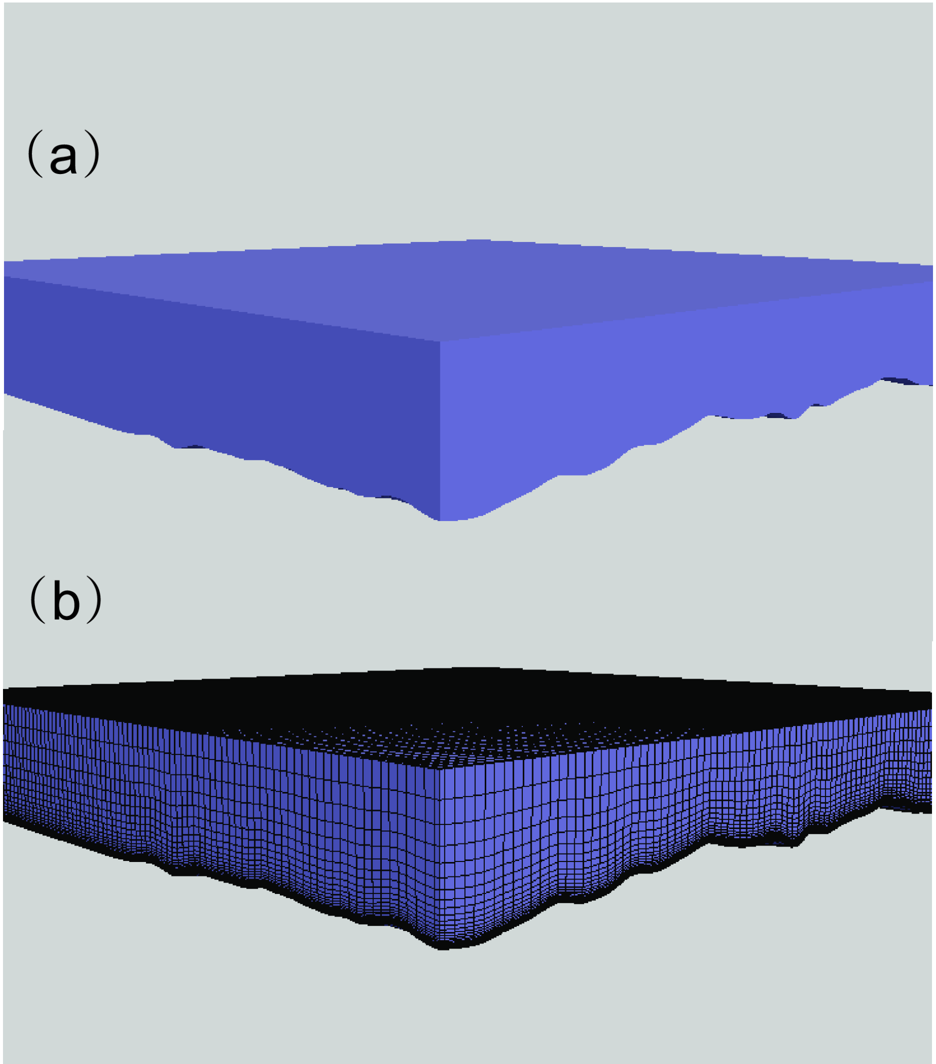

Figure 5.5 Detailed view of a clip (a) of the whole computational domain and the corresponding grids (b).

5.2.3 Boundary conditions

As known, in the CFD modeling, the boundary condition types and values must be specified appropriately in order to get a reasonable solution. In the following sub-sections, details of setting boundary condition for every patch will be presented briefly.

- **Wall boundary condition**
  In our study, only the *ground* patch is considered as *wall*. The no-slip condition is set for the velocity **U** (a constant value of (0, 0, 0) is fixed at the wall). For the pressure **p**, the condition is set as *zerogradient*. The *nutRoughWallfunction* which can take into account roughness is used to calculate the μ_t.

- **Top boundary condition**
  As stated in the study [43], the choice of the top boundary condition is very important for sustaining equilibrium boundary layer profiles. Here the slip condition is chosen. For a scalar, it represents a *zerogradient* condition while for a vector the normal component is fixed to zero and the tangential component is set as *zerogradient*.

- **Inlet boundary condition**
  The velocity **U** is prescribed by the real data from meso-scale WRF model, while p is still treated as *zerogradient* condition. It should be noted that there is not just one inflow patch for the real applications due to the wind direction always changes hourly. In other words, the decision of which patch has *inlet* attribute is made by the real wind direction. Specifically, if the wind direction is during 0-90 degree, the patches of *north* and *east* are set the *inlet* attribute; when it comes to 90-180, *east* and *south* will be chosen; for 180-270 degree, the patches of *south* and *west* should be *inlet* boundary condition and finally the patches of *west* and *north* are regarded as *inlet* boundary when the wind direction is from 270 degree to 360 degree.

- **Outlet boundary condition**
  The *zerogradient* condition is applied for all variables except for the pressure p, where

a fixed value of zero is set. As similar to the **Inlet boundary condition**, the *outlet* patches are also decided by the real wind direction.

5.2.4 Roughness modeling

Usually, the roughness of the terrain surface is vital to the calculation of the kinetic turbulence viscosity. In the OpenFOAM package, it is included in the *nut* file as a boundary condition of kinetic turbulence viscosity, being the aerodynamic roughness length which is a parameter of a wall function for calculating the fluid kinetic viscosity μ_t. As mentioned in the study [3], the way of terrain roughness modeling has three types. The first one (or the easiest one) is considering the aerodynamic roughness length over all terrain surface is a uniform value, which can be set as an empirical coefficient based on the different surface types as displayed in the Table 5.2. The second option is developing an appropriate roughness model based on some specific situations, for example, different ground height with different roughness. The third type is using the real roughness map data. For the simplification, the constant roughness is adopted to solve our problems.

Table 5.2 Typical roughness parameters for non-urban homogeneous terrain [113].

Surface type	Roughness (m)
Sea, sand and snow	0.0002
Flat desert	0.0002-0.0005
Short grass	0.008-0.03
Long grass	0.02-0.06
Low crops	0.04-0.09
High crops	0.12-0.18
Continuous bush land	0.35-0.45
Mature pine forest	0.8-1.6

5.3 Results of the coupled system

In this section, we initially intend to check the rationality (e.g., the assumption, setting of boundary conditions or the choice of some coefficients) of the coupled system by simulating an arbitrary case. Then forecasting ability of the multi-scale system is validated with the nacelle wind observations at each wind turbine site over a 8-day period (192 cases).

5.3.1 One case validation

The simulation case at 00:00 UTC 01 October, 2013, is selected to preliminarily validate the performance of multi-scale forecasting system for simulating wind flow over complex terrain. Due to the prevail wind is northwest, thus the patches *north* and *west* are set as *inlet* while *east* and *south* are regarded as *outlet* boundaries. The boundary condition of velocity for both *north* and *west* patch is derived from meso-scale WRF model, which is displayed in the Figure 5.6. The solver *simplefoam* and turbulence $k-\varepsilon$ model are chosen. Throughout all this study, we assume that the wind flow turns to steady state when the residual of U_x, U_y, U_z and k are all smaller than 10^{-3}, meanwhile the residual of p should be smaller than 10^{-2}.

