PREDICTION OF SHORT-PERIOD ENERGY PRODUCTION FOR WIND FARMS

BY

MR. NATAPOL KORPRASERTSAK

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (ENGINEERING AND TECHNOLOGY)
SIRINDHORN INTERNATIONAL INSTITUTE OF TECHNOLOGY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2019
COPYRIGHT OF THAMMASAT UNIVERSITY

Ref. code: 25625822300561TJA

PREDICTION OF SHORT-PERIOD ENERGY PRODUCTION FOR WIND FARMS

BY

MR. NATAPOL KORPRASERTSAK

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (ENGINEERING AND TECHNOLOGY)

SIRINDHORN INTERNATIONAL INSTITUTE OF TECHNOLOGY

THAMMASAT UNIVERSITY

ACADEMIC YEAR 2019

COPYRIGHT OF THAMMASAT UNIVERSITY

Ref. code: 25625822300561TJA

THAMMASAT UNIVERSITY
SIRINDHORN INTERNATIONAL INSTITUTE OF TECHNOLOGY

DISSERTATION

BY

MR. NATAPOL KORPRASERTSAK

ENTITLED

PREDICTION OF SHORT-PERIOD ENERGY
PRODUCTION FOR WIND FARMS

was approved as partial fulfillment of the requirements for
the degree of Doctor of Philosophy (Engineering and Technology)

on September 18, 2020

Chairperson _____
(Colonel Assistant Professor Anotai Suksangpanomrung, Ph.D.)

Member and Advisor _____
(Professor Thananchai Leephakpreeda, Ph.D.)

Member _____
(Associate Professor Supachart Chungpaibulpatana, D.Eng.)

Member _____
(Assistant Professor Chung-Hao Hsu, Ph.D.)

Member _____
(Assistant Professor Thunyaseth Sethaput, Ph.D.)

Director _____
(Professor Pruettha Nanakorn, D.Eng.)

Dissertation Title	PREDICTION OF SHORT-PERIOD ENERGY PRODUCTION FOR WIND FARMS
Author	Mr. Natapol Korprasertsak
Degree	Doctor of Philosophy (Engineering and Technology)
Faculty/University	Sirindhorn International Institute of Technology/ Thammasat University
Dissertation Advisor	Professor Thananchai Leephakpreeda, Ph.D.
Academic Years	2019

ABSTRACT

This dissertation introduces a novel approach to improve the reliability in the short-term prediction of power generation for wind farms. The wind forecasting models need to be improved, and appropriate methods of wind measurement are also essential to yield accurate data for wind power prediction. Therefore, in this work, the study is divided into two parts, which are the study of wind measurement and the study of wind power forecasting.

In wind measurement, the sampling rate is a crucial factor in wind data acquisition, for good accuracy in wind analysis. A high sampling rate is preferable for wind speed measurement. However, when a measurement at a high sampling rate is performed, a large amount of data is obtained for storage and computation. The Nyquist-based adaptive sampling rate method adapts the sampling rate to be the Nyquist frequency, according to actual wind conditions. In this study, the wind data at a high sampling rate of 10 Hz is used as a benchmark. The proposed Nyquist-based methodology is capable of providing high accuracy of analytical results, with percentage relative differences of less than 1% in wind analysis. In addition, the amount of wind data is significantly decreased (by 4000 times) from the benchmark.

In forecasting wind power, the predictions of the autoregressive moving average model, the artificial neural network model, and the grey prediction model are comparatively studied for wind power generation. In this study, the weighting method

Ref. code: 25625822300561TJA

systematically combines the predicted values of those three predictive models over time, based on their forecasting performance by the root mean square errors (RMSEs) between the actual values and the predicted values. The multiple forecasting models are applied to predict the wind power generation of a wind farm 1 h, 3 h, and 6 h ahead. The RMSEs of the multiple forecasting models are significantly the lowest values among those three predictive models and the benchmark by the persistence model. Furthermore, the prediction interval around the predicted value is statistically determined, to indicate the feasible range of wind power generation with a prescribed percentage of confidence under uncertainty. Uncertainty caused the historical prediction errors.

Keywords: Wind energy, Renewable energy, Energy system, Wind measurement, Wind analysis, Sampling rate, Missing data, Stochastic process, Energy planning, Time series forecasting, Neural network, Multiple forecasting models

Ref. code: 25625822300561TJA

ACKNOWLEDGEMENTS

The author would like to thank the Sirindhorn International Institute of Technology (SIIT) for the education fee, thesis support, and accommodation fees under the Excellent Thai Students (ETS) scholarship program of SIIT. The author thanks the school of manufacturing systems and mechanical engineering (MSME) for providing a working office and the equipment for conducting this research. The author thanks the research's advisor, Prof. Dr. Thananchai Leephakpreeda, including the following dissertation committee members: Assoc. Prof. Supachart Chungpaibulpatana, Asst. Prof. Anotai Suksangpanomrung, Asst. Prof. Chung-Hao Hsu, and Asst. Prof. Thunyaseth Sethaput for providing invaluable comments and suggestions which helped to achieve the research's goal, and the SIIT technician, Mr. Nikhom Meedet, for assistance in mechanic work. Finally, the author would like to give special thanks to the Sotavento experimental wind farm and relevant persons for supporting wind data and wind power generation data, which were essential for this research.

Mr. Natapol Korprasertsak

Ref. code: 25625822300561TJA

TABLE OF CONTENTS

	Page
ABSTRACT	(1)
ACKNOWLEDGEMENTS	(3)
LIST OF TABLES	(6)
LIST OF FIGURES	(7)
LIST OF SYMBOLS/ABBREVIATIONS	(9)
CHAPTER 1 INTRODUCTION	1
1.1 Statement of problem	1
1.1.1 Wind measurement	2
1.1.2 Wind forecasting	4
1.2 Objectives and scope of study	7
1.3 Significant of study	7
CHAPTER 2 REVIEW OF LITERATURE	9
2.1 Missing wind data problem	9
2.2 Conventional methods of wind measurement	10
2.3 Nyquist theorem and Fourier analysis	11
2.4 Short-term wind forecasting	12
2.5 Wind forecasts with predictive models	14
CHAPTER 3 METHODOLOGY	21
3.1 Determination of sampling rate for wind measurement	21
3.1.1 Wind analysis	24
3.1.2 Nyquist-based adaptive sampling rate	27
3.2 Short-term wind forecasting model	29

Ref. code: 25625822300561TJA

3.2.1 Persistence model 30
3.2.2 Autoregressive moving average (ARMA) model 30
3.2.2 Artificial neural network (ANN) model 32
3.2.3 Grey prediction (GP) model 34
3.3 Robust prediction of multiple forecasting model under uncertainty 35

CHAPTER 4 RESULTS AND DISCUSSION 42
4.1 Results and discussion on the study of wind measurement 42
4.1.1 Effect of sampling rate on wind data 43
4.1.2 Sensitivity analysis on parameters of Weibull distribution 44
4.1.3 Wind analysis from wind data at different sampling rates 47
4.2 Results and discussion on the study of wind power forecasting 54
4.2.1 Performance investigation of forecasting models 56
4.2.2 Prediction of power generation with multiple forecasting models 59

CHAPTER 5 CONCLUSION AND FUTURE RESEARCH 66

REFERENCES 69

APPENDICES
APPENDIX A 74
APPENDIX B 75
APPENDIX C 76
APPENDIX D 78

BIOGRAPHY 79

Ref. code: 25625822300561TJA

LIST OF TABLES

Tables	Page
2.1 Summary of literature reviews	16
4.1 Statistical analysis of 24-h wind data	44
4.2 Statistical analysis of wind data at different sampling rates (mean speed and standard deviation)	48
4.3 Statistical analysis of wind data at different sampling rates (power density, AEP, and capacity factor)	49
4.4 Statistical analysis of wind data at different sampling rates (wind direction)	51
4.5 Statistical analysis of wind data at different sampling rates (skewness and kurtosis)	52
4.6 Specification of wind turbines	55
4.7 Root mean square error of forecasting models	57
4.8 Performances of multiple forecasting models with various time horizons	60
4.9 Summary of results	64
A1 Conversion from unequally-sampled data to equivalent equally-sampled data	74
B1 Performances of ANN model with different activation functions	75

Ref. code: 25625822300561TJA

LIST OF FIGURES

Figures	Page
3.1 Wind measurement with a sampling time during a given cycle period	22
3.2 Fourier transform of wind data from (a) time domain to (b) frequency domain	23
3.3 Determination of cut-out amplitude α at maximum frequency component of $f_{\max}$	27
3.4 Collection of wind speed sampled at $f_{s,m}^{(N)}$ from wind speed data v at sampling rate f_s	28
3.5 Forecasting with ANN model	32
3.6 Schematic diagram of multiple forecasting models	39
3.7 Prediction of ARMA model, ANN model, and GP model based on actual input data.	40
3.8 Prediction of multiple forecasting models based on historical performance	40
3.9 Schematic diagram of determination of the prediction interval	41
3.10 Determination of the prediction interval, based on historical performance	41
4.1 Installation of three-cup type anemometer and wind vane on building rooftop	42
4.2 Illustration of wind speeds, sampled at a rate of 10 Hz	43
4.3 Sensitivity of Weibull parameters to changes in mean wind speed and standard deviation: (a) shape parameter with mean wind speed, (b) scale parameter with mean wind speed, (c) shape parameter with standard deviation, and (d) scale parameter with standard deviation	45
4.4 Yearly wind data at a sampling rate of 10 Hz: (a) probability distribution and (b) wind rose	46
4.5 Power curve of wind turbine for AEP calculation	47
4.6 Adaptation of sampling rate with Nyquist frequency against time	53
4.7 Location of Sotavento experimental wind farm	54
4.8 Power curves of wind turbines	55
4.9 Layout of wind turbines and meteorological masts	56

Ref. code: 25625822300561TJA

4.10 Plots of hourly energy output data against time 56
4.11 Plot of hourly predicted energy output data against time 57
4.12 Plot of weighting multiplier against time at high variance 58
4.13 Plot of weighting multiplier against time at low variance 59
4.14 One-week power generation of the wind farm 59
4.15 Hourly power generation: (a) predicted values and (b) actual values 60
4.16 Determination of number of error values n_e, at the present and past times 62
4.17 Plots of hourly predicted values and prediction intervals with 95% confidence, with respect to time 62
4.18 Plots of 3-hourly predicted values and prediction interval with 95% confidence, with respect to time 63
4.19 Plots of 6-hourly predicted values and prediction interval with 95% confidence, with respect to time 63
C1 Plots of hourly predicted values and prediction intervals with 95% confidence in continuous trend line. 76
C2 Plots of 3-hourly predicted values and prediction intervals with 95% confidence in continuous trend line. 76
C3 Plots of 6-hourly predicted values and prediction intervals with 95% confidence in continuous trend line. 77
D1 Illustration of wind speeds at various altitude. 78

Ref. code: 25625822300561TJA

LIST OF SYMBOLS/ABBREVIATIONS

Symbols/Abbreviations	Terms
α	Cut-out amplitude
AEP	Annual energy production
AGO	Accumulated generating operation
AIC	Akaike information criterion
ANN	Artificial neural network
ARMA	Autoregressive moving average
ARIMA	Autoregressive integrated moving average
BIC	Bayesian information criterion
b	Bias
C_f	Capacity factor
c	Scale parameter
DFT	Discrete Fourier transform
Δt	Sampling time
ε	White noise
e	Forecasting error
e_L	Upper limit and lower limit of forecasting errors
$\bar{e}$	Mean value of forecasting error
ϕ	Parameter of autoregressive part
F	Frequency
$f(\ )$	Probability density function
$f_{\max}$	Maximum frequency
f_s	Sampling rate
$f_s^{(N)}$	Nyquist rate
$\Gamma(\ )$	Gamma function
φ	Parameter of moving average part
GP	Grey prediction

Ref. code: 25625822300561TJA

Hz	Hertz
h	hour
IAGO	Inverse accumulated generating operation
IEC	International Electrotechnical Commission
kW	Kilowatt
kWh	Kilowatt hour
k	Shape parameter
$L(\ )$	Likelihood function
M	Total number of cycle periods
MSE	Mean square error
MW	Megawatt
$\hat{M}$	Total number of cycle periods in a year
m	meter
m/s	meter per second
min	minute
N	Number of measurement data
N_t	Number of training data
n	Total number of wind values
n_e	Number of error data
n_h	Number of hidden nodes
n_s	Total number of wind values for statistical analysis
P	Power density
P_r	Rated power output of wind turbines
P_t	Power curve of wind turbines
P_W	Power density based on Weibull probability density function
PI	Prediction interval
p	Order of autoregressive part

Ref. code: 25625822300561TJA

q	Order of moving average part
θ	Vector of autoregressive moving average model parameters
θ_{ARMA}	Parameter of ARMA model
ρ	Air density
RMSE	Root mean square error
$\text{RMSE}^{(\text{ANN})}$	Root mean square error of artificial neural network model
$\text{RMSE}^{(\text{ARMA})}$	Root mean square error of autoregressive moving average model
$\text{RMSE}^{(\text{GP})}$	Root mean square error of grey prediction model
RNN	Ridgelet neural network
σ	Standard deviation of wind speed
s	Second
s	Standard deviation of forecasting error
T	Cycle period
t	t multiplier
t	Time
$\tanh(\ )$	Hyperbolic tangent function
V	Complex number of data
V_a	Amplitude of sinusoids
v	Wind speed
v_{ci}	Cut-in wind speed
v_{co}	Cut-out wind speed
v_r	Reference wind speed
$\bar{v}$	Mean value of wind speed
W	Watt
W/m^2	Watt per square meter

Ref. code: 25625822300561TJA

Wh	Watt hour
w	Weight
$w^{(\mathrm{ANN})}$	Weighting multipliers of artificial neural network model
$w^{(\mathrm{ARMA})}$	Weighting multipliers of autoregressive moving average model
$w^{(\mathrm{GP})}$	Weighting multipliers of grey prediction model
w'	Weight of delayed outputs
X	Vector of time-series data
x	Measurement data
$x^{(1)}$	Data in accumulated generating operation series
$\hat{x}$	Predicted value
$\hat{x}^{(\mathrm{ANN})}$	Predicted value of artificial neural network model
$\hat{x}^{(\mathrm{ARMA})}$	Predicted value of autoregressive moving average model
$\hat{x}^{(\mathrm{GP})}$	Predicted value of grey prediction model
$\hat{x}^{(\mathrm{P})}$	Predicted value of persistence model
y	Output of hidden layers
y'	Delayed output of hidden layer
z	Altitude
z_r	Reference altitude

Ref. code: 25625822300561TJA

CHAPTER 1
INTRODUCTION

1.1 Statement of problem

In the present, humans consume energy from fossil fuels at a high rate. Fossil fuels are the majority of energy resources in various types of industries. Since fossil fuels, which refers to coal, petroleum, and natural gas, are not a type of energy resource that can regenerate itself, the depletion of fossil fuels due to overconsumption is a serious problem in the world, nowadays. An effective way to mitigate this critical situation is to increase the use of renewable energy resources. Renewable energy resources are considered as free energy resources. Thus, these resources are not depleted, no matter how much of them are consumed.

Wind energy is a renewable energy resource that is currently used as an alternative energy source in various applications, especially in the production of electricity. In Thailand, the high growth of electricity consumption has increasingly demanded the effective production of electricity from wind energy (Thaku & Mithulananthan, 2010). This trend occurs similarly on a worldwide scale. Recently, there has been research in harvesting wind energy under extreme wind conditions, such as low availability due to low wind speeds and minimal reliability due to the high variance of wind. However, wind power generation remains favorable because wind energy is clean and available as a free resource. Therefore, many practical techniques have been proposed to overcome the limitations of wind power generation. For example, a wind booster was proposed as an augmentation device that suitably directed wind with low speeds through guiding nozzles to efficiently impact blades of a vertical-axis wind turbine (Korprasertsak & Leephakpreeda, 2015; Korprasertsak & Leephakpreeda, 2016). The sensitivity analysis of wind power generation to design the parameters of wind turbines strongly recommended wind turbines with low cut-in speeds, to be installed for low wind speed regions (Quan & Leephakpreeda, 2015). Inconsistent wind in nature results in the uncertainty in the energy production of wind farms. Therefore, the forecasting of wind power generation under wild wind conditions is required for the preparation of real-time grid operations, ancillary service costs,

Ref. code: 25625822300561TJA

power quality, and the stability and reliability of power systems (Chang, 2014; Lund, 2005).

As wind integration grows dramatically, improvement of wind forecasting is required for the reliable operation of wind farms. The forecasting models need to be improved, and appropriate methods of wind measurement are also essential for the accuracy of data used in wind prediction. Therefore, in this work, the study is divided into two parts, which are the study of wind measurement and the study of wind power forecasting.

1.1.1 Wind measurement

In wind measurement, all-inclusive data of the speed and direction are important to accurately quantify the statistical interpretation and wind energy resources, such as wind power density and annual energy production. For wind analysis, missing wind data results in the decreased accuracy of wind analysis and wind resource assessment. The causes are failures in data acquisition during wind measurement. Up to now, most studies of missing wind data have focused on solving the missing data problems. It was reported that 10% of wind speed data, which were not recorded by a data acquisition system due to icing accretion, might cause a 3.8% bias of determination on annual energy production. A seasonality model was proposed to eliminate the bias of wind data in winter (Coville et al., 2011). Alternatively, an adaptive neuro-fuzzy inference system model was used to simulate the dynamic changes in wind velocity for interpolating missing wind data (Yang et al., 2011). With long-term observation, spectral analysis was capable of forecasting wind speed data, consisting of missing values over long horizons (Akçay & Filik, 2017).

Implicitly, missing wind data can be caused by an insufficient sampling rate in wind measurement under varying wind conditions. In other words, the wind data are collected with not enough values to obtain all-inclusive information on the dynamic behaviors of wind. Technically, this effect is called aliasing. To prevent this problem, a high sampling rate of 10 Hz was recommended. The wind speed is measured for accurate estimation of wind turbulence intensity and turbulence power spectral density in small wind turbine analysis (Tabrizi et al., 2015). Therefore, along with instrumentation in good condition, the sampling rate is one of the most crucial factors

Ref. code: 25625822300561TJA

in data acquisition, to yield genuine characteristics of wind signals. In Ref. (Kwon et al., 2016), the measurement data at short sampling rates caused severe distortions of a continuous signal, which was reconstructed from samples. Those aliasing signals are not able to form the original information of the signals. Apart from wind measurement, many studies have aliasing problems in signals, in various applications. For example, it was found that aliased content resulted in appreciable errors for the arithmetic computation of integration, degradation of radiography in medical imaging, oversight in an ultrasonic gas flowmeter, and deterioration in high accuracy measurement (Edwards, 2007; Hatagawa & Delsing, 1994; Hatagawa et al., 1997; Vasconcellos & Campos, 2012).

A high sampling rate is preferable for any application, as it captures the data without aliasing. However, when measurement at a high sampling rate is performed, a large amount of data is obtained for storage and computation. For wind measurement, the IEC 61400-12-1 standard recommends that a single sample of wind measurements be carried out per second. The wind data is averaged every 10 min (1.7×10^{-3} Hz) for recording, to reduce the amount of wind data. Statistical methods are applied for wind analysis where the wind data are usually treated as random variables in the Weibull probability distribution (Ozay & Celiktas, 2016). Although the Weibull probability distribution is widely used to represent the statistical distribution of wind data, the wind data may not always be modeled well (Akgül et al., 2016; Harris & Cook, 2014; Mohammadi et al., 2016). This study assesses how accurately the wind analysis is retrieved from the prepared data by this conventional practice. The statistical results of wind data at a high sampling rate are used as the benchmark for this investigation. For a tradeoff between accuracy and the amount of wind data, the Nyquist-based adaptive sampling rate is used to adjust the sampling rate in real-time measurement. The sampling rate is adapted to be suitable for actual wind conditions in a given cycle period by utilizing the Nyquist sampling theorem. The Nyquist sampling theorem has been widely applied in various applications that are related to data analytics in signal processing. In neuroscience, the variability of waveforms was reduced by Nyquist-based interpolation in an efficient computation time (Blanche & Swindale, 2006). In Ref. (Faust et al., 2012), the Nyquist rate was used as a benchmark for compressed samples in heart rate measurements. The Nyquist criterion was implemented to analyze

Ref. code: 25625822300561TJA

the stability of thermoacoustic systems, for when the low-order formulation was not appropriate or not available (Kopitz & Polifke, 2008). In this study, a novel and significant contribution is that the Nyquist sampling theorem is applied to systematically determine the optimal sampling rates for wind measurement under varying wind conditions. Unlike the IEC 61400-12-1 standard with a fixed sampling rate, the proposed methodology solves the excessive/missing wind data problem due to insufficient wind data causing the inaccuracy of wind analysis and energy resource assessment.

Theoretically, the sampling rate is equal to or greater than the Nyquist frequency, to prevent aliasing in signals (Diniz et al., 2002). From Fourier analysis (Bracewell, 2000), the wind data signal has harmonic components in the frequency domain. The sampling rate is chosen to maintain the dominant information of wind data while the amount is reduced. There have been studies of Fourier analysis for wind applications. For example, floating vertical-axis wind turbine aerodynamic loads were analyzed for frequency-domain characteristics (Borg & Collu, 2015). A short-time Fourier transform algorithm was used to identify the dominant frequency of wind response in buildings (Nagarajaiah & Varadarajan, 2005). In the literature, there have not been any relevant studies about: 1. the effects of sampling rate on wind measurement via spectrum analysis, and 2. Nyquist-based adaptations of the sampling rate under varying wind conditions, to improve the accuracy of wind analysis and wind resource assessment. These are the concerns of this work.

1.1.2 Wind forecasting

Recently, there have been advanced studies in wind forecasting where the predicted variables, such as the wind speed, wind direction, and wind power generation, were obtained. In Ref. (Hu & Wang, 2015), the wind speed was predicted hourly by an empirical wavelet transform coupled with the Gaussian regression method. The spatial-statistical technique was developed to estimate wind speeds (Liu et al., 2010). The effects of wind speed and direction were simultaneously analyzed for the forecasting of power generation (Dadkhah et al., 2018). The historical wind data, numeric weather prediction, and geographical property data of wind farms were used in the probabilistic forecasting of wind energy resources (Kim & Hur, 2018). A neural network was

Ref. code: 25625822300561TJA

extensively used in the prediction of wind power generation (Mason et al., 2018; Wang et al., 2018). In Ref. (Wang et al., 2016), the forecasting of a chaotic wind power series was obtained by a neural network combined with the Bernstein polynomial. In addition, fuzzy inputs to a neural network were used to forecast hourly the wind power and wind speed (Hong et al., 2010).

