UNCERTAINTY MODELING IN DISTRIBUTED POWER SYSTEMS

HAO QUAN

NATIONAL UNIVERSITY OF SINGAPORE

2014

UNCERTAINTY MODELING IN DISTRIBUTED POWER SYSTEMS

HAO QUAN

(B.Eng, Huazhong University of Science & Technology, China)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

August, 2014

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Hao Quan

August, 2014

Acknowledgments

First and foremost, I would like to thank my Ph.D. supervisor, Associate Professor Dipti Srinivasan for her patient guidance and support. It is my greatest honor to have had an opportunity to do my research with her. She is a very kind and experienced person. During the study and research, I had total freedom to discuss the research and manage my own time. She always respects students and their ideas, which is the greatest style I admire. Whenever I encountered difficulties in the research, her guidance and suggestions inspired me a lot. It is her patience and professional guidance that improved me step by step in my whole Ph.D. study. Not only did Prof. Dipti teach me the professional knowledge, but also she taught me how to plan the future research development and how to be a useful person to the whole society.

My great thanks also go to Dr. Abbas Khosravi in Deakin University, Australia, who is also the third author in my research papers. He is a nice person and always with a lot of ideas and patience. He provided me with the datasets for case studies, helped to correct my draft papers and taught me how to address the comments from the reviewers. It was my great pleasure to cooperate with him.

In addition, I would like to acknowledge the financial and academic support of the Department of Electrical and Computer Engineering and the National University of Singapore. This work was supported by National Research Foundation programme grant, NRF-2007EWT-CERP01-0954 (R-263-000-522-272) and R-263-000-A66-279. I would like to thank the thesis examiners for their comments, which help me to further improve the thesis. I sincerely appreciate the assistance of all faculty and staff.

I feel fortunate to have the opportunity to study in the Energy Management and Microgrid Laboratory. I wish to express my great thanks to the lab officers Mr. Seow and Mr. Looi, and my seniors Balaji Gokulan Parasumanna, Thillainathan Logenthiran, Vishal Sharma, Xiong Peng, David Zhao and Anupam Trivedi. They have shared their experience and helped me from time

to time. Thank the buddies and friends in this lab very much. They are Saranga Abeygunawardane, Shu Zhen, Lv Wenjing, Sumith Sreechithra, Bordin Bordeerath, Yang Dazhi, Shang Ce, Gu Chaojun, Zhang Yifan, Anurag Sharma, Dhivya Sampath Kumar, Bal Satarupa, Bharat Menon Radhakrishnan, Rahul Mehta, Rajasekharareddy Chilipi and Devendra Ramesh Patil. In addition, thank my friends from the nearby Electrical Machines and Drives Laboratory, Tan Sicong, Wang Xue, Guo Zhaoqin and Sangit Sasidharetc. I have spent most of the time with them, and although they come from diverse backgrounds and behaviors, they have made my university life memorable and enjoyable.

Moreover, I want to thank Chen Zongrui, who helped me a lot when I first came to Singapore. Thank Anurag Sharma. It was happy to play basketball and discuss questions with him. I also want to mention the happy time with two exchange students Bi Yunrui and Sun Zhe. We played tennis together and had a pleasure trip.

Last but not least, I would like to express my deep appreciation to my family: my father, mother, my sisters, my uncles, aunts and my cousins etc. When I had just came to Singapore, my mother was ill in hospital, and it was a really difficult period for me. However, everything went well soon. I want to thank all the financial and emotional support from my family. It is their love and encouragement that has always been a source of courage and strength and pushed me to move forward. It is my greatest honor to dedicate this thesis to my warm family.

Contents

Abstract

Penetration of intermittent renewable energy resources (IRESs), such as wind and solar power, into power systems significantly increases the uncertainties on system operation, economics, stability and reliability. Adequate handling and modeling of these uncertainties become urgent for smart grid applications. This thesis aims to develop uncertainty modeling methodologies in distributed power systems (DPSs) for decision making and risk assessment.

Firstly, a novel neural network (NN)-based prediction interval (PI) method was developed to handle uncertainties for forecasting. A new PI construction method named lower upper bound estimation (LUBE) method was applied and extended. The LUBE method adopts a NN with two outputs to directly generate the upper and lower bounds of PIs without any assumption about data distributions. A new PI evaluation index, which is suitable for NN training, was proposed. Further a new cost function was developed for the comprehensive evaluation of PIs based on their width and coverage probability. PSO with the mutation operator was used as a training algorithm and to minimize the cost function. Results from six case studies show that the proposed PSO-based LUBE method is very efficient in constructing higher quality PIs in a short time.

The PSO-based LUBE method was further improved and applied to uncertainty handling for electrical load and wind power forecasting in DPSs. The primary problem for construction of intervals was firstly formulated as a constrained single-objective problem. The width of PIs was treated as the key objective and their calibration was considered as the constraint. Compared to cost function method, advantages of the new formulation are its closeness to the primary problem and require fewer parameters. PSO enhanced by the mutation operator was then used to optimally tune NN parameters subject to constraints set on the quality of PIs. Results of both load and wind case studies clearly show that the proposed probabilistic forecasting method generates well-calibrated and informative PIs. Furthermore, comparative results demonstrate that the proposed PI construction method greatly outperforms three widely used benchmark methods.

Up to now, the forecast uncertainty has been well addressed. The following investigation is to incorporate wind power interval forecast uncertainties into stochastic unit commitment (UC) for decision making. Instead of a single level PI, wind power uncertainties were captured by a list of PIs. A new scenario generation method was proposed. For each hour, an empirical cumulative distribution function (ECDF) was fitted to these interval points. The Monte Carlo simulation method was used to generate scenarios from the ECDF. Then wind power scenarios were incorporated into a stochastic security-constrained UC (SCUC) problem. Compared to PSO which is powerful for parameter (real numbers) optimization, genetic algorithm (GA) is more suitable for the binary representation UC problem. Therefore, GA was proposed to solve the stochastic SCUC. Five deterministic and four stochastic case studies were implemented. Generation costs and available reserves from different UC strategies were discussed. Comparative results show the differences between the planned and real time economic dispatch reserves. The stochastic models are more robust than deterministic ones.

Previous studies mainly focused on one or two aspects (load and/or wind) of the uncertainties. A comprehensive computational framework was proposed in this thesis for quantification and integration of uncertainties in DPSs with IRESs. Different sources of uncertainties in DPSs such as electrical load, wind and solar power forecasts and generator outages are covered by the proposed framework. Both the deterministic and stochastic SCUC were conducted. Generation costs as well as different reserve strategies were investigated from the decision making and risk assessment perspectives. Simulation results show power systems run higher level of risk during peak load hours. Compared with the only wind case, although the overall costs are lower due to the penetration of solar power, the involved risks are also higher in this integration framework. The stochastic models indicate better robustness than deterministic ones.

Overall, this study mainly investigates uncertainty modeling methodologies for forecasting and renewable energy integration. The proposed methodologies and framework could provide potential solution for uncertainty quantification and integration in DPSs with IRESs.