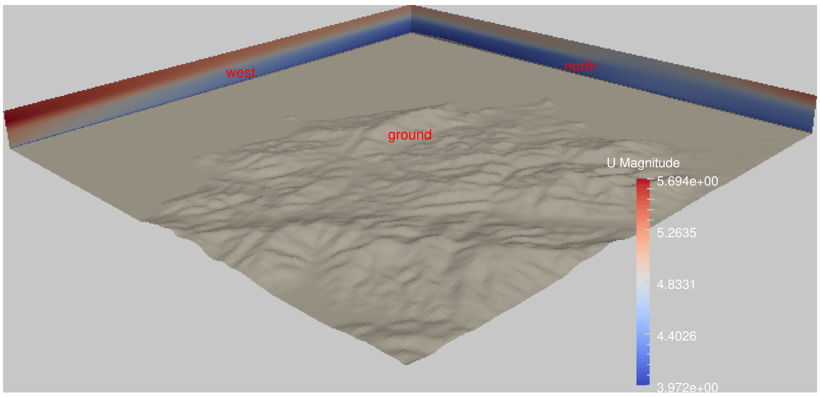

Figure 5.6 Generated boundary conditions for patches with *inlet* attribute from WRF forecasts at 00:00 UTC 01 October, 2013.

Figure 5.7 and Figure 5.8 show the simulation results of three variables. Figure 5.7 displays the information of a slice in the y direction ($x = 470307$) around the Awaji wind farm. It is easy to find from Figure 5.7 (a) that the wind velocity (shaded) over the hills tends to be significantly increased and it reaches the maximum around the top of the hills, while decreased at the lee of hills. The characteristic of streamlines indicate that there are no flow separation occurred behind the downwind hills in this case. The turbulent kinetic

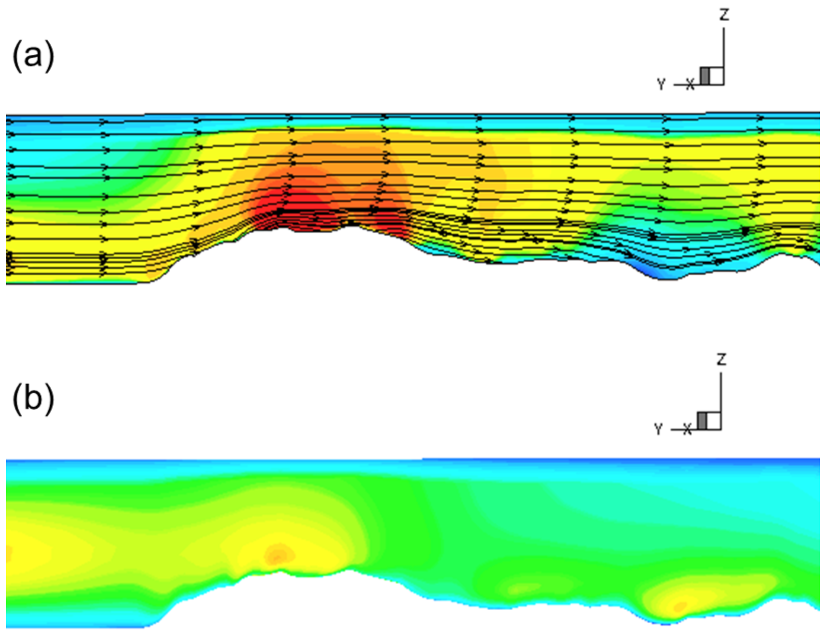

Figure 5.7 Two snapshots of a x-plane (where $x = 470307$) around the Awaji wind farm. (a) The horizontal component of the velocity (shaded) and streamlines over hills; (b) Magnitude of the turbulent kinetic energy k.

energy k produced by the mean strain rates is displayed in the Figure 5.7 (b). Clearly, the k is greatly enhanced over the hills at the first part and then generally smoothed out by the viscous dissipation as the flow moves toward downstream. Then with the terrain height higher, the increase of k at the lee of the hill is found. Those results seem to be reasonable following the theory of fluid dynamics and comparable with existing studies. More results of k and pressure over the whole ground are described in the Figure 5.8. It can be seen, there is also a y-plane in the Figure 5.8 aiming to clarify the performance of pressure and turbulent kinetic energy at the $x-z$ direction. Obviously, low pressure has been found at the top of the terrains while high pressure is found at the front of the hills (Figure 5.8 (a)). Meanwhile, Figure 5.8 (b) shows the reasonable distribution of the turbulent kinetic energy over the complex terrain.

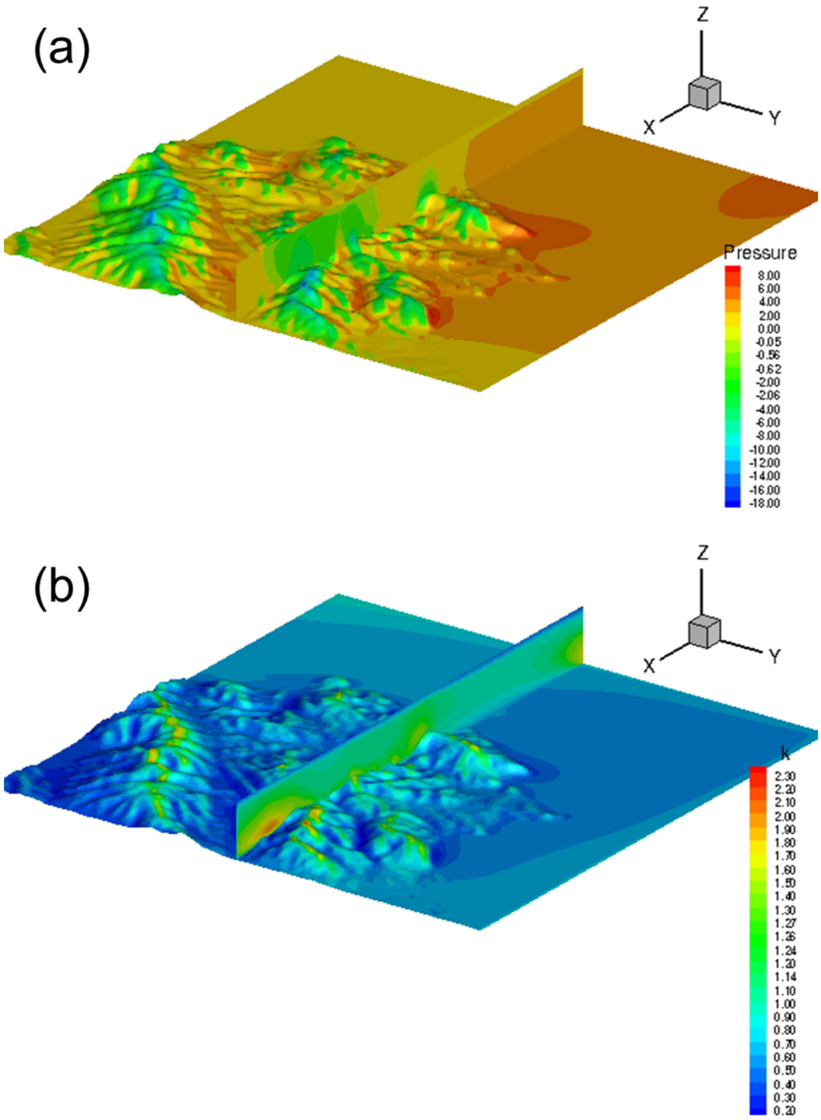

Figure 5.8 The distribution of the pressure (a) and the turbulent kinetic energy (b) over the whole ground patch as well as a y-plane where $y = 3795441$ for a real case (the dominant wind is northwestern) at 00:00 UTC 01 October, 2013.

5.3.2 Validating the ability of multi-scale forecasting system

It has been demonstrated in the above section that the coupled system especially the micro-scale component OpenFOAM has strong ability of handling the wind flow over the complex terrain. However, for the final purpose of providing operational predictions at some specific turbine sites, this forecasting system need to be further validated using the real observations collected at each turbine site. Obviously, only one case (i.e., described in the section 5.3.1) is far from enough. Thus, in order to fully validate whether the multi-scale forecasting system has an advantage against the WRF based system, additional 192 cases (from 00:00 UTC 02 October to 23:00 UTC 09 October, 2013; hourly interval) are computed and the results of the multi-scale forecasting system for 15 turbine sites are compared with the corresponding nacelle wind observations separately. The intuitive comparisons are displayed in the Figure 5.9 and Figure 5.10. In those two figures, both WRF forecasts symboled with "WRF_fore" and the prediction of the multi-scale system characterized by "WRF+OpenFOAM" are compared with the nacelle observations ("OBS""), during the period of 00:00 UTC 02 October to 23:00 UTC 09 October, 2013 (192 cases).