In short-term wind forecasting of up to 6 h ahead, the predictive models are frequently applied due to simplicity and affordability of implementation, compared with the theoretical models of atmospheric phenomena (Wang et al., 2011). The persistence model is used as a simple forecasting technique. The predicted value at the next time is assumed as the present value during a short time horizon. Actually, the predicted values of the persistence model are regarded as the benchmark, to evaluate the forecasting performance of other forecasting models (Xie et al., 2014). The predictive models, based on a statistical approach, are usually implemented for short-term wind forecasting. The fundamental Bayesian method was presented to develop an autoregressive model in 1-h ahead wind forecasting (Miranda & Dunn, 2006). In wind energy, the forecasting of wind speed and direction was employed with the autoregressive moving average (ARMA) method (Erdem & Shi, 2011; Kavasseri & Seetharaman, 2009). The ARMA model is widely used for prediction from the knowledge of past data in a time series (Lahouar & Slama, 2017). The ARMA model yields a higher accuracy of forecasting than the persistence model (Jiang et al., 2017). Due to linear characteristics, the prediction of the ARMA model is effective for the given operating conditions (Torres et al., 2005). In addition, the accuracy of the ARMA model is acceptable for a few hours ahead (Korprasertsak & Leephakpreeda, 2018). The ARMA model is limited from dealing with high-variance prediction data even though the ARMA model, integrated with differencing (ARIMA), can be used for non-stationary data (Alencar et al., 2017). In short-term wind forecasting, the data are considered to be stationary where the values of the mean and variance do not change significantly over the considered period. Besides, the degree of differencing in the ARIMA model needs to be subjectively determined, at present. The ARMA model or the ARIMA with the zero-degree integrated part is applied in the statistical analysis of time series. Recently, an artificial neural network (ANN) has been extensively used for prediction based on the best fit between the predicted value and the past data. Unlike

Ref. code: 25625822300561TJA

the ARMA model, the ANN model possesses a non-linear activation function in the model, which effectively emulates wind data with high variance (Dong et al., 2013). For instance, the Ridgelet neural network (RNN) owning ridge functions were introduced for the short-term forecasting of wind power (Amjady et al., 2011). However, the RNNs lack the memory of past events to process sequences of inputs. Their study found that the ANN model fails to predict data of power generation during calm wind conditions due to the intensely non-linear conditions. The grey prediction (GP) model is an ordinary differential equation of accumulated generating operations for predicting exponential growths of data in nature (El-Fouly, 2006; Kayacan et al., 2010, Leephakpreeda, 2008). The GP model requires fewer past data values for prediction than other forecasting models. In this study, the GP model is capable of efficiently mimicking wind data with exponential variation.

To obtain accurate forecasting, each predictive model should be selected to match the actual data of wind power generation under uncertainty over time. In other words, a single forecasting model predicts effectively for a certain period while it does not estimate well in other periods. To solve this problem, a combination of forecast models was suggested for forecasting time-series data. This can reduce large errors between the actual values and the predicted values that can be produced from a single model (Bates & Granger, 1969; Hibon & Evgeniou, 2005). This study uses the weighting method to systematically combine the predicted values of the ARMA model, the ANN model, and the GP model, based on the historical forecasting performance of each model. The multiple forecasting models are applied to predict the wind power generation of a wind farm, 1 h, 3 h, and 6 h ahead. Additionally, the prediction interval is determined statistically from historical prediction errors. The prediction interval indicates a feasible range of the actual power generation with a prescribed percentage of confidence. The predicted value with the prediction interval is meaningful, to robustly predict the future power generation under uncertainty.

The proposed methodology of multiple forecasting models is capable of reducing errors in wind prediction compared to a single statistical model. However, it should be remarked that forecasting accuracy still immensely decease at long time horizon, due to the limitation of statistics-based method. Thus, the proposed methodology can be performed effectively in short-term prediction up to 6 h ahead.

Ref. code: 25625822300561TJA

Besides, high reliable wind measurement device is required to obtain accurate data of power generation, since accurate historical error is needed in this method.

1.2 Objectives and scope of study

- To investigate the effects of sampling on the accuracy and number of recorded values in varying wind conditions.
- To provide a tradeoff between the accuracy of wind analysis and the number of recorded values by the Nyquist-based adaptive sampling rate.
- To propose a weighting method, to combine the multiple forecasting models used in predicting wind power.
- To implement the multiple forecasting model for the short-term prediction of wind power, up to 6 hours ahead.
- To determine the prediction interval, indicating the feasible range of the wind power generation with a prescribed percentage of confidence.

1.3 Significance of this study

In the field of wind measurement, this work investigates the effects of sampling rate when recording wind data on the accuracy of wind analysis and number of wind data values. The Nyquist-based adaptive sampling rate approach is proposed, to provide a tradeoff between the accuracy of wind information and number of wind records. Hence, this research yields some benefits to any application where the accuracy of wind data is important while the cost and time to deal with a large amount of data are limited. For example, wind prediction by statistical forecasting models running on a standard personnel computer is limited by the CPUs and memory. This scenario requires a high accuracy of historical wind data but not a high number of recorded data values for computation.

In wind farms, the short-term prediction of power generation is necessary for reliable operation. The multiple forecasting models are more reliable than the single forecasting model in the prediction of power generation under uncertainty of wind conditions. A weighting method to combine multiple forecasting models is proposed in this work. The proposed methodology has improved the accuracy of prediction over an individual model. In addition, this work presents a method to determine the prediction

Ref. code: 25625822300561TJA

interval around the point forecasts. The prediction interval indicates the feasible range where future data are in a prescribed percentage of confidence. This makes the prediction results more informative and more practical for applications in real-world situations.

Ref. code: 25625822300561TJA

CHAPTER 2
REVIEW OF LITERATURE

2.1 Missing wind data problem

During wind measurement, it is likely that missing wind data can occur due to many causes, such as a bad-condition measurement device, extreme weather conditions, etc. The missing wind data directly affects the accuracy of wind data. The more wind data losses during measurement, the lower the accuracy of the wind data that represent the wind in the real world. Up to now, most studies of missing wind data have focused on solving missing data problems.

Coville et al. (2011) reported that 10% of wind speed data, which were not recorded by a data acquisition system due to icing accretion, might cause a 3.8% bias of determination on annual energy production. This issue causes a bias in the estimation of wind energy resources for a potential wind turbine site. Their study proposed a seasonality-based model to diminish the bias in wind data (in winter).

Yang et al. (2011) mentioned the significance of wind data for wind farm design and planning. The missing wind data problem is found during wind measurement where it complicates wind energy resource assessment. An adaptive neuro-fuzzy inference system model was used to simulate the dynamic changes of wind velocity for interpolating missing wind data.

Akçay & Filik (2017) found that missing wind data was due to a failure of the measuring sensor. With long-term observation, spectral analysis was capable of forecasting wind speed data, consisting of missing values over long horizons. It was reported that the proposed method outperforms the persistence method.

Implicitly, missing wind data can be caused by an insufficient sampling rate in wind measurement under varying wind conditions.

Tabrizi et al. (2015) investigated the effect of sampling rate on the estimation of wind turbulence intensity and turbulence power spectral density for small wind turbine applications. Different sampling rates were tested in wind measurements. It was found that a high sampling rate of 10 Hz was recommended to measure wind speeds.

Ref. code: 25625822300561TJA

In other words, wind data are usually collected with not enough values to obtain all-inclusive information on the dynamic behaviors of wind. Technically, this effect is called aliasing. This problem is found in various applications.

Håkansson & Delsing (1994) revealed that the sampling rate is a major parameter when using an ultrasonic gas flowmeter for flow measurements. A mismatch between the sampling parameters in the flowmeter and the frequency of the pulsations can result in large errors due to aliasing.

In computed radiography for medical imaging, Hatagawa et al. (1997) stated that aliasing errors occur when the image is sparsely sampled. This causes the quality of the image to degrade on the imaging plate.

Edwards (2007) demonstrated that an aliasing problem could cause an appreciable error in the numerical integration of acceleration data to obtain estimates of velocity or displacement.

Vasconcellos & Campos (2012) made an error analysis of high-accuracy digital measurements. Their study found that the aliasing errors due to the low pass filtering performed prior to digital sampling caused the deterioration of high accuracy measurement.

Kwon et al (2016) found that measurement data at short sampling rates caused severe distortions of a continuous signal that was reconstructed from samples. Those aliasing signals are not able to form the original information of signals.

2.2 Conventional methods of wind measurement

For wind measurement, the International Electrotechnical Commission (2005) recommends that a single sample of wind measurements be carried out per second. The wind data is averaged every 10 min (1.7×103 Hz) for recording, to reduce the amount of wind data. This approach is the IEC 61400-12-1 standard.

Ozay & Celiktas (2016) applied a statistical method, the Weibull probability distribution. The wind data are usually treated as random variables in the Weibull distribution with two parameters: shape and scale. Their study supports that the Weibull distribution can provide general knowledge of wind energy potential in regions.

Ref. code: 25625822300561TJA

Nevertheless, although the Weibull probability distribution is widely used to represent the statistical distribution of wind data, the wind data may not always be modeled well.

Harris & Cook (2014) stated that the Weibull probability distribution is a purely empirical form, which cannot physically describe why the wind speed should follow the distribution.

Akgül et al. (2016) mentioned that the wind speed data might not always be effectively modeled by using the Weibull distribution in real-world situations. In other words, it may not represent all wind speed characteristics encountered in nature.

Mohammadi et al. (2016) found that by changing the methods to estimate the shape and scale parameter of the Weibull distribution, the accuracy of the calculated wind power density changes with a different method.

To sum up, the conventional method of wind measurement may not provide accurate information on wind data in certain cases. Inaccurate wind data can cause uncertainty in the parameters of wind analysis such as power density, annual energy production, and capacity factor.

2.3 Nyquist theorem and Fourier analysis

To obtain accurate wind data, a high sampling rate is preferable for any application, including wind measurement. The high sampling rate means a large amount of data is recorded. Hence, this can prevent an aliasing problem in the data. However, when measurement at a high sampling rate is performed, a large amount of data is obtained for storage and computation. Therefore, the Nyquist theorem is applied in this work for providing a tradeoff between the accuracy and the amount of wind data in real-time measurement. The Nyquist sampling theorem has been widely applied in various applications that are related to data analytics in signal processing.

Blanche & Swindale (2006) applied the Nyquist theorem in neuroscience. The variability of a waveform was reduced by Nyquist-based interpolation in an efficient computation time. Besides, the bandwidth and storage requirements can be greatly reduced by using data acquisition rates at or slightly above the Nyquist frequency.

Ref. code: 25625822300561TJA

Kopitz & Polifke (2008) implemented the Nyquist criterion for thermo-acoustic instabilities. Their study analyzed the stability of thermoacoustic systems for when the low-order formulation was not appropriate or not available.

Faust et al. (2012) used the Nyquist rate as a benchmark for compressed samples in heart rate measurements.

Fourier analysis is a useful tool for finding the frequency components of a signal. A wind data signal can be considered harmonic components in the frequency domain where the sampling rate is chosen to maintain the dominant information of the wind data while the amount is reduced.

There have been studies of Fourier analysis for wind applications.

Nagarajaiah & Varadarajan (2005) applied the short-time Fourier transform algorithm to identify the dominant frequency of wind response to buildings.

Mukhopadhyay et al. (2014) used the fast Fourier transform for wind speed data to observe the distribution of Fourier coefficients in the Fourier plane.

Borg & Collu (2015) analyzed floating vertical-axis wind-turbine aerodynamic loads in the frequency domain.

2.4 Short-term wind forecasting

In nature, the characteristics of wind are inconsistent. They cause uncertainty in the energy production by wind farms. The technique of wind forecasting is useful for providing knowledge of wind and energy production in the future, so that wind farms are able to operate reliably.

Chang (2014) mentioned that the unpredictability and variability of wind power generation due to a change of wind affect the reliability of an electric system. Effective wind forecasting can reduce risks from the unreliability of the electricity supply. The forecasting of wind power generation under wild wind conditions is required for the preparation of real-time grid operations, ancillary service costs, power quality, and the stability and reliability of power systems.

Wind forecasting can be classified based on the forecasting horizon into two categories: 1. short-term wind forecasting of up to several hours ahead and 2. long-term wind forecasting of up to several days ahead.

Ref. code: 25625822300561TJA

The methodology of wind forecasting can also be classified into two categories: 1. Physical approach and 2. Statistical approach. The physical approach is based on numerical weather prediction, which is directly obtained from the measurement of weather parameters in the atmosphere. Although the physical approach is usually more reliable, especially for long-term forecasting, than the statistical approach, it is costly to perform. In contrast, the statistical approach, which is based on knowledge of historical data, can be done with a low budget.

Wang et al. (2011) stated that the predictive models are frequently applied to short-term forecasting of up to 6 hours ahead. Using short-term models is simple and affordable to implement, compared to the theoretical models of atmospheric phenomena.

Recently, there have been advanced studies in wind forecasting where the predicted variables, such as the wind speed, wind direction, and wind power generation, were obtained.

Liu et al. (2010) developed a spatial statistical technique called "Kriging" to improve the performance of short-term wind speed prediction. The proposed method outperforms the autoregressive integrated moving average method for the mean absolute error and the root mean square error.

Hong et al. (2010) developed the multi-layer feedforward neural network with fuzzy input to forecast the hourly wind power and wind speed. It yields a lower forecasting error when compared to the traditional forecasting method.

Hu & Wang (2015) utilized an empirical wavelet transform coupled with the Gaussian regression method, to deal with the high uncertainty of wind speed data in hourly wind speed forecasting. Their proposed model can respond to wind speed changes and the forecasting environment.

Wang et al., 2016 proposed wind power prediction based on the forecasting of a chaotic wind power series using neural networks combined with the Bernstein polynomial. Their model has lower forecasting errors compared to many predictive models in the short-term prediction of wind power generation.

Dadkhah et al. (2018) analyzed simultaneously the wind speed and wind direction for power output, for the forecasting of power generation.

Ref. code: 25625822300561TJA

Kim & Hur (2018) used the ensemble method in the forecasting of wind energy resources. Their method yields predictions from historical wind data, numeric weather predictions, and geographical data of wind farms.

Mason et al. (2018) proposed a novel machine learning method called "evolutionary neural networks", based on a recurrent neural network. The proposed method yields reasonable forecasts in the prediction of wind power up to 2.5 hours ahead.

2.5 Wind forecasts with predictive models

In wind forecasting, the persistence model is the simplest forecasting technique. It is suitable for predicting the future data in a short time horizon. The model is typically regarded as the benchmark to evaluate the performance of other forecasting models.

Xie et al. (2014) used the persistence method as the benchmark to evaluate how much their proposed novel method is able to improve the short-term forecasting accuracy.

The predictive models, based on a statistical approach, are usually implemented for short-term wind forecasting. For wind energy, the forecasting of wind speed and direction was employed with the autoregressive moving average (ARMA) method. Lahouar & Slama (2017) stated that the ARMA model is one of the most widely-used models to predict the next data points in a time series. Jiang et al (2017) mentioned that an ARMA model usually yields a higher forecasting accuracy than the persistence model for a long time horizon.

Due to linear characteristics, the prediction of the ARMA model is effective for the given operating conditions. Recently, many studies employ the ARMA model in the forecasting of wind data and wind power generation.

Torres et al. (2005) forecasted the average wind speed with the ARMA model up to 10 hours ahead. The results show that the ARMA model behaves significantly better in the forecast than the persistence model. However, The ARMA model is only valid for short-term forecasting.

Kavasseri & Seetharaman (2009) used the fractional ARMA model integrated with differencing, to predict the wind speed 24 hours ahead and 48 hours ahead. It was

Ref. code: 25625822300561TJA

found that the proposed model yields a significant improvement from the persistence model.

Erdem & Shi (2011) presented predictive approaches based on the ARMA model for forecasting wind speed and wind direction. The proposed approaches were tested, to forecast wind an hour ahead for given observation sites. The results show that the approaches improve the accuracy of the traditional ARMA model.

Korprasertsak & Leephakpreeda (2018) investigated the performance of the ARMA model in wind speed forecasting for an airborne wind turbine. Their study suggests that the prediction accuracy of the ARMA model is acceptable for a few hours ahead.

However, the ARMA model is limited from dealing with high-variance prediction data. Unlike the ARMA model, the ANN model possesses a non-linear activation function in the model, which effectively emulates wind data with high variance.

Dong et al. (2013) mentioned that most of the time, the ANN model is superior to other kinds of statistical models, as many real systems, including wind data, possess nonlinearity and uncertainty.

Applications of the ANN model in wind forecasts have been widely studied. For example, Amjady et al. (2011) introduced the Ridgelet neural network (RNN) owning ridge functions for the short-term forecasting of wind power generation. The proposed RNN model provides acceptable results for the forecasting of the wind power output of wind farms and the aggregated wind generation of power systems. However, the RNNs lack the memory of past events, to process sequences of inputs.

Korprasertsak & Leephakpreeda (2018) applied ANN models in the short-term wind speed prediction of up to 6 hours. They found that the ANN model outperforms the ARMA model and the persistence model for single-step and multi-step forecasting. However, it is found in their study that the ANN model fails to predict the data of power generation during calm wind conditions.

The GP model is a mathematical model, which is based on an ordinary differential equation of an Accumulated Generating Operation (AGO) series for predicting exponential growths of data in nature. The advantage of the GP model is that it requires fewer past data values for prediction than other forecasting models. In this

Ref. code: 25625822300561TJA

study, it is observed that the GP model is capable of efficiently mimicking wind data with exponential variation. Although there are only a few studies that employ the GP model in wind forecasting, the GP model has been often used in other kinds of forecasting.

El-Fouly (2006) applied the GP model in wind energy conversion systems for output power prediction. The presented results reveal the effectiveness and the accuracy of the technique for wind speed forecasting and wind power prediction. The GP model significantly outperforms the persistence method.

Leephakpreeda (2008) employed the GP model to determine the indoor comfort temperature for an efficient heating, ventilating, and air-conditioning system under a dynamic environment. It was found that the GP model has viability in this application.

Kayacan et al. (2010) proposed a forecasting model based on the grey system theory for time-series prediction. The proposed model was used to predict the highly noisy data of the United States dollar to Euro parity.

In this chapter 2, all literature reviews by each author are summarized as listed in Table 2.1.

Table 2.1 Summary of literature reviews

Author	Main finding
Coville et al. (2011)	Some periods of wind speed data were not recorded due to icing accretion. This might cause bias of determination on annual energy production.
Yang et al. (2011)	The missing wind data problem is found during wind measurement where it complicates wind energy resource assessment.
Akçay & Filik (2017)	Missing wind data happens due to a failure of the measuring sensor.
Tabrizi et al. (2015)	A high sampling rate of 10 Hz was recommended to measure wind speeds for estimation of wind turbulence intensity and turbulence power spectral density for small wind turbine applications

Ref. code: 25625822300561TJA

Håkansson & Delsing (1994)	A mismatch between the sampling parameters in the ultrasonic flowmeter and the frequency of the pulsations can result in large errors due to aliasing.
Hatagawa et al. (1997)	Aliasing errors causes the quality of the image to degrade on the imaging plate in computed radiography for medical imaging.
Edwards (2007)	Aliasing problems could cause an appreciable error in the numerical integration of acceleration data to obtain estimates of velocity or displacement.
Vasconcellos & Campos (2012)	Aliasing errors due to the low pass filtering performed prior to digital sampling caused the deterioration of high accuracy measurement.
Kwon et al (2016)	Measurement data at short sampling rates caused severe distortions of a continuous signal that was reconstructed from samples.
International Electrotechnical Commission (2005)	A single sample of wind measurements is recommended to be carried out per second. The wind data is averaged every 10 min (1.7×10^3 Hz) for recording.
Ozay & Celiktas (2016)	Although the Weibull probability distribution is widely used to represent the statistical distribution of wind data, the wind data may not always be modeled well.
Harris & Cook (2014)	Weibull probability distribution is a purely empirical form, which cannot physically describe why the wind speed should follow the distribution.
Akgül et al. (2016)	Wind speed data might not always be effectively modeled by using the Weibull distribution in real-world situations.
Mohammadi et al. (2016)	Different methods to estimate the shape and scale parameter of the Weibull distribution yield different accuracy of the calculated wind power density.
Blanche & Swindale (2006)	In neuroscience, the variability of a waveform was reduced by Nyquist-based interpolation in an efficient computation

	time. The bandwidth and storage requirements can be greatly reduced by using data acquisition rates at or slightly above the Nyquist frequency.
Kopitz & Polifke (2008)	The author implement Nyquist criterion for thermo-acoustic instabilities.
Faust et al. (2012)	The author used the Nyquist rate as a benchmark for compressed samples in heart rate measurements.
Nagarajaiah & Varadarajan (2005)	The short-time Fourier transform algorithm to identify the dominant frequency of wind response to buildings
Mukhopadhyay et al. (2014)	Fast Fourier transform is used for wind speed data to observe the distribution of Fourier coefficients in the Fourier plane.
Borg & Collu (2015)	The author analyzed floating vertical-axis wind-turbine aerodynamic loads in the frequency domain.
Chang (2014)	The forecasting of wind power generation under wild wind conditions is required for the preparation of real-time grid operations, ancillary service costs, power quality, and the stability and reliability of power systems.
Wang et al. (2011)	Using short-term models is simple and affordable to implement, compared to the theoretical models of atmospheric phenomena.
Liu et al. (2010)	A spatial statistical technique called "Kriging" is developed to improve the performance of short-term wind speed prediction.
Hong et al. (2010)	A Multi-layer feedforward neural network with fuzzy input is developed to forecast the hourly wind power and wind speed.
Hu & Wang (2015)	Empirical wavelet transform coupled with the Gaussian regression method is used to deal with the high uncertainty of wind speed data in hourly wind speed forecasting.

Wang et al., 2016	The author proposes wind power prediction based on the forecasting of a chaotic wind power series using neural networks combined with the Bernstein polynomial.
Dadkhah et al. (2018)	The wind speed and wind direction are analyzed simultaneously for the forecasting of power generation.
Kim & Hur (2018)	The author combines predictions from historical wind data, numeric weather predictions, and geographical data of wind farms in the forecasting of wind energy.
Mason et al. (2018)	Evolutionary neural networks based on a recurrent neural network is develop for prediction of wind power.
Xie et al. (2014)	The persistence method is used as the benchmark to evaluate short-term forecasting accuracy.
Lahouar & Slama (2017)	An ARMA model is one of the most widely-used models to predict the next data points in a time series.
Jiang et al (2017)	An ARMA model usually yields a higher forecasting accuracy than the persistence model for a long time horizon.
Torres et al. (2005)	The ARMA model behaves significantly better in the wind speed forecast than the persistence model. However, The ARMA model is only valid for short-term forecasting.
Kavasseri & Seetharaman (2009)	The fractional ARMA model integrated with differencing is used to predict wind speed up to 48 h. It yields a significant improvement from the persistence model.
Erdem & Shi (2011)	Predictive approaches based on the ARMA model is developed for forecasting wind speed and wind direction.
Korprasertsak & Leephakpreeda (2018)	ARMA model is acceptable for a few hours ahead in wind speed forecasting for an airborne wind turbine.
Dong et al. (2013)	The ANN model is superior to other kinds of statistical models, as many real systems, including wind data, possess nonlinearity and uncertainty.

Ref. code: 25625822300561TJA

Amjady et al. (2011)	The Ridgelet neural network (RNN) owning ridge functions is introduced for the short-term forecasting of wind power generation.
Korprasertsak & Leephakpreeda (2018)	ANN model outperforms the ARMA model and the persistence model for single-step and multi-step forecasting. However, ANN model fails to predict the data of power generation during calm wind conditions.
El-Fouly (2006)	The GP model significantly outperforms the persistence method in wind power prediction.
Leephakpreeda (2008)	The GP model is viable to determine the indoor comfort temperature for an efficient heating, ventilating, and air-conditioning system under a dynamic environment.
Kayacan et al. (2010)	Forecasting model based on the grey system theory is proposed for time-series prediction.

Ref. code: 25625822300561TJA

CHAPTER 3
METHODOLOGY

3.1 Determination of sampling rate for wind measurement

The data of wind velocity in wind analysis is usually obtained from measurements at the anticipated location and hub height of wind turbines. After installation, a few years of wind measurement are performed, long enough for confidence in wind data with seasonal and inter-annual changes.