List of Tables

List of Figures

List of Symbols and Abbreviations

Mathematical Symbols in Unit Commitment Formulation:

i	Index of generators, i = 1, ..., N.
t	Index of scheduled hours, t = 1, ..., H.
s	Index of scenarios, s = 1, ..., S.
p_s	The probability of scenario s.
$(.)^s$	Variable related to scenario s.
$X_{i,t}$	The scheduled state (on/off) of unit i at time t.
$P_{i,t}$	The output power of unit i at time t.
$E(X,P)$	The objective function, the expected product costs.
$F_i(P_{i,t})$	The fuel cost of unit i when its output power is $P_{i,t}$.
$SU_{i,t}$	Startup cost of unit i at time t.
$CSU_{i,t}$	Cold startup cost of unit i at time t.
$HSU_{i,t}$	Hot startup cost of unit i at time t.
R_t	The spinning reserve at time t.
D_t	The system demand at time t.
$IRES_t$	The output power of intermittent renewable energy source at time t.
ENS_t^s	The energy not served at time t in scenario s.
RNS_t^s	The reserve not served at time t in scenario s.
C_{ens}	The cost of energy not served.
C_{rns}	The cost of reserve not served.
UR_i	Ramp up rate limit of unit i.
DR_i	Ramp down rate limit of unit i.
W_t^s	Wind generation at time t in scenario s.
PV_t^s	Solar generation at time t in scenario s.
FO_t^s	Generator forced outage at time t in scenario s.
$P_{i,max}$	Maximum real power generation of unit i.
$P_{i,min}$	Minimum real power generation of unit i.

L_{km}^{max}	Transmission flow limits from bus k to bus m.
$T_{i,t}^{off}$	The continuously off time of unit i at time t.
$T_{i,t}^{on}$	The continuously on time of unit i at time t.
T_i^{Up}	The minimum up time of unit i.
T_i^{Down}	The minimum down time of unit i.
T_i^{cold}	The cold start hours of unit i.
a_i, b_i, c_i	Coefficients for the quadratic cost curve of unit i;
$FLAC(i)$	The full load average cost of unit i.
A	The system-dependent constant in the fitness function.
q	The parameter in the q-tournament selection.

Abbreviations:

ACF	Autocorrelation Function.
ARIMA	Autoregressive Integrated Moving Average.
CDF	Cumulative Distribution Function.
CI	Confidence Interval.
COPT	Capacity Outage Probability Table.
CWC	Coverage Width-based Criterion.
DE	Differential Evolution.
DP	Dynamic Programming.
DPS	Distributed Power System.
ECDF	Empirical Cumulative Distribution Function.
ED	Economic Dispatch.
EENS	Expected Energy Not Supplied/Served.
EMS	Energy Management System.
ENS	Energy Not Served.
ES	Exponential Smoothing.
FLS	Fuzzy Logic System.
GA	Genetic Algorithm.
IRES	Intermittent Renewable Energy Source.

LMP Locational Marginal Price.

LOLD Loss of Load Duration.

LOLF Loss of Load Frequency.

LOLP Loss of Load Probability.

LR Lagrangian Relaxation.

LUBE Lower Upper Bound Estimation.

MAPE Mean Absolute Percentage Error.

MSE Mean Square Error.

NN Neural Network.

NSW New South Wales.

NW Nguyen-Widrow.

PACF Partial Autocorrelation Function.

PDF Probability Density Function.

PI Prediction Interval.

PICP Prediction Interval Coverage Probability.

PINAW Prediction Interval Normalized Average Width.

PINRW Prediction Interval Normalized Root-mean-square Width.

PL Priority Listing.

PSO Particle Swarm Optimization.

PV Photovoltaic.

RES Renewable Energy Source.

RNS Reserve Not Served.

SCADA Supervisory Control and Data Acquisition.

SCUC Security-Constrained Unit Commitment.

SD Standard Deviation.

SG Singapore.

SR Spinning Reserve.

SSE Sum of Squared Error.

STLF Short-Term Load Forecasting.

SUC Stochastic Unit Commitment.

UC Unit Commitment.

WDCF Weight Decay Cost Function.

Chapter 1

Introduction

Electrical power systems are evolving from today's centralized bulk systems, with generation plants connected to the transmission network, to more decentralized distributed power systems (DPSs), with smaller generating units connected directly to distribution networks close to demand consumption [1]. Due to the random nature of weather, intermittent power generation sources such as wind and solar systems involve high uncertainties. With sudden weather changes, the output power of a wind farm can fluctuate largely between zero and its capacity. Penetration of renewable energy sources (RESs) into power systems significantly increases the uncertainties on both generation and demand sides. These uncertainties bring challenges to system operation economics, stability, security and reliability of traditional power systems. Forecasting and unit commitment (UC) scheduling are important and challenging tasks in power systems. However, traditional methods of forecasting and UC scheduling in existing energy management systems (EMSs) cannot adequately address these uncertainties arising from RESs. Advanced methods for forecast uncertainty modeling and new algorithms to incorporate forecast uncertainty into UC are therefore needed in DPSs [2, 3].

The organization of this chapter is as follows. Section 1.1 and Section 1.2 provide the background information and general issues of uncertainty modeling separately. Section 1.3 briefly introduces the DPSs, such as distributed generation and microgrid. Then uncertainty represen-

tations in DPSs with intermittent RESs (IRESs) are introduced in Section 1.4. The motivations, objectives and main contributions are presented in Section 1.5 and Section 1.6 respectively.

1.1 Background

To protect the environment and reduce consumption of conventional energy resources, the installed capacity of renewable energy, wind power in particular, is growing at a significant rate in many countries around the world. For example, the share of wind power generation in the United States has been increasing with an annual rate of 25% since 1990. It is estimated that by 2020, about 12% of the world's electricity will be supplied by wind generation. Wind generation installation has reached considerable percentages (in the range of 5% to 20%) of the whole installed capacity in some European countries in recent years, such as Germany, Spain, and Denmark. On the other hand, the capital costs ($\$/KW$) of wind energy decrease gradually with the increment of the overall installed capacities. Thus, wind energy is becoming an important component in the supply mix to meet the growing demand for electric energy [4].

The uncertainties involved in DPSs are much more than traditional power systems. Most of the electrical load and wind power forecasting are point forecasts. However, point forecasts cannot properly handle the uncertainties associated with data sets. From the uncertainty handling point of view, probabilistic forecasting of load and wind power needs to be explored, such as quantile forecast or prediction intervals (PIs). PIs are excellent tools for the quantification of uncertainties associated with point forecasts and predictions [5].

UC and economic dispatches (ED) are essential and basic tasks in operation and economics of power systems. UC problem is usually formulated to minimize the total costs of generations under some operating constraints. Generally speaking, the task is to determine the on/off status and output power of each generator. Penetration of IRESs increases the uncertainties in generation, and this brings a big challenge to traditional UC scheduling. It is crucial to build uncertainty handling models to incorporate the IRES forecast uncertainties into the UC problem.

Among these uncertainty models, the stochastic model which represents the wind power forecast uncertainty through scenarios, is one of the most popular ones. Further these uncertainty models can be applied to decision making and risk assessment in DPSs.

1.2 Overview of Uncertainty Modeling

The definition, classification and sources of uncertainty are firstly introduced. Challenges, issues and general approaches for uncertainty modeling are then presented.