At first glance, it is not difficult to find that two kinds of forecasts both could capture the wind speed variations well for all 15 turbines, especially for the No.4, No.5, No.7, No.10, No.11 and No.15 turbines. For other turbines, the WRF forecasts of some cases (e.g., 30-60 and 120-150) have relatively large errors. Obviously, after coupling with the micro-scale OpenFOAM, this situation has been changed. The performance of the cases from 120-150 and 180-192 have been largely improved. Similar but relatively slight improvement for the case 30-60 also can be seen for all 15 turbines. Except for those three groups of cases, though the improvement brought by coupling technique is limited, worsening cases are barely obtained at least. Therefore, based on above discussions, we may conclude that the accuracy of wind speed forecasts of WRF based system can be significantly improved when a micro-scale OpenFOAM model is coupled. This conclusion is further confirmed from the Figure 5.11, where the statistical parameters as well as relative improvements of themselves are drawn. Firstly, we can find that the value of all three parameters ME, RMSE and CC of the WRF forecasts is in an acceptable range, due to the largest ME is around 3 m/s and the values of CC are almost larger than 60%. Despite all this, the multi-scale forecasting system indeed has an ability of improving the performance of the WRF modeling system. The distribution of the red lines in the three panels of Figure 5.11 notably demonstrates the positive effects brought by coupling the micro-scale OpenFOAM with the meso-scale WRF model.

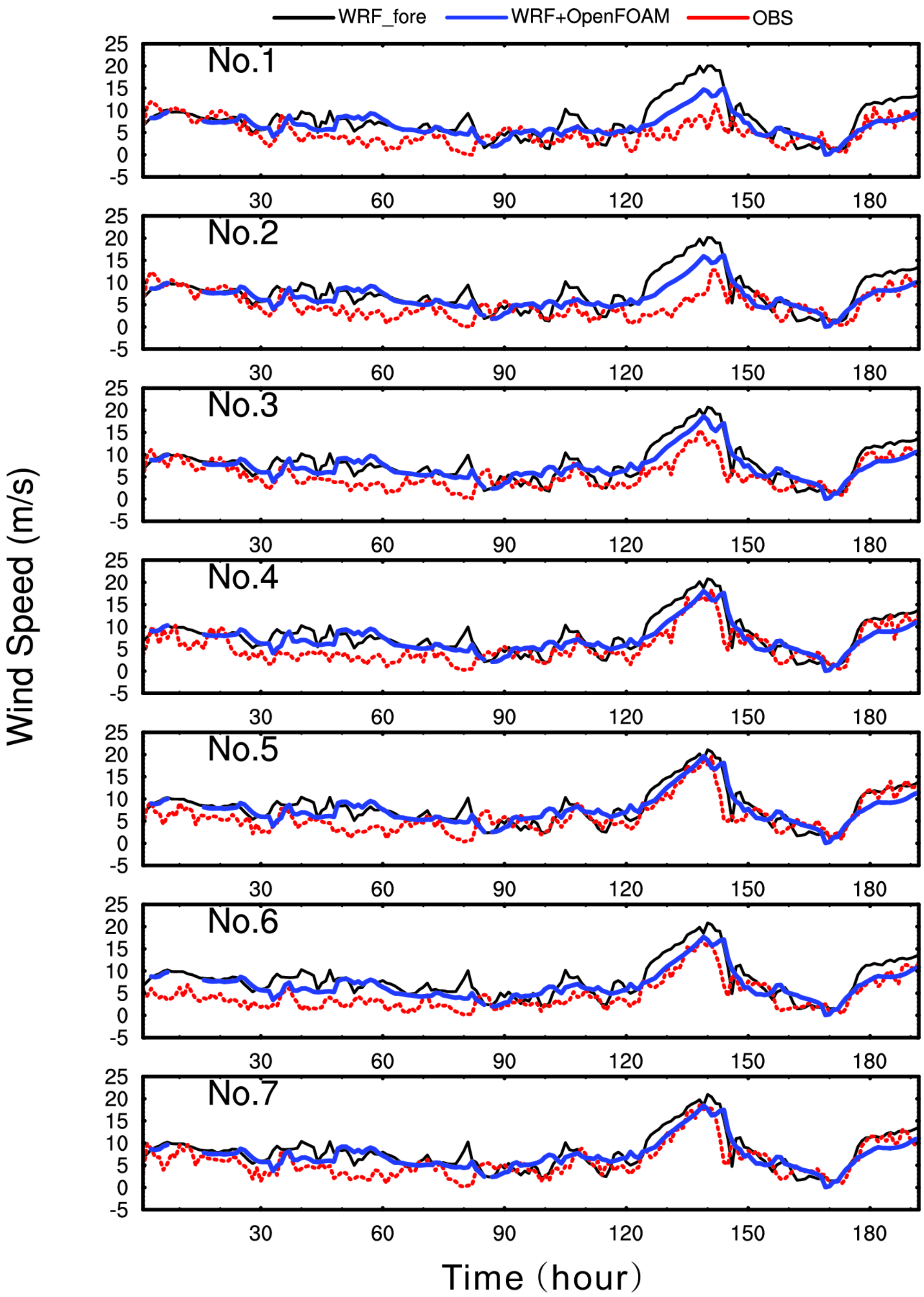

Figure 5.9 The WRF forecasts (“WRF_fore”), the predictions of multi-scale system (“WRF+OpenFOAM”) and the corresponding observations (“OBS"") of 7 turbines for 192 cases, from 00:00 UTC 02 October to 23:00 UTC 09 October, 2013.