In the IEC 61400-12-1 standard (International Electrotechnical Commission, 2005), a 10-min recording is typically recommended for typical wind measurements. In fact, how frequently measurements are taken depends on atmospheric variations, irregular terrain, or oriented obstructions. For example, if potential sites consist of mountainous and rough lands, there may be severe wind variations within the duration of measurements. The number of measured values per unit time, which is called the sampling rate, should be implemented according to particular wind conditions for given cycle periods. A high sampling rate is employed to capture the true characteristics of wind in those cases. The extraction of relevant information from data can be completed from wind analysis.

Effective decision making on wind farm development can be compromised. However, the more frequently data are recorded, the more expensive the data acquisition, and the more data values there are to analyze. Hence, data that are collected with a well-predetermined sampling rate can be used to effectively determine wind resources, wind prediction, wind turbine selection, and economic values with quantitative and qualitative accuracy. In particular, the sampling rate must be set according to the real dynamic behaviors of wind.

From common observation, the wind speed is a more variable measurement than wind direction. A sampling rate for wind speed can be conveniently implemented for wind direction, in practice. In this study, the sampling rate of wind speed measurement is studied. Figure 3.1 shows an analog signal of wind speed, which is sampled as values of measured wind speed v at every sampling time Δt during a given cycle period T

Ref. code: 25625822300561TJA

. A set of wind speed data at time $t_0,\ldots,t_{i-1},t_i,t_{i+1},\ldots t_{n-1}$, respectively, can be written as $v=\{v_0,\ldots,v_{i-1},v_i,v_{i+1},\ldots v_{n-1}\}$ where n is the total number of wind values.

For a given sampling time Δt, the sampling rate is expressed as:

$$f_s=\frac{1}{\Delta t} \tag{3.1}$$

The total number of wind speed values is proportional to the sampling rate. It can be calculated by:

$$n=(t_{n-1}-t_0+1)f_s \tag{3.2}$$

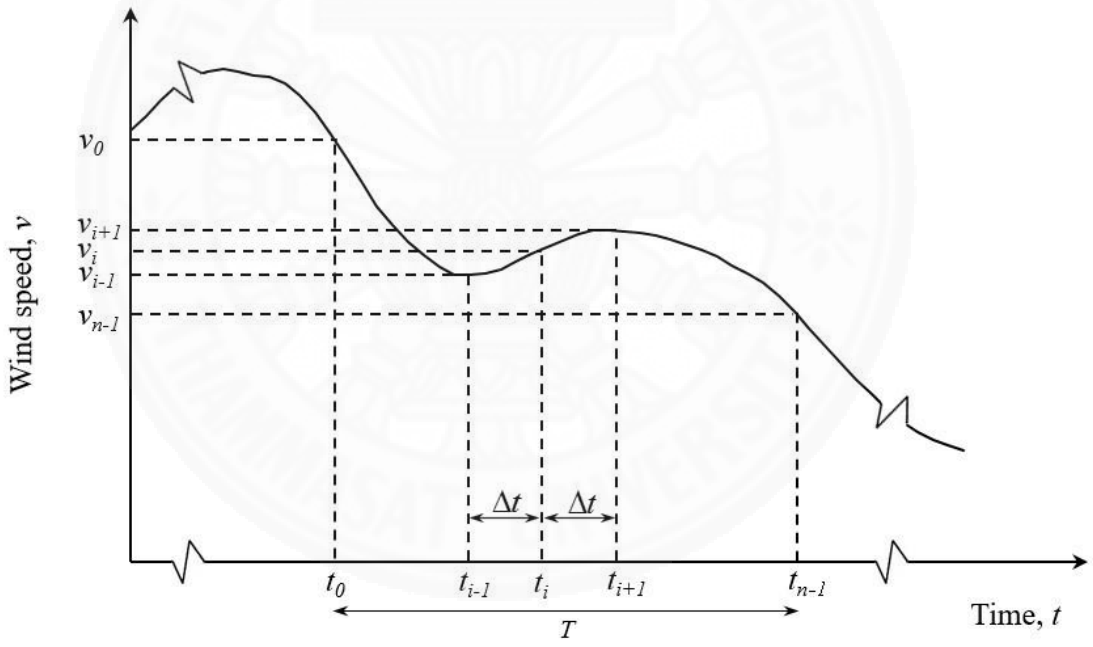

Figure 3.1 Wind measurement with a sampling time during a given cycle period.

In the time domain, there are observed effects of sampling rate on completeness of measured data for wind analysis. A signal of wind speed is transformed from the time domain to the frequency domain. The Fourier transformation is applied to a continuous-time signal of wind speed, which is explicitly decomposed into harmonic components at different frequencies, as illustrated in Figure 3.2. In other words, the

Ref. code: 25625822300561TJA

measured signal of wind speed is the algebraic sum of sinusoids with different amplitudes and frequencies.

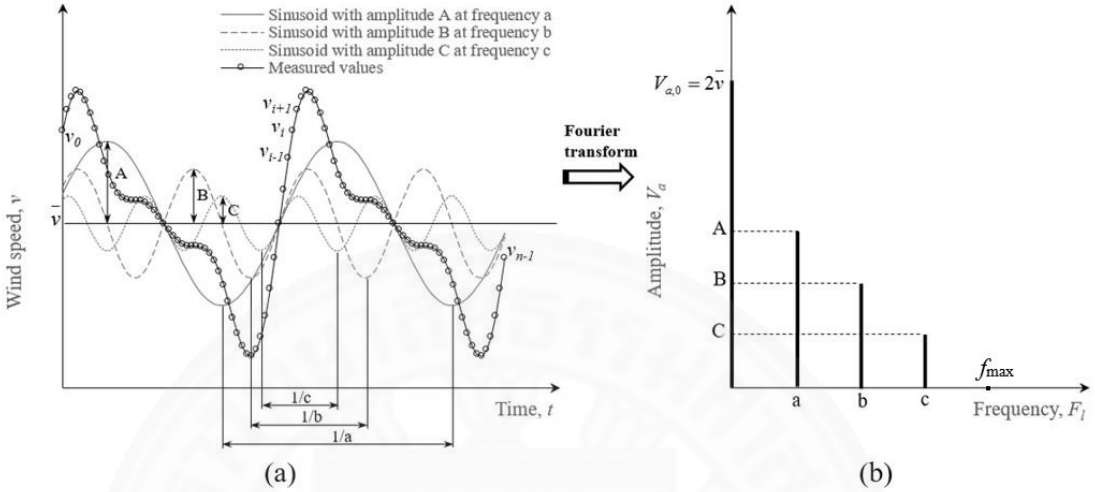

Figure 3.2 Fourier transform of wind data from **(a)** time domain to **(b)** frequency domain.

Since the wind speed is sampled at every sampling time, the Discrete Fourier Transformation is performed by using the Discrete Fourier Transform (DFT) equation over the collecting cycle period.

$$V_l = \sum_{i=0}^{n-1} v_i \left[\cos\left(\frac{2\pi i l}{n}\right) - j \sin\left(\frac{2\pi i l}{n}\right) \right] \tag{3.3}$$

where V_l is a complex number of data at l, which is from 0 to $n-1$, and $j = \sqrt{-1}$.

The amplitudes of each sinusoid $V_{a,l}$ at a generic frequency F_l is determined by:

$$V_{a,l} = \frac{2|V_l|}{n} \tag{3.4}$$

with

$$F_l = \frac{l f_s}{n} \tag{3.5}$$

Ref. code: 25625822300561TJA

In Figure 3.2(b), plots show a graphical frequency-domain representation of sinusoidal amplitudes of wind data against frequencies. To acquire the dynamic characteristics of wind, the sampling rate has to be high enough to record all dominant amplitudes up to the maximum frequency. The Nyquist sampling theorem states that all information in a measured signal with a maximum frequency component of $f_{\max}$ can be captured with a minimum sampling rate, the so-called "Nyquist rate" of:

$$f_s^{(N)} \geq 2 f_{\max} \tag{3.6}$$

In other words, an aliasing signal that is a combination of harmonics with low frequencies, can be obtained when the wind signal is sampled at less than the Nyquist rate, $f_s^{(N)}$. During wind measurement, it is necessary that the sampling rate is selected and adapted based on the Nyquist rate in accordance of wind condition in each cycle period, to be genuinely representative of wind from measurements. This is because the quality of statistical results from wind data affects the final findings of wind analysis, as explained in the next section.

3.1.1 Wind analysis

The wind has inherent variances. Its speed may vary within a second or a minute, a season of the year, and to some extent from year to year. In a statistical approach (Quan & Leephakpreeda, 2015), the wind data are treated as random variables, which are drawn from on-site measurements during a finite cycle period. Usually, the frequency of wind speed in the cycle period is not symmetric. Strong winds are relatively rare compared to moderate and weak wind speeds. The Weibull distribution is useful to govern the wind-speed frequency distribution. Unlike the normal distribution, Weibull distribution provide flexibility of the distribution curve with the shape parameter and the scale parameter. Even though similar distribution function, such as log-normal distribution and Burr distribution, may be used to fit wind speeds, Weibull distribution is the norm when fitting wind speed distribution in most wind study.

The Weibull probability density function is given as:

Ref. code: 25625822300561TJA

$$f(v)=\left(\frac{k}{c}\right)\left(\frac{v}{c}\right)^{k-1} e^{-\left(\frac{v}{c}\right)^{k}} \tag{3.7}$$

where $f(v)$ is the probability density function of wind speed v, k is the shape parameter, and c is the scale parameter.

The shape parameter and scale parameter of the Weibull distribution function can be, respectively, determined by Eqs. (3.8) and (3.9):

$$k=\left(\frac{\sigma}{v}\right)^{-1.086} \tag{3.8}$$

$$c=\frac{\bar{v}}{\Gamma\left(1+\frac{1}{k}\right)} \tag{3.9}$$

where σ is the standard deviation of wind speed, $\bar{v}$ is the mean value of wind speed, and $\Gamma(\ )$ is the gamma function.

The mean value and the standard deviation of wind speed are, respectively, determined by Eqs. (3.10) and (3.11):

$$\bar{v}=\frac{1}{n_s}\sum_{h=1}^{N} v_h \tag{3.10}$$

$$\sigma=\sqrt{\frac{1}{n_s-1}\sum_{h=1}^{N}\left(v_h-\bar{v}\right)^2} \tag{3.11}$$

where n_s is the total number of wind values for statistical analysis.

From Equation (3.7), a high value of the shape parameter indicates a small deviation from the mean wind speed, and vice versa. The scale parameter is directly proportional to the mean wind speed, as governed by Equation (3.9). The wind characteristics are defined by the shape parameter and the scale parameter, which correspond to the operating range of a wind turbine. The shape parameter and scale parameter of the Weibull distribution are important for reliability and maintainability analysis, for determining electricity generation of a wind turbine under statistical wind

Ref. code: 25625822300561TJA

conditions. In general, the power density of wind flowing through a wind turbine is determined by:

$$P = \frac{1}{2}\rho\int_{0}^{\infty} v^3 f(v)dv \tag{3.12}$$

where ρ is the air density.

With the shape parameter and scale parameter, the power density in Equation (3.12), based on the Weibull probability density function, is calculated by:

$$P_W = \frac{1}{2}\rho c^3 \Gamma\left(1+\frac{3}{k}\right) \tag{3.13}$$

The power density is the energy availability of wind, which is independent of the capacity of wind turbines. In wind energy analysis, mechanical power, which is generated from a wind turbine under annual wind probability, is transformed into electrical power via an electric generator. Annual energy production (AEP) is used to indicate how much electrical power an individual wind turbine can potentially produce in a year. The AEP is determined by:

$$\text{AEP} = 8760\int_{v_{ci}}^{v_{co}} P_t(v) f(v)dv \tag{3.14}$$

where v_{ci} is the cut-in wind speed of the wind turbine, v_{co} is the cut-out wind speed of the wind turbine, and P_t is the power curve of the wind turbine.

The capacity factor is regarded as the efficiency of electrical energy generation of a wind turbine under real wind conditions, compared with rated speed. The capacity factor is defined by:

$$C_f = \frac{\text{AEP}}{8760P_r} \tag{3.15}$$

where C_f is the capacity factor and P_r is the rated power output of the wind turbine.

Ref. code: 25625822300561TJA

From Equations (3.7) - (3.15), the accuracy and reliability of the parametric indices are dependent on the quality of wind data, which is collected from measurements. The wind data is deemed of high quality when it completely represents real wind variations, in terms of the wind speed frequency distribution. The wind speed frequency distribution is a statistical function of the mean wind speed and standard deviation, as defined in Equations (3.8) - (3.9). Therefore, it is important that the values of the measured wind velocity be sampled at a well-predetermined rate according to real wind conditions.

3.1.2 Nyquist-based adaptive sampling rate

The objective of Nyquist-based adaptive sampling rate is to determine the Nyquist rate $f_s^{(N)}$ in Equation (3.6) at given cycle period T, as illustrated in Figure 3.1. This determination can be done with Fourier analysis for wind speed v at sampling rate f_s, which is proposed in Section 3.1 from Equations (3.3) - (3.5). The sampling rate f_s is implemented for data acquisition in real time to ensure that information on all wind conditions can be extracted. To reduce the amount of wind data in the given cycle period T, unnecessary data is disregarded.

As illustrated in Figure 3.3, the approach is to remove all small amplitudes at high frequencies where the cut-out amplitude is defined as α at a maximum frequency component of $f_{\max}$. This is because those small amplitudes at high frequencies are insignificant in wind analysis.

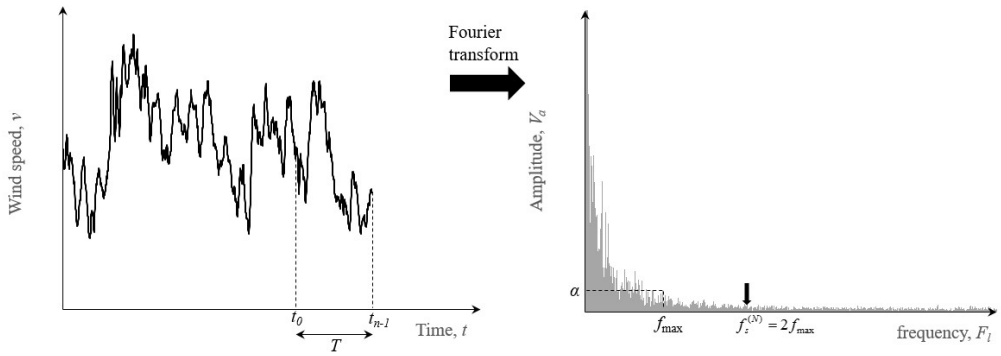

Figure 3.3 Determination of cut-out amplitude α at maximum frequency component of $f_{\max}$.

Ref. code: 25625822300561TJA

With $f_s^{(N)} \geq 2 f_{\max}$ in Equation (3.6), the Nyquist rate $f_s^{(N)}$ can be less than or equal to the sampling rate f_s. However, it is greater than or equal to $\frac{1}{T}$ since wind speed data in that cycle period must be selected. At the given cycle period T_m in time sequence m, define the wind speed $v_{\dot{p},m}$. $v_{\dot{p},m}$ is sampled from the wind data v with the Nyquist rate $f_{s,m}^{(N)}$, where $\dot{p} = \{0, \Delta t_m, 2\Delta t_m, 3\Delta t_m, \ldots, \eta_m\}$, $\Delta t_m = \frac{f_s}{f_{s,m}^{(N)}}$, and $\eta_m = (n-1) - [(n-1) \bmod \Delta t_m]$, as shown in Figure 3.4.

It should be noted that $[(n-1) \bmod \Delta t_m]$ is the remainder after division of $(n-1)$ by Δt_m.

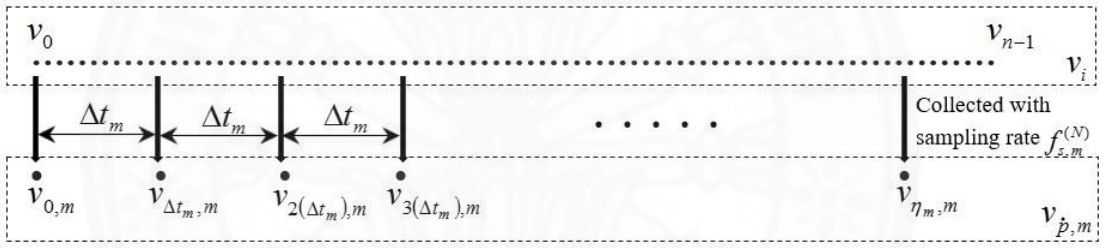

Figure 3.4 Collection of wind speed sampled at $f_{s,m}^{(N)}$ from wind speed data v at sampling rate f_s.

In the Nyquist-based adaptive sampling rate approach, the Nyquist rate $f_{s,m}^{(N)}$ is different at each cycle period T_m. With the weighed mean and standard deviation method (Watkins, n.d.), the mean wind speed and standard deviation of wind speed data are determined, respectively, by Eqs. (3.16) and (3.17):

$$\bar{v} = \frac{\sum_{m=0}^{M-1} \left(\sum_{\dot{p}=0}^{\eta_m} \frac{v_{\dot{p},m}}{f_{s,m}^{(N)}} \right)}{MT_m} \tag{3.16}$$

Ref. code: 25625822300561TJA

$$\sigma = \sqrt{\frac{\sum_{m=0}^{M-1}\left(\sum_{\dot{p}=0}^{\eta_m}\frac{\left(v_{\dot{p},m}-\bar{v}\right)^2}{f_{s,m}^{(N)}}\right)}{MT_m}} \tag{3.17}$$

where M is the total number of cycle periods, T_m.

Likewise, the power density P and AEP for a given wind turbine are, respectively, determined by Eqs. (3.18) and (3.19):

$$P = \frac{1}{2}\rho\frac{\sum_{m=0}^{M-1}\left(\sum_{\dot{p}=0}^{\eta_m}\frac{v_{\dot{p},m}^3}{f_{s,m}^{(N)}}\right)}{MT_m} \tag{3.18}$$

$$\text{AEP} = \frac{\hat{M}}{M}\sum_{m=0}^{M-1}\left(\sum_{\dot{p}=0}^{\eta_m}\frac{P_T\left(v_{\dot{p},m}\right)}{3600\times f_{s,m}^{(N)}}\right) \tag{3.19}$$

where $\hat{M}$ is the total number of cycle periods, T_m, in a year.

The influences of various sampling rates are considered by Equations (3.18) and (3.19), which are applied to determine the power density and annual energy production, respectively. However, the wind data are unequally sampled in each cycle period. This data collection form may not be applied conveniently to conventional statistical analysis when equally-sampled data is required. In Appendix A, Table A1 recommends an alternative conversion from the unequally-sampled data of wind speed to the equivalent equally-sampled data, according to a weighting factor in each cycle period.

3.2 Short-term wind forecasting models

Wind energy is usually integrated with other renewable energy sources into electricity grids for variable demands in time. Due to fluctuations in wind velocity, forecasting of wind power generation is needed to quantify the electricity production of wind turbines for optimal energy management. Forecasts can be expressed in terms of available wind power generation in MW or kWh in a given period. There are two main forecasting methods for resolving this: theoretical modeling and predictive modeling. In the former, forecasting models are structured based on the principles of

Ref. code: 25625822300561TJA

meteorological physics where numerical outcomes are obtained, according to model inputs of atmospheric conditions from observations. In the latter, forecasting models are statistically formed from time-series equations where numerical outcomes are obtained, according to model inputs of past and present outcomes. Unlike theoretical modeling, huge computation costs and advanced simulation techniques are not required for the efficiency of predictive modeling. Therefore, most well-known predictive models are selected in this study according to their advantages in short-term forecasting for wind power generation in an hour to 6 h ahead.

3.2.1 Persistence model

The persistence model is the simplest forecasting model. It is assumed that data at the next time is estimated by data at the present time. It can be mathematically written as:

$$\hat{x}_{t+\Delta t}^{(\mathrm{P})} = x_t \tag{3.20}$$

where $\hat{x}_{t+\Delta t}^{(\mathrm{P})}$ is the data at the next time $t+\Delta t$, x_t is the data at the present time t, and Δt is the sampling time, or equivalently, the forecasting horizon.

From using only the data at the present time, the persistence model lacks knowledge about the tendency or rate of change in data for a prediction (with accuracy) at a long forecasting horizon. The persistence model can be applied with acceptable accuracy when a short forecasting horizon is needed since the data at the next time is potentially close to the present data. The performance from the predicted results is used as a benchmark of the worst case within the same forecasting horizon.

3.2.2 Autoregressive moving average (ARMA) model

The ARMA model is a time-series model where the data at the next time has statistical correlation to the past and present data. The ARMA model has an autoregressive part and white noise with a moving average part (Box et al., 1994). In the autoregressive part, data at the next time can be extrapolated from linear regression of the past and present data in a time series. The moving average part involves the linear

Ref. code: 25625822300561TJA

regression of white noise, causing deviation of prediction. The $\mathrm{ARMA}(p,q)$ model is written as:

$$\hat{x}_{t+\Delta t}^{(\mathrm{ARMA})} = \sum_{i=1}^{p} \phi_i x_{t-(i-1)\Delta t} + \sum_{i=1}^{q} \varphi_i \varepsilon_{t-(i-1)\Delta t} + \varepsilon_{t+\Delta t} \tag{3.21}$$

where ϕ is the parameter of the autoregressive part, φ is the parameter of the moving average part, and ε is the zero-mean white noise. p is the order of autoregressive part, or equivalently, the total number of input data at the present time and past time. q is the order of the moving average part.

The forecasting performance of the ARMA model is based on how to find the best curve fit to a set of data, corresponding to the parameters of the ARMA model. Akaike information criterion (AIC) and Bayesian Information Criterion (BIC) are viable methods to determine the parameters of ARMA model. However, AIC has tendency of overfitting the dataset compared to the other, theoretically. Hence, the orders p and q are determined by using the Bayesian Information Criterion (BIC) where the performance index, $\mathrm{BIC}(p,q)$, in Equation (3.22) is minimized (Breid et al., 1991).

$$\mathrm{BIC}(p,q) = N_t \log(\mathrm{MSE}) + (p+q+1)\log(N_t) \tag{3.22}$$

where MSE is the mean square error between the actual data and predicted data at the next time, and N_t is the number of training data values for parametric determination.

The parameters of the ARMA model, ϕ and φ, are determined using the maximum likelihood estimation, which is expressed as Equation (3.23). The likelihood function $L(\theta, X)$ is defined as the probability density function of θ for a given X (Claeskens & Hjort, 2008).

$$\theta_{\mathrm{ARMA}} = \arg\max_{\theta} L(\theta, X) \tag{3.23}$$

where θ is the vector of model parameters, $\{\phi_1, \phi_2, \ldots, \phi_p, \varphi_1, \varphi_2, \ldots, \varphi_p\}$, X is the vector of time-series data, $\{x_t, x_{t-\Delta t}, x_{t-2\Delta t}, \ldots, x_{t-(p-1)\Delta t}\}$, and θ_{ARMA} is the parameters of the ARMA model.

Ref. code: 25625822300561TJA

The ARMA model is used for information on the behavior of data that slightly deviates from a mean value at each forecasting horizon. In other words, the ARMA model is valid in the vicinity of the mean value. If the mean values vary widely, other high-order models must be used.

3.2.3 Artificial neural network (ANN) model

The ANN model is a mathematical model for computing quantitative information, based on a collection of high-order functional operations that mimic the learning of neurons in the human brain (Atthajariyakul & Leephakpreeda, 2005). The recurrent neural network is proposed as an ideal type for approximating time-series data. As shown in Figure 3.5, the ANN model is inputted with the past and present data from an input layer through a hidden layer. At the hidden layer, the delayed outputs of the hidden layer are also cycled, as memory in the ANN model. Subsequently, the data at the next time is determined at an output layer. The numerical relations between the past and present data and data at the next time are captured by the weights and biases of the ANN model.

