



THE HONG KONG  
POLYTECHNIC UNIVERSITY

香港理工大學

Pao Yue-kong Library

包玉剛圖書館

---

## Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

**By reading and using the thesis, the reader understands and agrees to the following terms:**

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

### IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact [lbsys@polyu.edu.hk](mailto:lbsys@polyu.edu.hk) providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

UNCERTAINTY ANALYSIS, MODELLING AND  
PROGNOSIS FOR POWER SYSTEM OPERATION  
AND PLANNING

CAN WAN

Ph.D

The Hong Kong Polytechnic University

2015



The Hong Kong Polytechnic University

Department of Electrical Engineering

---

UNCERTAINTY ANALYSIS, MODELLING AND  
PROGNOSIS FOR POWER SYSTEM OPERATION  
AND PLANNING

**CAN WAN**

A thesis submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy

June 2014

# **CERTIFICATE OF ORIGINALITY**

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written or material which has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

\_\_\_\_\_ (Signed)

\_\_\_\_\_ WAN Can (Name of student)

## ***Abstract***

The introduction of smart grid, renewable energy, and electricity market, etc., has led to tremendous changes to modern power systems. Particularly, as one of the most important renewable energy, wind power has experienced dramatic growth worldwide. However, due to the intermittency, wind power generation also brings challenges to many aspects of power system operation and planning. Wind power is directly related to wind speed. Proper probability model for describing the stochastic wind speed would be essential for wind farm planning, power system analysis, and so forth. Traditional Weibull and Rayleigh distributions cannot accurately capture the wind speed properties that exhibit significant non-stationarities and complexities. In this thesis, an advanced generalized Lambda distribution is proposed to statistically model wind speed that can achieve a superior performance.

Considering the short-term operational planning horizon, such as from 1 hour to 48 hours, accurate forecasting of wind power generation becomes crucial to ensure secure and reliable management of power system operation. Traditionally, research efforts are paid to developing point prediction methodologies of wind power, which corresponds to the exact mathematical expectation of the stochastic wind power series at a given prediction horizon. Because of the chaotic nature of weather systems, wind power prediction errors cannot be avoided and can be significant in some conditions. Therefore, transformation from traditional point forecasts to probabilistic interval forecasts can be of great importance to quantify the uncertainties of future forecasts, thus effectively supporting the decision making activities against uncertainties and risks ahead. To this end, this thesis has proposed novel probabilistic forecasting methodologies to quantify the uncertainty involved in wind power forecasting. Theoretical background of probabilistic forecasting and the state-of-the-art approaches for wind power prediction are thoroughly reviewed. Subsequently, a parametric probabilistic forecasting approach, bootstrap-based extreme learning machine, is developed to generate prediction intervals, which can be trained at an extremely fast speed. In addition, two nonparametric probabilistic forecasting approaches using extreme

learning machine based forecaster are also developed, and possess significant advantages such as a much simplified problem formulation and no need of prior knowledge about point forecasting errors. The first one is the direct interval forecasting approach that is proposed to directly produce optimal prediction intervals in terms of the novel cost function combining reliability and interval score. The second one is the Pareto optimal interval forecasting approach which aims to construct optimal prediction intervals through reaching the Pareto front of two quality index reliability and sharpness.

The electricity market has important influence on the management of modern power systems, e.g., encouraging the development of renewable energy and the participation of consumers in smart grids. Precise electricity price prediction would assist market participants to properly deal with various decision making problems. Similar to wind power, electricity price also demonstrates significant nonstationarities and is fairly difficult to accurately forecast. An advanced hybrid approach is proposed for electricity price forecasting in this thesis to estimate prediction uncertainty of electricity price.

Probabilistic load flow computation is implemented considering the power system integrated with large wind farm. Based on the developed generalized Lambda distribution model for wind speed, the probability property of wind power can be described accordingly. Comprehensive numerical studies demonstrate that the point estimate method can give accurate estimation of probabilistic load flow in the environment with wind power integration, which can help to investigate the impacts of wind power and facilitate wind farm planning, power system operational and expansion planning, etc.

## *Acknowledgements*

This work is accomplished during my Ph.D. candidature in The Hong Kong Polytechnic University (HKPU) and my visit in Technical University of Denmark (DTU) and Tsinghua University, with the immense support and guidance of various personalities. First, I offer my sincere appreciation to my current chief supervisor, Dr. Zhao Xu, for his kind support and precious guidance to my research during the past three years. I would like to express my sincere thanks to ex-chief supervisor of my Ph.D. candidature, Prof. Zhao Yang Dong, for his invaluable supervision that indeed guides me to enter the research fields of power system engineering. I also would like to express my deep gratitude to my co-supervisor Prof. Kit Po Wong, for his continuous help and constructive advices during my Ph.D. study. I must thank Prof. Pierre Pinson, at the Centre for Electric Power and Energy (CEE) of DTU, who provides really valuable advices and expertise in wind power forecasting. My heartfelt thanks go to Prof. Yonghua Song, for his kind support and help to my visit and research at State Key Laboratory of Power Systems, Department of Electrical Engineering, Tsinghua University.

The path I was following does not certainly guide me to pursue my Ph.D. degree. I must express the most sincere thanks to them again for their priceless help and support. Specially, I also acknowledge the support from Hong Kong Ph.D. Fellowship awarded by Research Grants Council (RGC) of Hong Kong.

Further, I would like to thank Prof. Jacob Østergaard, Prof. Arne Hejde Nielsen and Dr. Guangya Yang, for their supports during my half-year visit in CEE of DTU. I am thankful to Dr. Zechun Hu for his kind help during my visit in Tsinghua University. I must express my sincere appreciation to Dr. Dibo Hou for his considerable help during my undergraduate study in Zhejiang University.

I am thankful to my fellow research mates especially Mr. Ming Niu, Mr. Jian Zhao, Ms. Jia Hou and Mr. Yelei Wang, for giving precious time to help me, share experiences and make progress together. Your friendship and accompany is really cherished. I hope you all will be successful with your Ph.D. studies. I also

would like to extend my thanks to all staffs of the Department of Electrical Engineering in HKPU, for their help and assistance provided, and also to my paper reviewers and thesis examiners for their valuable comments to this research.

Special thanks go to my relatives and people that I consider as relatives. My genuine appreciation and love are given to all my family for their support. More particularly, I thank my parents and grandparents for their unselfish love and encouragement, whenever and whatever, which have been deeply engraved on my heart. I would like to express my special thanks to my girlfriend, who has always been there for me. It was her understanding, encouragement and loving support that kept me focused and motivated. I would like to thank my childhood friends for the happy, unforgettable and helpful growth experiences. My friends in Zhejiang University, I will remember your encouragements and our experiences of pursuing dreams forever. Last but not least, I wish to thank all helpful people and fortunate events.

## ***Publications Arising from This Thesis***

### **Refereed Journal Publications**

1. **C. Wan**, Z. Xu, P. Pinson, Z.Y. Dong, and K.P. Wong, “Probabilistic forecasting of wind power generation using extreme learning machine,” **Invited paper**, *IEEE Trans. Power Systems*, vol.29, no.3, pp.1033-1044, May 2014.
2. **C. Wan**, Z. Xu, P. Pinson, Z.Y. Dong, and K.P. Wong, “Optimal prediction intervals of wind power generation,” *IEEE Trans. Power Systems*, vol.29, no.3, pp.1166-1174, May 2014.
3. **C. Wan**, Z. Xu, and P. Pinson, “Direct interval forecasting of wind power,” *IEEE Trans. Power Systems*, vol. 28, no. 4, pp. 4877-4878, Nov. 2013.
4. **C. Wan**, Z. Xu, Y. L. Wang, Z.Y. Dong, and K.P. Wong, “A hybrid approach for probabilistic forecasting of electricity price,” *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 463-470, Jan. 2014.
5. **C. Wan**, Z. Xu, J. Østergaard, Z. Y. Dong, and K. P. Wong, “Discussion on ‘Combined nonparametric prediction intervals for wind power generation’,” *IEEE Trans. Sustainable Energy*, vol. 5, no. 3, pp. 1021, Jul. 2014.
6. Z. Xu, and **C. Wan**, “The key technology for grid integration of wind power: Direct probabilistic interval forecasts of wind power,” *Southern Power System Technology*, vol. 7, no. 5, pp. 1-9, 2013.
7. E.B. Cao, **C. Wan**, M.Y. Lai, “Coordination of a supply chain with one manufacturer and multiple competing retailers under simultaneous demand and cost disruptions,” *In. J. Production Economics*, vol. 141, no. 1, pp. 425-433, Jan. 2013.
8. E. B. Cao, Y. J. Ma, **C. Wan**, and M. Y. Lai, “Contracting with asymmetric cost information in a dual-channel supply chain”, *Operations Research Letters*, vol. 41, no. 4, pp. 410-414, 2013.
9. M. Niu, **C. Wan**, and Z. Xu, “A review on applications of heuristic optimization algorithms for optimal power flow in modern power systems”, *J.*

*Modern Power Systems and Clean Energy*, vol. 2, no. 4, pp. 289-297, Dec. 2014.

#### **Journal papers in progress or under review**

10. **C. Wan**, Z. Xu, J. Østergaard, Z. Y. Dong, and K. P. Wong, “Estimation of wind speed probability distributions using generalized Lambda distribution”, *Applied Energy*, 2013. Submitted
11. **C. Wan**, Z. Xu, G.Y. Yang, P. Pinson, A.H. Nielsen, and J. Østergaard, “Probabilistic wind power forecasting with hybrid artificial neural networks,” *Wind Energy*, 2013. Submitted
12. **C. Wan**, Z. Xu, and P. Pinson, “Comments on ‘Comprehensive review of neural network-based prediction intervals and new advances’ and ‘Lower upper bound estimation method for construction of neural network-based prediction intervals’,” *IEEE Trans. Neural Networks and Learning Systems*, 2014. Submitted
13. **C. Wan**, R. Wang, Z. Xu, and P. Pinson, “Optimal construction of neural network based prediction intervals,” *IEEE Trans. Neural Networks and Learning Systems*, 2014. Submitted
14. **C. Wan**, M. Niu, Z. Xu, P. Pinson, Z.Y. Dong, J. Østergaard, and K. P. Wong, “Pareto optimal interval forecasts of wind power generation,” *IEEE Trans. Smart Grid*, 2014. Submitted
15. J. Zhao, **C. Wan**, Z. Xu, J. Wang, and K.P. Wong, “Risk-based day-ahead scheduling of electric vehicle aggregator using information gap decision theory,” to be submitted to *IEEE Trans. Power Systems*.
16. M. Niu, **C. Wan**, and Z. Xu, “A novel formulation of multi-objective optimal power flow based on the concept of data envelopment analysis,” to be submitted to *IEEE Trans. Power Systems*.

#### **Refereed papers in Conference Proceedings**

17. **C. Wan**, Z. Xu, Z.Y. Dong, and K.P. Wong, “Probabilistic load flow computation using first-order second-moment method”, in *Proc. IEEE Power & Energy Society General Meeting*, Jul. 2012.

18. **C. Wan**, Z.Y. Dong, F. J. Luo, M. Lu, P. Zhang, and K.P. Wong, “Cumulant-based probabilistic load flow calculation in a market environment”, in *Proc. 2011 4<sup>th</sup> Int. Conf. on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, pp. 1190-1195, Jul. 2011.
19. **C. Wan**, Z.Y. Dong, F. J. Luo, K. Meng, P. Zhang, K.P. Wong and L. Zhuang, “A point estimate method based probabilistic load flow in an electricity market environment”, *In. Conf. on Power and Energy Systems*, Jun. 2011.
20. F.J. Luo, Z.Y. Dong, **C. Wan**, Y.Y. Chen, K. Meng, and K.P. Wong, “A computational grid platform for large-scale power system analysis”, *In. Conf. on Power and Energy Systems*, Jun. 2011.
21. **C. Wan**, Z.Y. Dong, F.J. Luo, Y.Y. Chen, K. Meng, and K.P. Wong, “A novel XML-based resource modeling framework for power system heterogeneous data integrating”, *Applied Mechanics and Materials*, vol. 58-60, pp. 1476-1481, Jun. 2011.
22. F.J. Luo, Z.Y. Dong, **C. Wan**, Y.Y. Chen, K. Meng, and K.P. Wong, “Applying computational grid technology to power system”, *Applied Mechanics and Materials*, vol. 58-60, pp. 1442-1447, Jun. 2011.

#### **Patent**

23. **C. Wan** and Z. Xu, “Adaptive and Generalized Probabilistic Forecasting Method and System.” U.S. Patent, 2014, under application.



# *Table of Contents*

<b>ABSTRACT .....</b>	<b>I</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>III</b>
<b>PUBLICATIONS ARISING FROM THIS THESIS .....</b>	<b>V</b>
<b>TABLE OF CONTENTS .....</b>	<b>IX</b>
<b>LIST OF ABBREVIATIONS .....</b>	<b>XIII</b>
<b>LIST OF TABLES.....</b>	<b>XV</b>
<b>LIST OF FIGURES.....</b>	<b>XVII</b>
<b>1 INTRODUCTION .....</b>	<b>1</b>
1.1 GENERAL CONTEXT .....	1
1.2 PURPOSE OF THE WORK.....	5
1.3 PRIMARY CONTRIBUTIONS.....	6
1.4 STRUCTURE OF THE THESIS .....	8
<b>2 ESTIMATION OF WIND SPEED PROBABILITY DISTRIBUTIONS USING GENERALIZED LAMBDA DISTRIBUTION .....</b>	<b>11</b>
2.1 INTRODUCTION .....	11
2.2 CONVENTIONAL PROBABILITY MODELS .....	13
2.2.1 <i>Probability Models</i> .....	13
2.2.2 <i>Parameter Estimation of Conventional Probability Models</i> .....	17
2.3 GENERALIZED LAMBDA DISTRIBUTION.....	17
2.3.1 <i>Model Definition</i> .....	17
2.3.2 <i>Parameter space and support region</i> .....	18
2.3.3 <i>Parameter estimation</i> .....	19
2.4 FITTING PERFORMANCE TEST AND EVALUATION.....	23
2.4.1 <i>Kolmogorov–Smirnov Test</i> .....	23
2.4.2 <i>Chi-Square Test</i> .....	24
2.4.3 <i>Root Mean Square Error</i> .....	24
2.4.4 <i>Comprehensive Fitting Quality Evaluation</i> .....	25
2.5 NUMERICAL RESULTS AND ANALYSIS .....	26
2.5.1 <i>Wind speed data</i> .....	26
2.5.2 <i>Comparison of Fitting Performance</i> .....	27
2.6 CONCLUSIONS.....	38
<b>3 PROBABILISTIC FORECASTING: PRINCIPLE AND EVALUATION .....</b>	<b>39</b>
3.1 INTRODUCTION .....	39
3.2 PREDICTION INTERVALS.....	42
3.3 EVALUATION CRITERIA .....	44
3.3.1 <i>Reliability</i> .....	44
3.3.2 <i>Sharpness</i> .....	45

3.3.3	<i>Skill Score</i> .....	46
3.4	CONCLUSIONS .....	48
<b>4</b>	<b>STATE-OF-THE-ART IN PROBABILISTIC WIND POWER FORECASTING</b> .....	<b>51</b>
4.1	CHARACTERISTICS AND POINT FORECASTING OF WIND POWER GENERATION .....	51
4.1.1	<i>Characteristics of Wind Power Generation</i> .....	51
4.1.2	<i>Major Aspects of Wind Power Prediction</i> .....	53
4.1.3	<i>Wind power forecasting models</i> .....	54
4.2	PROBABILISTIC WIND POWER FORECASTING APPROACHES .....	57
4.2.1	<i>Statistical Approaches</i> .....	57
4.2.2	<i>Ensemble Forecasts</i> .....	59
4.3	PRACTICAL APPLICATIONS IN POWER SYSTEMS .....	60
4.4	CONCLUSIONS .....	63
<b>5</b>	<b>AN ADVANCED STATISTICAL APPROACH FOR PROBABILISTIC FORECASTING OF WIND POWER GENERATION</b> .....	<b>65</b>
5.1	INTRODUCTION .....	65
5.2	EXTREME LEARNING MACHINE .....	67
5.3	FORMULATION OF PREDICTION INTERVALS .....	70
5.3.1	<i>Uncertainty Analysis</i> .....	70
5.3.2	<i>Prediction Interval Formulation</i> .....	71
5.4	PREDICTION INTERVALS CONSTRUCTION .....	73
5.4.1	<i>Bootstrap Methods</i> .....	73
5.4.2	<i>Model Uncertainty Variance</i> .....	74
5.4.3	<i>Data Noise Variance</i> .....	75
5.5	NUMERICAL STUDIES .....	77
5.5.1	<i>Description of Experiment Data</i> .....	77
5.5.2	<i>Determination of ELM Hidden Nodes Number</i> .....	77
5.5.3	<i>Comparison of Bootstrap Methods</i> .....	79
5.5.4	<i>Analysis of Forecasting Results</i> .....	79
5.6	CONCLUSIONS .....	89
<b>6</b>	<b>DIRECT INTERVAL FORECASTING OF WIND POWER GENERATION</b> .....	<b>91</b>
6.1	INTRODUCTION .....	91
6.2	FORMULATION OF PIS .....	93
6.3	OPTIMAL CONSTRUCTION OF PIS .....	95
6.3.1	<i>Objective Function</i> .....	95
6.3.2	<i>Particle Swarm Optimization</i> .....	96
6.3.3	<i>DIF Algorithm for PI Optimization</i> .....	97
6.4	CASE STUDIES .....	100
6.4.1	<i>Introduction of Experiment Data</i> .....	100
6.4.2	<i>Experiment Result and Analysis</i> .....	100
6.5	CONCLUSIONS .....	107
<b>7</b>	<b>PARETO OPTIMAL INTERVAL FORECASTS OF WIND POWER: A CASE STUDY ON BORNHOLM ISLAND, DENMARK</b> .....	<b>109</b>
7.1	INTRODUCTION .....	109
7.2	OPTIMIZATION FORMULATION AND PROCEDURES .....	111
7.2.1	<i>Formulation of Optimization Objectives</i> .....	111
7.2.2	<i>Pareto Optimality</i> .....	111

7.2.3	<i>NSGA-II</i> .....	112
7.2.4	<i>Reliability Precision</i> .....	114
7.2.5	<i>Overall Framework and Model Selection</i> .....	115
7.3	CASE STUDIES .....	118
7.3.1	<i>Wind Farm Data</i> .....	118
7.3.2	<i>Numerical Analysis and Discussions</i> .....	119
7.4	CONCLUSIONS.....	127
<b>8</b>	<b>A HYBRID APPROACH FOR PROBABILISTIC FORECASTING OF ELECTRICITY PRICE.....</b>	<b>129</b>
8.1	INTRODUCTION .....	129
8.2	PIS FORMULATION AND CONSTRUCTION .....	132
8.2.1	<i>Formulation</i> .....	132
8.2.2	<i>Variance of Model Uncertainty</i> .....	133
8.2.3	<i>Variance of Noise</i> .....	134
8.2.4	<i>Overall Procedure</i> .....	135
8.3	NUMERICAL STUDIES.....	135
8.3.1	<i>Electricity Market and Experiment Data</i> .....	135
8.3.2	<i>Comparison Analysis</i> .....	136
8.4	CONCLUSIONS.....	143
<b>9</b>	<b>PROBABILISTIC LOAD FLOW CONSIDERING THE UNCERTAINTY OF WIND POWER .....</b>	<b>145</b>
9.1	INTRODUCTION .....	145
9.2	LOAD FLOW SOLUTIONS.....	148
9.2.1	<i>AC Load Flow</i> .....	148
9.2.2	<i>Linearized Load Flow Equation</i> .....	149
9.3	MODEL OF INPUT UNCERTAINTY .....	150
9.3.1	<i>Injected Power</i> .....	150
9.3.2	<i>Bus Load</i> .....	150
9.3.3	<i>Wind Power Output</i> .....	151
9.4	MATHEMATICAL BACKGROUND .....	152
9.4.1	<i>Monte Carlo Simulation</i> .....	152
9.4.2	<i>First-order Second-moment Method</i> .....	153
9.4.3	<i>Point Estimate Method</i> .....	155
9.5	SIMULATION TEST RESULTS.....	157
9.6	CONCLUSIONS.....	162
<b>10</b>	<b>GENERAL CONCLUSIONS .....</b>	<b>165</b>
10.1	CONCLUSIONS.....	165
10.2	PERSPECTIVES .....	168
	<b>REFERENCES .....</b>	<b>171</b>



## *List of Abbreviations*

ACE	Average coverage error
AEMO	Australian Energy Market Operator
ANEM	Australian National Electricity Market
ANNs	Artificial neural networks
ARIMA	Autoregressive integrated moving average
AWPI	Average width of prediction intervals
AWPPS	Armines Wind Power Prediction System
AWPT	Advanced Wind Power Prediction Tool
BELM	Bootstrap based extreme learning machine
BNN	Bootstrap based traditional neural networks
CDFs	Cumulative distribution functions
CMS	Combined moment and starship
CWC	Coverage width-based criterion
DIF	Direct interval forecasting
DLF	Deterministic load flow
ECDF	Empirical cumulative distribution function
ECP	Empirical coverage probability
ESM	Exponential smoothing method
ELM	Extreme learning machine
FOSM	First-order second-moment method
GARCH	Generalized autoregressive conditional heteroscedasticity
GENCO	Generation company
GLD	Generalized Lambda distribution

KDE	Kernel density estimation
LUBE	Lower upper bound estimation
MAE	Mean absolute error
MCPs	Market clearing prices
MCS	Monte Carlo simulation
MLE	Maximum likelihood estimation
MUIs	Model uncertainty intervals
NCP	Nominal coverage probability
NDBC	National Data Buoy Centre
NSW	New South Wale
NNs	Neural networks
NSGA-II	Non-dominated sorting genetic algorithm-II
NWP	Numerical weather prediction
PDFs	Probability density functions
PEM	Point estimate method
PIs	Prediction intervals
PLF	Probabilistic load flow
PSO	Particle swarm optimization
QR	Quantile regression
RMSE	Root mean square error
SLFNs	Single-hidden layer feedforward neural networks
SVM	Support vector machine
TSO	Transmission system operator
WPPT	Wind Power Prediction Tool

## *List of Tables*

<b>Table 2.1</b>	The parameter space and support region of GLD [68].	19
<b>Table 2.2</b>	The first four moments of observed samples.	21
<b>Table 2.3</b>	Specific information of the fifteen stations	26
<b>Table 2.4</b>	K-S test results of different probability models for historical wind speed data measured at the fifteen stations.	28
<b>Table 2.5</b>	$\chi^2$ test results of different probability models for historical wind speed data measured at the fifteen stations.	29
<b>Table 2.6</b>	RMSE test results of different probability models for historical wind speed data measured at the fifteen stations.	30
<b>Table 2.7</b>	Comprehensive goodness-of-fit results of different probability models for historical wind speed data measured at the fifteen stations.	31
<b>Table 5.1</b>	Evaluation results of different bootstrap methods.	79
<b>Table 5.2</b>	Results of PIs in summer 2012.	81
<b>Table 5.3</b>	Results of PIs in autumn 2011.	81
<b>Table 5.4</b>	Results of PIs in winter 2011.	82
<b>Table 5.5</b>	Results of PIs in spring 2010.	82
<b>Table 5.6</b>	Reliability and efficiency of different bootstrap replicates	86
<b>Table 5.7</b>	Effects of bootstrapping on resultant PIs	87
<b>Table 5.8</b>	Results of multi-step intra-hour forecasting	88
<b>Table 6.1</b>	Results of different methods in Challicum Hills wind farm.	102
<b>Table 6.2</b>	Results of different methods in Starfish Hill wind farm	103
<b>Table 7.1</b>	Prediction results with look-ahead time of 30 minutes	123
<b>Table 7.2</b>	Prediction results with look-ahead time of one hour.	124
<b>Table 7.3</b>	Prediction results with look-ahead time of two hours	124
<b>Table 8.1</b>	Evaluation results of PIs on MCP <sub>1</sub>	137

<b>Table 8.2</b>	Evaluation results of PIs on MCP <sub>2</sub> .....	137
<b>Table 8.3</b>	Evaluation results of PIs on MCP <sub>3</sub> .....	138
<b>Table 8.4</b>	Evaluation results of PIs on MCP <sub>4</sub> .....	138
<b>Table 8.5</b>	Training time comparison .....	140
<b>Table 9.1</b>	Mean and standard deviation of active power flow for Case 1 .....	159
<b>Table 9.2</b>	Mean and standard deviation of active power flow for Case 2.....	160
<b>Table 9.3</b>	Mean and standard deviation of active power flow for Case 3.....	160
<b>Table 9.4</b>	Average error indices for studied cases .....	161
<b>Table 9.5</b>	Computation time of different methods.....	162

## *List of Figures*

<b>Figure 1.1</b> Global cumulative installed wind power capacity from 1996 to 2013 (Data from [17]).	2
<b>Figure 1.2</b> Top ten cumulative installed wind power capacities in 2013.	3
<b>Figure 2.1</b> Histogram and probability distributions for wind speed at the station Settlement Point.	32
<b>Figure 2.2</b> Histogram and probability distributions for wind speed at the station Passage Island.	33
<b>Figure 2.3</b> Histogram and probability distributions for wind speed at the station Delaware Bay.	33
<b>Figure 2.4</b> Histogram and probability distributions for wind speed at the station Newport.	33
<b>Figure 2.5</b> Histogram and probability distributions for wind speed at the station South Bass Island.	34
<b>Figure 2.6</b> Histogram and probability distributions for wind speed at the station Dunkirk.	34
<b>Figure 2.7</b> Histogram and probability distributions for wind speed at the station Folly Island.	34
<b>Figure 2.8</b> Histogram and probability distributions for wind speed at the station Keaton Beach.	35
<b>Figure 2.9</b> Histogram and probability distributions for wind speed at the station Flat Island Light.	35
<b>Figure 2.10</b> Histogram and probability distributions for wind speed at the station Five Finger.	35
<b>Figure 2.11</b> Histogram and probability distributions for wind speed at the station Grand Isle.	36
<b>Figure 2.12</b> Histogram and probability distributions for wind speed at the station Fowey Rocks.	36
<b>Figure 2.13</b> Histogram and probability distributions for wind speed at the station Smith Island.	36

<b>Figure 2.14</b> Histogram and probability distributions for wind speed at the station Lake Ontario.....	37
<b>Figure 2.15</b> Histogram and probability distributions for wind speed at the station Christmas Island. ....	37
<b>Figure 3.1</b> Point forecasting and probabilistic forecasting of wind power over a realistic wind farm in Australia .....	40
<b>Figure 4.1</b> Wind power curve of the Vestas V80 (2.0 MW) wind turbine installed at wind farm on Bornholm Island of Denmark.....	52
<b>Figure 5.1</b> The typical structure of ELM.....	68
<b>Figure 5.2</b> The framework for PIs construction of the proposed BELM approach. ....	76
<b>Figure 5.3</b> Validation test for ELMs with different numbers of hidden neurons. .	78
<b>Figure 5.4</b> PIs with NCP 90% in summer 2012 obtained by the proposed BELM approach. ....	83
<b>Figure 5.5</b> PIs with NCP 90% in autumn 2011 obtained by the proposed BELM approach. ....	83
<b>Figure 5.6</b> PIs with NCP 90% in winter 2011 obtained by the proposed BELM approach. ....	84
<b>Figure 5.7</b> PIs with NCP 90% in spring 2010 obtained by the proposed BELM approach. ....	85
<b>Figure 6.1</b> The ELM model for PIs generation by the proposed DIF approach. ...	94
<b>Figure 6.2</b> The overall implementation procedures of the proposed DIF approach. ....	99
<b>Figure 6.3</b> PIs with NCP 90% in March 2010 of the Challicum Hills wind farm obtained the proposed DIF approach.....	104
<b>Figure 6.4</b> PIs with NCP 90% in June 2010 of the Challicum Hills wind farm obtained by the proposed DIF approach.....	105
<b>Figure 6.5</b> PIs with NCP 90% in February 2010 of the Starfish Hill wind farm obtained by the proposed DIF approach.....	105
<b>Figure 6.6</b> PIs with NCP 90% in May 2010 of the Starfish Hill wind farm obtained by the proposed DIF approach.....	106
<b>Figure 7.1</b> The overall flowchart of the proposed algorithm. ....	116
<b>Figure 7.2</b> The framework of the developed forecasting system by the proposed approach. ....	117

<b>Figure 7.3</b> The wind farm of Bornholm Island (courtesy of Centre for Electric Power and Energy, Technical University of Denmark). .....	118
<b>Figure 7.4</b> Time series with wind power (10-min resolution).....	119
<b>Figure 7.5</b> The Pareto front and identified optimal models of PIs with NCP 90% and look-ahead time 30 minutes. ....	121
<b>Figure 7.6</b> The Pareto front and identified optimal models of PIs with NCP 90% and look-ahead time 1 hour.....	122
<b>Figure 7.7</b> The Pareto front and identified optimal models of PIs with NCP 90% and look-ahead time 2 hours. ....	122
<b>Figure 7.8</b> PIs with NCP 90% and 30min look-ahead time in December 2012 in the wind farm of Bornholm Island. ....	125
<b>Figure 7.9</b> PIs with NCP 90% and 1hour look-ahead time in December 2012 in the wind farm of Bornholm Island.....	126
<b>Figure 7.10</b> PIs with NCP 90% and 2 hour look-ahead time in December 2012 in the wind farm of Bornholm Island. ....	126
<b>Figure 8.1</b> The overall framework of PIs construction by the proposed hybrid approach. ....	135
<b>Figure 8.2</b> PIs with nominal confidence 90% of $MCP_1$ obtained by the proposed hybrid approach. ....	141
<b>Figure 8.3</b> PIs with nominal confidence 90% of $MCP_2$ obtained by the proposed hybrid approach. ....	141
<b>Figure 8.4</b> PIs with nominal confidence 90% of $MCP_3$ obtained by the proposed hybrid approach. ....	142
<b>Figure 8.5</b> PIs with nominal confidence 90% of $MCP_4$ obtained by the proposed hybrid approach .....	142
<b>Figure 9.1</b> Diagram of IEEE 39 test system.....	158



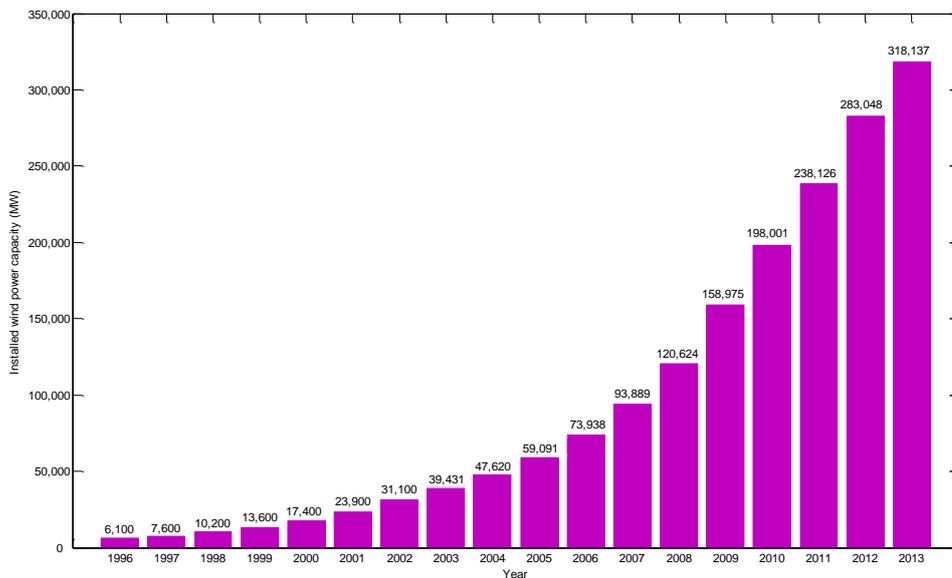
# ***1 Introduction***

## **1.1 General Context**

Electric power is the backbone of modern industrial society and closely related to economy development and people's common life, such as transport, heating, lighting, communications, and computation, and so forth [1]. Power system is a network of electrical components used to supply, transmit and use electric power, which can be separated into three parts: 1) the generators that supply the power, 2) the transmission system that carries the electric power from the generators to the load points, 3) the distribution system that feeds the power to nearby homes and industries [2].

Recently, smart grid has attracted a wide attention, which is defined as an advanced electrical network that collects information about all entities connected to it, including the behaviors of suppliers and consumers, and the acts of generators, transmission and distribution equipment, to improve the efficiency, reliability, economy, and sustainability of the production and distribution of electricity [3, 4]. It enables the implementation of advanced metering infrastructure for load management, the integration of distributed renewable energy generation and the formation of microgrids, and so on [5]. Flexible and reliable integration of renewable energy should be one of the basic characteristics of modern power systems. Due to the climate change and energy crisis, renewable energy has drawn many attentions nowadays [6-8]. Particularly, prospective studies towards 100% renewable energy-systems for Denmark, Australia, and Ireland have been simulated in [6, 9, 10]. So far, renewable energy technologies adopted for generation mainly include wind power, hydropower, photovoltaic power, biomass, biofuel, geothermal energy, and so forth [11]. Among others, wind power is regarded one of the most efficient and important

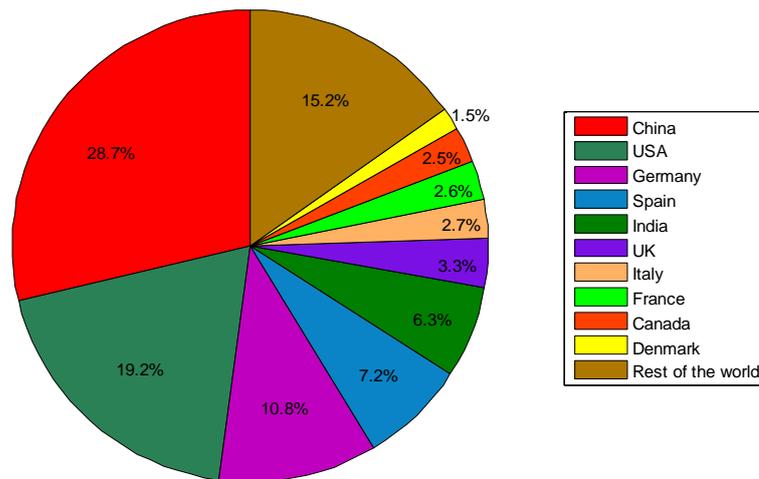
renewable energy resources for electricity generation, penetrated to modern power systems with large scale [12-15]. Many attentions have been drawn to wind energy technology development under the smart grid environment [16]. In the past decades, wind power generation has experience a dramatic growth all over the world. The cumulative installed wind capacity reached nearly 320GW by the end of 2013. Figure 1.1 demonstrates the cumulative wind power from 1996 to 2013, increasing more than 50 times [17]. The ten countries including China, USA, Germany, Spain, India, UK, Italy, France, Canada, and Denmark have the highest installed wind power capacity, and the corresponding percentage is explicitly depicted in Figure 1.2 [17].



**Figure 1.1** Global cumulative installed wind power capacity from 1996 to 2013 (Data from [17]).

In Denmark, the power system has reached a wind share up to approximately 20% for more than ten years [18]. Danish power system would expand the penetration of wind energy to 50% by 2025 according to the national strategy [19]. As an intermittent energy, wind power generation introduces high level of uncertainties, which makes the power system planning and operation much harder than before. A number of approaches have been proposed for power system expansion planning considering the integration and uncertainty of wind

power [20-23]. To overcome the challenges involved in power system expansion, wind energy evaluation and wind farm planning, the stochastic nature of wind speed should be modeled by means of proper probability distribution models to accurately estimate long-term wind power uncertainty and assess potential wind energy [24-26].



**Figure 1.2** Top ten cumulative installed wind power capacities in 2013.

The inherent volatility of wind power has significant impacts on the secure operation of power systems, concerned by system operator [27-29]. Particularly, large-scale integration of wind power can seriously affect the balance and stability of power systems and cause difficulties in scheming proper regulation strategies to coordinating overall energy systems [30, 31].

To well accommodate the fluctuations of wind power, other generation units need to cooperate with wind farms in power systems to ensure the balance and security of the whole system. Therefore, wind power forecasting becomes critical to power system operation and control and certainly produces economic benefits [32, 33]. Wind power is directly related with wind speed that demonstrates chaotic nature due to the extremely complicated weather systems. Like weather prediction, exact wind power forecasting is nearly impossible and the forecasting errors can be significant from time to time [34]. Quantification of the uncertainty

involved in wind power forecasting is indeed needed to meet the requirements of various decision making problems in power systems [35]. The uncertainty estimation becomes more meaningful for the purpose of constructing robust and reliable power systems when considering the large-scale penetration of wind generation.

The operation of different types of units in power systems should be more difficult when accounting for potential power trades in electricity market and corresponding tariff determined by government energy policy [18, 36]. An electricity market is a system for facilitating purchases, sales, and short-term trade of electricity conducted by different participants in power systems, where electricity is a commodity traded through free competitions [37]. Electricity owns the unique properties such as large scale storage is difficult and the supply and demand balance has to be kept in real time. Consequently it is impossible to massively keep it in stock and sale it like other common products in supply chains. In modern power systems, the deregulation of power grids can introduce more investments to electricity network and generation, and improves the efficiency of running power systems. The integration of renewable energy especially for wind power brings new factors to the running framework of electricity market [38, 39]. During the past decade, electricity market in Denmark, as part of the Nordic Market, has been developed to properly accommodate penetrated wind power as well as trading electricity with neighboring countries, such as Norway, Sweden, and Germany [18]. A few researches focus on investigating how to optimally offer its variable power into a competitive electricity market as an independent wind power producer [39, 40]. Electricity price plays a critical role in electricity market operation to guide actions of different participants in power systems [41, 42]. Therefore, electricity price forecasting is important to facilitate individual market participant to optimize their bidding strategies [43, 44]. On the other hand, with the price uncertainty estimation, demand side management in electricity market environment becomes more realistic and efficient [45, 46].

## 1.2 Purpose of the Work

This thesis aims at analyzing, modeling, and predicting the uncertainties involved in modern power systems. The uncertainties involved in wind power generation and electricity price are mainly investigated in this thesis, due to their particular importance in context of smart grids and electricity markets. Specifically, the thesis develops several new methods for probabilistic distribution analysis of wind speed, probabilistic interval forecasts of wind power generation and electricity market price, and probabilistic power flow analysis with wind penetration.

**Wind speed distribution analysis:** Various traditional distribution models have been adopted to estimate the wind speed distribution, the most popular ones of which are Weibull and Rayleigh distributions. However, wind speed usually has high complicated statistics properties, e.g., it demonstrates different probability shapes in different places. It is difficult to have a generalized performance for common probability distribution models. In this thesis, a universal distribution model, known as generalized Lambda distribution, is proposed to accurately fit the distribution of wind speed data measured at different locations. This universal model can be helpful for wind farm planning and reliability analysis of wind penetrated power system in the future.

**Interval forecasts of wind power:** Significantly high level of uncertainties have been introduced to modern power systems, accompanying with more and more wind penetration. These uncertainties have non-ignorable impacts on the balance and stability of power systems. Therefore, accurate quantification of the wind power uncertainty is expected not only by the wind power producer but also by the transmission operator. Since the perfect forecasting of wind generation is impossible, estimation of the forecasting errors becomes critical to decision-makers in power systems. In this thesis, three novel interval forecasts methodologies are developed to generate reliable prediction intervals of wind power generation. These research outcomes could be meaningful to efficient and secure management of power systems, especially in the future smart grid environment.

**Interval forecasts of electricity price:** The future electricity price is highly concerned by the generation companies, consumers, etc., since it is closely related to the interests and profits. Electricity price prediction has important effects on electricity market running, such as day-ahead bidding, and is also be critical in the future smart grid environment to influence the end-consumers' behaviors. Reliable and accurate electricity price forecasting can help to manage the risks faced by individuals in electricity market. Traditional studies on interval forecasting of electricity price are lack of comprehensive evaluation index, which makes the results not very convincing. A hybrid approach is developed to efficiently generate prediction intervals of electricity price, using the advantages of extreme learning machine. The study focuses on day-ahead electricity price forecasting, as the day-head electricity market is more important for bidding of generation companies. In addition, a systematic assessment framework is proposed for the evaluation of prediction intervals of electricity prices

**Probabilistic load flow analysis:** Probabilistic load flow analysis based on point estimation approach is implemented considering the uncertainty of wind power forecasts. The uncertainty of wind power is formulated on basis of the generalized Lambda distribution for wind speed proposed in this thesis. Besides, the effectiveness of point estimate approach is verified through comprehensive studies and comparisons. The computation results of probabilistic load flow can provide meaningful insights to the operational planning and risk assessment of power systems.

### 1.3 Primary Contributions

To achieve the objectives of the research, namely uncertainty analysis, modeling and prognosis for power system operation and planning, the thesis investigates into several different new methods to facilitate wind power penetration into power systems. The scientific contributions achieved in this thesis are summarized as follows:

1. A novel probability distribution model, i.e. the generalized Lambda distribution, is proposed for estimating the probability distribution of wind speed. The proposed model owns the high flexibility of fitting stochastic variable of high complexities in comparison with conventional

distribution models for wind speed such as the Weibull and Rayleigh distributions.

2. Based on the Bootstrap and extreme learning machine, a new method is proposed for the construction of prediction intervals, and applied for wind power forecasting. This approach has the advantages of superior forecasting performance, extremely fast computation speed, and flexible extendability, etc.
3. A direct interval forecasting approach is proposed to directly generate the optimal lower and upper bounds of prediction intervals for wind power through one optimization procedure based on extreme learning machine and particle swarm optimization. A novel objective function is formulated to combine both reliability and interval score for generated prediction intervals. This performance-based objective function can always ensure the sound quality of produced prediction intervals through optimization.
4. A Pareto optimal forecasting model based on extreme learning machine and NSGA-II is proposed to obtain Pareto optimal prediction intervals with respect to the two quality objectives of prediction intervals including both reliability and sharpness. The reliability precision is proposed to select desired optimal forecasters from the Pareto front to form the final forecasting model.
5. A hybrid approach combining extreme learning machine and maximum likelihood estimation is developed for intervals forecasting of electricity price, which has the advantage of high flexibility and efficiency. In addition, a comprehensive evaluation framework is proposed for evaluation of electricity price prediction.
6. A point estimation based probabilistic load flow analysis method is proposed to take into account the uncertainty of wind power generation, where the uncertainty of wind power generation is modeled using the proposed generalized Lambda distribution. The method has proved to have robust performance and fast computation speed to analyze the results of probabilistic load flow.

## 1.4 Structure of the Thesis

The remainder of this thesis is organized as follows,

Chapter 2 proposes a novel probability distribution model named generalized Lambda distribution to estimate wind speed probability. The proposed distribution model has high flexibility and adaptability to fit different shapes of probability density. The state-of-the-art of wind speed distribution modeling is surveyed in the first part of this chapter. A modified starship approach is developed to significantly improve the efficiency of approximating parameters of generalized Lambda distribution. Comprehensive comparisons of results for nine conventional distribution of wind speed probability are carried out to demonstrate the superiority of the proposed model.

Chapter 3 introduces the fundamental principle of probabilistic forecasting (also known as probabilistic interval forecasts) in details, and gives the systematic evaluation criteria. Due to the more and more complicated artificial systems, conventional point forecasting cannot consistently ensure the accuracy. Probabilistic forecasting can provide confidence level of point forecasting and quantify the uncertainty involved in traditional point forecasting. This chapter gives basic theoretical knowledge of the interval forecasts for wind power and electricity price fulfilled in Chapters 5-8.

In Chapter 4, this chapter surveys the state-of-the-art for probabilistic wind power forecasting. The mathematical background and common techniques for point forecasting of wind power are investigated as the prerequisites for probabilistic wind power forecasting. The motivations and benefits of probabilistic forecasting of wind generation are also introduced from different perspectives of power system operation and planning.

In Chapter 5, a novel prediction interval construction approach based on extreme learning machine and bootstrap is developed, which can be proper for the case of nonlinear, nonstationary, chaotic and bounded process such as wind power generation. Mathematical background of extreme learning machine and formulation of prediction interval construction are depicted. The proposed approach has extremely fast speed and flexible expansion capability. It is tested

on wind farm in Australia for multiple stage prediction with hourly and intra-hour resolution, proving the effectiveness of the developed approach.

Chapter 6 presents a novel nonparametric direct interval forecasting approach to construct optimal prediction interval via constructing optimal extreme learning machine based forecaster, applied for wind power forecasting. The optimal forecaster is trained through optimizing the objective function formulated based on the evaluation criteria described in Chapter 3, combining reliability and overall skill. It should be extremely difficult to accurately estimate the probability distribution of wind power prediction error, due to absence of perfect knowledge. The proposed approach successfully avoids the distribution model assumption and can promise the quality as the performance-oriented optimization.

Different from the direct interval forecasting approach proposed in Chapter 6, Chapter 7 proposes a Pareto optimal prediction interval construction method via Pareto optimization algorithm in terms of the two objectives reliability and sharpness that measure/guarantee the quality of constructed prediction intervals. This approach is verified on the wind farm on Bornholm Island of Denmark.

Chapter 8 presents a hybrid approach combining extreme learning machine and maximum likelihood estimation approach for the construction of prediction intervals, which is successfully applied in electricity price forecasts. Like wind power process, electricity price series also demonstrate significant nonstationarity and is nearly impossible to obtain accurate prediction, though the prediction results are pretty important for designing bidding strategies and other actions such as load management in electricity markets. The efficiency and effectiveness of the developed method is proved based on realistic data from Australian electricity market.

In Chapter 9, probabilistic load flow computation is implemented considering the uncertainty of wind power generation. Point estimate method and first-order second-moment approach are used to approximate the mean and variance of load flow variables. Based on the numerical test, point estimate method demonstrates superior performance. Generally, through the uncertainty analysis of the whole system, useful information can be provided for operation and planning of wind power penetrated power systems.

Finally, Chapter 10 summarizes the overall conclusions from the presented researches in this thesis and provides perspectives for further work.

## ***2 Estimation of Wind Speed Probability Distributions Using Generalized Lambda Distribution***

### **2.1 Introduction**

As one of the most important renewable energy nowadays, wind energy is widely utilized for power generation in many countries such as United States, China, Germany, Spain, France, India and Denmark [15]. Theoretically, the energy content in the wind is proportional to the cubic of wind speed, so the probability distribution of wind speed determines the wind power density. In practice, the average wind turbine power  $\bar{P}_w$  can be calculated in terms of the probability density function (PDF) of wind speed  $v$ , according to the following equation

$$\bar{P}_w = \int_0^{\infty} P_w(v) f(v) dv \quad (2.1)$$

where  $f(v)$  is the PDF of wind speed  $v$  and  $P_w(v)$  denotes power curve of the wind turbines indicating power output with respect to wind speed [12]. Therefore, the study of wind speed is essential to wind energy assessment and is the most critical process to wind power project planning which mainly consists of wind farm location selection, wind turbine design to optimize the generation systems and improve the energy conversion efficiency, etc.

Many probability density functions (PDFs) have been proposed to describe the probability distributions of wind speed, which have been reviewed in [24]. These

wind speed distribution models mainly include Weibull [47-53], Rayleigh [52, 53], Beta [24], Burr [54], Gamma [25, 55], Lognormal [56, 57], Inverse Gaussian [58], Gumbel Maximum [59, 60]. So far, the Weibull distribution is the most widely used probability model to estimate the wind speed probability distribution. However, in practice, Weibull distribution cannot universally fit the wind speed data well since the wind speeds in different areas or different seasons can have different characteristics. Several other conventional statistical distributions have the similar limitations. Nonparametric kernel density method [61] and parametric mixture probability functions [62] are proposed to estimate wind speed distribution to overcome the disadvantage of conventional probability models, but with high computational complexity.

This chapter proposes a novel four-parameter continuous probability distribution model, i.e., the generalized Lambda distribution to formulate wind speed probability distribution. The generalized Lambda distribution originates from Tukey's Lambda distribution [63, 64]. With many years' development, the GLD evolves into the traditional RS GLD proposed by Ramberg and Schmeiser [65, 66] and FMKL GLD proposed by Freimer et al. [67].

Practically, the GLD can accommodate a wide range of shapes and is very flexible to fit statistical data. It has been proved that GLD could be used to approximate some well-known distributions, including but are not limited to Normal, Uniform, Student's  $t$ , Exponential, Chi-Square, Gamma, Weibull, Lognormal, Beta, Inverse Gaussian, Logistic, Pareto, Extreme value, and  $F$  distributions [68]. Due to the high flexibility and generality, GLD can be used to statistically model the chaotic wind speed and can provide a more accurate description of its distribution than conventional probability models.

The RS GLD is used to estimate the wind speed distribution. The combined moment and starship method is developed to efficiently and accurately estimate the parameters of GLD based on the historical wind speed data. To demonstrate the performance of GLD, nine conventional PDFs are employed to fit the wind speed data for comparison purposes. Three goodness-of-fit measure approaches including K-S test,  $\chi^2$  test and RMSE are used as the performance judgment criteria. A comprehensive fitting quality evaluation method based on the results

of the applied three goodness-of-fit measures is proposed to rank the overall performance of nine conventional probability models and GLD.

## 2.2 Conventional Probability Models

### 2.2.1 Probability Models

In this chapter, nine conventional probability models including two-parameter Weibull (Weibull-2), three-parameter Weibull (Weibull-3), Rayleigh, Beta, Burr, Gamma, Lognormal, Inverse Gaussian (Inverse G.), and Gumbel Maximum (Gumbel M.) distributions are used to fit the historical wind speed data. The PDFs, cumulative distribution functions (CDFs) and corresponding parameter regions of the nine conventional distributions are given as below.

#### 1) *Two-parameter Weibull Distribution*

The Weibull distribution is one of the most widely applied probability distribution for modeling uncertainty of wind speed [48-50, 52, 53]. The probability density function of a Weibull random variable  $x$  is

$$f(x|\alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha}, \quad x \geq 0 \quad (2.2)$$

where  $\alpha > 0$  is the shape parameter and  $\beta > 0$  is the scale parameter of the distribution. Therefore, the CDF of Weibull distribution can be derived and given as the following equation

$$F(x|\alpha, \beta) = 1 - e^{-\left(\frac{x}{\beta}\right)^\alpha} \quad (2.3)$$

#### 2) *Three-parameter Weibull Distribution*

The three-parameter Weibull distribution is a generalization of Weibull distribution, with one more location parameter. It can estimate the distribution of wind speed more accurately in some places [55, 69]. The PDF of three-parameter Weibull distribution is given as

$$f(x|\alpha, \beta, \gamma) = \frac{\alpha}{\beta} \left(\frac{x-\gamma}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x-\gamma}{\beta}\right)^\alpha} \quad (2.4)$$

where  $\alpha > 0$  is the shape parameter,  $\beta > 0$  is the scale parameter and  $\gamma$  is the location parameter of the distribution, location parameter  $-\infty < \gamma < \infty$ . When  $\gamma = 0$ , this reduces to the 2-parameter distribution. The CDF of Weibull distribution is given as

$$F(x | \alpha, \beta, \gamma) = 1 - e^{-\left(\frac{x-\gamma}{\beta}\right)^\alpha} \quad (2.5)$$

### 3) *Rayleigh Distribution*

The Rayleigh distribution can be taken as a special case of Weibull distribution with  $\beta = 2$ . It is the simplest probability distribution commonly employed for fitting the statistics of wind speed [52, 53]. The PDF of Rayleigh distribution is defined by the following equation

$$f(x | \sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad x \geq 0 \quad (2.6)$$

where parameter  $\sigma > 0$ . The CDF of Rayleigh distribution is given as

$$F(x | \sigma) = 1 - e^{-\frac{x^2}{2\sigma^2}} \quad (2.7)$$

### 4) *Beta Distribution*

The generalized Beta distribution consisting of shape parameters and boundary parameters is also used for wind speed modeling [24]. The PDF of Beta distribution is expressed as

$$f(x | \alpha_1, \alpha_2, a, b) = \frac{1}{B(\alpha_1, \alpha_2)} \frac{(x-a)^{\alpha_1-1} (b-x)^{\alpha_2-1}}{(b-a)^{\alpha_1+\alpha_2-1}}, \quad a \leq x \leq b \quad (2.8)$$

where the shape parameters  $\alpha_1 > 0$  and  $\alpha_2 > 0$ , and the boundary parameters satisfy the inequality  $a < b$ . Accordingly, the CDF of the generalized Beta distribution can be computed as the following equation

$$F(x | \alpha_1, \alpha_2, a, b) = \frac{B_z(\alpha_1, \alpha_2)}{B(\alpha_1, \alpha_2)} \quad (2.9)$$

where  $z = \frac{x-a}{b-a}$ ,  $B(\cdot)$  is the Beta function, and  $B_z$  is the incomplete Beta function.