1.2.1 Definition and Classification of Uncertainty

Uncertainty in engineering analysis and design is commonly defined as knowledge incompleteness due to inherent deficiencies in acquired knowledge [6]. Uncertainty can be classified based on its sources into three types: ambiguity, approximations, and likelihood. The ambiguity comes from the possibility of having multioutcomes for processes or systems. The process of approximation can involve the use of vague semantics in language, approximate reasoning, and dealing with complexity by emphasizing relevance. Approximations can be viewed to include vagueness, coarseness and simplification. The likelihood can be defined in the context of chance, odds and gambling. Likelihood has primary components of randomness and sampling [7].

Uncertainty can also be used to characterize the state of a system as being unsettled or in doubt, such as the uncertainty of the outcome. Uncertainty is an important dimension in the analysis of risks. Traditionally, uncertainty in risk analysis processes is classified as follows [6]:

1) Inherent randomness (i.e., aleatory uncertainty): Some events and modeling variables are perceived to be inherently random and are treated to be nondeterministic in nature. The uncertainty in this case is attributed to the physical world because it cannot be reduced or eliminated by enhancing the underlying knowledge base. This type of uncertainty is sometimes referred to as aleatory uncertainty. It is representative of unknowns that differ each time we run the same experiment. An example of this uncertainty type is strength properties of materials such

as steel and concrete, and structural load characteristics such as wave loads on an offshore platform.

2) Subjective (or epistemic) uncertainty: In many situations, uncertainty is also present as a result of a lack of complete knowledge. In this case, the uncertainty magnitude could be reduced as a result of enhancing the state of knowledge by expending resources. Sometimes, this uncertainty cannot be reduced due to resource limitations, technological infeasibility, or sociopolitical constraints. This type of uncertainty, sometimes referred to as epistemic uncertainty, is the most dominant type in risk analysis. A subjective estimate of this probability can be used in risk analysis; however, the uncertainty in this value should be recognized. With some additional modeling effort, this value can be treated as a random variable bounded using probability intervals or percentile ranges. By enhancing our knowledge base about this potential event, these ranges can be updated.

In real-world applications, these two kinds of uncertainties may present separately or together. Uncertainty quantification intends to work towards reducing epistemic uncertainties to aleatory uncertainties.

1.2.2 Sources of Uncertainty

The sources of uncertainty in modeling and computation of engineering aspects of power systems include, but not limited to the following [8, 9]:

1) Parameter uncertainty, which comes from the model parameters that are inputs to the computer model (mathematical model) but their exact values are unknown to experimentalists, cannot be controlled in physical experiments or subject to error.

2) Model inadequacy, also called as structural uncertainty, model bias or model discrepancy, which comes from the lack of knowledge of the underlying true physics. It depends on how accurately a mathematical model describes the true system for a real-life situation, consider-

ing the fact that models are almost always only approximations to reality.

3) Parametric variability, which comes from the variability of input variables of the model. Inputs into the model may be subject to noise. For example, the value of a given system load may only be representable as an uncertain variable.

4) Observation error, which comes from the variability of experimental measurements.

5) Numerical uncertainty, numerical modeling using finite arithmetic introduces errors in the modeling process that can, under some conditions, overwhelm the accuracy of a model.

6) Interpolation uncertainty, which comes with a lack of available data collected from computer model simulations and/or experimental measurements. For other input settings that don't have simulation data or experimental measurements, one must interpolate or extrapolate in order to predict the corresponding responses.

1.2.3 Challenges and Issues of Uncertainty Modeling

Uncertainty, which was considered synonymous with random, stochastic, and probabilistic processes, has grown to incorporate many more uncertain tools and methodologies. Today the questions with which many practitioners are struggling are [7]:

1) What is uncertainty? Is it just a lack of knowledge and limited information?

2) What are the correct approaches for addressing, analyzing, and modeling uncertainty?

3) How does the quality and quantity of information affect uncertainty analysis and modeling?

4) How robust are answers obtained from uncertainty analysis and modeling?

5) How to incorporate the uncertainty modeling methods into practices for decision making and risk assessment?

Since the late 1990's, computational intelligence, or soft computing is attracting increasing attention to both the theory and practical applications. These computational intelligence methods,

such as neural networks, fuzzy logic theory and evolutionary algorithms, have been successfully applied as a tool in uncertainty analysis and modeling.

1.2.4 General Approaches for Uncertainty Modeling

There are two major types of problems in uncertainty quantification: one is the forward propagation of uncertainty and the other is the inverse assessment of model uncertainty and parameter uncertainty. The uncertainty propagation of the two problems is different. Forward uncertainty propagation quantifies the uncertainties of system output(s) propagated from uncertain inputs. The focus is on the response of system outputs to the uncertain inputs. Inverse uncertainty quantification attempts to estimate the discrepancy (bias correction) and the unknown parameters (parameter calibration), given some experimental measurements of a system and some computer simulation results from its mathematical model.

Existing forward uncertainty propagation approaches include probabilistic approaches and non-probabilistic approaches. There are basically five categories of probabilistic approaches for uncertainty propagation [10]:

- Simulation/sampling-based methods: Monte Carlo simulations [11], importance sampling [12], adaptive sampling [13], etc.
- Local expansion-based methods: Taylor series, perturbation method, etc. Methods in this category are weak against the large variability of inputs and nonlinearity of performance functions (outputs) [10].
- Functional expansion-based methods: Neumann expansion [14], polynomial chaos expansion (PCE) [15]. The PCE approach has been gaining more attention in uncertainty representations, stochastic mechanics, solution of stochastic differential equations etc.
- Most probable point (MPP)-based methods : first-order reliability method (FORM) and second-order reliability method (SORM) [16].

- Numerical integration-based methods: Full factorial numerical integration (FFNI) [17] and dimension reduction (DR) [18].

For non-probabilistic approaches, interval analysis [19], fuzzy theory [20], possibility theory and evidence theory are most widely used. Probabilistic approach is widely applied to uncertainty analysis in engineering design due to its consistency with the theory of decision analysis. Taking this thesis as an example, probabilistic forecasting results have been incorporated into a stochastic UC modeling for uncertainty quantification and decision making.

Existing methodologies for inverse uncertainty quantification are mostly under Bayesian framework [9]. The complicated direction is to solve problems with both bias correction and parameter calibration. The challenges of such problems include not only the influences from model inadequacy but also parameter uncertainty.

1.3 Distributed Power Systems (DPSs)

The distributed generation and microgrid are main characteristics of DPSs. They are introduced separately in this section.

1.3.1 Distributed Generation

Traditional power systems are centralized power systems with centralized control and management. They usually have large power stations with large generators, which are far away from the end-users. Thus the transmission lines from the generation side to the distribution side are very long. Supervisory Control and Data Acquisition (SCADA) serves as a data gathering and device control center. Traditional power systems suffer from reliability, environment, flexibility and efficiency problems.

Renewable energy sources, with its advantages of inexhaustibility and nonpollution, become the key to a sustainable energy supply infrastructure. A number of renewable energy technologies are now commercially available, the most notable being wind power, solar power,

biomass, geothermal systems, fuel cells and various forms of hydraulic power [1]. Most of these technologies offer clean, efficient and cost-effective electric energy. Their sizes are usually small, thus they are flexible enough to connect to distributed grids.