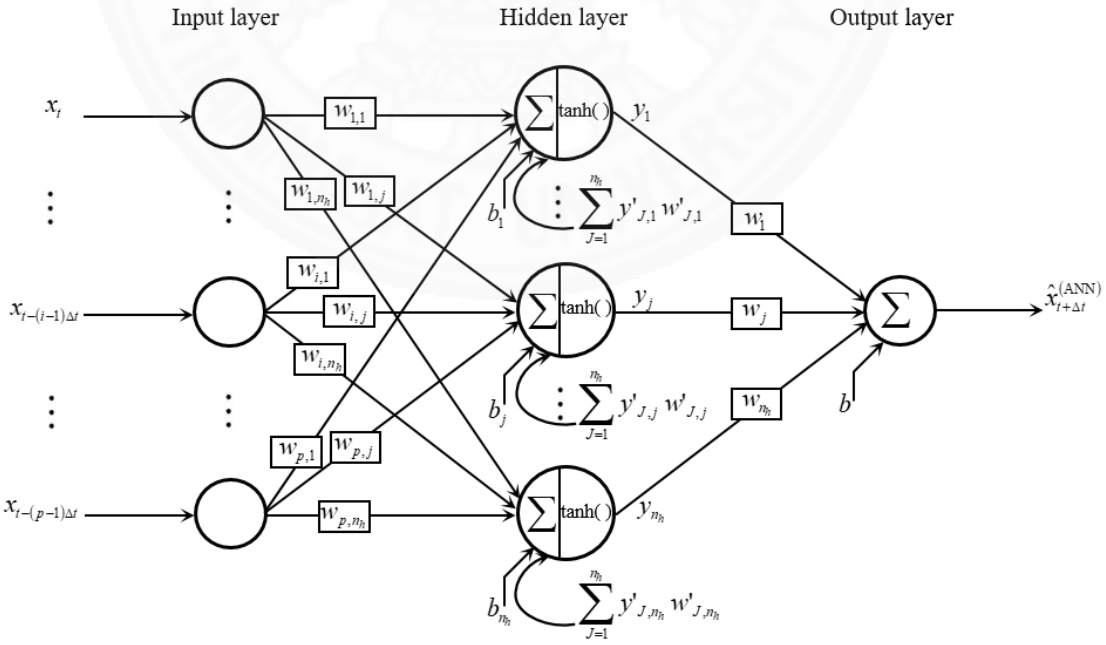

Figure 3.5 Forecasting with ANN model.

Ref. code: 25625822300561TJA

With past and present data $\{x_t, x_{t-\Delta t}, \ldots, x_{t-i\Delta t}, \ldots, x_{t-(p-1)\Delta t}\}$, the output y_i at each hidden node j in the hidden layer is obtained by:

$$y_j = \tanh\left(\sum_{i=1}^{p} x_{t-(i-1)\Delta t} w_{i,j} + \sum_{J=1}^{n_h} y'_{J,j} w'_{J,j} + b_j \right) \tag{3.24}$$

where $w_{i,j}$ is the weight between input node i and hidden node j, n_h is the number of hidden nodes j in the hidden layer, $y'_{J,j}$ is the delayed output of the hidden layer from node J to node j, $w'_{J,j}$ is the weight of the delayed output between hidden node J and hidden node j, and b_j is the bias of hidden node j.

In the hidden layer of the ANN model, the hyperbolic tangent function is selected from a performance investigation of the activation functions: hyperbolic tangent function, log-sigmoid function, rectified linear units' function, and softmax function. The setup details of the ANN model are listed in Table B1 in Appendix B.

From the hidden layer to the output layer, the data at the next time is determined by:

$$\hat{x}_{t+\Delta t}^{(\mathrm{ANN})} = \sum_{j=1}^{n_h} y_j w_j + b \tag{3.25}$$

where w_j is the weights between hidden node j and the output node, and b is the bias of the output node.

The weights and biases of the ANN model in Equations (3.24) and (3.25) are determined by the Levenberg-Marquardt backpropagation algorithm. This algorithm minimizes the mean squared errors between the predicted data and actual data at the next time for given past and present data. The ANN model is iteratively trained with sets of those predetermined input/output data (Levenberg, 1944; Marquardt, 1963). The ANN model is applied for data with large variation; therefore, it may fail to predict data at the next time that has a dynamic behavior around a given mean value due to overfitting of the model.

Ref. code: 25625822300561TJA

3.2.4 Grey prediction (GP) model

The GP model is a mathematical model, which is based on an ordinary differential equation of an Accumulated Generating Operation (AGO) series. It is used for predicting exponential growths of data in nature (Deng, 1989). If the input data at the present time and past time of the GP model is obtained as $\left\{x_t, x_{t-\Delta t}, \ldots, x_{t-(i-1)\Delta t}, \ldots, x_{t-(p-1)\Delta t}\right\}$, then the AGO series of the input data is defined as $\left\{x_t^{(1)}, x_{t-\Delta t}^{(1)}, \ldots, x_{t-(i-1)\Delta t}^{(1)}, \ldots, x_{t-(p-1)\Delta t}^{(1)}\right\}$ with:

$$x_{t-(i-1)\Delta t}^{(1)} = \sum_{j=i}^{p} x_{t-(j-1)\Delta t} \tag{3.26}$$

The ordinary differential equation for the GP model is written as:

$$\frac{dx^{(1)}}{dt} + \dot{a}x^{(1)} = \dot{b} \tag{3.27}$$

with

$$\begin{bmatrix} \dot{a} \\ \dot{b} \end{bmatrix} = \left(X^T X\right)^{-1} X^T Y \tag{3.28}$$

$$X = \begin{bmatrix} \dfrac{x_{t-(p-2)\Delta t}^{(1)} + x_{t-(p-1)\Delta t}^{(1)}}{2} & 1 \\ \dfrac{x_{t-(p-3)\Delta t}^{(1)} + x_{t-(p-2)\Delta t}^{(1)}}{2} & 1 \\ \vdots & \vdots \\ \dfrac{x_t^{(1)} + x_{t-\Delta t}^{(1)}}{2} & 1 \end{bmatrix} \tag{3.29}$$

$$Y = \begin{bmatrix} x_{t-(p-2)\Delta t} \\ x_{t-(p-3)\Delta t} \\ \vdots \\ x_t \end{bmatrix} \tag{3.30}$$

The Inverse Accumulated Generating Operation (IAGO) is applied to Equation (3.27). The data at the next time is calculated by:

Ref. code: 25625822300561TJA

$$\hat{x}_{t+\Delta t}^{(\mathrm{GP})} = \left(x_{t-(p-1)\Delta t} - \frac{\dot{b}}{\dot{a}} \right) e^{-\dot{a}p} \left(1 - e^{\dot{a}} \right) \tag{3.31}$$

It should be noted that Equation (3.28) is valid for a mathematical operation when the number of data values p must not be less than three. For observations in this study, the number of data values is from four to eight. In addition, the GP model in Equation (3.31) requires a few model parameters to predict the data at the next time. The GP model has a short computation time. In addition, the GP model is implemented as a forecasting model in this study since it is mostly neutral to the dynamic behaviors of data, which can be governed by the rate of change or a differential equation in time. In other words, the GP model is capable of predicting data at the next time under moderate variations, to bridge the gap between the ARMA model and the ANN model.

3.3 Robust prediction of multiple forecasting model under uncertainty

For the conventional forecasting procedure, predictive models such as the persistence model, ARMA model, ANN model, and GP model, are individually applied to predict the data at the next time, according to the input data at the present time and past time. Furthermore, there are comparative studies of forecasting performance among those predictive models under different conditions where the persistence model is usually applied as the benchmark but not as a usable forecasting model. In Section 3.2, the forecasting performance of each model is dependent on the allowable capability of the functional characteristics in describing the dynamic behaviors of data at that time. Therefore, the real-time implementation of a single model may be applicable to data prediction at one time, but it may not be effective at another time. To overcome this problem, a prediction from multiple forecasting models is proposed, to determine a single number taken as a balanced representation of those predictive values by using weighting multipliers according to historical forecasting performance. This technique quantitatively improves the accuracy of the multiple forecasting models from accessible knowledge of the uncertainty of each model in the past. The predicted value $\hat{x}_{t+\Delta t}$ of multiple forecasting models is determined by:

Ref. code: 25625822300561TJA

$$\hat{x}_{t+\Delta t} = w^{(\mathrm{ARMA})}\hat{x}^{(\mathrm{ARMA})}_{t+\Delta t} + w^{(\mathrm{ANN})}\hat{x}^{(\mathrm{ANN})}_{t+\Delta t} + w^{(\mathrm{GP})}\hat{x}^{(\mathrm{GP})}_{t+\Delta t} \tag{3.32}$$

where $w^{(\mathrm{ARMA})}$, $w^{(\mathrm{ANN})}$, and $w^{(\mathrm{GP})}$ are the weighting multipliers of the ARMA model, ANN model, and GP model, respectively.

The weighting multiplier of the ARMA model $w^{(\mathrm{ARMA})}$ is determined by the root mean square errors of the ARMA model, $\mathrm{RMSE}^{(\mathrm{ARMA})}$, ANN model, $\mathrm{RMSE}^{(\mathrm{ANN})}$, and GP model, $\mathrm{RMSE}^{(\mathrm{GP})}$, as written in Equation (3.33). The RMSE is used as a performance index. The accuracy of the forecasting models is indicated by the RMSEs that are used in prediction.

$$w^{(\mathrm{ARMA})} = \frac{\left(\mathrm{RMSE}^{(\mathrm{ARMA})}\right)^{-1}}{\left(\mathrm{RMSE}^{(\mathrm{ARMA})}\right)^{-1} + \left(\mathrm{RMSE}^{(\mathrm{ANN})}\right)^{-1} + \left(\mathrm{RMSE}^{(\mathrm{GP})}\right)^{-1}} \tag{3.33}$$

with the $\mathrm{RMSE}^{(\mathrm{ARMA})}$, which is expressed as:

$$\mathrm{RMSE}^{(\mathrm{ARMA})} = \sqrt{\frac{1}{p}\sum_{i=1}^{p}\left(x_{t-(i-1)\Delta t} - \hat{x}^{(\mathrm{ARMA})}_{t-(i-1)\Delta t}\right)^2} \tag{3.34}$$

where x is the actual value of $\hat{x}$ from the ARMA model.

Likewise, the weighting multipliers and RMSEs of the ANN model and GP model are determined by changing the numerators corresponding to the predictive models in Equations (3.33) and (3.34), respectively. Equation (3.33) can be interpreted as: the less the magnitude of $\mathrm{RMSE}^{(\mathrm{ARMA})}$, the more the weighting multiplier $w^{(\mathrm{ARMA})}$ contributes to the predicted value in Equation (3.32), and vice versa. Therefore, the use of weighting multipliers for the predicted values compensates for the uncertainty in forecasting, based on historical performance around the present time.

ARMA model, ANN model, and GP model are widely-used statistical method for wind prediction. The idea of multiple forecasting models with weighting multiplier is to raise the strength and diminish the weakness of each predictive model according to the wind condition in nature. With combined prediction from ARMA model, ANN model, and GP model, the model can be assumed to deal with wind at any speed and variation, effectively. However, it should be remarked that the proposed methodology

Ref. code: 25625822300561TJA

of multiple forecasting model is not limited to ARMA model, ANN model, and GP model. Other predictive model can be added or substituted, in the same manner of Equation 3.2.

To be confident in forecasting wind power generation under uncertainty, the prediction interval is reported with a prescribed percentage of confidence. The level of confidence can be determined from statistical occurrences of the forecasting error. Define the forecasting error as the difference between the actual value x_t and the predicted value $\hat{x}_t$ at a given time t, as written in Equation (3.35).

$$e_t = x_t - \hat{x}_t \tag{3.35}$$

The mean value of the forecasting error and the standard deviation of the forecasting error can be, respectively, defined by Eqs. (3.36) and (3.37):

$$\bar{e} = \frac{1}{n_e}\sum_{i=1}^{n} e_{t-(i-1)\Delta t} \tag{3.36}$$

$$s = \sqrt{\frac{1}{n_e - 1}\sum_{i=1}^{n}\left(e_{t-(i-1)\Delta t} - \bar{e}\right)^2} \tag{3.37}$$

where n_e is the number of error data values.

The upper limit and lower limit of forecasting errors for $\bar{e}$ are determined as:

$$e_L = \bar{e} \pm \mathrm{t}_{\beta/2,n_e-1} s\sqrt{1+\frac{1}{n_e}} \tag{3.38}$$

where $\mathrm{t}_{\beta/2,n_e-1}$ is the t multiplier, corresponding to the percentage of confidence 100%(1-β) with $n_e - 1$ degrees of freedom.

Hence, the prediction interval $\mathbf{PI}$ is written as:

$$\mathrm{PI}(\beta) = \hat{x}_{t+\Delta t} + e_L \tag{3.39}$$

Under uncertainty in forecasting, the prediction interval is statistically determined from the most recent errors at the present time and past time. This yields

Ref. code: 25625822300561TJA

the upper value and lower value of the predictive value, which is expected to have a specified probability and a prescribed percentage of confidence.

Figure 3.6 shows a schematic diagram of the proposed methodology in determining the predicted values of multiple forecasting models. In step 1, a forecasting horizon is specified as a requirement of a prediction ahead of time, e.g., one hour ahead of the present time. A set of N measurement data values of power generation from a wind farm is collected to determine orders p and q, model parameters of the ARMA model, weights, and biases of the ANN model. In step 2, the order p indicates the total number of input data values at the present time and past time for the ARMA model, ANN model, and GP model, as graphically presented in Figure 3.7. The p input data values are used by each model to determine the predicted values of the ARMA model, ANN model, and GP model. In step 3, the predicted values are obtained, and they also stored as the predicted values at the past time after passing the time delay. The predicted values at the past time are used to determine the weighting multipliers. Finally, the predicted values of the ARMA model, ANN model, and GP model (including the corresponding weighting multipliers) are used to determine the predicted values of the multiple forecasting models, as illustrated in Figure 3.8.

Ref. code: 25625822300561TJA

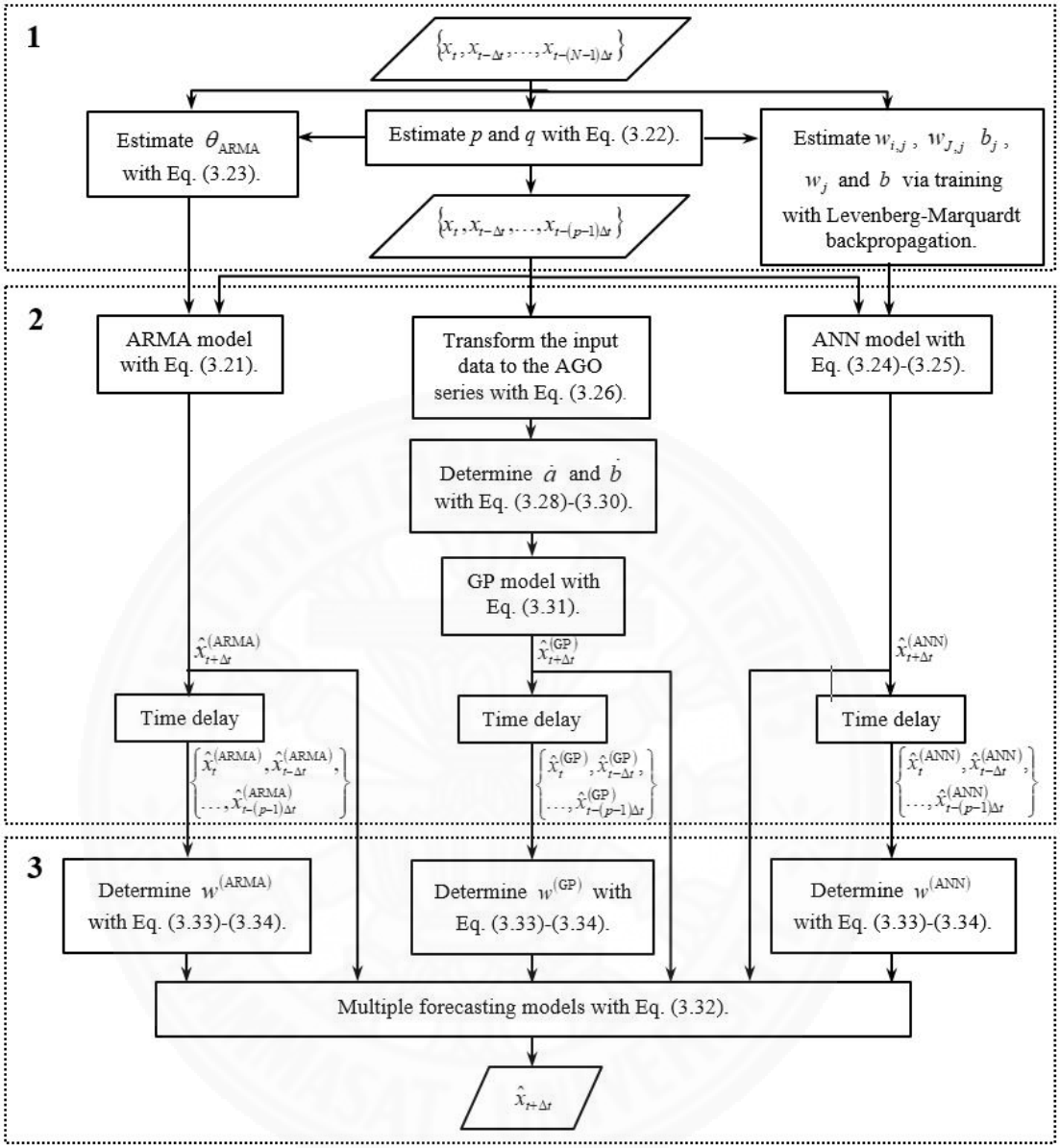

Figure 3.6 Schematic diagram of multiple forecasting models.

Ref. code: 25625822300561TJA

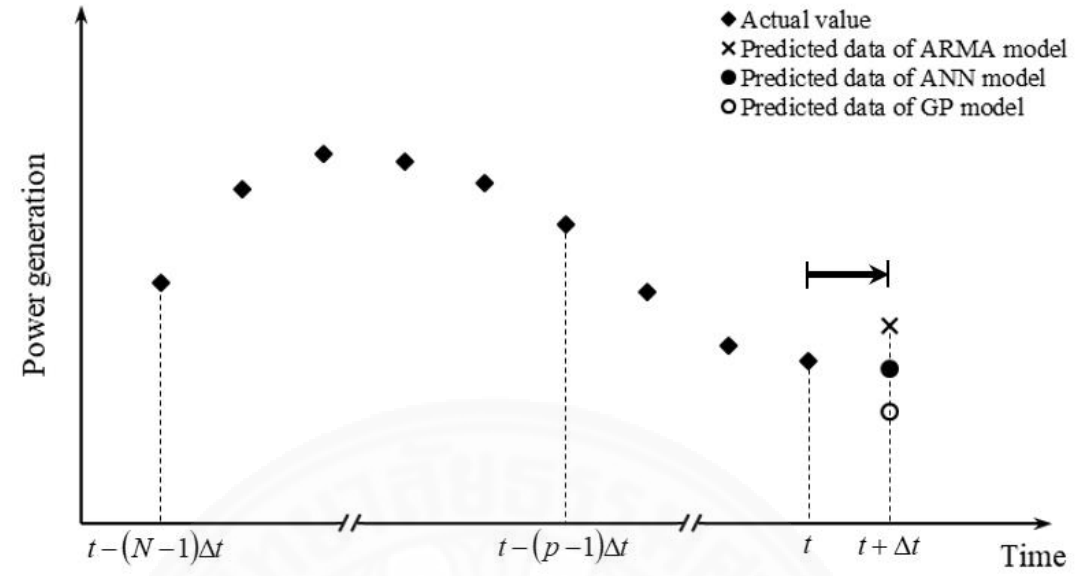

Figure 3.7 Prediction of ARMA model, ANN model, and GP model, based on actual input data.

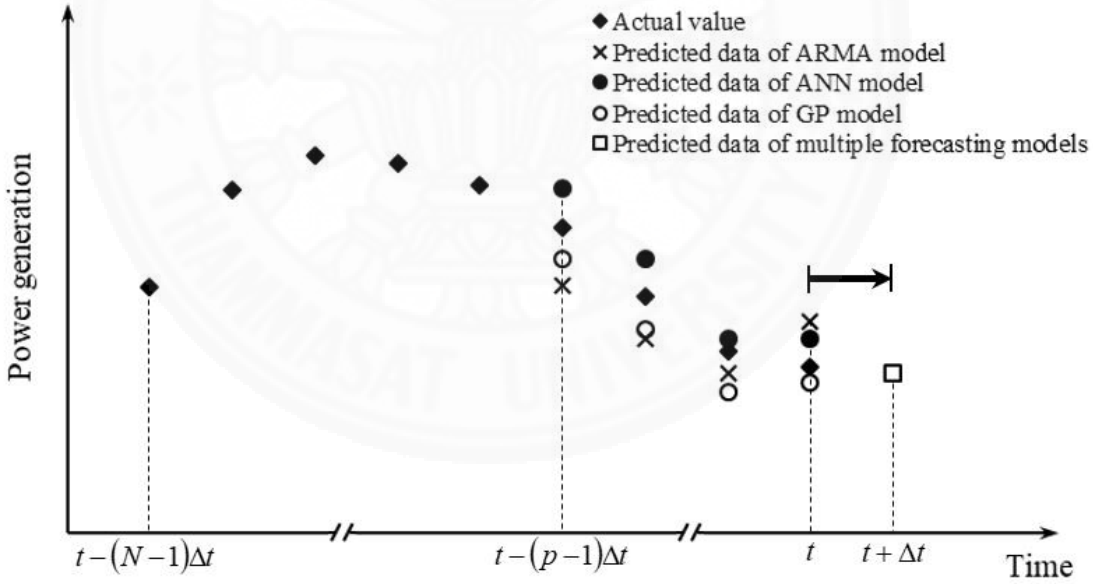

Figure 3.8 Prediction of multiple forecasting models, based on historical performance.

Figure 3.9 shows the procedure steps for the determination of the prediction interval. The forecasting errors between the actual values and predicted values at the present time and past time are determined as statistical data for determining the

Ref. code: 25625822300561TJA

prediction interval under uncertainty. The predicted value of multiple forecasting models is in the prediction interval with a probability of the percentage of confidence 100%(1-β), as depicted in Figure 3.10. The proposed methodology of multiple forecasting models and procedural steps of calculation are applied to the measurement data of power generation, which are obtained from a given wind farm under actual wind conditions.