### 5) *Burr Distribution*

The Burr distribution, also known as Singh-Maddala distribution, is a continuous probability distribution for a non-negative random variable. Recently it has been used to approximate wind speed distribution with good performance [54].

$$f(x|k, \alpha, \beta) = \frac{\alpha k \left(\frac{x}{\beta}\right)^{\alpha-1}}{\beta \left(1 + \left(\frac{x}{\beta}\right)^\alpha\right)^{k+1}}, \quad x \geq 0 \quad (2.10)$$

where shape parameters  $\alpha > 0$ ,  $k > 0$ , scale parameter  $\beta > 0$ .

$$F(x|k, \alpha, \beta) = 1 - \left(1 + \left(\frac{x}{\beta}\right)^\alpha\right)^{-k} \quad (2.11)$$

### 6) *Gamma Distribution*

The probability density function of the gamma distribution can be expressed in terms of the gamma function parameterized with respect to a shape parameter  $\alpha$ , scale parameter  $\beta$ , and the location parameter  $\gamma$ . The Gamma distribution has been applied for wind speed estimation [25]. The probability density function of a gamma distributed random variable  $x$  is defined through the following equation

$$f(x|\alpha, \beta, \gamma) = \frac{1}{\beta^\alpha} \frac{1}{\Gamma(\alpha)} (x-\gamma)^{\alpha-1} e^{-\frac{(x-\gamma)}{\beta}}, \quad x \geq \gamma \quad (2.12)$$

where both  $\alpha$  and  $\beta$  will be positive values, location parameter  $-\infty < \gamma < \infty$ , and  $\Gamma(\cdot)$  denotes Gamma function. The CDF of the Gamma distribution can be calculated and given as the following equation

$$F(x|\alpha, \beta, \gamma) = \frac{1}{\Gamma(\alpha)} \Upsilon\left(\alpha, \frac{x-\gamma}{\beta}\right) \quad (2.13)$$

where  $\Upsilon(\cdot)$  is the lower incomplete gamma function which can be expressed as

$$\Upsilon(s, x) = \int_0^x t^{s-1} e^{-t} dt \quad (2.14)$$

### 7) *Lognormal Distribution*

The Lognormal distribution describes a probability distribution of a random variable whose logarithm is normally distributed, which follows the variable

transformation rule on the density function of a normal distribution. It is extensively employed to model wind speed distribution [56, 57]. The Lognormal distribution derived from a normal distribution with mean  $\mu$  and standard deviation  $\sigma$  is represented as the following equation

$$f(x | \mu, \sigma, \gamma) = \frac{1}{(x - \gamma)\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x-\gamma)-\mu)^2}{2\sigma^2}}, \quad x \geq \gamma \quad (2.15)$$

where the mean value  $-\infty < \mu < \infty$ , the standard deviation  $\sigma > 0$ , and location parameter  $-\infty < \gamma < \infty$ . The cumulative probability function of the Lognormal distribution can be obtained and expressed as

$$F(x | \mu, \sigma, \gamma) = \frac{1}{2} \operatorname{erfc}\left(-\frac{\ln x - \mu}{\sigma\sqrt{2}}\right) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right) \quad (2.16)$$

where  $\Phi(\cdot)$  is the standard normal CDF,  $\operatorname{erfc}(\cdot)$  is the complementary Gauss error function and is defined as

$$\operatorname{erfc}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt \quad (2.17)$$

### 8) *Inverse Gaussian Distribution*

In probability theory, the inverse Gaussian distribution, also named as Wald distribution, is a three-parameter family of continuous probability distributions with support on region  $(0, \infty)$ . It has been proposed to be an alternative to Weibull distribution for modeling of wind speed uncertainty [58]. The probability density function of the inverse Gaussian distribution is given by the following equation

$$f(x | \mu, \lambda, \gamma) = \left(\frac{\lambda}{2\pi(x-\gamma)^3}\right)^{1/2} e^{-\frac{\lambda(x-\gamma-\mu)^2}{2\mu^2(x-\gamma)}}, \quad x > \gamma \quad (2.18)$$

for  $x > 0$ , where  $\mu > 0$  is the mean and  $\lambda > 0$  is the shape parameter, location parameter  $-\infty < \gamma < \infty$ . Then the cumulative probability function of the inverse Gaussian distribution can be expressed as the following equation

$$F(x | \mu, \lambda, \gamma) = \Phi\left(\sqrt{\frac{\lambda}{x-\gamma}}\left(\frac{x-\gamma}{\mu} - 1\right)\right) + e^{\frac{2\lambda}{\mu}} \Phi\left(-\sqrt{\frac{\lambda}{x-\gamma}}\left(\frac{x-\gamma}{\mu} + 1\right)\right) \quad (2.19)$$

where  $\Phi(\cdot)$  is the CDF of the standard normal (standard Gaussian) distribution.

### 9) *Gumbel Maximum Distribution*

In probability theory, the Gumbel maximum distribution is actually a particular case of the generalized extreme value distribution, which has been introduced to model the wind speed frequency [59, 60]. The probability density function of the Gumbel maximum distribution is defined by the following equation

$$f(x|\sigma, \mu) = \frac{1}{\sigma} e^{-\frac{x-\mu}{\sigma}} e^{-\frac{x-\mu}{\sigma}} \quad (2.20)$$

where scale parameter  $\sigma > 0$ , and location parameter  $-\infty < \mu < \infty$ . Then the cumulative probability function of the Gumbel maximum distribution is given as

$$F(x|\sigma, \mu) = e^{-e^{-\frac{x-\mu}{\sigma}}} \quad (2.21)$$

#### 2.2.2 Parameter Estimation of Conventional Probability Models

The most commonly used numerical methods for estimating the parameters of the conventional distributions described in Section 2.2.1 are the method of moments, the maximum likelihood estimation (MLE) and the least square method. Particularly, since Weibull is the most popular statistical distribution for wind speed distribution modeling, the parameters estimation methods of Weibull distribution have been investigated comprehensively. The MLE method is proved to be the most efficient parameters estimating method and is recommended in [49, 50]. Actually, the majority of the related literatures on the wind speed probability distribution analysis use the MLE method for deriving the values of PDF parameters. Therefore, in this chapter the MLE method is employed to estimate the parameters of the nine conventional distributions.

## 2.3 Generalized Lambda Distribution

### 2.3.1 Model Definition

The generalized Lambda distribution is a continuous probability distribution with four parameters  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ ,  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ , which is defined in terms of the inverse of its distribution function, and is different from conventional distributions. In this study, for the relatively simple distribution function, we use

the Ramberg- Schmeiser generalized Lambda distribution (RS GLD) proposed by Ramberg and Schmeiser [65, 66]. The RS  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  could be described as

$$Q(u) = \lambda_1 + \frac{u^{\lambda_3} - (1-u)^{\lambda_4}}{\lambda_2} \quad (2.22)$$

where  $0 \leq u \leq 1$ ,  $Q(u)$  denotes the value of  $x$  satisfying cumulative frequency  $F_X(x) = u$ ,  $\lambda_1$  is the location parameter,  $\lambda_2$  is the scale parameter, while  $\lambda_3$  and  $\lambda_4$  are the shape parameters jointly determining the skewness and kurtosis of  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ .

Given the inverse distribution function (2.22), the PDF of  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  can be derived based on the relationship between PDF and cumulative distribution function (CDF). From (2.22), we have  $x = Q(u)$  and  $u = F_X(x)$ . By differentiating  $u$  with respect to  $x$ , the PDF can be obtained

$$f(x) = \frac{du}{dx} = \frac{du}{d(Q(u))} = \frac{1}{\frac{d(Q(u))}{du}} \quad (2.23)$$

Since  $Q(u)$  is given, from (2.22), the following is obtained

$$\frac{d(Q(u))}{du} = \frac{d}{du} \left( \lambda_1 + \frac{u^{\lambda_3} - (1-u)^{\lambda_4}}{\lambda_2} \right) = \frac{\lambda_3 u^{\lambda_3-1} + \lambda_4 (1-u)^{\lambda_4-1}}{\lambda_2} \quad (2.24)$$

Substituting (2.24) to (2.23), the PDF of  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  can be expressed as

$$f(x) = \frac{\lambda_2}{\lambda_3 u^{\lambda_3-1} + \lambda_4 (1-u)^{\lambda_4-1}}, \quad x = Q(u) \quad (2.25)$$

### 2.3.2 Parameter space and support region

For any formula that specifies a probability distribution, the parameters of the model have corresponding domain to ascertain that the formula holds. In particular, a specific function  $f(x)$  can be used as a PDF if and only if it satisfies the conditions defined by

$$f(x) \geq 0 \quad \text{and} \quad \int_{-\infty}^{\infty} f(x) dx = 1 \quad (2.26)$$

As such, the four parameter  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$ , and  $\lambda_4$  of GLD also have respective spaces to ensure the distribution function holds. Theoretically, the formula of  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  cannot always specify a valid probability distribution. Based on (2.25), the  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  is validated when the conditions in (2.26) are satisfied simultaneously.

$$\frac{\lambda_2}{\lambda_3 u^{\lambda_3-1} + \lambda_4 (1-u)^{\lambda_4-1}} \geq 0 \text{ and } \int_{-\infty}^{\infty} f(Q(u))dQ(u) = 1 \quad (2.27)$$

From (2.23), we know that

$$f(Q(u))dQ(u) = du \quad (2.28)$$

where  $0 \leq u \leq 1$ . Accordingly, for any  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$ , and  $\lambda_4$  the PDF  $f(x)$  of GLD will integrate to 1. Thus the second condition in (2.27) always holds.

The space of parameters is determined by the satisfaction of the first condition given in (2.27). According to Karian and Dudewicz [68], the parameter space and the support region of  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  are as explicitly listed in Table 2.1.

**Table 2.1** The parameter space and support region of GLD [68].

$\lambda_3$	$\lambda_4$	Support region
$\lambda_3 > 0$	$\lambda_4 > 0$	$[\lambda_1-1/\lambda_2, \lambda_1+1/\lambda_2]$
$\lambda_3 > 0$	$\lambda_4 = 0$	$[\lambda_1, \lambda_1+1/\lambda_2]$
$\lambda_3 = 0$	$\lambda_4 > 0$	$[\lambda_1-1/\lambda_2, \lambda_1]$
$\lambda_3 < 0$	$\lambda_4 < 0$	$[-\infty, \infty]$
$\lambda_3 < 0$	$\lambda_4 = 0$	$[-\infty, \lambda_1+1/\lambda_2]$
$\lambda_3 = 0$	$\lambda_4 < 0$	$[\lambda_1-1/\lambda_2, \infty]$

### 2.3.3 Parameter estimation

The probability function of GLD is a particularly implicit function. Several approaches can be used to estimate the parameters of GLD, including the moment based method and the starship method [70]. In this study, the combined moment and starship method is newly proposed to fit GLD from the historical wind speed samples. The proposed CMS approach combines the advantages of the moment based method and the starship method to guarantee the accuracy and efficiency simultaneously.

### 1) The Moment-matching Method

The moment-matching method is a time-saving method used for GLD estimation for a long time [66, 68, 71]. However, the moment-matching method usually causes considerable errors for fitting GLD. In the study, the moment based method is applied in the first step of the proposed CMS approach. Given the parametric GLD distribution with inverse distribution function  $Q(u)$ , the moment-matching method can be particularly described as follows:

1. Obtain the function of the first four moments, including the mean  $\alpha_1$ , variance  $\alpha_2$ , skewness  $\alpha_3$ , and kurtosis  $\alpha_4$ , for the theoretical function  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ ;
2. Calculate the mean  $\hat{\alpha}_1$ , variance  $\hat{\alpha}_2$ , skewness  $\hat{\alpha}_3$ , and kurtosis  $\hat{\alpha}_4$  of the observed samples;
3. Derive the four parameters  $\lambda_1, \lambda_2, \lambda_3$ , and  $\lambda_4$  of the GLD such that the first four moments of the theoretical GLD match the corresponding moments of the observed samples.

The first four moments of the theoretical  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  are specified in (2.29)-(2.36).

$$\alpha_1 = \mu = E(X) = \lambda_1 + \frac{A}{\lambda_2} \quad (2.29)$$

$$\alpha_2 = \sigma^2 = E[(X - \mu)^2] = \frac{B - A^2}{\lambda_2^2} \quad (2.30)$$

$$\alpha_3 = \frac{E[(X - \mu)^3]}{\sigma^3} = \frac{C - 3AB + 2A^3}{\lambda_2^3 \sigma^3} \quad (2.31)$$

$$\alpha_4 = \frac{E[(X - \mu)^4]}{\sigma^4} = \frac{D - 4AC + 6A^2B - 3A^4}{\lambda_2^4 \sigma^4} \quad (2.32)$$

where

$$A = \frac{1}{1 + \lambda_3} - \frac{1}{1 + \lambda_4} \quad (2.33)$$

$$B = \frac{1}{1 + 2\lambda_3} + \frac{1}{1 + 2\lambda_4} - 2\beta(1 + \lambda_3, 1 + \lambda_4) \quad (2.34)$$

$$C = \frac{1}{1+3\lambda_3} - \frac{1}{1+3\lambda_4} - 3\beta(1+2\lambda_3, 1+\lambda_4) + 3\beta(1+\lambda_3, 1+2\lambda_4) \quad (2.35)$$

$$D = \frac{1}{1+4\lambda_3} - \frac{1}{1+4\lambda_4} - 4\beta(1+3\lambda_3, 1+\lambda_4) + 6\beta(1+2\lambda_3, 1+2\lambda_4) - 4\beta(1+\lambda_3, 1+3\lambda_4) \quad (2.36)$$

It can be observed that  $A, B, C, D$  in (2.33)-(2.36) are free of the parameters  $\lambda_1$  and  $\lambda_2$ . According to (2.30), we can obtain

$$\lambda_2 = \left( \frac{B - A^2}{\sigma^2} \right)^{1/2} \quad (2.37)$$

Substituting (2.37) to (2.31) and (2.32),  $\lambda_1$  and  $\lambda_2$  can be eliminated from the functions of  $\alpha_3$  and  $\alpha_4$  respectively. Therefore, both  $\alpha_3$  and  $\alpha_4$  are just determined by  $\lambda_3$  and  $\lambda_4$ . The first four moments of the observed random data are given in Table 2.2. Then  $\lambda_3$  and  $\lambda_4$  can be obtained through the following equation:

$$\alpha_3 = \hat{\alpha}_3 \quad \text{and} \quad \alpha_4 = \hat{\alpha}_4 \quad (2.38)$$

The appropriate initial values for solving (2.31) and (2.32) have been given in [68]. Then,  $\lambda_2$  and  $\lambda_1$  can be easily derived by solving (2.29) and (2.30) accordingly.

**Table 2.2** The first four moments of observed samples.

Mean	$\hat{\alpha}_1 = \hat{\mu} = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^N x_i$
Variance	$\hat{\alpha}_2 = \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{\mu})^2$
Skewness	$\hat{\alpha}_3 = \frac{1}{N\hat{\sigma}^3} \sum_{i=1}^N (x_i - \hat{\mu})^3$
Kurtosis	$\hat{\alpha}_4 = \frac{1}{N\hat{\sigma}^4} \sum_{i=1}^N (x_i - \hat{\mu})^4$

The moment-matching method described above is extremely time-saving, but the accuracy can be compromised slightly. For the empirical samples with large variability, it will be difficult to derive accurate GLD by means of the moment-matching method. The approximation errors of the four parameters  $\lambda_1, \lambda_2, \lambda_3$  and  $\lambda_4$  may have important impacts on the accuracy of the distribution fitting. The

obtained GLD may not match the empirical distribution effectively, though the first four moments are matched exactly.

### **2) The Starship Method**

The “starship” method for fitting the parameters of GLD is proposed by King and MacGillivray [72]. It is a computer intensive estimation method and based on that the GLD is a transformation of the uniform (0, 1) distribution. Generally, the starship method is implemented via three steps:

(1). For a set of random data and a range of parameters  $\lambda_1, \lambda_2, \lambda_3$  and  $\lambda_4$  values, use the reverse transformation function to transfer the random data value  $x$  to the cumulative frequency  $F_X(x)$ ;

(2). Compute the value of a proper goodness-of-fit measure that assesses the closeness of the obtained values in Step (1) to the uniform (0, 1) distribution;

(3). Obtain the optimal values  $\lambda_1, \lambda_2, \lambda_3$  and  $\lambda_4$  that minimize the chosen goodness-of-fit measure to the uniform fitting performance, as the fitted parameters.

In Step (1), the range of  $\lambda_1, \lambda_2, \lambda_3$  and  $\lambda_4$  values consists of a  $20 \times 20 \times 20 \times 20$  grid according to the parameter space given in Table 2.2. The starship method can be used to approximate the parameters covering the full parameters space and is flexible according to the fitting requirement. However, in practical terms, large computational efforts, always a number of hours time, are needed for the starship method. Due to the grid resolution, it is difficult to obtain the really optimal results.

### **3) The Combined Moment and Starship Method**

To overcome the drawbacks of the moment-matching and the starship methods described above, the combined moment and starship method is newly developed in this chapter to ensure both the efficiency and accuracy simultaneously. The specific procedures of the CMS method are described as follows:

(a). Given the observed samples, obtain the parameters  $\lambda_1, \lambda_2, \lambda_3$  and  $\lambda_4$  of GLD based on the moment-matching method described in Section 2.3.1;

(b). Generate a set of initial values for  $\lambda_1, \lambda_2, \lambda_3$  and  $\lambda_4$  that consists of an appropriate four-dimensional space surrounding the values of the four parameters derived from Step (a);

(c). Use K-S test as the goodness-of-fit evaluation criterion, and compute the goodness-of-fit measure for assessing fitting quality of GLD with the generated values  $\lambda_1, \lambda_2, \lambda_3$  and  $\lambda_4$  to the empirical distribution of the observed samples;

(d). Obtain the optimal four parameters of GLD to minimize the goodness-of-fit measure through Nelder-Mead Simplex algorithm.

The optimal results obtained in the Step (d) are the parameters of GLD fitted by the proposed CMS method for the observed samples. The range of parameters can be greatly reduced according the initial results derived from the moment based method. The computation process can be greatly facilitated by applying the Nelder-Mead Simplex algorithm. In general, the CMS method significantly improves the efficiency with time consumption in minutes and the accuracy through the goodness-of-fit measure.

## 2.4 Fitting Performance Test and Evaluation

In the study, the performance judgment criterion is the goodness-of-fit of the fitted distribution to the observed wind speed data. The quality of the distribution fitting can be examined via three goodness-of-fit test methods described in this section. Then the fitting performance can be identified according to the values of the goodness-of-fit measures.

### 2.4.1 Kolmogorov–Smirnov Test

In statistics, the K-S test is a nonparametric test popularly used as a goodness of fit measure free of bins. The K-S test quantifies the maximum vertical distance between the empirical cumulative distribution function (ECDF) calculated from the given samples and the CDF of the fitted distribution. The ECDF  $F_n$  for  $n$  observed samples  $x_1, x_2, \dots, x_n$  is defined by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I_{x_i \leq x} \quad (2.39)$$

where  $I_{X_i \leq x}$  is taken as the indicator function that equals to 1 if  $X_i \leq x$ , and equals to 0 otherwise, used to compute the number of samples less than  $x$ . The K-S test for a given CDF  $F(x)$  is

$$KS = \sup |F_n(x) - F(x)| \quad (2.40)$$

where  $\sup$  represents the maximum value of the set of distances.

### 2.4.2 Chi-Square Test

The  $\chi^2$  test is widely employed in statistical test for wind speed distribution estimation, which is based on the difference between the expected and observed frequencies with bins. The observed samples are divided into  $n$  groups. Chi-Square error is defined by

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (2.41)$$

where  $O_i$  is the observed frequencies of the  $i^{th}$  group, and  $E_i$  is the frequencies of the expected distribution of the  $i^{th}$  group. Given the CDF of the expected distribution  $F(x)$ ,  $E_i$  can be obtained through

$$E_i = F(U_i) - F(L_i) \quad (2.42)$$

where  $U_i$  and  $L_i$  denote the maximum and minimum value of the samples in group  $i$ , separately.

### 2.4.3 Root Mean Square Error

RMSE is also used as statistical test based on the difference between the observed and expected frequencies, similar to  $\chi^2$  test. The numerical values of RMSE can be used to compare the difference between the histogram of observed random data and selected theoretical distributions. When the observed samples are divided into  $N$  bins, RMSE can be expressed as

$$RMSE = \left( \frac{1}{N} \sum_{i=1}^N (O_i - E_i)^2 \right)^{1/2} \quad (2.43)$$

where  $O_i$  is the observed probability frequencies in interval of the  $i^{th}$  bin, and  $E_i$  is the frequencies of the expected distribution within the  $i^{th}$  bin. A smaller RMSE

indicates a higher fit quality between the observed random samples and the expected distribution.

#### 2.4.4 Comprehensive Fitting Quality Evaluation

Generally, the smaller resulted goodness-of-fit measures for all the three test methods mean the higher performance of the distribution fitting. K-S test is independent of bin width. By contrast, RMSE and  $\chi^2$  test both rely on the width of the group division and they also need to make the sample size large enough. K-S value is a more straightforward approach compared with the other two.

Each goodness-of-fit statistics described above has its own merits and demerits. To rank the performance of the conventional probability models and GLD more systematically, a comprehensive evaluation method is developed here. This evaluation method is based on the normalized statistical test values of K-S test,  $\chi^2$  test and RMSE. For a specific statistical test, the normalized statistical test value  $S'_i$  is defined by

$$S'_i = \frac{S_i^{test} - S_{\min}^{test}}{S_{\max}^{test} - S_{\min}^{test}} \quad (2.44)$$

where  $S_i^{test}$  is the test value of the  $i^{th}$  probability model,  $S_{\max}^{test}$  and  $S_{\min}^{test}$  are the maximum and minimum test value of all the comparison probability models under the given test approach. The normalized individual test value ranges from 0 to 1. Obviously, the smaller value indicates the higher accuracy and performance. Therefore, given the test values, the comprehensive evaluation index  $S_i$  of the  $i^{th}$  probability model in a specific station is defined by

$$S_i = \frac{S_i^{KS} - S_{\min}^{KS}}{S_{\max}^{KS} - S_{\min}^{KS}} + \frac{S_i^{\chi^2} - S_{\min}^{\chi^2}}{S_{\max}^{\chi^2} - S_{\min}^{\chi^2}} + \frac{S_i^{RMSE} - S_{\min}^{RMSE}}{S_{\max}^{RMSE} - S_{\min}^{RMSE}} \quad (2.45)$$

where  $S_i$  is the summation of the three normalized statistical test values. With the comprehensive evaluation index  $S_i$ , the performance of the nine conventional distributions and GLD can be compared and ranked reasonably.

**Table 2.3** Specific information of the fifteen stations

Station	Geographical information			Years	Wind speed			
	Latitude	Longitude	Elevation (m)		Mean (m/s)	Variance ( $m^2/s^2$ )	Skewness	Kurtosis
Settlement Point (STP)	26°42'16" N	78°59'40" W	1.5	2005-2011	5.7189	7.9452	0.6528	3.3322
Passage Island (PSI)	48°13'24" N	88°22'0" W	195.4	2006-2011	6.7627	12.719	0.7665	3.4757
Delaware Bay (DLB)	38°27'49" N	74°42'7" W	0	2004-2011	6.3622	11.529	0.6229	3.0633
Newport (NPT)	44°36'48" N	124°40" W	9.1	2006-2011	5.008	12.214	1.2128	4.7674
South Bass Island (SBI)	41°37'42" N	82°50'30" W	176.8	2006-2011	5.9446	12.583	0.9261	3.8103
Dunkirk (DKK)	42°29'36" N	79°21'12" W	182.9	2006-2011	5.2249	12.955	1.3077	4.9839
Folly Island (FLI)	32°41'6" N	79°53'18" W	3	2005-2011	4.78	6.1912	0.6781	3.3016
Keaton Beach (KTB)	29°49'2" N	83°35'30" W	1.8	2005-2011	3.9457	4.8795	1.0292	4.5193
Flat Island Light (FIL)	59°19'54" N	151°59'42" W	17.9	2006-2011	5.726	14.808	0.9816	3.7207
Five Finger (FFG)	57°16'18" N	133°37'48" W	6.7	2006-2011	5.9021	16.993	0.8276	3.1561
Grand Isle (GRI)	29°16'0" N	89°57'24" W	1.8	1999-2005	4.6786	6.0244	1.0344	4.6273
Fowey Rocks (FWR)	25°35'25" N	80°5'48" W	0	2005-2011	6.7914	9.6741	0.4365	3.3986
Smith Island (SMI)	48°19'6" N	122°50'36" W	15.2	2005-2011	5.4813	15.453	1.1585	4.3198
Lake Ontario (LKO)	43°37'9" N	77°24'18" W	74.7	2002-2011	5.3301	9.4764	0.8143	3.9157
Christmas Island (CRI)	0°0'1" N	153°54'46" W	0	2000-2001 2004-2008	6.3254	2.7995	-0.2128	3.0435

## 2.5 Numerical results and analysis

### 2.5.1 Wind speed data

According to the International Electrotechnical Commission (IEC) [73], the design of wind turbines and planning of a wind power plant in a specific station require 10-minute wind speed for various assessments such as the turbulence intensity assessment. Fortunately, the National Data Buoy Center (NDBC) of National Oceanic and Atmospheric Administration (NOAA) provides historical 10-minute wind speed data for more than one decade [74]. In the study, about six year to seven-year long historical wind speed data from 15 stations are used for the wind speed probability distribution fitting. The geographical information, and observed wind speed information including time span, mean, variance, skewness, and kurtosis are given in Table 2.3.

### 2.5.2 Comparison of Fitting Performance

The three statistical test methods applied in the study are introduced in Section 2.4. Based on the estimation results, statistical tests of all the applied probability distributions are carried out and the results are given in Tables 2.4-2.6, respectively. For all the three statistical tests, it is obvious that the proposed GLD has much higher performance for describing the wind speed data than conventional distributions at all the 15 stations.

Since the K-S test is free of the bin width and group division, K-S test value can be more reliable compared with the other two approaches. Table 4 shows that the K-S test value of the proposed GLD is much smaller than the other nine distributions. With the wind speed samples at all fifteen stations, the K-S test values of GLD estimation are totally less than 0.01 with ten of which are less than 0.006. This result indicates that GLD performs much more effectively than the conventional distributions. Especially, the performance differences are very significant at stations NPT, DKK, FFG and SMI. In these stations the test values of nine conventional distributions are all larger than 0.2.

Practically, the three goodness-of-fit statistics have their own emphasis. From Tables 2.4-2.6, the ranking of the accuracy of the studied distributions measured by the three statistical test methods may be different from each other. Based on the goodness-of-fit test results, the comprehensive statistical test values and ranking orders of the conventional probability models and GLD are obtained and listed in Table 2.7. The ranking orders are given in square brackets.

**Table 2.4** K-S test results of different probability models for historical wind speed data measured at the fifteen stations.

Model	Station														
	STP	PSI	DIB	NPT	SBI	DKK	FLI	KTB	FLU	FFG	GRI	FWR	SMI	LKO	CRI
Weibull-2	0.02054	0.0241	0.007075	0.02461	0.01919	0.04654	0.02223	0.03049	0.02313	0.02553	0.04286	0.01068	0.02328	0.03865	0.01007
Weibull-3	0.02208	0.02602	0.01669	0.03343	0.02296	0.04875	0.02563	0.03179	0.02846	0.03027	0.04979	0.008081	0.02641	0.0124	0.01637
Rayleigh	0.0274	0.02508	0.01376	0.09909	0.05325	0.112	0.01827	0.0404	0.08587	0.08769	0.04251	0.05205	0.09621	0.02156	0.2276
Beta	0.01598	0.01437	0.01219	0.02931	0.01571	0.03721	0.01726	0.01291	0.02445	0.02555	0.0223	0.01071	0.02473	0.009249	0.00831
Burr	0.0194	0.01808	0.006957	0.02432	0.01698	0.03493	0.02232	0.01443	0.02298	0.02536	0.02069	0.0107	0.02298	0.03869	0.009401
Gamma	0.02054	0.01308	0.01205	0.0283	0.01432	0.03735	0.02476	0.01286	0.01897	0.0299	0.02228	0.0145	0.02273	0.008428	0.01893
Lognormal	0.02539	0.01592	0.01761	0.02461	0.02103	0.02899	0.03067	0.01588	0.02626	0.04078	0.01591	0.01571	0.02421	0.01273	0.01632
Inverse G.	0.07964	0.08945	0.1095	0.08715	0.1026	0.06635	0.08223	0.08384	0.1012	0.1095	0.07339	0.1298	0.08972	0.105	0.08807
Gumbel M.	0.02076	0.01675	0.01561	0.0589	0.0311	0.06578	0.02954	0.02006	0.05036	0.05962	0.01886	0.03396	0.05916	0.0113	0.08256
GLD	0.005313	0.003467	0.002703	0.008108	0.003783	0.009369	0.005576	0.006886	0.006615	0.005303	0.004795	0.005072	0.006361	0.003263	0.0025

**Table 2.5**  $\chi^2$  test results of different probability models for historical wind speed data measured at the fifteen stations.

Model	Station														
	STP	PSI	DLB	NPT	SBI	DKK	FLI	KTB	FL	FFG	GRI	FWR	SMI	LKO	CRI
Weibull-2	0.005887	0.007754	0.001721	0.01486	0.007641	0.0355	0.006973	0.007282	0.01001	0.01085	0.02393	0.008992	0.01019	0.01645	0.004637
Weibull-3	0.006491	0.008574	0.00248	0.01625	0.008698	0.03672	0.008425	0.007697	0.01155	0.012	0.03057	0.004712	0.01155	0.0030015	0.002537
Rayleigh	0.01128	0.008461	0.002135	0.07919	0.02489	0.1138	0.007842	0.00933	0.06367	0.06846	0.02521	0.03672	0.09276	0.003553	0.5093
Beta	0.003121	0.002557	0.001094	0.008638	0.004106	0.01675	0.004692	0.0012	0.008144	0.005588	0.004298	0.006683	0.007038	0.002072	0.0026
Burr	0.004672	0.004748	0.00171	0.01459	0.006029	0.02002	0.006632	0.000934	0.009992	0.01084	0.005076	0.009175	0.01016	0.01646	0.004797
Gamma	0.005983	0.002402	0.003812	0.008299	0.003925	0.01681	0.009112	0.001199	0.008808	0.02013	0.004291	0.008537	0.00563	0.001783	0.01964
Lognormal	0.009078	0.00409	0.006614	0.01192	0.006727	0.0125	0.01451	0.002272	0.01748	0.04006	0.002258	0.009192	0.01115	0.002742	0.01737
Inverse G.	0.1599	0.1804	0.185	0.08459	0.1254	0.05568	0.09852	0.06411	0.1052	0.1509	0.07679	0.3543	0.08454	0.18	0.4679
Gumbel M.	0.01348	0.005302	0.009466	0.04097	0.01204	0.05797	0.01753	0.003002	0.04049	0.05062	0.002544	0.04112	0.04722	0.005365	0.2134
GJD	0.001839	0.001175	0.000345	0.005869	0.001787	0.01024	0.002783	0.000118	0.00532	0.002427	0.001731	0.004007	0.002755	0.000824	0.0015

**Table 2.6** RMSE test results of different probability models for historical wind speed data measured at the fifteen stations.

Model	Station														
	STP	PSI	DLB	NPT	SBI	DKK	FLI	KTB	FIL	FFG	GRI	FWR	SMI	LKO	CRJ
Weibull-2	0.004942	0.004906	0.002346	0.001632	0.005064	0.01079	0.005757	0.008322	0.005642	0.005938	0.01032	0.003991	0.001018	0.007463	0.002698
Weibull-3	0.005171	0.005128	0.003159	0.001786	0.005346	0.01094	0.006358	0.008605	0.006147	0.006371	0.01064	0.00358	0.001126	0.003433	0.004281
Rayleigh	0.005428	0.005082	0.002904	0.006932	0.008463	0.01731	0.005714	0.009998	0.01359	0.01382	0.01053	0.008645	0.006344	0.003764	0.05334
Beta	0.003586	0.002831	0.002009	0.000956	0.003742	0.007795	0.004949	0.002806	0.005157	0.004637	0.004514	0.004088	0.000726	0.002202	0.002298
Burr	0.004587	0.003969	0.002327	0.001602	0.004558	0.008332	0.005709	0.00349	0.005631	0.00593	0.005127	0.003999	0.001016	0.007466	0.002505
Gamma	0.004804	0.002761	0.003133	0.000889	0.003339	0.007796	0.006867	0.002801	0.004683	0.00749	0.004511	0.004727	0.0005928	0.002087	0.007876
Lognormal	0.006056	0.00375	0.004481	0.000899	0.004444	0.005978	0.008626	0.003317	0.006247	0.0107	0.00318	0.004962	0.0009356	0.003008	0.005652
Inverse-G.	0.01615	0.01551	0.02093	0.008027	0.01835	0.01295	0.01802	0.02439	0.01752	0.02029	0.01552	0.02155	0.007243	0.02302	0.0299
Gumbel M.	0.006391	0.004137	0.004809	0.003808	0.006318	0.01278	0.009216	0.004939	0.01084	0.01265	0.003556	0.007925	0.003834	0.003652	0.02662
GJD	0.002372	0.001786	0.00117	0.00044	0.002208	0.004823	0.003566	0.002564	0.003661	0.002626	0.002355	0.003384	0.000209	0.001697	0.001965

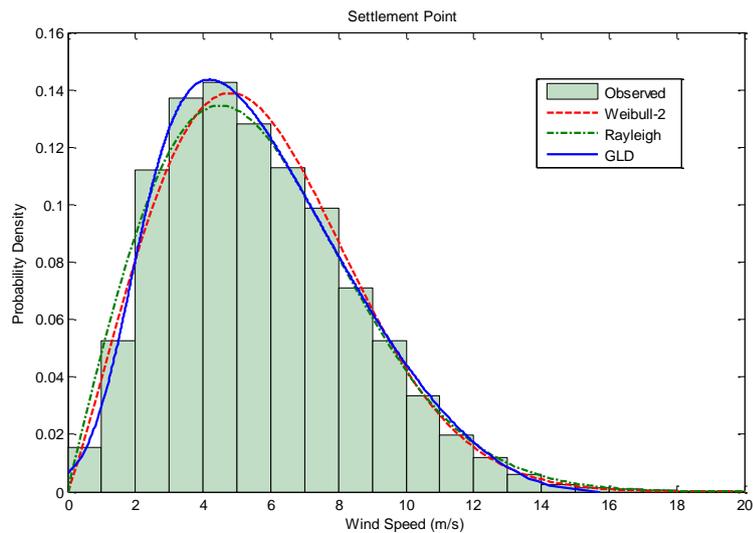
From Table 2.7, it can be easily seen that GLD always gives the most accurate results for all the fifteen stations. Traditionally, Weibull distribution is the most popular probability model for describing the wind speed. However, only the three-parameter Weibull distribution (Weibull-3) can outperform other conventional probability models at the station FWR. At other stations, Weibull even demonstrates relatively poor accuracy than other conventional distributions. The performance of Weibull-3 and Weibull-2 distributions for fitting the random wind speeds also can be compared through Table 2.7. Though Weibull-2 distribution has one less location parameter than Weibull-3 distribution, except stations FWR and LKO, Weibull-2 outperforms Weibull-3 at other stations. Rayleigh distribution is also popularly used to describe the distribution of wind speed. However, comparing with other conventional probability models, Rayleigh only ranks the second at FLI and performs much worse elsewhere.

**Table 2.7** Comprehensive goodness-of-fit results of different probability models for historical wind speed data measured at the fifteen stations

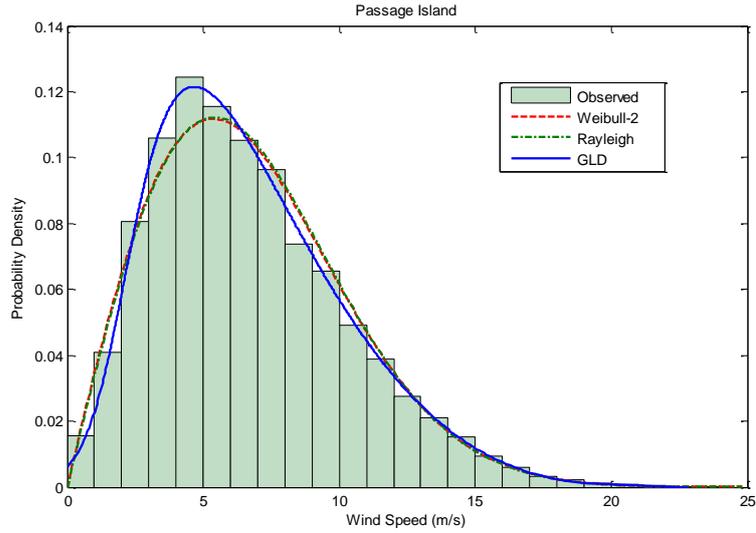
Station	Probability Model										
	Weibull-2	Weibull-3	Rayleigh	Beta	Burr	Gamma	Lognormal	Inverse G.	Gumbel M.	GLD	
STP	0.417 [5]	0.4582 [6]	0.5787 [8]	0.2396 [2]	0.3683 [3]	0.4075 [4]	0.5833 [9]	3.0 [10]	0.5731 [7]	0.0 [1]	
PSI	0.5041 [7]	0.5471 [9]	0.5322 [8]	0.2106 [3]	0.349 [6]	0.1897 [2]	0.3042 [4]	3.0 [10]	0.3489 [5]	0.0 [1]	
DLB	0.1079 [3]	0.2432 [7]	0.201 [5]	0.1354 [4]	0.1058 [2]	0.2056 [6]	0.3411 [8]	3.0 [10]	0.3545 [9]	0.0 [1]	
NPT	0.5787 [6]	0.7205 [7]	2.839 [9]	0.4132 [4]	0.5665 [5]	0.3816 [2]	0.3896 [3]	2.8687 [10]	1.5978 [8]	0.0 [1]	
SBI	0.3802 [6]	0.4444 [7]	1.075 [9]	0.2345 [3]	0.3134 [4]	0.194 [2]	0.353 [5]	3.0 [10]	0.6143 [8]	0.0 [1]	
DKK	1.0841 [6]	1.1291 [7]	3.0 [10]	0.572 [3]	0.6244 [5]	0.5742 [4]	0.3055 [2]	1.6446 [8]	1.6481 [9]	0.0 [1]	
FLI	0.4126 [5]	0.5138 [6]	0.3671 [3]	0.268 [2]	0.4069 [4]	0.5448 [7]	0.8 [8]	3.0 [10]	0.8576 [9]	0.0 [1]	
KTB	0.6826 [7]	0.7188 [8]	0.9201 [9]	0.1063 [3]	0.1532 [4]	0.1054 [2]	0.1942 [5]	3.0 [10]	0.3251 [6]	0.0 [1]	
FIL	0.3645 [5]	0.4727 [6]	2.1386 [9]	0.3248 [3]	0.362 [4]	0.2393 [2]	0.5161 [7]	3.0 [10]	1.3331 [8]	0.0 [1]	
FFG	0.4384 [4]	0.5161 [5]	1.8694 [9]	0.3295 [2]	0.4362 [3]	0.6306 [6]	1.0511 [7]	3.0 [10]	1.4134 [8]	0.0 [1]	
GRI	1.4556 [7]	1.6691 [9]	1.4833 [8]	0.4534 [5]	0.4867 [6]	0.4527 [4]	0.2317 [2]	3.0 [10]	0.3071 [3]	0.0 [1]	
FWR	0.0926 [4]	0.0369 [2]	0.7595 [9]	0.0916 [3]	0.0937 [5]	0.1624 [6]	0.1869 [7]	3.0 [10]	0.5875 [8]	0.0 [1]	
SMI	0.5178 [5]	0.5911 [7]	2.9228 [10]	0.4288 [3]	0.5137 [4]	0.3541 [2]	0.5202 [6]	2.8365 [9]	1.7534 [8]	0.0 [1]	
LKO	0.7053 [8]	0.1833 [5]	0.2919 [7]	0.0895 [3]	0.7059 [9]	0.0744 [2]	0.1652 [4]	3.0 [10]	0.1957 [6]	0.0 [1]	
CRI	0.0536 [4]	0.1071 [5]	3.0 [10]	0.0342 [2]	0.0473 [3]	0.2195 [7]	0.1617 [6]	1.8221 [9]	1.2348 [8]	0.0 [1]	

According to Table 2.7, it also can be seen that the overall performance of the other seven conventional probability models differ significantly across the

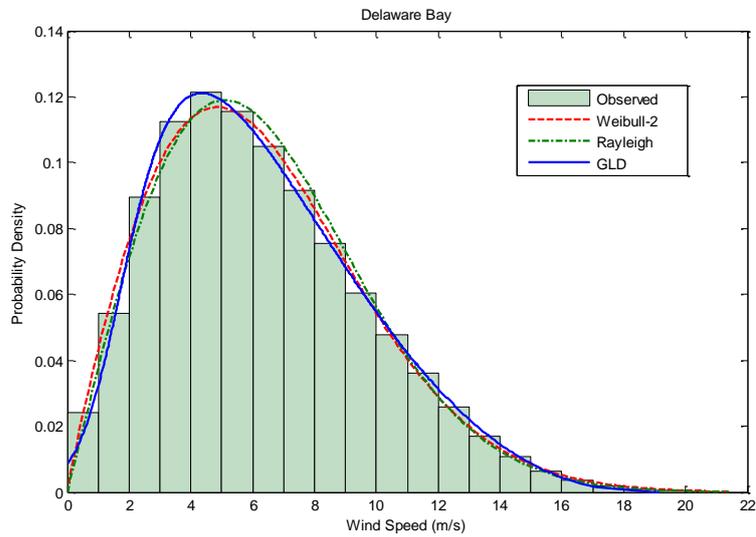
stations. Comparing the performances of the nine conventional distributions, Burr only ranks the first at DLB; Beta distribution reaches the highest fitting accuracy at stations STP, FLI, FFG and CRI; whereas, Lognormal has the top performance at DKK and GRI; Gamma presents the best overall performance at seven stations PSI, NPT, SBI, KTB, FIL, SMI and LKO and is more accurate than other conventional distributions. Then it can be concluded that any conventional probability model in consideration cannot provide consistent acceptable performance for all the fifteen stations. In contrast, GLD can achieve the highest performance at all the stations. It can be concluded that GLD generally can provide much higher accuracy on statistical wind speed fitting than conventional probability models.



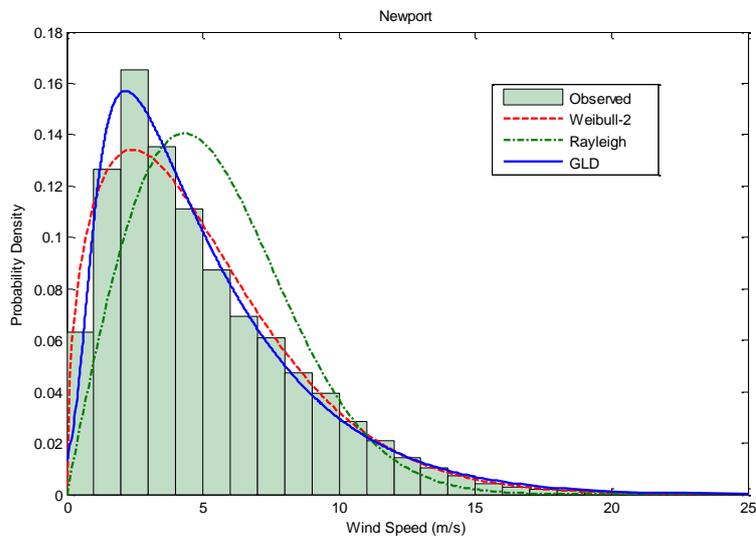
**Figure 2.1** Histogram and probability distributions for wind speed at the station Settlement Point.



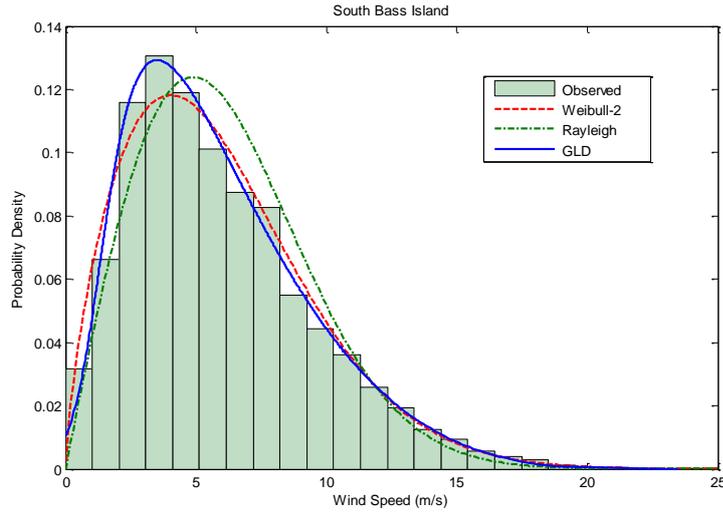
**Figure 2.2** Histogram and probability distributions for wind speed at the station Passage Island.



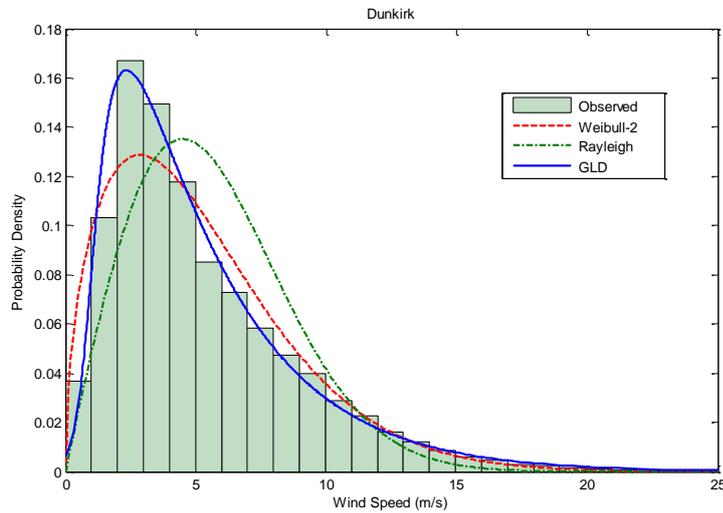
**Figure 2.3** Histogram and probability distributions for wind speed at the station Delaware Bay.



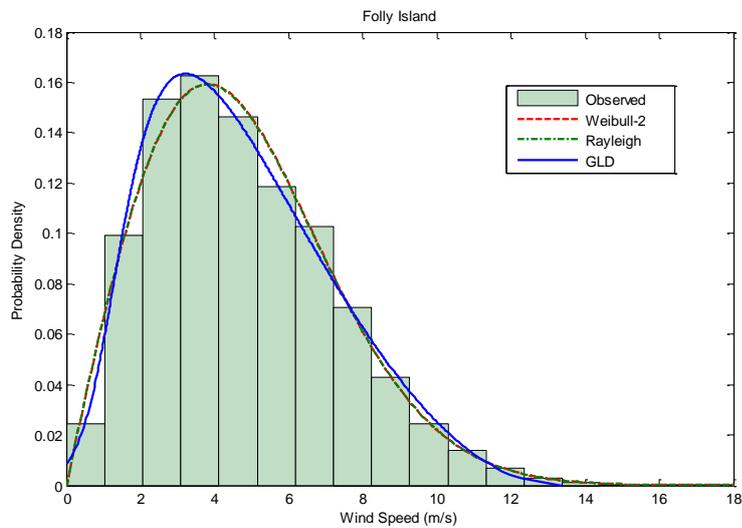
**Figure 2.4** Histogram and probability distributions for wind speed at the station Newport.



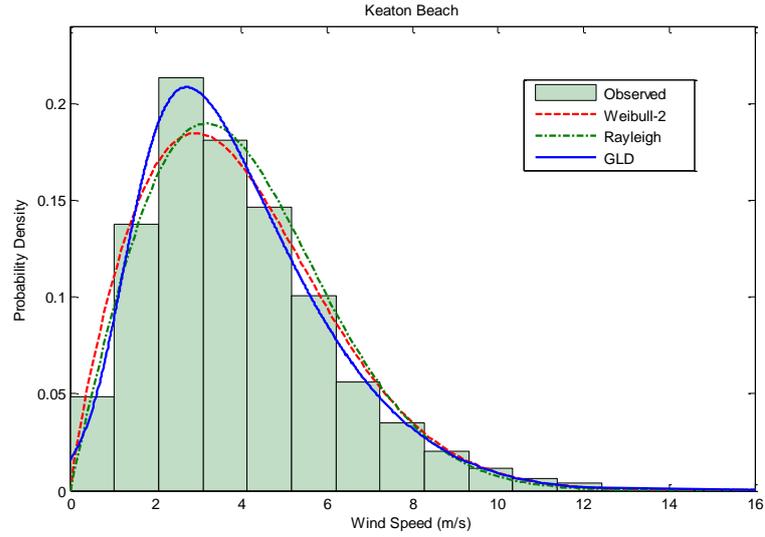
**Figure 2.5** Histogram and probability distributions for wind speed at the station South Bass Island.



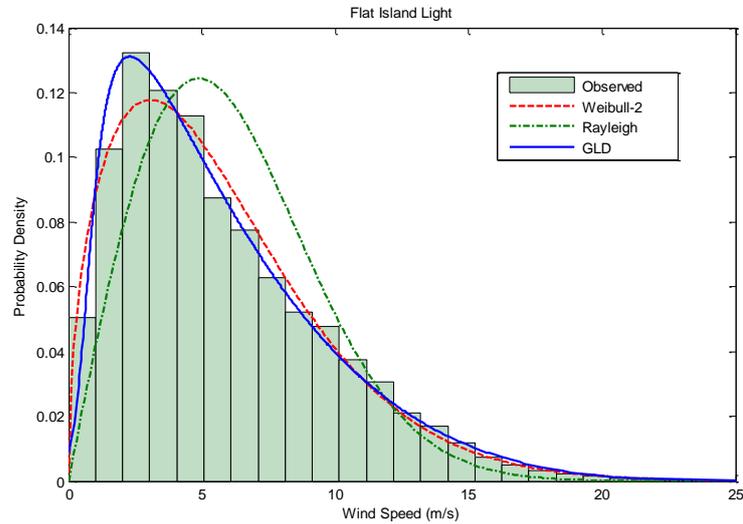
**Figure 2.6** Histogram and probability distributions for wind speed at the station Dunkirk.



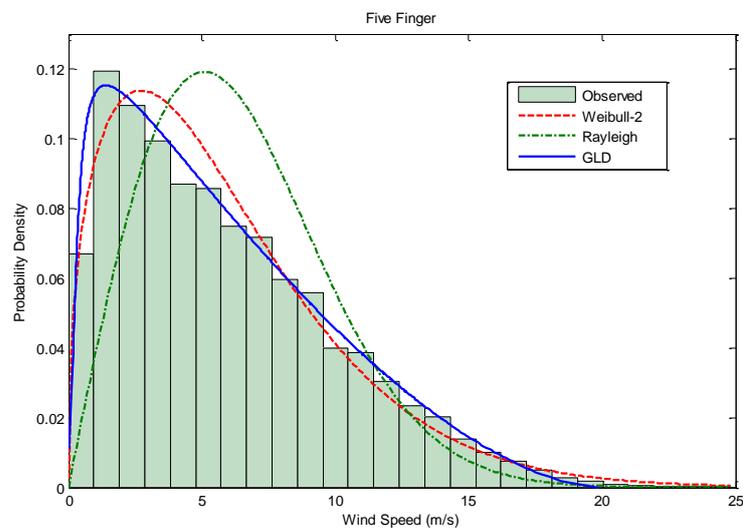
**Figure 2.7** Histogram and probability distributions for wind speed at the station Folly Island.



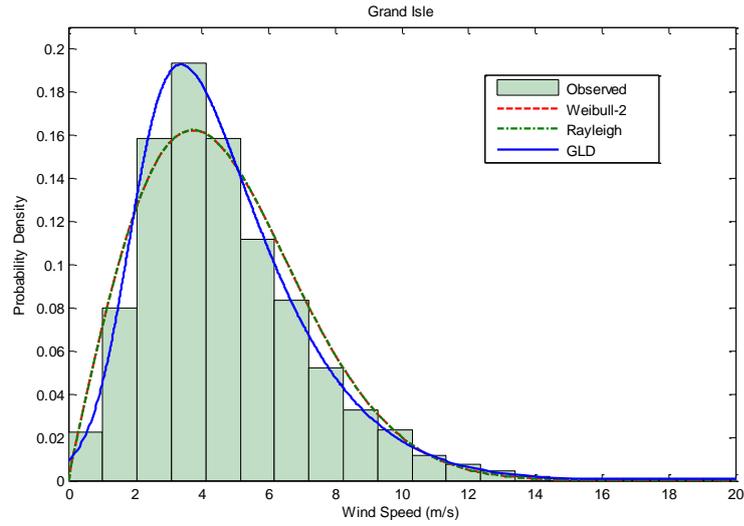
**Figure 2.8** Histogram and probability distributions for wind speed at the station Keaton Beach.



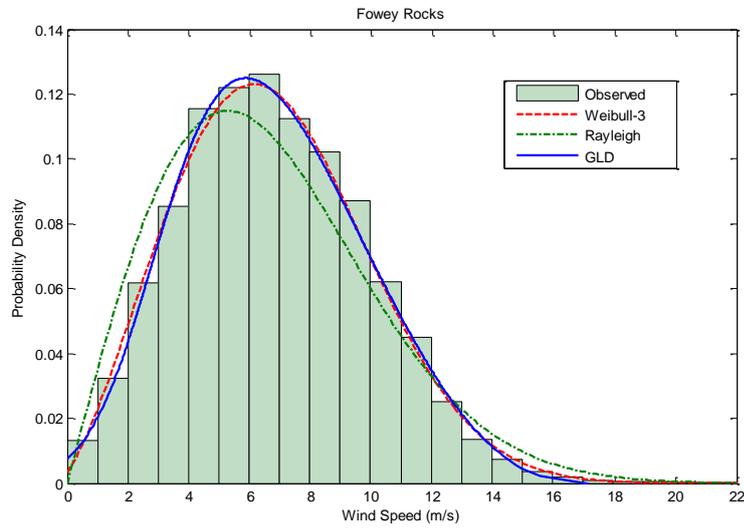
**Figure 2.9** Histogram and probability distributions for wind speed at the station Flat Island Light.



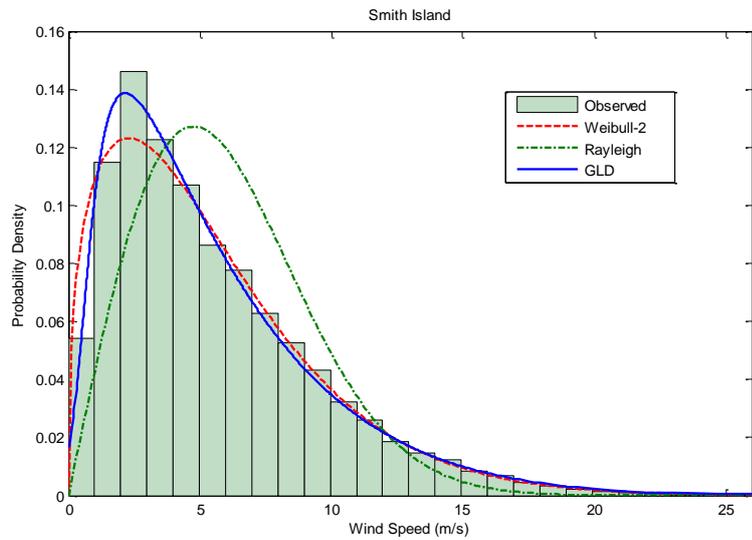
**Figure 2.10** Histogram and probability distributions for wind speed at the station Five Finger.