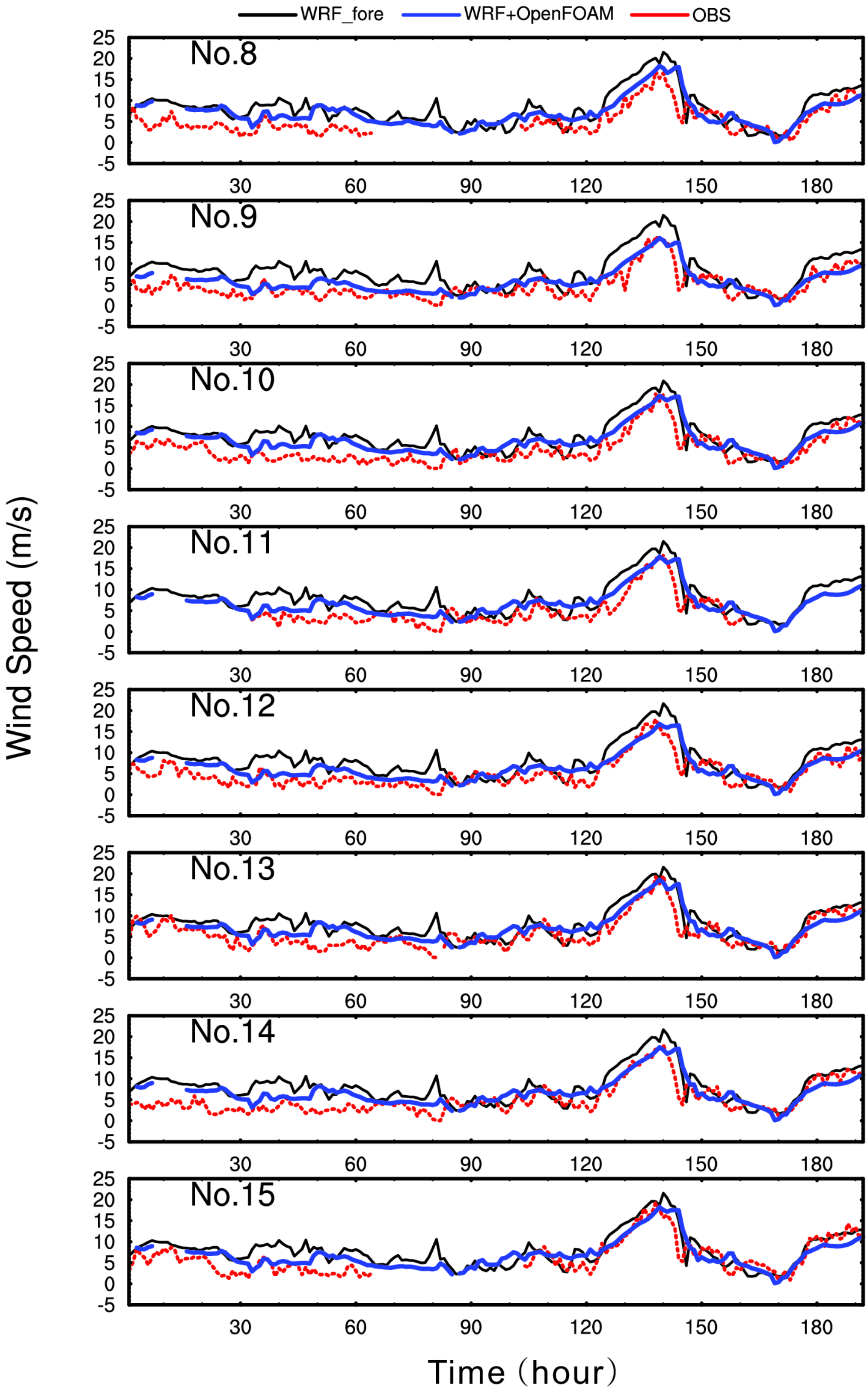

Figure 5.10 Same as the Figure 5.9, but for the turbines from No.8 to No.15.

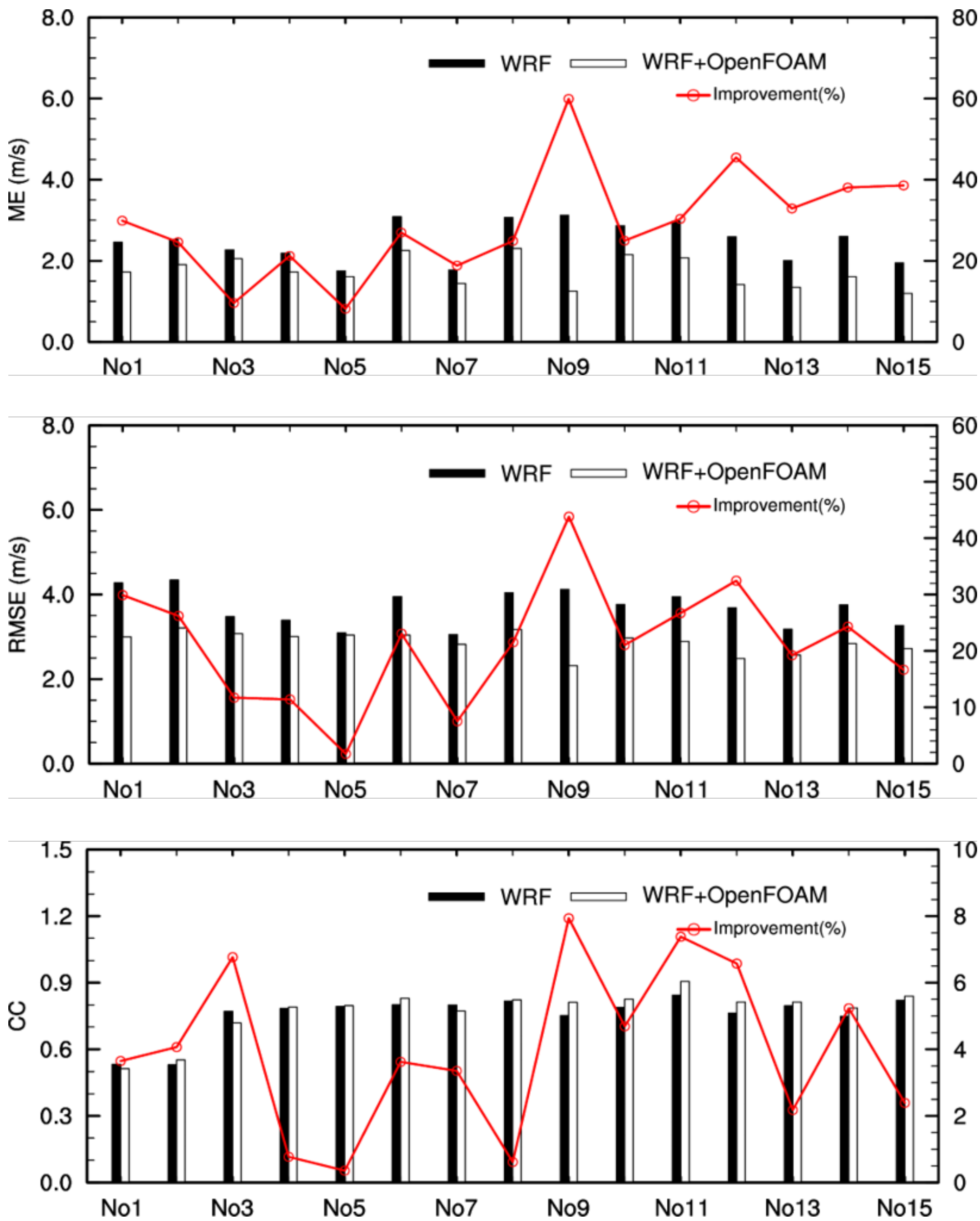

Figure 5.11 A comparison of the ME (a), RMSE (b) and CC (c) of the WRF (solid black bar) and the multi-scale system forecasts (solid white bar) for the 15 turbines of the Awaji-island wind farm. The red line stands for the relative improvement after implementing the coupling procedure against the WRF raw forecasts of the wind speed.

5.4 Summary

This chapter has described in detail the start-point and procedure of coupling the meso-scale WRF system with the micro-scale OpenFOAM model modified as a multi-scale forecasting system for operational use, to forecast short-term hub-height wind accurately for a wind farm of interest, where the geographic condition is complex.

Typically, the coupling is achieved by applying the WRF forecasted low-level wind as the initial and boundary conditions for the OpenFOAM, to calculate the steady state of the airflow affected by the local terrain feature. During this process, several useful programs have been developed to generate structured mesh and automatically change the *inlet* patches based on the prevail wind direction. This work made the OpenFOAM no longer just a simulation tool, but can be used to do the real-time prediction in the wind energy forecasting area.

The developed multi-scale forecasting system has a strong ability of predicting wind flow under the complex terrain conditions, which is not very accurate using WRF based system solely. This conclusion has been fully validated with nacelle wind observations of 15 turbines at the Awaji wind farm in Japan.

Chapter 6

Conclusions and Future work

Global wind energy capacity has been doubling nearly every three and a half years since 1990, due to its clean, renewable and sustainable characteristics. However, wind makes up only 0.56% of the total power supply in Japan by the end of year 2016, where the development of wind energy is limited by a number of factors, such as complex geographic features, high population density and the government policy. Among them, the topographical complexity may be an important reason that complicates wind flow, and thus causes greater fluctuation in power output, which makes the integration of wind power into the electric power grids more challenging than other regions. An effective solution to stabilize the wind power output is to make use of supplementary electric sources/sinks through active operations before delivering the power to the grid systems. In order to optimize the operation plan, accurate predictions of both the wind speed and wind power for the targeted turbines and wind farms are of crucial importance. However, to our best knowledge, most of the wind energy companies do not have a synthetic forecasting system yet.