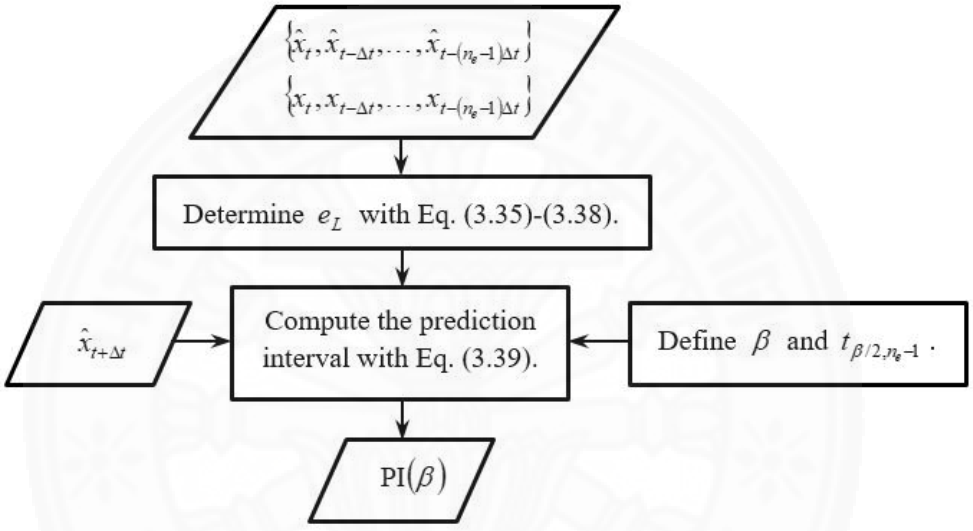

Figure 3.9 Schematic diagram of determination of the prediction interval.

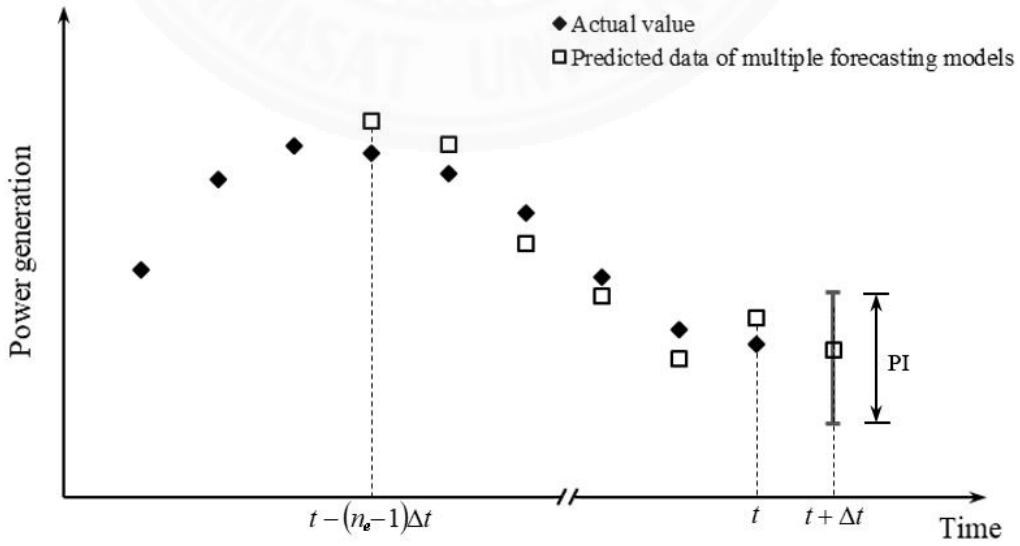

Figure 3.10 Determination of the prediction interval, based on historical performance.

Ref. code: 25625822300561TJA

CHAPTER 4
RESULTS AND DISCUSSION

4.1 Results and discussion on study of wind measurement

Wind measurements were performed from October 2015 to October 2016 on the rooftop of a laboratory building with a height of 20 m, as shown in Figure 4.1. A three-cup anemometer and a wind vane (Theodor Friedrichs™ 4035.0000BG) were installed for measuring the wind speed and wind direction, respectively. The measured range of wind speed was 0-60 m/s with an accuracy of ±0.2 m/s. A sampling rate of 10 Hz for data recording was set in order to perceive the actual dynamic behaviors of wind for this study. The data recovery rate of annual wind velocity measurements is 99.9%. There were a few missing wind data values during short periods of data transfer from the data acquisition system (probably due to a power failure).

Figure 4.1 Installation of three-cup type anemometer and wind vane on building rooftop.

Ref. code: 25625822300561TJA

A plot of wind speeds against time for 30 s is shown in Figure 4.2, where the evolution of wind speed in time can be observed explicitly. The wind data with a high sampling rate was representative of real continuous wind. A high sampling rate can lead to accurate results and interpretation in wind analysis.

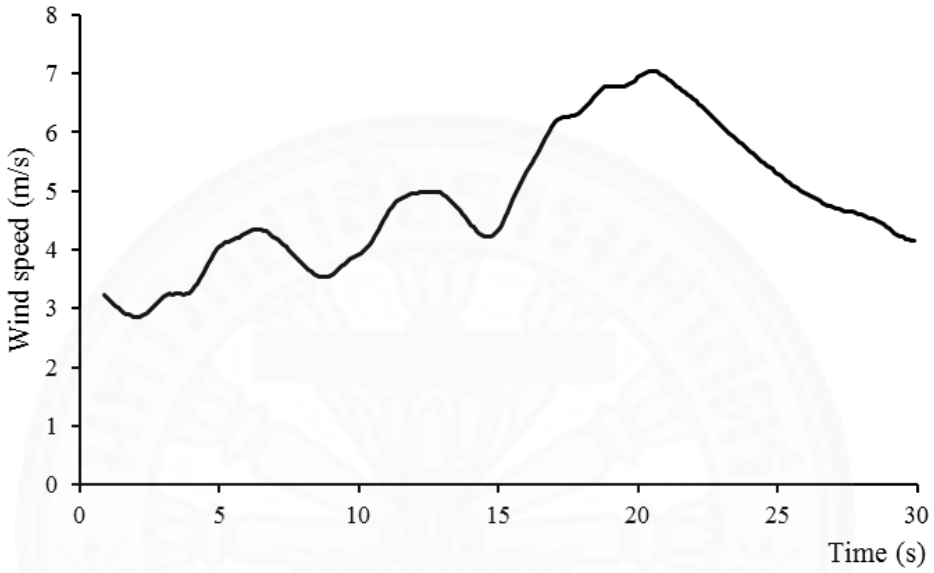

Figure 4.2 Illustration of wind speeds, sampled at a rate of 10 Hz.

4.1.1 Effects of sampling rate on wind data

The sampling rate in wind measurement has strong influences on the mean wind speed and standard deviation of wind data. Without loss of generality, one-day wind speed data is randomly selected from the annual wind data. As reported in Table 4.1, the mean wind speeds, standard deviations, and number of values are determined from the data at different sampling rates. It is noticed that the values of the mean wind speed and standard deviation at lower sampling rates tend to diverge from the values at a sampling rate of 10 Hz.

Ref. code: 25625822300561TJA

Table 4.1 Statistical analysis of 24-h wind data.

Sampling rate (Hz)	Mean speed (m/s)	Standard deviation (m/s)	Number of values
10	6.43	1.82	864000
1	6.43	1.82	86400
10^{-1}	6.43	1.82	8640
10^{-2}	6.45	1.80	864
$1.7\times10^{-3*}$	6.43	1.35	86400
10^{-3}	6.21	1.62	86
10^{-4}	4.91	1.63	8

* IEC 61400-12-1 standard

The characteristics of wind data are fuzzy when low sampling rates are implemented. However, the higher the sampling rate, the more data values. It can be seen that the number of wind data values at a sampling rate of 10 Hz is relatively high in this case. There is an unnecessary amount of wind data, to be used in the analysis. The optimal sampling rate, in this case, may be 0.01 Hz. This is a tradeoff between the accuracy and number of data values since the values of mean wind speed and standard deviation are equal to the values at a sampling rate of 10 Hz while the amount of data is reduced significantly.

For the IEC 61400-12-1 standard, wind speed measurements are performed at a sampling rate of 1 Hz. The wind data values are recorded, and averaged every 10 min (1.7×10^{-3} Hz). The results have less accuracy due to using averaged values of wind data.

4.1.2 Sensitivity analysis on parameters of Weibull distribution

In Figure 4.3, a sensitivity analysis is used to determine how different values of the mean wind speed and standard deviation affect the shape parameter and scale parameter of the Weibull distribution function, which are determined by Equations (3.8) and (3.9), respectively.

As illustrated in Figures 4.3(a) and 4.3(b), the percentage changes in the mean wind speed affect the percentage changes in the shape parameter and scale parameter. The percentage increases in the shape parameter and scale parameter are proportional to the percentage increase in the mean wind speed. In Figure 4.3(b), the percentage increase in the scale parameter is enlarged when the magnitude of the standard deviation

Ref. code: 25625822300561TJA

rises. In contrast, the percentage increases in the shape parameter and scale parameter are proportional to the percentage decrease in the standard deviation, as depicted in Figures 4.3(c) and 4.3(d). However, the relation of the scale parameter to the standard deviation in Figure 4.3(d) is reversed for low magnitudes of standard deviation, such as $\sigma = 0.4\bar{v}$ and $\sigma = 0.2\bar{v}$. This shift is caused by a critical point of the Gamma function between the negative and positive slopes.

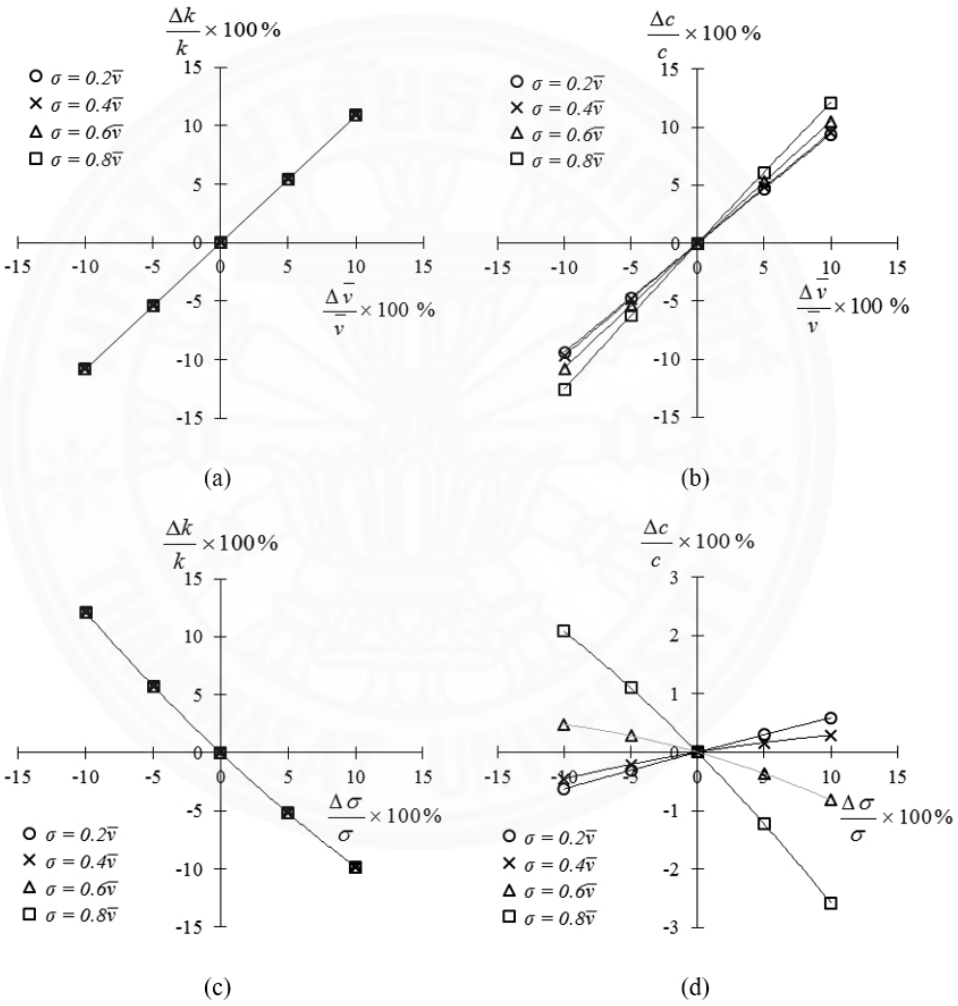

Figure 4.3 Sensitivity of Weibull parameters to changes in mean wind speed and standard deviation: **(a)** shape parameter with mean wind speed, **(b)** scale parameter with mean wind speed, **(c)** shape parameter with standard deviation, and **(d)** scale parameter with standard deviation.

Ref. code: 25625822300561TJA

In general, the variances in the shape parameter and scale parameter are sensitive to changes/uncertainties in the mean wind speed and standard deviation. It is important that the sampling rate is determined to yield the most accurate statistical results while obtaining the least number of wind data values. In Figure 4.4, the probability distribution of the wind speed and wind rose are plotted with yearly wind data at a sampling rate of 10 Hz. Around the measurement field, low wind speeds (<6 m/s) are frequently observed. The wind direction is from the tropical southwest monsoon and trade winds. The wind speed data at a sampling rate of 10 Hz are used to prepare the wind speed data according to the IEC 61400-12-1 standard for a case study. Consequently, the shape and scale parameters are determined to be 1.29 and 2.51 m/s, respectively, for the Weibull probability density function in Equation (3.7). In Figure 4.4(a), the Weibull probability density function does not match the frequency distribution of wind data at low wind speeds.

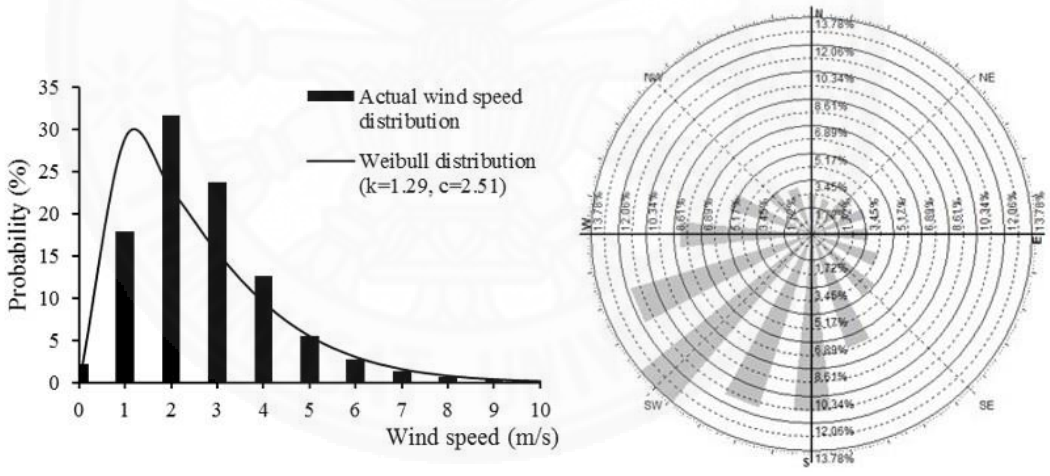

Figure 4.4 Yearly wind data at a sampling rate of 10 Hz: **(a)** probability distribution and **(b)** wind rose.

The assumptions of the Weibull distribution limit the extent of statistical results even though the wind data are sampled at different rates. This is because the variations of wind data are assumed to be governed by the Weibull distribution function with a shape parameter and a scale parameter. Therefore, power estimation without the

Ref. code: 25625822300561TJA

distributional assumptions yields genuine values of power production from the information of the sampled wind data, as determined by Equations (3.18) - (3.19).

4.1.3 Wind analysis from wind data at different sampling rates

The power curve of the low-speed horizontal-axis wind turbine from Ref. (Kishore & Priya, 2013) is proposed for the AEP calculation, as shown in Figure 4.5. Tables 4.2 and 4.3 list the corresponding statistical values of the wind speed, power density, AEP, and capacity factor. From Equations (3.7) - (3.15), the numerical values tend to deviate from the values at a sampling rate of 10 Hz. The percentage relative differences of analytical results in Tables 4.2 and 4.3 quantitatively indicate the information loss of wind data at different sampling rates, compared with the benchmark.

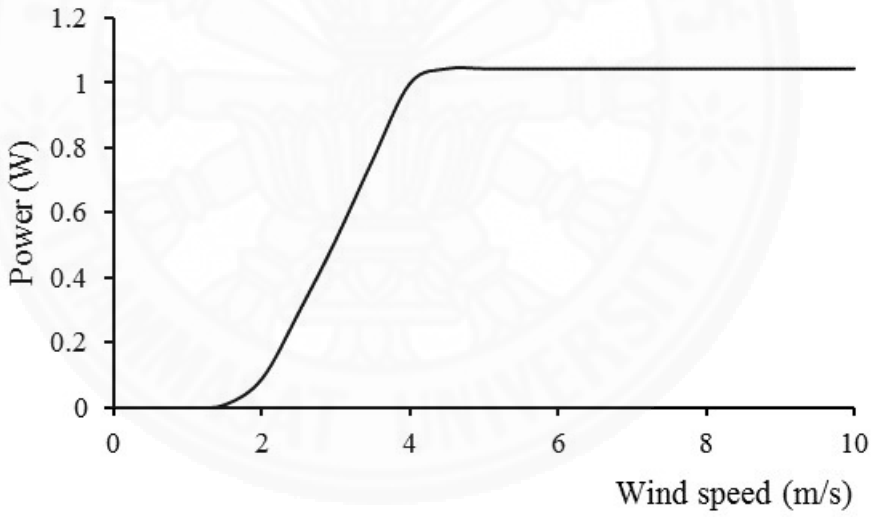

Figure 4.5 Power curve of wind turbine for AEP calculation.

Ref. code: 25625822300561TJA

Table 4.2 Statistical analysis of wind data at different sampling rates (mean speed and standard deviation).

	Cycle period	Number of values	Mean speed (m/s)	Standard deviation (m/s)
10 Hz	-	3.15×10^{8}	2.32	1.84
IEC 61400-12-1 standard	-	3.15×10^{7}	2.32 (0.00%)	1.68 (-8.70%)
Adaptive $\alpha = 0.05$	10 min	8.80×10^{6}	2.32 (0.00%)	1.84 (0.00%)
	hour	2.80×10^{6}	2.32 (0.00%)	1.84 (0.00%)
	day	4.00×10^{5}	2.33 (+0.43%)	1.85 (+0.54%)
	week	6.33×10^{4}	2.37 (+2.16%)	1.88 (+2.17%)
	month	$9.30\text{x}10^{3}$	2.36 (+1.72%)	1.92 (+4.35%)
Adaptive $\alpha = 0.1$	10 min	2.76×10^{6}	2.32 (0.00%)	1.84 (0.00%)
	hour	1.02×10^{6}	2.32 (0.00%)	1.84 (0.00%)
	day	7.44×10^{4}	2.33 (+0.43%)	1.85 (+0.54%)
	week	8.69×10^{3}	2.35 (+1.29%)	1.88 (+2.17%)
	month	2.53×10^{3}	2.37 (+2.16%)	1.86 (+1.09%)
Adaptive $\alpha = 0.15$	10 min	1.51×10^{6}	2.32 (0.00%)	1.84 (0.00%)
	hour	5.06×10^{5}	2.32 (0.00%)	1.84 (0.00%)
	day	2.23×10^{4}	2.34 (+0.86%)	1.86 (+1.09%)
	week	4.16×10^{3}	2.35 (+1.29%)	1.90 (+3.26%)
	month	1.38×10^{3}	2.34 (+0.86%)	1.96 (+6.52%)

Remark: Percentage of relative difference to wind data at a 10-Hz sampling rate, in parentheses.

Ref. code: 25625822300561TJA

Table 4.3 Statistical analysis of wind data at different sampling rates (power density, AEP, and capacity factor).

	Cycle period	Number of values	Power density (W/m^2)	Annual Energy Production (Wh)	Capacity factor
10 Hz	-	3.15×10^8	32.50	2460.0	0.2808
IEC 61400-12-1 standard	-	3.15×10^7	22.63 (-30.37%)	2793.1 (+13.54%)	0.3188 (+13.54%)
Adaptive $\alpha = 0.05$	10 min	8.80×10^6	32.50 (0.00%)	2462.6 (+0.11%)	0.2811 (+0.11%)
	hour	2.80×10^6	32.43 (-0.22%)	2465.8 (+0.24%)	0.2815 (+0.24%)
	day	4.00×10^5	32.86 (+1.11%)	2452.8 (-0.29%)	0.2800 (-0.29%)
	week	6.33×10^4	34.36 (+5.72%)	2351.3 (-4.62%)	0.2684 (-4.62%)
	month	$9.30\text{x}10^3$	35.58 (+9.48%)	1984.1 (-23.99%)	0.2265 (-23.99%)
Adaptive $\alpha = 0.1$	10 min	2.76×10^6	32.52 (+0.06%)	2469.1 (+0.37%)	0.2819 (+0.37%)
	hour	1.02×10^6	32.54 (+0.12%)	2469.5 (+0.38%)	0.2819 (+0.38%)
	day	7.44×10^4	32.79 (+0.89%)	2439.9 (-0.82%)	0.2785 (-0.82%)
	week	8.69×10^3	34.32 (+5.60%)	2327.7 (-5.68%)	0.2657 (-5.68%)
	month	2.53×10^3	32.44 (-0.18%)	2049.2 (-20.05%)	0.2339 (-20.05%)
Adaptive $\alpha = 0.15$	10 min	1.51×10^6	32.51 (+0.03%)	2476.3 (+0.66%)	0.2827 (+0.66%)
	hour	5.06×10^5	32.58 (+0.25%)	2478.2 (+0.73%)	0.2829 (+0.73%)
	day	2.23×10^4	33.05 (+1.69%)	2486.3 (+1.06%)	0.2838 (+1.06%)
	week	4.16×10^3	35.75 (+10.00%)	2307.8 (-6.60%)	0.2634 (-6.60%)
	month	1.38×10^3	37.56 (+15.57%)	1847.0 (-33.19%)	0.2108 (-33.19%)

Remark: Percentage of relative difference to wind data at a 10-Hz sampling rate, in parentheses.

Ref. code: 25625822300561TJA

With IEC 61400-12-1 standard, the percentages of relative differences are listed as 30.37% and 13.54% for the power density and AEP, respectively, while the differences with less wind speed data values are smaller. The averaged values of wind data at every 10 min may cause flaws in wind information. However, the mean wind speed is close to the mean value at a sampling rate of 10 Hz because the information about the mean values every 10 min is kept in the wind data. To obtain more accuracy and fewer values, the Nyquist-based adaptive sampling rate approach is implemented for various cycle periods T_m on the collected wind data. The cut-out amplitude α is varied from 0.05 to 0.15 m/s, and the cycle period T_m is 10 min, an hour, a day (highlighted), a week, and a month, as indicated in Tables 4.2 and 4.3. It is found that the higher the accuracy of results and number of wind speed values, the lower the values of α and T_m. The Nyquist rate $f_{s,m}^{(N)}$ is chosen at a high rate when the values of α and T_m are low. Therefore, the tradeoff between the number of wind values and the accuracy can be specified by the Nyquist rate $f_{s,m}^{(N)}$. It is found that a recommended α of 0.1 and T_m of a day provide acceptable results (highlighted). The percentages of errors are less than 1% for all statistical results at a sampling rate of 10 Hz. The number of values can be decreased by approximately 4000 times.

Table 4.4 illustrates the corresponding analysis for wind direction. The measurement data of wind direction are recorded with the same sampling rates that are implemented for wind speed. It can be seen that the percentage differences in most cases are significantly low, compared with wind speed. This result shows that there is less information loss in wind direction since the variation of wind direction is not high, as expected. Therefore, the sampling rates of wind direction can be determined from wind speed, in practice. It is noted that the variation of wind direction is low. In this case, a mode wind direction is more statistically suitable, referring to the prevailing wind direction, than a mean wind direction, which is usually determined as an indicator in wind analysis.

Ref. code: 25625822300561TJA

Table 4.4 Statistical analysis of wind data at different sampling rates (wind direction).