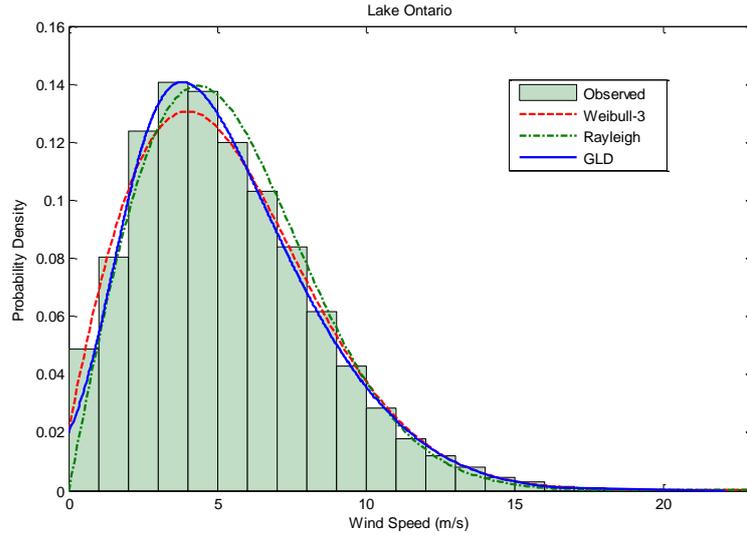
**Figure 2.11** Histogram and probability distributions for wind speed at the station Grand Isle.



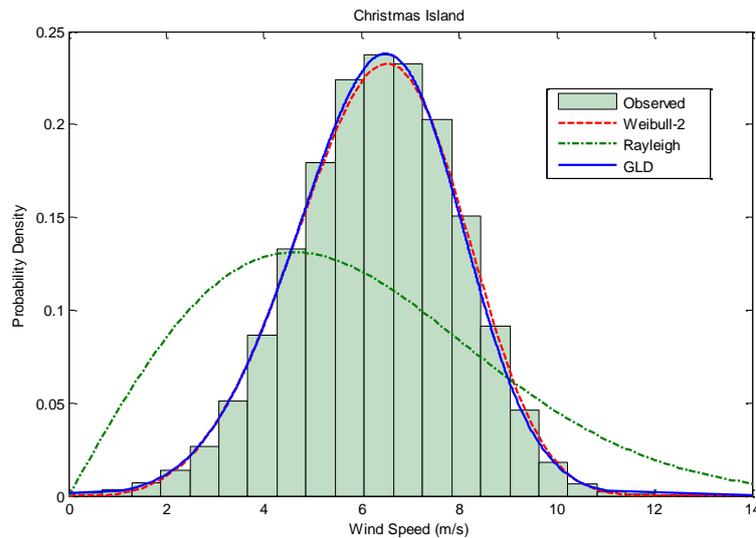
**Figure 2.12** Histogram and probability distributions for wind speed at the station Fowey Rocks.



**Figure 2.13** Histogram and probability distributions for wind speed at the station Smith Island.



**Figure 2.14** Histogram and probability distributions for wind speed at the station Lake Ontario.



**Figure 2.15** Histogram and probability distributions for wind speed at the station Christmas Island.

For the fifteen stations, the histogram and probability density graphs of the fitted GLD and the most commonly used models including Weibull (Weibull-2 or Weibull-3) and Rayleigh distributions are displayed in Figures 2.1-2.15, where the GLD consistently outperforms the other probability models. According to the graphs, it can be easily seen that wind speed at different locations has different statistical distribution characteristics. The PDF of GLD fits the best with the histogram of the observed wind speed samples. The accuracy of wind speed probability estimation is significantly improved by means of GLD model. From

Figures 2.1-2.15, GLD demonstrates much higher performance than Weibull distribution at stations PSI, NPT, SBI, DKK, KTB, FIL, FFG, GRI and SMI, which echoes the numerical results of the goodness-of-fit tests.

As can be seen above, generalized Lambda distribution with only four parameters has demonstrated fairly significant adaptability with high accuracy in describing the probability distribution of wind speed at the fifteen locations. It could be an effective alternative to the conventional Weibull distribution to model the wind speed distribution.

## **2.6 Conclusions**

Wind speed assessment plays a key role in wind turbine design and wind farm planning. This chapter has proposed the use of generalized Lambda distribution to model the wind speed distribution for higher fitting accuracy than the conventional models. Furthermore, the combined moment and starship method has also been developed in this chapter to efficiently and accurately estimate the parameters of GLD.

To compare the proposed model GLD with the conventional models, a comprehensive fitting performance evaluation method has been developed to assess the overall performance of all the used probability models. The study results on the wind speed data from NDBC have indicated that the proposed GLD features the highest accuracy and adaptability at all the fifteen stations. In contrast, conventional probability models cannot provide consistently sound performances at different stations. Furthermore, some conventional models maybe not fulfill satisfactory accuracy. The study has shown that the proposed GLD with only four parameters can provide significantly accurate wind speed description for wind power development in the future with high reliability and adaptability. For the limitation of parametric probability model, the proposed GLD cannot perfectly fit the wind speed probability distributions that are not unimodal. It can be overcome through mixed GLD distribution for the future research.

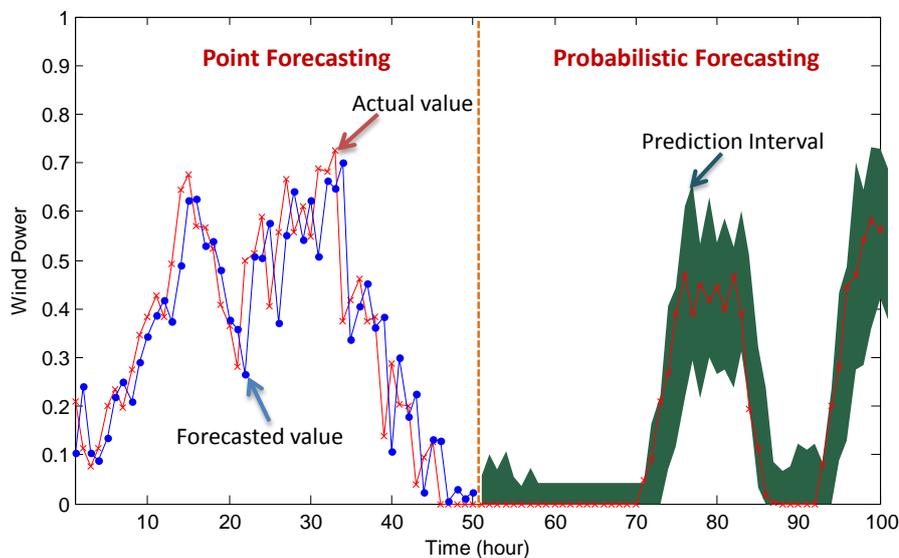
# **3** *Probabilistic Forecasting: Principle and Evaluation*

## **3.1 Introduction**

In practice, decision makers in power systems always desire to obtain a reliable and accurate prediction for an uncertain future [75]. Traditionally, point forecasts are preferred to directly guide the forecast user, since it is relatively easy to derive and understand. It should be emphasized that point forecasts only describe one possible outcome, single-valued forecasts, which might not be sufficient in practical applications. The high complexity of industrial or natural systems in our realistic world contributes to the high levels of uncertainty in the process of operation, control, etc. For instance, precise prediction of variables is extremely difficult in the chaotic weather system. Likewise, wind power forecasting is challenging in the modern power systems due to the impacts of chaotic weather systems. In electricity market environment, electricity price series has significant nonstationarities and is also difficult to forecast. In those applications, it is extremely difficult to improve the performance of the point prediction by simply adjusting forecasting models. Further, deterministic point forecasts cannot provide reliability and confidence of the prediction outcomes.

Therefore, probabilistic forecasting, also referred to as prediction intervals (PIs) or interval forecasts, has drawn tremendous attentions recently. A probabilistic forecaster produces a predictive probability distribution over future events or quantities, which can be parametric or nonparametric [76, 77]. With the likely range of outcomes, it is permitted to implement comprehensive

contingency planning for the end-user. Actually, the transformation from point estimation to distribution estimation from the nineteenth century was described in [78]. For modern weather forecasting, probabilistic forecasts have been used for a long time to generate probabilistic solutions of the meteorological problem [79]. In the previous two decades, developing good probabilistic forecasting methodologies has been the focus in the field [80]. Nowadays, probabilistic forecasting has been applied in various fields, including finance and market [81-83], population [84], project management [85, 86], software engineering [87], space study [88], hydrology[89-91], earthquake [92, 93], environment [94], as well as weather and climate prediction [80, 95], and so on. In particular, in the study area of power systems and renewable energy, it has been employed to predict wind speed [96], wind power [97], electricity load [98], electricity price [99], etc., which can provide extremely meaningful information to the corresponding decision-making processes.



**Figure 3.1** Point forecasting and probabilistic forecasting of wind power over a realistic wind farm in Australia

Figure 3.1 explicitly shows point forecasting and probabilistic forecasting of wind power generation on basis of Cathedral Rocks wind farm in South Australia with rated capacity 66 MW and visually explains the difference between

traditional deterministic point prediction and probabilistic interval forecasts. It can be found from Figure 3.1 that prediction errors can be very large at some circumstances, which can be considered as an inherent nature to some extent. In contrast, probabilistic forecasting generates the prediction intervals with nominal confidence 90% can well cover the actual value, which will assist relevant decision makers to conduct risk management. Compared to the traditional point forecast, the probabilistic forecast methods can provide prediction intervals with associated confidence levels to effectively quantify the uncertainties of future wind production, thus enabling all power system players to do e.g. beforehand preparation for possible scenarios, which can significantly reduce the risks due to high wind penetration in various operation and planning activities.

In addition to providing prediction intervals to the end-users, it is essential to evaluate the predictive ability of forecasters through comparisons of different forecasting approaches. The relevant works have been in some specific research fields, such as meteorology [100] and economy [101, 102]. It is fundamental to provide an accurate and general evaluation of the performance of probabilistic forecasts, which should be independent of any specific end-user. As analyzed in [103], probabilistic forecasting has two basic properties, calibration (also known as reliability) and sharpness. In general, probabilistic forecasting aims to maximize the sharpness of the forecasted prediction intervals or predictive distributions, on the basis of ensuring the reliability. To rank the performance of prediction intervals, proper skill score is required for evaluating the overall skill of probabilistic forecasts to involve different perspectives of PIs including both reliability and sharpness. Various skill score rules have been proposed in the past decades [104]. Information theory is used to evaluate comprehensive performance of probabilistic prediction [105]. Continuous ranked probability score (CRPS) is proposed to measure the probabilistic forecasts with respect to the predictive CDF [106]. Scoring methodologies for quantiles and prediction intervals are developed in [107, 108]. These well formulated scoring rules can provide effective criteria for developing advanced probabilistic forecasting approaches in different applications, such as wind power forecasting [109] and electricity price forecasting [99] in power systems that are focused in the thesis.

### 3.2 Prediction Intervals

From traditional point of view, forecasting is rarely considered as a purely statistical study. According to the aforementioned descriptions, probabilistic forecasting aims at estimating the probability distribution of prediction targets. The forecasting error of realistic processes can be caused by limited amount of available knowledge or chaotic characteristics of complicated systems, such as natural weather systems, and man-made systems including power systems, economic market, etc. Traditional point forecasts of a continuous variable  $t_{p+h}$  (prediction target) at time point  $p$  for  $h$  ( $h \geq 1$ ) step ahead are considered as an explanatory case herein. It focuses on finding a function  $f_h$  with parameters  $\theta$  to be approximated over the training dataset. The input variables of the forecasting model always include a vector of lagged measurements of prediction targets  $\mathbf{t}_{p,h}^*$ , and a vector of lagged measurements and forecasting value of explanatory variables  $\mathbf{u}_{p,h}$ , given by (3.1) and (3.2) respectively.

$$\mathbf{t}_{p,h}^* = (t_p, t_{p-1}, \dots, t_{p-l}) \quad (3.1)$$

$$\mathbf{u}_{p,h} = (u_p, u_{p-1}, \dots, u_{p-m}, \hat{u}_{p+h|p}) \quad (3.2)$$

For instance, these explanatory variables can be exogenous variables informing about the instantaneous weather conditions around the measurement location, e.g., wind speed, ambient temperatures. Then, the two datasets can be integrated to one vector  $\mathbf{x}_{p,h}$ , expressed by

$$\mathbf{x}_{p,h} = (\mathbf{t}_{p,h}^*, \mathbf{u}_{p,h}) \quad (3.3)$$

Therefore, the forecasting model  $f_h$  can be defined by

$$t_{p+h} = f_h(\mathbf{x}_{p,h}; \theta) + \varepsilon_{p+h} \quad (3.4)$$

where  $\varepsilon_{p+h}$  denotes the corresponding forecasting error with zero mean. According to the theory of model training and regression, the forecaster  $f_h(\mathbf{x}_{p,h}; \theta)$  will generate an approximation of the mean of the probability distribution of target. For simplicity, the datasets collected and used to train the formulated forecaster can be expressed as

$$D_t = \{\mathbf{x}_i, t_i\}_{i=1}^{N_t} = \{\mathbf{x}_{p,h}, t_{p+h|p}\}_{i=1}^{N_t} \quad (3.5)$$

The uncertainty or forecasting error in the estimation of the mean of the target distribution can be caused by various factors, including the finite training data for constructing the prediction model, a noise component coming from data acquisition devices, the choice of the model that may not reflect the true behavior of the process, the way the model parameters are estimated, and so on. It should be clarified on the difference between *confidence intervals* and *prediction intervals* [110, 111]. Confidence intervals quantify the uncertainty of the estimation of regression model. In contrast, prediction intervals correspond to the confidence in the estimation of the expected target. A prediction interval should be wider than confidence interval and encloses the corresponding confidence interval. This will be discussed in the following chapters. Quantification of the uncertainty associated with deterministic forecasting, i.e., the prediction intervals, is focused in this thesis.

Due to the complexity of probability distribution of forecasting errors, it can be approximated by the predictive distribution with non-parametric quantiles [112, 113]. Predicting the  $\alpha$ -quantile  $\hat{q}_t^\alpha(\mathbf{x}_i)$  with nominal proportion  $\alpha \in [0,1]$  of  $t_{p+h}$  (also denoted as  $t_i$  in this thesis) given the information at time point  $p$  is defined as finding the smallest value such that

$$P(t_i \leq \hat{q}_t^\alpha(\mathbf{x}_i)) = \alpha \quad (3.6)$$

If given  $\mathbf{x}_i$ , the cumulative distribution function  $F$  of  $t_i$  is increasing and known beforehand, then one specific prediction quantile  $\hat{q}_t^\alpha(\mathbf{x}_i)$  can be mathematically expressed as

$$\hat{q}_t^\alpha(\mathbf{x}_i) = F^{-1}(\alpha) \quad (3.7)$$

In practice, the decision makers are always inclined to derive the prediction intervals with given confidence  $(1 - \alpha)$ , named as *nominal coverage probability (NCP)*,

$$\hat{I}_t^\alpha(\mathbf{x}_i) = [\hat{L}_t^\alpha(\mathbf{x}_i), \hat{U}_t^\alpha(\mathbf{x}_i)] \quad (3.8)$$

Therefore, the future target is expected to lie in the PI with the probability of NCP, expressed as

$$P(t_i \in \hat{I}_t^\alpha(\mathbf{x}_i)) = 1 - \alpha \quad (3.9)$$

Usually, prediction intervals are defined as central prediction intervals, centered on the median of the predictive probability distribution. The lower and upper bounds of prediction intervals correspond to the quantiles with proportion  $\alpha/2$  and  $(1 - \alpha/2)$  respectively, given as

$$\begin{cases} \hat{L}_t^\alpha(\mathbf{x}_i) = \hat{q}_t^{\alpha/2}(\mathbf{x}_i) \\ \hat{U}_t^\alpha(\mathbf{x}_i) = \hat{q}_t^{1-\alpha/2}(\mathbf{x}_i) \end{cases} \quad (3.10)$$

$$\begin{cases} P(t_i \leq \hat{L}_t^\alpha(\mathbf{x}_i)) = \alpha / 2 \\ P(t_i \leq \hat{U}_t^\alpha(\mathbf{x}_i)) = 1 - \alpha / 2 \end{cases} \quad (3.11)$$

For the task of probabilistic forecasting, this thesis aims to develop new advanced approach including parametric and non-parametric approaches for construction of prediction intervals. It will be applied for continuous variables of practical processes in power systems, such as wind power generation and electricity price.

### 3.3 Evaluation Criteria

In principle, comprehensive assessment of PIs must include reliability as well as sharpness aspects to ensure the quality of resultant PIs [103]. However, a number of existing works on probabilistic forecasts fail to consider the sharpness aspect [114-119], which would inevitably result in prejudicial forecasts. A proper evaluation framework of PIs becomes extremely important for fair quality assessment and developing new methodologies of probabilistic forecasts. In this section, general performance indices are introduced to assess the quality of PIs, independent of specific approaches and applications.

#### 3.3.1 Reliability

Reliability refers to the statistical consistency between the probabilistic forecasts and the empirical observations. By definition from (3.8) and (3.9), the

future targets are expected to lie within the bounds of constructed PIs with a prescribed probability NCP ( $1-\alpha$ ). It is expected that the coverage probability of obtained PIs will asymptotically reach the nominal level of confidence over the full test data. PI *empirical coverage probability* (*ECP*) is a critical measure for the reliability of the constructed PIs, which is defined by

$$ECP = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} c_i \quad (3.12)$$

where  $N_{test}$  is the number of test samples, and  $c_i$  is the indicator of ECP and is defined as

$$c_i = \begin{cases} 1 & t_i \in \hat{I}_t^\alpha(\mathbf{x}_i) \\ 0 & t_i \notin \hat{I}_t^\alpha(\mathbf{x}_i) \end{cases} \quad (3.13)$$

For reliable PIs, the examined ECP should be close to its corresponding NCP. Another assessment index, *average coverage error* (*ACE*), is defined by

$$ACE = ECP - NCP \quad (3.14)$$

Generally, to ensure PIs with high reliability, the ACE should be as close to zero as possible, i.e. smaller absolute ACE indicates more reliable PIs.

### 3.3.2 Sharpness

Sharpness refers to the property of the concentration of the predictive probability distribution. The sharper forecasts have the more concentrated predictive distributions and the better prediction performance, subject to the reliability. Obviously, ECP is directly related with the sharpness of PIs. High level ECP can be easily reached via widening PIs. However, such PIs are meaningless in practice since they do not express the actual variation of the measured wind power. The width of PI for the  $i$ th target,  $\delta_t^\alpha(\mathbf{x}_i)$ , is expressed as

$$\delta_t^\alpha(\mathbf{x}_i) = \hat{U}_t^\alpha(\mathbf{x}_i) - \hat{L}_t^\alpha(\mathbf{x}_i) \quad (3.15)$$

*Average width of PIs* (*AWPI*)  $\bar{\delta}_{p+h|p}^\alpha$  quantifies the average width of PIs, defined by

$$AWPI = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \delta_t^\alpha(\mathbf{x}_i) \quad (3.16)$$

which can be used to measure the sharpness of constructed PIs. If the PIs have similar reliability, PIs with the smallest AWPI are to be preferred undoubtedly. It

should be highlighted that the training and testing data are normalized in the study.

Based on the properties of PIs described above, both the reliability and sharpness should be considered for accurate and comprehensive evaluation of constructed PIs. As a crucial feature indicating the accuracy of PIs, the reliability is usually prioritized in the evaluation process.

### 3.3.3 Skill Score

#### 1) Proper Skill Score

Various evaluation frameworks for probabilistic forecasts have been proposed [105, 107, 120]. Essentially, they employ relatively simple skill scores to comprehensively measure both reliability and sharpness. In practice, more generalized skill scores are required to assess the performance of probabilistic forecasts by forecasting developers and end-users. Proper skill scores are always preferred, since improper scores will produce prejudiced results [121]. It should be easily understood that good probabilistic forecast should have high reliability, i.e., the empirical coverage probability should be as close to the nominal probability as possible, and high sharpness, i.e., mean width of PIs should be as narrow as possible under the promise of reliability. Reasonably, fulfilling these properties is the prerequisite for realistic applications of any developed technique. Notably, the two requirements are somewhat incompatible with each other. Therefore, proper performance assessment should take both the reliability and sharpness into account and evaluate the overall skill of forecasted PIs.

Theoretically, the skill score is a function of our forecast interval  $I_t^\alpha(\mathbf{x}_i)$  and realistic targets  $t_i$ , denoted by  $Sc(I_t^\alpha(\mathbf{x}_i), t_i)$ . The empirical skill is a sample mean, given as

$$Sc(I_t^\alpha; t) = \frac{1}{N} \sum_{i=1}^N Sc(I_t^\alpha(\mathbf{x}_i); t_i) \quad (3.17)$$

We assume that the perfect interval forecast  $\widehat{I}_t^{(\alpha)}(\mathbf{x}_i)$  can be obtained and expressed as

$$\widehat{I}_t^{(\alpha)}(\mathbf{x}_i) = [\widehat{L}_t^{(\alpha)}(\mathbf{x}_i), \widehat{U}_t^{(\alpha)}(\mathbf{x}_i)] \quad (3.18)$$

Perfect PIs are defined as forecasted PIs have the best quality with exact reliability  $ACE = 0$ , and highest sharpness, i.e., wider or narrower PIs would result in lower reliability. According to the principle of defining a general skill score in [104, 121], should a generalized score for PIs be proper for any prediction interval, the following equality must hold

$$Sc(\hat{I}_t^\alpha; t) \geq Sc(I_t^\alpha; t) \quad (3.19)$$

It can be easily observed that the skill score is defined as positively oriented here.

## 2) Interval Score

When a predictive distribution  $\hat{F}$  is characterized by its quantiles,  $q_1, q_2, \dots, q_k$ , at confidence levels  $\alpha_1, \alpha_2, \dots, \alpha_k$ , if  $u_i(\cdot)$  is a non-decreasing function, and  $g(\cdot)$  can be arbitrarily chosen, the skill score defined by

$$Sc(q_1, q_2, \dots, q_k; t) = \sum_{i=1}^k [\alpha_i u_i(q_i) + (u_i(t) - u_i(q_i)) \mathbf{1}\{t \leq q_i\}] + g(t) \quad (3.20)$$

can be a proper score for assessing the set of quantiles, which is proved in [104].

In this thesis, the interval forecasts primarily aims at obtaining PIs with two quantiles (lower and upper bounds) at particular confidences. The interval score is proposed to assess the overall skill of constructed PIs from the developed interval forecasting approaches. For interval forecasting, central PIs with predictive quantiles at confidence levels  $\alpha/2$  and  $1-\alpha/2$  are focused. Combining the scoring rule for lower and upper quantiles of central PIs with nominal coverage rate  $(1-\alpha)$  and reversing the sign of the scoring rule, we can obtain the negatively oriented interval score. For simplicity, the two functions in (3.20) can be specifically defined as

$$u(x) = 4x, \quad g(x) = -2x \quad (3.21)$$

The interval score can be employed to comprehensively evaluate the overall skill of wind power PIs to assess the sharpness [104]. The interval score of the PI  $I_t^\alpha(\mathbf{x}_i)$  with NCP  $(1-\alpha)$  is defined by

$$S_t^\alpha(\mathbf{x}_i) = \begin{cases} -2\alpha\delta_t^\alpha(\mathbf{x}_i) - 4[L_t^\alpha(\mathbf{x}_i) - t_i], & \text{if } t_i < L_t^\alpha(\mathbf{x}_i) \\ -2\alpha\delta_t^\alpha(\mathbf{x}_i), & \text{if } t_i \in I_t^\alpha(\mathbf{x}_i) \\ -2\alpha\delta_t^\alpha(\mathbf{x}_i) - 4[t_i - U_t^\alpha(\mathbf{x}_i)], & \text{if } t_i > U_t^\alpha(\mathbf{x}_i) \end{cases} \quad (3.22)$$

Based on the test dataset, the overall score value  $\bar{S}_t^\alpha$  can be obtained and given as

$$\bar{S}_t^\alpha = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} S_t^\alpha(\mathbf{x}_i) \quad (3.23)$$

where  $N_{test}$  is the size of test dataset.

The interval score rewards the narrow PI and gives penalty if the target does not lie within the estimated PI. The score  $\bar{S}_t^\alpha$  can be used to evaluate the overall skill of constructed PIs by taking all aspects of PIs quality into consideration. Generally, given a particular NCP and similar ECPs, PIs with the larger interval score have the relatively higher overall skill. However, the unique interval score does not distinguish the specific contributions of reliability or sharpness to the skill. In the evaluation process, we reasonably give a higher priority to the reliability, since it is the key feature reflecting the correctness of the constructed PIs. Based on the prior analysis of PIs reliability, the interval score can be used to assess PIs from the perspective of sharpness.

### 3.4 Conclusions

Forecasting is an important research topic in the field of power system engineering. The accuracy of traditional point forecasts cannot be ensured if there is significant uncertainty involved in the complicated natural and artificial systems, such as the weather system related to intermittent renewable energy generation and electricity market. This is challenging to different decision-makers in power systems. Probabilistic forecasting is able to quantify the uncertainty associated with traditional point forecasts and provide reliable information to subsequent decision-making activities. It has attracted many attentions recently and becomes more and more important due to the high shares of renewable energies in modern power systems.

In this chapter, the backgrounds and principles of probabilistic forecasting have been mathematically introduced in detail. Comprehensive evaluation indices of prediction intervals, including reliability, sharpness and skill score, have been presented in this chapter. These evaluation criteria will be employed to systematically assess the new developed probabilistic forecasting techniques. Chapters 5-8 will focus on developing advanced probabilistic forecasting

techniques applied for wind power generation and electricity price. Generally, this chapter gives sufficient technical background information for the further studies.



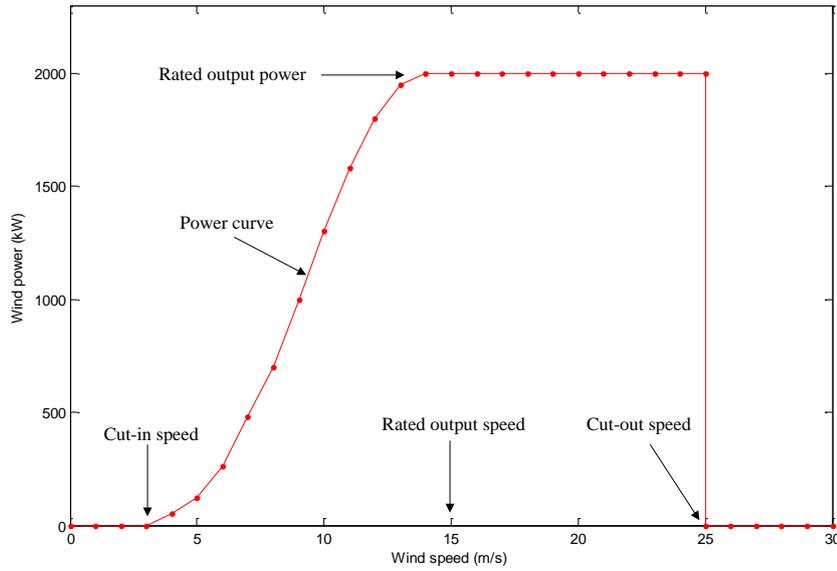
# ***4 State-of-the-art in Probabilistic Wind Power Forecasting***

## **4.1 Characteristics and Point Forecasting of Wind Power Generation**

### **4.1.1 Characteristics of Wind Power Generation**

Wind power generation demonstrates high intermittency, as it is directly related with wind speed. Wind speed has significant uncertainty, e.g., its long-term distribution has been described in Chapter 2. Theoretically, wind turbine converts the kinetic energy of wind into mechanical energy for operating an electrical generator. The energy conversion process for a single wind turbine can be typically characterized by the wind turbine power curve. Characteristic wind power curves have roughly similar shape to wind turbines of different manufactures and types.

Figure 4.1 shows the typical power curve of wind turbine of Vestas V80 which is installed on Bornholm Island in Denmark. Cut-in wind speed depicted in Figure 4.1 is the wind speed at which the wind turbine will start to produce power, about 4 m/s. The wind turbine generates the rated power when wind reaches the rated wind speed of around 15 m/s. Above the rated speed, a constant rated power is maintained until it reaches the cut-out speed. Cut-out wind speed is the highest wind speed of the wind turbine's working status, about 25 m/s, above which the turbine is stopped to avoid damage.



**Figure 4.1** Wind power curve of the Vestas V80 (2.0 MW) wind turbine installed at wind farm on Bornholm Island of Denmark

Theoretically, the power/energy  $P_w$  available in the flowing stream of wind is a cubic function of the wind speed  $v$ , which can explain the sharp increase of the power curve in Figure 4.1, defined as

$$P_w = \frac{1}{2} \rho_{air} A_w v^3 \quad (4.1)$$

where  $A_w = \frac{\pi d_t^2}{4}$  ( $d_t$  is the rotor diameter) is the swept area of the turbine blades exposed to the wind and  $\rho_{air}$  indicates the air density.

In practice, only about 20%-30% of the original power available in the wind can be captured and output finally, because of the Betz limit [122], the efficiencies of the generator and gearbox, etc. The power generation of a wind turbine  $P_T$  can be expressed by the following equation,

$$P_T = \frac{1}{2} \rho_{air} C_p \eta_g \eta_b A_w v^3 \quad (4.2)$$

where  $\eta_g$  and  $\eta_b$  denote the turbine efficiency ratio for the generator and the gearbox respectively, and  $C_p$  is the coefficient of performance for the wind turbine, which is essentially a type of power conversion efficiency, defined by

$$C_p = \frac{P_R}{P_w} \quad (4.3)$$

where  $P_R$  denotes the power actually captured by the wind turbine rotor,  $P_w$  is the available power defined in (4.1). The maximum theoretical value of the coefficient of performance  $C_p$  is 59.3%, which is determined by a fluid mechanics constraint well known as the Betz limit [122].

#### 4.1.2 Major Aspects of Wind Power Prediction

Wind power brings much more uncertainties than conventional generation. Due to the significant intermittency of wind generation, running the power systems would be nearly impossible without wind power forecasting, especially when high penetration of wind power [123]. Accurate and reliable wind power forecasting becomes extremely important to optimize the operation cost and improve the reliability of the power system with increased wind penetration [124]. In practice, different forecasting horizons have different applications corresponding to the requirements of decision-makers in power systems [33, 35, 125], given as

- *Very short-term* (From a few seconds to minutes): Very short-term forecasts can be used for the wind turbine control, electricity market clearing, such as 5 minutes for Australia Electricity Market [126], and continuous units dispatch for power system balance [127].
- *Short-term* (Up to 48-72 hours ahead): Such forecasts are crucial for different decision making problems involved in power system operation and electricity market, including unit commitment [128, 129], economic dispatch [128], dynamics security assessment [130], reserve determination [131], wind power trading [132], and so forth. Most European countries focus on the time scale from 0-48 hours [33].
- *Medium-term* (Up to one week ahead): Medium-term forecasting can be valuable for e.g. maintenance scheduling of wind farms, conventional power plants, transformer, and transmission lines [125]. For the special case of offshore wind farms maintenance costs may be high, and thus an optimal planning of maintenance schedule is of particular importance.

- *Long-term* (Up to months to years): Long-term predictions can be used for long-term wind energy assessment [125]. Like in Chapter 2, yearly wind power or wind speed description would be critical for wind farm planning.

In general, the relevant information available as inputs to the prediction models of wind power generation can include but not limited to,

- Historical measurements of wind power generation
- Historical measurements of explanatory variables, such as relevant meteorological variables including wind speed and direction, temperature, etc.
- Forecasts of explanatory variables, e.g., Numerical Weather Prediction (NWP)

#### 4.1.3 Wind power forecasting models

In the past decade, various approaches have been developed for improving the accuracy of wind power forecasting [32, 33, 125]. In general, these approaches can be divided into four categories, including the reference persistence approach, the physical approach, the statistical approach, and the hybrid approach.

*a) Persistence approach:* This prediction approach is always regarded as a naive predictor, widely used for meteorology-related forecasting [133]. This simple prediction method assumes that the wind production in the future  $P_{t+k|t}$  will be the latest measurement  $P_t$ .

$$P_{t+k|t} = P_t \quad (4.4)$$

Though with significant simplicity, the persistence approach is difficult to outperform for the look-ahead times shorter than a few hours. Especially, it can be more accurate than most physical and statistical approaches for very short-term forecasts in practice [134]. The generalized persistence method is defined that the future prediction target is the average of the last  $n$  measured values, expressed as

$$P_{t+k|t} = \frac{1}{n} \sum_{i=0}^{i=n-1} P_{t-i} \quad (4.5)$$

which is also known as moving average. Despite the simple formulation, the persistence approach is the most simple and important reference model for wind power forecasting [135]. The new developed forecast method should be examined against classical benchmarks including the persistence method to investigate the degree of improvement.

*b) Physical approach:* The physical models are directly based on meteorological forecasts through a detailed physical mathematical formulation of the chaotic atmosphere to account for various physical factors, such as roughness, orography and obstacles [136]. Firstly, wind speed and direction are obtained from NWP on a coarse grid surrounding geographical points of the wind farm. The wind speed and direction from the relevant NWP level is transformed to the onsite location of wind farm and the hub height of the wind turbine [137], which is known as *downscaling*. Then the derived wind speed is *converted to power* through a power curve, which can be the manufacture power curve or the empirical power curve estimated from the forecasted wind speed and direction and measured power. To take into account systematic prediction errors from the NWP model or other physical formulations, Model Output Statistics (MOS) or different relatively simple statistical techniques are applied for the prediction data to reduce the remaining residual errors.

It should be noted that the temporal resolution of NWP is usually within 1 and 3 hours. Therefore, it is not possible to accurately estimate the power output inside this time interval using the NWP information alone. Instead, stochastic methods can be used to predict the power output depending on historical explanatory variables, and historical measured wind power.

*c) Statistical approach:* Statistical approach is to find out the mapping between historical measurements of wind power as well as historical and forecast values of explanatory variables, and wind power measures. The forecasting models are trained over the collected measurement data through minimizing the difference between the predicted and the actual values based on proper formulated loss function.

In general, statistical approaches can be classified as direct time series-based models including Autoregressive Moving Average (ARMA) [138, 139], Kalman Filter [140, 141], wavelet [142], orthogonal fitting [143], etc., and machine

learning based approaches, including those based on artificial neural networks (ANNs) [144], support vector machine (SVM) [145], local recurrent neural networks [146], data mining algorithms [147], clustering algorithm [148], etc. In addition, a hybrid approach combining linear autoregressive model and an adaptive fuzzy logic based model is proposed for cases of wind power prediction in Crete and Shetland [149]. In [150], various models including feed-forward NNs, radial basis function networks, adaptive network based fuzzy inference system, etc., are compared to forecast hourly mean wind speed, indicating that all nonlinear models perform better than any linear models.

*d) Hybrid approach:* In most commercial models, a combination of physical and statistical approaches is used, since both approaches can obtain good forecasting performance. The physical approach relies on the NWP information, which is always time-consuming and expensive and also difficult to forecasting with look-ahead time shorter than the NWP temporal resolution. Typically, prediction models using NWP outperform direct time series approaches for longer than 3-6 hours look-ahead time, consequently the NWP should be taken as the input of statistical approaches to ensure the forecasting performance [33]. Recent studies indicate that a combination of these model for aggregating wind power over a large region and shorter horizons can reduce the average error of the prediction [151]. An adaptive combination of forecasts is investigated to produce efficient wind power forecast [152]. An advanced hybrid prediction system composed of an ensemble of NWP forecasts and persistence, autoregressive, and autoregressive moving-average models is developed to provide accurate wind speed forecasts with 1 hour ahead of time, which is more efficient and accurate than traditional NWP models and can provide an alternative for common statistical approach for several hours ahead wind prediction [153].

Many industrial systems for wind power forecasting have been developed and applied worldwide, including *WPPT* used in Denmark, *AWPT* used in Germany, *Sipre dlico* used in Spain, *AWPPS* used in Ireland, *Prediktor* used in Spain, Denmark, Ireland, Germany, and USA, *Zephyr* used in Denmark, and so on [154].

## 4.2 Probabilistic Wind Power Forecasting Approaches

Traditionally, most researches focus on developing accurate point forecasting methods for wind power [32, 33, 155-157]. Due to the chaotic nature of the weather system, the processes of wind power demonstrate significantly non-stationary and nonlinear characteristics, and consequently errors in wind power forecasting are simply unavoidable and quite often can be significant. Wind power forecasting errors have significant impacts on power system operation [158], electricity market [34, 159], etc. Wind power forecasting error is statistically analyzed and explicitly observed in [160]. It is found that conditional distribution of wind power prediction error can be approximately modeled by normal distribution [161]. The distribution of wind power forecasting error is proven to have the shape of fat tail with large kurtosis and modeled with Beta distribution [162, 163]. A mixed probability distribution is proposed for estimation of the probability of wind power forecasting error and used for penalties analysis in the deregulated electricity market [159].

Because of unavoidable and sometimes significant forecasting errors of wind power, heavy attentions are paid to developing probabilistic forecasting techniques for wind power generation to construct PIs quantifying the prediction uncertainty [97].

### 4.2.1 Statistical Approaches

#### *1) Non-parametric Probability Density Functions*

According to the previous researches introduced above, it is very difficult to accurately describe the probability distribution of wind power forecasting error, due to the extreme complexity of weather systems. In order to avoid assumptions on the shape of predictive distributions of wind power, probabilistic predictions produced from non-parametric methods are preferred in practice, which aim at obtaining quantiles of the predictive distribution.

*Quantile regression (QR)* approach can be used to estimate different wind power forecasting quantiles without the assumption of distribution shape. Local quantile regression (LQR) has been introduced for the estimation of a finite number of quantiles of the conditional wind power distribution [112]. For the

LQR model, the dependence of the forecasted quantiles on the explanatory variables is modeled via linear regression. In addition, this approach has the disadvantage that only one quantile can be estimated once, which is indeed time consuming. This approach is compared with two other quantile estimators, Local Gaussian and Nadaraya-Watson methods, demonstrating similar performance [164]. Spline quantile regression (SQR) approach is proposed for probabilistic forecasting of wind power, of which the quantiles are formulated as a set of nonlinear functions [165]. Time-adaptive SQR approach based on simplex algorithm is developed and can significantly save the computational time [166]. In [167], various variability indices are extracted from wind power series and employed as explanatory variables for QR models to involve the influence of changing weather regimes and reach a more accurate performance.

*Kernel density estimation* (KDE) method is an advanced and popular nonparametric approach for PDF estimation. The probability density curve can be derived by adding up all curves of probability density at each point. The explanatory variables, such as meteorological information including wind speed, wind direction, etc., are directly related with wind power. This nonlinear correlation should be identified before using the KDE approach, and the actual shape of the PDF curve of predicted wind power is actually estimated with a discrete-continuous mixed model [168]. Then, prediction intervals of wind power can be obtained on basis of the obtained PDF. The Nadaraya-Watson estimator based KDE approach is proposed for wind power forecasting to improve the reliability [169]. A novel time-adaptive quantile-copula estimator is developed for kernel density forecasts of wind power generation [170]. A conditional KDE approach is developed to account for the measurements of wind power with an exponential process and conditionally update the prognostic PDF of wind power generation [171].

*Adaptive resampling* approach is developed to generate nonparametric quantiles conditioning on the point forecasting results of, e.g., AWPPS, WPPT and Siprońico [113]. The wind power is classified into three levels using fuzzy logic sets based on the power curve. With nonparametric probability forecasts of wind power, the statistical scenarios of short-term wind power can be generated [172].

Due to the excellent nonlinear regression capability, *neural networks* (NNs) based approaches are developed for probabilistic forecasting of wind power. In [173], radial basis function has been implemented to derive PIs of wind power based on point prediction results, weather conditions, and etc. In [174], the Coverage Width-based Criterion (CWC) approach is proposed to generate the ‘optimal’ lower and upper bounds of wind power PIs through the trained NN using the CWC-based cost function. However, this approach will generate too wider or narrower PIs due to the flawed evaluation index CWC, i.e., the true optimal PIs can never be obtained by this approach, which has been mathematically analyzed and discussed in [175].

*Evaluation framework* for nonparametric quantile forecasts is systematically proposed in [109], consisting of measures and diagrams, to provide comprehensive information of the required and desirable properties, including reliability, sharpness, resolution and skill.

## **2) Parametric Probability Density Functions**

Parametric probabilistic forecasts approaches are also developed for wind power prediction, with the advantage of fast computation speed. *Exponential smoothing method* (ESM) is proposed for multi-step density forecasts of wind power based on the truncated normal distribution modeling of forecasting errors [176]. *Markov-switching autoregressive* (MSAR) model is used for 10-min resolution time series of wind power to capture the regime-switching behavior which cannot be explained by the evolution of some explanatory variables [177]. Autoregressive and conditional parametric autoregressive models together with the proposed generalized logit-normal distribution are used for very short-term wind power forecasting [127]. *Neural network* based approaches with normal distribution assumption are proposed for wind power forecasting [178, 179].

## **4.2.2 Ensemble Forecasts**

Ensembles of forecasts methodology can be used to quantify the inherent uncertainty involved in the NWP processes instead of just increasing the resolution more and more, through generating multi-scenario of meteorological variables. In general, it can be divided into *multi-model ensemble* using different NWP models produced by different parameterizations of the same NWP model,

or different NWP models developed by different institutes, and *single-model ensemble*, varying the initial conditions of the NWP model or collecting NWP outputs for the prediction time point determined by the NWP temporal resolution and the look-ahead time [180]. Then, the ensemble forecasts of meteorological variables are converted to wind power ensemble forecasts. In practice, wind power forecasting is inclined to use different kinds of ensembles, e.g. Zephyr developed in Denmark [181]. The ensemble forecasts can provide medium-term (up to one week) probabilistic prediction [180]. Prediction Risk Index (PRI) is defined in [160] to comprehensively describe the expected level of forecast uncertainty and used to measure wind power ensemble predictions derived from ECMWF and NCEP, as well as a lagged average approach [182]. A complete ensemble-based probabilistic forecasts technique is proposed and applied for the practical case of Horns Rev wind farm in Denmark, and the NWP ensembles are converted to wind power using the novel orthogonal fitting method [183]. Wind power point forecasts and probability density function of associated uncertainty are obtained from weather ensemble predictions generated by an atmospheric model incorporating calibration and kernel smoothing approaches in [184]. A hybrid approach combining Bayesian model averaging (BMA) and ensemble approach is proposed to fulfill probabilistic prediction of wind power [96], where the statistical BMA approach is used to calibrate the forecast ensembles to derive better calibrated predictive PDFs.

### **4.3 Practical Applications in Power Systems**

The increased penetration of wind power brings significant challenges to modern power systems, because of the intermittency and uncertainty of wind power generation. With high penetration of wind power, the knowledge of uncertainties ahead provided by probabilistic forecasts can be extremely valuable to a number of power system operation and management procedures. Based on the PIs with associated confidence level, the quantified uncertainties of wind power forecasts can provide useful information to decision makers, including wind power producer, transmission system operator (TSO), traditional generation company (GENCO), etc., to well prepare for the worst and the best conditions ahead. Probabilistic forecasting of wind power generation can effectively assist

different aspects of power system operation and planning, including electricity market running, reserve determination, energy storage sizing, unit commitment, economic dispatch, and so on.

The integration of wind generation can have important effects on electricity price in the liberalized electricity market, which have been discussed and analyzed on basis of different realistic electricity markets including Denmark [185], UK [186], Australia [187], USA [188] and so on. In electricity market, the wind power producer suffers from penalties related to regulation costs because of the serious fluctuation of wind power. Wind generation prediction errors are closely related to such penalties [34, 159]. A general methodology accounting for predictive distributions obtained by probabilistic forecasts of wind generation is developed to design advanced bidding strategies, which is applied for Danish wind farms with significant benefits [132]. The prospective applications of probabilistic predictions of wind generation in electricity markets are discussed in [189], including system scheduling and dispatch and other relevant decision-making problems to TSO. It is found that probabilistic forecasts can reduce the cost and improve the reliability of power systems.

Under a large scale penetration of wind generation, determining adequate operating reserve is highly concerned by the TSO because of security and economy reasons. Traditional deterministic reserve management tools using e.g. the famous  $N-1$  criterion cannot satisfy the requirement of the large integration of wind power. The new reserve management tools that can involve the uncertainty information given by probabilistic forecasting methodologies are needed to help to make operational decisions. A risk evaluation methodology is formulated to describe the consequences of each possible reserve level through a set of risk indices for decision making [131, 190], where the decision strategies are searched via a given risk level compromising between economic issues and the risk of loss of load and the effectiveness is demonstrated on Portuguese power system. A probabilistic method using the probabilistic predictions is developed to help with determining the operating reserve requirements of a TSO in Portugal [130]. Meanwhile, a fuzzy power flow tool for system steady state security assessment is also developed to identify potential congestion conditions and

voltage violations in the transmission system based on the probabilistic forecasts of wind power [130].

Energy storage is considered as an efficient solution to reduce wind power fluctuations and improve the reliability and performance of wind penetrated power grids. A novel energy storage sizing approach is proposed to plan an energy storage system (ESS) and then reduce the uncertainty of short-term wind power forecasts up to 48 hours [191]. The probabilistic forecasts based approach significantly reduces the capacity of the ESS and gives useful information about the uncompensated energy. A dynamic storage sizing methodology is proposed for a dynamic assessment of the necessary storage capacity for different delivery periods, on basis of the risk of the wind power prediction [192], which reaches a significant reduction of the storage capacity needed and ensures the profit of wind power producer.

As wind power is increasingly integrated into power systems, it is challenging for TSO to ensure the security of power grids under uncertain circumstances and to maintain the maximum utilization of renewable energy. A probabilistic methodology is proposed to incorporate multiple sources of uncertainty especially for wind power forecasts in power system operation, dispatch, and unit commitment procedures, integrated with an energy management system (EMS) for realistic electricity grids operation [193].

A comprehensive computational framework is proposed to include wind power uncertainty that is quantified via ensemble NWP model in formulations of stochastic unit commitment/economic dispatch [129]. In order to accommodate wind power output uncertainty, a robust optimization approach is proposed in [194] to make optimal unit commitment schemes for traditional thermal units in the day-ahead market. The proposed robust optimization formulates uncertainty sources through a set of random samples including the worst-case scenario and reaches the minimal cost under the worst wind power output scenario. In addition, the intermittency and volatility of forecasted wind generation is also taken into consideration in the security-constrained unit commitment [195]. In context of smart grid and electricity market, demand-side management of price responsive demand combining with the uncertainty information of probabilistic wind power

predictions can facilitate the efficient and flexible power system operation and accommodate large scale integration of wind power [128].

With more and more wind power penetration, much more uncertainties than before have been brought into the decision-making processes in power systems. As introduced above, probabilistic forecasts may either be integrated in the decision process of energy management or bidding strategies in an electricity market. Generally, quantitative uncertainty information of wind power prediction can provide an alternative yet effective input for different decision-making problems in power systems.

#### **4.4 Conclusions**

Wind power forecasting is crucial to the operation of power system and electricity market. Traditional point prediction of wind power can produce significant forecasting errors. To overcome the limitation, probabilistic wind power forecasting has received a lot of attentions. This chapter provides a comprehensive review on the state-of-the-art of probabilistic wind power forecasting for the related studies in Chapters 5-7.

The essential characteristics of wind power generation are introduced, including wind power curve and wind power function, etc. In practice, there are several different prediction horizons for wind power forecasting according to corresponding practical applications in decision-making activities. Relevant variables commonly used as the inputs of prediction model are also explicitly depicted in this chapter. In general, point forecasts of wind power can be considered as the basis of probabilistic wind power forecasting. A series of techniques for wind power point forecasting have been developed and can be divided into persistence approach, physical approach, statistical approach and hybrid approach. Then, probabilistic wind power forecasting methodologies are systematically investigated and classified into statistical approach and ensemble forecasts approach. The practical applications of probabilistic forecasting of wind production in various aspects of power systems and electricity market are comprehensively surveyed, which strongly indicates the significant meanings.



# **5** *An Advanced Statistical Approach for Probabilistic Forecasting of Wind Power Generation*

## **5.1 Introduction**

It is difficult to accurately forecast wind power because of the non-stationary and chaotic nature of wind power series. Theoretically, the multi-layered feedforward NN can be a universal function approximator to map any complex nonlinear relationships to any degree of desired accuracy [196]. Therefore, NNs are widely applied for forecasting in many different fields of business, industry and science featuring high tremendous performance and high flexibility [197]. Due to the excellent approximation and generalization capabilities, NNs are also popularly used for wind power forecasts [144-146, 198], irrespective of some drawbacks like local minima, overtraining, and high computational costs. However, NN-based forecasting methods cannot provide satisfactory predictions if the training data are chaotic or too noisy. Usually the prediction performance cannot be improved by changing the NN structure or increasing the training iteration.

To effectively account for forecasting uncertainties, several approaches have been developed to obtain PIs for NN based methods, including delta, Bayesian, bootstrap, and mean-variance estimation methods [110, 116, 199, 200]. The Bayesian and delta approaches can be only used for construction of PIs of homogeneous noise with constant variance, and the accuracy cannot be ensured for the heterogeneous noise. Besides, they suffer intensive computation due to

the involved complex derivatives and Hessian matrix. Among these techniques, the bootstrap approach can flexibly approximate the non-constant variance and has proved to produce more reliable PIs [201]. Moreover, it prevents complicated formula derivatives in the delta and Bayesian approaches. Due to relatively simple implementation and reliable performance, the bootstrap approach is taken as the most commonly used approach for NN-based PIs construction [175, 202]. However, due to the limitations of traditional NNs, the bootstrap approach suffers from significantly high computational burden, especially for large datasets.

Bayesian based NNs are used for very short-term probabilistic wind power forecasting [179]. It cannot well quantify the heteroscedasticity wind power series. Bootstrap based traditional NNs are employed to constructed PIs of wind generation in [178], which suffers from extremely high computational burden, due to the application of traditional gradient based NNs. Moreover, in [178] the applied evaluation index of PIs, CWC index, is not mathematically rigorous and cannot provide precise evaluation of constructed PIs, consequently misleading results can be produced. This has been explicitly proven in [175].

In this chapter, a new probabilistic wind power forecasting approach is proposed based on the extreme learning machine (ELM) which is a novel learning algorithm proposed for training single-hidden layer feedforward neural networks (SLFNs). It randomly chooses the input weights of hidden layer neurons and analytically determines the output weights through simple matrix computations, therefore featuring an extremely faster learning speed than for most popular learning algorithms such as Back-propagation [203]. ELM has also demonstrated excellent generalization capability and outperformed traditional NNs. In practice, ELM has been used in many different applications including both regression and classification tasks [204-206].

The bootstrap technique for traditional NNs cannot be applicable to the case of ELM, since the associated learning process is very different from that for conventional NN learning algorithms. Therefore a bootstrap based ELM approach (BELM) is newly developed to construct PIs taking the heteroscedasticity of wind power time series into account. The proposed BELM method can rapidly formulate the PIs through extremely fast learning by ELM.

Notably, though with high extendibility, the work in this chapter focuses on a simplified approach with fast speed, using the historical wind power data alone while providing satisfactory performance for hourly ahead and intra-hour forecasting, which is significant for power system operation and control in practice. For instance, in the Nord Pool market in Scandinavia, the hourly market plays a key role in maintaining system balance [29].

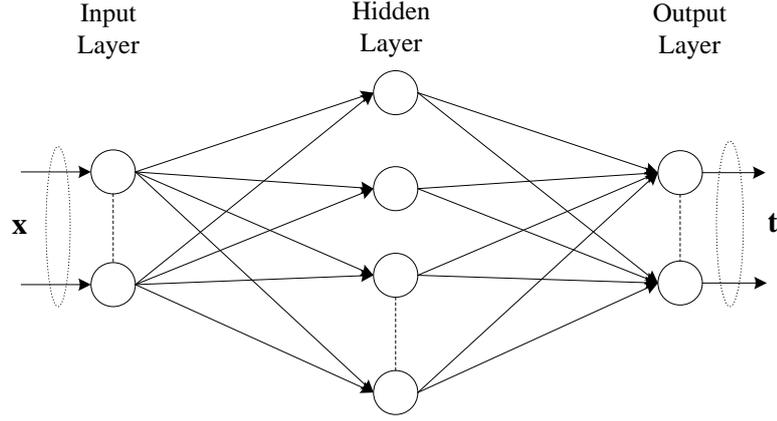
The proposed BELM method has been tested using the measurement data of a wind farm in Australia. The reliability and overall skill of the forecasting results have been comprehensively evaluated to demonstrate the effectiveness of the proposed method. The proposed approach gives a general framework for short-term probabilistic wind power forecasting. With high reliability, efficiency and flexibility, the proposed approach can potentially provide an online tool for power system operation and planning, e.g. to assist TSO in determining the required reserves in advance to avoid either higher costs or excessive risks under traditional deterministic reserve dispatch, and to help suppliers manage risks facing in electricity market trading through strategic bidding.