Therefore, considering the specific situation in Japan, we have established a preliminary forecasting system initially for wind power prediction based on the meso-scale meteorological WRF model and a power curve. The global-scale GFS dataset is adopted as both initial and boundary conditions for the regional-scale and high resolution WRFv3.6 model through a 4-level nesting refining the horizontal grid resolution down to 1 $km \times 1$ km for the Awaji wind farm. On one hand, the WRF model has been tuned based on some basic studies in order to make it suitable for Awaji wind farm with complex terrain conditions, and the ACM2 PBL and the corresponding parameterization schemes were chosen for predicting the hub height wind speed of each turbine (15 turbines in total) in that wind farm. On the other hand, based on the historical observation of power output, a novel power curve model is proposed, which is different from the one provided by the turbine manufacturers and it can be expected to be an efficient model predicting power output for each turbine separately. Combining this developed power curve with WRF model, a basic forecasting system has been built. Its forecasting ability of both wind and power are validated by comparing with the observed wind speed and power of 15 turbines in the target wind farm, from 1 August 2013 to 31 January 2014. The results indicate this basic forecasting system indeed has a relatively high forecasting skills for both wind and power. In fact, this system has been installed by a wind energy company to provide operational prediction twice each day. However, there is no doubt that many errors and uncertainties exist in this system, for example the simplification of physical complexity, spatial/temporal discretization and errors in the initial/boundary condition, which could affect the forecasting skill of the system, for example the consistently overestimating the power in the section 2.4. Those problems lead to the possibility to further improve the prediction of wind speed and power by developing Kalman filter, data assimilation modules and coupling a CFD model.

Initially, the Kalman filter module is developed as an effective module to reduce the errors and uncertainties in the basic system we have already built. The Kalman filter actually is an estimation algorithm named after Rudolf E Kálmán, which operates recursively on streams of input data (containing random variations) to produce a statistically optimal estimate of the underlying system state. In this study, the Kalman filter is used as a predictor of forecasting errors of WRF and power model. Compared to the traditional formulation of the Kalman filter algorithm, the bias of predicted field is chosen in this study instead of the field itself. In addition, the white noises in the forecasting model and observations are estimated with the Kalman algorithm itself, which is better than the way used in the previous studies [45, 69]. After integrating with the basic forecasting system, the impacts on both wind and power forecasts against the raw forecasts derived from the basic system are validated, compared to the corresponding observations. The results show that the predicted wind field can be substantially improved by the Kalman filter as a post-processing procedure. In specific, the 15-turbine averaged improvements of ME , RMSE and CC are 97%, 22% and 10% respectively. Meanwhile, the Kalman filter also demonstrates a promising capability of reducing the uncertainties in the power curve model. More specifically, Kalman filter could significantly improve the raw model prediction of power by 92%, 33% and 15% in ME, RMSE and CC respectively. Systematic validations regarding both wind speed and power output were carried out against the observations for the target wind farm, which show that the integrated power forecasting system presented here can be an effective and practical tool for short-term predictions of wind speed and power output in Japan area.

Though the function of Kalman filter module for reducing the errors in the existing forecasting system has been demonstrated, it somehow cannot deal well with the random errors caused by many factors, for example, the inaccurate initial conditions of the WRF model. That is the main reason we try to integrate data assimilation as another key module for our forecasting system. The main function of the data assimilation is to obtain an optimal estimate of the state of a physical system by combining all available information from models, background and observations. In other words, it can provide accurate initial condition for the WRF model, which is vital to the short-term forecasts. In this thesis, one of data assimilation methods the so-called 3DVAR (3D variational) is chosen to assimilating the nacelle wind observations to improve the wind speed forecasts, to our best knowledge, which has never been done in Japan. As displayed schematically in the Figure 4.3, the final analysis field at 18:00 UTC each day was cyclically assimilated three times with a 6-hour interval. The cyclical way could take more information from the observations which are regarded as the representative of the real atmosphere state. Similar to the Kalman filter module, its role also has been validated with the historical observations. The validation results indicate that the WRF model forecasts can be markedly improved after assimilating nacelle wind data, with the relative improvements of 34%, 24% and 9% in ME, RMSE and IA respectively. It is also worth to mention that the data assimilation module can handle part of random errors which cannot be eliminated by Kalman filter module. This kind of study also indicates an import information that the nacelle wind data seems very reliable after data quality control, though the previous studies always use upwind meteorological tower measurement instead of the nacelle wind observations.

Therefore, so far, integrating those two modules with the basic system, the best forecasting performance can be obtained for the Awaji wind farm. However, as mentioned in the chapter 5, the current system only focus on capturing the relatively large scale weather or flow systems. Due to its coarse resolution (500 m of horizontal direction), the complex terrain features always limit the forecasting ability of meso-scale WRF based system for predicting wind flow around hub-height, which means coupling a smaller scale model might resolve the flow driven by the complex terrain. In this thesis, the micro-scale OpenFOAM model is chosen to build a multi-scale forecasting system by coupling with the WRF system, to forecast short-term hub-height wind accurately for a wind farm of interest, where the geographic condition is complex. We build this kind of multi-scale system assuming the CFD model could generate more detailed and reasonable information to reveal the wind flow over the complex terrain in the ABL. The Reynolds Averaged Navier-Stokes (RANS) equations combined with a turbulence model ($k-\varepsilon$), which has a fair compromise between computational burden and modeling accuracy, is adopted. Typically, the coupling is achieved by using the mesoscale WRF forecasts as the initial and boundary conditions for the micro-scale OpenFOAM. Thus, the OpenFOAM will be provided instantaneous boundary values representing the real atmosphere state. The ability of this multi-scale system for simulating wind flow the complex terrain is firstly validated with an arbitrary case. It is found that this system can capture reasonable distribution of the velocity and turbulent kinetic energy at the atmospheric boundary layer compared with other researchers' work. By using several external programs we have developed, the OpenFOAM can be used to do the real-time prediction in the wind energy forecasting area. A 8-day series of the forecasts from the multi-scale forecasting system is validated with observations at 15 turbine sites. It is found, fortunately, that this system indeed has a strong ability of forecasting wind flow under the complex terrain conditions.

In summary, we have combined the meso-scale WRF model, a novel power curve, Kalman filter, data assimilation modules and the micro-scale OpenFOAM model together to build an advanced and integrated forecasting system for wind energy forecasting under complex terrain conditions, of which the performance has been validated at the Awaji wind farm in Japan. Part of this system has been installed and used by a wind energy company in Japan to do the realistic and operational prediction, meanwhile, others also show strong potential to be realistic application. Actually, this research has not exhausted all possible experiments that could beneficially be performed. We suggest the following aspects which could provide insights and possibly lead to improvements of our current forecasting system.

- The sensitivity of turbulence models, schemes and mesh for conducting OpenFOAM component should be further tested, in order to find the best combination to possibly improve the current system;
- Uncertainties of the coefficients in the turbulence model might be studied using Ensemble Kalman filter (EnKF);
- In the process of data assimilation, we always lack of data with high quality. Thus assimilating the output from the multi-scale system we have already built in turn could possibly improve the initial conditions of the WRF model at some extent.

Bibliography

[1] Al-Hamadi, H. and Soliman, S. (2004). Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model. *Electric Power Systems Research*, 68(1):47–59.

[2] Al-Yahyai, S., Charabi, Y., and Gastli, A. (2010). Review of the use of numerical weather prediction (NWP) models for wind energy assessment. *Renewable and Sustainable Energy Reviews*, 14(9):3192–3198.

[3] Azevedo, J. M. d. S. (2013). *Development of procedures for the simulation of atmospheric flows over complex terrain, using OpenFOAM*. PhD thesis, Instituto Politécnico do Porto. Instituto Superior de Engenharia do Porto.