Distributed generation is an electric power source connected directly to the distribution network or on the customer site of the meter [21]. The size of distributed generation can vary from a few kilowatts to a few megawatts. Nowadays, distributed generation is gaining growing interest among smart grids, particularly as on-site generation for business and homeowners, which provides better power quality, higher reliability and fewer environment problems. Distributed generation technology is often lumped with distributed storage, and the combination is referred to as distributed energy resource that represents a modular electrical generation or storage installed at customer site. Distributed generation is operated in parallel with the utility system or islanded from the utility system [22]. Advantages of distributed generation over the traditional large-scale power generation are their characteristics of being less expensive, more reliable, flexible and environmentally friendly etc.

Shiguo, Luo in [23] introduced eight basic characteristics of distributed power systems (DPSs): thermal management and packaging, module size reduction, reduced electromagnetic interference and harmonics, modularity and standardized designs, redundancy and reliability, availability and maintainability, point-of-load regulation, flexible system structure and layout.

1.3.2 Microgrid

One important component of DPSs is microgrid. Microgrid [24] is a low voltage electrical network that interconnects small, modular generation sources such as photovoltaic, wind turbine, fuel cell and micro-turbines together with storage elements such as flywheel, super capacitors and batteries, and controllable loads. Microgrids can be operated in interconnected mode or islanded mode if it is disconnected from the main power grid.

From the grid's point of view, microgrids can be considered as controllable entities within

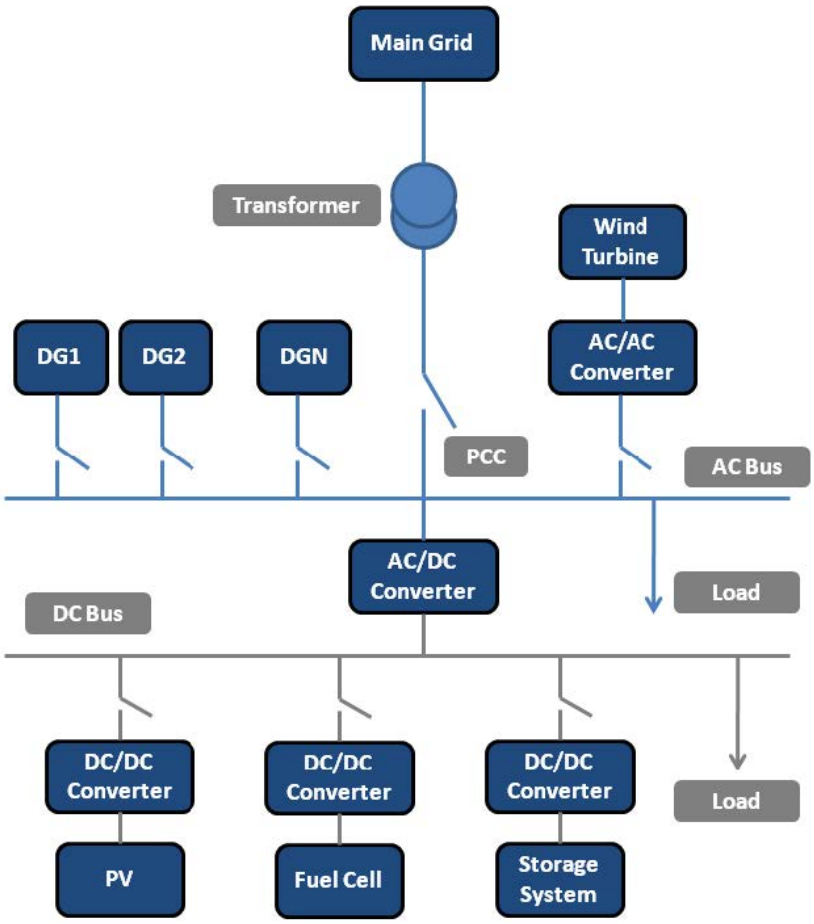

Figure 1.1: Schematic diagram of a microgrid

the electrical power system that behave as aggregated loads or sources of power and also provide ancillary services to the supporting networks which depend on the status of both the microgrid and the main distribution grid. From the customers' point of view, microgrids are similar to traditional distribution networks that provide electricity to the customers. Microgrids enhance local reliability, reduce emissions, improve power quality and potentially lower the cost of energy supply. This denotes the capability of a microgrid in the smart grid development at distribution level. Figure. 1.1 shows the schematic diagram of a microgrid [22].

Microgrid extends and adapts the concepts of distributed generation and micro-generation. It also increases the presence of renewable power sources. The key characteristics of microgrids are intelligent, efficient, resilient, dynamic, load integrated and flexible [25].

1.4 Uncertainty Representation in DPSs with Intermittent Renewable Energy Sources

Distributed power systems operations are mainly subject to four types of uncertainties: generation uncertainties, transmission uncertainties, distribution uncertainties and load uncertainties. These types of uncertainties have been summarized in Fig. 1.2. Generation uncertainties consist of generation hardware uncertainties and IRESs. The solar irradiation and wind speed fluctuate significantly with weather changes. These result in high uncertainties of the output power from IRESs. Regarding transmission uncertainties, the unreliable equipment leads to discrete events, such as the transmission line outages [26]. Distribution uncertainties include the microgrid uncertainties and power flow direction uncertainty etc. Departures from load forecasts are often continuous in nature, and include uncertain load, demand response and area interchange. In the reminder of this section, several main uncertainties are introduced one by one.

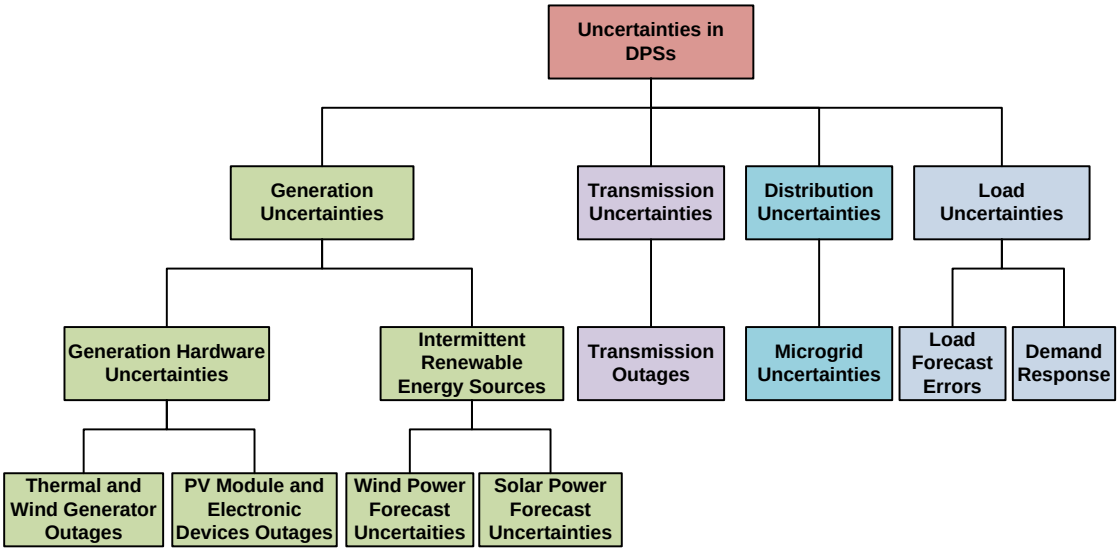

Figure 1.2: Uncertainties in DPSs.