## 5.2 Extreme Learning Machine

ELM is simply a single hidden-layer feedforward neural network [203, 205]. Instead of using traditional gradient based learning methods that can involve many iterations, the ELM randomly chooses the input weights and biases, and subsequently determines the output weights through simple matrix computations. Given  $N$  arbitrary distinct samples  $\{(\mathbf{x}_i, \mathbf{t}_i)\}_{i=1}^N$ , where  $\mathbf{x}_i \in \mathbf{R}^n$  with  $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$  and  $\mathbf{t}_i \in \mathbf{R}^m$  with  $\mathbf{t}_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T$ , ELM with  $K$  hidden nodes and activation function  $g(\cdot)$  can be mathematically modeled as

$$f(\mathbf{x}_j; \mathbf{w}, b, \beta) = \sum_{i=1}^K \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i), \quad j = 1, \dots, N. \quad (5.1)$$

where  $\mathbf{w}_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$  is the weight vector connecting the  $i^{\text{th}}$  hidden node and the input nodes and  $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$  is the weight vector connecting the  $i^{\text{th}}$  hidden node and the output nodes, and  $b_i$  is the threshold of the  $i^{\text{th}}$  hidden node. The typical structure of an ELM is shown in Figure 5.1.



**Figure 5.1** The typical structure of ELM.

The standard ELM with  $K$  hidden nodes and activation function  $g(\cdot)$  can approximate the  $N$  samples with zero error means that

$$\sum_{i=1}^K \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) = \mathbf{t}_j, \quad j = 1, \dots, N. \quad (5.2)$$

The above  $N$  equations can be rewritten as

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \quad (5.3)$$

where  $\mathbf{H}$  is the hidden layer output matrix of the ELM,

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_K \cdot \mathbf{x}_1 + b_K) \\ \vdots & \cdots & \vdots \\ g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_K \cdot \mathbf{x}_N + b_K) \end{bmatrix}_{N \times K} \quad (5.4)$$

The  $k$ th column of  $\mathbf{H}$  denotes the output vector of the  $l$ th hidden node corresponding to the inputs  $\mathbf{x}_j$ . In (5.3),  $\boldsymbol{\beta}$  is the matrix of output weights and  $\mathbf{T}$  is the matrix of targets, respectively expressed as,

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \quad \text{and} \quad \mathbf{T} = \begin{bmatrix} \mathbf{t}_1^T \\ \vdots \\ \mathbf{t}_N^T \end{bmatrix}_{N \times m} \quad (5.5)$$

The input weights  $\mathbf{w}_i$  and the hidden layer biases  $b_i$  are randomly generated using continuous probability distributions and in fact not necessarily tuned. The hidden layer output matrix  $\mathbf{H}$  can actually remain unchanged once random values

have been assigned to these parameters in the beginning of learning. Find specific parameters  $\hat{\mathbf{w}}_i$ ,  $\hat{b}_i$  and  $\hat{\beta}$ , such that

$$\begin{aligned} & \left\| \mathbf{H}(\hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_K, \hat{b}_1, \dots, \hat{b}_K) \hat{\beta} - \mathbf{T} \right\| \\ & = \min \left\| \mathbf{H}(\mathbf{w}_1, \dots, \mathbf{w}_K, b_1, \dots, b_K) \beta - \mathbf{T} \right\| \end{aligned} \quad (5.6)$$

which is equivalent to minimizing the cost function of the traditional gradient-based learning algorithms used in back-propagation (BP) learning,

$$C_{BP} = \sum_{j=1}^N \left[ \sum_{i=1}^K \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i) - t_j \right]^2 \quad (5.7)$$

Given the input weights and the hidden layer biases are randomly assigned and fixed, training an SLFN is simply equivalent to finding a least-squares solution of the linear system. The smallest norm least-squares solution of the above linear system is

$$\hat{\beta} = \mathbf{H}^\dagger \mathbf{T} \quad (5.8)$$

where  $\mathbf{H}^\dagger$  is the Moore-Penrose generalized inverse of matrix  $\mathbf{H}$ . The singular value decomposition (SVD) method is generally used to obtain  $\mathbf{H}^\dagger$ .

The advantages of the ELM algorithm are significant [203, 205, 206]. Without iterative gradient based training, it avoids many limitations of conventional gradient-based NN training algorithms, such as the local minima, the overtraining, and the high computing burdens, etc. For any infinitely differentiable activation function, the ELM with  $N$  hidden layer neurons can learn  $N$  distinct samples exactly with zero error. In addition, ELM training can obtain the best results according to the assigned input weights. The training speed is extremely fast due to the simple matrix operation in (5.8). ELM also distinguishes from traditional NNs in superior generalization capability without the overtraining issue.

## 5.3 Formulation of Prediction Intervals

### 5.3.1 Uncertainty Analysis

The uncertainty of NN based prediction is mainly due to the noise of training data and the misspecification of NN model for regression.

#### 1) *Uncertainty in NN Model*

Misspecifications in model structure and parameters should account for the uncertainty of neural network forecasting, which may be caused by the local minima in training process, the randomly generated input weights, etc. Besides, even if the global minimum can be reached, the misspecification of model structure also introduces innegligible uncertainties in prediction results. The model uncertainty also comes from another fact that training based on finite samples can never guarantee consistent generalization performance of NN for the unseen future. Particularly, in the study of wind power forecasting herein, it is impossible to find perfect information to reduce uncertainties of predictions. These factors are collectively termed as model uncertainty. Because of the model uncertainty, the output uncertainty of neural network should be well addressed in order to produce accurate estimation.

#### 2) *Uncertainty in Data*

Except for the model uncertainty, the data noise also contributes to the prediction uncertainty. If the data exhibit stochastic characteristics, it is extremely difficult to model them in a deterministic manner. Especially, when dealing with non-stationary time series, the data noise has significant influences on the prediction results. In the study, wind power data is highly chaotic. Determining the variance of the data noise is critical in constructing prediction intervals.

Both model misspecification and data noise are the major sources of uncertainties that affect the forecasting results. Therefore the main task of probabilistic forecasting is to quantify the prediction intervals with associated confidence taking the two uncertainties into account.

### 5.3.2 Prediction Interval Formulation

Theoretically, multilayered feedforward neural networks are universal approximators and, as such, have an excellent ability of approximating any nonlinear mapping to any degree of accuracy [196]. In this chapter, the SLFN based ELM is applied for the regression task to estimate the underlying mathematical relationship between input and output variables based on a finite set of training data possibly corrupted by noises. Given a set of distinct pairs  $\{(\mathbf{x}_i, t_i)\}$ , the measured data can be modeled by

$$t_i = f(\mathbf{x}_i; \mathbf{w}, b, \beta) + \varepsilon(\mathbf{x}_i) = y(\mathbf{x}_i) + \varepsilon(\mathbf{x}_i) \quad (5.9)$$

where  $t_i$  is the  $i$ th measured target,  $\mathbf{x}_i$  denotes relevant input variables that can include historical wind power and wind speed, numerical weather predictions (NWP) including wind speed and wind direction, and so forth for wind power forecasts,  $\varepsilon(\mathbf{x}_i)$  denotes the noise with zero mean, and  $y(\mathbf{x}_i) = f(\mathbf{x}_i; \mathbf{w}, b, \beta)$  is the true regression mean. The error term moves the target away from its true regression mean  $y(\mathbf{x}_i)$  toward the measured value  $t_i$ . We assume that the noise is more or less normal distributed with variance  $\sigma_\varepsilon^2$  that may depend on the input vector  $\mathbf{x}_i$ , i.e.,

$$\varepsilon(\mathbf{x}_i) \sim N_G(0, \sigma_\varepsilon^2(\mathbf{x}_i)) \quad (5.10)$$

Actually, in the study the censored normal distribution is used to model the wind power prediction uncertainty, with potential concentration of probability mass at the bounds of the unit interval  $[0, 1]$ , which maintains PIs within the wind power capacity range [177]. To some extent, the censored normal distribution can fit different skewnesses, i.e. different shapes of probability distributions [207]. In addition, it has been studied in [75] that, even if the actual error distribution is non-normal, the time series models based on normal distribution assumption can still be applied with satisfactory performance. In the following sections, the censored normal assumption will also be proved to be reasonable and acceptable by generating reliable PIs based on actual wind power data.

In practice, the trained neural network  $\hat{y}(\mathbf{x}_i)$  could be regarded as an estimation of the true regression  $y(\mathbf{x}_i)$ . In principle, NNs generate the averaged values of targets conditioned on input variables vector  $\mathbf{x}_i$ ,  $E[t_i | \mathbf{x}_i]$  [208].

$$\hat{y}(\mathbf{x}_i) = f(\mathbf{x}_i; \hat{\mathbf{w}}, \hat{b}, \hat{\beta}) = E[t_i | \mathbf{x}_i] \quad (5.11)$$

According to the two uncertainties discussed in the preceding section, we can divide the prediction errors into two components including the one involved in the estimation of the true regression and the other involved in the estimation of the measured targets. Then the prediction error can be expressed as

$$t_i - \hat{y}(\mathbf{x}_i) = [y(\mathbf{x}_i) - \hat{y}(\mathbf{x}_i)] + \varepsilon(\mathbf{x}_i) \quad (5.12)$$

where  $t_i - \hat{y}(\mathbf{x}_i)$  denotes the total prediction error, and  $y(\mathbf{x}_i) - \hat{y}(\mathbf{x}_i)$  denotes the error of the neural network estimation of the true regression. To account for model uncertainties, model uncertainty intervals (MUIs) can be used to quantify the uncertainty between the neural network estimation and the true regression  $P(y(\mathbf{x}_i) | \hat{y}(\mathbf{x}_i))$ . In contrast, prediction intervals aim to quantify the uncertainty associated with the difference between the measured values  $t_i$  and the predicted values  $\hat{y}(\mathbf{x}_i)$ , i.e.  $P(t_i | \hat{y}(\mathbf{x}_i))$ . Accordingly, PIs will be wider than MUIs and will enclose them.

Assuming two error components in (5.12) are statistically independent, the variance of the total prediction errors  $\sigma_t^2$  can be mathematically obtained based on the variance of model uncertainty  $\sigma_y^2$  and the variance of data noise  $\sigma_\varepsilon^2$ ,

$$\sigma_t^2(\mathbf{x}_i) = \sigma_y^2(\mathbf{x}_i) + \sigma_\varepsilon^2(\mathbf{x}_i) \quad (5.13)$$

Given a real process, an  $(1-\alpha)$  confidence level PI of the measured target  $t_i$  is a stochastic interval  $I_t^\alpha(\mathbf{x}_i)$  expressed in (5.14), such that the coverage rate  $P(t_i \in I_t^\alpha) = 100(1-\alpha)\%$ .

$$I_t^\alpha(\mathbf{x}_i) = [L_t^\alpha(\mathbf{x}_i), U_t^\alpha(\mathbf{x}_i)] \quad (5.14)$$

where the lower bound  $L_t^\alpha(\mathbf{x}_i)$  and the upper bound  $U_t^\alpha(\mathbf{x}_i)$  can be obtained by

$$L_t^\alpha(\mathbf{x}_i) = \hat{y}(\mathbf{x}_i) - z_{1-\alpha/2} \sqrt{\sigma_t^2(\mathbf{x}_i)} \quad (5.15)$$

$$U_t^\alpha(\mathbf{x}_i) = \hat{y}(\mathbf{x}_i) + z_{1-\alpha/2} \sqrt{\sigma_i^2(\mathbf{x}_i)} \quad (5.16)$$

where  $z_{1-\alpha/2}$  is the critical value of the standard normal distribution, which depends on the desired confidence level  $(1-\alpha)$ . When the bounds of  $I_t^\alpha(\mathbf{x}_i)$  go beyond the unit interval  $[0, 1]$ , they should be adjusted to the corresponding lower or upper constraint bounds to ensure the constructed wind power PIs within the capacity range, and the corresponding probability mass is added to the adjusted bounds.

## 5.4 Prediction Intervals Construction

In this section, the prediction intervals for ELM forecasting are developed based on the bootstrap method. Several bootstrap methods have been compared to identify the most suitable one for PIs construction of ELM based wind power forecasting.

### 5.4.1 Bootstrap Methods

Bootstrap is regarded as a general approach of statistical inference based on building a sampling distribution by uniform sampling with replacements from the original data [209, 210]. It is widely applied as a robust alternative to the statistical inference based on the parametric assumptions, which can be unreliable and even impossible due to the sophistications involved in computing the standard errors in some conditions.

Three different bootstrap algorithms can be applied for regression analysis [111, 209], including the pairs bootstrap, the standard residuals bootstrap and the wild residuals bootstrap (wild bootstrap). The pairs bootstrap can be applied according to the algorithm shown as the following steps:

- Step 1 Obtain the training samples  $\{(\mathbf{x}_i, t_i)\}_{i=1}^N$ .
- Step 2 Generate bootstrapped pairs  $\{(\mathbf{x}_i^*, t_i^*)\}_{i=1}^N$  by uniform sampling with replacement from the original training data  $\{(\mathbf{x}_i, t_i)\}_{i=1}^N$ .
- Step 3 Estimate the ELM  $\hat{y}_i(\mathbf{x}_i^*)$  from the  $l^{\text{th}}$  bootstrapped dataset  $\{(\mathbf{x}_i^*, t_i^*)\}_{i=1}^N$ .
- Step 4 Repeat Steps 2-3 to obtain  $B$  bootstrap replicates.

The two other methods differ from the pairs bootstrap method mainly in sampling the residuals of which details can be found in [209]. These bootstrap methods have been implemented in our case study to identify the best one for ELM based wind power forecast.

For the three bootstrap approaches, when training an ELM on particular bootstrap samples, the model parameters are estimated in order to minimize the errors on the training data. Based on the  $B$  bootstrap replicates, we can train and obtain  $B$  ELMs ready for wind power forecasting.

#### 5.4.2 Model Uncertainty Variance

The MUIs can quantify the confidence in the network estimation  $\hat{y}(\mathbf{x}_i)$  for the true regression  $y(\mathbf{x}_i)$ . The bootstrap-based approach assumes that an ensemble of NN models will reach a relatively less biased approximation of true regression of the measured targets.

$$D_t = \{(\mathbf{x}_i, \mathbf{t}_i)\}_{i=1}^N \quad (5.17)$$

The  $B_M$  training data sets are resampled from the original training data with replacement. The average output of the ensemble of  $B_M$  ELMs is taken as the estimation of the true regression, expressed as

$$\hat{y}(\mathbf{x}_i) = \frac{1}{B_M} \sum_{l=1}^{B_M} \hat{y}_l(\mathbf{x}_i) \quad (5.18)$$

where  $\hat{y}_l(\mathbf{x}_i)$  is the prediction value of the input samples generated by the  $l$ th bootstrapped ELM.

The variance of model misspecification uncertainty  $P(y(\mathbf{x}_i) | \hat{y}(\mathbf{x}_i))$  can be estimated from the variance in the outputs of the trained  $B_M$  ELMs

$$\sigma_{\hat{y}}^2(\mathbf{x}_i) = \frac{1}{B_M - 1} \sum_{l=1}^{B_M} (\hat{y}_l(\mathbf{x}_i) - \hat{y}(\mathbf{x}_i))^2 \quad (5.19)$$

Following the bootstrap procedures, MUIs of ELM forecasts,  $I_M^\alpha(\mathbf{x}_i) = [L_M^\alpha(\mathbf{x}_i), U_M^\alpha(\mathbf{x}_i)]$ , can be obtained through

$$L_M^\alpha(\mathbf{x}_i) = \hat{y}(\mathbf{x}_i) - z_{1-\alpha/2} \sqrt{\sigma_{\hat{y}}^2(\mathbf{x}_i)} \quad (5.20)$$

$$U_M^\alpha(\mathbf{x}_i) = \hat{y}(\mathbf{x}_i) + z_{1-\alpha/2} \sqrt{\sigma_y^2(\mathbf{x}_i)} \quad (5.21)$$

### 5.4.3 Data Noise Variance

In addition to the model uncertainty of ELM forecasting, the uncertainty caused by the data noise is analyzed in this section. Due to the heteroscedasticity, with only one observation of wind power at each time point, it is challenging to estimate the data uncertainty. According to the variance definition in [211], the variance of the measured target  $t_i$  conditioned on the input variables  $\mathbf{x}_i$ , can be calculated from

$$\sigma_\varepsilon^2(t_i | \mathbf{x}_i) = E[(t_i - E[t_i | \mathbf{x}_i])^2 | \mathbf{x}_i] \quad (5.22)$$

Given the training data  $\{(\mathbf{x}_i, t_i)\}_{i=1}^N$ , as can be seen from (5.11), the outputs of ELMs produce averaged values of the targets conditioned on input variables  $\mathbf{x}_i$ , i.e.  $\hat{y}(\mathbf{x}_i) = E[t_i | \mathbf{x}_i]$ . Then the values of  $E[t_i | \mathbf{x}_i]$  in (5.22) can be derived based on the trained ELM.

Keeping the input  $\mathbf{x}_i$  and replacing the targets  $t_i$  with  $(\hat{y}(\mathbf{x}_i) - t_i)^2$ , we can obtain the transformed training dataset

$$D_\varepsilon = \left\{ \left( \mathbf{x}_i, (\hat{y}(\mathbf{x}_i) - t_i)^2 \right) \right\}_{i=1}^N \quad (5.23)$$

The objective variance can be estimated by training a separate ELM  $h_\varepsilon(\mathbf{x}_i; \mathbf{w}, b, \beta)$ , mathematically expressed as

$$h_\varepsilon(\mathbf{x}_i; \mathbf{w}, b, \beta) = (\hat{y}(\mathbf{x}_i) - t_i)^2, i = 1, \dots, N. \quad (5.24)$$

The output of the trained ELM,  $\hat{r}(\mathbf{x}_i)$ , can be represented as

$$\hat{r}(\mathbf{x}_i) = E[(\hat{y}(\mathbf{x}_i) - t_i)^2 | \mathbf{x}_i] \quad (5.25)$$

The model uncertainty associated with ELM  $h_\varepsilon(\mathbf{x}_i; \mathbf{w}, b, \beta)$ , represented by  $\sigma_r^2(\mathbf{x}_i)$ , also should be taken into account. It can be calculated through the bootstrap-based method similar to the procedure of deriving the model uncertainty variance. Supposing  $B_N$  bootstrap replicates are implemented, we can

obtain the estimated noise variance  $\hat{\sigma}_\varepsilon^2(t_i | \mathbf{x}_i)$  and the variance of regression model uncertainty  $\sigma_r^2(\mathbf{x}_i)$  respectively,

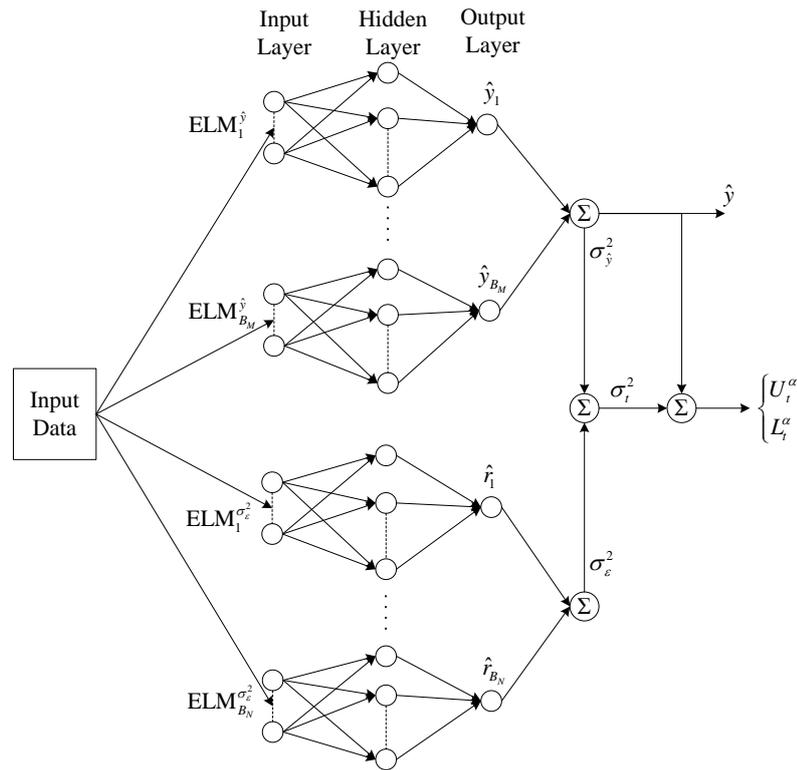
$$\hat{\sigma}_\varepsilon^2(t_i | \mathbf{x}_i) = \hat{r}(\mathbf{x}_i) = \frac{1}{B_N} \sum_{l=1}^{B_N} \hat{r}_l(\mathbf{x}_i) \quad (5.26)$$

$$\sigma_r^2(\mathbf{x}_i) = \frac{1}{B_N - 1} \sum_{l=1}^{B_N} (\hat{r}_l(\mathbf{x}_i) - \hat{r}(\mathbf{x}_i))^2 \quad (5.27)$$

The variance of data noise can be obtained through

$$\sigma_\varepsilon^2(\mathbf{x}_i) = \hat{\sigma}_\varepsilon^2(t_i | \mathbf{x}_i) + \sigma_r^2(\mathbf{x}_i) \quad (5.28)$$

With the model uncertainty variance and data noise variance, the total variance of the prediction intervals can be obtained based on (5.13).



**Figure 5.2** The framework for PIs construction of the proposed BELM approach.

For PIs construction of ELM forecasting using the proposed algorithm,  $B_M + B_N$  ELM models are required in total. The overall framework for the proposed

bootstrap based approach for ELM probabilistic forecasting is explicitly displayed in Figure 5.2. If traditional NNs are used, intensive computational efforts are required since the bootstrap based forecasting approach involves a great number of bootstrap replicates. With the extremely fast learning speed, the proposed BELM approach can effectively and efficiently provide the probabilistic forecasting for wind power production.

## **5.5 Numerical Studies**

### **5.5.1 Description of Experiment Data**

In the study, the proposed BELM approach has been tested using the wind power data from Cathedral Rocks wind farm, South Australia. The wind farm has nominal generation capacity  $P_c$  of 66MW combined with 33 wind turbines of 2MW. The wind power data with one hour temporal resolution from Jun. 2008 until Jun. 2012 are used for the case study.

Operational planning and scheduling in modern power systems with wind power requires the forecasts of the future wind power generation according to the planning horizons. Generally, wind power forecasting can be divided into four categories of different timeframes: very short-term, short-term, medium-term and long-term forecasts [125]. Short-term and very short-term forecasts are important because of their significances to both generation and reserve dispatches and etc. As introduced in [33], statistical methods would outperform NWP-based methods for forecasting wind power with look-ahead times less than a few hours. Though having external NWP information, statistical models using historical measurements only should be preferred for such short look-ahead times [33]. Therefore, the proposed BELM approach takes only historical wind power data as inputs for hourly ahead forecasting, which is essential for e.g. dispatching ancillary service market in practice. Other data such as the weather information can be easily included in our future work.

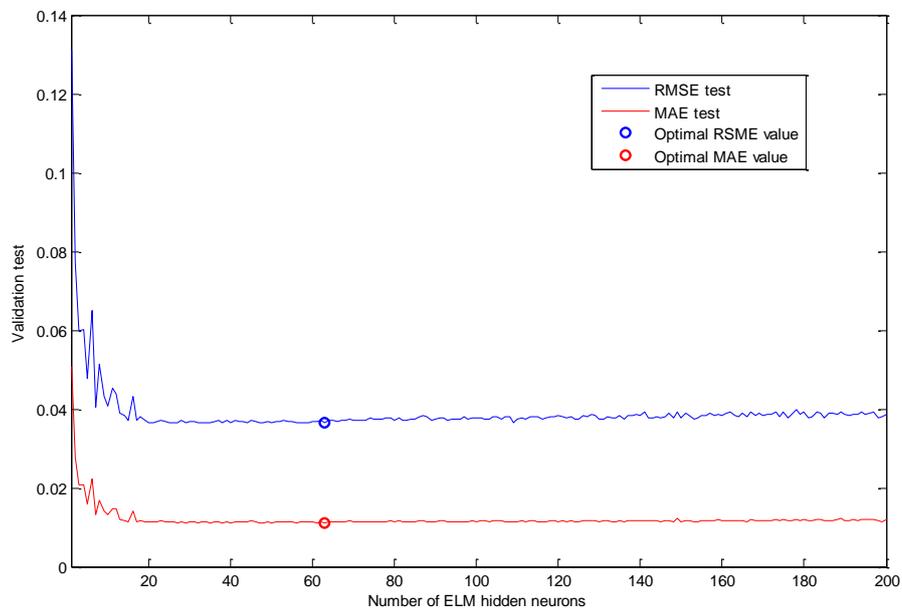
### **5.5.2 Determination of ELM Hidden Nodes Number**

The extreme learning machine is based on SLFNs, of which the number of hidden layer neurons for ELM models need to be determined properly. Since different forecasts may have different needs and properties, optimization of the

ELM structure is necessary and critical to minimize the uncertainties due to model misspecifications and ensure the efficiency simultaneously. The hidden nodes number of ELMs is determined based on the cross-validation approach [203]. The ELMs' generalization performance of different structures over the validation dataset is assessed by both root-mean-square error (RMSE) and mean absolute error (MAE),

$$MAE = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} |t_i - \hat{y}(\mathbf{x}_i)| \quad (5.29)$$

$$RMSE = \sqrt{\frac{1}{N_{test}} \sum_{i=1}^{N_{test}} [t_i - \hat{y}(\mathbf{x}_i)]^2} \quad (5.30)$$



**Figure 5.3** Validation test for ELMs with different numbers of hidden neurons.

Figure 5.3 shows the results of validation test for hourly ahead forecasting based on the actual wind farm data. It can be seen that ELMs will have stable generalization performance once the hidden nodes exceed a certain threshold. The ELM with 63 hidden neurons can sufficiently ensure the optimal MAE and RMSE simultaneously.

### 5.5.3 Comparison of Bootstrap Methods

The MUIs of ELM regression are approximated based on the bootstrap methods. The commonly applied bootstrap methods, including pairs bootstrap, standard residuals bootstrap and wild bootstrap, can provide different performances for different applications. The performances of the different bootstrap approaches used to construct the PIs of wind power forecasting with ELMs are compared. ECP with corresponding NCP, ACE and interval score of different bootstrap approaches are given in Table 5.1.

**Table 5.1** Evaluation results of different bootstrap methods

Method	NCP	ECP	ACE	Score
Pairs bootstrap	90%	93.59%	3.59%	-6.02%
	95%	96.17%	1.17%	-3.67%
	99%	98.25%	-0.75%	-1.17%
Standard residuals bootstrap	90%	92.85%	2.85%	-6.11%
	95%	95.92%	0.92%	-3.72%
	99%	98.09%	-0.91%	-1.21%
Wild bootstrap	90%	93.43%	3.43%	-6.16%
	95%	95.92%	0.92%	-3.78%
	99%	97.67%	-1.33%	-1.25%

From Table 5.1, it can be found that the pairs bootstrap provides the most reliable PIs of the measured wind power. The residuals based bootstrap relies on the errors that are representative of the true model errors. However, the nonlinear relationship for prediction is always unknown, and the model misspecification is unavoidable. If the model is either misspecified or overfitted, the pairs based bootstrap approach can be more robust [111]. As expected, the pairs bootstrap method outperforms the other two methods in the tests using the chaotic wind power data. Based on the comparisons, the pair bootstrap is applied for the proposed BELM approach.

### 5.5.4 Analysis of Forecasting Results

High complexity of chaotic climate systems contributes to high level of uncertainties in wind power generation. The patterns of weather conditions and wind speeds vary very much in different seasons. To examine the effectiveness and applicability of the proposed approach, the four seasons in Australia, summer (December to February.), autumn (March to May), winter (June to

August), and spring (September to November) are considered respectively. Models are separately constructed for different seasons.

Considering the seasonal difference and diversity, the proposed BELM method is tested using wind power data in summer 2012, autumn 2011, winter 2011 and spring 2010. The wind power data before these test dates are used as the training data. These datasets are normalized with respect to the capacity of Cathedral Rocks wind farm before applying to the proposed models. To evaluate the proposed approach, the climatology and the persistence approaches are used for benchmarking the forecasting performance [109, 113, 173]. Climatology predictive distribution is formed based on all available wind power observations, and is a unique and unconditional probabilistic prediction. The climatology is relatively easy to outperform for the short-term probabilistic wind power forecasting. In deterministic point forecasts of wind power generation, the persistence method is considered as the most common benchmark and difficult to outperform for short look-ahead time forecasting. The persistence based probabilistic forecast model is used as a benchmark for comparisons in the study as well. Its mean is given by the last available power measurement, and the variance is computed using the latest observations. Both the climatology and persistence methods are relatively simple. To benchmark the proposed BELM approach, an advanced model, the ESM method, proposed in [176] is applied in the study. In addition, to evaluate the impacts of forecasting error distribution model, the proposed method has also been tested using the Beta distribution for forecasting error modeling, termed as BELM-Beta [162].

The major objective of the proposed BELM method is to derive reliable PIs. Furthermore, power system operation requires useful information with high confidence levels. Therefore it should be more practically meaningful to obtain high confidence level PIs to fulfill the needs of power system operation. Different levels of NCP ( $1-\alpha$ ) ranging from 90%-99% are considered in the study. For the PIs quality test, corresponding ECPs, ACEs and interval score are of different approaches in four seasons are given in Tables 5.2-5.5 respectively.

As seen in Tables 5.2-5.5, in all four seasons, the proposed method outperforms other approaches with the resultant ECPs consistently closer to the corresponding nominal confidence levels. All ACEs of the proposed method are

close to zero, especially for the higher confidence levels 95% and 99%, which indicates the high reliability of the constructed PIs. E.g. in autumn, the proposed method has absolute ACEs at confidences 95% and 99% around 1%, smaller than the other four benchmarks. Particularly, in summer, PIs obtained by the Beta distribution model have similar reliability with the censored normal distribution model of the proposed BELM approach and demonstrate much better quality than in other seasons, which means that the Beta distribution modeling is much more proper to summer than to other three seasons.

**Table 5.2** Results of PIs in summer 2012

NCP	Method	ECP	ACE	Score
90%	BELM	92.46%	2.46%	-7.61%
	BELM-Beta	89.40%	-0.60%	-8.39%
	Persistence	86.93%	-3.07%	-8.78%
	Climatology	98.01%	8.01%	-16.08%
	ESM	88.55%	-1.45%	-8.54%
95%	BELM	95.09%	0.09%	-4.59%
	BELM-Beta	92.46%	-2.54%	-5.37%
	Persistence	90.70%	-4.30%	-5.49%
	Climatology	98.65%	3.65%	-8.50%
	ESM	92.09%	-2.91%	-5.25%
99%	BELM	98.08%	-0.92%	-1.41%
	BELM-Beta	95.30%	-3.70%	-2.21%
	Persistence	95.35%	-3.65%	-1.96%
	Climatology	99.72%	0.72%	-1.81%
	ESM	96.64%	-2.36%	-1.76%

**Table 5.3** Results of PIs in autumn 2011

NCP	Method	ECP	ACE	Score
90%	BELM	93.01%	3.01%	-5.92%
	BELM-Beta	72.88%	-17.12%	-5.97%
	Persistence	83.99%	-6.01%	-6.50%
	Climatology	95.84%	5.84%	-15.10%
	ESM	88.98%	-1.02%	-6.36%
95%	BELM	95.59%	0.59%	-3.64%
	BELM-Beta	76.46%	-18.54%	-3.61%
	Persistence	88.36%	-6.64%	-4.03%
	Climatology	98.59%	3.59%	-7.85%
	ESM	92.90%	-2.10%	-3.89%
99%	BELM	97.67%	-1.33%	-1.19%
	BELM-Beta	79.45%	-19.55%	-1.09%
	Persistence	92.36%	-6.64%	-1.46%
	Climatology	100%	1.00%	-1.71%
	ESM	97.04%	-1.96%	-1.30%

**Table 5.4** Results of PIs in winter 2011

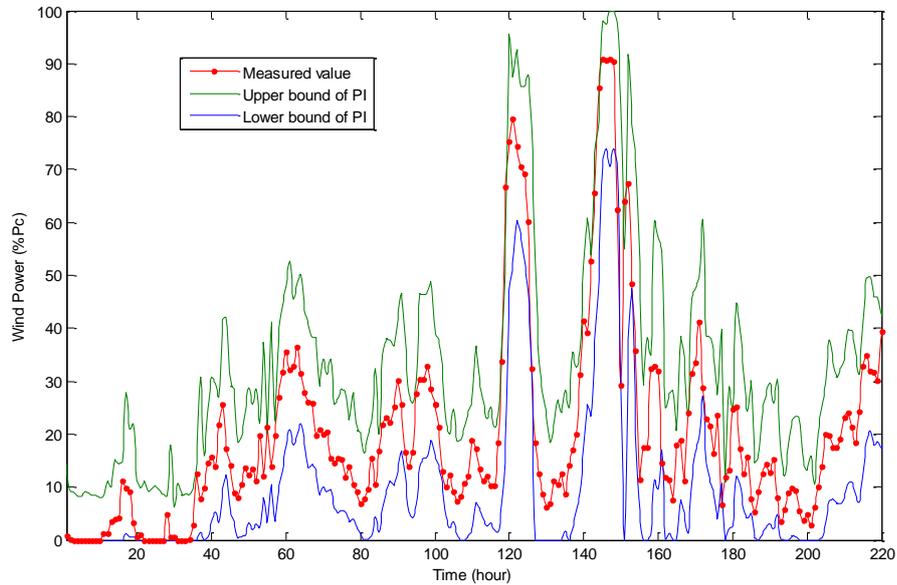
NCP	Method	ECP	ACE	Score
90%	BELM	91.30%	1.30%	-6.91%
	BELM-Beta	80.22%	-9.78%	-6.85%
	Persistence	86.83%	-3.17%	-7.60%
	Climatology	96.34%	6.34%	-15.66%
	ESM	88.97%	-1.03%	-7.36%
95%	BELM	94.41%	-0.59%	-4.12%
	BELM-Beta	83.61%	-11.39%	-4.06%
	Persistence	90.88%	-4.12%	-4.76%
	Climatology	99.45%	4.45%	-8.19%
	ESM	92.94%	-2.06%	-4.47%
99%	BELM	97.89%	-1.11%	-1.28%
	BELM-Beta	87.45%	-11.55%	-1.18%
	Persistence	95.30%	-3.70%	-1.65%
	Climatology	100%	1.00%	-1.78%
	ESM	96.99%	-2.01%	-1.40%

**Table 5.5** Results of PIs in spring 2010

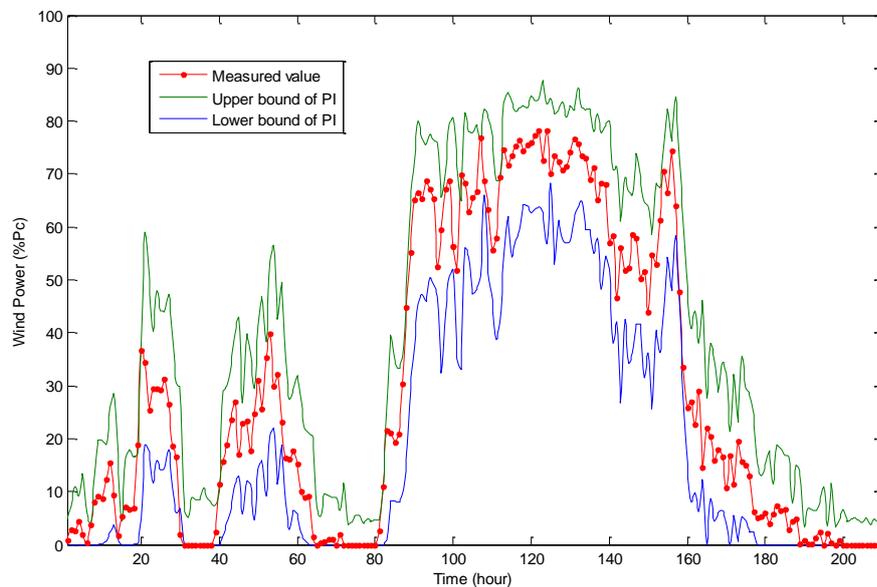
NCP	Method	ECP	ACE	Score
90%	BELM	93.19%	3.19%	-7.10%
	BELM-Beta	84.03%	-5.97%	-7.19%
	Persistence	86.52%	-3.48%	-7.95%
	Climatology	95.46%	5.46%	-16.31%
	ESM	88.36%	-1.64%	-7.71%
95%	BELM	96.12%	1.12%	-4.15%
	BELM-Beta	86.39%	-8.61%	-4.16%
	Persistence	90.46%	-4.54%	-4.96%
	Climatology	97.54%	2.54%	-8.55%
	ESM	91.89%	-3.11%	-4.72%
99%	BELM	98.68%	-0.32%	-1.12%
	BELM-Beta	89.79%	-9.21%	-1.09%
	Persistence	95.00%	-4.00%	-1.75%
	Climatology	99.62%	0.62%	-1.81%
	ESM	96.24%	-2.76%	-1.55%

According to Tables 5.2-5.5, interval scores of the proposed approach are larger than the climatology, the persistence, and the ESM methods, indicating the proposed BELM approach outperforms these three benchmarks from the perspective of sharpness and overall skill. In addition, the proposed approach also can have similar or higher skill than the approach with Beta distribution based error modeling in some cases. It can be proved that the average interval score of the proposed method still outperform the Beta distribution considering

all the four seasons. Considering the reliability and overall skill, the proposed approach shows much better results in terms of comprehensive performance than the four benchmarks.



**Figure 5.4** PIs with NCP 90% in summer 2012 obtained by the proposed BELM approach.

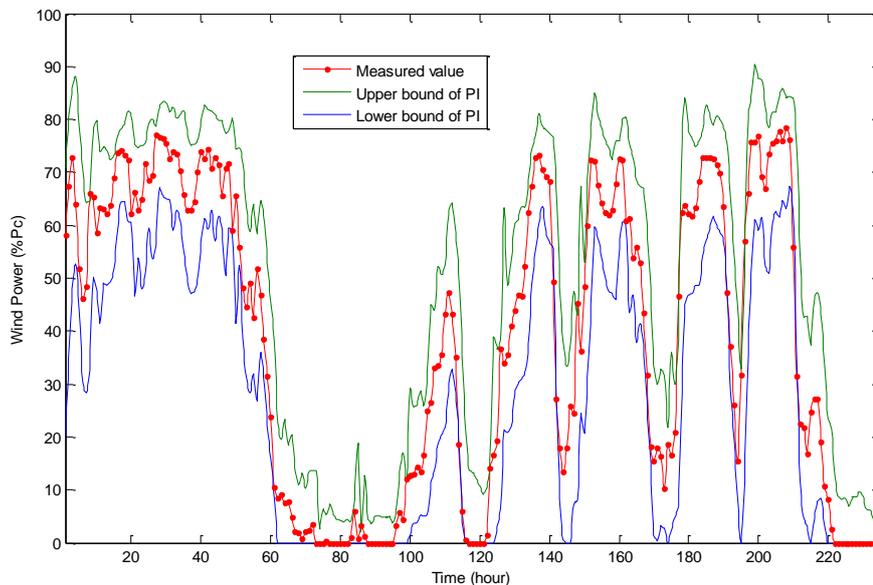


**Figure 5.5** PIs with NCP 90% in autumn 2011 obtained by the proposed BELM approach.

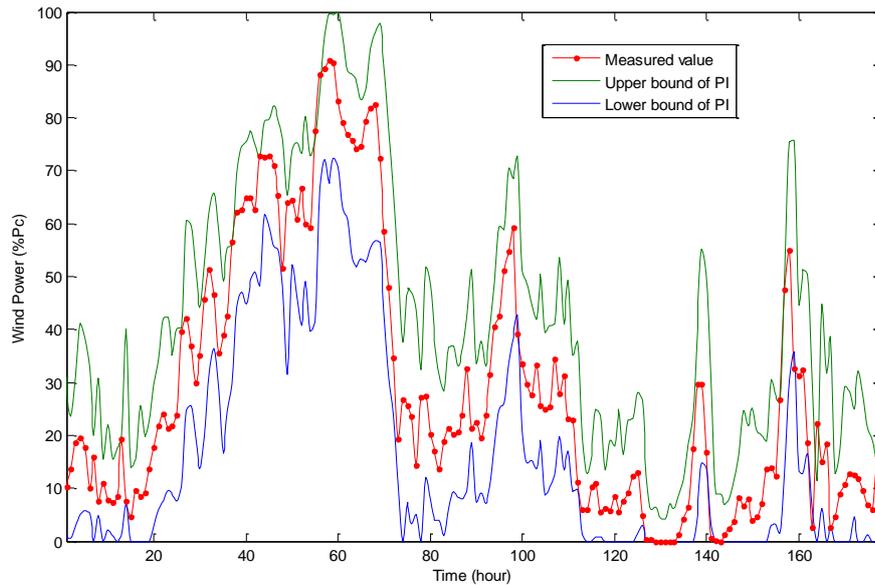
This is not unreasonable as Beta distribution can well reflect the long-term statistics of wind power forecasting errors [162], but it is unable to reflect

seasonal variations in detail. In addition, the Beta distribution can model the forecasting errors well given different levels of wind power outputs, which is not the case in our approach, where the forecasting errors are modeled statistically dependent of the input variables. Therefore the Beta distribution based approach provides less satisfactory results at high confidence levels of 90% and above. Furthermore, in the autumn season, the Beta distribution based approach gives very poor results.

The climatology is a simple unconditional prediction approach and does not consider the heteroscedasticity of wind power data. Therefore large widths of PIs are resulted at high confidence levels which are barely useful in practice. The ECP for persistence approach varies significantly in different seasons, indicating the significant seasonal variations of wind power. Due to the simple mapping, the persistence cannot obtain sufficiently satisfactory PIs. In comparisons, the ESM approach has fair results with respect to both reliability and sharpness. From Tables 5.2-5.5, it can be seen that in summer PIs reliability of the proposed method are slightly lower than the rest seasons. This is understandable as weather conditions in summer are relatively more chaotic.



**Figure 5.6** PIs with NCP 90% in winter 2011 obtained by the proposed BELM approach.



**Figure 5.7** PIs with NCP 90% in spring 2010 obtained by the proposed BELM approach.

The 90% confidence PIs obtained by the proposed BELM method and the actually measured wind power in the four seasons are visually displayed in Figures 5.4-5.7 respectively. For all four seasons, the measured wind power data are perfectly enclosed by the PIs generated by the proposed method, indicating an excellent performance that can fulfill the needs of power system operation. These graphs also clearly demonstrate the non-stationary characteristics of wind power series. Notably, some PIs can have abnormal values out of capacity range of the wind farm. Therefore resultant predictive densities have been censored to concentrate probability mass outside the interval on the bounds [113].

To investigate the influence of the bootstrap replicates number on the resulted PIs, the proposed method is further tested on the wind power data in autumn 2011. Each test with given bootstrap replicates is conducted for 100 repetitive times using a PC with Intel Core Duo 3.16GHz CPU and 4GB RAM. The mean ECP (MECP), the standard deviation of ECP (SDECP) and the needed training time with different bootstrap replicates are given in Table 5.6. With bootstrap replicates varying between 20 and 1000, the BELM method can produce reliable PIs. Considering both the accuracy and efficiency, the 200 bootstrap replicates for generating PIs are considered the best option for the case study. Although the size of the training data set is not small, the total time needed for ELMs training using the proposed BELM approach only accounts for about 30 seconds,

indicating a significantly high efficiency and potential for online application. On the contrary, training traditional BP NNs for hourly ahead wind power forecasting using the similar size of data can take thousands of times longer. The extremely fast model construction should benefit practical applications from several aspects under the precondition of ensuring satisfactory performance. Foremost, it saves the efforts in offline model construction which could otherwise be much time-consuming and computationally intensive. It should not be unreasonable that the characteristics of wind power series could be changed continuously or suddenly, similarly as the chaotic weather systems. Therefore, continuous online model updating can be significantly meaningful to maintain and improve the forecasting performance as far as possible, especially for the very short-term forecasting.

**Table 5.6** Reliability and efficiency of different bootstrap replicates

Bootstrap Replicates	NCP	MECP	SDECP	Training Time (s)
20	90%	92.09%	0.38%	3.26
	95%	95.18%	0.28%	
50	90%	92.09%	0.28%	7.89
	95%	95.17%	0.21%	
100	90%	92.04%	0.24%	15.78
	95%	95.17%	0.18%	
200	90%	92.03%	0.18%	31.43
	95%	95.14%	0.16%	
300	90%	92.03%	0.18%	47.16
	95%	95.17%	0.14%	
500	90%	92.02%	0.17%	78.15
	95%	95.17%	0.15%	
1000	90%	92.01%	0.17%	166.78
	95%	95.16%	0.14%	

Traditionally, the model uncertainty is always ignored in point forecasting [198]. Further, there have existed theoretical results for uncertainty estimation for the normal linear regression [212]. However, the linear regression system of ELM is based on the randomly assigned input weights and biases. As introduced in Section 5.3.1, the randomly generated input weights and biases are also one source of the model uncertainty of NNs. To comprehensively involve the model misspecifications and improve the forecasts accuracy, the bootstrap is applied for the proposed ELM based PI construction approach. To investigate the influence

of the bootstrapping, the proposed approach with (200 replicates) or without bootstrapping are conducted for 100 times to obtain the mean forecasting reliability measured by ECP and ACE, and mean sharpness measured by interval score. The test results are shown in Table 5.7. The proposed approach without bootstrapping just uses ELM for mean and variance regression (MVR), termed as MVR-ELM here.

**Table 5.7** Effects of bootstrapping on resultant PIs

NCP	Method	ECP	ACE	Score
90%	BELM	92.03%	2.03%	-6.20%
	MVR-ELM	88.56%	-1.44%	-6.13%
95%	BELM	95.14%	0.14%	-3.97%
	MVR-ELM	91.00%	-4.00%	-3.94%
99%	BELM	97.75%	-1.25%	-1.33%
	MVR-ELM	93.56%	-5.44%	-1.38%

We can find that the model uncertainty does have observable impacts on the resultant PIs from Table 5.7. Though similar sharpness can be obtained, ECP can be reduced by more than 4%, if the model uncertainty is not considered by the application of bootstrap in PI formulation. Similarly, degradations of ACE due to no bootstrapping involvements are also observed in Table 5.7. It echoes the descriptions in Section 5.3.1 that the model uncertainty is one indispensable aspect of uncertainty sources for the proposed BELM-based forecasting, and should be considered to form reliable and accurate PIs.

The comprehensive numerical studies have indicated the effectiveness of the proposed BELM approach. Actually, the multi-step forecasts with look-ahead times two and three hours have been implemented using the proposed approach, and satisfactory PIs can be obtained. In addition to wind power forecasts with hourly resolution, the intra-hour prediction results are also highly concerned by TSO and wind farm controller. The higher resolution wind power e.g. 10-min measures are very crucial to wind farm control, continuous generation and reserve dispatch and etc., and would have higher volatility than hourly data. Practically, in Denmark the 10-min lead time is regarded as the most important very short-term horizon by the TSO since power fluctuations at this time horizon have the most serious impacts on the balance of power systems [213]. We study

10-min resolution forecasting with different look-ahead horizons including 10 min, 30 min, and 1h in autumn 2009 of Cathedral Rocks wind farm. The resultant PIs with NCPs 90% and 95% are obtained and evaluated respectively, given in Table 5.8.

**Table 5.8** Results of multi-step intra-hour forecasting

Horizon	Method	NCP 90%			NCP 95%		
		ECP	ACE	Score	ECP	ACE	Score
10 min	BELM	91.51%	1.51%	-3.10%	95.72%	0.72%	-1.99%
	BELM-Beta	71.89%	-18.11%	-3.11%	75.14%	-19.86%	-2.01%
	Persistence	83.83%	-6.17%	-3.38%	87.25%	-7.75%	-2.17%
	Climatology	96.94%	6.94%	-14.83%	98.50%	3.50%	-7.75%
	ESM	90.48%	0.48%	-3.38%	93.72%	-1.28%	-2.12%
30 min	BELM	91.69%	1.69%	-5.33%	94.72%	-0.28%	-3.27%
	BELM-Beta	70.89%	-19.11%	-5.38%	73.70%	-21.30%	-3.32%
	Persistence	83.06%	-6.94%	-5.94%	86.90%	-8.10%	-3.73%
	Climatology	96.94%	6.94%	-14.83%	98.50%	3.50%	-7.75%
	ESM	90.09%	0.09%	-5.86%	93.47%	-1.53%	-3.61%
1 h	BELM	90.54%	0.54%	-6.81%	93.82%	-1.18%	-4.09%
	BELM-Beta	70.21%	-19.79%	-6.86%	73.45%	-21.55%	-4.11%
	Persistence	82.44%	-7.56%	-7.52%	86.45%	-8.55%	-4.66%
	Climatology	96.94%	6.94%	-14.83%	98.50%	3.50%	-7.75%
	ESM	89.47%	-0.53%	-7.38%	93.31%	-1.69%	-4.45%

From Table 5.8, it can be found that the proposed BELM has superior performance than the ESM method and other benchmarks for intra-hour wind power forecasting. The persistence approach demonstrates relatively lower reliability than that in hourly forecasting, indicating the higher violability of wind power with 10-min resolution. Comparing with the ESM approach, which is a well-established time series model for short-term probabilistic forecasting of wind power, the proposed BELM approach has high flexibility due to the non-linear mapping capability of ELM. With successful application to short term probabilistic wind power forecasting in this chapter, the proposed method can perform longer term forecasting by including NWP information as additional inputs to ensure the performance. In practice, the system-level aggregated wind power is also highly concerned by the TSO. Due to the flexibility, the proposed BELM approach provides a generalized framework for probabilistic wind power forecasting. Therefore, local NWP and historical wind power of individual wind farms can be taken as the inputs to the proposed model to forecast the aggregated

wind generation involving the farm-level information. With the fast speed and high flexibility, the proposed model can provide an online tool to facilitate various decision making activities by TSO and generation companies to determine the needed reserve, design proper bidding strategies against risks, etc.

## **5.6 Conclusions**

Wind power forecasting is critical to power system operation, planning and control. However, wind power forecasting errors are unavoidable to some extent due to the nonlinear and stochastic nature of the weather system. Traditional neural network based forecasting models cannot provide satisfactory performances with respect to both accuracy and computing time needed. In this chapter, extreme learning machine is successfully applied for probabilistic interval forecasting of wind power. A novel statistical approach BELM is developed to construct the PIs of ELM based regression. Accurate PIs can be obtained by combining the variances of regression model uncertainty and residual noise. Different bootstrap methods have been compared and analyzed to select the best one for the developed forecasting model. The influence of bootstrap replicates on the efficiency and the quality of the constructed PIs has also been carefully investigated. Further, the effect of the model uncertainty (bootstrapping process) on resultant PIs is examined in the study, verifying its indispensability. Because of the extremely fast learning, the training of the proposed BELM forecasting method can be extremely faster than traditional NNs based approaches, demonstrating a high potential for online application. Comprehensive experiments using practical wind farm data of different seasons have demonstrated the highly satisfactory results, which indicate that efficient and accurate short term wind power forecasts can be achieved using the proposed BELM method. With fast speed, high reliability and high flexibility, the proposed BELM approach is a generalized framework for probabilistic forecasting of wind power and can provide an efficient and meaningful online tool for power system applications including probabilistic reserve determination, generation dispatch, market trading, etc.



# **6** *Direct Interval Forecasting of Wind Power Generation*

## **6.1 Introduction**

In Chapter 5, the BELM approach is developed to estimate the predictive distribution of wind power with extremely fast speed and effective performance. The predictive PDF of the proposed BELM method is modeled with censored normal distribution. Classical NNs based PIs construction methods always assume that prediction errors are normally distributed [110, 116, 199, 200]. These normal distribution based approaches may not guarantee satisfactory results, especially when the noise is heterogeneous [117]. In practice, the probability distribution of prediction errors of realistic processes can be complicated, such as the wind power generation in this study [159]. Many attentions have been paid to the study of probabilistic wind power forecasts with non-parametric predictive PDF as the high complexity of the statistics of wind power prediction error. Different approaches have been proposed for driving quantiles of wind power, including meteorological ensembles [181, 192], quantile regression [112, 165], adaptive resampling approach [113], kernel density estimation [169, 171], radial basis function [173], etc.

In this chapter, a direct interval forecasting (DIF) approach is newly developed to produce prediction intervals of wind power generation based on ELM [203] and particle swarm optimization (PSO) [214, 215]. The proposed DIF method aims to obtain optimal PIs without the prior knowledge, statistical inference or distribution assumption of forecasting errors required in most traditional approaches. ELM applied in the proposed approach is a novel learning algorithm

proposed for SLFNs featuring extremely fast learning speed and superior generalization capability. ELM successfully avoids the limitations of traditional NNs learning algorithms, such as local minima, overtraining, high computation costs and so forth.

Traditional PIs construction methods for wind power rely on quantile analysis of point forecast errors with or without prior distribution assumptions [112, 113, 165, 169, 171, 173], where the procedures of PIs formulation and final performance assessment are usually separated. E.g. in [113], PIs can be achieved through a conditional probabilistic modeling between point forecast outputs and associated errors. In contrast, the proposed DIF approach integrates the two procedures holistically to formulate the PIs directly to pursue the best quality of resultant PIs, without the need of prior knowledge and distribution assumption of point forecasts errors. As investigated in [107] as early as 1970s, with a properly constructed cost function, PI estimation could be considered as a Bayesian decision-making procedure to acquire an optimal PI that minimizes the expected cost. Lower upper bound estimation (LUBE) method is proposed for non-parametric PI construction based on traditional NNs in [216], applied in load forecasting [217] and wind power forecasting [174, 178]. However, traditional NNs employed in the LUBE method would cause several inevitable limitations, such as overtraining, high computation burden, and so forth. The coverage width-based criterion (CWC) score accounting for both the coverage probability and the average width of PIs. The CWC score cannot accurately measure the overall skill of constructed PIs, which has been mathematically proven in [175]. The CWC index gives more weights to the ECP for reliability due to exponential term for penalization. The extremely narrow PIs can outperform any constructed PIs when the CWC is used as the performance measure. The CWC-based cost function can mislead the construction of optimal PIs, resulting in wider or narrower PIs [175], since the CWC-base cost function gives biased weights to reliability and sharpness. The problems of the application of LUBE approach in probabilistic wind power forecasting have been discussed in [218]. The objective function of the developed DIF approach is specially formulated to address both the coverage probability and sharpness of PIs simultaneously, and is optimized through PSO featuring fast convergence and gradient-free optimization.