[4] Bannister, R. N. (2008). A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics. *Quarterly Journal of the Royal Meteorological Society*, 134(637):1971–1996.

[5] Barbounis, T. G., Theocharis, J. B., Alexiadis, M. C., and Dokopoulos, P. S. (2006). Long-term wind speed and power forecasting using local recurrent neural network models. *IEEE Transactions on Energy Conversion*, 21(1):273–284.

[6] Barker, D. M., Huang, W., Guo, Y.-R., Bourgeois, A., and Xiao, Q. (2004). A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Monthly Weather Review*, 132(4):897–914.

[7] Benjamin, S. G., Grell, G. A., Brown, J. M., Smirnova, T. G., and Bleck, R. (2004). Mesoscale weather prediction with the ruc hybrid isentropic–terrain-following coordinate model. *Monthly Weather Review*, 132(2):473–494.

[8] Blocken, B., van der Hout, A., Dekker, J., and Weiler, O. (2015). CFD simulation of wind flow over natural complex terrain: case study with validation by field measurements for

Ria de Ferrol, Galicia, Spain. *Journal of Wind Engineering and Industrial Aerodynamics*, 147:43–57.

[9] Bowen, A. (2003). Modelling of strong wind flows over complex terrain at small geometric scales. *Journal of wind engineering and industrial aerodynamics*, 91(12):1859–1871.

[10] Brown, A. R., Hobson, J., and Wood, N. (2001). Large-eddy simulation of neutral turbulent flow over rough sinusoidal ridges. *Boundary-Layer Meteorology*, 98(3):411–441.

[11] Cadenas, E., Jaramillo, O., and Rivera, W. (2010). Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method. *Renewable Energy*, 35(5):925–930.

[12] Cadenas, E. and Rivera, W. (2007). Wind speed forecasting in the south coast of Oaxaca, Mexico. *Renewable energy*, 32(12):2116–2128.

[13] Castro, F. A., Palma, J., and Silva Lopes, A. (2003). Simulation of the Askervein Flow. Part 1: Reynolds Averaged Navier–Stokes Equations (k-ε Turbulence Model). *Boundary-Layer Meteorology*, 107(3):501–530.

[14] Catalão, J. P. d. S., Pousinho, H. M. I., and Mendes, V. M. F. (2009). An artificial neural network approach for short-term wind power forecasting in Portugal. In *Intelligent System Applications to Power Systems, 2009. ISAP'09. 15th International Conference on*, pages 1–5. IEEE.

[15] Center, E. (2003). The global forecast system (gfs)-global spectral model (gsm)(gfs version 11.0. 6). Technical report, Tech. rep., Environmental Modeling Center [Disponible en http://www. emc. ncep. noaa. gov/GFS/doc. php].

[16] Charytoniuk, W. and Chen, M.-S. (2000). Very short-term load forecasting using artificial neural networks. *IEEE transactions on Power Systems*, 15(1):263–268.

[17] Chaudhari, A., Hellsten, A., Agafonova, O., and Hämäläinen, J. (2014). Large eddy simulation of boundary-layer flows over two-dimensional hills.

[18] Chen, F. and Dudhia, J. (2001). Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Review*, 129(4):569–585.

[19] Chock, G. Y. and Cochran, L. (2006). Erratum to "modeling of topographic wind speed effects in hawaii":[journal of wind engineering & industrial aerodynamics: Volume 93, issue 8, august 2005, pp. 623–638]. *Journal of Wind Engineering and Industrial Aerodynamics*, 94(3):173–187.

[20] Conan, B., Chaudhari, A., Aubrun, S., van Beeck, J., Hämäläinen, J., and Hellsten, A. (2016). Experimental and numerical modelling of flow over complex terrain: the bolund hill. *Boundary-Layer Meteorology*, 158(2):183–208.

[21] Conan, B., van Beeck, J., and Aubrun, S. (2012). Sand erosion technique applied to wind resource assessment. *Journal of Wind Engineering and Industrial Aerodynamics*, 104:322–329.

[22] Cosme, E. (2010). Data assimilation based on estimation theory: Some Fundamentals. *Lecture notes. Université Joseph Fourier, Grenoble*.

[23] Costa, A., Crespo, A., Navarro, J., Lizcano, G., Madsen, H., and Feitosa, E. (2008a). A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews*, 12(6):1725–1744.

[24] Costa, A., Crespo, A., Navarro, J., Lizcano, G., Madsen, H., and Feitosa, E. (2008b). A review on the young history of the wind power short-term prediction. *Renewable and Sustainable Energy Reviews*, 12(6):1725–1744.

[25] Costello, R., Mccoy, D., O'Donnell, P., Dutton, A., and Kariniotakis, G. (2002). Potential benefits of wind forecasting and the application of More-Care in Ireland. In *Proceedings of the 3rd MED power conference*.

[26] Council, G. W. E. (2015). Global wind report 2015. *Online: www.gwec.net/wp-content/uploads/vip/GWEC-Global-Wind-2015-Report_April-2016_22_04.pdf*.

[27] Croonenbroeck, C. and Dahl, C. M. (2014). Accurate medium-term wind power forecasting in a censored classification framework. *Energy*, 73:221–232.

[28] Croonenbroeck, C. and Stadtmann, G. (2015). Minimizing asymmetric loss in medium-term wind power forecasting. *Renewable Energy*, 81:197–208.

[29] Cutler, N. J., Outhred, H. R., and MacGill, I. F. (2012). Using nacelle-based wind speed observations to improve power curve modeling for wind power forecasting. *Wind Energy*, 15(2):245–258.

[30] Damousis, I. G., Alexiadis, M. C., Theocharis, J. B., and Dokopoulos, P. S. (2004). A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. *IEEE Transactions on Energy Conversion*, 19(2):352–361.

[31] Davies, T., Cullen, M., Malcolm, A., Mawson, M., Staniforth, A., White, A., and Wood, N. (2005). A new dynamical core for the Met Office's global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 131(608):1759–1782.

[32] Delle Monache, L., Nipen, T., Deng, X., Zhou, Y., and Stull, R. (2006). Ozone ensemble forecasts: 2. A Kalman filter predictor bias correction. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 111(D5).

[33] Delle Monache, L., Nipen, T., Liu, Y., Roux, G., and Stull, R. (2011). Kalman filter and analog schemes to postprocess numerical weather predictions. *Monthly Weather Review*, 139(11):3554–3570.

[34] Delle Monache, L., Wilczak, J., Mckeen, S., Grell, G., Pagowski, M., Peckham, S., Stull, R., Mchenry, J., and Mcqueen, J. (2008). A Kalman-filter bias correction method applied to deterministic, ensemble averaged and probabilistic forecasts of surface ozone. *Tellus B*, 60(2):238–249.

[35] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38.

[36] DiMego, G. J. (1988). The national meteorological center regional analysis system. *Monthly weather review*, 116(5):977–1000.

[37] Draxl, C., Hahmann, A. N., Peña, A., and Giebel, G. (2014). Evaluating winds and vertical wind shear from weather research and forecasting model forecasts using seven planetary boundary layer schemes. *Wind Energy*, 17(1):39–55.

[38] Dudhia, J. (1989). Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *Journal of the Atmospheric Sciences*, 46(20):3077–3107.

[39] Dudhia, J. (1996). A multi-layer soil temperature model for mm5. In *Preprints, The Sixth PSU/NCAR mesoscale model users' workshop*, pages 22–24.

[40] El-Fouly, T., El-Saadany, E., and Salama, M. (2006). Grey predictor for wind energy conversion systems output power prediction. *IEEE Transactions on Power Systems*, 21(3):1450–1452.

[41] El Kasmi, A. and Masson, C. (2010). Turbulence modeling of atmospheric boundary layer flow over complex terrain: a comparison of models at wind tunnel and full scale. *Wind Energy*, 13(8):689–704.