	Cycle period	Number of values	Mean direction (degrees)	Standard deviation (degrees)
10 Hz	-	3.15×10^{8}	211.37	96.38
IEC 61400-12-1 standard	-	3.15×10^{7}	211.37 (0.00%)	92.26 (-4.27%)
Adaptive $\alpha = 0.05$	10 min	8.80×10^{6}	211.37 (0.00%)	96.38 (0.00%)
	hour	2.80×10^{6}	211.37 (0.00%)	96.38 (0.00%)
	day	4.00×10^{5}	211.38 (0.00%)	96.38 (0.00%)
	week	6.33×10^{4}	211.44 (+0.03%)	96.36 (-0.02%)
	month	$9.30\text{x}10^{3}$	211.68 (+0.15%)	96.33 (-0.05%)
Adaptive $\alpha = 0.1$	10 min	2.76×10^{6}	211.37 (0.00%)	96.38 (0.00%)
	hour	1.02×10^{6}	211.37 (0.00%)	96.38 (0.00%)
	day	7.44×10^{4}	211.32 (-0.02%)	96.37 (-0.01%)
	week	8.69×10^{3}	211.04 (-0.15%)	96.48 (+0.10%)
	month	2.53×10^{3}	211.69 (+0.15%)	96.58 (+0.21%)
Adaptive $\alpha = 0.15$	10 min	1.51×10^{6}	211.37 (0.00%)	96.38 (0.00%)
	hour	5.06×10^{5}	211.37 (0.00%)	96.38 (0.00%)
	day	2.23×10^{4}	211.46 (+0.04%)	96.33 (-0.05%)
	week	4.16×10^{3}	211.33 (-0.02%)	96.27 (-0.11%)
	month	1.38×10^{3}	212.52 (+0.54%)	95.36 (-1.05%)

Remark: Percentage of relative difference to wind data at a 10-Hz sampling rate, in parentheses.

The mean wind speed and standard deviation are commonly reported to perceive the statistical characteristics of wind in engineering practice. The analysis of wind data in terms of higher-order statistics, such as skewness and kurtosis, yields a similar

Ref. code: 25625822300561TJA

tendency of information loss at different sampling rates, which is listed for completeness of information in Table 4.5.

Table 4.5 Statistical analysis of wind data at different sampling rates (skewness and kurtosis).

	Cycle period	Number of values	Skewness	Kurtosis
10 Hz	-	3.15×10^{8}	1.66	7.67
IEC 61400-12-1 standard	-	3.15×10^{7}	1.79 (7.83%)	9.26 (20.73%)
Adaptive $\alpha = 0.05$	10 min	8.80×10^{6}	1.66 (0.00%)	7.69 (0.26%)
	hour	2.80×10^{6}	1.66 (0.00%)	7.67 (0.00%)
	day	4.00×10^{5}	1.65 (-0.60%)	7.57 (-1.30%)
	week	6.33×10^{4}	1.68 (1.20%)	7.74 (0.91%)
	month	$9.30\text{x}10^{3}$	1.81 (9.04%)	8.65 (12.78%)
Adaptive $\alpha = 0.1$	10 min	2.76×10^{6}	1.66 (0.00%)	7.70 (0.39%)
	hour	1.02×10^{6}	1.66 (0.00%)	7.65 (-0.26%)
	day	7.44×10^{4}	1.65 (-0.60%)	7.60 (-0.91%)
	week	8.69×10^{3}	1.62 (-2.41%)	7.35 (-4.17%)
	month	2.53×10^{3}	1.62 (-2.41%)	7.22 (-5.87%)
Adaptive $\alpha = 0.15$	10 min	1.51×10^{6}	1.66 (0.00%)	7.68 (0.13%)
	hour	5.06×10^{5}	1.67 (0.60%)	7.67 (0.00%)
	day	2.23×10^{4}	1.71 (3.01%)	7.82 (1.96%)
	week	4.16×10^{3}	1.63 (-1.81%)	7.47 (-2.61%)
	month	1.38×10^{3}	1.98 (19.28%)	9.75 (27.12%)

Remark: Percentage of relative difference to wind data at a 10-Hz sampling rate, in parentheses.

Ref. code: 25625822300561TJA

Correspondingly, Figure 4.6 shows a plot of the Nyquist rate, which is determined from the proposed methodology for the whole year. The sampling rate changes day by day to acquire all-inclusive wind information, based on the Nyquist sampling theorem. It can be observed that the sampling rate is defined to be well-suited to varying wind conditions in each day. This implementation ensures that excessive wind data is not collected under a low variation of wind speed, whereas necessary wind data is fully recorded under a high variation of wind speed, without missing data.

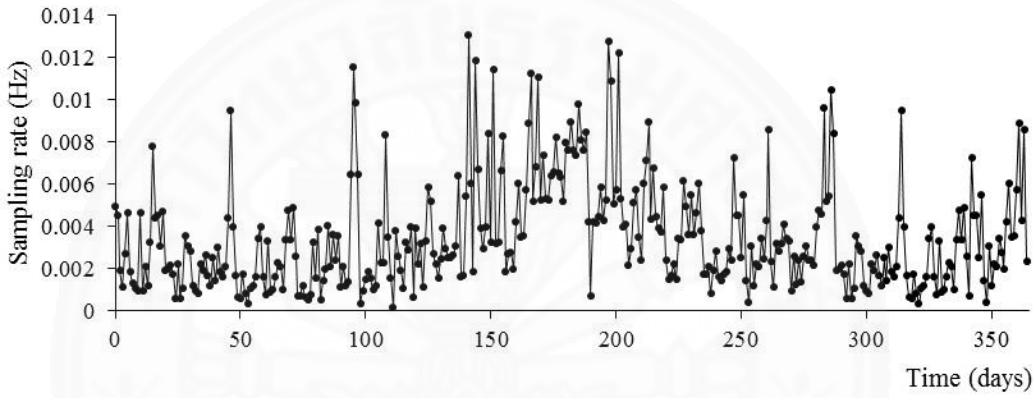

Figure 4.6 Adaptation of sampling rate with Nyquist frequency against time.

It is remarked that, due to the limitation of apparatuses and sites, the rooftop of building where the wind measurement is performed is 20 m height, approximately. However, the proposed Nyquist-based adaptive sampling rate approaches can be applied regardless of altitude. The results are in the same way at any altitude, which high sampling rate is preferred for high wind speed, and vice versa. As explained in Appendix D, wind speed and altitude is proportional according to the wind profile power law. The characteristic of wind speeds at different altitude do not change, as shown in Figure D1.

Ref. code: 25625822300561TJA

4.2 Results and discussion on the study of wind power forecasting

Measurement data of power generation from the 17.8-MW Sotavento experimental wind farm was retrieved from the webpage of Sotavento Galicia, S.A. at http://www.sotaventogalicia.com from November to December, 2018. The data was retrieved in real time at a sampling time of 2 seconds to obtain accurate information of the data. The wind farm is located in Galicia, A Coruña, Spain, which is approximately 700 m above sea level, as shown in Figure 4.7. There are twenty-four wind turbines of nine models with generation capacities from 640 kW to 1320 kW, as listed in Table 4.6. The power curves against wind speed are graphically presented in Figure 4.8. Figure 4.9 shows a layout of wind turbines and meteorological masts, which are aligned along a roadway. It is reported that the annual average wind speed at the site is 6.4 m/s (prevailing east-west wind direction). The electricity production of this wind farm over time is used as case scenarios to verify the real-time implementation of multiple forecasting models under varying wind conditions.

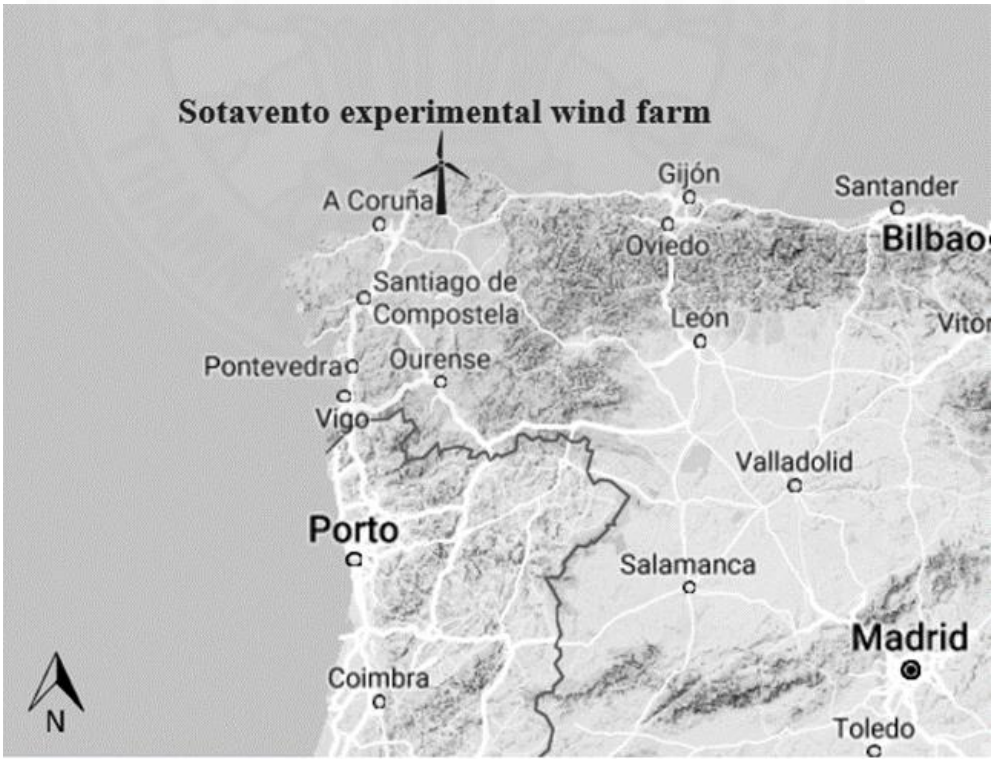

Figure 4.7 Location of Sotavento experimental wind farm.

Ref. code: 25625822300561TJA

Table 4.6 Specification of wind turbines.

Model	Rotor diameter (m)	Hub height (m)	Number	Rated power (kW)
Ecotecnia 44/640	44	46	4	640
Gamesa G-47	47	45	4	660
Izar-Bonus 1.3 MW	62	49	1	1300
Izar-Bonus MK-IV	44	40	4	660
Made AE-46	46	45	4	660
Made AE-52	52	50	1	800
Made AE-61	61	60	1	1320
Neg Micon NM-48 750	48	45	4	750
Neg Micon NM-52 900	48	45	1	900

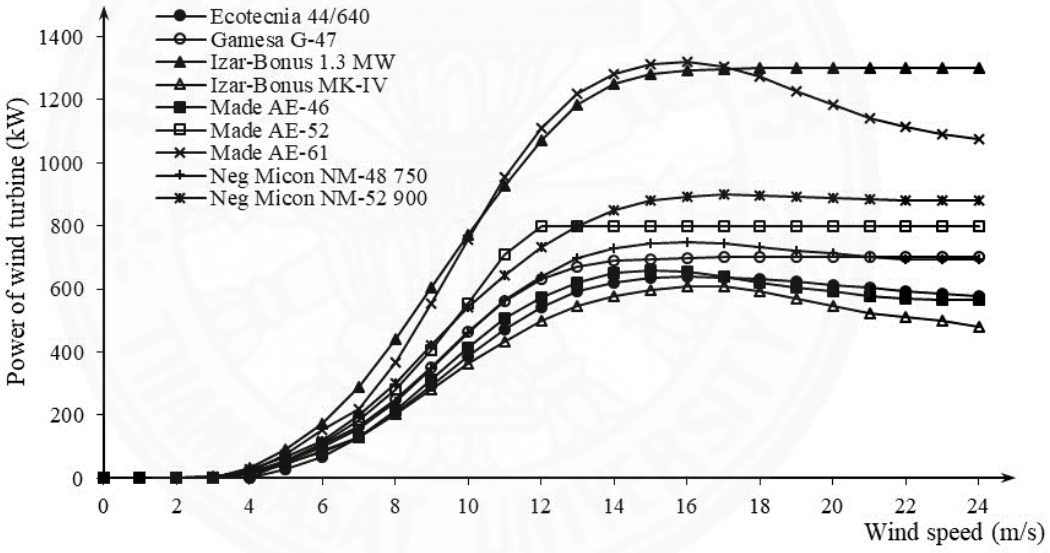

Figure 4.8 Power curves of wind turbines.

Ref. code: 25625822300561TJA

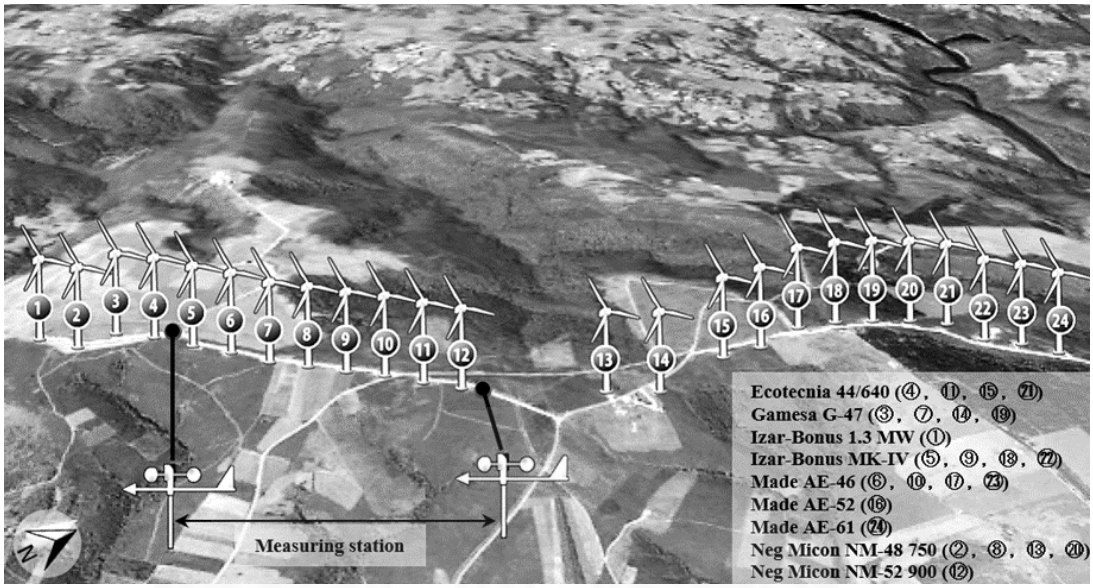

Figure 4.9 Layout of wind turbines and meteorological masts.

4.2.1 Performance investigation of forecasting models

To demonstrate the capability of each predictive model, one-week measurement data of hourly power generation is subjectively selected from the Sotavento experimental wind farm, as shown in Figure 4.10. Two different periods are used. One period has a high variance with a standard deviation of 830 kWh, while the other period has a low variance with a standard deviation of 395 kWh.

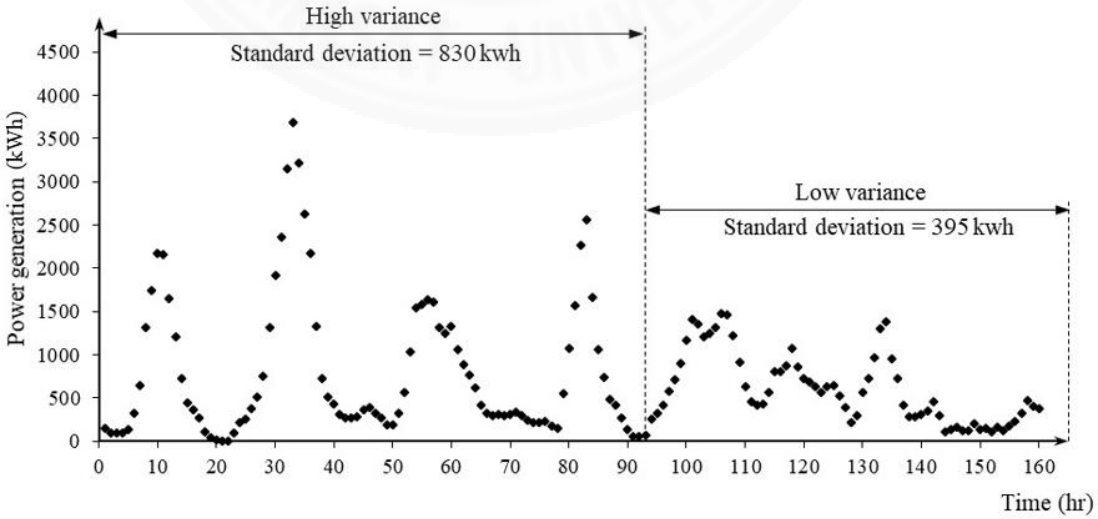

Figure 4.10 Plots of hourly energy output data against time.

Ref. code: 25625822300561TJA

According to the procedure steps of calculation in Figure 3.6, Figure 4.11 presents a time-series plot of one-hour ahead forecasting by the persistence model, ARMA model, ANN model, GP model, and multiple forecasting models. The differences between the actual data in Figure 4.10 and predicted data in Figure 4.11 are used to determine RMSEs for forecasting performance. RMSEs of each forecasting model at two different periods are listed in Table 4.7. As expected, RMSEs of the persistence model are the highest values among all the predictive models at two different periods. In the conventional approach of a single predictive model, the ANN model is superior in dealing with high variation as it yields the lowest RMSE, compared to the ARMA model and GP model. In contrast, the RMSE of the ARMA model is the lowest value at a low variance of power generation. This can be interpreted that there is no most-effective forecasting model at all times. However, the multiple forecasting models yield predicted data with the lowest RMSEs at two different periods. The proposed method has the capability to weigh the predicted values from the ARMA model, ANN model, and GP model based on historical errors of prediction. Compared to the persistence model as the benchmark, the accuracy of the multiple forecasting models is significantly improved by 57% and 39% during periods of high variance and low variance, respectively.

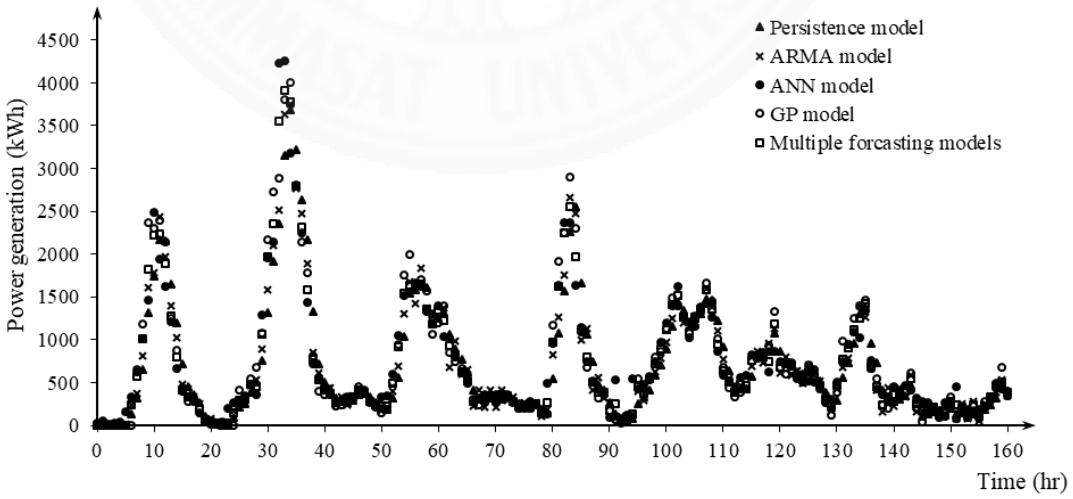

Figure 4.11 Plot of hourly-predicted energy output data against time.

Ref. code: 25625822300561TJA

Table 4.7 Root mean square error of forecasting models.

Forecasting model	RMSEs (kWh)	
	At high variance	At low variance
Persistence model	301	153
ARMA model	217	129
ANN model	165	133
GP model	193	136
Multiple forecasting model	128	93

To understand the functionality of multiple forecasting models, Figures 4.12 and 4.13 show adjustments of the weighting multipliers during forecasting at high variance and low variance, respectively. The linear regressions of weighting multipliers from three models are determined, to present average weighting multipliers over a period. In Figure 4.12, the predicted data of the ANN model is dominant at a high-variance period since the weighting multipliers are the highest values. In Figure 4.13, the weighting multipliers of the ARMA model are the most dominant at a low variance period. This is in good agreement with the forecasting performance from the RMSEs in Table 4.7.

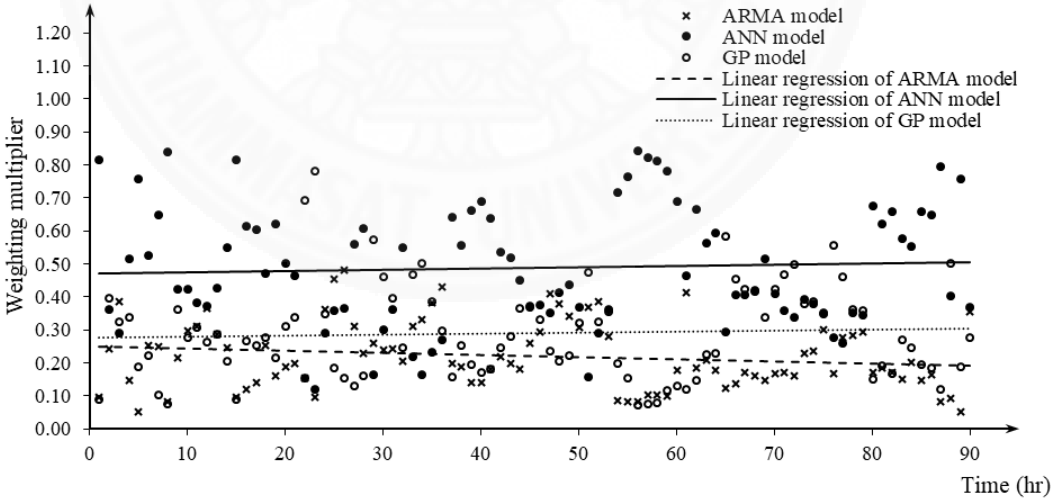

Figure 4.12 Plot of weighting multiplier against time at high variance.

Ref. code: 25625822300561TJA

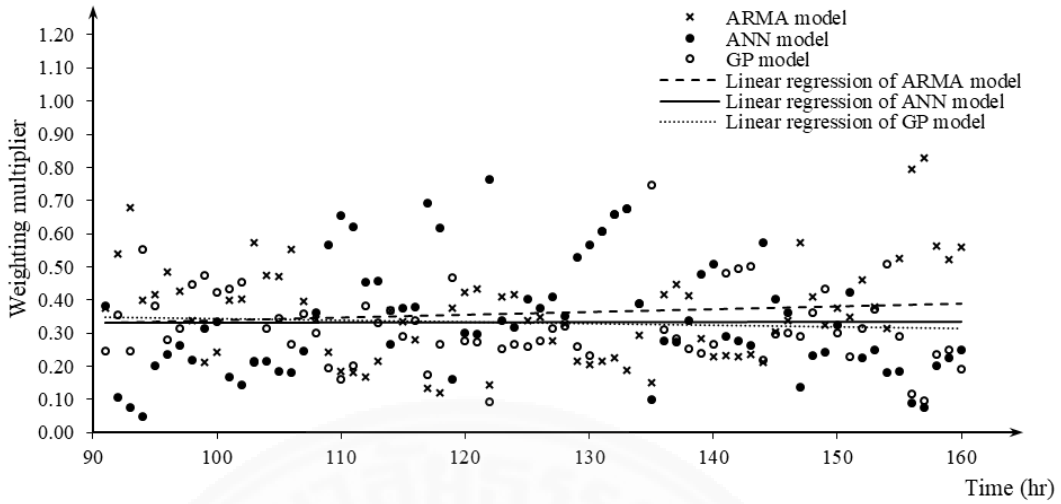

Figure 4.13 Plot of weighting multiplier against time at low variance.