Furthermore, the proposed method is able to generate multiple optimal PIs of different confidence levels in one single optimization process.

Generally, different decision-makers in power systems have different look-ahead time preferences ranging from minutes to days for wind power forecasts according to their own operational requirements. Very short-term wind power prediction is needed to wind farm control [219, 220], the temporal operation of wind storage systems associated with temporal market regulations such as Australian National Electricity Market with 5 minutes resolution [114], and the TSO which aims to optimally dispatch reserves for the continuous balance of the power system [31, 127]. Hourly ahead forecast is crucial for power system and electricity market balance, e.g. Nord pool market [29]. Longer term forecasts up to days ahead are very meaningful for unit commitment [193], day-ahead market trading [132], etc.

The proposed DIF method has been tested using the practical data of two wind farms in Australia. Without loss of generality, in the case study we focus on the hourly forecast on an hourly basis though with extendibility. Comparing with benchmarks, the effectiveness of the proposed method has been proved through comprehensive evaluations with respect to both the reliability and overall skill of the forecasting results. By accurate quantification of the uncertainties of wind generation forecasts, the proposed interval forecasting approach has a high potential to support various operation and planning activities in power systems, such as to provide reliable information for dispatching e.g. the hourly Nord pool market. Particularly, the interval forecasting results can also be used to develop new operation and planning tools for TSO to probabilistically determine the needed reserves in advance [130, 190], and to facilitate Gencos' risk management through strategic bidding [132].

## 6.2 Formulation of PIs

PIs quantify the uncertainty associated with forecasts. Given a set of process pairs,

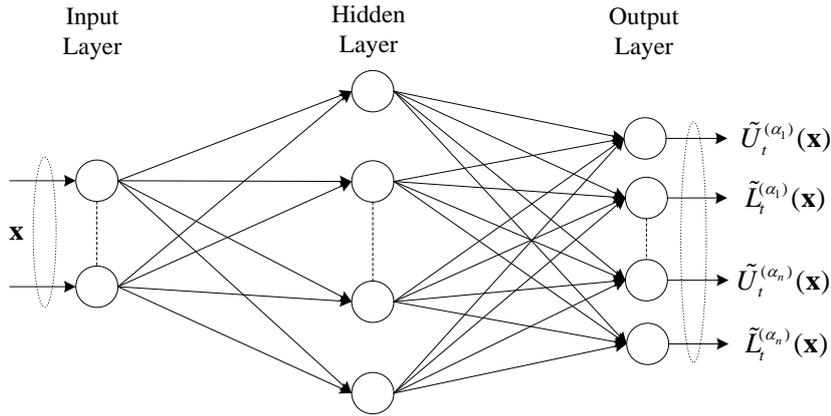
$$D_i = \{(\mathbf{x}_i, t_i)\}_{i=1}^N \quad (6.1)$$

where  $t_i$  is the future target to forecast, and  $\mathbf{x}_i$  denotes relevant input variables that can include historical wind power and wind speeds, numerical weather predictions and so on for wind power forecasting in the study. PI with nominal confidence  $(1-\alpha)$  of the future target  $t_i$ , represented as  $\tilde{I}_t^{(\alpha)}(\mathbf{x}_i)$ , can be expressed as the following equation

$$\tilde{I}_t^{(\alpha)}(\mathbf{x}_i) = [\tilde{L}_t^{(\alpha)}(\mathbf{x}_i), \tilde{U}_t^{(\alpha)}(\mathbf{x}_i)] \quad (6.2)$$

where  $\tilde{L}_t^{(\alpha)}(\mathbf{x}_i)$  and  $\tilde{U}_t^{(\alpha)}(\mathbf{x}_i)$  denote the lower and upper bounds of PI  $\tilde{I}_t^{(\alpha)}(\mathbf{x}_i)$  respectively, such that the future target  $t_i$  is expected to be enclosed by  $\tilde{I}_t^{(\alpha)}(\mathbf{x}_i)$  with coverage probability,

$$P(t_i \in \tilde{I}_t^{(\alpha)}(\mathbf{x}_i)) = 100(1-\alpha)\% \quad (6.3)$$



**Figure 6.1** The ELM model for PIs generation by the proposed DIF approach.

The proposed method aims to directly generate the lower and upper bounds of the expected PIs by ELM. It should be pointed out that the proposed method actually provides an unique framework capable of generating multiple pairs of PI bounds with different nominal coverage probabilities  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$  simultaneously through a single optimization approach. The overall structure of the proposed ELM model is shown in Figure 6.1, where the ELM takes the inputs and outputs the corresponding PI bounds of different confidence levels.

## 6.3 Optimal Construction of PIs

### 6.3.1 Objective Function

The proposed DIF method adopts an ELM to predict the PIs and pursues the optimal quality of produced PIs without statistical inferences and distribution assumptions for forecasting errors. Because of the unique properties of ELM described in Section 5.2, training the ELM based forecasters is equivalent to analytically determining the output weights alone. Comprehensive PIs evaluation criteria are well established and described in Section 3.3. To ensure the quality of produced PIs, ELM output weights are optimized to account for both reliability and sharpness of the generated PIs simultaneously, which can be considered as a multi-objective optimization problem (MOOP) [221].

A multi-objective function for training the proposed model is developed based on well-established PIs evaluation criteria introduced in the previous Section 3.3 to produce optimal PIs. It should be highlighted that though the interval score accounts for reliability and sharpness, it cannot quantitatively distinguish the contributions of the two aspects. However, the interval score can provide an evaluation from the perspective of sharpness given a prior analysis of reliability. Under the same nominal confidence and similar reliability, PIs with the smaller the absolute score  $\bar{S}_t^{(\alpha)}$  have the higher sharpness and the higher quality. The interval score is not a dedicated index for reliability assessment anyhow. As the primary requirement of probabilistic forecasting, the reliability of PIs should be given a prior analysis in the assessment process. Therefore to specifically quantify and emphasize the reliability aspect, ELM output weights  $\beta$  are optimized with respect to the objective  $F$  combining ACE  $A_t^{(\alpha)}$  and overall score  $\bar{S}_t^{(\alpha)}$  to optimize both reliability and sharpness of PIs at particular confidence levels  $(1-\alpha_i)$ ,  $i = 1, 2, \dots, n$ ,

$$\min_{\beta} F = \sum_{i=1}^n \left[ \gamma_i |A_t^{(\alpha_i)}| + \lambda_i |\bar{S}_t^{(\alpha_i)}|_{norm} \right] \quad (6.4)$$

$$\text{s.t. } \tilde{L}_t^{(\alpha_i)}(\mathbf{x}) \geq \tilde{L}_t^{(\alpha_j)}(\mathbf{x}) \text{ and } \tilde{U}_t^{(\alpha_i)}(\mathbf{x}) \leq \tilde{U}_t^{(\alpha_j)}(\mathbf{x}), \text{ if } \alpha_i \leq \alpha_j \quad (6.5)$$

$$\tilde{L}_t^{(\alpha_i)}(\mathbf{x}) \leq \tilde{U}_t^{(\alpha_i)}(\mathbf{x}) \quad (6.6)$$

where  $|\cdot|$  is the absolute value function,  $A_t^{(\alpha_i)}$  is ACE of PIs with corresponding NCP  $(1-\alpha_i)$ ,  $|\bar{S}_t^{(\alpha_i)}|_{norm}$  denotes the normalized absolute interval score  $|\bar{S}_t^{(\alpha_i)}|$  which is normalized over the corresponding maximum score  $|\bar{S}_t^{(\alpha_i)}|_{max}$  and minimum score  $|\bar{S}_t^{(\alpha_i)}|_{min}$ , defined by

$$|\bar{S}_t^{(\alpha_i)}|_{norm} = \frac{|\bar{S}_t^{(\alpha_i)}| - |\bar{S}_t^{(\alpha_i)}|_{min}}{|\bar{S}_t^{(\alpha_i)}|_{max} - |\bar{S}_t^{(\alpha_i)}|_{min}} \quad (6.7)$$

and  $\gamma_i$  and  $\lambda_i$  are importance weights of the reliability and overall skill (including sharpness), respectively. With the normalized objectives, the importance weights  $\gamma_i$  and  $\lambda_i$  are set as unit values in the study. The compatibility of the resultant PIs with different confidence levels can be assured through the constraints given in (6.5) and (6.6). The minimum value  $|\bar{S}_t^{(\alpha_i)}|_{min}$  is set to 0, which means the perfect condition with exact forecasting results. The maximum value  $|\bar{S}_t^{(\alpha_i)}|_{max}$  is set to  $2\alpha$ , which indicates the most conservative PIs with the maximum width.

### 6.3.2 Particle Swarm Optimization

Particle swarm optimization is a heuristic and population based optimization method and has proved to be an efficient, robust and gradient-free optimization algorithm [214]. PSO also distinguishes itself from other heuristic optimization methods by its fast convergence speed. It can be seen that the objective function in (6.4) is non-differentiable with respect to the ELM output weights. Therefore PSO is applied for objective function minimization to obtain the optimized ELM.

Given that the search space of PSO is  $S$ -dimensional and the size of the particles population is  $N_p$ , the  $i$ th particle of the swarm can be represented by the  $S$ -dimensional vector  $x_i = [x_{i1}, x_{i2}, \dots, x_{iS}]^T$  and the best particle in the swarm, i.e. the particle generating the smallest objective function value, is expressed by  $P_g^b$ . The previous best position, i.e. the position with the smallest objective function value of the  $i$ th particle, is stored in a vector and expressed as  $P_i^b = [P_{i1}^b, P_{i2}^b, \dots, P_{iS}^b]^T$ , and the position velocity of the  $i$ th particle is represented

as  $v_i = [v_{i1}, v_{i2}, \dots, v_{iS}]^T$ . In each iteration of PSO, the velocity of each particle is computed, and the particles are manipulated accordingly,

$$v_i = wv_i + c_1R_1(P_i^b - x_i) + c_2R_2(P_g^b - x_i) \quad (6.8)$$

$$x_i = x_i + \phi v_i \quad (6.9)$$

where  $i = 1, 2, \dots, N_p$ ;  $w$  is the inertia weight;  $\phi$  is a constriction factor controlling and keeping the velocity within the range  $[-v_{max}, +v_{max}]$ ;  $c_1$  and  $c_2$  are two positive constants;  $R_1$  and  $R_2$  are random numbers within  $[0, 1]$ . The velocity of the  $i$ th particle is a function with respect to three components: the particle's previous velocity, the distance between the previous best position of the particle and its current position, and the distance between the swarm's best success and the particle's current location. The performance of each particle is evaluated through the objective function modeled.

### 6.3.3 DIF Algorithm for PI Optimization

The proposed DIF method aims to achieve the PIs of the best quality through directly optimizing the ELM with respect to the objective function (6.4) using PSO. The core idea underneath is simply to directly approximate the PIs through a regression procedure using the PSO based optimization, where the objective function strictly measures the quality of resultant PIs including both reliability and sharpness. The major steps of the developed algorithm are described as follows:

Step 1) With the historical data of wind generation, wind speed and numerical weather prediction information and so forth, formulate the dataset  $D_t = \{(\mathbf{x}_i, t_i)\}_{i=1}^N$ , based on which two training datasets  $D_t^+ = \{(\mathbf{x}_i, \hat{t}_i^+)\}_{i=1}^N$  and  $D_t^- = \{(\mathbf{x}_i, \hat{t}_i^-)\}_{i=1}^N$  respectively for the upper and lower bounds of the PI should be prepared for ELM initialization. The targets of bounds including  $\hat{t}_i^+$  and  $\hat{t}_i^-$  can be generated by slightly increasing or decreasing original  $t_i$  by e.g.  $\pm\rho$ ,  $0 < \rho < 1$ , respectively. This manipulation is based on the knowledge that the actual wind power should be enclosed by the potential PIs.

Step 2) Given the randomly determined the input weights  $\mathbf{w}_i$  and biases  $b_i$ ,

establish an ELM to initialize the output weights  $\beta_{int}$  which is  $S$ -dimensional, using the modified training datasets obtained in Step 1).

Step 3) Initialize a population array of particles  $Pop$  with random positions around the output weights  $\beta_{int}$  of the ELM obtained in the Step 2) and velocities  $V$  in the  $S$ -dimensional search space.

Step 4) Set the iteration counter  $L = 0$ .

WHILE maximum number of iterations or sufficiently good fitness has not been reached, do

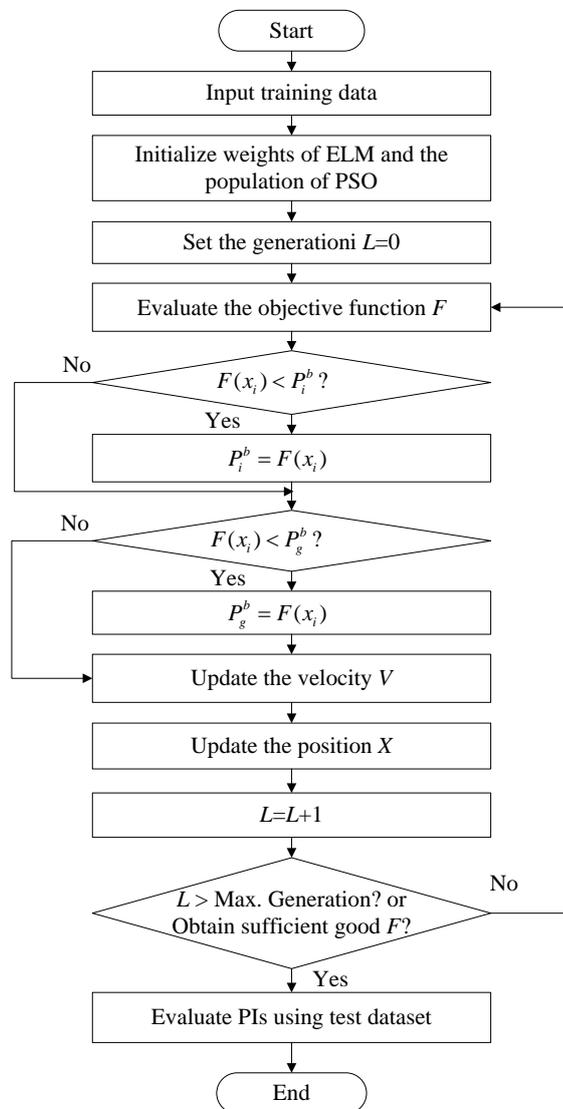
- a) For each particle in  $Pop$ , evaluate the objective function according to the PIs generated by ELM with the output weights over the original training data  $D_t$ .
- b) Compare the particle's evaluation through value of objective function (6.4) satisfying the constraints (6.5) and (6.6) with its previous best position  $P_i^b$ . If current value is better than that of  $P_i^b$ , then set  $P_i^b$  equal to the current location.
- c) Identify the particle in the swarm better than the best experience and update the smallest value of objective function (6.4) and the best position  $P_g^b$ .
- d) Change the velocities and move the positions of particles according to (6.8) and (6.9).
- e) Keep the particles in the given search space in case that they exceed their valid boundaries, and when the decision variable is out of its lower or upper boundary, takes the value of its corresponding boundary.
- f) Increment the iteration counter  $L = L + 1$ .

Step 5) END WHILE

Step 6) Based on the test data, evaluate the PIs generated by the ELM with optimized parameters  $\beta$ .

The overall flowchart of the DIF algorithm is shown in Figure 6.2. According to the detailed procedures of the proposed algorithm introduced above, the

proposed DIF approach can construct an optimized ELM to directly generate the bounds of PIs with different confidences of the best quality, avoiding the efforts needed for statistical inference and distribution assumption of point forecasting errors for traditional approaches. The application of ELM provides an extremely fast initialization procedure and significantly reduces the complexity of optimizing decision variables. The proposed algorithm demonstrates high flexibility due to the high mapping capability of ELM. The proposed DIF approach is indeed performance-oriented, and the quality of constructed PIs can be ensured through optimization on the formulated objective function.



**Figure 6.2** The overall implementation procedures of the proposed DIF approach.

## **6.4 Case Studies**

### **6.4.1 Introduction of Experiment Data**

The highly chaotic climate systems are responsible for the high level of uncertainties in wind power generation. To comprehensively validate the effectiveness of the proposed approach, it is tested by two wind farms the Chalicum Hills wind farm and the Starfish Hill wind farm in Australia. The weather conditions and wind speeds vary significantly in the two regions where the wind farms locate. Therefore forecasting models and case studies are separately constructed and conducted for the two wind farms respectively.

The Chalicum Hills wind farm locates near Ararat in western Victoria, Australia, with coordinate latitude  $-37.38^{\circ}$  S and longitude  $143.09^{\circ}$  E. The wind farm has a combined generating capacity  $P_c$  52.5 MW consisting of 35 wind turbines of 1.5 MW. Wind power generation data with one-hour resolution of this wind farm used in the study covers the period from September 2008 to August 2010.

The second wind farm Starfish Hill is near Cape Jervis on the Fleurieu Peninsula, South Australia, with coordinate latitude  $-35.57^{\circ}$  S and longitude  $138.16^{\circ}$  E. It consists of 23 wind turbines of 1.5 MW each, with a total installed capacity of 34.5 MW. Wind power generation data with one-hour resolution of Starfish Hill wind farm used in the study covers the period from January 2009 to May 2010.

To ensure both forecasting performance and computation efficiency, in the case study the wind power series is used as the inputs alone to the proposed DIF approach to conduct hourly ahead forecasting, of which the results can be significant to generation and ancillary service dispatch and so on in practice, e.g. in the Nord Pool market in Scandinavia, the hourly market plays a key role in maintaining system balance [29].

### **6.4.2 Experiment Result and Analysis**

To evaluate the forecast performance of the proposed approach, five other PI forecasting methods including the climatology method, the constant forecast method, the persistence method, the exponential smoothing method (ESM), and

the quantile regression (QR) approach are employed to compute PIs using the same training and testing data for benchmarking.

The climatology is the most commonly used benchmark for probabilistic forecasts of meteorological or weather-related processes. It is the unconditional predictive distribution computed from all historical observations available. The constant forecast takes the form of normal distribution, and the mean and variance are derived from the observed wind power data. Since the climatology and constant approaches are fairly easy to outperform for short look-ahead time forecasting, other three methods also are applied for comparisons. For point forecasting, the persistence forecast method is a widely used benchmark and is known to be difficult to outperform for short look-ahead time. The persistence based probabilistic forecast model is used as benchmark herein, of which the forecast error is assumed to be random and normally distributed. Its mean is given by the last available power measurement, and the variance is computed using the latest observations. In addition, a nice benchmark the exponential smoothing method is employed for comparisons as well, which applies a normal predictive density with its conditional mean based on exponential smoothing of past measured values and its conditional variance determined from exponential smoothing of previous squared residuals [127]. It is obvious that both the persistence and ESM approaches are based on the normal assumption of forecasting uncertainty. To better demonstrate the effectiveness of the proposed approach, quantile regression approach is employed as an advanced benchmark, which does not need the assumption of probability distribution for forecasting errors [165, 222].

The proposed model mainly aims to optimally compute reliable PIs with expected confidences. In practice, power system operation always requires accurate information with high confidence levels, e.g. state estimation always pursues higher confidence level like in [223] to ensure operation security. Therefore it is much more practically meaningful to produce high confidence level PIs to satisfy the requirements of power system operation. In our case study, PIs with different NCPs involving 90%, 95% and 99% are constructed to evaluate the performance of the proposed approach, i.e.  $\alpha = [0.1, 0.05, 0.01]^T$  and  $n = 3$  in the optimization objective function defined by (6.4). The parameter  $\rho$  in

Step 1) of the DIF algorithm is set to 0.3 in the case studies. The proposed method and applied benchmarks are tested for the two wind farms for detailed analysis and comparisons. For the Challicum Hills wind farm, the wind power generation data from March 2010 to August 2010 are used for testing the forecasting methods. For the Starfish Hill wind farm, the wind power generation data from January 2010 to May 2010 are used for testing the forecasting methods. The rest data of the two wind farms are used for training the applied methods separately.

**Table 6.1** Results of different methods in Challicum Hills wind farm

NCP	Methods	ECP	ACE	Score
90%	Proposed Method	90.80%	0.80%	-6.61%
	Constant	89.31%	-0.69%	-19.77%
	Climatology	92.57%	2.57%	-18.54%
	Persistence	85.87%	-4.13%	-7.24%
	ESM	89.10%	-0.90%	-7.33%
	QR	91.43%	1.43%	-7.30%
95%	Proposed Method	95.50%	0.50%	-3.94%
	Constant	92.84%	-2.16%	-10.41%
	Climatology	94.81%	-0.19%	-9.76%
	Persistence	89.61%	-5.39%	-4.55%
	ESM	92.43%	-2.57%	-4.57%
	QR	95.83%	0.83%	-4.42%
99%	Proposed Method	98.48%	-0.52%	-1.14%
	Constant	99.91%	0.91%	-1.94%
	Climatology	97.73%	-1.27%	-2.05%
	Persistence	94.14%	-4.86%	-1.65%
	ESM	96.76%	-2.24%	-1.67%
	QR	99.14%	0.14%	-1.23%

The detailed testing results from the two wind farms, including the PIs evaluation indices ECP, ACE and overall score, are given in Tables 6.1 and 6.2 respectively. It can be observed that the proposed method can provide fairly satisfactory performances for both wind farms from Tables 6.1 and 6.2. At all confidence levels in the case studies, the ECPs of the proposed method are close to the corresponding nominal confidences. The absolute ACEs obtained from the proposed method at different nominal confidence levels for the two farms are smaller than 1%, indicating a significantly high reliability of the generated PIs. E.g., at the confidence level with NCP = 90%, the proposed method produces ECPs of 90.80% and 90.91% for the Challicum Hills wind farm and the Starfish

Hill wind farm respectively, which outperform all other methods. As an advanced approach, quantile regression method provides comparable reliability as the proposed approach, better than the other four benchmarks. Nevertheless, the proposed method has the smallest absolute interval scores for all studied cases in the two wind farms, which indicates the best overall skill and the highest sharpness of the PIs generated by the proposed approach compared to other methods. E.g., at the nominal confidence level 90%, the proposed method produces PIs with absolute interval score 6.43% for the Starfish Hill wind farm, which outperforms the applied five benchmarks. Accounting for both reliability and overall skill, the proposed DIF approach produces the best PIs in terms of comprehensive performance against the other five benchmarks.

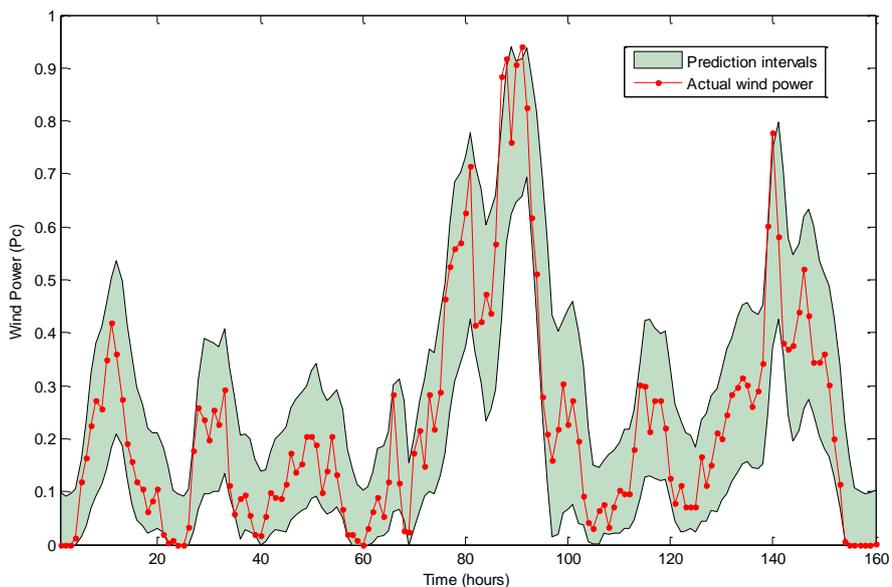
**Table 6.2** Results of different methods in Starfish Hill wind farm

NCP	Methods	ECP	ACE	Score
90%	Proposed Method	90.91%	0.91%	-6.43%
	Constant	93.48%	3.48%	-16.33%
	Climatology	97.65%	7.65%	-16.47%
	Persistence	86.93%	-3.07%	-7.17%
	ESM	89.06%	-0.94%	-7.12%
	QR	91.60%	1.60%	-7.15%
95%	Proposed Method	94.90%	-0.10%	-4.00%
	Constant	99.01%	4.01%	-8.44%
	Climatology	99.59%	4.59%	-8.62%
	Persistence	90.33%	-4.67%	-4.62%
	ESM	92.65%	-2.35%	-4.48%
	QR	96.05%	1.05%	-4.31%
99%	Proposed Method	99.28%	0.28%	-1.08%
	Constant	100.0%	1.00%	-2.00%
	Climatology	100.0%	1.00%	-1.83%
	Persistence	94.88%	-4.12%	-1.85%
	ESM	96.11%	-2.89%	-1.67%
	QR	99.28%	0.28%	-1.16%

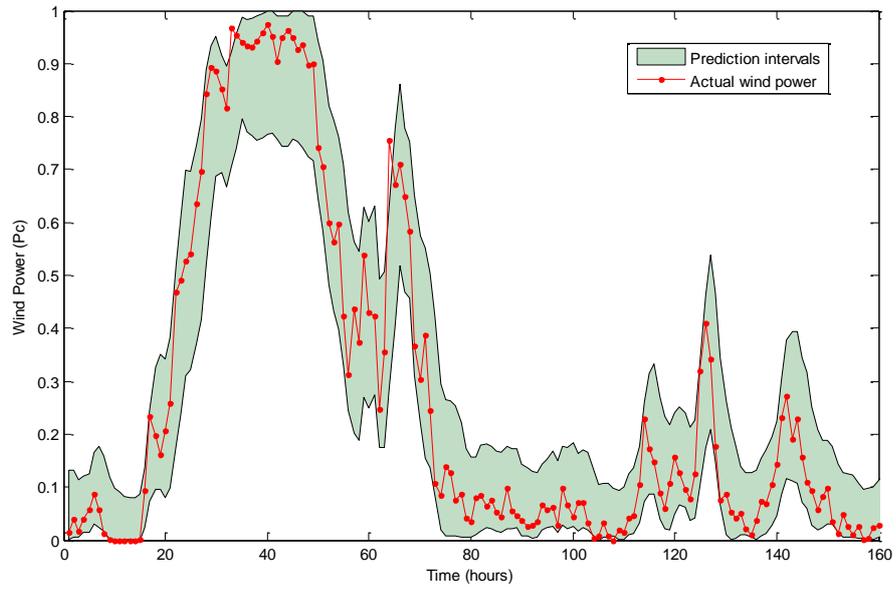
The climatology and constant approaches are unconditional forecasts and do not take into account the non-stationarity and heteroscedasticity of wind power series. Though PIs derived by the climatology and constant forecasts demonstrate fair reliability at the tested high confidence levels, they are generally too wide with low sharpness and therefore not meaningful for practical applications. ESM and persistence based interval forecasting approaches are difficult to outperform for short-term forecasts. According to the experiment results, the ESM and

persistence forecasts cannot generate PIs to best fit the expected confidences especially for the NCPs larger than 95%. According to the experiment results, quantile regression approach can derive relatively comparable PIs to the proposed approach, especially from the aspect of reliability. Generally, it performs better than the other four benchmarks from the perspectives of both reliability and sharpness. This should not be unreasonable since the quantile regression approach does not require any distribution assumption of forecasting errors, as a conditional forecasting approach.

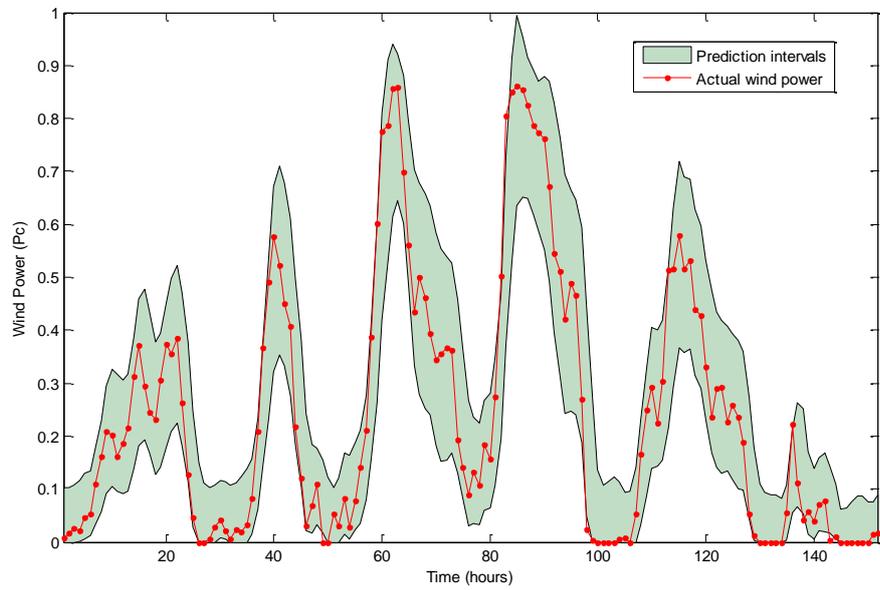
PIs with NCP 90% obtained by the proposed method and the corresponding actual wind power are displayed in Figures 6.3-6.6, where the actual measured wind farm outputs are perfectly covered by the constructed PIs in the tested two wind farms. Figures 6.3-6.6 visually demonstrate the highly satisfactory performance of the proposed approach in different months for the two wind farms. It also can be easily found that the wind power series have different nonstationary characteristics at different time and different regions. In consideration of that some generated PIs may have abnormal values beyond the possible generation range of the wind farms, the resultant predictive densities shown in Figures 6.3-6.6 have been censored to concentrate probability of abnormal conditions mass on the bounds.



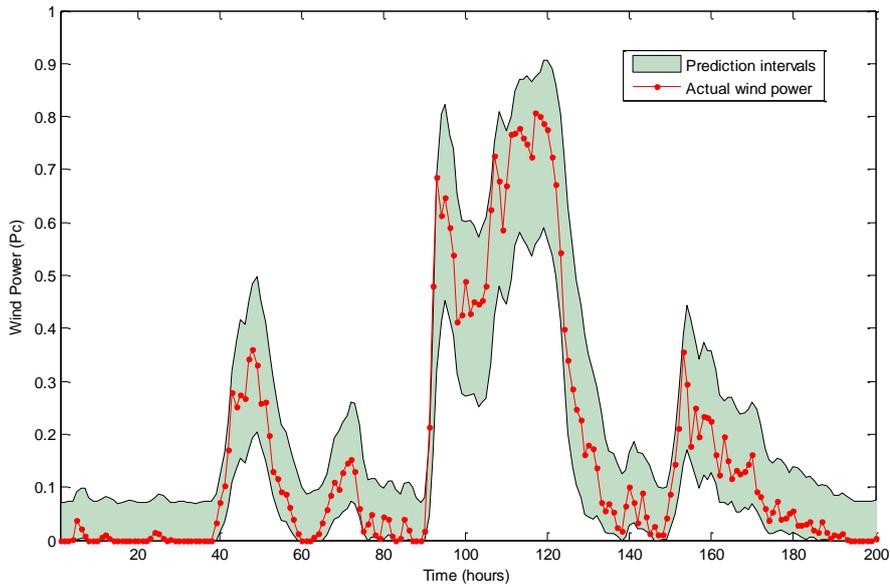
**Figure 6.3** PIs with NCP 90% in March 2010 of the Challicum Hills wind farm obtained the proposed DIF approach.



**Figure 6.4** PIs with NCP 90% in June 2010 of the Chalicum Hills wind farm obtained by the proposed DIF approach.



**Figure 6.5** PIs with NCP 90% in February 2010 of the Starfish Hill wind farm obtained by the proposed DIF approach.



**Figure 6.6** PIs with NCP 90% in May 2010 of the Starfish Hill wind farm obtained by the proposed DIF approach.

The experimental results demonstrate that the proposed method is highly satisfactory for short-term probabilistic wind power forecasting in comparisons with other five benchmarks including both time series and statistical models. Though wind power series is taken as the input alone to produce hourly ahead PIs in the case study, the proposed DIF approach in this chapter gives a generalized forecasting framework having the advantages of flexible extendibility in terms of inputs, outputs and look-ahead time window, because of the high mapping ability of ELM. It is well known that wind power generation fluctuates due to the volatility of the wind speed, wind direction, etc. For wind power prediction with longer than a few hours look-ahead time, it is necessary to involve numerical weather prediction data as the forecasting model inputs. Certainly, this can be easily included to the proposed model.

In most existing interval forecasting methods, it is necessary to conduct quantile analysis of point forecast errors involving statistical inferences, with or without prior assumption of the forecast error distribution. For instance, in the case study the ESM and persistence rely on the normal assumption of wind power forecasting errors. Comparing with quantile regression without the need of

distribution assumption, the proposed approach shows flexible and higher regression ability due to the universal mapping capability of ELM. The DIF approach focuses on PIs quality and offers a novel framework that does not require any information of point forecast results or the associated errors at all. Moreover, since the proposed method provides a performance oriented optimization model, the quality of PIs can be ensured through the optimization directly. Due to the optimization and flexibility, it has high potential practical applications to power systems operation, including reserve determination, wind power trading, wind farm control, unit commitment and so on.

## 6.5 Conclusions

Wind power forecasting is critical to modern power system operation with increased wind penetration. Traditional probabilistic wind power forecasting approaches are usually based on prior knowledge or assumption of forecasting errors. In Chapter 5, a novel BELM based parametric approach is proposed for probabilistic forecasting of wind power, based on the censor normal distribution for forecasting error. In this chapter, a novel DIF approach combining extreme learning machine and particle swarm optimization is developed and successfully applied for interval forecasting of wind power without the prior knowledge of forecasting errors. A novel objective function accounting for PIs coverage probability and overall skill is constructed to obtain optimal PIs at multiple confidence levels simultaneously through one single performance-oriented optimization process to ensure both reliability and sharpness. The effectiveness of the proposed method for short term forecast has been successfully verified through tests and comparisons with several well-established benchmarks using practical wind farm data. The proposed DIF approach provides a general framework of probabilistic wind power forecasting, with high flexibility. With large scale of wind power integration in modern power systems, the proposed DIF approach indicate high potential in practical applications in power systems operations, e.g. reserve determination by TSO to meet the load and safely and economically operate the systems.



# **7** *Pareto Optimal Interval Forecasts of Wind Power: A Case Study on Bornholm Island, Denmark*

## **7.1 Introduction**

Smart grid aims at construction of reliable, secure, deregulated, flexible, and efficient power systems and has attracted increased attentions recently. However, the uncertainty introduced by wind generation is indeed a big challenge for large scale integration into the smart grid [16]. For instance, the penetration of wind power of microgrid in Bornholm Island of Denmark reaches more than 50%. With increased penetration of wind power, inaccurate forecast can result in both technical and economic consequences [31, 34]. Accurate forecasting of wind power becomes more and more important than before to power system operation planning and control.

From the preceding chapters, it is well known that probabilistic forecasting can provide a possible range of wind power with associated probabilities, i.e. confidence levels, compared to traditional point prediction, which can help the system operator to effectively hedge against potential risks in all kinds of decision making activities such as energy and reserve dispatch. Generally, most researches focus on developing new techniques for probabilistic forecasting of wind power, including quantile regression [165], ensemble-based forecasts [184], adaptive resampling approach [113], etc., as well as accounting for more relevant variables spatio-temporal information to improve the prediction accuracy [126].

In principle, the reliability and sharpness are the two key quality indicators of

PIs [103]. However, in the many researches, the reliability is considered as the only index for measuring the PIs quality, and the sharpness is often ignored in the evaluation process though it is indispensable [114, 115, 117, 202]. Traditionally, NN based PIs construction approaches including bootstrap, delta, and Bayesian based ones [110, 116, 199] are based on normal distribution assumption for the prediction errors, and the evaluation indices are not involved in the formulation process. In Chapter 5, a parametric BELM approach is developed for probabilistic wind power prediction with the censor normal assumption of forecasting errors. To avoid the parametric assumption of probability distribution of prediction errors, in Chapter 6, an ELM-based nonparametric DIF approach is proposed to directly generate optimal PIs of wind power, combing both the objectives reliability and interval score.

The LUBE approach is proposed to generate the bounds of PIs through the CWC based cost function [216] and has been used for wind power forecasting in [178, 224]. However, it has been proved that the LUBE approach can derive wider or narrower PIs due to the flawed formulation of the cost function [175]. This can be caused by the improper and unfair treatment to the reliability and sharpness. Multi-optimization based PIs construction is proposed for wind speed forecasting in [225] using traditional neural networks (NNs) and the flawed evaluation index coverage-width criteria (CWC), which has been discussed in [175, 218]. According to [76, 103], probabilistic forecasting aims at maximizing the sharpness of the prediction intervals, subject to reliability, based on the available training and test data. Actually, the reliability and sharpness can be incompatible to some extent. In view of this, a Pareto optimal ELM-based forecaster is developed to optimize the two objectives reliability and sharpness of the generated PIs. This new formulation will directly address the key properties of PIs which is superior to the biased formulation of the CWC based cost function.

In this chapter, the optimal PIs of wind power is constructed through a multi-objective optimization procedure for training ELMs [203]. The Pareto optimal front of the generated PIs is obtained through the NSGA-II algorithm subject to different constraints [226], which has been successfully used in different multi-objective optimization problems [227]. Reliability precision analysis via

reliability diagram is introduced to select the desired optimal decision variables from the derived Pareto optimal front. The proposed approach is tested using realistic wind power generation data of Bornholm Island, which represents a practical microgrid with high wind penetration in Denmark. The proposed method is further compared with other mature models, of which the results demonstrate the superior performance of the proposed approach.

## 7.2 Optimization Formulation and Procedures

### 7.2.1 Formulation of Optimization Objectives

The proposed approach aims to construct an ELM to generate the bounds of PIs depending on inputs, pursuing the best quality and avoiding any assumption of probability distribution. According to the evaluation criteria described in Section 3.3, optimal PIs should have the best reliability with the minimal absolute ACE defined by (3.14) and the best sharpness with the minimal AWPI defined by (3.16). Given NCP  $(1-\alpha)$ , the parameters (output weights) of ELM is optimized to reach the two objectives given as

$$\begin{aligned} & \mathbf{Min} \quad |\text{ACE}| \\ & \mathbf{Min} \quad \text{AWPI} \\ & \text{s.t.} \quad \text{AWPI} > 0 \end{aligned} \tag{7.1}$$

where  $|\cdot|$  is the absolute function. To acquire optimal decision variables, the two objectives should be optimized simultaneously.

It would be very difficult to ensure that the performance of optimal decision variable must be better than any other solutions with regard to the two objectives [228]. Proper trade-off has to be involved in any multi-objective problem. Therefore, the Pareto optimality will be introduced to handle the multi-objective problem and derive the Pareto front of the optimal decision variables.

### 7.2.2 Pareto Optimality

Many real world problems involve multiple measures of performance or objectives, which should be optimized simultaneously, such as in the study of reliability and sharpness of PIs. The simultaneous optimization of multiple, possibly conflicting objectives deviates from single function optimization in that

it seldom admits a single, perfect solution. Instead, multi-objective optimization problems tend to be characterized by a family of alternatives, which must be considered to be equivalent in the absence of prior knowledge of the relevance of each objective to the others.

The family of solution of a multiobjective optimization problem is composed of all decision vectors, for which the corresponding objective vectors cannot be improved in any dimension without degradation in another. This is known as the concept of Pareto optimality [229]. Without loss of generality, a minimization problem is considered here and a set of optimal decision variables are found with respect to the formulated multiobjectives  $f_i(\cdot), i = 1, 2, \dots, k$ . Let  $x_l$  and  $x_m$  denote decision vectors. Then,  $x_l$  dominates  $x_m$  (also written as  $x_l \succ x_m$ ) if and only if both the conditions (7.2) and (7.3) are satisfied

$$\forall i \in \{1, 2, \dots, k\}, f_i(x_l) \leq f_i(x_m) \quad (7.2)$$

$$\exists j \in \{1, 2, \dots, k\}, f_j(x_l) < f_j(x_m) \quad (7.3)$$

All decision vectors which are not dominated by any other vectors of a given set are Pareto Optimal solutions. The set of all Pareto Optimal solutions of a multiobjective problem is called Pareto Optimal Set and we denote it as  $P^*$ . The Pareto-optimal front consists of decision vectors that are nondominated within the entire search space, expressed as

$$\mathcal{PF}^* = \{ (f_1(x), \dots, f_k(x)) \mid x \in P^* \} \quad (7.4)$$

### 7.2.3 NSGA-II

The NSGA-II algorithm has been proved as one of the most efficient algorithms for multiobjective optimization. Its fast nondominated sorting approach successfully reduces the computational complexity of nondominated sorting process to  $O(MN^2)$  (where  $M$  is the number of objectives and  $N$  is the population size). Its crowded-comparison approach ensures a good spread of solutions in the obtained set of solutions without any extra parameter tuning compared with some other diversity preservation methods (e.g. the well-known sharing function approach) [226]. A brief description of the core operators of

NSGA-II is given as follows:

### 1) *Fast Non-dominated Sorting Approach*

For each solution, two entities are calculated: 1) domination count  $n_p$ , the number of solutions which dominate the solution  $p$ , and 2)  $S_p$ , a set of solutions that the solution  $p$  dominated.

All solutions in the first non-dominated front ( $R_1$ ) will have their domination count as zero. For each solution in  $R_1$ , each member of its set  $S_p$  will be visited and its domination count will be reduced by one. By doing so, those members whose domination count becomes zero comprise the second non-dominated front ( $R_2$ ). The above procedure is continued with each member of  $R_2$  and the third front is identified. This process continues until all fronts are identified.

### 2) *Diversity Preservation*

To promote the solutions in the sparse region, crowding distance  $i_{distance}$  is assigned to each candidate solution. The crowding distance computation requires sorting the population according to each objective value in ascending order of magnitude. Thereafter, for each objective, the boundary solutions (solutions with smallest function value  $f_m^{min}$  and largest function values  $f_m^{max}$ ) are assigned an infinite distance value. For a specific objective  $m$ , a distance value of  $i$  is calculated as follow:

$$i_{distance.m} = (f_{[i+1].m} - f_{[i-1].m}) / (f_m^{max} - f_m^{min}) \quad (7.5)$$

where  $f_{[i+1].m}$  and  $f_{[i-1].m}$  are the function values of the two adjacent solutions of the  $i$ th solution. This calculation is continued with other objectives. The overall crowding distance value  $i_{distance}$  is calculated as the sum of individual distance values corresponding to each objective.

### 3) *Crowded-comparison Operator*

After fast nondominated sorting and diversity preservation operation, every individual in the population would obtain two attributes: 1) nondomination rank  $i_{rank}$ , 2) crowding distance  $i_{distance}$ . A partial order  $\prec_n$  has been defined in [226] as:

$$i \prec_n j \text{ if } (i_{rank} < j_{rank}) \quad (7.6)$$

$$\text{or } ((i_{rank} = j_{rank}) \text{ and } (i_{rank} < j_{rank}))$$

Between two solutions with different nondomination ranks, the solution with the better rank survives. Otherwise, if both solutions belong to the same front, the solution which is located in a lesser crowded region survives.

According to the description above, the detailed implementation procedure of NSGA-II is given as follows:

### The Procedure of NSGA-II

---

**Begin**

$t \leftarrow 0$

- i) Initialize  $P_t$  (size  $N$ )
- ii) Create an offspring population  $Q_t$  (size  $N$ ) by binary tournament selection, recombination and mutation operator of GA

**while** (not termination condition) **do**

**begin**

$t \leftarrow t+1$

- iii) Form a combine population  $R_{t-1} = P_{t-1} \cup Q_{t-1}$  (size  $2N$ )
- iv) Sort  $R_{t-1}$  according to nondomination
- v) Select the new population  $P_t$  from  $R_{t-1}$  according to the crowded-comparison operator
- vi) Create an offspring population  $Q_t$

**end**

**end**

---

#### 7.2.4 Reliability Precision

The reliability diagram is a common diagnostic for verifying the property of probabilistic forecasts [230]. A reliable predictor should generate a reliability diagram close to the diagonal. However, even a perfectly reliable prediction system is not anticipated to have an exactly reliability, i.e., zero ACE, due to the effects of limited training samples. Therefore, to assess the correctness of a developed forecaster, it is needed to quantify how far the observed coverage rate is anticipated to be from nominal confidence to a reliable forecast system [230], [231]. This interval of the empirical coverage rate can be regarded as consistency bar. The consistency resampling technique is used to quantify the range of

expected frequency if the prediction probability was in fact reliable, and is described in the following procedures

- Step 1 Set forecasts  $X$  to a constant interval in the uniform domain for a given nominal coverage rate. For instance, 90% intervals are equal to [0.05, 0.95].
- Step 2 Randomly generate uniformly distributed values  $Z$  with the number  $N$  equivalent to our test dataset size.
- Step 3 Compute the relative frequency  $f$ , satisfying  $Z$  in  $X$  (it should be 90% in theory, but here it will vary owing to sampling effects).
- Step 4 Repeat Steps 1-3 for  $N_{boot}$  times.
- Step 5 Obtain the consistency bar.

Let  $F_B$  denote the cumulative function of the empirical distribution of the sampled frequency. Given the  $(1-\rho)$  confidence level, the lower and upper bounds of consistency bar are estimated in the study, given as

$$Bar = [P_r^{(\rho)}, P_r^{(\bar{\rho})}] \quad (7.7)$$

where

$$P_r^{(\rho)} = F_B^{-1}(\rho / 2) \quad (7.8)$$

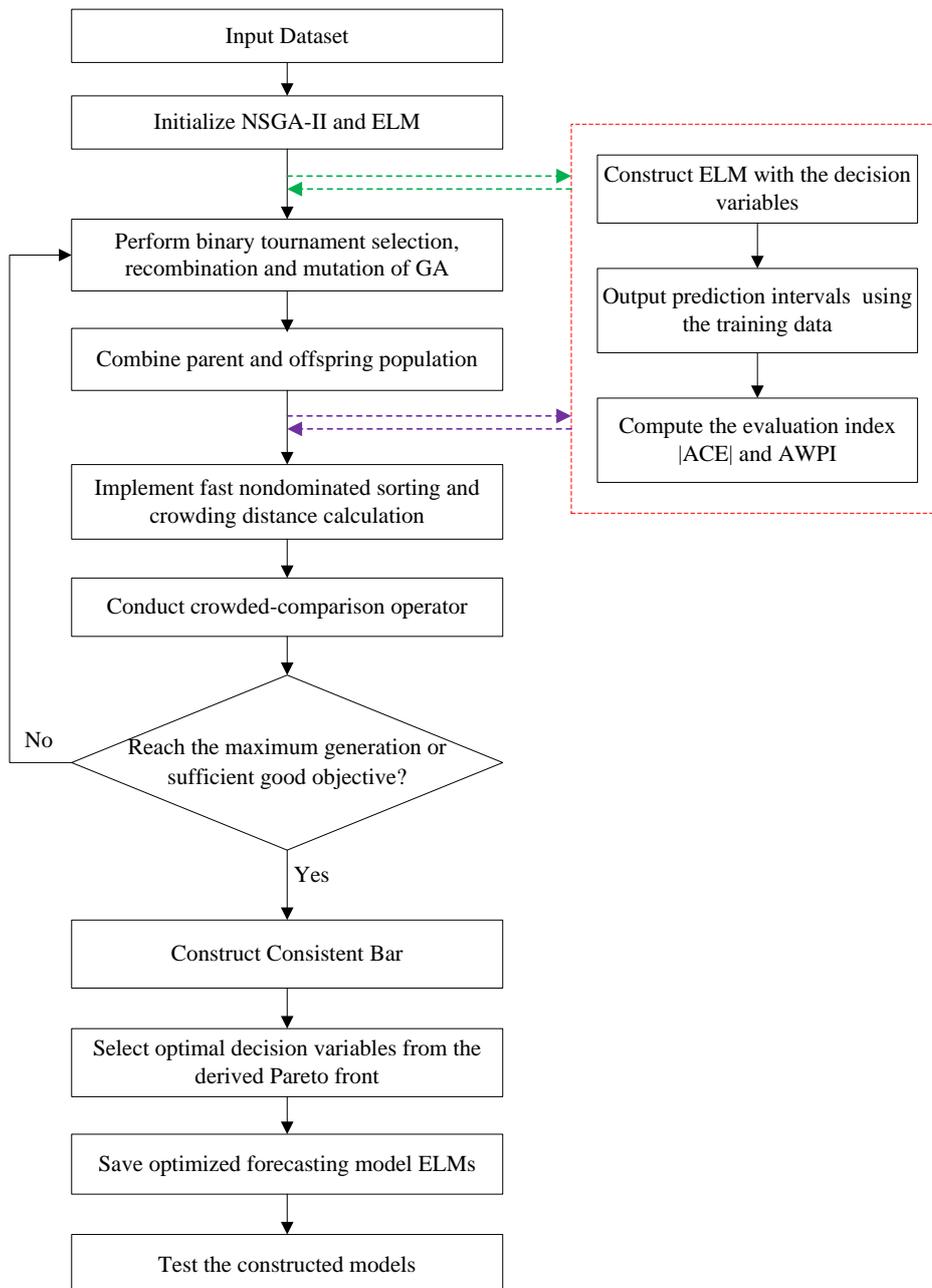
$$P_r^{(\bar{\rho})} = F_B^{-1}(1 - \rho / 2) \quad (7.9)$$

It could give the range of potentially observed coverage rate for a perfectly reliable probabilistic forecasting system. The consistency bar not only helps to measure the correct reliability of constructed prediction intervals but also provides an efficient tool for selecting the desired models from the Pareto front.

### 7.2.5 Overall Framework and Model Selection

The main algorithm of the proposed approach can be divided into three major steps, 1) Initialize prediction models and evolutionary algorithm, collect dataset divided into training data and test data and configure the inputs and outputs of formulated prediction models, 2) Optimize the parameters of prediction models with respect to the reliability and sharpness of corresponding constructed PIs based on the training dataset, 3) Identify the best models from the derived Pareto front and conduct forecasting verification over the test dataset.

The detailed procedures of the proposed Pareto optimization algorithm combining NSGA-II and ELM can be depicted in Figure 7.1. The decision variables are the output weights of the ELM based forecaster. The proposed approach aims at optimizing the overall quality of prediction intervals involving two objectives of the reliability, i.e.,  $|ACE|$ , and the sharpness, i.e., AWPI. The Pareto optimal results can be obtained through simultaneously optimizing the two objectives over the training dataset.



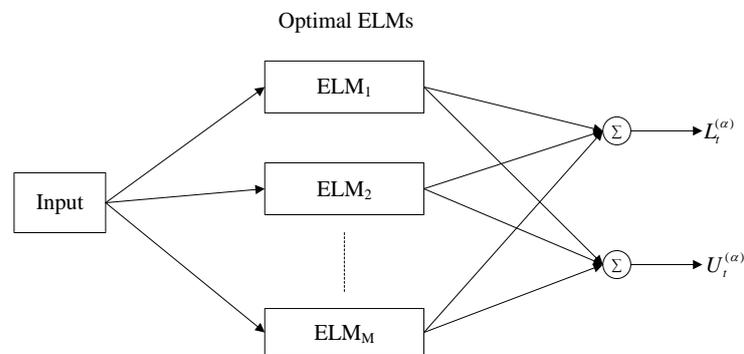
**Figure 7.1** The overall flowchart of the proposed algorithm.

After the NSGA-II based optimization, the main task is to find the desired best decision variables from the set of solutions on the Pareto front and then construct and save the desired forecasting models. According to the reliability precision analysis via the reliability diagram introduced in Section 7.2.4, if the ECP of a forecaster stays in the consistency bar, the PIs can be considered as perfectly reliable. Given certain confidence level, the corresponding consistency bar defined in (7.7)-(7.9) is computed over the training data. From the view point of practice, all the selected  $M$  models with ECPs within the range of probability interval can be used as the final forecaster. Then the best  $M$  forecasting models in terms of reliability can be identified from the Pareto front. Therefore, to make the forecasting more reliable and robust, the built forecasting system is composed of a family of ELMs satisfying the reliability precision, expressed as  $ELM_i, i = 1, 2, \dots, M$ . The lower and upper bounds of the PIs generated by the forecasting system are calculated as the mean values of the outputs of the all ELMs according to (7.10) and (7.11).

$$L_t^{(\alpha)} = \frac{1}{M} \sum_{i=1}^M L_i^{(\alpha)} \quad (7.10)$$

$$U_t^{(\alpha)} = \frac{1}{M} \sum_{i=1}^M U_i^{(\alpha)} \quad (7.11)$$

Accordingly, the overall structure of the developed forecasting system is displayed in Figure 7.2.