[42] Fan, S., Liao, J. R., Yokoyama, R., Chen, L., and Lee, W.-J. (2009). Forecasting the wind generation using a two-stage network based on meteorological information. *IEEE Transactions on Energy Conversion*, 24(2):474–482.

[43] Franke, J., Hellsten, A., Schlünzen, H., and Carissimo, B. (2007). Best practice guideline for the CFD simulation of flows in the urban environment. cost action 732: Quality assurance and improvement of microscale meteorological models. *Hamburg, Germany*.

[44] Freedman, J. M., Manobianco, J., Schroeder, J., Ancell, B., Brewster, K., Basu, S., Banunarayanan, V., Hodge, B.-M., and Flores, I. (2014). The Wind Forecast Improvement Project (WFIP): A Public/Private Partnership for Improving Short Term Wind Energy Forecasts and Quantifying the Benefits of Utility Operations. The Southern Study Area, Final Report. Technical report, AWS Truepower, LLC, Albany, NY (United States).

[45] Galanis, G., Louka, P., Katsafados, P., Pytharoulis, I., and Kallos, G. (2006). Applications of Kalman filters based on non-linear functions to numerical weather predictions. In *Annales Geophysicae*, volume 24, pages 2451–2460. Copernicus GmbH.

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

[47] Giebel, G., Landberg, L., Nielsen, T., and Madsen, H. (2001). The zephyr project-The next generation prediction system.

[48] Glahn, H. R. and Lowry, D. A. (1972). The use of model output statistics (MOS) in objective weather forecasting. *Journal of applied meteorology*, 11(8):1203–1211.

[49] Hayes, B. P., Ilie, I., Porpodas, A., Djokic, S. Z., and Chicco, G. (2011). Equivalent power curve model of a wind farm based on field measurement data. In *PowerTech, 2011 IEEE Trondheim*, pages 1–7. IEEE.

[50] Hong, S.-Y., Dudhia, J., and Chen, S.-H. (2004). A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Monthly Weather Review*, 132(1):103–120.

[51] Hong, S.-Y. and Lim, J.-O. J. (2006). The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc*, 42(2):129–151.

[52] Hong, S.-Y., Noh, Y., and Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, 134(9):2318–2341.

[53] Janjić, Z. I. (2002). Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP Meso model. *NCEP office note*, 437:61.

[54] Japan for Sustainability (2016). Current Status of Renewable Energy in Japan. *Online: www.japanfs.org/en/news/archives/news_id035124.html.*

[55] Kain, J. S. (2004). The Kain-Fritsch convective parameterization: an update. *Journal of Applied Meteorology*, 43(1):170–181.

[56] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45.

[57] Kalnay, E. (2003). *Atmospheric modeling, data assimilation and predictability*. Cambridge university press.

[58] Kariniotakis, G., Stavrakakis, G., and Nogaret, E. (1996). Wind power forecasting using advanced neural networks models. *IEEE transactions on Energy conversion*, 11(4):762–767.

[59] Kavasseri, R. G. and Seetharaman, K. (2009). Day-ahead wind speed forecasting using f-ARIMA models. *Renewable Energy*, 34(5):1388–1393.

[60] Kessler, E. (1995). On the continuity and distribution of water substance in atmospheric circulations. *Atmospheric research*, 38(1-4):109–145.

[61] Khan, A. and Shahidehpour, M. (2009). One day ahead wind speed forecasting using wavelets. In *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES*, pages 1–5. IEEE.

[62] Kim, H. G., Patel, V., and Lee, C. M. (2000). Numerical simulation of wind flow over hilly terrain. *Journal of wind engineering and industrial aerodynamics*, 87(1):45–60.

[63] Landberg, L. and Watson, S. J. (1994). Short-term prediction of local wind conditions. *Boundary-Layer Meteorology*, 70(1-2):171–195.

[64] Laprise, R. (1992). The Euler equations of motion with hydrostatic pressure as an independent variable. *Monthly weather review*, 120(1):197–207.

[65] Launder, B. E. and Spalding, D. B. (1974). The numerical computation of turbulent flows. *Computer methods in applied mechanics and engineering*, 3(2):269–289.

[66] Lazić, L., Pejanović, G., and Živković, M. (2010). Wind forecasts for wind power generation using the Eta model. *Renewable Energy*, 35(6):1236–1243.

[67] Lin, Y.-L., Farley, R. D., and Orville, H. D. (1983). Bulk parameterization of the snow field in a cloud model. *Journal of Climate and Applied Meteorology*, 22(6):1065–1092.

[68] Liu, H., Tian, H.-Q., Chen, C., and Li, Y.-f. (2010). A hybrid statistical method to predict wind speed and wind power. *Renewable energy*, 35(8):1857–1861.

[69] Louka, P., Galanis, G., Siebert, N., Kariniotakis, G., Katsafados, P., Pytharoulis, I., and Kallos, G. (2008). Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *Journal of Wind Engineering and Industrial Aerodynamics*, 96(12):2348–2362.

[70] Marras, S., Kelly, J. F., Moragues, M., Müller, A., Kopera, M. A., Vázquez, M., Giraldo, F. X., Houzeaux, G., and Jorba, O. (2015). A review of element-based galerkin methods for numerical weather prediction: Finite elements, spectral elements, and discontinuous galerkin. *Archives of Computational Methods in Engineering*, pages 1–50.

[71] Martinez, B. (2011). Wind resource in complex terrain with OpenFOAM. *Risø DTU, National Laboratory for Sustainable Energy*.

[72] Maurizi, A., Palma, J., and Castro, F. (1998). Numerical simulation of the atmospheric flow in a mountainous region of the North of Portugal. *Journal of wind engineering and industrial aerodynamics*, 74:219–228.

[73] McTaggart-Cowan, R. and Zadra, A. (2015). Representing Richardson Number Hysteresis in the NWP Boundary Layer. *Monthly Weather Review*, 143(4):1232–1258.

[74] Mellor, G. L. and Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, 20(4):851–875.

[75] Meroney, R. N. (1980). Wind-tunnel simulation of the flow over hills and complex terrain. *Journal of Wind Engineering and Industrial Aerodynamics*, 5(3-4):297–321.

[76] Miao, Y., Liu, S., Chen, B., Zhang, B., Wang, S., and Li, S. (2013). Simulating urban flow and dispersion in beijing by coupling a CFD model with the WRF model. *Advances in atmospheric sciences*, 30(6):1663.

[77] Miao, Y., Liu, S., Zheng, Y., Wang, S., and Chen, B. (2015). Numerical study of the effects of topography and urbanization on the local atmospheric circulations over the Beijing-Tianjin-Hebei, China. *Advances in Meteorology*, 2015.

[78] Milligan, M., Schwartz, M., and Wan, Y. (2003). Statistical wind power forecasting models: Results for US wind farms. *National Renewable Energy Laboratory, Golden, CO*.

[79] Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 102(D14):16663–16682.

[80] Moreno, P., Gravdahl, A. R., and Romero, M. (2003). Wind flow over complex terrain: application of linear and CFD models. In *European wind energy conference and exhibition*, pages 16–19.

[81] Mori, H. and Kurata, E. (2008). Application of gaussian process to wind speed forecasting for wind power generation. In *Sustainable Energy Technologies, 2008. ICSET 2008. IEEE International Conference on*, pages 956–959. IEEE.

[82] Neal, D., Stevenson, D., and Lindley, D. (1981). A wind tunnel boundary-layer simulation of wind flow over complex terrain: Effect of terrain and model construction. *Boundary-Layer Meteorology*, 21(3):271–293.

[83] O'Sullivan, J. (2012). *Modelling wind flow over complex terrain*. PhD thesis, ResearchSpace@ Auckland.