1.4.1 Load Forecast Uncertainty

Load forecast is of great importance in power system planning and operating. Although various forecasting methods are being developed to improve the forecasting accuracy, such as time series forecasting, artificial NN models etc, some degree of load forecast uncertainty still exists in practice [27]. The level of uncertainty becomes higher if the forecasting period is longer, from several weeks to decades.

Valenzuela et al. [28] estimated power generation production costs considering the influence of temperature and load forecast uncertainty. Monte Carlo simulation was used to estimate the contributions of uncertainties resulting from load and generator availabilities to the variance of production costs. R. Billinton et al. in [27] investigated the effects of load forecast uncertainty on bulk electric system reliabilities. Load forecast uncertainty could be described by a probability distribution whose parameters can be estimated from past experience and future considerations [29]. In [27] the tabulating technique of sampling was applied to determine the uncertainty by a normal distribution with a given standard deviation. System and load point reliability indices and reliability index probability distributions were affected by load forecast uncertainty and the distribution ranges increased with increase in the uncertainty. In [2] distribution fitting method and the empirical distribution method were used to evaluate forecast uncertainties. Load and wind forecast errors were assumed to follow truncated normal distribution. In [30] different levels of standard deviation of load forecasting were examined on locational marginal price (LMP) forecasts. It has shown that more accurate load forecast should lead to less deviation in forecast price.

1.4.2 Uncertainty of Generation and Transmission Line Outages

A generator unit can be either available or not available in its role of delivering power on demand. When a generator is available for delivering power, it can be operated at its maximum continuous rating or in a derated state due to operational constraints (i.e., operating at less than its rated

capacity) or remain idle due to insufficient demand (i.e., reserve shut down state). There are many causes that make a generating unit unavailable to meet the demands imposed on it (e.g., forced outage, planned maintenance, scheduled maintenance, failure to start, etc.) [31]. Eleven distinctive states and eleven reliability indices of generators are also introduced in [31].

Frequent forced outages of transmission equipment have significant influences on the reliability and control of industrial and commercial power systems. Combinations of transmission line outages can easily lead to massive blackouts. According to the Canadian Electricity Association outage data in [31], the top three causes of transmission line "line-related" sustained forced outages are adverse weather, defective equipment and foreign interference.

In [32], load forecast errors, random outages of generators and transmission lines, and fuel price fluctuations are simulated by the Monte Carlo method. These uncertainties are accounted together for the optimal maintenance outage scheduling. Jae Hyung Roh et al. [33] also applied the Monte Carlo method to represent the random outages of generators and transmission lines as well as the inaccuracies in the long-term load forecasting. But their objective was to present a stochastic coordination of generation and transmission expansion planning model in a competitive electricity market. Manuel A. Matos et al. in [34] proposed a probabilistic model to present various uncertainties for setting the operating reserve. These uncertainties include load forecast and wind power generation forecast uncertainty, unplanned outages of conventional generators and wind turbines etc. Capacity outage probability table (COPT) was applied to represent the discrete probability distribution of conventional generation. COPT gives the probability of occurrence for each possible outage capacity level.

In power systems, operating reserve margin, percentage reserve, expected energy not supplied (EENS), loss of load probability (LOLP), loss of load frequency (LOLF), loss of load duration (LOLD) are the reliability measures. Uncertainties of generation and transmission line outages are also related to the power system reliability, especially for the EENS, LOLP and LOLF [34].

1.4.3 Uncertainty Modeling of Solar Generating Sources

Statistical solar irradiance distribution is described. The theoretical relationship between the output power of PV modules and solar irradiance are also provided.

1.4.3.1 Statistical Solar Irradiance Distribution

The basic building block of photovoltaic (PV) technology is the solar "cell", a solar cell (also called a PV cell) converts solar energy into electrical energy. Multiple PV cells are connected to form a PV "module", the smallest PV component sold commercially. These PV modules can be connected in series to increase the voltage, in parallel to increase the current or both series and parallel connection to increase power. Multiple PV modules connected together to form PV arrays. The model of the ith solar generator cell consists of two parts: the solar irradiation function and the power generation function which links the solar irradiation to the power output of the PV solar generator. To describe the random phenomenon of the irradiance data, a Beta probability density function (PDF) is utilized for statistical description as shown in Fig. 1.3 [35]:

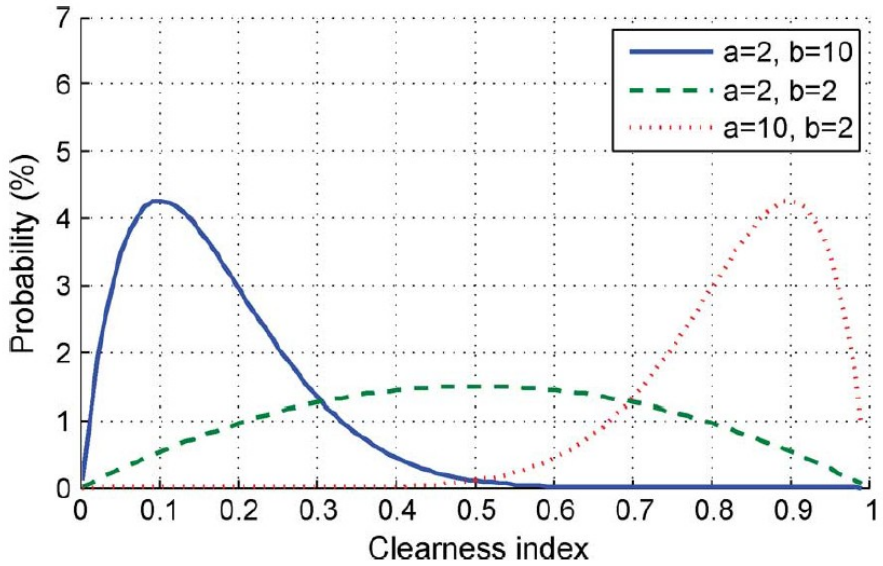

Figure 1.3: Beta probability distributions

$$f(s_i) = \begin{cases} \dfrac{\Gamma(a_i + b_i)}{\Gamma(a_i)\Gamma(b_i)} * s_i^{(a_i-1)} * (1 - s_i)^{(b_i-1)}, & 0 \leq s_i \leq 1, a_i \geq 0, b_i \geq 0 \\ 0. & otherwise \end{cases} \tag{1.1}$$

where $s_i \in [0, 1]$ is the solar irradiance (measured in KW/m^2) received by the ith solar generator, $f(s_i)$ is the Beta PDF of s_i, a_i and b_i are the parameters of the Beta PDF which can be inferred from estimates of the mean μ_{ci} and standard deviation σ_{ci} of clearness index in a predefined time period.

$$a_i = \mu_{ci} \left(\frac{\mu_{ci}(1 - \mu_{ci})}{\sigma_{ci}^2} - 1 \right) \tag{1.2}$$

$$b_i = (1 - \mu_{ci}) \left(\frac{\mu_{ci}(1 - \mu_{ci})}{\sigma_{ci}^2} - 1 \right) \tag{1.3}$$