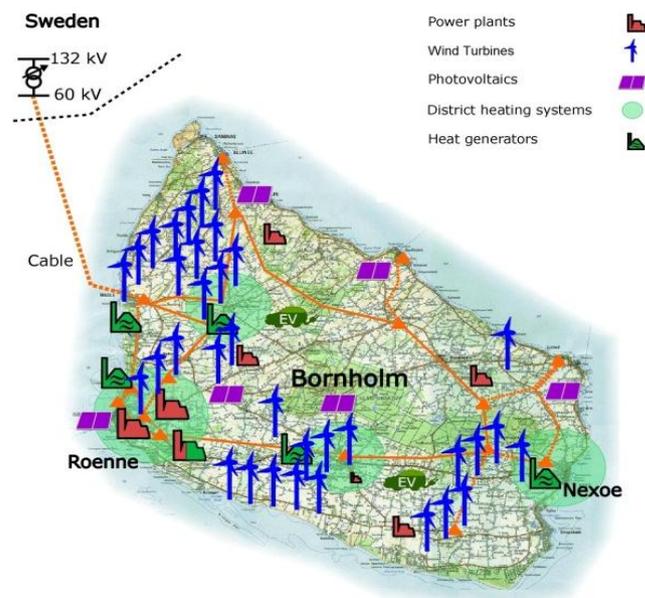


**Figure 7.2** The framework of the developed forecasting system by the proposed approach.

## 7.3 Case Studies

### 7.3.1 Wind Farm Data

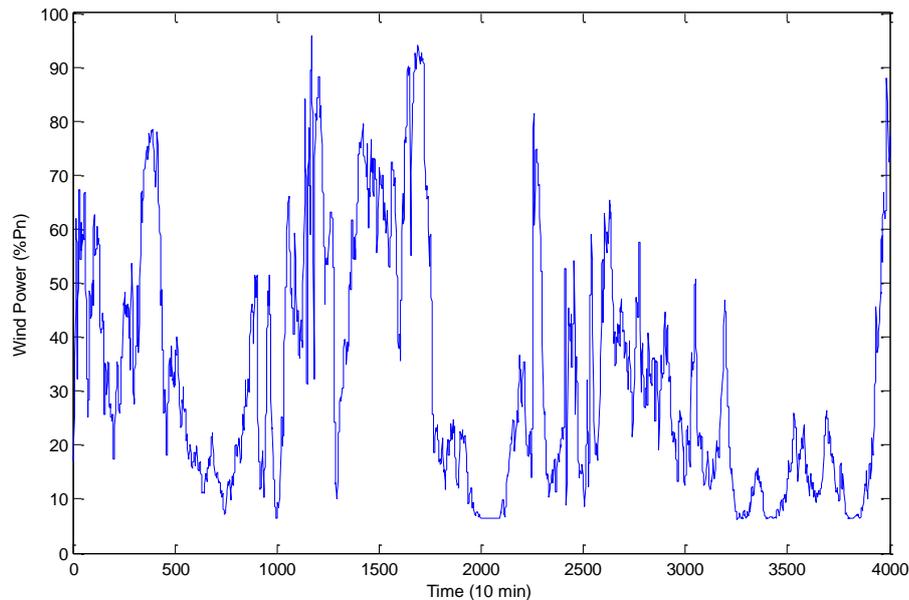
The proposed approach is tested using wind generation data of Bornholm Island in Denmark, which is a practical microgrid and has been used as test beds in several large scale projects including e.g. EU More Microgrid project, EDISION project and so forth [232]. The Bornholm electricity network has a more than 50% penetration level of wind power. Moreover, the Bornholm Island is as a base for developing and testing numerous smart grid technologies nowadays. Figure 7.3 shows the Bornholm power system where the layout of the 60 kV backbone network as well as traditional and wind generations are clearly depicted. The wind farm in Bornholm Island possesses a total installation capacity  $P_n$  of about 30 MW.



**Figure 7.3** The wind farm of Bornholm Island (courtesy of Centre for Electric Power and Energy, Technical University of Denmark).

Traditionally, hourly wind power data are favored to conduct and develop wind power forecasting systems with different forecasting horizons [33]. In practice, different look-ahead horizons satisfy different needs of operation and planning in power systems [125]. With the rapid development of smart grid and the higher and higher penetration of wind generation, the very short-term

forecasting becomes essential for optimal control and operation of wind farms and the whole electricity network [233]. Particularly, the 10-min wind power is taken as the most critical one having serious influences on the balance of power systems, according to the TSO in Denmark [213].



**Figure 7.4** Time series with wind power (10-min resolution).

Figure 7.4 shows the time series of normalized wind power with 10-min measurements in microgrids of Bornholm Island. It can be found that wind power production at this resolution demonstrates seriously high volatility. Wind generation data of Bornholm Island with a 10-min resolution between October and December 2012 are used for the reported case studies.

### 7.3.2 Numerical Analysis and Discussions

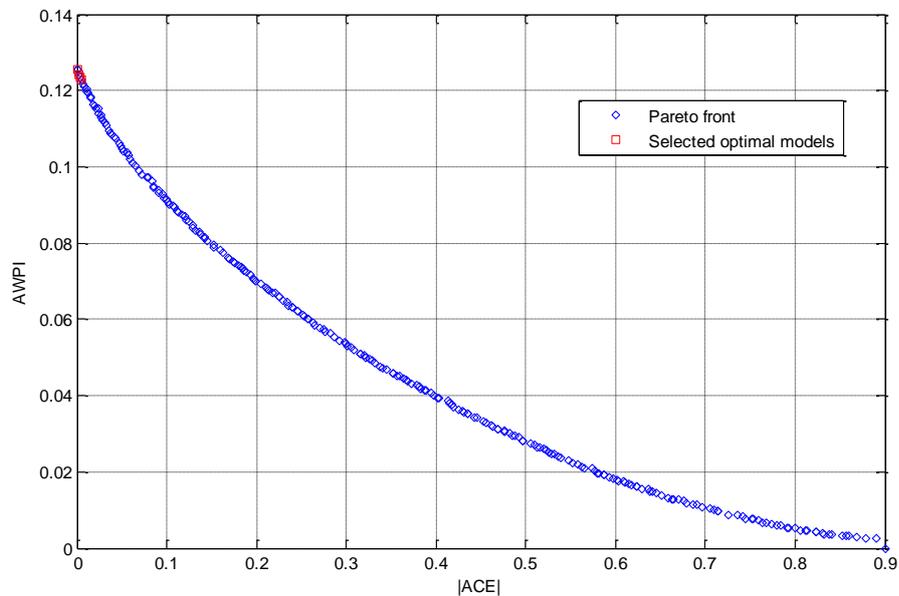
Accounting for the significant variations of wind power especially in short time scale, our case study covers the short-term forecasts with different look-ahead horizons including 30 minutes, one hour and two hours, which would be meaningful for wind farm control, storage dispatch and so forth in practice. According to [33], for wind generation forecast with a lead time shorter than a few hours the statistical forecasting models using only the historical

measurements as the inputs alone can perform better than NWP-based prediction. What's more, implementation of the NWP models based on 10-min resolutions is computationally intensive because of too frequent need of information retrieving from NWP system, which is hardly possible. Therefore, only historical wind generation data are taken as inputs to the proposed models in the case study. For forecast with longer look-ahead time, other data such as the weather information can be easily included in the proposed model to enhance the accuracy.

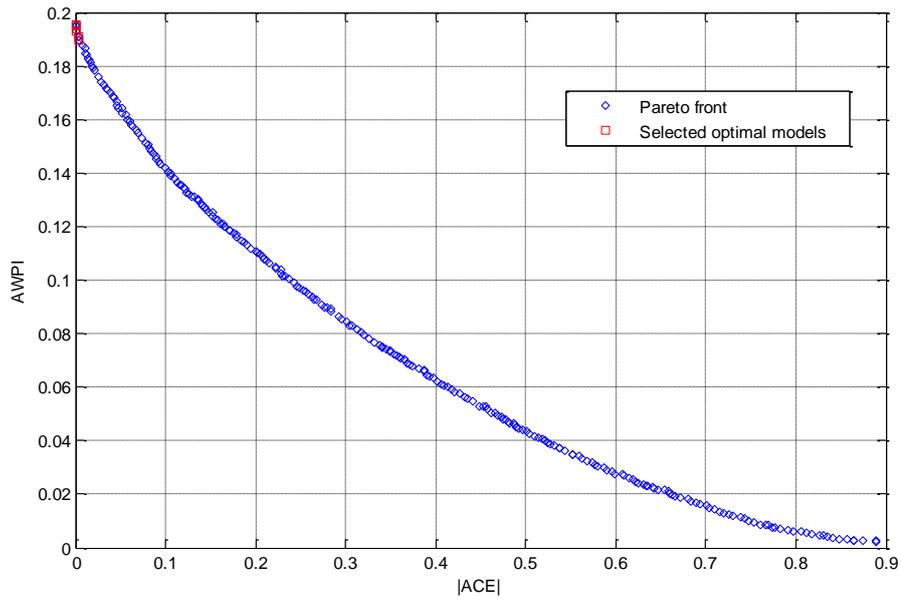
To verify the effectiveness of the proposed approach, 75% of the wind power data are used for training the model and the rest 25% are employed to test the constructed prediction model. The climatology and the normal distribution are used for implementing comparisons [173, 234], to assess the performance of the proposed approach. Climatology predictive distribution is a widely used benchmark for weather-related processes, and is modeled via all observed wind power measurements. The normal distribution assumes that the future wind power is normally distributed, and the mean and variance of the distribution are estimated from the historical wind generation data. Obviously, the climatology and normal distribution are unconditional probabilistic forecasts approaches. For the deterministic prediction of wind production, the persistence approach is regarded as the most widely applied benchmark and hard to be overtaken for short prediction horizon [33]. The persistence based interval forecasts model is employed for comparison analysis as well. The mean of the persistence model is given by the last available measurement, and the variance is calculated through the latest observations. In addition, an advanced NN-based LUBE approach is used to benchmark the proposed model as well [178, 216], the bounds of the prediction intervals are approximated through optimizing the CWC-based cost function.

Probabilistic forecasts aim to generate PIs as reliable as possible. Practically, it should be more preferable to obtain PIs of relatively high confidence level to satisfy the requirements of operation and control in smart grids and ensure the reliability and security of power systems. In the numerical studies, the two nominal confidence levels at 80% and 90% are considered respectively. The Pareto fronts of PIs with NCP 90% and different look-ahead horizons 30 minutes, 1 hour, and 2 hours are given in Figures 7.5-7.7 respectively, where the Pareto

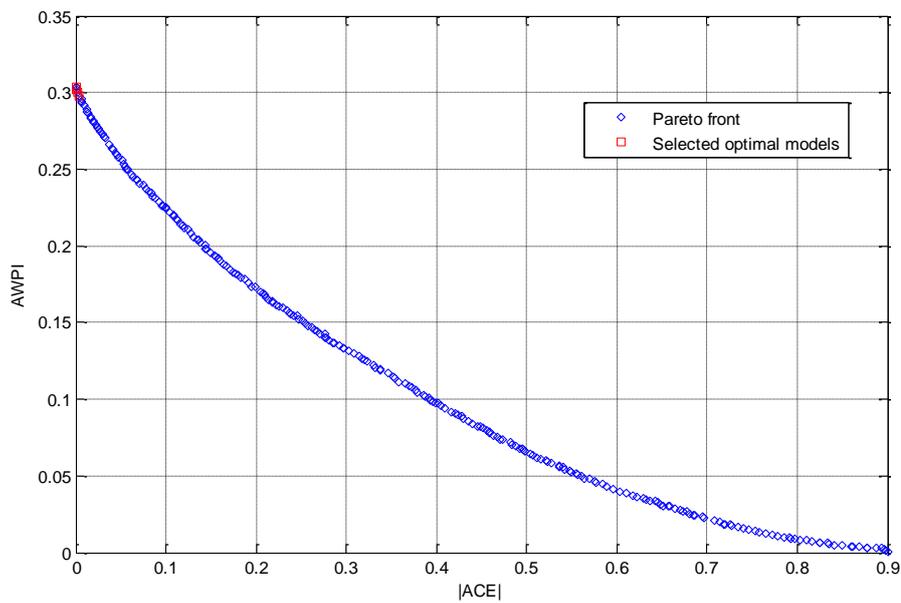
solutions are well dispersed along the smooth fronts. Each solution along the Pareto front corresponds to an optimal forecasting model. In order to select the most reliable models among the many Pareto ones, the reliability precision reflected by the consistency bar is approximated through sampling technique on the training data. The confidence level for quantifying consistency bar defined in (7.7)-(7.9) is set to 90%. The solutions that have reliability objective  $|ACE|$  satisfying the reliability precision are identified as the best models to form the whole forecasting system. The corresponding best forecasting models selected from the Pareto front according to the given identification criteria are shown in Figures 7.5-7.7. It can be seen that PIs the smaller  $|ACE|$  have the larger AWPI, which should be consistent with the fact that the higher ECP corresponds to relatively wider PIs under the optimal condition. Among the cases of different look-ahead horizons, PIs of the 2-hour forecasting horizon have the largest AWPI, and the 1-hour ahead forecasting gives the smallest AWPI from Figures 7.5-7.7. This should not be difficult to understand that the forecasting with longer horizon can have higher uncertainty due to the chaotic nature and nonstationarity of wind power.



**Figure 7.5** The Pareto front and identified optimal models of PIs with NCP 90% and look-ahead time 30 minutes.



**Figure 7.6** The Pareto front and identified optimal models of PIs with NCP 90% and look-ahead time 1 hour.



**Figure 7.7** The Pareto front and identified optimal models of PIs with NCP 90% and look-ahead time 2 hours.

Based on the obtained Pareto fronts, the identified best ELMs finally constitute the desired forecasting model. It is examined on the test dataset. The evaluation results of constructed PIs with look-ahead times of 30 minutes, 1 hour and 2

hours, including NCPs and corresponding consistency bar, ECP, ACE and AWPI are given in Tables 7.1-7.3 respectively.

According to Tables 7.1-7.3, it can be seen that the ECPs of the proposed approach can well stay within the computed consistency bar for different NCPs and forecasting horizons. This indicates the highly reliable performance of the proposed interval forecasts approach over the test data. Nearly all the absolute ACEs are smaller than 1%. In contrast, the ECPs of the four used benchmarks can fall outside the consistency bar, except for some cases of normal distribution model, which indicates the relatively lower reliability of these approaches. For example, considering the case of NCP at 80%, the proposed method can provide ECPs of 81.09%, 80.99% and 80.34% for the look-ahead times of 30 minutes, 1 hour, and 2 hours respectively, which are better than all other applied approaches. Moreover, the proposed approach provides PIs with the smallest AWPIs comparing with the four benchmarks, demonstrating the best sharpness. For example, when NCP = 90%, the proposed approach has AWPIs of 11.71%, 17.94% and 26.74% for the look-ahead times of 30 minutes, 1 hour, and 2 hours respectively, which outperform the other four models applied. Taking reliability, sharpness, and reliability precision into consideration, the proposed approach outperforms the four benchmarks and can produce the best PIs over the test data set.

**Table 7.1** Prediction results with look-ahead time of 30 minutes

NCP	Methods	Bar	ECP	ACE	AWPI
80%	Proposed Method		81.09%	1.09%	7.70%
	Normal	78.82%	85.76%	5.76%	61.72%
	Climatology		75.74%	-4.26%	64.03%
	Persistence	81.15%	83.26%	3.26%	8.66%
	LUBE		86.82%	6.82%	8.92%
90%	Proposed Method		90.54%	0.54%	11.71%
	Normal	89.11%	90.04%	0.04%	73.01%
	Climatology		85.70%	-4.30%	74.29%
	Persistence	90.88%	87.95%	-2.05%	11.11%
	LUBE		93.92%	3.92%	13.68%

**Table 7.2** Prediction results with look-ahead time of one hour

NCP	Methods	Bar	ECP	ACE	AWPI
80%	Proposed Method		80.99%	0.99%	13.01%
	Normal	78.85%	85.75%	5.75%	61.71%
	Climatology		75.71%	-4.29%	64.05%
	Persistence	81.12%	82.22%	2.22%	14.02%
	LUBE		85.87%	5.87%	17.32%
90%	Proposed Method		90.06%	0.06%	17.94%
	Normal	89.10%	90.00%	0.00%	72.97%
	Climatology		85.86%	-4.14%	74.29%
	Persistence	90.87%	88.62%	-1.38%	17.99%
	LUBE		93.66%	3.66%	22.02%

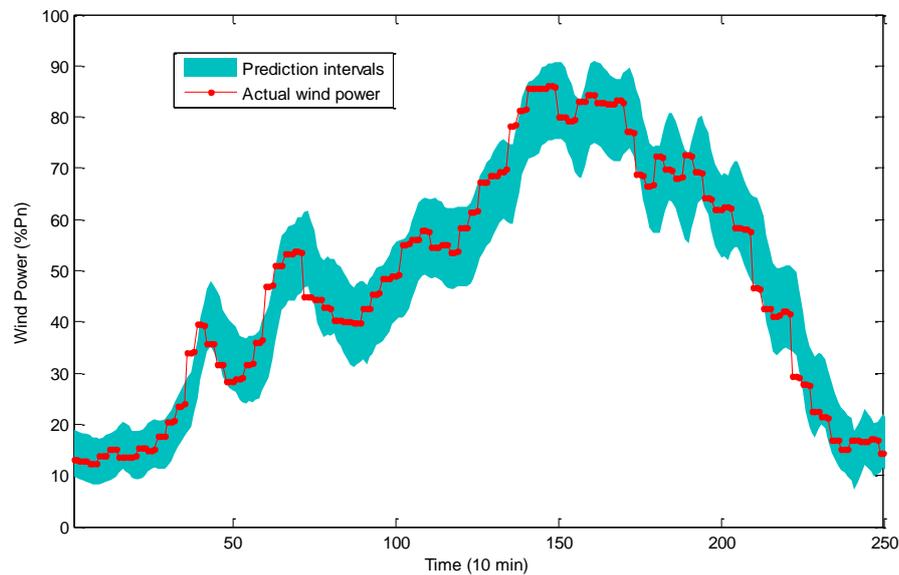
**Table 7.3** Prediction results with look-ahead time of two hours

NCP	Methods	Bar	ECP	ACE	AWPI
80%	Proposed Method		80.34%	0.34%	20.71%
	Normal	78.79%	85.63%	5.63%	61.69%
	Climatology		75.77%	-4.23%	64.06%
	Persistence	81.15%	82.25%	2.25%	21.64%
	LUBE		85.07%	5.07%	23.77%
90%	Proposed Method		89.30%	-0.70%	26.74%
	Normal	89.14%	89.98%	-0.02%	72.92%
	Climatology		85.66%	-4.32%	74.29%
	Persistence	90.82%	88.31%	-1.69%	27.77%
	LUBE		94.71%	4.71%	34.33%

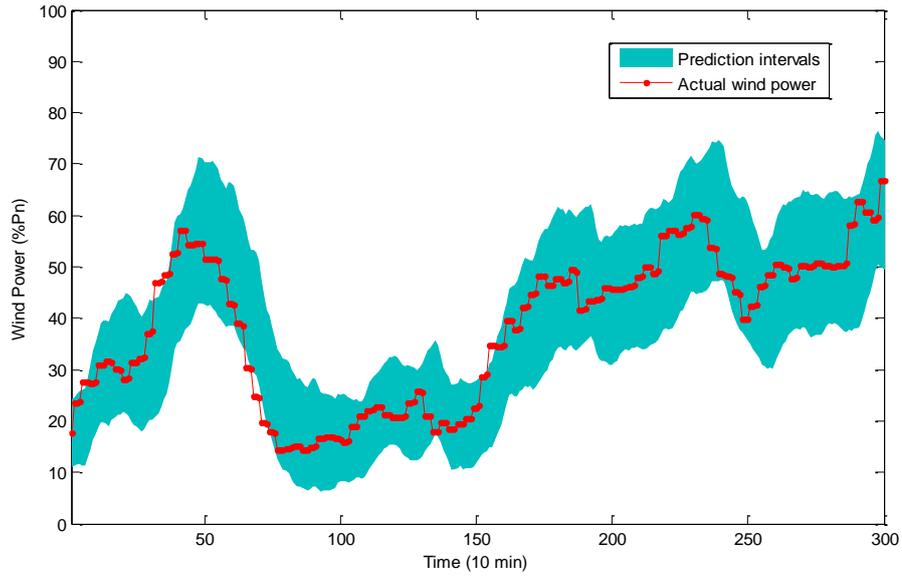
From Tables 7.1-7.3, the climatology and normal distribution generate too wide PIs comparing with other approaches. These two approaches provide simply unconditional probabilistic forecasts and cannot account for the serious volatility of the processes of wind power generation. The extremely wide PIs hardly meet the requirements of any practical applications. The ECPs of the persistence based probabilistic forecasting approach presents relatively fair performance comparing with the climatology and normal distribution approaches. With the time-series based simple mapping and normal assumption, the persistence approach cannot be flexible and satisfactory forecasting tool since the longer-term forecasting inevitably needs the meteorological information to enhance the accuracy. The LUBE approach always produces wider PIs according to the case study results, with ACEs ranging from about 4% to 6%. This can be caused by that the PIs with ECP larger than NCP are much preferable in [216]

and [178]. Actually, the wider PIs would decrease the sharpness and make unnecessary costs of decision makers in smart grids.

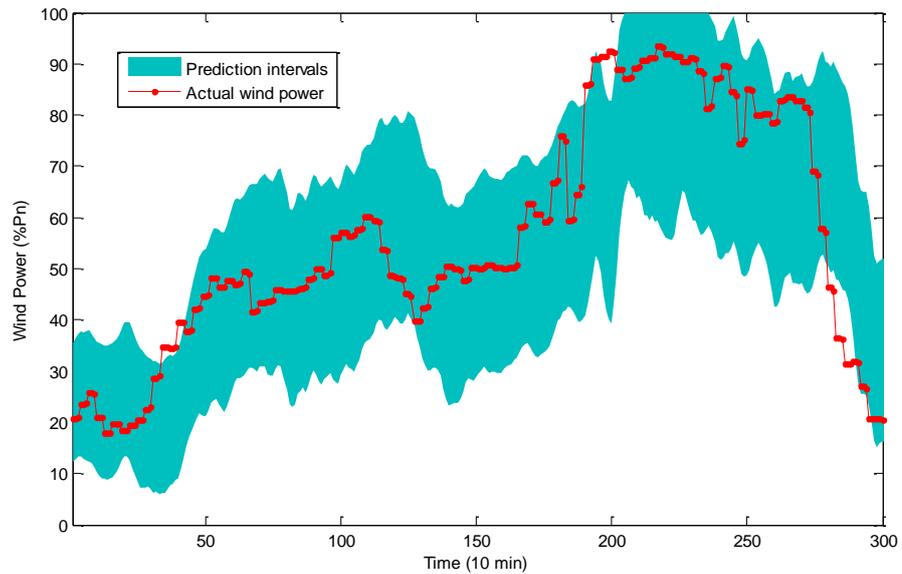
PIs with NCP at 90% constructed by the proposed method and the actually measured wind power are shown in Figures 7.8-7.10 with look-ahead horizons of 30 minutes, 1 hour, and 2 hours respectively. For different forecasting horizons, the actual observed wind generation values are well enclosed by the PIs produced by the proposed method, indicating a satisfactory performance especially for shorter horizon cases. Besides, the cases with longer look-ahead time turn out to have wider PIs, which clearly indicate that the longer term forecasting can involve higher uncertainty. It should be noted that the prediction probability densities have been censored to take the potential probability mass outside the capacity range on the bounds if PIs have abnormal values beyond the actual capacity interval of the realistic wind farm [126, 234].



**Figure 7.8** PIs with NCP 90% and 30min look-ahead time in December 2012 in the wind farm of Bornholm Island.



**Figure 7.9** PIs with NCP 90% and 1hour look-ahead time in December 2012 in the wind farm of Bornholm Island.



**Figure 7.10** PIs with NCP 90% and 2 hour look-ahead time in December 2012 in the wind farm of Bornholm Island.

Through above analysis, the proposed approach has been proved to be effective for the interval forecasting of wind power. Traditional probabilistic forecasting approaches rely on the formulation of the point forecast error distributions, the sharpness of which is limited by the quality of point forecasts.

Besides, more accurate modeling of the probability distribution would result in the sharper resultant forecasted PIs. The proposed approach does not need the point forecasts as the input, and can straightly produce the bounds of PIs without the requirements of the modeling of forecasting error distribution. Moreover, the proposed approach gives the Pareto front of wind power PIs. Pareto optimization included in the proposed approach successfully avoids the quantified ratio assigned to the reliability and sharpness that would be difficult to any decision maker. Although only short term forecast of wind power is studied in this chapter, it has the potential to provide satisfactory performance of wind power forecasting considering longer look-ahead time through e.g. incorporating meteorological conditions as additional inputs to the developed model for performance enhancement.

With the non-ignorable uncertainty and subsequent impacts, traditional point forecasts cannot meet the requirements for all kinds of decision making activities for operation and control of power systems and electricity market. In view of fast development of smart grid in the coming future, it can be expected that interval forecasts of wind power will be utilized in various decision-making problems facilitating the integration of wind power into power systems, such as wind farm control, wind power trading, reserve determination, storage dispatch, etc.

## **7.4 Conclusions**

In this chapter, a Pareto optimal interval forecasts approach combining NSGA-II and ELM is proposed for probabilistic interval forecasting of wind power generation in the study. The reliability and sharpness are particularly considered as two objectives for optimizing forecasting models in line with the evaluation criteria for PIs. The Pareto optimal front of the decision variables with respect to the two objectives are gauged through the NSGA-II based optimization. Reliability precision analysis using the reliability diagram is conducted to identify the best results among the many Pareto optimal solutions to build the forecasting system, which is also used for PIs verification in comparison analysis. Wind power data with 10-min resolution from a realistic microgrid system of Bornholm Island in Denmark are used to test the developed forecasting system. The validity of the proposed method for short term forecast of wind generation

has been examined through comprehensive comparisons with several benchmarks. The proposed approach provides a generic framework for multiobjective optimization based interval forecasting of wind power production which can reasonably identify the best possible solution among many Pareto alternatives. It can have high potential in practical application in smart grids. In our future work the proposed model will be extended to have a longer look-ahead horizon of probabilistic wind power forecasting.

# 8 *A Hybrid Approach for Probabilistic Forecasting of Electricity Price*

## 8.1 Introduction

With the deregulation of modern power systems, electricity market price forecasting becomes more and more important to the running of electricity market which plays a key role in practical operation of smart grids [44]. In the future smart grids, the electricity market price is essential to guide the behaviors of both consumers and suppliers, facilitate demand side participation, etc. Particularly, smart grids will advocate and facilitate consumers' participations directly into wholesale and other electricity markets as outlined in EU's vision for smart grids [235]. Accordingly, the electricity consumers will have strong interests in market participations by e.g. actively reacting to forecasted electricity prices through flexible management of their consumption patterns that can lead to significant benefits in economy, environmental protection, system reliability, and so forth. In general, accurate and reliable electricity price forecasting is meaningful to facilitate various decision making activities of power system short-term operation and long-term planning analysis by market participants, such as formulating bidding schemes, making investment decision, etc.

By far most researches in this field focus on point forecasting of electricity price series. Time series model approaches including autoregressive integrated moving average (ARIMA) [236] and generalized autoregressive conditional heteroscedasticity (GARCH) [237] were employed to forecast electricity price value. Since the ARIMA model is basically a linear model, wavelet techniques were integrated to improve the forecasting performance [74]. To overcome the

limitations of linear time series models, NNs were applied in electricity price forecasting [238-240].

Due to the nonstationarity of electricity price series, forecasting errors are simply unavoidable and sometimes can be significant. Recognizing the inherent limitations of traditional point forecasting methods, probabilistic forecasting recently attracts increased research interests. It can offer prediction intervals (PIs) to quantify the uncertainty associated with point forecasts of electricity clearing price. With quantified PIs, electricity market participants are able to prepare for the best and the worst conditions. ARIMA models were employed to forecast the price value and associated forecasting error respectively in [241]. The prediction intervals were obtained based on the forecasted price values and the corresponding forecasted errors. However, the ARIMA model is a linear one and difficult to accurately model the heteroskedasticity of electricity price series. MCP prediction and corresponding confidence interval (CI) were estimated by an integrated adaptive learning and CI estimation method combining neural networks and extended Kalman filter (EKF) [242]. SVM based nonlinear conditional heteroscedastic forecasting (NCHF) model was proposed to conduct interval forecasting of electricity price [119]. ELM was applied for interval forecasting of MCP in [118] combining with a wild bootstrap approach, termed as ELM-Bootstrap herein. However, the uncertainty of data noise was not considered. Though ELM has extremely fast learning speed, a great number of bootstrap iterations would cause high computation burden. To the best of our knowledge, most related works including [118] just evaluate the reliability of PIs, but do not systematically consider the sharpness and the overall skill of PIs, which is an essential quality index for the resultant PIs. Without considering the sharpness, the constructed PIs can become too wide and therefore meaningless to be used in any decision making activities in practice.

Because of the excellent approximation capability [196], NNs were widely used for electricity market price forecasts [119, 238-240, 242]. However, traditional NNs have the drawbacks like the time consuming training process, local minima, etc. Moreover, if the data are very chaotic, the performance of classical NNs based prediction methods cannot be satisfactory, and this cannot be improved by e.g. changing the NN structure or increasing the training iterations.

Several NN based methods for PI formulation have been developed, including the Delta, Bayesian, Bootstrap as well as the mean-variance estimation methods [110, 116, 199, 200, 243]. Among others, the bootstrap approach is simple and easy for implementation and proved to provide more reliable PIs [201]. It does not require calculations of complicated derivatives and the Hessian matrix involved in Delta and Bayesian methods. Nevertheless, the major disadvantage of the bootstrap based approach is the low computational efficiency for large datasets especially when a great number of NNs are involved.

In this chapter, a hybrid approach is proposed for probabilistic interval forecasting of electricity price based on extreme learning machine (ELM) and maximum likelihood estimation (MLE). ELM is a novel learning algorithm for single layer feedforward neural network featuring extremely fast learning speed [203, 244]. It has also demonstrated superior generalization capability than traditional NNs and therefore attracted more and more attentions recently. The ELM has been successfully used for probabilistic forecasting of wind power generation in Chapters 5-7 of this thesis. Because of many advantages, ELM is used to conduct point forecasts of electricity price and estimate the model uncertainties via a bootstrapping approach in this chapter. Subsequently, MLE method is used to train a NN to estimate the noise variance of forecasting results. Then the forecasting PIs can be obtained based on the variances of model uncertainty and noise. In addition, a generalized PIs evaluation framework for probabilistic forecasting of electricity price is provided in this chapter to comprehensively involve different aspects of PIs quality.

The proposed method aims to develop a PIs forecasting model for day-ahead electricity price with superior performance and fast speed. It is tested using Australian electricity market data. The obtained PIs have been comprehensively evaluated with respect to both reliability and sharpness. The numerical results indicate the effectiveness and high efficiency of the proposed hybrid approach, which can provide a meaningful online tool for probabilistic forecasting of electricity price with high potential in practical applications of smart grids.

## 8.2 PIs Formulation and Construction

### 8.2.1 Formulation

The uncertainties of artificial neural network prediction include the model uncertainty mainly caused by the misspecification of NN structure and parameters, and the noise of training data due to stochastic characteristics of regression data. Given a set of distinct pairs  $\{(\mathbf{x}_i, t_i)\}_{i=1}^N$ , the prediction target can be expressed as

$$t_i = g(\mathbf{x}_i) + \varepsilon(\mathbf{x}_i) \quad (8.1)$$

where  $t_i$  is the  $i$ th prediction target,  $\mathbf{x}_i$  is the vector of the inputs which can include historical MCPs and demands for the electricity price forecasting in the study,  $\varepsilon(\mathbf{x}_i)$  denotes the noise with zero mean,  $g(\mathbf{x}_i)$  is the true regression mean. The noise is assumed more or less normally distributed with mean zero and variance  $\hat{\sigma}_\varepsilon^2$  that may depend on the input variables  $\mathbf{x}_i$ . This normal assumption is practically not unreasonable according to [119] and [75]. The numerical case studies using actual electricity price data also prove that reliable PIs can be obtained based on the normal assumption.

In practice, the trained neural network  $\hat{g}(\mathbf{x}_i)$  could be regarded as an estimate of the true regression  $g(\mathbf{x}_i)$ . Then the prediction error can be expressed as

$$t_i - \hat{g}(\mathbf{x}_i) = [g(\mathbf{x}_i) - \hat{g}(\mathbf{x}_i)] + \varepsilon(\mathbf{x}_i) \quad (8.2)$$

where  $t_i - \hat{g}(\mathbf{x}_i)$  denotes the total prediction error, and  $g(\mathbf{x}_i) - \hat{g}(\mathbf{x}_i)$  denotes the error of the neural network estimate with respect to the true regression.

Assuming that the estimation error and noise are statistically independent, the variance of the total prediction errors  $\hat{\sigma}_t^2(\mathbf{x}_i)$  can be obtained through the summation of the variance of model uncertainty  $\hat{\sigma}_g^2(\mathbf{x}_i)$  and the variance of noise  $\hat{\sigma}_\varepsilon^2(\mathbf{x}_i)$ ,

$$\hat{\sigma}_t^2(\mathbf{x}_i) = \hat{\sigma}_g^2(\mathbf{x}_i) + \hat{\sigma}_\varepsilon^2(\mathbf{x}_i) \quad (8.3)$$

Given confidence level  $(1-\alpha)$ , the lower and upper bounds of the PI  $[L_t^{(\alpha)}(\mathbf{x}_i), U_t^{(\alpha)}(\mathbf{x}_i)]$  can be defined by the following equations

$$L_t^{(\alpha)}(\mathbf{x}_i) = \hat{g}(\mathbf{x}_i) - z_{1-\alpha/2} \sqrt{\hat{\sigma}_t^2(\mathbf{x}_i)} \quad (8.4)$$

$$U_t^{(\alpha)}(\mathbf{x}_i) = \hat{g}(\mathbf{x}_i) + z_{1-\alpha/2} \sqrt{\hat{\sigma}_t^2(\mathbf{x}_i)} \quad (8.5)$$

where  $z_{1-\alpha/2}$  is the critical value of the standard normal distribution, and depends on the expected confidence level  $(1-\alpha)$ .

### 8.2.2 Variance of Model Uncertainty

The overall forecasting uncertainty should cover two parts, the model uncertainty and the data noise. The proposed approach aims at a full quantification of both two uncertainties.

For neural network regression, the model uncertainty is mainly caused by the random initialization of network parameters and model structure misspecification. The uncertainty in the neural network estimate  $\hat{g}(\mathbf{x}_i)$  of the true regression  $g(\mathbf{x}_i)$  can be quantified by the bootstrap approach. The bootstrap approach can obtain an ensemble of ELM models, which will derive a less biased estimation of true regression of the future targets. The  $B$  training datasets are uniformly re-sampled from the original dataset  $D_t$  with replacement,

$$D_t = \{(\mathbf{x}_i, t_i)\}_{i=1}^N \quad (8.6)$$

The mean of the bootstrapped  $B$  ELMs outputs is regarded as the approximation of the true regression.

$$\hat{g}(\mathbf{x}_i) = \frac{1}{B} \sum_{q=1}^B \hat{g}_q(\mathbf{x}_i) \quad (8.7)$$

where  $\hat{g}_q(\mathbf{x}_i)$  is the prediction value of the input sample  $\mathbf{x}_i$  generated by the  $q$ th bootstrapped ELM.

The variance of model uncertainty can be estimated from the variance in the outputs of the trained  $B$  ELMs

$$\hat{\sigma}_g^2(\mathbf{x}_i) = \frac{1}{B-1} \sum_{q=1}^B \left( \hat{g}_q(\mathbf{x}_i) - \hat{g}(\mathbf{x}_i) \right)^2 \quad (8.8)$$

### 8.2.3 Variance of Noise

After determining the model uncertainty variance, to construct PIs, we need to estimate the variance of noise  $\hat{\sigma}_\varepsilon^2$ . According to (8.3), the estimated  $\hat{\sigma}_\varepsilon^2$  can be obtained through

$$\hat{\sigma}_\varepsilon^2 \sim E[(t - \hat{g})^2] - \hat{\sigma}_g^2 \quad (8.9)$$

A set of squared residuals  $r^2(\mathbf{x}_i)$  is calculated to estimate a model to fit the remaining residuals.

$$r^2(\mathbf{x}_i) = \max([t_i - \hat{g}(\mathbf{x}_i)]^2 - \hat{\sigma}_g^2(\mathbf{x}_i), 0) \quad (8.10)$$

where  $\hat{g}(\mathbf{x}_i)$  and  $\hat{\sigma}_g^2(\mathbf{x}_i)$  can be obtained from (8.7) and (8.8). A separate dataset is generated by the residuals and corresponding inputs,

$$D_{r^2} = \left\{ (\mathbf{x}_i, r^2(\mathbf{x}_i)) \right\}_{i=1}^N \quad (8.11)$$

Based on the normal assumption in (8.1), the noise is normally distributed with zero mean,

$$P(r^2(\mathbf{x}_i); \hat{\sigma}_\varepsilon^2(\mathbf{x}_i)) = \frac{1}{\sqrt{2\pi\hat{\sigma}_\varepsilon^2(\mathbf{x}_i)}} \exp\left(-\frac{r^2(\mathbf{x}_i)}{2\hat{\sigma}_\varepsilon^2(\mathbf{x}_i)}\right) \quad (8.12)$$

The formulation of noise variance can effectively approximate the noise uncertainty and is also proved through the real electricity price data. Maximum likelihood estimation approach can be used to train a new neural network for the residuals noise variance approximation. The logarithm of the likelihood function is given as

$$L = \sum_{i=1}^N \log \left[ \frac{1}{\sqrt{2\pi\hat{\sigma}_\varepsilon^2(\mathbf{x}_i)}} \exp\left(-\frac{r^2(\mathbf{x}_i)}{2\hat{\sigma}_\varepsilon^2(\mathbf{x}_i)}\right) \right] \quad (8.13)$$

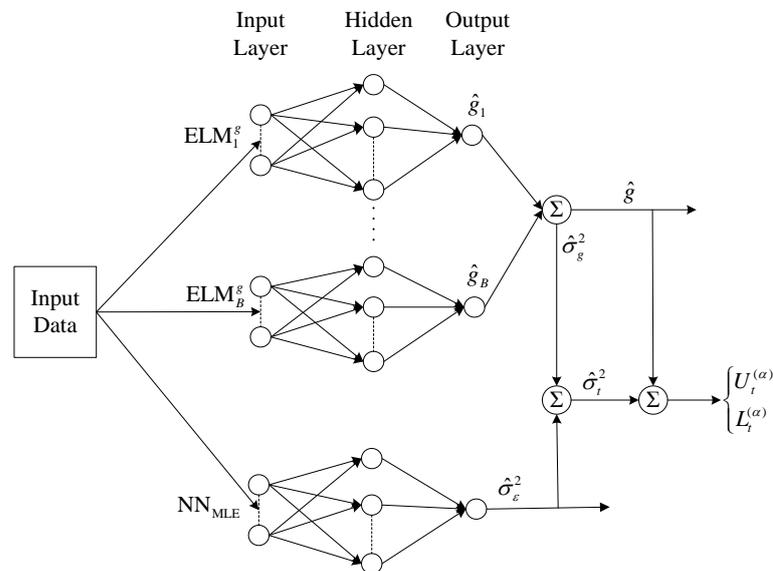
A separate NN model can be indirectly trained to estimate the unknown noise variance  $\hat{\sigma}_\varepsilon^2$ , so as to maximize the probability of the observing samples in  $D_{r^2}$ . Since maximizing a variable is the same as minimizing the negative of that variable, after ignoring the constant part in (8.13), the cost function for training the NN model can be obtained and expressed as

$$C_N = \frac{1}{2} \sum_{i=1}^N \left[ \frac{r^2(\mathbf{x}_i)}{\hat{\sigma}_\varepsilon^2(\mathbf{x}_i)} + \log(\hat{\sigma}_\varepsilon^2(\mathbf{x}_i)) \right] \quad (8.14)$$

The output activation function of the new NN is set to be exponential to ensure that the estimated variance is always positive. The minimization of the cost function  $C_N$  can be reached through a traditional gradient decent method.

### 8.2.4 Overall Procedure

Generally, there are two steps in the construction process of the PIs by the proposed hybrid approach. With the quantified model uncertainty variance and noise variance, the variance of point forecasts uncertainty can be calculated from (8.3). Then the prediction intervals can be derived according to (8.4) and (8.5). The overall framework of the proposed probabilistic forecasting approach is depicted in Figure 8.1. In general, the proposed method needs  $B + 1$  NN models in total to construct the prediction intervals.



**Figure 8.1** The overall framework of PIs construction by the proposed hybrid approach.

## 8.3 Numerical Studies

### 8.3.1 Electricity Market and Experiment Data

The proposed forecast model is tested for Australian National Electricity

Market (ANEM) to validate its effectiveness and efficiency. The ANEM consists of five regional market jurisdictions, including Queensland, New South Wales, Victoria, South Australia and Tasmania. The Australian Energy Market Operator (AEMO) operates the whole power system [114].

In ANEM, the electricity trading is settled down based on a half hour trading interval. Generators submit their offers, and then are dispatched with corresponding dispatch prices every five minute. Market clearing price is calculated by averaging the six consecutive 5-minute dispatch prices for each half hour trading interval, based on the bids and offers of scheduled generators and consumers. A separate spot price is determined in this way for each of the five regions of ANEM. To help electricity producers and consumers for their decision making, day-ahead 48 MCPs in total are concerned and required to be forecasted for the coming trading day.

The New South Wale (NSW) regional market in Australia Electricity Market is employed in the study. The electric energy price and demand data applied cover the period from January 2007 to December 2009. Since MCPs of each trading interval demonstrate similar properties reflecting the variations of demand and operation restrictions, the proposed model is constructed separately for each trading interval.

### **8.3.2 Comparison Analysis**

To evaluate the forecast performances of the proposed approach, we compare the results with other three benchmarks, the persistence (random walk) approach, bootstrap based traditional NNs (BNN) approach, and ELM-Bootstrap approach [118]. The persistence forecast is a simple way of producing a time series forecast, and is widely used as a benchmark in forecasting research, especially for chaotic weather prediction [133]. This approach simply takes the forecasted electricity price of day  $d$  exactly equivalent to the electricity price of day  $d-1$ . The variance of forecasting errors can be obtained based on the latest observations. The BNN approach is newly applied for electricity price forecasting in this chapter, and has been proved effective for reliable PIs construction [110]. Different from the proposed hybrid approach, the BNN approach applies the traditional NNs to implement the PIs construction. Due to

the non-linear regression ability of NNs, the BNN approach can have relatively satisfactory performance in forecasting electricity price but with high computing costs. In addition, a mature ELM-Bootstrap approach combining ELM and bootstrap proposed in [118] is also employed to validate the effectiveness of the developed hybrid approach.

**Table 8.1** Evaluation results of PIs on MCP<sub>1</sub>

NCP	Method	ECP	ACE	Score
90%	BNN	94.41%	4.41%	-3.06
	Persistence	87.11%	-2.89%	-4.33
	ELM-Bootstrap	87.19%	-2.81%	-10.47
	Hybrid Method	89.11%	-0.89%	-2.75
95%	BNN	97.21%	2.21%	-1.79
	Persistence	91.41%	-3.59%	-2.64
	ELM-Bootstrap	94.06%	-0.94%	-5.66
	Hybrid Method	94.13%	-0.87%	-1.63
99%	BNN	99.32%	0.32%	-0.49
	Persistence	95.70%	-3.30%	-0.85
	ELM-Bootstrap	97.81%	-1.19%	-1.39
	Hybrid Method	97.77%	-1.23%	-0.48

**Table 8.2** Evaluation results of PIs on MCP<sub>2</sub>

NCP	Method	ECP	ACE	Score
90%	BNN	89.63%	-0.37%	-8.39
	Persistence	82.91%	-7.09%	-125.52
	ELM-Bootstrap	93.78%	3.78%	-28.26
	Hybrid Method	91.29%	1.29%	-7.56
95%	BNN	91.70%	-3.30%	-5.34
	Persistence	84.87%	-10.13%	-118.01
	ELM-Bootstrap	97.10%	2.10%	-15.71
	Hybrid Method	93.36%	-1.64%	-4.74
99%	BNN	95.85%	-3.15%	-1.99
	Persistence	86.83%	-12.17%	-105.83
	ELM-Bootstrap	99.59%	0.59%	-3.69
	Hybrid Method	96.27%	-2.73%	-1.75

For power system and electricity market operation, it is more preferable to have forecasted information of high confidence levels to reduce the risks. Therefore, PIs of electricity price at high confidence levels including 90%, 95% and 99% are obtained and analyzed in the study. MCPs of several typical trading intervals are selected to test the forecasting models, including the trading interval 04:00-04:30 (MCP<sub>1</sub>), the trading interval 12:00-12:30 (MCP<sub>2</sub>), the trading

interval 17:00-17:30 (MCP<sub>3</sub>), and the trading interval 23:00-23:30 (MCP<sub>4</sub>). The selected trading intervals involve the peak load and low load conditions during daily operation of NSW market. The proposed model and benchmarks are tested on the price series of 2009, and the rest data are used for training the models. The evaluation results including ECP, ACE and interval score of PIs obtained by the proposed hybrid approach and the other three benchmark models over the four typical MCPs(1-4) are given in Tables 8.1-8.4.

**Table 8.3** Evaluation results of PIs on MCP<sub>3</sub>

NCP	Method	ECP	ACE	Score
90%	BNN	93.75%	3.75%	-7.37
	Persistence	79.55%	-10.45%	-90.98
	ELM-Bootstrap	87.89%	-2.11%	-27.68
	Hybrid Method	91.80%	1.80%	-7.08
95%	BNN	97.66%	2.66%	-4.44
	Persistence	82.35%	-12.65%	-84.83
	ELM-Bootstrap	93.36%	-1.64%	-15.89
	Hybrid Method	95.70%	0.70%	-4.44
99%	BNN	98.83%	-0.17%	-1.56
	Persistence	84.87%	-14.13%	-76.85
	ELM-Bootstrap	98.05%	-0.95%	-4.37
	Hybrid Method	98.05%	-0.95%	-1.66

**Table 8.4** Evaluation results of PIs on MCP<sub>4</sub>

NCP	Method	ECP	ACE	Score
90%	BNN	96.43%	6.43%	-4.55
	Persistence	82.82%	-7.18%	-18.86
	ELM-Bootstrap	89.78%	-0.22%	-16.54
	Hybrid Method	89.29%	-0.71%	-3.81
95%	BNN	97.73%	2.73%	-2.79
	Persistence	85.90%	-9.10%	-15.85
	ELM-Bootstrap	96.00%	1.00%	-9.01
	Hybrid Method	96.43%	1.43%	-2.43
99%	BNN	99.03%	0.03%	-0.89
	Persistence	89.45%	-9.55%	-12.55
	ELM-Bootstrap	100%	1.00%	-2.06
	Hybrid Method	97.73%	-1.27%	-0.89

As seen in Tables 8.1-8.4, the proposed hybrid method consistently outperforms the persistence method, the BNN method, and the ELM-Bootstrap approach. ECPs of the hybrid method are significantly close to the corresponding nominal confidence NCPs. All the absolute ACEs of the hybrid method are

smaller than 3%, indicating a satisfactory reliability higher than the other three approaches. For instance, at the confidence level 95% of MCP<sub>3</sub>, the proposed method generates ACE 0.70%, which significantly outperforms the other three methods, especially for the persistence method. According to the interval scores expressed in Tables 8.1-8.4, the hybrid method shows the largest scores in general, indicating the best sharpness and overall skill of the constructed PIs. E.g. for confidence level 90% of MCP<sub>2</sub>, the hybrid approach produces interval score -7.56, larger than that of the other three approaches. From Tables 8.1-8.4, the ELM-Bootstrap approach can obtain satisfactory reliability. However, it demonstrates much lower sharpness and overall skill than the proposed hybrid approach. Particularly, considering PIs with NCP 90% of the MCPs(1-4), the interval scores of the proposed hybrid method are about four times larger than that of the ELM-Bootstrap approach, indicating the significantly higher overall skill and sharpness. Considering both reliability and overall skill, the proposed hybrid approach provides the best PIs with respect to comprehensive performance comparing with the other three benchmarks.

The persistence method is a simple model for short-term time series forecasting and demonstrates much lower performance than the hybrid approach from the perspectives of both reliability and sharpness. This can be explained as it cannot model the nonstationarity and heteroscedasticity of the electricity price series. By contrast, the ELM-Bootstrap approach can obtain PIs with relatively high reliability as the proposed approach. However, it cannot ensure satisfactory sharpness since it does not include the data noise to formulate PIs. This study result also confirms the importance of sharpness, as an indispensable aspect for the quality of constructed PIs, echoing PIs evaluation indices introduced in Section 3.3.

In general, the BNN approach has relatively lower performance comparing with the proposed hybrid approach, especially with respect to the sharpness of PIs. We can find that the BNN approach also can obtain much better forecasting results than the persistence method and ELM-Bootstrap approach. Due to the nonlinear mapping capability of NNs, it should not be a surprise that the BNN approach can have relatively comparable performance with the proposed hybrid approach in some cases according to PIs evaluation. However, the traditional

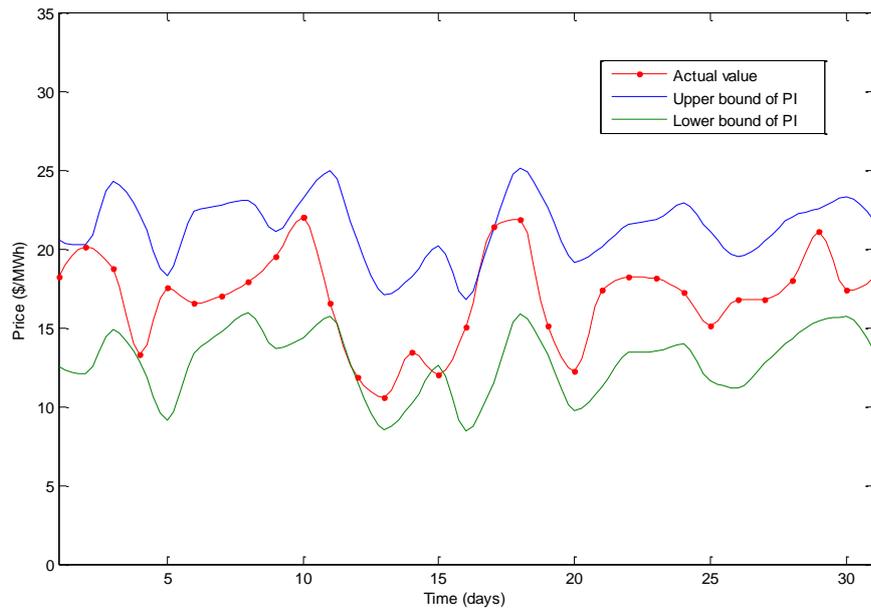
NNs have the significant disadvantage of high computational burden. In the case study, the bootstrap replicates number  $B$  is set to 100 for both the hybrid approach and the BNN approach. The two approaches are implemented using a PC with Intel Core Duo i5-3470T CPU @2.9GHz CPU and 8.0G RAM. The average training time of the BNN method and the hybrid method tested on the MCPs within one specific trading interval is compared in Table 8.5.

Table 8.5 shows that the proposed hybrid approach performs more than 100 times faster than the BNN method, demonstrating a significantly higher efficiency. Considering all the 48 MCPs, the BNN method averagely requires 48 times the computation time shown in Table 8.5, which indicates a significantly heavy computation burden and the infeasibility for realistic applications. With the fast computation speed, the hybrid method has high potential for practical applications in the future.

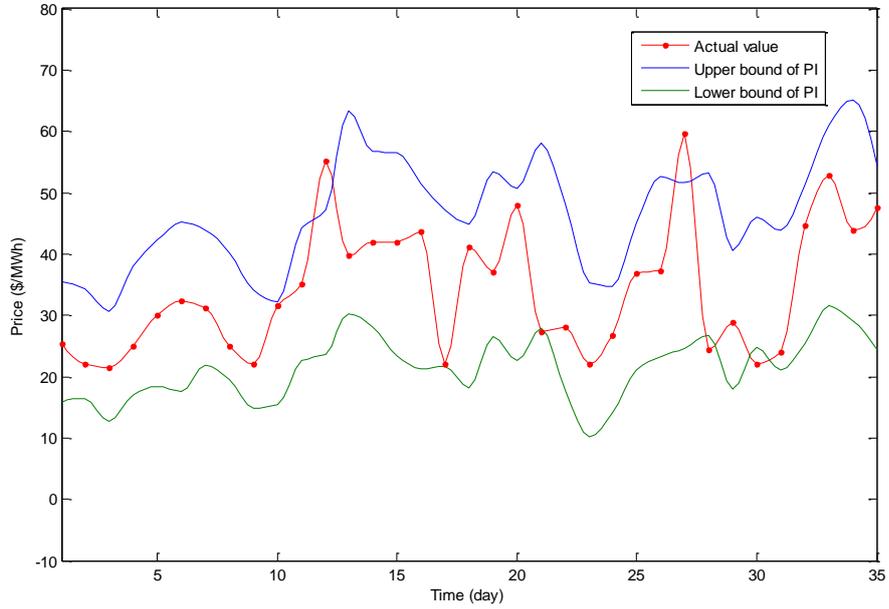
**Table 8.5** Training time comparison

Method	Time (s)
Hybrid method	36.70
BNN	4725.03

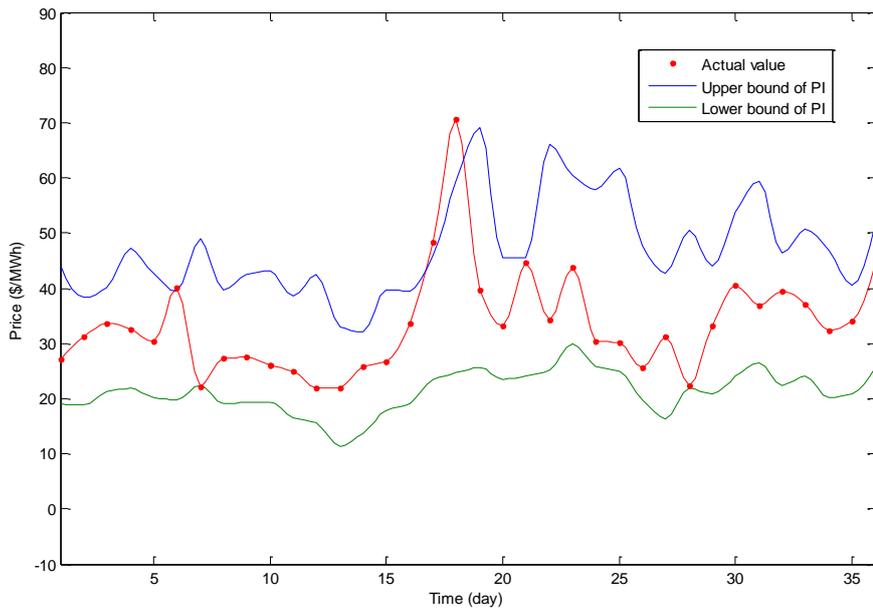
PIs with nominal confidence 90% obtained by the proposed method and the corresponding actual MCPs in the four tested trading intervals crossing around 30 trading days are displayed in Figures 8.2-8.5 respectively. It can be seen that the actual MCPs are mostly well enclosed by the constructed PIs for all the tested trading intervals. We also can observe that in the different trading intervals the electricity price exhibits fairly different variations. Figures 8.2-8.5 explicitly demonstrate the high performance of the proposed hybrid approach.