4.2.2 Prediction of power generation with multiple forecasting models

To demonstrate the effectiveness of the proposed methodology for short-term prediction, the multiple forecasting models are applied to predict the power generation with time horizons of 1 hour, 3 hours, and 6 hours. In Fig. 4.14, the one-week power generation (in kW) is collected from the Sotavento experimental wind farm under open field conditions. It can be seen that the wind power varies drastically with respect to time where the following case studies of power prediction are worthy of investigation.

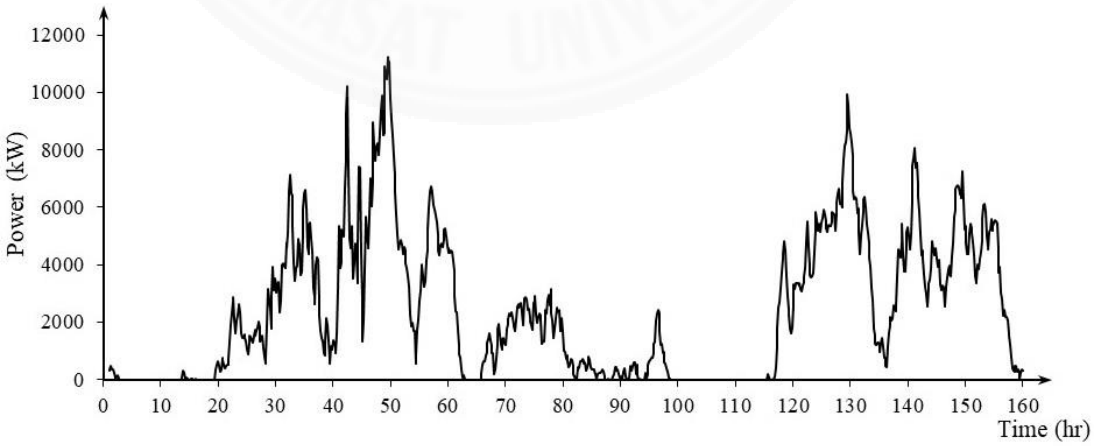

Figure 4.14 One-week power generation of the wind farm.

Ref. code: 25625822300561TJA

According to the flowchart in Figure 3.6, Figure 4.15(a) shows the hourly prediction of power generation with respect to time. The actual power generation (in kWh) from Figure 4.14 is presented in Figure 4.15(b) for comparison. The multiple forecasting models yield predicted values that are close to the actual values with an RMSE of 321 kWh over a week, as listed in Table 4.8. The benchmark of prediction from the persistence model is obtained with an RMSE of 695 kWh, which is significantly inferior, compared to the multiple forecasting models.

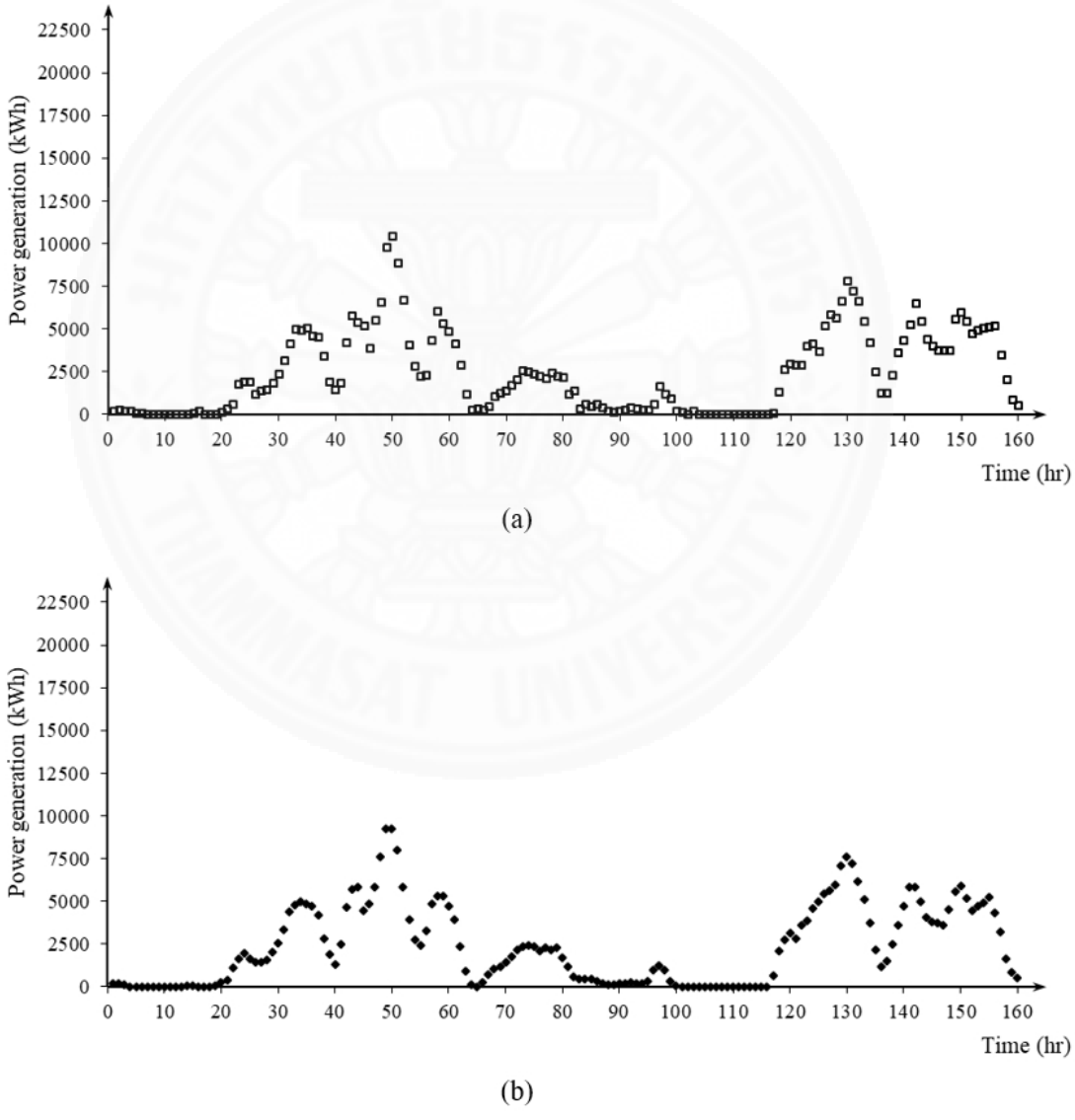

Figure 4.15 Hourly power generation: **(a)** predicted values and **(b)** actual values.

Ref. code: 25625822300561TJA

Table 4.8 Performances of multiple forecasting models with various time horizons.

Time horizon (h)	RMSEs (kWh)	
	Persistence model	Multiple forecasting models
1	695	321
3	4340	2532
6	10445	5021

To determine the corresponding prediction intervals of multiple forecasting models, errors in forecasting are needed with the number of data values n_e at the present and past times for Equations (3.35) - (3.37), or $n_e - 1$ degrees of freedom for Equation (3.38). The number of data values n_e at the present and past times are chosen so that the actual value at the next time is within the small prediction interval. The prediction interval is enlarged by decreasing the number of data values n_e according to Equation (3.38) while increasing the number of data values n_e results in large computation time and less sensitivity to the prediction interval. In this study, the number of data values n_e at the present and past times are defined from the least number that has the smallest PI. For example, the different number of data values n_e at the present and past times are considered at the 50th h in Figure 4.15. As shown in Figure 4.16, n_e is recommended to be 15 for the narrowest prediction interval of 2250 kWh. In this case, the prediction interval is comparatively large due to the acute increase of power generation at this time where the errors of prediction are significantly large.

Ref. code: 25625822300561TJA

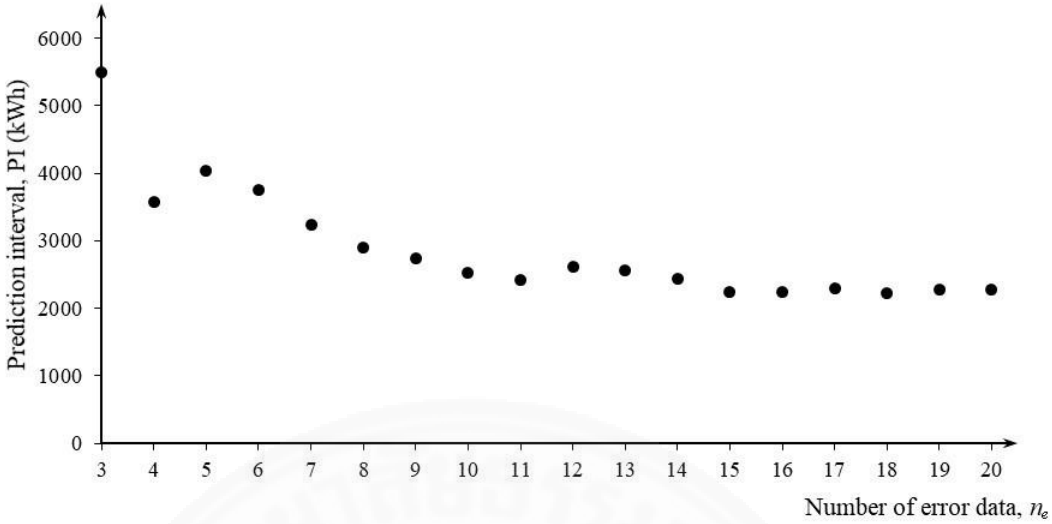

Figure 4.16 Determination of number of error values n_e, at the present and past times.

After the number of error data values is determined at each time, Figure 4.17 shows the predicted values of the multiple forecasting models, and the corresponding prediction intervals with 95% confidence during the whole week. It can be observed that the prediction intervals are large when the predicted values change rapidly. This result indicates that the prediction intervals cover the feasible ranges of the actual values.

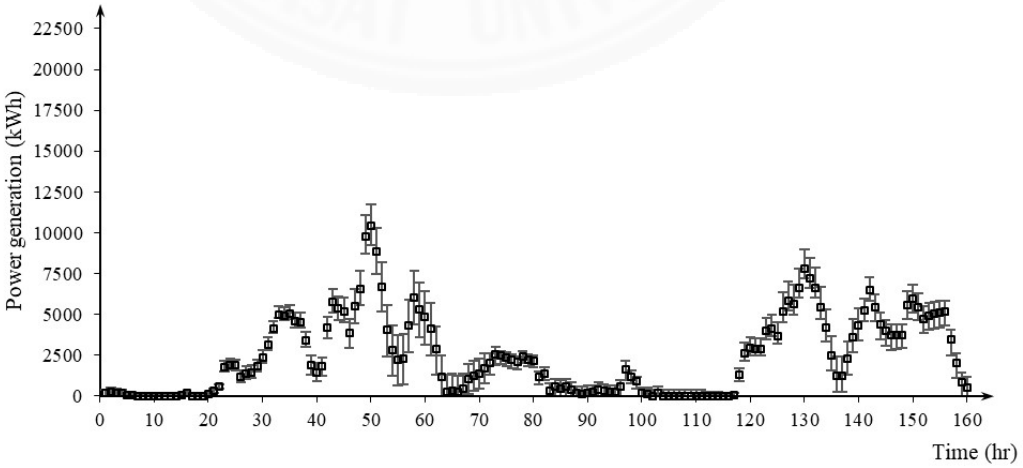

Figure 4.17 Plots of hourly predicted values and prediction intervals with 95% confidence, with respect to time.

Ref. code: 25625822300561TJA

Figure 4.18 and Figure 4.19 present the predicted values and the prediction interval (95% confidence) for the one-week power generation with time horizons of 3 hours and 6 hours, respectively. It is noticed that the widths of PIs increase as the time horizon increases. This result is explained by the increase of RMSEs at a high time horizon, as listed in Table 3. Additionally, most of the actual values of power generation are included within the PIs. For three cases, it is found that 96.25% of the actual values are within the PIs.

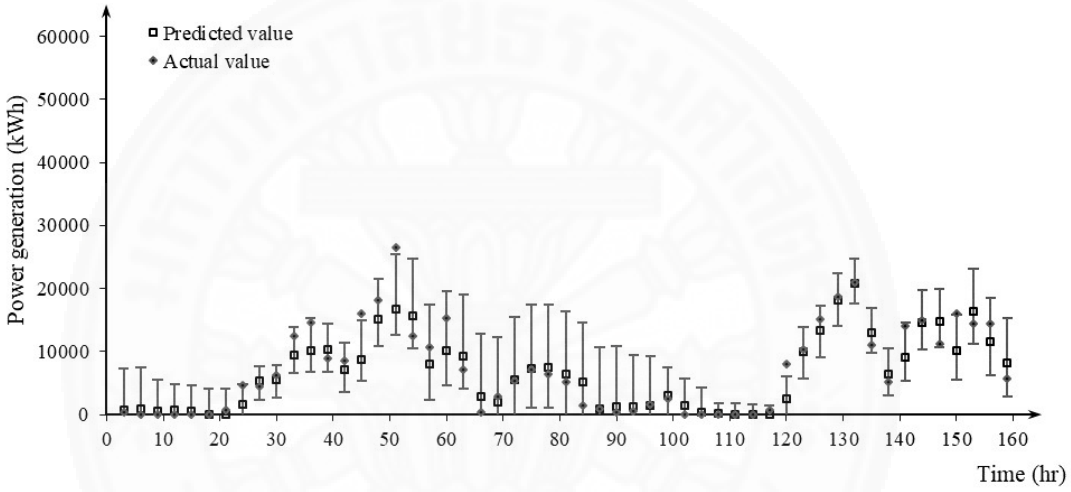

Figure 4.18 Plots of 3-hourly predicted values and prediction interval with 95% confidence, with respect to time.

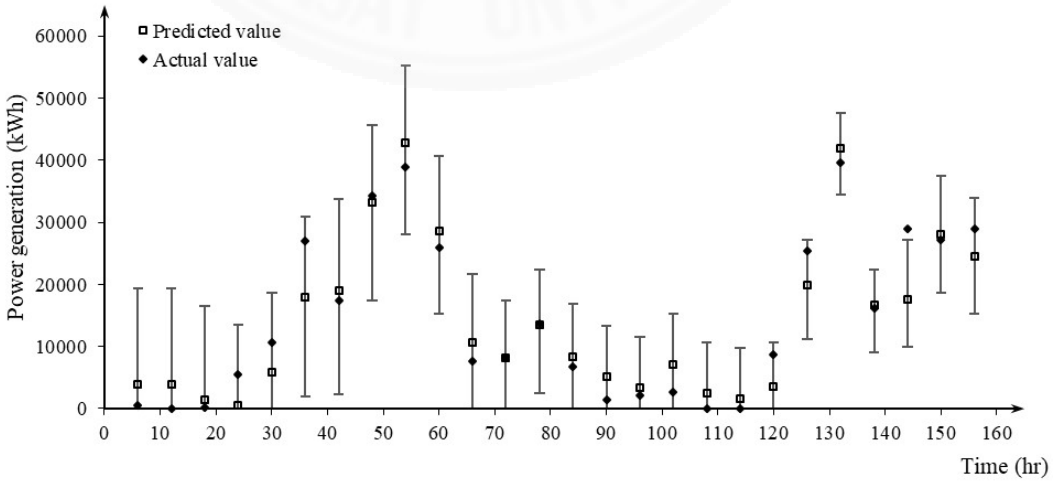

Figure 4.19 Plots of 6-hourly predicted values and prediction interval with 95% confidence, with respect to time.

Ref. code: 25625822300561TJA

In addition, the prediction intervals in Figure 4.17, 4.18, and 4.19 can be alternatively presented in a continuous series along with the predicted values, as shown in Figure C1, C2, and C3 in Appendix C, respectively. These figures show the interpolated trend line of the prediction intervals in-between the forecasting horizon.

To be concise, all the main results in this Chapter 4 are summarized as listed on Table 4.9.

Table 4.9 Summary of results

Section	Summary
4.1.1 Effects of sampling rate on wind data	The sampling rate has significant impact on wind analysis. As the sampling rate decreases, the value of mean speed and standard deviation tend to deviate from the benchmark of 10 Hz. In turn, the amount of values is reduced.
4.1.2 Sensitivity analysis on parameters of Weibull distribution	The variances in the shape parameter and scale parameter are sensitive to changes/uncertainties in the mean wind speed and standard deviation. The Weibull probability density function does not match the frequency distribution of wind data at low wind speeds in this case study.
4.1.3 Wind analysis from wind data at different sampling rates	IEC 61400-12-1 standard approach yields significant percentages of relative differences are listed as 30.37% and 13.54% for the power density and AEP, respectively. The tradeoff between the number of wind values and the accuracy can be specified by the Nyquist rate $f_{s,m}^{(N)}$. It is found that a recommended α of 0.1 and T_m of a day provide acceptable results. The percentages of errors are less than 1% for all statistical results at a

Ref. code: 25625822300561TJA

	sampling rate of 10 Hz. The number of values can be decreased by approximately 4000 times.
4.2.1 Performance investigation of forecasting models	The RMSEs of the persistence model are highest among all the predictive models. Except the multiple forecasting model, ANN model yields the lowest RMSE at a high variance of power generation and the ARMA model yields the lowest RMSE at a variance of power generation. Multiple forecasting model yields the lowest RMSEs at both different periods. Compared to the persistence model, the accuracy of the multiple forecasting models is significantly improved by 57% and 39% during periods of high variance and low variance, respectively.
4.2.2 Prediction of power generation with multiple forecasting models	Compared to the benchmark of persistence model, the multiple forecasting models significantly improve the accuracy in short-term prediction of power generation at time horizons of 1, 3, and 6 hours. The widths of PIs with 95% confidence increase as the time horizon increases since it yields high RMSE. Most of the actual values of power generation are included within the PIs. It is found that 96.25% of the actual values are within the PIs.

Ref. code: 25625822300561TJA

CHAPTER 5
CONCLUSION AND FUTURE RESEARCH

The wind measurements were performed from October 2015 to October 2016. The wind data at a high sampling rate of 10 Hz is defined as the benchmark. The data recovery rate of the annual wind speed data measurement is 99.9%. In the IEC 61400-12-1 standard, the wind data are sampled every second, and the mean values are recorded every 10 min. The standard wind analysis yields an accurate mean wind speed but an inaccurate standard deviation, compared with the benchmark. In sensitivity analysis, increasing the standard deviation by 5% results in changes of the shape parameter by -5%, and the scale parameter by -1% to 0.5% for the Weibull wind distribution function for wind energy assessment. It is observed that the percentage deviations of the power density, AEP, and capacity factor from the corresponding results of the benchmark are 30.37%, 13.54%, and 13.54%, respectively, while the differences with less wind speed data are small. Without consideration of wind conditions, it is confirmed that there are various deviations of analytical results, with excessive or missing wind data at high and low sampling rates. Conventional wind measurements with a fixed sampling rate cause significant excessive or missing wind data since the wind velocity changes differently in time. It is necessary to adjust the sampling rates of measurement, according to the wind conditions. The sampling rate is determined by the Nyquist frequency, covering all dominant amplitudes of wind speed. The Nyquist-based adaptive sampling rate method is proposed to systematically adapt the sampling rate to be the optimal Nyquist frequency, according to varying wind conditions. For time series analysis, the measurement data are formed by an array of cycle periods with optimal sampling rates, which determine the necessary amount of wind data. The proposed methodology is capable of providing high accuracy of analytical results with a percentage relative difference of less than 1% in wind analysis with the recommended parameters of the cut-out amplitude and cycle period. In addition, the amount of wind data is significantly decreased by 4000 times from the benchmark.

This study of wind measurement addresses the problem of sampling rate in the measured wind data and how the common approach of IEC 61400-12-1 standard and

Ref. code: 25625822300561TJA

Weibull probability density function does not yield acceptable results in wind analysis. Very few studies in the literature reviews provide enough investigation on how significance the sampling rate in wind measurement is. The study in this part focuses on obtaining tradeoff between the accuracy of wind data and the amount of data from wind measurement. The novel approach for selecting the sampling rate in wind measurement called Nyquist-based adaptive sampling rate is proposed. Unlike the IEC 61400-12-1 standard, the sampling rate is determined and adapted according to the wind condition where the wind measurement is performed. The research can provide benefits for any application where the accuracy of wind data is important while the cost and time to deal with a large amount of data are limited.

For wind power forecasting, comparative studies of well-known predictive models, such as the ARMA model, the ANN model, and the GP model, for the short-term forecasting of power generation from a wind farm are presented in this work. It is confirmed that each model has a limited capability of prediction in a certain variance of wind power generation. In contrast to this, the ANN model has been applied as a universal approximation even though much effort is required in computation. The traditional implementation of a single predictive model is not effective, to forecast wind power generation under uncertainty all the time. The proposed methodology of the multiple forecasting models is applied by weighting the predicted value of each individual model for the best-predicted value, based on the preceding performances of those predictive models. The RMSEs of the multiple forecasting models are significantly lower than the benchmark from the persistence model in all cases of one hour, three hours, and six hours ahead. In addition, the prediction intervals with 95% confidence are statistically determined from the historical errors of prediction where it is found that 96.25% of the actual values are in the prediction intervals. The proposed multiple forecasting models with the prediction interval yield reliable forecasts with statistical confidence under variances of wind power generation.

The study of wind power forecasting shows that the traditional implementation of a single predictive model is not effective to predict short-term wind power generation in real world. This research develops a novel combined prediction with weighting method called multiple forecasting models. The forecasts are presented in form of the prediction intervals to make the prediction results more informative and more practical

Ref. code: 25625822300561TJA

to applications. The multiple forecasting models yields significant improvement compared to the traditional approach. For the contribution to wind farms, the proposed methodology can be implemented to for effective prediction of power generation up to 6 hours ahead. This provides reliable operation in wind farms for the preparation of real-time grid operations, ancillary service costs, power quality, and the stability and reliability of power systems.

In future research, there is still some adaptation that can be tested and experimented in multiple forecasting models scheme for improvement. More advanced predictive models can be added/substituted to the method of multiple forecast models. Variant of ARMA model, ANN model, and high order of GP model may be considered to be included, for example, fractional-ARIMA model, Radial Basis Function ANN model, Convolutional ANN model, etc. For wind prediction at longer time horizon, it is a good idea to combine prediction from physical approaches or numerical weather prediction into the multiple forecasting model schemes. Since, the physical approach of wind forecasting is much more superior at long time horizon. Anyway, deep analysis and performance investigation is needed. Both proposed method of wind prediction and wind power forecasting may not limited to the wind energy area. Similar implementation can be consider to be used in applications of other renewable energy, such as solar energy, hydro energy, etc. However, specific investigation and experiment for the corresponding type of data is necessary.

Ref. code: 25625822300561TJA

REFERENCES

Akçay, H. & Filik, T. (2017). Short-term wind speed forecasting by spectral analysis from long-term observations with missing values. *Applied Energy, 191*, 653-662.

Akgül, F. G., Senoglu, B. S., & Arslan, T. (2016). An alternative distribution to Weibull for modeling the wind speed data: inverse Weibull distribution. *Energy Conversion and Management, 114*, 234-240.

Alencar, D. B., Affonso, C. M., & Oliveira, R. C. L. (2017). Different models for forecasting wind power generation: case study. *Energies, 10*(12), 1976.