1.4.3.2 Output Power of PV Modules

The output power of the PV module is dependent on the solar irradiance and ambient temperature of the site as well as the characteristics of the module itself. Therefore, once the Beta PDF is generated for a specific time segment, the output power during the different states is calculated for this segment using the following [1, 35]:

$$P_{Sy}(s_{ay}) = N * FF * V_y * I_y \tag{1.4}$$

$$T_{Cy} = T_A + s_{ay} \left(\frac{N_{OT} - 20}{0.8} \right) \tag{1.5}$$

$$I_y = s_{ay} \left[I_{sc} + K_i(T_C - 25) \right] \tag{1.6}$$

$$V_y = V_{OC} - K_v * T_{Cy} \tag{1.7}$$

$$FF = \frac{V_{MPP} * I_{MMP}}{V_{OC} * I_{SC}} \tag{1.8}$$

where T_{Cy} is the cell temperature $^\circ C$ during state y, T_A is the ambient temperature $^\circ C$, K_v is voltage temperature coefficient $V/^\circ C$, K_i is current temperature coefficient $A/^\circ C$, N_{OT} is nominal operating temperature of cell in $^\circ C$, FF is the fill factor, I_{sc} is the short circuit current in A, V_{oc} is the open-circuit voltage in V, I_{MPP} is current at maximum power point in A, V_{MPP} is voltage at maximum power point in V, P_{Sy} is the output power of the PV module during state y, and s_{ay} is the average solar irradiance of state y [1].

1.4.4 Uncertainty Modeling of Wind Generating Sources

The uncertainties of wind generating sources are mainly from two aspects: the intermittent and volatile wind speed and the uncertain power curve [36].

1.4.4.1 Statistical Wind Speed Distribution

Wind speed, as an essential measurement for wind power generation, is influenced by many factors such as the weather conditions, the land terrain, and the height above the ground surface. A large number of statistics show that wind speed in most regions approximately follows the Weibull distribution [35, 37]. The PDF of Weibull distribution is defined as:

$$f(v) = \frac{k}{c}(\frac{v}{c})^{k-1} \, exp\left[-(\frac{v}{c})^{k}\right] \tag{1.9}$$

where k is the shape parameter and c is the scale parameter. When $k = 2$, the PDF is called a Rayleigh PDF. Fig. 1.4 shows the PDF of Weibull distribution for different sets of parameters.

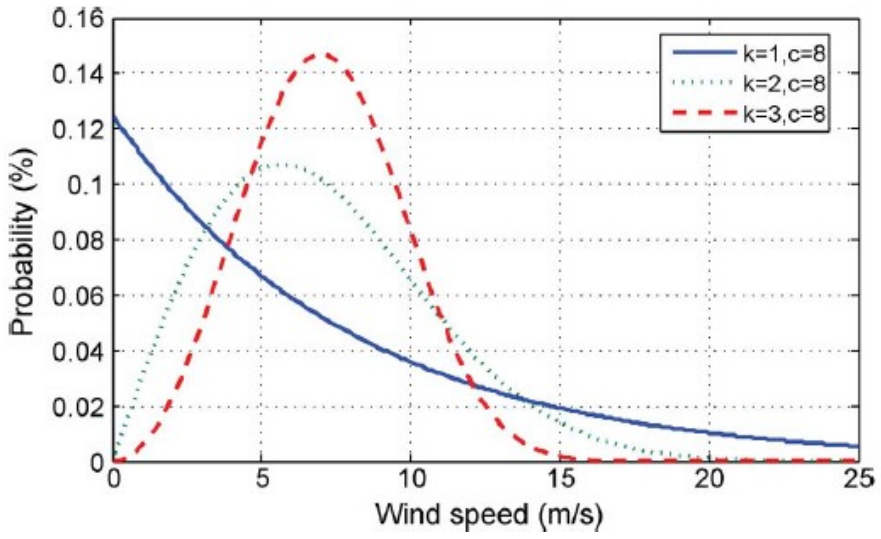

Figure 1.4: Weibull probability distributions

$$f(v) = \frac{2v}{c^2} \, exp\left[-(\frac{v}{c})^{2}\right] \tag{1.10}$$

When the distribution is Rayleigh distribution, the scale parameter c can be estimated as:

$$c = 1.128\,\overline{v} \tag{1.11}$$

where $\overline{v}$ is the average wind speed.

1.4.4.2 Output Power of Wind Turbines

The power curve of a wind turbine is a graph that indicates the mathematical mapping from different wind speeds to the electrical power output of wind turbines. After the wind speed distribution is determined, the relationship between the output power of a wind generating unit and the wind speed can be formulated as [37]:

$$P_w = \begin{cases} 0, & 0 \le v \le v_{ci} \text{ or } v_{co} \le v \\ P_{rated}\,\dfrac{(v - v_{ci})}{(v_r - v_{ci})}, & v_{ci} \le v \le v_r \\ P_{rated}. & v_r \le v \le v_{co} \end{cases} \tag{1.12}$$

where v is the wind speed at the hub height of the wind unit; v_{ci}, v_{co} and v_r are the cut-in wind speed, cut-out wind speed, and the rated wind speed respectively; P_{rated} is the rated output power of the wind unit.

The above Formula (1.12) is an ideal deterministic power curve. However, the deterministic power curve is different from the empirical power curve. In fact, the obtained empirical power curve in the real operation involves uncertainties and this curve is never unique. For comparison purpose, both the deterministic and empirical power curves of a real wind farm are shown in Fig. 1.5 and Fig. 1.6 [38] respectively. The possible deviation may be because of well-known factors like wake effects, environmental effects, hysteresis, and curtailments in the wind farms [38].

1.4.5 System Other Uncertainties

Different from the distribution network in traditional power systems, microgrid and power flow direction uncertainties are involved in the distribution systems in DPSs. A microgrid can exist in