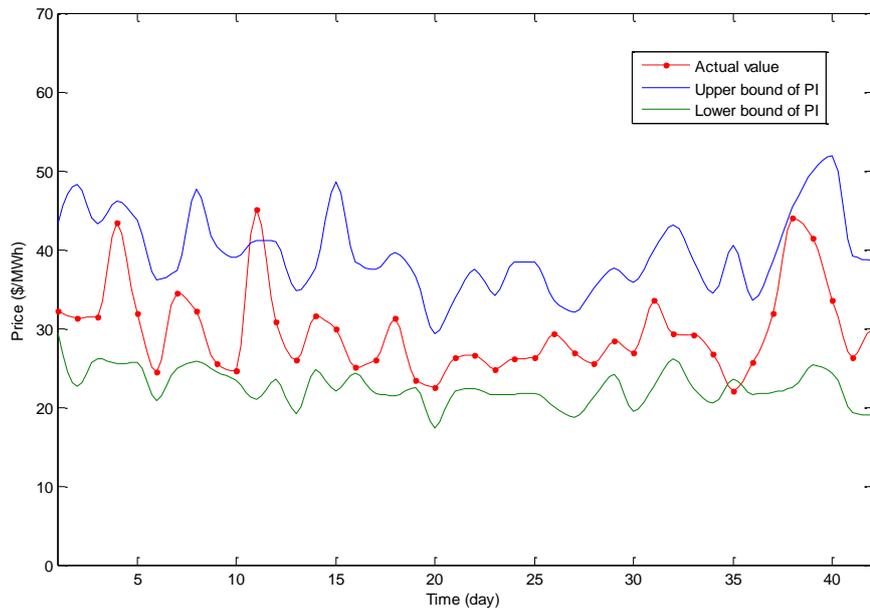
**Figure 8.2** PIs with nominal confidence 90% of  $MCP_1$  obtained by the proposed hybrid approach.



**Figure 8.3** PIs with nominal confidence 90% of  $MCP_2$  obtained by the proposed hybrid approach.



**Figure 8.4** PIs with nominal confidence 90% of  $MCP_3$  obtained by the proposed hybrid approach.



**Figure 8.5** PIs with nominal confidence 90% of  $MCP_4$  obtained by the proposed hybrid approach

According to the case studies, the proposed hybrid approach has satisfactory performance for probabilistic interval forecasting of electricity price. The experiment results also prove the necessity and importance of sharpness in our proposed PIs evaluation framework. With the fast training speed, the hybrid

approach can save computation efforts significantly, thus potentially can be implemented as an efficient online tool. Meanwhile, the hybrid method has high flexibility due to the nonlinear mapping ability of the ELM. In our case study, we take the historical MCPs and demands as the only inputs to the proposed model. In fact, the electricity price can also be affected by other factors, such as weather conditions, maintenance schedules, etc. To obtain optimal prediction results, all the relevant factors should be taken into account in the electricity price forecasting modeling. With the high flexibility of the proposed hybrid approach, extensions to have additional inputs can be conducted easily to improve the forecast performance further. The proposed approach can potentially provide an efficient and meaningful tool for different participants in power systems to assist various decision making activities.

## **8.4 Conclusions**

The electricity market plays a critical role in realizing the economic benefits of smart grids. Electricity price forecasting is crucial to electricity market operation and decision making by electricity market participants. However, electricity price forecasting errors are unavoidable to some extent due to the nonstationarity of electricity price series. Traditional NNs based forecasting approaches cannot provide satisfactory performance. In this chapter, a hybrid approach combining extreme learning machine and maximum likelihood estimation is proposed for probabilistic forecasting of electricity price. A bootstrap based approach is used to quantify the model misspecification uncertainties, and maximum likelihood estimation is applied to estimate the residuals noise. Reliable PIs can be computed by combining regression model uncertainty and residual noise. A comprehensive and systematic evaluation framework for probabilistic forecasting of electricity price is proposed in this chapter, where both the reliability and sharpness are considered in the PIs evaluation process. Numerical studies based on the practical electricity market data in Australian electricity market have been conducted with satisfactory results achieved. Due to the extremely fast learning speed, the proposed hybrid approach can be more than hundred times faster than bootstrap-based traditional NNs approach, which indicates its high potential for online application in future smart grids.



# **9** *Probabilistic Load Flow Considering the Uncertainty of Wind Power*

## **9.1 Introduction**

Power flow computation is the foundation for power system steady analysis and essential in all kinds of operation and planning analysis, such as expansion and operational planning. Recent changes to power system, such as introduction of electricity markets, open access of transmission systems, renewable energy sources and flexible operation according to the smart grid concept have caused many unpredictable operating uncertainties. These uncertainties pose much more risks and challenges to power system operation and expansion planning [22, 245].

Traditional deterministic load flow (DLF) is effectively used in power system steady-state analysis for a long time. However, DLF is not sufficient for modern power system planning study, where emerging uncertainties due to e.g. wind generation need to be considered. Specific values of generation, loads, and network configurations should be given for DLF computation. Due to increased use of renewable energies and other emerging technologies, substantial uncertainties are introduced into modern power system operation that have to be addressed in especially different operation and planning stages.

Therefore, probabilistic approaches are introduced to power flow analysis to consider uncertainties in power system. Probabilistic load flow (PLF) was first proposed by Borkowska in 1974 [246]. The probabilistic load flow (PLF) deals with the stochastic loads, generation and network configurations in power system, and can be applied in both real-time operation [247, 248], long term transmission

expansion planning [249-251], transmission loss evaluation [252], voltage quality assessment [253, 254], voltage and reactive power control [255, 256], etc. There are several techniques developed to cope with PLF problems under uncertainties. Usually the PLF problems are solved by three main techniques: Monte Carlo simulation (MCS), analytical approaches, and approximate approaches.

Monte Carlo approaches are considered as the simplest approach dealing with PLF problem and widely used in PLF computation [257] and reliability assessment [258]. MCS solves the PLF problems by a series of deterministic routines. Theoretically, meaningful and accurate results can be obtained without any constraint if the limitation of computational hardware is neglected. Therefore, the results obtained by MCS are always set as comparison reference for developing other PLF approaches. However, the major disadvantage of the MCS is the requirement of large computational efforts. Generally, thousands of simulations are needed to obtain sufficiently accurate and reliable results. To overcome this limitation, other advanced sampling approaches such as Latin hypercube sampling are proposed for PLF approximation to reduce the size of sampling dataset [259-261].

In order to improve the efficiency, analytical methods and approximate methods are proposed for PLF computation. Basic convolution techniques [262, 263] and Fast Fourier Transform (FFT) [264] are used to solve the PLF problems. The two methods are computationally intensive and require large data storage. A PLF method considering uncertainties of bus power injection is proposed in [265] to employ cumulants and Gram-Charlier expansion to obtain the PDF and CDF of branch power flows. The computation time is greatly reduced. Maximum Entropy (ME) methods are proposed for PDF reconstructions of PLF analysis of large-scale networks and demonstrate statistically accurate and computationally efficient, better than Gram-Charlier (GC) approach [266]. The main disadvantages of the analytical approaches are that they require much more assumptions and complicated mathematical computations. Method combining analytical approaches and MCS techniques is introduced in [267] to enhance the accuracy of load flow results and improve the computational efficiency by means of simplifying the computational process. PLF algorithms combining MCS with

multi-linearized power flow equations are proposed to reduce computational burden and efficiently evaluate the power flow results [268, 269].

Besides, point estimate method (PEM) has been proposed to solve the PLF problems and obtain approximate probability distribution of the random output without complicated mathematical computations [270, 271], which has been proved to be efficient and accurate. In comparisons to analytical approaches, PEM has the merits that it uses deterministic routines for PLF computation and requires relatively less information of the input random variables. The two-point estimate method has been used as an approximate method for probabilistic analysis proposed by Rosenblueth [272, 273]. However, when Rosenblueth's technique is applied to solving PLF problems in power system, the computational burden would be very heavy for the large number of input variables. Then, a new and efficient point estimate method proposed by Hong [274] could be used directly without huge iterations and complicated mathematical transformation.

In general, uncertain factors such as generation injections, load demands, line parameters, and network configuration are taken into account in PLF computation problems [265, 270, 271, 275]. The dramatic growth of intermittent renewable energy generation, including wind power, photovoltaic generation, etc., and relatively unpredictable probabilistic loads such as plug-in hybrid electric vehicles (PEVs) introduce more and more uncertainties and increase the complexity of modern power systems. PLF computation considering the uncertainty of wind power is investigated in [276-278]. The integration of photovoltaic resources in transmission or distribution systems are considered for PLF computation [279-281]. The uncertainties brought by PEVs are also estimated in [282, 283].

As described in the preceding chapters, wind power has become a crucial uncertainty to power systems. In Chapter 2, the stochastic wind speed is modeled by means of the proposed generalized Lambda distribution. It will be applied for formulating the uncertainty of wind power in PLF analysis. In addition, the uncertainties of loads and generators are also considered in the study. In this chapter, point estimate method is applied in the study to estimate the mean and variance of the voltage magnitude and angle, branch flow. To investigate influences of wind power, wind farm is connected to power systems at different

nodes. Through comprehensive case studies, the effectiveness of the PEM method is benchmarked with first-order second-moment method (FOSM) and Monte Carlo simulation. Efficient and accurate PLF computation can be helpful for power system operation and planning considering the large integration of wind power generation.

## 9.2 Load Flow Solutions

### 9.2.1 AC Load Flow

If the active and reactive power at each bus and the node admittance matrix are specified, the injection power of each bus can be given as:

$$P_i + jQ_i = U_i \sum_{j=1}^n Y_{ij}^* U_j^* \quad (9.1)$$

The complex equation can be represented as two real equations in terms of two real variables instead of one complex variable [2]. Therefore,  $P$  and  $Q$  at each bus are functions of voltage magnitude  $V$  and angle  $\delta$  of all buses. The non-linear load flow equation could be expressed as:

$$P_i = V_i \sum_{j \in i} V_j (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) \quad (9.2)$$

$$Q_i = V_i \sum_{j \in i} V_j (G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}) \quad (9.3)$$

where  $\delta_{ij}$  denotes the voltage angle difference between the voltages at bus  $i$  and  $j$ .

The application of Newton-Raphson method can be described as

$$\begin{bmatrix} \Delta \mathbf{P} \\ \Delta \mathbf{Q} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{P}}{\partial \boldsymbol{\delta}} & \frac{\partial \mathbf{P}}{\partial \mathbf{V}} \\ \frac{\partial \mathbf{Q}}{\partial \boldsymbol{\delta}} & \frac{\partial \mathbf{Q}}{\partial \mathbf{V}} \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{\delta} \\ \Delta \mathbf{V} \end{bmatrix} = \mathbf{J} \begin{bmatrix} \Delta \boldsymbol{\delta} \\ \Delta \mathbf{V} \end{bmatrix} \quad (9.4)$$

The line flow can be obtained based on the function of voltage magnitudes and angles,

$$P_{ij} = V_i V_j (G_{ij} \cos \delta_{ij} + B_{ij} \sin \delta_{ij}) - k_{ij} G_{ij} V_i^2 \quad (9.5)$$

$$Q_{ij} = V_i V_j (G_{ij} \sin \delta_{ij} - B_{ij} \cos \delta_{ij}) + (k_{ij} B_{ij} + B_{ij0}) V_i^2 \quad (9.6)$$

where  $P_{ij}$  and  $Q_{ij}$  are the active and reactive power flows at line from bus  $i$  to bus  $j$ ;  $k_{ij}$  is transformer off nominal turns ratio;  $B_{ij0}$  equals 1/2 of total line charging susceptance.

### 9.2.2 Linearized Load Flow Equation

In general, load flow computation can be expressed by two sets of nonlinear equations.

$$\mathbf{S} = \mathbf{g}(\mathbf{X}) \quad (9.7)$$

$$\mathbf{Z} = \mathbf{h}(\mathbf{X}) \quad (9.8)$$

where  $\mathbf{S}$  is the input vector of real and reactive power injections;  $\mathbf{Z}$  is the output vector of line flows;  $\mathbf{X}$  is the state vector of nodal voltage magnitudes and angles;  $\mathbf{g}$  and  $\mathbf{h}$  are nodal power injection and branch flow functions, respectively.

In the probabilistic load flow study, the input variables  $\mathbf{S}$  are known and considered probabilistically distributed. In this paper, uncertainties of PLF study mainly include random load variations and unit forced outages.

Assuming total independence of all input variables, the conventional load flow equations can be linearized around the expected operating point by Taylor series expansion with neglecting the terms of higher powers.

$$\mathbf{X} = \mathbf{X}_0 + \mathbf{J}_0^{-1} \Delta \mathbf{S} \quad (9.9)$$

$$\mathbf{J}_0 = \left. \frac{\partial \mathbf{g}(\mathbf{X})}{\partial \mathbf{X}} \right|_{\mathbf{x}=\mathbf{x}_0} \quad (9.10)$$

where  $\mathbf{X}_0$  is expected value of  $\mathbf{X}$ .

Then line flows can be given by function of nodal voltage state vector.

$$\mathbf{Z} = \mathbf{Z}_0 + \mathbf{H}_0 \Delta \mathbf{X} \quad (9.11)$$

$$\mathbf{H}_0 = \left. \frac{\partial \mathbf{h}(\mathbf{X})}{\partial \mathbf{X}} \right|_{\mathbf{x}=\mathbf{x}_0} \quad (9.12)$$

Then substitute (9.9) into (9.11), line flows are expressed as linearized functions in terms of power injections.

$$\mathbf{Z} = \mathbf{Z}_0 + \mathbf{R}_0 \Delta \mathbf{S} \quad (9.13)$$

$$\mathbf{R}_0 = \mathbf{H}_0 \mathbf{J}_0^{-1} \quad (9.14)$$

The analytical method applied in probabilistic load flow solution become feasible as the traditional load flow equations have been linearized around the expected value.

### 9.3 Model of Input Uncertainty

#### 9.3.1 Injected Power

The injected power considered here is generated by other power plants except for wind generators, and is the key input of power flow computation. On basis of the formulation in [284-286], a discrete distribution can be applied to describe the traditional generation power involved in PLF analysis. Accordingly, the injected power of one generator can be one of fixed finite  $n$  possible outputs. The PDF of the injected power is defined by the following equation,

$$f_G(x) = \sum_{i=1}^n \tau_i \delta(x - \eta_i) \quad (9.15)$$

where  $\delta(\cdot)$  is Dirac Delta function;  $x$  represents the injected power; and  $\tau_i$  denotes the probability of the power to be equal to  $\eta_i$ . The unit outage uncertainty is also included in the finite scenarios of the formulated discontinuous distribution.

#### 9.3.2 Bus Load

The bus loads in the PLF study can be formulated as normal distribution as [265]. The PDF and CDF can be expressed as

$$f_L(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (9.16)$$

$$F_L(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \quad (9.17)$$

where  $x$  denotes the bus load, and  $\mu$  and  $\sigma$  denote the mean and variance of the normal distribution respectively.

In practice, the PDF and CDF could be obtained from the statistical analysis of the historical measurement data to compute the mean  $\mu$  and standard deviation  $\sigma$  of the bus loads.

### 9.3.3 Wind Power Output

To conduct PLF analysis of power system with integration of wind energy, the statistics of wind power is formulated by a quadratic model to describe wind power as a function of wind speed, through combining the knowledge of the PDF of wind speed and the power curve of the specific type of wind turbines. In this chapter, the formulation of wind power will be involved in different simulation methods, including the MCS technique, and the analytical FOSM approach, and the PEM approach.

#### 1) Wind Speed

Traditionally, the Weibull and Rayleigh distributions are used to approximate the stochastic behavior of wind speed in the process of PLF computation. As the quantitative analysis given in Chapter 2 of this thesis, Weibull and Rayleigh distribution cannot adaptively and accurately describe the stochastic process of wind speed. Generalized Lambda distribution is applied to formulate the PDF of wind speed in the study. According to GLD defined in (2.22), the  $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$  for wind speed distribution can be expressed as

$$v = \lambda_1 + \frac{u^{\lambda_3} - (1-u)^{\lambda_4}}{\lambda_2} \quad (9.18)$$

where  $0 \leq u \leq 1$  represents cumulative frequency,  $v$  denotes the wind speed value satisfying cumulative frequency  $F(v) = u$ .

Then, the PDF of wind speed can be expressed as

$$f(v) = \frac{\lambda_2}{\lambda_3 u^{\lambda_3-1} + \lambda_4 (1-u)^{\lambda_4-1}} \quad (9.19)$$

The mean value  $\mu$  and variance  $\sigma^2$  of wind speed are defined in (2.29) and (2.30) respectively.

#### 2) Power Output of Wind Turbine

The typical power curve of wind turbine is shown in Figure 4.1. Since wind speed is treated as a random variable, the wind power is also characterized as a stochastic variable with respect to wind speed. As expressed in [287], the quadratic model of power curve function of a wind turbine can be simplified into a piecewise linear function, defined by the following equation,

$$P_w = \begin{cases} 0, & \text{if } v < v_{in} \text{ or } v > v_{out} \\ P_r \frac{v - v_{in}}{v_r - v_{in}}, & \text{if } v_{in} \leq v \leq v_r \\ P_r, & \text{if } v_r \leq v \leq v_{out} \end{cases} \quad (9.20)$$

where  $P_w$  denotes the output wind power of a wind turbine,  $v_{in}$  is the cut-in wind speed,  $v_r$  is the rated wind speed,  $v_{out}$  is the cut-out wind speed, and  $P_r$  is the rated power of a single wind turbine.

In the study, the wind turbines are considered as Vestas V80 with rated power 2.0 MW given in Chapter 4, which has been installed at the wind farm on Bornholm Island in Denmark. The PDFs and CDFs of wind power output of a single wind turbine can be represented on basis of the model of wind speed probability distribution and power curve of the wind turbines. In the study, the impacts of the fluctuating wind power on the transmission systems are focused to investigate the resultant probabilistic characteristics of load flows. The active power of the large wind farm  $P_w$  integrated to the power system is the sum of the active power generated by each single wind turbine. Ignoring the differences between each wind turbine, the large wind farm can be formulated via an aggregated model that is simply characterized equivalent to a single wind turbine.

## 9.4 Mathematical Background

### 9.4.1 Monte Carlo Simulation

Let a random variable  $Y$  denote the function of stochastic variable  $x$  with probabilistic density function  $p(x)$ ,

$$Y = f(x) \quad (9.21)$$

Monte Carlo simulation method will generate random samples  $x_1, x_2, \dots, x_n$  according to its probability density function  $p(x)$  to estimate the probability distribution of  $Y$ . Based on the simulated results, mean and variance of  $Y$  can be expressed as

$$E(Y) = \frac{1}{N} \sum_{i=1}^N f(x_i) \quad (9.22)$$

$$\text{var}(Y) = \frac{1}{N-1} \sum_{i=1}^N (f(x_i) - E(Y))^2 \quad (9.23)$$

Generally, the estimation process of MCS is extremely time-consuming and computationally demanding.

#### 9.4.2 First-order Second-moment Method

The FOSM method applies the first-order terms of the Taylor series expansion about the expectation of each input stochastic variable [288]. In addition, the second moments of the uncertain variables are required for the mean and variance estimation. The FOSM method is an efficient approximation method to estimate the mean and variance of any targeted random variable with high speed and accuracy, and is widely used in engineering analysis [289-291].[285]

In general, the integrals, as in (9.24) and (9.25), are always used to compute the mean and variance if a variable is continuously distributed with a known PDF  $p(x)$ . However, the computation process of integrals is particularly time-consuming.

$$E(Y) = \int_{-\infty}^{+\infty} f(x)p(x)dx \quad (9.24)$$

$$\text{var}(Y) = \int_{-\infty}^{+\infty} [f(x) - E(Y)]^2 p(x)dx \quad (9.25)$$

If several random variables are included, the convolution method has to be applied, which will make the computation much more complicated. Without loss of generality, a function of several random variables  $x_1, x_2, \dots, x_n$  is considered.

$$Y = f(x_1, x_2, \dots, x_n) \quad (9.26)$$

The function could be expanded using Taylor series around the value  $x_{0_i}$  ( $i = 1, 2, \dots, n$ ):

$$Y = f(x_{0_1}, x_{0_2}, \dots, x_{0_n}) + \sum_{i=1}^n (x_i - x_{0_i}) \left. \frac{\partial f}{\partial x_i} \right|_{\mathbf{x}_0} + \frac{1}{2} \sum_{i=1}^n \left( (x_1 - x_{0_1}) \frac{\partial}{\partial x_1} + (x_2 - x_{0_2}) \frac{\partial}{\partial x_2} + \dots + (x_n - x_{0_n}) \frac{\partial}{\partial x_n} \right)^2 f \quad (9.27)$$

When the terms with orders higher than two are neglected in the Taylor series expansion (9.27), the general nonlinear function can be linearized and expressed as.

$$Y \approx f(x_{0_1}, x_{0_2}, \dots, x_{0_n}) + \sum_{i=1}^n (x_i - x_{0_i}) \left. \frac{\partial f}{\partial x_i} \right|_{\mathbf{x}_0} \quad (9.28)$$

Expanding the function in a Taylor series about the mean values  $\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}$  of these variables  $x_1, x_2, \dots, x_n$ , the function in (9.26) can be transformed to

$$Y \approx f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) + \sum_{i=1}^n (x_i - \mu_{x_i}) \left. \frac{\partial f}{\partial x_i} \right|_{\boldsymbol{\mu}_x} \quad (9.29)$$

According to (9.29), the mathematical expectation of target variable  $Y$  can be taken as the following expression:

$$E(Y) = f(\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}) \quad (9.30)$$

The variance of  $Y$  is determined by variances of input variables and their covariance of input variables by pair.

$$\begin{aligned} \text{var}(Y) &= E \left[ \left( \sum_{i=1}^n (x_i - \mu_{x_i}) \left. \frac{\partial f}{\partial x_i} \right|_{\boldsymbol{\mu}_x} \right)^2 \right] \\ &= \sum_{i=1}^n \left( \left. \frac{\partial f}{\partial x_i} \right|_{\boldsymbol{\mu}_x} \right)^2 \text{var}(x_i) + 2 \sum_{i=1}^n \sum_{j \neq i}^n \left( \left. \frac{\partial f}{\partial x_i} \right|_{\boldsymbol{\mu}_x} \right) \left( \left. \frac{\partial f}{\partial x_j} \right|_{\boldsymbol{\mu}_x} \right) \text{cov}(x_i, x_j) \end{aligned} \quad (9.31)$$

where  $\text{cov}(x_i, x_j)$  denotes the covariance between variables  $x_i$  and  $x_j$ , given by the expression:

$$\text{cov}(x_i, x_j) = E[(x_i - \mu_{x_i})(x_j - \mu_{x_j})] \quad (9.32)$$

The second term in (9.31) expresses the correlation among the input variables. If the input statistical variables are mutually independent, the second term will become zero. Then the variance of  $Y$  becomes a function of the variance of input variables.

$$\text{var}(Y) = \sum_{i=1}^n \left( \left. \frac{\partial f}{\partial x_i} \right|_{\mu_{\mathbf{x}}} \right)^2 \text{var}(x_i) \quad (9.33)$$

Therefore, the standard deviation could be given as

$$\sigma_Y = \sqrt{\sum_{i=1}^n \left( \left. \frac{\partial f}{\partial x_i} \right|_{\mu_{\mathbf{x}}} \right)^2 \sigma_{x_i}^2} \quad (9.34)$$

In conclusion, the FOSM method can yield the expected value and standard deviation of the random variable in consideration through simple calculation. This method allows for estimating the stochastic property of the objective variable without the knowledge of specific PDF of the input variables beforehand. The computation effort is greatly reduced.

### 9.4.3 Point Estimate Method

Generally, the PEM approach is regarded as a weighting method rather than a probability distribution transformation method. Theoretically, the point estimate method concentrates the statistical information described by the first few central moments of input random variables on points of each variable determined by mathematical formulation. The information of probabilistic distribution of the output random variables can be provided by the corresponding central moments, which are estimated based on the concentrated points. In this study, Hong's efficient point method [274] is employed to solve PLF problems. PEM proposed by Hong is an extension of Rosenblueth's two-point concentration method [272, 273] and significantly improve the computation efficiency.

Let  $Z$  represents a random variable which is a function of  $n$  random variables  $x_k$ , the relationship between  $Z$  and  $x_k$  can be expressed by

$$Z = h(x_1, x_2, \dots, x_k, \dots, x_{n-1}, x_n) \quad (9.35)$$

The point estimate method in this paper uses  $m \times n$  point concentrations to obtain the probabilistic distribution of  $Z$  and  $m$  is the number of concentrations for each random variable  $x_k$ . Let  $\xi_{k,i}$  denote the standard location, the  $m$  concentrations  $(x_{k,i}, \xi_{k,i})$  of the  $n$  input random variables

$(x_1, x_2, \dots, x_k, \dots, x_{n-1}, x_n)$  can be obtained from the probability distribution function of each random variable  $x_k$ .

$$x_{k,i} = \mu_{x_k} + \xi_{k,i} \sigma_{x_k}, \quad k = 1, 2, \dots, n, \quad i = 1, 2, \dots, m. \quad (9.36)$$

where  $\mu_{x_k}$  and  $\sigma_{x_k}$  are the mean and standard deviation of random variable  $x_k$  with PDF  $f_{x_k}$ .

Let  $\omega_{k,i}$  denote the concentrations (or weights) at the chosen location  $(\mu_{x_1}, \mu_{x_2}, \dots, x_{k,i}, \dots, \mu_{x_{n-1}}, \mu_{x_n})$ . The standard location  $\xi_{k,i}$  and the weight  $\omega_{k,i}$  of  $m$  concentrations of each random variable  $x_k$  are provided by [274], given as

$$\sum_{k=1}^n \sum_{i=1}^m \omega_{k,i} = 1 \quad (9.37)$$

$$\sum_{i=1}^m \omega_{k,i} = \frac{1}{n}, \quad k = 1, 2, \dots, n. \quad (9.38)$$

$$\sum_{i=1}^m \omega_{k,i} (\xi_{k,i})^j = \lambda_{k,j}, \quad k = 1, 2, \dots, n, \quad j = 1, 2, \dots, 2m-1. \quad (9.39)$$

where  $\lambda_{k,j}$  denotes the ratio of  $M'_j(x_k)$  representing the  $j$ th order central moment of random variable  $x_k$  to  $\sigma_{x_k}^j$ .

$$\lambda_{k,j} = \frac{M'_j(x_k)}{\sigma_{x_k}^j}, \quad k = 1, 2, \dots, n, \quad j = 1, 2, \dots, 2m-1. \quad (9.40)$$

$$M'_j(x_k) = \int_{-\infty}^{\infty} (x_k - \mu_{x_k})^j f_{x_k} dx_k \quad (9.41)$$

where  $\lambda_{k,1}$  equals to zero,  $\lambda_{k,2}$  equals to one,  $\lambda_{k,3}$  and  $\lambda_{k,4}$  are the coefficients of skewness and of kurtosis of  $x_k$  respectively. The  $j$ th raw moment of  $Z$  can be approximated by

$$\begin{aligned} \alpha_j = E[Z^j] &\cong \sum_{k=1}^n \sum_{i=1}^m \omega_{k,i} \times [Z_j(k,i)]^j \\ &= \sum_{k=1}^n \sum_{i=1}^m \omega_{k,i} \times [h(\mu_{x_1}, \mu_{x_2}, \dots, x_{k,i}, \dots, \mu_{x_{n-1}}, \mu_{x_n})]^j \end{aligned} \quad (9.42)$$

The standard deviation of the  $Z_i$  can be obtained though

$$\sigma_{Z_i} = \sqrt{\text{var}(Z_i)} = \sqrt{E(Z_i^2) - [E(Z_i)]^2} \quad (9.43)$$

The two-point estimate method, with two concentrations ( $m=2$ ) for each random variable, is applied in the study. Then the standard location  $\xi_{k,i}$  and the weight  $\omega_{k,i}$  can be obtained by solving the equations (9.37)-(9.39) and expressed by

$$\xi_{k,1} = \lambda_{k,3} / 2 + \sqrt{n + (\lambda_{k,3} / 2)^2} \quad (9.44)$$

$$\xi_{k,2} = \lambda_{k,3} / 2 - \sqrt{n + (\lambda_{k,3} / 2)^2} \quad (9.45)$$

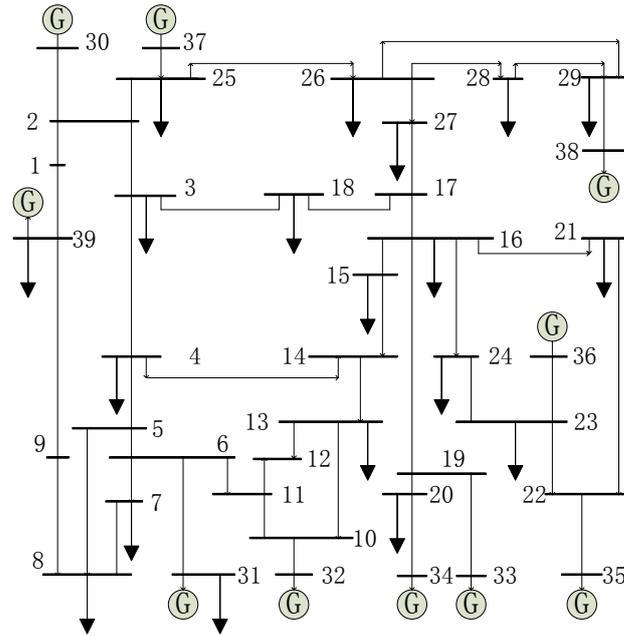
$$\omega_{k,1} = -\frac{1}{m} \frac{\xi_{k,2}}{\xi_{k,1} - \xi_{k,2}} \quad (9.46)$$

$$\omega_{k,2} = \frac{1}{m} \frac{\xi_{k,1}}{\xi_{k,1} - \xi_{k,2}} \quad (9.47)$$

There are  $2n$  concentrations need to be calculated in total. Finally, the PDF and CDF computing the mean and standard deviation of random variable  $Z$  could be obtained from (9.42) and (9.43).

## 9.5 Simulation Test Results

In the study, PLF analysis considering the uncertainty of wind generation is implemented on the IEEE 39 test system consisting of 39 nodes and 46 lines with the base power 100 MVA. The diagram of the electricity network is depicted in Figure 9.1. According to the detailed studies in Chapter 2, the wind speed is modeled with a GLD distribution. The wind farm is designed to be composed by 150 Vestas V80 2.0MW wind turbines with total capacity 300 MW. The cut-in, rated and cut-out wind speeds are 4 m/s, 16 m/s and 25 m/s, respectively. In addition, the uncertainty of other generation and load is also taken into account and following the formulation in Section 9.3. In order to demonstrate the effects of wind power uncertainty, three cases are considered in the study, including Case 1 without integration of wind farm, Case 2 with wind farm connected to node 14, and Case 3 with wind farm connected to node 27.



**Figure 9.1** Diagram of IEEE 39 test system

PLF computation aims at estimate the mean and standard variance of different load flow variables, including active and reactive power flow, voltage magnitude and angle [270, 271, 292]. In the study, the results obtained by MCS approach are reasonably taken as benchmark for the accuracy analysis. In order to verify the accuracy and efficiency of the applied PEM and FOSM approaches, the results obtained are compared with those from Monte Carlo simulation. The MCS approach is implemented with 10000 samples. This number of simulations is large enough to ensure the convergence of Monte Carlo approach. Two performance indices, the relative errors of mean  $\varepsilon_{\mu}$  and standard deviation  $\varepsilon_{\sigma}$ , are used to assess the performance of the applied FOSM and PEM approaches. The two indices are defined as

$$\varepsilon_{\mu} = \left| \frac{\mu^* - \mu_{MCS}}{\mu_{MCS}} \right| \times 100\% \quad (9.48)$$

$$\varepsilon_{\sigma} = \left| \frac{\sigma^* - \sigma_{MCS}}{\sigma_{MCS}} \right| \times 100\% \quad (9.49)$$

where  $\mu_{MCS}$  and  $\sigma_{MCS}$  indicate mean values and standard deviations obtained MCS respectively,  $\mu^*$  and  $\sigma^*$  denote the mean values and standard deviations approximated by PEM and FOSM.

**Table 9.1** Mean and standard deviation of active power flow for Case 1

Line	Index	MCS	FOSM	PEM
1-2	$\mu$	-1.6866	-1.6913	-1.6912
	$\sigma$	0.5392	0.5353	0.5829
4-5	$\mu$	-2.6837	-2.6745	-2.6796
	$\sigma$	0.8806	0.8769	0.8877
5-6	$\mu$	-6.1921	-6.1811	-6.1881
	$\sigma$	1.0933	1.0945	1.1057
6-31	$\mu$	-8.3707	-8.3458	-8.3613
	$\sigma$	2.1135	2.1161	2.1403
10-13	$\mu$	3.5980	3.5892	3.5943
	$\sigma$	0.5528	0.5479	0.5560
17-18	$\mu$	1.7037	1.7130	1.7116
	$\sigma$	0.4630	0.4576	0.4637
23-24	$\mu$	3.3881	3.3898	3.3390
	$\sigma$	0.2549	0.2497	0.2555
26-29	$\mu$	-1.7817	-1.7815	-1.7809
	$\sigma$	0.2799	0.2724	0.2785

The mean and standard deviation of active power flow  $P$  [p.u.] of the selected 8 lines at the “from” bus for Cases 1-3 are listed in Tables 9.1-9.3 respectively, which have been considered as representatives of the general results. From Table 9.1, it can be observed that FOSM and PEM approaches have similar performance for Case 1 without wind farm integration. By comparing the results with MCS, both FOSM and PEM approaches provide relatively acceptable results for Case 1. In Case 2 and Case 3, wind farm is connected to the test system at different nodes. Obviously, the two cases have higher uncertainty than Case 1. The integration of intermittent wind power introduces significant impacts on the majority of system branches, shown by means of the corresponding mean and standard deviation of line flow given in Tables 9.2 and 9.3. Particularly, some branches such as line 23-24 and line 26-29 close to traditional generators are not significantly influenced by the integration of wind farm. The accuracy of FOSM on Cases 2 and 3 decreases significantly comparing with the results of Case 1. For instance, considering Case 2, line 6-31 has an actual mean active power flow -7.5345 p.u. obtained by MCS. The mean value estimated by the

PEM approach is -7.5475 p.u., close to that of MCS. In comparisons, the value estimated by the FOSM method is -6.4029 p.u., demonstrating a significant deviation to the reference value of MCS.

**Table 9.2** Mean and standard deviation of active power flow for Case 2

Line	Index	MCS	FOSM	PEM
1-2	$\mu$	-1.7357	-1.7925	-1.7335
	$\sigma$	0.5377	0.7008	0.5440
4-5	$\mu$	-2.2972	-1.7803	-2.3053
	$\sigma$	0.9766	0.9268	0.9666
5-6	$\mu$	-5.8187	-5.3164	-5.8260
	$\sigma$	1.1733	1.1850	1.1654
6-31	$\mu$	-7.5345	-6.4029	-7.5475
	$\sigma$	2.3052	2.2798	2.2931
10-13	$\mu$	3.2261	2.7264	3.2329
	$\sigma$	0.6752	0.5753	0.6661
17-18	$\mu$	1.7651	1.8370	1.7635
	$\sigma$	0.4686	0.4607	0.4666
23-24	$\mu$	3.3887	3.3898	3.3898
	$\sigma$	0.2562	0.2497	0.2555
26-29	$\mu$	-1.1772	-1.7816	-1.7810
	$\sigma$	0.2841	0.2724	0.2785

**Table 9.3** Mean and standard deviation of active power flow for Case 3

Line	Index	MCS	FOSM	PEM
1-2	$\mu$	-1.8141	-1.9725	-1.8090
	$\sigma$	0.5527	0.6999	0.5550
4-5	$\mu$	-2.2565	-1.6864	-2.2658
	$\sigma$	0.9806	0.9241	0.9815
5-6	$\mu$	-5.7443	-5.1402	-5.7520
	$\sigma$	1.1864	1.1830	1.1897
6-31	$\mu$	-7.5379	-6.4176	-7.5531
	$\sigma$	2.2823	2.2735	2.2881
10-13	$\mu$	3.3375	2.9934	3.3444
	$\sigma$	0.6087	0.5695	0.6096
17-18	$\mu$	1.9778	2.3216	1.9664
	$\sigma$	0.5317	0.4642	0.5311
23-24	$\mu$	3.3880	3.3898	3.3896
	$\sigma$	0.2569	0.2497	0.2555
26-29	$\mu$	-1.7850	-1.7819	-1.7811
	$\sigma$	0.2773	0.2724	0.2785

The overall performance can be measured by means of the average relative error to different types of variables. Table 9.4 shows the average error indices in terms of the estimation results of FOSM and PEM methods applied on the IEEE

39-bus system. The estimation system parameters include active and reactive power flow of all branches, voltage magnitude and angle at all nodes. In general, the PEM approach can derive more accurate results than the FOSM method. For Case 1 without wind farm penetration, FOSM and PEM have relatively close performance. The gap of performance between PEM and FOSM becomes much wider when the wind farm is connected to the system. In Case 3, the average estimation error of mean active power flow by PEM approach reaches about 0.6%, approximately 50 times better than FOSM with average error about 30%. From the aspect of standard deviation of active power, the PEM with 0.5% performs 15 times better than FOSM with 7.3%. In both Cases 2 and 3, the estimation of voltage angle, PEM has average error smaller than 1% and performs much better than FOSM with average error larger than 52%.

**Table 9.4** Average error indices for studied cases

Case	Method	Index	$P_{i,j}$	$Q_{i,j}$	$V$	$\delta$
Case 1	FOSM	$\bar{\varepsilon}_{\mu}$	0.7243	5.9251	0.0624	2.2566
		$\bar{\varepsilon}_{\sigma}$	1.5000	7.8803	12.1214	1.3841
	PEM	$\bar{\varepsilon}_{\mu}$	0.4974	0.2566	0.0006	0.4322
		$\bar{\varepsilon}_{\sigma}$	0.6516	2.5476	3.127	1.2241
Case 2	FOSM	$\bar{\varepsilon}_{\mu}$	11.0162	14.5481	0.1295	53.7988
		$\bar{\varepsilon}_{\sigma}$	6.7087	10.2853	15.4442	3.0396
	PEM	$\bar{\varepsilon}_{\mu}$	0.2044	0.2790	0.0028	0.4058
		$\bar{\varepsilon}_{\sigma}$	0.8034	2.2576	3.0434	0.4788
Case 3	FOSM	$\bar{\varepsilon}_{\mu}$	29.8713	16.9076	0.1373	52.2126
		$\bar{\varepsilon}_{\sigma}$	7.2553	11.7477	16.2279	4.4081
	PEM	$\bar{\varepsilon}_{\mu}$	0.6050	0.3032	0.0015	0.8867
		$\bar{\varepsilon}_{\sigma}$	0.4981	3.3293	4.6567	0.3503

According to the formulation given in Section 9.4.2, FOSM relies on the linearized load flow equation around the operation point. The linear function is obtained by means of neglecting orders higher than two in the Taylor series expansion. It is based on the assumption that the actual variable is not far from the operation condition. Wind power introduces much higher uncertainties than traditional generation power. Higher fluctuations deviating from operation point lead to higher estimation errors of FOSM. It can be found that the PEM approach

yields good results comparing with MCS method for both the mean and standard deviation at all the tested three cases, which also demonstrates its robustness. Similar with MCS, PEM approach employs deterministic routines to resolve probabilistic load flow problems. However, it needs a much lower computation time due to the much smaller size of computation samples. In practice, it is hard even impossible to perfectly determine the probability functions of stochastic variables due to the lack of sufficient information. To some extent, PEM method can avoid this problem involved in the absence of perfect probability functions through using only their first few statistical moments, i.e., mean, variance and skewness, instead of specific quantitative functions. Therefore, less uncertainty information is required for PLF computation.

**Table 9.5** Computation time of different methods

Method	Time (s)
MCS	54.0680
FOSM	0.0312
PEM	0.3276

Table 9.5 gives the computation time needed to calculate the PLF for PEM, FOSM, as well as MCS with 10000 samples. FOSM shows the fastest speed due to the once deterministic load flow computation. PEM also demonstrates significantly satisfactory performance, though more time required than FOSM. Considering both accuracy and computation efficiency, PEM can be an effective for probabilistic load flow analysis of power system with wind power integration. The probabilistic load flow estimated by PEM approach can provide meaningful information for power system planning and operation with large scale integration wind power.

## 9.6 Conclusions

The main theme of the thesis is to conduct uncertainty analysis, modeling and prognosis for power system operation, planning, and etc. In this chapter, probabilistic load flow is studied considering the uncertainty of wind power generation. Computation of power flows is one of the critical tasks for power system operation and planning. More and more uncertainties, such as wind power

generation, have been introduced into power systems. Applying probabilistic approaches in the power flow computation provides more accurate analysis of modern power system. In the study, PLF is computed using two-point PEM and FOSM approaches to estimate the mean and standard deviation of variables. On basis of the case studies on IEEE 39-bus test system, the results of load flow obtained by the PEM approach match the results of MCS very well, performing better than FOSM approach. When the uncertainty has relatively low variance, the FOSM approach indicates fair performance with extremely fast computation speed. In addition, test results indicate that large wind power integration has considerable influences on PLF analysis.

Due to the robust performance, PEM can be an effective approach to accurately and efficiently calculate probabilistic load flow for power system with wind farm integration. Only one wind farm is involved in the study. Multiple wind farms and the corresponding correlation can be considered in the future work. Probabilistic information of load flow provided by the proposed method is meaningful to the power system real time operation and control, operational planning and expansion planning.



# ***10 General Conclusions***

## **10.1 Conclusions**

Power system is considered as one of the most complicated man-made systems. Many uncertainties have been introduced into modern power systems, by renewable energy such as wind power, solar energy, etc., and clearing prices in the electricity market. The interests of smart grid would accommodate these uncertainties in modern power system to ensure the security and economy. This thesis focuses on developing advanced approaches for quantitative analysis, modeling and prognosis of these uncertainties in power systems for assisting to implement proper actions for operation and planning to reduce risks as far as possible.

Wind power is regarded as the most important and efficient renewable energy in modern power systems. From the long-term perspective, accurate modeling the probability distribution of stochastic wind speed would be meaningful for wind farm planning, analysis and planning of power systems with wind power penetration, etc. Traditionally, Weibull distribution and Rayleigh distribution are popularly applied for describing the stochastic behavior of wind speed. However, wind speed demonstrates complicated diversity in different places and circumstances, and consequently simple models cannot well formulate this complex property. In this thesis, an adaptive and generalized probability distribution model, generalized Lambda distribution, is proposed for the approximation of stochastic wind speed. Based on studies on wind data collected from a number of measurement stations, the proposed GLD distribution demonstrates significant superiority compared with traditional probability models.

From the short-term perspective of power system operation and planning, decision makers are interested in accurately predicting the future power output of wind farm with look-ahead time from minutes to several days. Different from traditional generation, wind power is directly related to natural wind speed and indeed an intermittent energy. Therefore, wind power prediction becomes extremely critical to power system operation, planning and control, such as dispatching traditional thermal unit, determining reserve, etc. Because of the chaotic nature of weather systems and wind speed, perfect prediction of wind power is impossible in practice, even though NNs with excellent regression ability are applied. In this thesis, several novel techniques of probabilistic forecasting have been developed for wind power generation in order to quantify the uncertainty of short-term wind power forecasting.

Firstly, a bootstrap based ELM approach is newly developed to generate the probabilistic prediction intervals of wind power. With the application of ELM, the proposed BELM approach overcomes the drawbacks of traditional NNs, including local minima, over training, high computational burden, and so forth. As a parametric approach, the stochastic forecasting errors are modeled using the censored normal distribution. Though not perfect, it is able to quantify different shapes of probability distribution to some extent. The parameters of the censored normal distribution are formulated through considering both the model regression uncertainty and data noise. The effectiveness of the developed BELM approach has been verified by means of case studies of an Australian wind farm, where the seasonality of wind power series is taken into consideration. Multi-step forecasting of hourly and intra-hour wind power has been conducted, indicating wide application potential. Due to the excellent regression capability of ELM, the BELM approach can provide a generalized prediction model easily to be extended to fit forecasting wind power with different look-ahead horizons and aggregated power of different wind farms. Furthermore, the prediction model of BELM technique can be constructed with extremely fast learning speed, and then can be used for real time practical applications in power systems.

To avoid the formulation of stochastic forecasting errors needed in parametric probabilistic forecasting approaches, an advanced non-parametric direct interval forecasting approach is developed to directly generate the optimal prediction

intervals of wind power generation applying ELM based forecaster. The novel cost function combining reliability and interval score is formulated to train the ELM. Different traditional nonparametric approaches quantile regression and resampling approach, the developed DIF approach does not need the knowledge of historical forecasting errors. The proposed DIF method successfully avoids the point forecasting of wind power and overcomes the problem that different point forecasting approaches would generate different patterns of wind power prediction errors. The performance-oriented cost function can ensure the quality of generated PIs, indicating high robustness. On basis of the universal mapping ability of ELM, the DIF approach can be easily adjusted to different forecasting horizons according to the requirements of decision makers.

It is well known that the quality of prediction intervals is assessed by means of reliability and sharpness. The two evaluation indices can be incompatible to some extent. Therefore, a Pareto optimization approach is developed to derive optimal prediction intervals via Pareto optimization of the two quality indices of prediction intervals, i.e., reliability and sharpness. This nonparametric approach seems more intuitive than DIF approach. The effectiveness of this Pareto optimization approach is verified on Bornholm Island in Denmark, a realistic microgrid for EU smart grid test, implying the high potential in the future smart grid environment.

The electricity market acts a crucial character in realizing the economic prophecy of smart grids. Accurate and reliable electricity market price forecasting is essential to facilitate various decision making activities of market participants in the future smart grid environment. However, due to the nonstationarities involved in MCPs, it is rather difficult to accurately predict MCPs in advance, like wind power. The challenge is getting intensified as more and more renewable energy and other new technologies emerged in smart grids. This thesis proposes a hybrid approach to construct PIs of MCPs with a two-stage formulation. In the first stage, ELM is applied to estimate point forecasts of MCPs and model uncertainties involved. In the second stage, the maximum likelihood method is used to estimate the noise variance. Based on the mathematical background introduced in this thesis, a generalized and comprehensive framework is proposed for the evaluation of probabilistic

electricity price forecasting. The effectiveness and efficiency of the proposed hybrid method has been validated through comprehensive tests using real price data from Australian electricity market.

Wind power has introduced serious uncertainties to power systems, due to its inherent intermittency. Therefore, traditional deterministic load flow requiring specific values of power generation, loads, as well as network configuration cannot be sufficiently accurate to describe the status of electricity networks. In contrast, probabilistic load flow deals with the stochastic loads, generation and network configurations in power system can provide more general information of the power systems. In this thesis, it is implemented to estimate the mean and variance of stochastic variables of load flow parameters considering the uncertainty of wind power. PEM and FOSM approaches are applied for probabilistic load flow analysis comparing with MCS. Numerical studies demonstrate the better performance of PEM approach. In general, probabilistic load flow can be an effective technique for the power system operation and expansion planning under the integration of large wind farm.

## **10.2 Perspectives**

In this thesis, the developed probabilistic interval forecasting approaches are just examined for wind power forecasting with look-ahead time shorter than a few hours, without loss of generality. However, in practice, TSO in power systems concern multi-step wind power forecasting up to 48 hours ahead for most European countries. As generalized prediction frameworks, theoretically the developed models should be able to handle this task. The meteorological information such as wind speed and direction will be included as the input of these forecasters to implement wind power forecasting up to 48 hours ahead. Outputs of different NWP models can be combined to improve the quality of wind power forecasts. It will be verified on various realistic wind farms worldwide.

It should be emphasized that systematic probabilistic forecasting algorithms and methodologies, including parametric and nonparametric techniques, have been developed in this thesis. With the encouraging outcomes of the study, a new wind power forecasting system and platform is expected to be developed in the

future. It is essential to make a close collaboration between wind power predictors and prediction users in the future work. The demands of industrial utilities and standards of system operators will be sufficiently considered to enhance the capability of the prediction system to derive optimal prediction interval according to the needs of users. This will certainly introduce potential commercial applications.

On basis of the satisfactory performance, the developed techniques have the high potential for the application to the operational context of power system and smart grid. The value of the uncertainty estimation provided probabilistic forecasting will be comprehensively investigated from different perspectives of power systems. Research attentions will be paid to the decision-making processes. Various types of decisions in power systems can be beneficial from probabilistic wind power forecasting, such as reserve determination, unit commitment, economic dispatch, wind power trading, wind farm control, etc. Optimal benefits from the application of probabilistic interval forecasts should be specifically investigated from the perspective of the end-user or decision-maker, by means of different decision making methodologies. In addition to probabilistic forecasting of wind power, probabilistic forecasting of electricity price will also be studied with application in different operational problems in power systems. In general, the study efforts will focus on enhance the meaningfulness and motivation of developing probabilistic forecasting.

In the thesis, the probabilistic load flow is only studied for the integration of a single wind farm. It obviously can be extended to consider multiple wind farms and the correlations between different wind farms in the future work. Focus also will be given to developing more advanced methodologies for probabilistic load flow computation for power system penetrated with different renewable energies. Furthermore, probabilistic load flow will be applied for analysis of microgrid, power system operation and planning with renewable energy integration, and so on.



## **References**

- [1] D. A. Jones, "Electrical engineering: The backbone of society," *IEE Proc. A, Sci. Meas. Tech.*, vol. 138, no. 1, pp. 1-10, Jan. 1991.
- [2] P. Kundur, *Power System Stability and Control*. New York: McGraw-hill, 1994.
- [3] H. Farhangi, "The path of the smart grid," *IEEE Power Energy Mag.*, vol. 8, no. 1, pp. 18-28, 2010.
- [4] K. Moslehi and R. Kumar, "A reliability perspective of the smart grid," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 57-64, Jun. 2010.
- [5] J. Torriti, "Demand side management for the European supergrid: Occupancy variances of European single-person households," *Energy Policy*, vol. 44, no. 0, pp. 199-206, May 2012.
- [6] D. Connolly, H. Lund, B. V. Mathiesen, and M. Leahy, "The first step towards a 100% renewable energy-system for Ireland," *Appl. Energy*, vol. 88, no. 2, pp. 502-507, Feb. 2011.
- [7] W. Liu, H. Lund, B. V. Mathiesen, and X. Zhang, "Potential of renewable energy systems in China," *Appl. Energy*, vol. 88, no. 2, pp. 518-525, Feb. 2011.
- [8] E. Denny and M. O'Malley, "Wind generation, power system operation, and emissions reduction," *IEEE Trans. Power Syst.*, vol. 21, no. 1, pp. 341-347, Feb. 2006.
- [9] H. Lund and B. V. Mathiesen, "Energy system analysis of 100% renewable energy systems —The case of Denmark in years 2030 and 2050," *Energy*, vol. 34, no. 5, pp. 524-531, May 2009.
- [10] B. Elliston, M. Diesendorf, and I. MacGill, "Simulations of scenarios with 100% renewable electricity in the Australian National Electricity Market," *Energy Policy*, vol. 45, no. 0, pp. 606-613, Jun. 2012.
- [11] J. Twidell and T. Weir, *Renewable Energy Resources*, 2nd ed.: Taylor & Francis, 2012.
- [12] J. F. Manwell, J. G. McGowan, and A. L. Rogers, *Wind Energy Explained: Theory, Design and Application*, 2nd ed. Chichester, UK: John Wiley & Sons, 2010.

- [13] T. Ackermann, *Wind Power in Power Systems*. Chichester, UK: John Wiley & Sons, 2005.
- [14] C. İlkılıç, H. Aydın, and R. Behçet, "The current status of wind energy in Turkey and in the world," *Energy Policy*, vol. 39, no. 2, pp. 961-967, Feb. 2011.
- [15] F. Dincer, "The analysis on wind energy electricity generation status, potential and policies in the world," *Renew. Sust. Energ. Rev.*, vol. 15, no. 9, pp. 5135-5142, Dec. 2011.
- [16] M. Glinkowski, J. Hou, and G. Rackliffe, "Advances in wind energy technologies in the context of smart grid," *Proc. IEEE*, vol. 99, no. 6, pp. 1083-1097, Jun. 2011.
- [17] "Global Wind Statistics 2013," Global Wind Energy Council (GWEC), Feb. 2013.
- [18] H. Lund, F. Hvelplund, P. A. Østergaard, B. Möller, B. V. Mathiesen, P. Karnøe, A. N. Andersen, P. E. Morthorst, K. Karlsson, M. Münster, J. Munksgaard, and H. Wenzel, "System and market integration of wind power in Denmark," *Energy Strategy Reviews*, vol. 1, no. 3, pp. 143-156, Mar. 2013.
- [19] Z. Xu, M. Gordon, M. Lind, and J. Ostergaard, "Towards a Danish power system with 50% wind - Smart grids activities in Denmark," in *Power & Energy Society General Meeting, IEEE*, Calgary, AB 2009, pp. 1-8.
- [20] K. Sperling, F. Hvelplund, and B. V. Mathiesen, "Evaluation of wind power planning in Denmark – Towards an integrated perspective," *Energy*, vol. 35, no. 12, pp. 5443-5454, Dec. 2010.
- [21] R. Billinton, G. Yi, and R. Karki, "Application of a joint deterministic-probabilistic criterion to wind integrated bulk power system planning," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1384-1392, Aug. 2010.
- [22] P. Heejung and R. Baldick, "Transmission planning under uncertainties of wind and load: Sequential approximation approach," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2395-2402, Aug. 2013.
- [23] G. A. Orfanos, P. S. Georgilakis, and N. D. Hatziargyriou, "Transmission expansion planning of systems with increasing wind power integration," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1355-1362, May 2013.