[84] Paiva, L. M., Bodstein, G. C., and Menezes, W. F. (2009). Numerical simulation of atmospheric boundary layer flow over isolated and vegetated hills using RAMS. *Journal of Wind Engineering and Industrial Aerodynamics*, 97(9):439–454.

[85] Panteri, E. and Papathanassiou, S. (2008). Evaluation of two simple wind power forecasting models. In *Proc. of the European Wind Energy Conf.*

[86] Pleim, J. E. (2007). A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *Journal of Applied Meteorology and Climatology*, 46(9):1383–1395.

[87] Potter, C. W. and Negnevitsky, M. (2006). Very short-term wind forecasting for Tasmanian power generation. *IEEE Transactions on Power Systems*, 21(2):965–972.

[88] Prospathopoulos, J., Politis, E., and Chaviaropoulos, P. (2012). Application of a 3d rans solver on the complex hill of bolund and assessment of the wind flow predictions. *Journal of Wind Engineering and Industrial Aerodynamics*, 107:149–159.

[89] Ramirez-Rosado, I. J., Fernandez-Jimenez, L. A., Monteiro, C., Sousa, J., and Bessa, R. (2009). Comparison of two new short-term wind-power forecasting systems. *Renewable Energy*, 34(7):1848–1854.

[90] Reynolds, O. (1894). On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Proceedings of the Royal Society of London*, 56(336-339):40–45.

[91] Salcedo-Sanz, S., Pérez-Bellido, A. M., Ortiz-García, E. G., Portilla-Figueras, A., Prieto, L., and Paredes, D. (2009). Hybridizing the fifth generation mesoscale model

with artificial neural networks for short-term wind speed prediction. *Renewable Energy*, 34(6):1451–1457.

[92] Senjyu, T., Yona, A., Urasaki, N., and Funabashi, T. (2006). Application of recurrent neural network to long-term-ahead generating power forecasting for wind power generator. In *Power Systems Conference and Exposition, 2006. PSCE'06. 2006 IEEE PES*, pages 1260–1265. IEEE.

[93] Shimada, S., Ohsawa, T., and Yatsu, K. (2009). A study on the ability of mesoscale model mm5 for offshore wind resource assessment in japanese coastal waters. In *sn, 2009 European Wind Energy Conference EWEC*, volume 2009.

[94] Shin, H. H. and Hong, S.-Y. (2011). Intercomparison of planetary boundary-layer parametrizations in the wrf model for a single day from cases-99. *Boundary-Layer Meteorology*, 139(2):261–281.

[95] Siddiqui, K., Hangan, H., and Rasouli, A. (2008). PIV technique implementation for wind mapping in complex topographies. *Measurement Science and Technology*, 19(6):065403.

[96] Sideratos, G. and Hatziargyriou, N. D. (2007). An advanced statistical method for wind power forecasting. *IEEE Transactions on power systems*, 22(1):258–265.

[97] Simmons, A. and Hollingsworth, A. (2002). Some aspects of the improvement in skill of numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 128(580):647–677.

[98] Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G. (2005). A description of the advanced research WRF version 2. Technical report, DTIC Document.

[99] Smaïli, A. and Masson, C. (2004). On the rotor effects upon nacelle anemometry for wind turbines. *Wind Engineering*, 28(6):695–714.

[100] Smith, B., Link, H., Randall, G., and McCoy, T. (2002). Applicability of nacelle anemometer measurements for use in turbine power performance tests. *Proceedings of the American Wind Energy Association (AWEA) WINDPOWER 2002*.

[101] Soman, S. S., Zareipour, H., Malik, O., and Mandal, P. (2010). A review of wind power and wind speed forecasting methods with different time horizons. In *North American Power Symposium (NAPS), 2010*, pages 1–8. IEEE.

[102] Sørensen, N. N. (1995). *General purpose flow solver applied to flow over hills*. PhD thesis, Technical University of DenmarkDanmarks Tekniske Universitet, Risø National Laboratory for Sustainable EnergyRisø Nationallaboratoriet for Bæredygtig Energi, Wind Energy DivisionAfdelingen for Vindenergi, Aeroelastic DesignAeroelastisk Design.

[103] Stensrud, D. J. and Skindlov, J. A. (1996). Gridpoint predictions of high temperature from a mesoscale model. *Weather and forecasting*, 11(1):103–110.

[104] Storm, B., Dudhia, J., Basu, S., Swift, A., and Giammanco, I. (2009). Evaluation of the weather research and forecasting model on forecasting low-level jets: implications for wind energy. *Wind Energy*, 12(1):81–90.

[105] Stull, R. B. (1988). *An Introduction to Boundary Layer Meteorology*, volume 13. Springer Science & Business Media.

[106] Sukegawa, H., Ishihara, T., Yamaguchi, A., Fukumoto, Y., Yamaguchi, A., Ishihara, T., and Fujino, Y. (2006). An assessment of offshore wind energy potential using mesoscale model. In *Proc. European Wind Energy Conference (CD-ROM), Athenes*.

[107] Sukoriansky, S., Galperin, B., and Perov, V. (2006). A quasi-normal scale elimination model of turbulence and its application to stably stratified flows. *Nonlinear Processes in Geophysics*, 13(1):9–22.

[108] Tamura, T., Cao, S., and Okuno, A. (2007). Les study of turbulent boundary layer over a smooth and a rough 2d hill model. *Flow, Turbulence and Combustion*, 79(4):405–432.

[109] Taylor, J. W. (2008). An evaluation of methods for very short-term load forecasting using minute-by-minute British data. *International Journal of Forecasting*, 24(4):645–658.

[110] Taylor, J. W., McSharry, P. E., and Buizza, R. (2009). Wind power density forecasting using ensemble predictions and time series models. *IEEE Transactions on Energy Conversion*, 24(3):775–782.

[111] Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7):7183–7192.

[112] Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., and Cuenca, R. (2004). Implementation and verification of the unified noah land surface model in the wrf model. In *20th conference on weather analysis and forecasting/16th conference on numerical weather prediction*, volume 1115.

[113] Tieleman, H. W. (2003). Roughness estimation for wind-load simulation experiments. *Journal of Wind Engineering and Industrial Aerodynamics*, 91(9):1163–1173.

[114] Torres, J. L., Garcia, A., De Blas, M., and De Francisco, A. (2005). Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Solar Energy*, 79(1):65–77.

[115] Wang, W., Bruyere, C., Duda, M., Dudhia, J., Gill, D., Lin, H., Michalakes, J., Rizvi, S., Zhang, X., Beezley, J., et al. (2010). ARW version 3 modeling system user's guide. *Mesoscale & Miscroscale Meteorology Division, National Center for Atmospheric Research, http://www. mmm. ucar. edu/wrf/users/docs/user_guide_V3/ARWUsersGuideV3. pdf*.

[116] Warner, T. T. (2011). Quality assurance in atmospheric modeling. *Bulletin of the American Meteorological Society*, 92(12):1601–1610.

[117] Warner, T. T., Peterson, R. A., and Treadon, R. E. (1997). A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bulletin of the American Meteorological Society*, 78(11):2599–2617.

[118] Welch, G. and Bishop, G. (2001). An introduction to the Kalman filter. *Proceedings of the Siggraph Course, Los Angeles*.

[119] Zack, J. (2003). Overview of wind energy generation forecasting. *Draft report for NY State Energy Research and Development Authority and for NY ISO, True Wind Solutions LLC, Albany, NY, USA*, 17.

[120] Zahle, F. and Sørensen, N. N. (2011). Characterization of the unsteady flow in the nacelle region of a modern wind turbine. *Wind Energy*, 14(2):271–283.

[121] Zajaczkowski, F. J., Haupt, S. E., and Schmehl, K. J. (2011). A preliminary study of assimilating numerical weather prediction data into computational fluid dynamics models for wind prediction. *Journal of Wind Engineering and Industrial Aerodynamics*, 99(4):320–329.