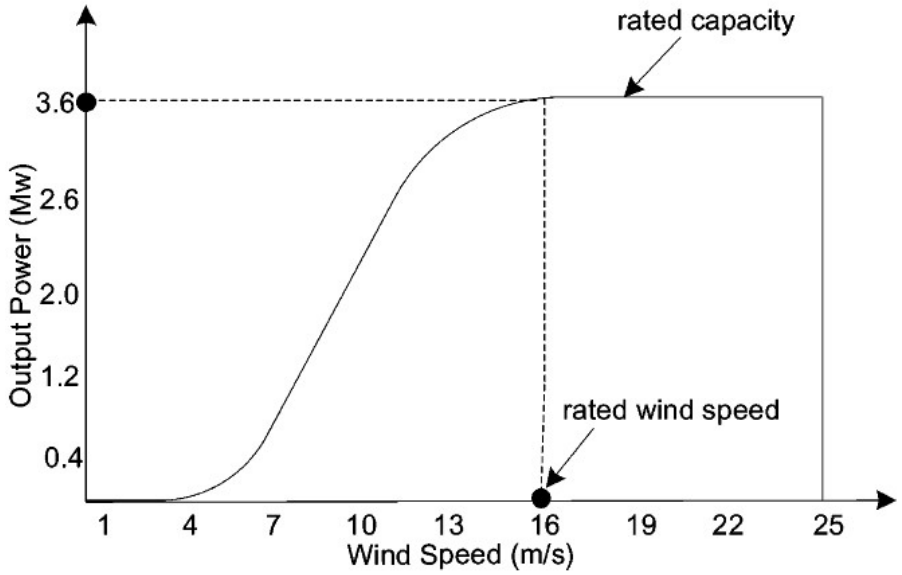

Figure 1.5: Typical machine power curve

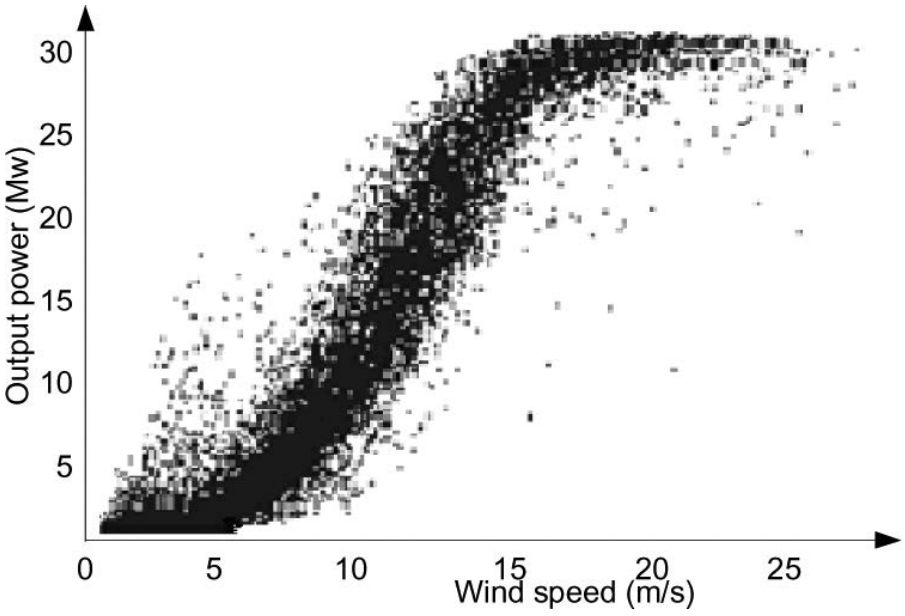

Figure 1.6: Empirical relation between speed and output power at an actual wind farm

the distribution systems in DPSs. As mentioned before, microgrids can be operated in interconnected mode or islanded mode. When the microgrids are connected into the main power grid, the direction of the power flow is determined by the power balance within the microgrids. If the wind speed is low or the solar irradiation decreases, the electricity shortage may happen in the microgrids. Then the main grid can compensate for this shortage and the power flows from the

main grid into the microgrids. If excess power is generated by the microgrids, these power will flow into the main grid. These power flow uncertainties never exist before, because the end users only consume the electricity in traditional distribution networks.

Other uncertainties such as the PV module and electronic devices outages, demand response and uncertain electricity markets will be investigated in future studies.

1.5 Motivations and Objectives

The complexity and level of uncertainty present in operation of power systems have significantly grown due to penetration of IRESs into the grid. High penetration of IRESs such as wind and solar power, will have significant impacts on power system stability, security, and reliability due to fast fluctuation and unpredictable characteristics of IRESs. The integration of a large number of wind and solar power generations can have either positive or negative impacts on DPSs. A detailed literature review can be found in Chapter 2. From the literature on incorporation of IRES uncertainties into DPSs, the main limitations and difficulties can be summarized as follows:

- The uncertainties of IRESs are not well represented in the existing forecasting methods. The forecasting results are directly from other agencies rather than being proposed and implemented by the authors themselves.

- The forecasting methods usually have special assumptions on data sets and are parametric methods. Nonparametric probabilistic methods are seldomly used for electrical load and wind power forecasting in this area of applications.

- It is difficult to apply the nonparametric forecasting results (eg. quantiles or PIs) to DPSs for decision making and risk assessment. The scenario generation approaches for stochastic models and the solution methods are very complicated.

- Previous studies mainly focus on one or two aspects of the uncertainties, such as the load and wind power forecast uncertainties. It is very important to address the uncertainty

problem comprehensively by including different sources of uncertainty for consideration.

The main objective of this study was to explore uncertainty modeling methods applied in DPSs. This main objective consisted of two parts:

- To develop advanced uncertainty modeling methods for forecasting;
- To incorporate IRESs forecast uncertainties into stochastic unit commitment (UC) for decision making and risk assessment.

In real practice of power system, uncertainties involved in DPSs are huge. The uncertainty modeling methods proposed in this research are mainly concerning forecasting and incorporating IRES forecast uncertainties into stochastic UC problems. In addition, the real power systems are very complicated and the test systems here only model the main issues. Thus some constraints in UC are omitted, such as the transmission line constraint and the ramping rate constraint of generators. The most important consideration is to validate the proposed algorithms for uncertainty quantification rather than build a very complex test system. Modeling all the constraints in real power systems for UC is beyond the scope of this research.

1.6 Main Research Contributions

The main contributions of this study are listed below:

1) Investigated the definition, classifications and sources of uncertainty. Reviewed challenges and general approaches for uncertainty modeling. General issues of uncertainty representation in DPSs with IRESs are studied.

2) Reviewed the uncertainty modeling methods for forecasting and unit commitment (UC) problems in DPSs, especially methods about the artificial neural network (NN) and the stochastic models. Reviewed formulation and solution methods of the UC problems.

3) Proposed a new and efficient uncertainty modeling method for forecasting using the NN-based PIs. This method is called the particle swarm optimization (PSO)-based lower upper bound estimation (LUBE) method.

4) With a new problem formulation, the improved PSO-based LUBE method was applied to PI construction for short-term load, wind and solar power forecasting to model the uncertainty.

5) Proposed a new scenario generation method which builds an important bridge between PIs and scenarios used in the stochastic model. The IRES forecast uncertainties were further incorporated into the stochastic UC and economic dispatch (ED) for decision making and risk assessment.

6) Proposed a computational framework for uncertainty integration in DPSs with IRESs. It was conducted to integrate load, wind and solar forecast uncertainties and generator outages uncertainty together.

The main limitations/difficulties raised in literatures (listed in Section 1.5) have been well addressed. The uncertainties of IRESs are well captured and represented by PIs using our proposed forecasting method. The PSO-based LUBE forecasting model is a nonparametric probabilistic method and the assumptions about data distributions have been avoided. The computation time and implementation difficulty are also decreased significantly. The new scenario generation method builds an important linkage between PIs and decision making. Compared to other scenario generation method, it relaxes assumptions for specific data distributions and can be easily implemented. In this way, nonparametric forecasting results (eg. quantiles or PIs) can be mathematically applied to DPSs for computation and decision making. Moreover, a computational framework has been proposed to consider different sources of uncertainty together rather than only one or two aspects of uncertainty.

The proposed uncertainty modeling methodologies would improve the existing EMS for forecasting and renewable integrations involving uncertainty. The methodologies and integration

framework can be shifted to other areas for smart grid applications. The software used for simulation and algorithm development is MATLAB on Windows XP platform. The major functions used are 'rand' (random number generator), 'plot' (figures) and 'mapminmax' (data normalization). However, the main algorithms such as the NN modeling, the PSO and genetic algorithm (GA)-based algorithms and the UC problem formulations are all programmed manually.