Amjady, N., Keynia, F., & Zareipour, H. (2011). Short-term wind power forecasting using ridgelet neural network. *Electric Power Systems Research, 81*, 2099-2107.

Atthajariyakul, S., & Leephakpreeda, T. (2005). Neural computing thermal comfort index for HVAC systems. *Energy Conversion and Management, 46*(15), 2553-2565.

Bates, J. M. & Granger, C. W. J. (1969). The combination of forecasts. *Operations Research Quarterly, 20*, 451-468.

Blanche, T. J. & Swindale, N. V. (2006). Nyquist interpolation improves neuron yield in multiunit recordings. *Journal of Neuroscience Methods, 155*, 81-91.

Borg, M. & Collu, M. (2015). Frequency-domain characteristics of aerodynamic loads of offshore floating vertical axis wind turbines. *Applied Energy, 155*, 629-636.

Box, G., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis: forecasting and control (3rd ed)*. New Jersey: Prentice-Hall.

Bracewell, R. (2000). *The Fourier Transform and its Applications (3rd ed.)*. Boston: McGraw-Hill.

Breid, F. J., Davis, R. A., Lh, K. S., & Rosenblatt, M. (1991). Maximum likelihood estimation for noncausal autoregressive processes. *Journal of Multivariate Analysis, 36*(2), 175-198.

Chang, W (2014). A literature review of wind forecasting methods. *Journal of Power and Energy Engineering, 2*, 161-168.

Ref. code: 25625822300561TJA

Claeskens, G., & Hjort, N. L. (2008). *Model selection and model averaging.* Cambridge: Cambridge University Press.

Coville, A., Siddiqui, A., & Vogstadd, K. (2011). The effect of missing data on wind resource estimation. *Energy, 36*(7), 4505-4517.

Dadkhah, M., Rezaee, M. J., & Chavoshi, A. Z. (2018). Short-term power output forecasting of hourly operation in power plant based on climate factors and effects of wind direction and wind speed. *Energy, 148*, 775-788.

Deng, J. L. (1989). Introduction to grey system theory. *Journal of Grey System, 1*, 1-24.

Diniz, P. S. R., Silva, E. A. B. D., & Netto, S. L. (2002). *Digital Signal Processing: System Analysis and Design (1st ed.)*. New York: Cambridge University Press.

Dong, G., Fataliyev, K., & Wang, L. (2013). One-step and multi-step ahead stock prediction using backpropagation neural networks. Proceeding of *9th international conference on information, communications & signal processing* (pp. 1-5). Tainan, Taiwan: IEEE.

Edwards, T. S. (2007). Effects of aliasing on numerical integration. *Mechanical Systems and Signal Processing, 21,* 165-176.

El-Fouly, T. H. M. (2006). Grey predictor for wind energy conversion systems output power prediction. *IEEE Transactions on Power Systems, 21*(3), 1450-1452.

Erdem, E. & Shi, J. (2011). ARMA based approaches for forecasting the tuple of wind speed and direction. *Applied Energy, 88*(4), 1405-1414.

Faust, O., Acharya, U. R., Ma, J., Min, L. C., & Tamura, T. (2012). Compressed sampling for heart rate monitoring. *Computer Methods and Programs in Biomedicine, 108,* 1191-1198.

Hatagawa, E. & Delsing, J. (1994). Effects of pulsating flow on an ultrasonic gas flowmeter. *Flow Measurement and Instrumentation, 5,* 93-101.

Harris, R. I. & Cook, N. J. (2014). The parent wind speed distribution: why Weibull?. *Journal of Wind Engineering and Industrial Aerodynamics, 131*, 72-87.

Hatagawa, M., Kawahata, H., Shogaki, M., Yoshida, R., & Kariya, K. (1997). The effects of aliasing error on computed radiography in medical imaging. *Measurement, 21,* 71-78.

Ref. code: 25625822300561TJA

Hibon, M. & Evgeniou, T. (2005). To combine or not to combine: selecting among forecasts and their combinations. *International Journal of Forecasting, 21*, 15-24.

Hong, Y. Y., Chang, H. L., & Chiu, C. S. (2010). Hour-ahead wind power and speed forecasting using simultaneous perturbation stochastic approximation (SPSA) algorithm and neural network with fuzzy inputs. *Energy, 35*, 3870-3876.

Hu, J. & Wang, J. (2015). Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression. *Energy, 93*, 1456-1466.

Jiang, P., Wang, Y., & Wang, J. (2017). Short-term wind speed forecasting using a hybrid model. *Energy, 119*, 561-577.

International Electrotechnical Commission (2005). *I.E.C. 61400-12-1 Standard, Wind Turbine- Part 21-1: Power Performance Measurements of Electricity Producing Wind Turbines (1st ed.).* Geneva, Switzerland.

Kavasseri, R. G., & Seetharaman, K. (2009). Day-ahead wind speed forecasting using f-ARIMA models. *Renewable Energy, 34*(5), 1388-1393.

Kayacan, E., Uluta, B., & Kaynak, O. (2010). Grey system theory-based models in time series prediction. *Expert Systems with Applications, 37*, 1784-1789.

Kishore, R. A. & Priya, S. (2013). Design and experimental verification of a high efficiency small wind energy portable turbine (SWEPT). *Journal of Wind Engineering and Industrial Aerodynamics, 118,* 12-19.

Kim, D. & Hur, J. (2018). Short-term probabilistic forecasting of wind energy resources using the enhanced ensemble method. *Energy, 157*, 211-226.

Kopitz, J. & Polifke, W. (2008). CFD-based application of the Nyquist criterion to thermoacoustic instabilities. *Journal of Computational Physics, 227*, 6754-6778.

Korprasertsak, N. & Leephakpreeda, T. (2015). Optimal design of wind boosters for low speed vertical axis wind turbines. *Applied Mechanics and Materials, 798*, 195-199.

Korprasertsak, N. & Leephakpreeda, T. (2016). Analysis and optimal design of wind boosters for vertical axis wind turbines at low wind speed. *Journal of Wind Engineering and Industrial Aerodynamics, 159*, 9-18.

Korprasertsak, N. & Leephakpreeda, T. (2018). Short-term forecasting models of wind

Ref. code: 25625822300561TJA

speed for airborne wind turbines: a comparative study. *International Journal of Mechanical Engineering and Robotics Research, 7*(3), 250-256.

Kwon, J. Y., Park, S. W., Park, M. K., & Kang, M. G. (2016). Aliasing artifacts reduction with subband signal analysis for demosaicked images. *Digital Signal Processing, 59*, 115-128.

Lahouar, A. & Slama, J. B. H. (2017). Hour-ahead wind power forecast based on random forests. *Renewable Energy, 109*, 529-541.

Leephakpreeda, T. (2008). Grey prediction on indoor comfort temperature for HVAC systems. *Expert Systems with Applications, 34*(4), 2284-2289.

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics, 2*, 164-168.

Liu, H., Shi, J., & Erdem, E. (2010). Prediction of wind speed time series using modified Taylor Kriging method. *Energy, 35*, 4870-4879.

Lund, H. (2005). Large-scale integration of wind power into different energy systems. *Energy, 30*(13), 2402-2412.

Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics, 11*(2), 431-441.

Mason, K., Duggan, J., & Howley, E. (2018). Forecasting energy demand, wind generation and carbon dioxide emissions in Ireland using evolutionary neural networks. *Energy, 155*, 705-720.

Miranda, M. S., & Dunn, R. W. (2006). One-hour-ahead wind speed prediction using a Bayesian methodology. Proceeding of *2006 IEEE power engineering society general meeting* (pp. 1-6). Montreal, Canada: IEEE.

Mohammadi, K., Alavi, O., Mostafaeipour, A., Goudarzi, N. & Jalilvand, M. (2016). Assessing different parameters estimation methods of Weibull distribution to compute wind power density. *Energy Conversion and Management, 108*, 322-335.

Mukhopadhyay, S., Dash, D., Mitra, A., & Bhattacharya, P. (2014). A comparative study between seasonal wind speed by Fourier and Wavelet analysis. Retrieved from https://arxiv.org/ftp/arxiv/papers/1407/1407.8476.pdf.

Nagarajaiah, S. & Varadarajan, N. (2005). Short time Fourier transform algorithm for wind response control of buildings with variable stiffness TMD. *Engineering*

Ref. code: 25625822300561TJA

Structures, 27, 431-441.

Ozay, C. & Celiktas, M. S. (2016). Statistical analysis of wind speed using two-parameter Weibull distribution in Alaçatı region. *Energy Conversion and Management, 121*, 49-54.

Quan, P. & Leephakpreeda, T. (2015). Assessment of wind energy potential for selecting wind turbines: an application to Thailand. *Sustainable Energy Technologies and Assessments, 11*, 17-26.

Tabrizi, A. B., Whale, J., Lyons, T., & Urmee, T. (2015). Rooftop wind monitoring campaigns for small wind turbine applications: effect of sampling rate and averaging period. *Renewable Energy, 77*, 320-330.

Thaku, D. & Mithulananthan, N. (2010). Wind energy in Thailand to enhance energy security: potential, status and barriers. *International Energy Journal, 11*, 203-212.

Torres, J., Garcia, A., Blas, M. D., & Francisco, A. D. (2005). Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Solar Energy, 79*(1), 65-77.

Vasconcellos, R. T. B. & Campos, M. L. R. (2012). Error analysis in high-accuracy digital measurements. *Measurement, 45*, 881-888.

Watkins, J. C. (n.d.). *An Introduction to the Science of Statistics: From Theory to Implementation (Preliminary ed.).* Retrieved from http://math.arizona.edu/~jwatkins/statbook.pdf/.

Wang, C., Zhang, H., Fan, W., & Fan, X. (2016). A new wind power prediction method based on chaotic theory and Bernstein Neural Network. *Energy, 117*, 259-271.

Wang, K., Qi, X., Liu, H., & Song, J. (2018). Deep belief network based k-means cluster approach for short-term wind power forecasting. *Energy, 165*, 840-852.

Wang, X., Guo, P., & Huang, X. (2011). A Review of wind power forecasting models. *Energy Procedia, 12*, 770-778.

Xie, L., Gu, Y., Zhu, X., & Genton, M. (2014). Short-term spatiotemporal wind power forecast in robust look-ahead power system dispatch Smart Grid. I*EEE Transactions on Smart Grid, 5*(1), 511-520.

Yang, Z., Liu, Y., & Li, C. (2011). Interpolation of missing wind data based on ANFIS. *Renewable Energy, 36*, 993-998.

Ref. code: 25625822300561TJA

APPENDICES

Ref. code: 25625822300561TJA

APPENDIX A
CONVERSION OF UNEQUALLY-SAMPLED DATA

Table A1 Conversion from unequally-sampled data to equivalent equally-sampled data.

Wind speed				
Sampled at 10 Hz	Sampled at Nyquist rate			Equivalent equally-sampled data
Data	Cycle period = 1 s	Data	Weight ($f_s / f_s^{(N)}$)	Data
2.05	1st cycle period with 5 Hz	2.05	2	2.05
2.16				2.05
2.11		2.11	2	2.11
2.09				2.11
2.15		2.15	2	2.15
2.22				2.15
2.26		2.26	2	2.26
2.18				2.26
2.17		2.17	2	2.17
2.20				2.17
2.18	2nd cycle period with 2 Hz	2.18	5	2.18
2.12				2.18
2.34				2.18
2.12				2.18
2.32				2.18
2.12		2.12	5	2.12
2.30				2.12
2.18				2.12
2.30				2.12
2.21				2.12
2.40	3rd cycle period with 10 Hz	2.40	1	2.40
2.36		2.36	1	2.36
2.36		2.36	1	2.36
2.47		2.47	1	2.47
2.51		2.51	1	2.51
2.47		2.47	1	2.47
2.47		2.47	1	2.47
2.53		2.53	1	2.53
2.54		2.54	1	2.54
2.62		2.62	1	2.62
• •	• •	• •	• •	• •

Ref. code: 25625822300561TJA

APPENDIX B
SETUP DETAILS OF ANN MODEL

In the ANN model, the layrecnet function in MATLAB™ software is used to run the recurrent neural network with a single hidden layer. The number of nodes in the hidden layer is set to be 5 for best practice. In training, the trainlm function is implemented with the data of training set, validation set, and test set by the ratio of 70:15:15. The maximum number of epoch is set to be 1000 to ensure that the minimization of errors in training is completed. The RMSEs of the ANN with the four activation functions are obtained from forecasting the power generation data in Figure 4.10 where the hyperbolic tangent function (tanh) is applied in the ANN model since yields the lowest RMSE, according to Table B1.

Table B1 Performances of ANN model with different activation functions.

Activation function	RMSEs (kWh)
Hyperbolic tangent	150
Log-sigmoid	156
Rectified linear units	196
Softmax	228

Ref. code: 25625822300561TJA

APPENDIX C
ALTERNATIVE PRESENTATION OF PREDICTION INTERVAL

The prediction intervals in Figure 4.17, 4.18, and 4.19 are alternatively presented in continuous series along with the predicted values as shown in Figure C1, C2, and C3 in respectively.

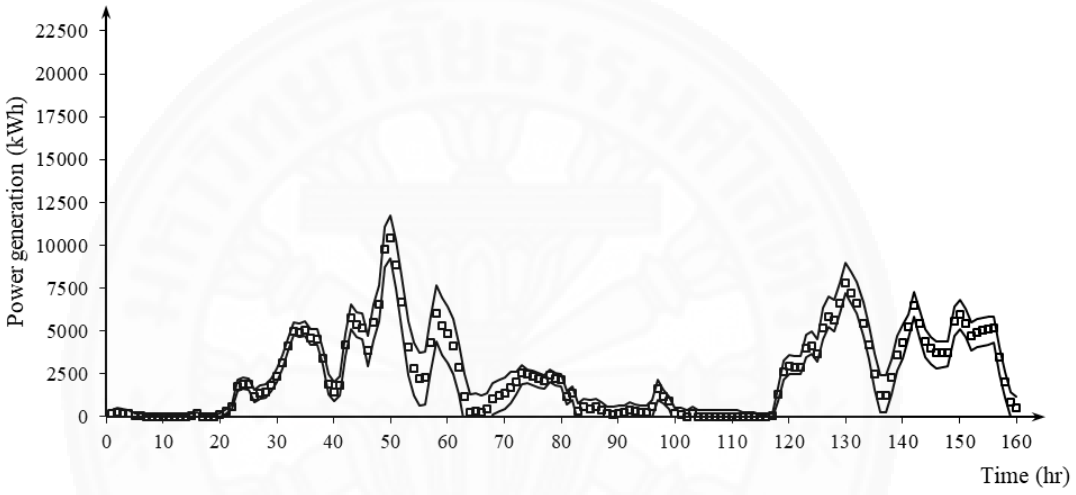

Figure C1 Plots of hourly predicted values and prediction intervals with 95% confidence in continuous trend line.

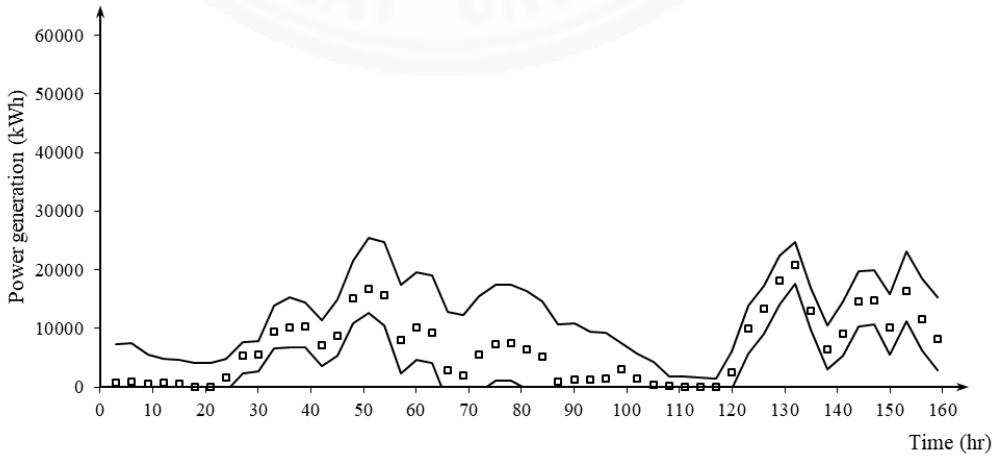

Figure C2 Plots of 3-hourly predicted values and prediction intervals with 95% confidence in continuous trend line.

Ref. code: 25625822300561TJA

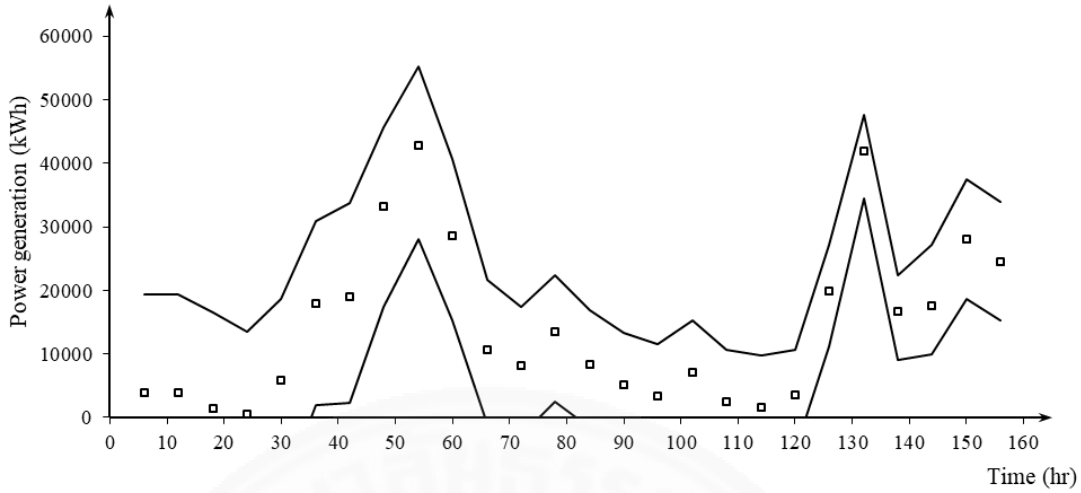

Figure C3 Plots of 6-hourly predicted values and prediction intervals with 95% confidence in continuous trend line.

Ref. code: 25625822300561TJA

APPENDIX D
WIND PROFILE POWER LAW

According to wind profile power law, wind speed at any altitude can be estimated from the known wind speed at the corresponding altitude. The profile power law is expressed by:

$$v = v_r\left(\frac{z}{z_r}\right)^{0.143} \tag{D1}$$

where v is the wind speed at the altitude z and v_r is the known wind speed at the reference altitude z_r.

The wind speed data in Figure 4.2 is extrapolated to various altitude as illustrated in Figure D1. The characteristic of the wind speed curves do not change by different altitude.

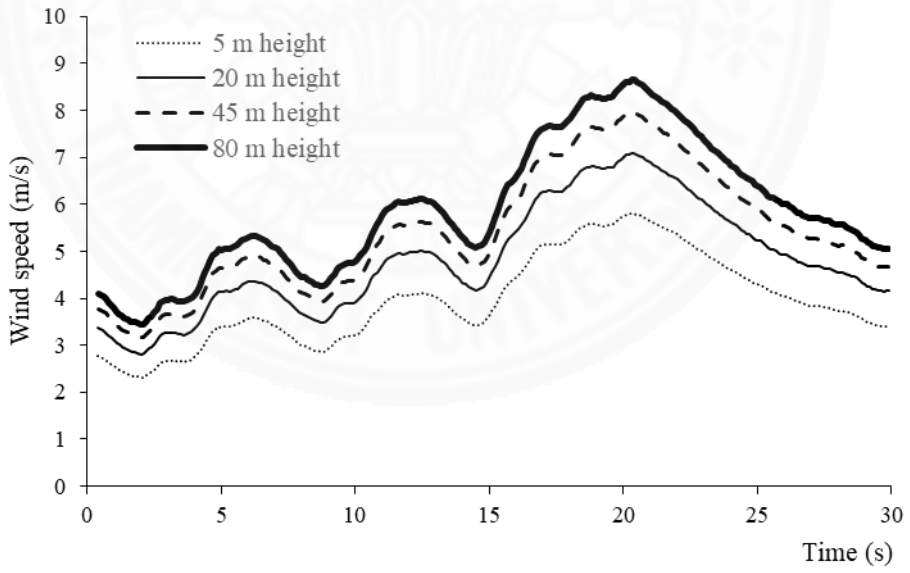

Figure D1 Illustration of wind speeds at various altitude.

Ref. code: 25625822300561TJA

BIOGRAPHY

Name	Mr. Natapol Korprasertsak
Date of Birth	October 22, 1990
Education	2012: Bachelor of Engineering (Mechanical Engineering) Sirindhorn International Institute of Technology/ Thammasat University 2015: Master of Science (Engineering and Technology) Sirindhorn International Institute of Technology/ Thammasat University

Publications

Korprasertsak, N., Korprasertsak, N, & Leephakpreeda, T. (2014). CFD modeling and design of wind boosters for low speed vertical axis wind turbines. *Advanced Materials Research, 1016,* 554-558.

Korprasertsak, N. & Leephakpreeda, T. (2014). CFD-based study of novel wind booster for low speed vertical axis wind turbine. Proceeding of *7th AUN/SEED-Net Regional Conference in Mechanical and Manufacturing Engineering* (pp. 301-304). Hanoi, Vietnam.

Korprasertsak, N. & Leephakpreeda, T. (2015). CFD-based power analysis on low speed vertical axis wind turbines with wind boosters. *Energy Procedia, 79,* 963-968.

Ngamlarp, S., Korprasertsak, N, & Leephakpreeda, T. (2015). Experimental study of novel wind booster for low speed vertical axis wind turbine. Proceeding of *2015 International Conference on Alternative Energy in Developing Countries and Emerging Economies* (pp. 179-180). Bangkok, Thailand.

Korprasertsak, N. & Leephakpreeda, T. (2015). Optimal design of wind boosters for low speed vertical axis wind turbines. *Applied Mechanics and Materials, 798,* 195-199.

Ref. code: 25625822300561TJA

Korprasertsak, N. & Leephakpreeda, T. (2016). Analysis and optimal design of wind boosters for vertical axis wind turbines at low wind speed. *Journal of Wind Engineering and Industrial Aerodynamics, 159,* 9-18.

Korprasertsak, N. & Leephakpreeda, T. (2017). Comparative investigation of short-term wind speed forecasting models for airborne wind turbines. Proceeding of *8th International Conference on Mechanical and Aerospace Engineering* (pp. 451-454). Prague, Czech Republic.

Korprasertsak, N. & Leephakpreeda, T. (2018). Nyquist-based adaptive sampling rate for wind measurement under varying wind conditions. *Renewable Energy, 119,* 290-298.

Korprasertsak, N. & Leephakpreeda, T. (2018). Short-term forecasting models of wind speed for airborne wind turbines: a comparative study. *International Journal of Mechanical Engineering and Robotics Research, 7*(3), 250-256.

Korprasertsak, N. & Leephakpreeda, T. (2019). Improving accuracy of wind analysis with multi sampling rates of wind measurement. Proceeding of *3rd International Conference on Power, Energy and Mechanical Engineering* (E3S web of conferences). Prague, Czech Republic.

Korprasertsak, N. & Leephakpreeda, T. (2019). Robust short-term prediction of wind power generation under uncertainty via statistical interpretation of multiple forecasting models. *Energy, 180,* 387-397.

Ref. code: 25625822300561TJA