- [24] J. A. Carta, P. Ramírez, and S. Velázquez, "A review of wind speed probability distributions used in wind energy analysis: Case studies in the Canary Islands," *Renew. Sust. Energ. Rev.*, vol. 13, no. 5, pp. 933-955, Jun. 2009.
- [25] J. Chadee and C. Sharma, "Wind speed distributions: a new catalogue of defined models," *Wind Eng.*, vol. 25, no. 6, pp. 319-337, Nov. 2001.
- [26] C. L. Vincent, P. Pinson, and G. Giebela, "Wind fluctuations over the North Sea," *Int. J. Climatol.*, vol. 31, no. 11, pp. 1584-1595, Sep. 2011.
- [27] Y. V. Makarov, C. Loutan, J. Ma, and P. de Mello, "Operational impacts of wind generation on California power systems," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 1039-1050, May 2009.
- [28] M. H. Albadi and E. F. El-Saadany, "Overview of wind power intermittency impacts on power systems," *Electr. Pow. Syst. Res.*, vol. 80, no. 6, pp. 627-632, Jun. 2010.
- [29] H. Holttinen, "Impact of hourly wind power variations on the system operation in the Nordic countries," *Wind Energy*, vol. 8, no. 2, pp. 197-218, Apr./Jun. 2005.
- [30] H. Lund, "Large-scale integration of wind power into different energy systems," *Energy*, vol. 30, no. 13, pp. 2402-2412, Oct. 2005.
- [31] P. Sorensen, N. A. Cutululis, A. Viguera-Rodriguez, L. E. Jensen, J. Hjerrild, M. H. Donovan, and H. Madsen, "Power fluctuations from large wind farms," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 958-965, Aug. 2007.
- [32] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa, "A review on the young history of the wind power short-term prediction," *Renew. Sust. Energ. Rev.*, vol. 12, no. 6, pp. 1725-1744, Aug. 2008.
- [33] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, "The state-of-the-art in short-term prediction of wind power: A literature review," Tech. Rep., ANEMOS.plus Project Deliverable Report D1.2, 2011.
- [34] A. Fabbri, T. G. S. Román, J. R. Abbad, and V. H. M. Quezada, "Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1440-1446, Aug. 2005.

- [35] P. Pierre, "Estimation of the uncertainty in wind power forecasting," Ph.D. thesis, Ecole des Mines de Paris, France, 2006.
- [36] Z. Hu, J. Wang, J. Byrne, and L. Kurdgelashvili, "Review of wind power tariff policies in China," *Energy Policy*, vol. 53, no. 0, pp. 41-50, Feb. 2013.
- [37] C. Harris, *Electricity Markets: Pricing, Structures and Economics*. Chichester, UK: John Wiley & Sons, 2011.
- [38] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, "Integrating Renewables in Electricity Markets: Operational Problems," ed. New York, US: Springer, 2014.
- [39] M. Zugno, J. M. Morales, P. Pinson, and H. Madsen, "Pool strategy of a price-maker wind power producer," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3440-3450, Aug. 2013.
- [40] E. Y. Bitar, R. Rajagopal, P. P. Khargonekar, K. Poolla, and P. Varaiya, "Bringing wind energy to market," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1225-1235, Aug. 2012.
- [41] A. J. Conejo, F. J. Nogales, and J. M. Arroyo, "Price-taker bidding strategy under price uncertainty," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 1081-1088, Nov. 2002.
- [42] W. Reinisch and T. Tezuka, "Market power and trading strategies on the electricity market: A market design view," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1180-1190, Aug. 2006.
- [43] D. W. Bunn, "Forecasting loads and prices in competitive power markets," *Proc. IEEE*, vol. 88, no. 2, pp. 163-169, Feb. 2000.
- [44] A. Motamedi, H. Zareipour, and W. D. Rosehart, "Electricity price and demand forecasting in smart grids," *IEEE Trans. Smart Grid*, vol. 3, no. 2, pp. 664-674, 2012.
- [45] D. S. Kirschen, "Demand-side view of electricity markets," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 520-527, May 2003.
- [46] H. Zareipour, C. A. Canizares, and K. Bhattacharya, "Economic impact of electricity market price forecasting errors: A demand-side analysis," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 254-262, Feb. 2010.

- [47] K. Conradsen, L. B. Nielsen, and L. P. Prahm, "Review of Weibull statistics for estimation of wind speed distributions," *J. Climate Appl. Meteor.*, vol. 23, no. 8, pp. 1173-1183, Aug. 1984.
- [48] I. Y. F. Lun and J. C. Lam, "A study of Weibull parameters using long-term wind observations," *Renew. Energy*, vol. 20, no. 2, pp. 145-153, Jun. 2000.
- [49] T. P. Chang, "Performance comparison of six numerical methods in estimating Weibull parameters for wind energy application," *Appl. Energy*, vol. 88, no. 1, pp. 272-282, Jan. 2011.
- [50] J. V. Seguro and T. W. Lambert, "Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis," *J. Wind Eng. Ind. Aerodyn.*, vol. 85, no. 1, pp. 75-84, Mar. 2000.
- [51] J. Zhou, E. Erdem, G. Li, and J. Shi, "Comprehensive evaluation of wind speed distribution models: A case study for North Dakota sites," *Energ. Convers. Manage.*, vol. 51, no. 7, pp. 1449-1458, Jul. 2010.
- [52] A. N. Celik, "A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey," *Renew. Energy*, vol. 29, no. 4, pp. 593-604, Apr. 2004.
- [53] B. Safari and J. Gasore, "A statistical investigation of wind characteristics and wind energy potential based on the Weibull and Rayleigh models in Rwanda," *Renew. Energy*, vol. 35, no. 12, pp. 2874-2880, Dec. 2010.
- [54] V. Lo Brano, A. Orioli, G. Ciulla, and S. Culotta, "Quality of wind speed fitting distributions for the urban area of Palermo, Italy," *Renew. Energy*, vol. 36, no. 3, pp. 1026-1039, Mar. 2011.
- [55] L. Van Der Auwera, F. De Meyer, and L. M. Malet, "The use of the Weibull three-parameter model for estimating mean wind power densities," *J. Appl. Meteorol.*, vol. 19, no. 7, pp. 819-825, Jul. 1980.
- [56] R. E. Luna and H. W. Church, "Estimation of long-term concentrations using a "universal" wind speed distribution," *J. Appl. Meteorol.*, vol. 13, no. 8, pp. 910-916, Dec. 1974.
- [57] A. Garcia, J. L. Torres, E. Prieto, and A. de Francisco, "Fitting wind speed distributions: A case study," *Sol. Energy*, vol. 62, no. 2, pp. 139-144, Feb. 1998.

- [58] W. E. Bardsley, "Note on the use of the inverse Gaussian distribution for wind energy applications," *J. Appl. Meteorol.*, vol. 19, no. 9, pp. 1126-1130, Sep. 1980.
- [59] E. V. A. Bauer, "Characteristic frequency distributions of remotely sensed in situ and modelled wind speeds," *Int. J. Climatol.*, vol. 16, no. 10, pp. 1087-1102, Oct. 1996.
- [60] G. Li and J. Shi, "Application of Bayesian model averaging in modeling long-term wind speed distributions," *Renew. Energy*, vol. 35, no. 6, pp. 1192-1202, Jun. 2010.
- [61] Z. Qin, W. Li, and X. Xiong, "Estimating wind speed probability distribution using kernel density method," *Electr. Pow. Syst. Res.*, vol. 81, no. 12, pp. 2139-2146, Dec. 2011.
- [62] T. P. Chang, "Estimation of wind energy potential using different probability density functions," *Appl. Energy*, vol. 88, no. 5, pp. 1848-1856, May 2011.
- [63] C. Hastings, Jr., F. Mosteller, J. W. Tukey, and C. P. Winsor, "Low moments for small samples: A comparative study of order statistics," *Ann. Math. Stat.*, vol. 18, no. 3, pp. 413-426, Sep. 1947.
- [64] J. W. Tukey, "The practical relationship between the common transformations of percentages of counts and amounts," Tech. Rep. 36, Statistical Techniques Research Group, Princeton University, 1960.
- [65] J. S. Ramberg and B. W. Schmeiser, "An approximate method for generating symmetric random variables," *Commun. ACM*, vol. 15, no. 11, pp. 987-990, Nov. 1972.
- [66] J. S. Ramberg and B. W. Schmeiser, "An approximate method for generating asymmetric random variables," *Commun. ACM*, vol. 17, no. 2, pp. 78-82, Feb. 1974.
- [67] M. Freimer, G. Kollia, G. S. Mudholkar, and C. T. Lin, "A study of the generalized Tukey lambda family," *Commun. Stat. Theor. M.*, vol. 17, no. 10, pp. 3547-3567, Jan. 1988.
- [68] Z. A. Karian and E. J. Dudewicz, *Fitting statistical distributions: The generalized lambda distribution and generalized bootstrap methods*. New York: CRC press, 2010.

- [69] D. A. Stewart and O. M. Essenwanger, "Frequency distribution of wind speed near the surface," *J. Appl. Meteorol.*, vol. 17, no. 11, pp. 1633-1642, Nov. 1978.
- [70] A. Lakhany and H. Mausser, "Estimating the parameters of the generalized lambda distribution," *Algo Res. Quart.*, vol. 3, no. 3, pp. 47-58, Dec. 2000.
- [71] Z. A. Karian, E. J. Dudewicz, and P. McDonald, "The extended generalized lambda distribution system for fitting distributions to data: History, completion of theory, tables, applications, the "final word" on moment fits," *Commun. Stat. Simul. Comput.*, vol. 25, no. 3, pp. 611-642, Jan. 1996.
- [72] R. A. R. King and H. L. MacGillivray, "Theory & Methods: A starship estimation method for the generalized  $\lambda$  distributions," *Aust. N.Z. J. Stat.*, vol. 41, no. 3, pp. 353-374, Sep. 1999.
- [73] IEC, "Power performance measurements of electricity producing wind turbines," ed. IEC 61400-12-1, 2005.
- [74] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1035-1042, May 2005.
- [75] J. D. Hamilton, *Time series analysis*. Princeton: Princeton Univ. Press, 1994.
- [76] T. Gneiting and M. Katzfuss, "Probabilistic forecasting," *Annu. Rev. Stat. Appl.*, vol. 1, no. 1, pp. 125-151, Jan. 2014.
- [77] B. Abramson and R. Clemen, "Probability forecasting," *Int. J. Forecast.*, vol. 11, no. 1, pp. 1-4, Mar. 1995.
- [78] S. Stigler, "The transition from point to distribution estimation," *Bull. Int. Stat. Inst.*, vol. 46, pp. 332-340, 1975.
- [79] S. Petterssen, *Weather Analysis and Forecasting*, 2nd ed. New York: McGraw-Hill, 1956.
- [80] T. Gneiting and A. E. Raftery, "Weather forecasting with ensemble methods," *Science*, vol. 310, no. 5746, pp. 248-249, Oct. 2005.

- [81] C.-A. S. Staß Von Holstein, "Probabilistic forecasting: An experiment related to the stock market," *Organ. Behav. Hum. Perform.*, vol. 8, no. 1, pp. 139-158, Aug. 1972.
- [82] J. F. Yates, P. C. Price, J.-W. Lee, and J. Ramirez, "Good probabilistic forecasters: The 'consumer's' perspective," *Int. J. Forecast.*, vol. 12, no. 1, pp. 41-56, Mar. 1996.
- [83] D. Önköl and G. Muradoglu, "Effects of task format on probabilistic forecasting of stock prices," *Int. J. Forecast.*, vol. 12, no. 1, pp. 9-24, Mar. 1996.
- [84] W. C. Sanderson, S. Scherbov, B. C. O'Neill, and W. Lutz, "Conditional probabilistic population forecasting," *Int. Stat. Rev.*, vol. 72, no. 2, pp. 157-166, Aug. 2004.
- [85] B. Kim and K. Reinschmidt, "Probabilistic forecasting of project duration using Bayesian inference and the Beta distribution," *J. Constr. Eng. Manag.*, vol. 135, no. 3, pp. 178-186, Mar. 2009.
- [86] B. Kim and K. Reinschmidt, "Probabilistic forecasting of project duration using Kalman filter and the earned value method," *J. Constr. Eng. Manag.*, vol. 136, no. 8, pp. 834-843, Aug. 2010.
- [87] P. C. Pendharkar, G. H. Subramanian, and J. A. Rodger, "A probabilistic model for predicting software development effort," *IEEE Trans. Softw. Eng.*, vol. 31, no. 7, pp. 615-624, Jul. 2005.
- [88] R. L. McPherron and G. Siscoe, "Probabilistic forecasting of geomagnetic indices using solar wind air mass analysis," *Space Weather*, vol. 2, no. 1, p. S01001, Jan. 2004.
- [89] R. Krzysztofowicz, "Bayesian theory of probabilistic forecasting via deterministic hydrologic model," *Water Resour. Res.*, vol. 35, no. 9, pp. 2739-2750, Sep. 1999.
- [90] R. Krzysztofowicz, "Bayesian system for probabilistic river stage forecasting," *J. Hydrol.*, vol. 268, no. 1-4, pp. 16-40, Nov. 2002.
- [91] S.-T. Chen and P.-S. Yu, "Real-time probabilistic forecasting of flood stages," *J. Hydrol.*, vol. 340, no. 1-2, pp. 63-77, Jun. 2007.
- [92] Y. Y. Kagan and D. D. Jackson, "Probabilistic forecasting of earthquakes," *Geophys. J. Int.*, vol. 143, no. 2, pp. 438-453, Nov. 2000.

- [93] Y. Y. Kagan and D. D. Jackson, "Long-term probabilistic forecasting of earthquakes," *J. Geophys. Res. Solid Earth*, vol. 99, no. B7, pp. 13685-13700, Jul. 1994.
- [94] A. J. Cannon, "Neural networks for probabilistic environmental prediction: Conditional density estimation network creation and evaluation (CaDENCE) in R," *Comput. and Geosci.*, vol. 41, no. 0, pp. 126-135, Apr. 2012.
- [95] R. Buizza, A. Hollingsworth, F. Lalaurette, and A. Ghelli, "Probabilistic predictions of precipitation using the ECMWF ensemble prediction system," *Weather Forecast.*, vol. 14, no. 2, p. 168, Apr. 1999.
- [96] J. M. Sloughter, T. Gneiting, and A. E. Raftery, "Probabilistic wind speed forecasting using ensembles and Bayesian model averaging," *J. Am. Stat. Assoc.*, vol. 105, no. 489, pp. 25-35, Mar. 2010.
- [97] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renew. Sust. Energ. Rev.*, vol. 32, no. 0, pp. 255-270, Apr. 2014.
- [98] P. E. McSharry, S. Bouwman, and G. Bloemhof, "Probabilistic forecasts of the magnitude and timing of peak electricity demand," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1166-1172, May 2005.
- [99] C. Wan, Z. Xu, Y. Wang, Z. Y. Dong, and K. P. Wong, "A hybrid approach for probabilistic forecasting of electricity price," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 463-470, Jan. 2014.
- [100] I. T. Jolliffe and D. B. Stephenson, *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 2nd ed. New York: John Wiley & Sons, 2011.
- [101] P. F. Christoffersen, "Evaluating interval forecasts," *Int. Econ. Rev.*, pp. 841-862, Nov. 1998.
- [102] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Stat.*, vol. 13, no. 3, pp. 253-263, Jul. 1995.
- [103] T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," *J. R. Stat. Soc., Ser. B Stat. Methodol.*, vol. 69, no. 2, pp. 243-268, Apr. 2007.

- [104] T. Gneiting and A. E. Raftery, "Strictly proper scoring rules, prediction, and estimation," *J. Am. Stat. Assoc.*, vol. 102, no. 477, pp. 359-378, Mar. 2007.
- [105] M. S. Roulston and L. A. Smith, "Evaluating probabilistic forecasts using information theory," *Mon. Weather Rev.*, vol. 130, no. 6, pp. 1653-1660, Jun. 2002.
- [106] J. E. Matheson and R. L. Winkler, "Scoring rules for continuous probability distributions," *Manage.Sci.*, vol. 22, no. 10, pp. 1087-1096, Jun. 1976.
- [107] R. L. Winkler, "A decision-theoretic approach to interval estimation," *J. Am. Stat. Assoc.*, vol. 67, no. 337, pp. 187-191, Mar. 1972.
- [108] J. W. Taylor, "Evaluating volatility and interval forecasts," *J. Forecasting*, vol. 18, no. 2, pp. 111-128, Mar. 1999.
- [109] P. Pinson, H. A. Nielsen, J. K. Møller, H. Madsen, and G. N. Kariniotakis, "Non-parametric probabilistic forecasts of wind power: Required properties and evaluation," *Wind Energy*, vol. 10, no. 6, pp. 497-516, Nov./Dec. 2007.
- [110] T. Heskes, "Practical confidence and prediction intervals," in *Advances in Neural Information Processing Systems*. vol. 9, M. Mozer, M. Jordan, and T. Petsche, Eds., ed Cambridge, MA: MIT Press, 1997, pp. 176–182.
- [111] R. Tibshirani, "A comparison of some error estimates for neural network models," *Neural Comput.*, vol. 8, no. 1, pp. 152-163, Jan. 1996.
- [112] J. B. Bremnes, "Probabilistic wind power forecasts using local quantile regression," *Wind Energy*, vol. 7, no. 1, pp. 47-54, Jan./Mar. 2004.
- [113] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1845-1856, Nov. 2010.
- [114] G. Papadopoulos, P. J. Edwards, and A. F. Murray, "Confidence estimation methods for neural networks: A practical comparison," *IEEE Trans. Neural Netw.*, vol. 12, no. 6, pp. 1278-1287, Nov. 2001.
- [115] G. Chryssolouris, M. Lee, and A. Ramsey, "Confidence interval prediction for neural network models," *IEEE Trans. Neural Netw.*, vol. 7, no. 1, pp. 229-232, Jan. 1996.

- [116] J. G. Hwang and A. A. Ding, "Prediction intervals for artificial neural networks," *J. Am. Stat. Assoc.*, vol. 92, no. 438, pp. 748-757, Jun. 1997.
- [117] A. A. Ding and X. He, "Backpropagation of pseudo-errors: Neural networks that are adaptive to heterogeneous noise," *IEEE Trans. Neural Netw.*, vol. 14, no. 2, pp. 253-262, Mar. 2003.
- [118] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, and H. W. Ngan, "Electricity price forecasting with extreme learning machine and bootstrapping," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2055-2062, Nov. 2012.
- [119] J. H. Zhao, Z. Y. Dong, Z. Xu, and K. P. Wong, "A statistical approach for interval forecasting of the electricity price," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 267-276, May 2008.
- [120] J. L. Cervera and J. Muñoz, "Proper scoring rules for fractiles," in *Bayesian Statistics 5*, J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Eds., ed Oxford, U.K.: Oxford University Press, 1996, pp. 513 - 519.
- [121] J. Bröcker and L. A. Smith, "Scoring probabilistic forecasts: The importance of being proper," *Weather Forecast.*, vol. 22, no. 2, pp. 382-388, Apr. 2007.
- [122] A. Betz, *Introduction to The Theory of Flow Machines*. Oxford:: Pergamon, 1966.
- [123] J. Lerner, M. Grundmeyer, and M. Garvert, "The role of wind forecasting in the successful integration and management of an intermittent energy source," *Energy Central Topics Newsletter*, vol. 3, no. 8, Jul. 2009.
- [124] R. J. Barthelmie, F. Murray, and S. C. Pryor, "The economic benefit of short-term forecasting for wind energy in the UK electricity market," *Energy Policy*, vol. 36, no. 5, pp. 1687-1696, May 2008.
- [125] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," in *North American Power Symposium (NAPS), 2010*, Arlington, TX 2010, pp. 1-8.
- [126] J. Tastu, P. Pinson, P. J. Trombe, and H. Madsen, "Probabilistic forecasts of wind power generation accounting for geographically dispersed

- information," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 480-489, Jan. 2014.
- [127] P. Pinson, "Very-short-term probabilistic forecasting of wind power with generalized logit-normal distributions," *J. R. Stat. Soc., Ser. C Appl. Stat.*, vol. 61, no. 4, pp. 555-576, Aug. 2012.
- [128] A. Botterud, Z. Zhou, J. Wang, J. Sumaili, H. Keko, J. Mendes, R. J. Bessa, and V. Miranda, "Demand dispatch and probabilistic wind power forecasting in unit commitment and economic dispatch: A case study of Illinois," *IEEE Trans. Sust. Energy*, vol. 4, no. 1, pp. 250-261, Jan. 2013.
- [129] E. M. Constantinescu, V. M. Zavala, M. Rocklin, L. Sangmin, and M. Anitescu, "A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 431-441, Feb. 2011.
- [130] R. J. Bessa, M. A. Matos, I. C. Costa, L. Bremermann, I. G. Franchin, R. Pestana, N. Machado, H. Waldl, and C. Wichmann, "Reserve setting and steady-state security assessment using wind power uncertainty forecast: A case study," *IEEE Trans. Sust. Energy*, vol. 3, no. 4, pp. 827-836, Oct. 2012.
- [131] M. A. Matos and R. J. Bessa, "Setting the operating reserve using probabilistic wind power forecasts," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 594-603, May 2011.
- [132] P. Pinson, C. Chevallier, and G. N. Kariniotakis, "Trading wind generation from short-term probabilistic forecasts of wind power," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 1148-1156, Aug. 2007.
- [133] S. D. Campbell and F. X. Diebold, "Weather forecasting for weather derivatives," *J. Am. Stat. Assoc.*, vol. 100, no. 469, pp. 6-16, Mar. 2005.
- [134] C. W. Potter and M. Negnevitsky, "Very short-term wind forecasting for Tasmanian power generation," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 965-972, May 2006.
- [135] H. Madsen, P. Pinson, G. Kariniotakis, H. A. Nielsen, and T. Nielsen, "Standardizing the performance evaluation of short-term wind power prediction models," *Wind Eng.*, vol. 29, no. 6, pp. 475-489, Dec. 2005.
- [136] U. Focken and M. Lange, *Physical Approach to Short-term Wind Power Prediction*: Springer, 2006.

- [137] L. Landberg, "A mathematical look at a physical power prediction model," *Wind Energy*, vol. 1, no. 1, pp. 23-28, Jan. 1998.
- [138] L. Kamal and Y. Z. Jafri, "Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan," *Sol. Energy*, vol. 61, no. 1, pp. 23-32, Jul. 1997.
- [139] J. L. Torres, A. Garc ía, M. De Blas, and A. De Francisco, "Forecast of hourly average wind speed with ARMA models in Navarre (Spain)," *Sol. Energy*, vol. 79, no. 1, pp. 65-77, Jul. 2005.
- [140] E. Bossanyi, "Short-term wind prediction using Kalman filters," *Wind Eng.*, vol. 9, no. 1, pp. 1-8, 1985.
- [141] P. Louka, G. Galanis, N. Siebert, G. Kariniotakis, P. Katsafados, I. Pytharoulis, and G. Kallos, "Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering," *J. Wind Eng. Ind. Aerodyn.*, vol. 96, no. 12, pp. 2348-2362, Dec. 2008.
- [142] K. Hunt and G. Nason, "Wind speed modelling and short-term prediction using wavelets," *Wind Eng.*, vol. 25, no. 1, pp. 55-61, Jan. 2001.
- [143] P. Pinson, H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Local linear regression with adaptive orthogonal fitting for the wind power application," *Stat. Comput.*, vol. 18, no. 1, pp. 59-71, 2008.
- [144] E. Cadenas and W. Rivera, "Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks," *Renew. Energy*, vol. 34, no. 1, pp. 274-278, Jan. 2009.
- [145] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, "Support vector machines for wind speed prediction," *Renew. Energy*, vol. 29, no. 6, pp. 939-947, May 2004.
- [146] T. G. Barbounis, J. B. Theocharis, M. C. Alexiadis, and P. S. Dokopoulos, "Long-term wind speed and power forecasting using local recurrent neural network models," *IEEE Trans. Energy Convers.*, vol. 21, no. 1, pp. 273-284, Mar. 2006.
- [147] A. Kusiak, H. Zheng, and Z. Song, "Short-term prediction of wind farm power: A data mining approach," *IEEE Trans. Energy Convers.*, vol. 24, no. 1, pp. 125-136, Mar. 2009.

- [148] A. Kusiak and W. Li, "Short-term prediction of wind power with a clustering approach," *Renew. Energy*, vol. 35, no. 10, pp. 2362-2369, Oct. 2010.
- [149] G. A. Dutton, G. Kariniotakis, J. A. Halliday, and E. Nogaret, "Load and wind power forecasting methods for the optimal management of isolated power systems with high wind penetration," *Wind Eng.*, vol. 23, no. 2, 1999.
- [150] A. Sfetsos, "A comparison of various forecasting techniques applied to mean hourly wind speed time series," *Renew. Energy*, vol. 21, no. 1, pp. 23-35, Sep. 2000.
- [151] B. Ernst, B. Oakleaf, M. L. Ahlstrom, M. Lange, C. Moehrlen, B. Lange, U. Focken, and K. Rohrig, "Predicting the wind," *IEEE Power Energy Mag.*, vol. 5, no. 6, pp. 78-89, Nov.-Dec. 2007.
- [152] I. Sánchez, "Adaptive combination of forecasts with application to wind energy," *Int. J. Forecast.*, vol. 24, no. 4, pp. 679-693, Oct./Dec. 2008.
- [153] J. J. Traiteur, D. J. Callicutt, M. Smith, and S. B. Roy, "A short-term ensemble wind speed forecasting system for wind power applications," *J. Appl. Meteorol.*, vol. 51, no. 10, pp. 1763-1774, Oct. 2012.
- [154] G. Giebel, L. Landberg, G. Kariniotakis, and R. Brownsword, "State-of-the-art methods and software tools for short-term prediction of wind energy production," in *Proc. 2003 European Wind Energy Conference and Exhibition (EWEC)*, Madrid, Spain, 2003.
- [155] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, G. Conzelmann, Decision, I. Sciences, and I. Porto, "Wind power forecasting : State-of-the-art 2009," Argonne National Laboratory (ANL), Tech. Rep., 2009.
- [156] M. Lei, L. Shiyang, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renew. Sust. Energ. Rev.*, vol. 13, no. 4, pp. 915-920, May 2009.
- [157] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation," *Renew. Energy*, vol. 37, no. 1, pp. 1-8, Jan. 2012.

- [158] C. Lowery and M. O'Malley, "Impact of wind forecast error statistics upon unit commitment," *IEEE Trans. Sust. Energy*, vol. 3, no. 4, pp. 760-768, Oct. 2012.
- [159] S. Tewari, C. J. Geyer, and N. Mohan, "A statistical model for wind power forecast error and its application to the estimation of penalties in liberalized markets," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2031-2039, Nov. 2011.
- [160] P. Pinson and G. Kariniotakis, "On-line assessment of prediction risk for wind power production forecasts," *Wind Energy*, vol. 7, no. 2, pp. 119-132, Apr./Jun. 2004.
- [161] M. Lange, "On the uncertainty of wind power predictions—Analysis of the forecast accuracy and statistical distribution of errors," *J. Sol. Energy Eng.*, vol. 127, no. 2, pp. 177-184, Apr. 2005.
- [162] H. Bludszweit, J. A. Dominguez-Navarro, and A. Llombart, "Statistical analysis of wind power forecast error," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 983-991, Aug. 2008.
- [163] H. Bludszweit, "Reduction of the uncertainty of wind power predictions using energy storage," Ph.D. thesis, Department of Electrical Engineering, University of Zaragoza, Zaragoza, Spain, 2009.
- [164] J. B. Bremnes, "A comparison of a few statistical models for making quantile wind power forecasts," *Wind Energy*, vol. 9, no. 1-2, pp. 3-11, Jan./Apr. 2006.
- [165] H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts," *Wind Energy*, vol. 9, no. 1-2, pp. 95-108, Jan./Apr. 2006.
- [166] J. K. Møller, H. A. Nielsen, and H. Madsen, "Time-adaptive quantile regression," *Comput. Stat. Data Anal.*, vol. 52, no. 3, pp. 1292-1303, Jan. 2008.
- [167] G. Anastasiades and P. McSharry, "Quantile forecasting of wind power using variability indices," *Energies*, vol. 6, no. 2, pp. 662-695, 2013.
- [168] J. Juban, L. Fugon, and G. Kariniotakis, "Probabilistic short-term wind power forecasting based on kernel density estimators," in *Proc. 2007 European Wind Energy Conference and Exhibition (EWEC)*, Milan, Italy, 2007.

- [169] R. J. Bessa, V. Miranda, A. Botterud, J. Wang, and E. M. Constantinescu, "Time adaptive conditional kernel density estimation for wind power forecasting," *IEEE Trans. Sust. Energy*, vol. 3, no. 4, pp. 660-669, Oct. 2012.
- [170] R. J. Bessa, V. Miranda, A. Botterud, Z. Zhou, and J. Wang, "Time-adaptive quantile-copula for wind power probabilistic forecasting," *Renew. Energy*, vol. 40, no. 1, pp. 29-39, Apr. 2012.
- [171] J. Jeon and J. W. Taylor, "Using conditional kernel density estimation for wind power density forecasting," *J. Am. Stat. Assoc.*, vol. 107, no. 497, pp. 66-79, Mar. 2012.
- [172] P. Pinson, H. Madsen, H. A. Nielsen, G. Papaefthymiou, and B. Klöckl, "From probabilistic forecasts to statistical scenarios of short-term wind power production," *Wind Energy*, vol. 12, no. 1, pp. 51-62, Jan. 2009.
- [173] G. Sideratos and N. D. Hatziargyriou, "Probabilistic wind power forecasting using radial basis function neural networks," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1788-1796, Nov. 2012.
- [174] A. Khosravi and S. Nahavandi, "Combined nonparametric prediction intervals for wind power generation," *IEEE Trans. Sust. Energy*, vol. 4, no. 4, pp. 849-856, Oct. 2013.
- [175] C. Wan, Z. Xu, and P. Pinson, "Comments on "Comprehensive review of neural network-based prediction intervals and new advances" and "Lower upper bound estimation method for construction of neural network-based prediction intervals"," *IEEE Trans. Neural Netw. Learn. Syst.*, 2014. Submitted.
- [176] A. Lau and P. McSharry, "Approaches for multi-step density forecasts with application to aggregated wind power," *Ann. Appl. Stat.*, vol. 4, no. 2010, pp. 1311-1341, 2010.
- [177] P. Pinson and H. Madsen, "Adaptive modelling and forecasting of offshore wind power fluctuations with Markov-switching autoregressive models," *J. Forecasting*, vol. 31, no. 4, pp. 281-313, Jul. 2012.
- [178] A. Khosravi, S. Nahavandi, and D. Creighton, "Prediction intervals for short-term wind farm power generation forecasts," *IEEE Trans. Sust. Energy*, vol. 4, no. 3, pp. 602-610, Jul. 2013.

- [179] R. Blonbou, "Very short-term wind power forecasting with neural networks and adaptive Bayesian learning," *Renew. Energy*, vol. 36, no. 3, pp. 1118-1124, Mar. 2011.
- [180] G. Giebel, L. Landberg, J. Badger, K. Sattler, H. Feddersen, T. S. Nielsen, H. A. Nielsen, and H. Madsen, "Using ensemble forecasting for wind power," in *Proc. 2003 European Wind Energy Conference and Exhibition (EWEC)*, Madrid, Spain, 2003.
- [181] G. Giebel (ed.), J. Badger, L. Landberg, H. A. Nielsen, T. S. Nielsen, H. Madsen, K. Sattler, H. Feddersen, H. Vedel, J. Tøfting, L. Kruse, and L. Voulund, "Wind power prediction using ensembles," *Risø-R-1527(EN)*, Sep. 2005.
- [182] P. Pinson, H. A. Nielsen, H. Madsen, and G. Kariniotakis, "Skill forecasting from ensemble predictions of wind power," *Appl. Energy*, vol. 86, no. 7–8, pp. 1326-1334, Jul. 2009.
- [183] P. Pinson and H. Madsen, "Ensemble-based probabilistic forecasting at Horns Rev," *Wind Energy*, vol. 12, no. 2, pp. 137-155, Mar. 2009.
- [184] J. W. Taylor, P. E. McSharry, and R. Buizza, "Wind power density forecasting using ensemble predictions and time series models," *IEEE Trans. Energy Convers.*, vol. 24, no. 3, pp. 775-782, Sep. 2009.
- [185] J. Munksgaard and P. E. Morthorst, "Wind power in the Danish liberalised power market—Policy measures, price impact and investor incentives," *Energy Policy*, vol. 36, no. 10, pp. 3940-3947, Oct. 2008.
- [186] G. Strbac, A. Shakoor, M. Black, D. Pudjianto, and T. Bopp, "Impact of wind generation on the operation and development of the UK electricity systems," *Electr. Pow. Syst. Res.*, vol. 77, no. 9, pp. 1214-1227, Jul. 2007.
- [187] N. J. Cutler, N. D. Boerema, I. F. MacGill, and H. R. Outhred, "High penetration wind generation impacts on spot prices in the Australian national electricity market," *Energy Policy*, vol. 39, no. 10, pp. 5939-5949, Oct. 2011.
- [188] C. K. Woo, I. Horowitz, J. Moore, and A. Pacheco, "The impact of wind generation on the electricity spot-market price level and variance: The Texas experience," *Energy Policy*, vol. 39, no. 7, pp. 3939-3944, Jul. 2011.

- [189] Z. Zhou, A. Botterud, J. Wang, R. J. Bessa, H. Keko, J. Sumaili, and V. Miranda, "Application of probabilistic wind power forecasting in electricity markets," *Wind Energy*, vol. 16, no. 3, pp. 321-338, Apr. 2013.
- [190] M. Matos and R. Bessa, "Decision support tool for power balance and reserve management," Tech. Rep., ANEMOS.plus Project Deliverable Rep. 3.3, 2008.
- [191] H. Bludszweit and J. A. Dominguez-Navarro, "A probabilistic method for energy storage sizing based on wind power forecast uncertainty," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1651-1658, Aug. 2011.
- [192] P. Pinson, G. Papaefthymiou, B. Klockl, and J. Verboomen, "Dynamic sizing of energy storage for hedging wind power forecast uncertainty," in *IEEE Power & Energy Society General Meeting, 2009*, Calgary, AB 2009, pp. 1-8.
- [193] Y. V. Makarov, P. V. Etingov, J. Ma, Z. Huang, and K. Subbarao, "Incorporating uncertainty of wind power generation forecast into power system operation, dispatch, and unit commitment procedures," *IEEE Trans. Sust. Energy*, vol. 2, no. 4, pp. 433-442, Oct. 2011.
- [194] R. Jiang, J. Wang, and Y. Guan, "Robust unit commitment with wind power and pumped storage hydro," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 800-810, May 2012.
- [195] Q. Wang, Y. Guan, and J. Wang, "A chance-constrained two-stage stochastic program for unit commitment with uncertain wind power output," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 206-215, Feb. 2012.
- [196] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359-366, Jan. 1989.
- [197] G. Zhang, B. Eddy Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," *Int. J. Forecast.*, vol. 14, no. 1, pp. 35-62, Mar. 1998.
- [198] G. N. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret, "Wind power forecasting using advanced neural networks models," *IEEE Trans. Energy Convers.*, vol. 11, no. 4, pp. 762-767, Dec. 1996.
- [199] D. J. C. MacKay, "The evidence framework applied to classification networks," *Neural Comput.*, vol. 4, no. 5, pp. 720-736, Sep. 1992.

- [200] D. A. Nix and A. S. Weigend, "Estimating the mean and variance of the target probability distribution," in *Proc. IEEE Int. Conf. Neural Netw.*, Orlando, FL, 1994, pp. 55-60.
- [201] R. Dybowski and S. Roberts, "Confidence intervals and prediction intervals for feed-forward neural networks," in *Clinical Applications of Artificial Neural Networks*, R. Dybowski and V. Grant, Eds., ed Cambridge, U.K.: Cambridge University Press, 2001, pp. 298-326.
- [202] F. Giordano, M. La Rocca, and C. Perna, "Forecasting nonlinear time series with neural network sieve bootstrap," *Comput. Stat. Data Anal.*, vol. 51, no. 8, pp. 3871-3884, May 2007.
- [203] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501, Dec. 2006.
- [204] A. H. Nizar, Z. Y. Dong, and Y. Wang, "Power utility nontechnical loss analysis with extreme learning machine method," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 946-955, Aug. 2008.
- [205] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst. Man Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513-529, Apr. 2012.
- [206] G.-B. Huang, X. Ding, and H. Zhou, "Optimization method based extreme learning machine for classification," *Neurocomputing*, vol. 74, no. 1-3, pp. 155-163, Dec. 2010.
- [207] B. Muthn, "Moments of the censored and truncated bivariate normal distribution," *Br. J. Math. Stat. Psychol.*, vol. 43, no. 1, pp. 131-143, May 1990.
- [208] C. M. Bishop, *Neural Networks for Pattern Recognition*. London, U.K.: Oxford university press, 1995.
- [209] B. Efron, "Bootstrap methods: Another look at the Jackknife," *The Annals of Statistics*, vol. 7, no. 1, pp. 1-26, Jan. 1979.
- [210] B. Efron and R. Tibshirani, *An Introduction to The Bootstrap*. New York: CRC Press, 1993.
- [211] H. Cremc, *Mathematical Methods of Statistics (PMS-9)*. Princeton: Princeton Univ. Press, 1999.

- [212] G. A. F. Seber and A. J. Lee, *Linear Regression Analysis*. New York: Wiley, 2003.
- [213] V. Akhmatov, "Influence of wind direction on intense power fluctuations in large offshore windfarms in the North Sea," *Wind Eng.*, vol. 31, no. 1, pp. 59-64, Jan. 2007.
- [214] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, Perth, WA 1995, pp. 1942-1948.
- [215] J. F. Kennedy, J. Kennedy, and R. C. Eberhart, *Swarm intelligence*. San Francisco, USA: Morgan Kaufmann, 2001.
- [216] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Lower upper bound estimation method for construction of neural network-based prediction intervals," *IEEE Trans. Neural Netw.*, vol. 22, no. 3, pp. 337-346, Mar. 2011.
- [217] A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of optimal prediction intervals for load forecasting problems," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1496-1503, Aug. 2010.
- [218] C. Wan, Z. Xu, J. Østergaard, Z. Y. Dong, and K. P. Wong, "Discussion of "Combined nonparametric prediction intervals for wind power generation", " *IEEE Trans. Sust. Energy*, vol. 5, no. 3, p. 1021, Jul. 2014.
- [219] P. Sørensen, A. D. Hansen, K. Thomsen, T. Buhl, P. E. Morthorst, L. H. Nielsen, F. Iov, F. Blaabjerg, H. A. Nielsen, H. Madsen, and M. H. Donovan, "Operation and control of large wind turbines and wind farms - Final report," Tech. Rep., Risø National Laboratories for Sustainable Energy, Technical University of Denmark, Risø-R-1532(EN), 2005.
- [220] J. Kristoffersen and P. Christiansen, "Horns Rev offshore windfarm: Its main controller and remote control system," *Wind Eng.*, vol. 27, no. 5, pp. 351-359, Sep. 2003.
- [221] K. E. Parsopoulos and M. N. Vrahatis, "Particle swarm optimization method in multiobjective problems," in *Proceedings of 2002 ACM Symposium on Applied Computing*, Madrid, Spain, 2002, pp. 603-607.
- [222] J. Juban, L. Fugon, and G. Kariniotakis, "Uncertainty estimation of wind power forecasts: Comparison of probabilistic modelling approaches," in *Proceedings of European Wind Energy Conference & Exhibition (EWEC) 2008*, , Brussels, Belgium 2008.

- [223] A. K. Al-Othman and M. R. Irving, "Analysis of confidence bounds in power system state estimation with uncertainty in both measurements and parameters," *Electr. Pow. Syst. Res.*, vol. 76, no. 12, pp. 1011-1018, Aug. 2006.
- [224] H. Quan, D. Srinivasan, and A. Khosravi, "Short-term load and wind power forecasting using neural network-based prediction intervals," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 2, pp. 303 - 315 Feb. 2014.
- [225] Ronay Aka, Yanfu Li, V. Vitellia, and E. Zioa, "Multi-objective genetic algorithm optimization of a neural network for estimating wind speed prediction intervals," 2013.
- [226] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182-197, Apr. 2002.
- [227] M. Niu and Z. Xu, "Efficiency ranking-based evolutionary algorithm for power system planning and operation," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1437 - 1438 May 2014.
- [228] R. L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge, U.K.: Cambridge University Press, 1993.
- [229] C. M. Fonseca and P. J. Fleming, "An overview of evolutionary algorithms in multiobjective optimization," *Evol. Comput.*, vol. 3, no. 1, pp. 1-16, 1995.
- [230] J. Bröcker and L. A. Smith, "Increasing the Reliability of Reliability Diagrams," *Weather Forecast.*, vol. 22, no. 3, pp. 651-661, Jun. 2007.
- [231] P. Pinson, P. McSharry, and H. Madsen, "Reliability diagrams for non-parametric density forecasts of continuous variables: Accounting for serial correlation," *Quart. J. Roy. Meteor. Soc.*, vol. 136, no. 646, pp. 77-90, Jan. 2010.
- [232] J. Østergaard, A. Foosnæs, Z. Xu, T. A. Mondorf, C. A. Andersen, S. Holthusen, T. Holm, M. F. Bendtsen, and K. Behnke, "Electric vehicles in power systems with 50% wind power penetration: the Danish case and the EDISON programme," in *European Conference Electricity & Mobility*, 2009.

- [233] J. P. Barton and D. G. Infield, "Energy storage and its use with intermittent renewable energy," *IEEE Trans. Energy Convers.*, vol. 19, no. 2, pp. 441-448, Jun. 2004.
- [234] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Optimal prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1166 - 1174, May 2014.
- [235] "Strategic research agenda for Europe's electricity networks of the future," EUR 22580, European Technology Platform SmartGrids, 2007.
- [236] J. Contreras, R. Espinola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014-1020, Aug. 2003.
- [237] R. C. Garcia, J. Contreras, M. van Akkeren, and J. B. C. Garcia, "A GARCH forecasting model to predict day-ahead electricity prices," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 867-874, May 2005.
- [238] B. R. Szkuta, L. A. Sanabria, and T. S. Dillon, "Electricity price short-term forecasting using artificial neural networks," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 851-857, Aug. 1999.
- [239] Z. Xu, Z. Y. Dong, and W. Q. Liu, "Neural network models for electricity market forecasting," in *Neural Networks Applications in Information Technology and Web Engineering*, D. H. Wang and N. K. Lee, Eds., ed Sarawak: Borneo Publishing, 2005, pp. 233–245.
- [240] N. Amjady, "Day-ahead price forecasting of electricity markets by a new fuzzy neural network," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 887-896, May 2006.
- [241] M. Zhou, Z. Yan, Y. X. Ni, G. Li, and Y. Nie, "Electricity price forecasting with confidence-interval estimation through an extended ARIMA approach," *IEE Proc. Gener. Transm. Distrib.*, vol. 153, no. 2, pp. 187-195, Mar. 2006.
- [242] Z. Li and P. B. Luh, "Neural network-based market clearing price prediction and confidence interval estimation with an improved extended Kalman filter method," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 59-66, Feb. 2005.

- [243] R. D. De VIEaux, J. Schumi, J. Schweinsberg, and L. H. Ungar, "Prediction intervals for neural networks via nonlinear regression," *Technometrics*, vol. 40, no. 4, pp. 273-282, Nov. 1998.
- [244] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proceedings of 2004 IEEE International Joint Conference on Neural Networks*, Budapest, Hungary, 2004, pp. 985-990.
- [245] S. de la Torre, A. J. Conejo, and J. Contreras, "Transmission expansion planning in electricity markets," *IEEE Trans. Power Syst.*, vol. 23, no. 1, pp. 238-248, Feb. 2008.
- [246] B. Borkowska, "Probabilistic load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 752-759, May 1974.
- [247] M. Ni, J. D. McCalley, V. Vittal, S. Greene, T. Chee-Wooi, V. S. Ganugula, and T. Tayyib, "Software implementation of online risk-based security assessment," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1165-1172, Aug. 2003.
- [248] M. Ni, J. D. McCalley, V. Vittal, and T. Tayyib, "Online risk-based security assessment," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 258-265, Feb. 2003.
- [249] T. De la Torre, J. W. Feltes, T. G. S. Román, and H. M. Merrill, "Deregulation, privatization, and competition: transmission planning under uncertainty," *IEEE Trans. Power Syst.*, vol. 14, no. 2, pp. 460-465, May 1999.
- [250] A. M. Leite da Silva, S. M. P. Ribeiro, V. L. Arienti, R. N. Allan, and M. B. Do Coutto Filho, "Probabilistic load flow techniques applied to power system expansion planning," *IEEE Trans. Power Syst.*, vol. 5, no. 4, pp. 1047-1053, Nov. 1990.
- [251] T. A. M. Sharaf and G. J. Berg, "Stochastic and probabilistic load flow analysis in system planning," *Can. Electr. Eng. J.*, vol. 8, no. 1, pp. 9-17, Jan. 1983.
- [252] A. P. Meliopoulos, X. Y. Chao, G. Cokkinides, and R. Monsalvatge, "Transmission loss evaluation based on probabilistic power flow," *IEEE Trans. Power Syst.*, vol. 6, no. 1, pp. 364-371, Feb. 1991.

- [253] J. M. Sexauer and S. Mohagheghi, "Voltage quality assessment in a distribution system with distributed generation — A probabilistic load flow approach," *IEEE Trans. Power Del.*, vol. 28, no. 3, pp. 1652-1662, Jul. 2013.
- [254] N. D. Hatziargyriou and T. S. Karakatsanis, "Probabilistic load flow for assessment of voltage instability," *IEE Proc. Gener. Transm. Distrib.*, vol. 145, no. 2, pp. 196-202, Mar. 1998.
- [255] N. D. Hatziargyriou and T. S. Karakatsanis, "Distribution system voltage and reactive power control based on probabilistic load flow analysis," *IEE Proc. Gener. Transm. Distrib.*, vol. 144, no. 4, pp. 363-369, Jul. 1997.
- [256] Y.-Y. Hong and Y.-F. Luo, "Optimal VAR control considering wind farms using probabilistic load-flow and gray-based genetic algorithms," *IEEE Trans. Power Del.*, vol. 24, no. 3, pp. 1441-1449, Jul. 2009.
- [257] P. Jorgensen, J. S. Christensen, and J. O. Tande, "Probabilistic load flow calculation using Monte Carlo techniques for distribution network with wind turbines," in *Proceedings of 8th International Conference On Harmonics and Quality of Power Athens 1998*, pp. 1146-1151.
- [258] A. Sankarakrishnan and R. Billinton, "Effective techniques for reliability worth assessment in composite power system networks using Monte Carlo simulation," *IEEE Trans. Power Syst.*, vol. 11, no. 3, pp. 1255-1261, Aug. 1996.
- [259] Y. Chen, J. Wen, and S. Cheng, "Probabilistic load flow method based on Nataf transformation and Latin Hypercube sampling," *IEEE Trans. Sust. Energy*, vol. 4, no. 2, pp. 294-301, Apr. 2013.
- [260] C. Defu, S. Dongyuan, and C. Jinfu, "Probabilistic load flow computation with polynomial normal transformation and Latin hypercube sampling," *IET Gener. Transm. Distrib.*, vol. 7, no. 5, pp. 474-482, May 2013.
- [261] H. Yu, C. Y. Chung, K. P. Wong, H. W. Lee, and J. H. Zhang, "Probabilistic load flow evaluation with hybrid Latin hypercube sampling and Cholesky decomposition," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 661-667, May 2009.
- [262] R. N. Allan and M. R. G. Al-Shakarchi, "Probabilistic a.c. load flow," *Proc. IEE.*, vol. 123, no. 6, pp. 531-536, Jun. 1976.

- [263] R. N. Allan and M. R. G. Al-Shakarchi, "Probabilistic techniques in a.c. load-flow analysis," *Proc. IEE.*, vol. 124, no. 2, pp. 154-160, Feb. 1977.
- [264] R. N. Allan, A. M. Leite da Silva, and R. C. Burchett, "Evaluation methods and accuracy in probabilistic load flow solutions," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 5, pp. 2539-2546, May 1981.
- [265] P. Zhang and S. T. Lee, "Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion," *IEEE Trans. Power Syst.*, vol. 19, no. 1, pp. 676-682, Feb. 2004.
- [266] T. Williams and C. Crawford, "Probabilistic load flow modeling comparing maximum entropy and Gram-Charlier probability density function reconstructions," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 272-280, Feb. 2013.
- [267] A. M. Leite da Silva, V. L. Arienti, and R. N. Allan, "Probabilistic load flow considering dependence between input nodal powers," *IEEE Power Eng. Rev.*, vol. PER-4, no. 6, pp. 67-68, Jun. 1984.
- [268] A. M. Leite da Silva and V. L. Arienti, "Probabilistic load flow by a multilinear simulation algorithm," *IEE Proc. C, Gener. Transm. Distrib.*, vol. 137, no. 4, pp. 276-282, Jul. 1990.
- [269] R. N. Allan and A. M. Leite da Silva, "Probabilistic load flow using multilinearisations," *IEE Proc. Gener. Transm. Distrib.*, vol. 128, no. 5, pp. 280-287, Sep. 1981.
- [270] J. M. Morales and J. Perez-Ruiz, "Point estimate schemes to solve the probabilistic power flow," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1594-1601, Nov. 2007.
- [271] C.-L. Su, "Probabilistic load-flow computation using point estimate method," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1843-1851, Nov. 2005.
- [272] E. Rosenblueth, "Point estimates for probability moments," *Proc. Natl. Acad. Sci.*, vol. 72, no. 10, pp. 3812-3814, Oct. 1975.
- [273] E. Rosenblueth, "Two-point estimates in probabilities," *Appl. Math. Model.*, vol. 5, no. 5, pp. 329-335, Oct. 1981.
- [274] H. P. Hong, "An efficient point estimate method for probabilistic analysis," *Reliab. Eng. Syst. Safe.*, vol. 59, no. 3, pp. 261-267, Mar. 1998.

- [275] Z. Hu and X. Wang, "A probabilistic load flow method considering branch outages," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 507-514, May 2006.
- [276] M. H. Ahmed, K. Bhattacharya, and M. M. A. Salama, "Probabilistic distribution load flow with different wind turbine models," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1540-1549, May 2013.
- [277] X. Ai, J. Wen, T. Wu, and W.-J. Lee, "A discrete point estimate method for probabilistic load flow based on the measured data of wind power," *IEEE Trans. Ind. Appl.*, vol. 49, no. 5, pp. 2244-2252, Sep./Oct. 2013.
- [278] N. D. Hatziargyriou, T. S. Karakatsanis, and M. Papadopoulos, "Probabilistic load flow in distribution systems containing dispersed wind power generation," *IEEE Trans. Power Syst.*, vol. 8, no. 1, pp. 159-165, Feb. 1993.
- [279] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Probabilistic power flow analysis with generation dispatch including photovoltaic resources," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1797-1805, May 2013.
- [280] M. Fan, V. Vittal, G. T. Heydt, and R. Ayyanar, "Probabilistic power flow studies for transmission systems with photovoltaic generation using cumulants," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2251-2261, Nov. 2012.
- [281] F. J. Ruiz-Rodriguez, J. C. Hernández, and F. Jurado, "Probabilistic load flow for radial distribution networks with photovoltaic generators," *IET Renew. Pow. Gen.*, vol. 6, no. 2, pp. 110-121, Mar. 2012.
- [282] G. Li and X.-P. Zhang, "Modeling of plug-in hybrid electric vehicle charging demand in probabilistic power flow calculations," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 492-499, Mar. 2012.
- [283] J. G. Vlachogiannis, "Probabilistic constrained load flow considering integration of wind power generation and electric vehicles," *IEEE Trans. Power Syst.*, vol. 24, no. 4, pp. 1808-1817, Nov. 2009.
- [284] R. N. Allan, C. H. Grigg, and M. R. G. Al-Shakarchi, "Numerical techniques in probabilistic load flow problems," *Int. J. Numer. Meth. Eng.*, vol. 10, no. 4, pp. 853-860, 1976.
- [285] C. Wan, Z. Xu, Z. Y. Dong, and K. P. Wong, "Probabilistic load flow computation using first-order second-moment method," in *Proceeding of*

2012 *IEEE Power and Energy Society General Meeting*, San Diego, CA,, 2012.

- [286] D. Villanueva, J. L. Pazos, and A. Feijoo, "Probabilistic load flow including wind power generation," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1659-1667, Aug. 2011.
- [287] S. H. Karaki, R. B. Chedid, and R. Ramadan, "Probabilistic performance assessment of wind energy conversion systems," *IEEE Trans. Energy Convers.*, vol. 14, no. 2, pp. 217-224, Jun. 1999.
- [288] F. S. Wong, "First-order, second-moment methods," *Comput. Struct.*, vol. 20, no. 4, pp. 779-791, 1985.
- [289] S.-J. Wang and K.-C. Hsu, "The application of the first-order second-moment method to analyze poroelastic problems in heterogeneous porous media," *J. Hydrol.*, vol. 369, no. 1-2, pp. 209-221, May 2009.
- [290] F. Yuan, "Analysis of stochastic behaviour of linear circuits using first-order second-moment and adjoint network techniques," *Electron. Lett.*, vol. 33, no. 9, pp. 766-768, Apr. 1997.
- [291] S. Sayed, G. Dodagoudar, and K. Rajagopal, "Reliability analysis of reinforced soil walls under static and seismic forces," *Geosynth. Int.*, vol. 15, no. 4, pp. 246-257, Aug. 2008.
- [292] P. Caramia, G. Carpinelli, and P. Varilone, "Point estimate schemes for probabilistic three-phase load flow," *Electr. Pow. Syst. Res.*, vol. 80, no. 2, pp. 168-175, Feb. 2010.