1.7 Outline of the Thesis

The rest of the thesis is organized as follows:

- In Chapter 2, a literature review is provided to investigate the uncertainty modeling methods in DPSs. The review is mainly on forecasting, UC with IRESs, as well as incorporation of forecast uncertainties into UC in DPSs.

- In Chapter 3, the PSO-based LUBE method is proposed for construction of PIs to handle uncertainty from forecast. Demonstrated results from six case studies indicate that the proposed method is very effective in constructing high quality PIs in a short time.

- Chapter 4 proposes a new problem formulation for construction of NN-based PIs. Using this new formulation, an improved PSO-based LUBE method is further applied to short-term load and wind power forecasting in DPSs.

- Chapter 5 proposes a new scenario generation method and incorporates the wind power interval forecast uncertainties into stochastic UC for decision making and risk assessment.

- Chapter 6 proposes a computational framework for uncertainty integration in DPSs with IRESs. In this framework, load, wind and solar power forecast uncertainties and generator outage uncertainty are integrated together.

- Chapter 7 presents the conclusion of this thesis and provides recommendations for future work in this area.

Chapter 2

Literature Review

A comprehensive literature review is provided to investigate the uncertainty modeling methods in DPSs. The review is mainly on forecasting, UC problem with IRESs and incorporating forecast uncertainties into UC scheduling for decision making. Four popularly used methods on uncertainty modeling for forecasting and UC problems are firstly reviewed. These methods are probabilistic method, PIs, fuzzy logic and stochastic models. Then a detailed investigation on construction of NN-based PIs is implemented. The traditional methods and the newly proposed LUBE method are introduced. Further, issues of UC with IRESs such as the deterministic UC formulation and solution methods are reviewed. At the end of this chapter, different studies on incorporating wind power generation forecast uncertainties into UC and other DPSs applications are investigated.

2.1 Uncertainty Modeling Methods for Forecasting and Unit Commitment Problems in DPSs

Forecasting and UC with IRESs are two challenging and important tasks in daily operations of DPSs. The popular methods used to quantify uncertainties from forecasting and UC are probabilistic method, PIs, fuzzy logic and stochastic models.

2.1.1 Probabilistic Method

Probabilistic method is the old and classical method for uncertainty modeling [39]. For the non-deterministic or uncertain events, the probabilistic theory represents them by the statistical probability.

The fundamental characterization of probability is the probability density function (PDF). In probability theory, a PDF, or density of a continuous random variable, is a function that describes the relative likelihood for this random variable to take on a given value. The probability for the random variable to fall within a particular region is given by the integral of this variables density over the region. The probability density function is nonnegative everywhere, and its integral over the entire space is equal to one.

A PDF is most commonly associated with absolutely continuous univariate distributions. A random variable X has density f, where f is a non-negative Lebesgue-integrable function, if [40]:

$$P[A \leq X \leq B] = \int_A^B f(x)dx. \tag{2.1}$$

Hence, if F is the cumulative distribution function (CDF) of X, then:

$$F(x) = \int_{-\infty}^{x} f(u)du, \tag{2.2}$$

and (if f is continuous at x),

$$f(x) = \frac{d}{dx}F(x). \tag{2.3}$$

Intuitively, one can think of $f(x)dx$ as being the probability of X falling within the infinitesimal interval $[x, x + dx]$. The uniform, Guassian and binomial PDF and CDF are shown in Fig. 2.1 [8].

If a random variable X is given and its distribution admits a probability density function f, then the expected value of X (if it exists) can be calculated as,

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx. \tag{2.4}$$

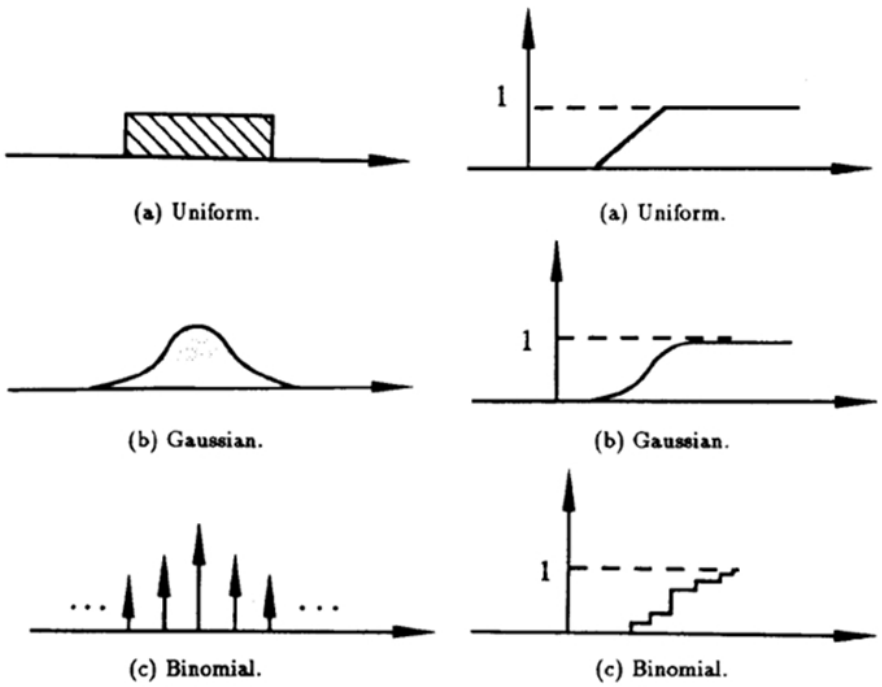

Figure 2.1: The uniform, Guassian and binomial PDF (left) and CDF (right).

There are many contents about probability theory, here only the basic PDF, CDF are listed, the detailed information are omitted in this thesis, one can easily find more from a probability book [40].

2.1.2 Prediction Intervals

In this subsection, the advantages of PIs compared to point forecasts are summarized from the uncertainty representation perspective. The difference between PIs and confidence intervals is reviewed.

2.1.2.1 Prediction Intervals VS. Point Forecasts

In regression application of NNs, they have been mainly used for generating point forecasts. However, there are two disadvantages of point forecasts compared with PIs. The first disadvantage is that the reliability of point forecasts significantly drops when the level of uncertainty increases. The problem becomes more severe when datasets are multivalued, sparse, noisy or the

target values are affected by some probabilistic events [41]. Nowadays, real world man-made systems become more and more complex, and they have a high level of uncertainty. Point forecasts are more likely to be unreliable and questionable for these applications. Compared with PIs, there is another disadvantage for point forecasts. NN point forecasts only provide point prediction values but convey no information about the prediction accuracy [41]. For example, point forecasts provide only the prediction error but tell nothing about the probability for correct predictions. This makes decision-making more problematic, as limited information is available by predicted values.

Unlike point forecasts, PI is a powerful tool for uncertainty modeling by its nature [42]. By definition, a PI consists of lower and upper bounds that bracket a future unknown target value with a certain probability $((1-\alpha)\%)$ called the confidence level [19]. A typical PI consists of three parts: the upper bound, lower bound and the coverage probability. PIs not only provide a range that targets are highly likely to lie within, but also have an indication of their accuracy called the confidence level. PIs are more reliable and informative than point forecasts for decision makers. Using high quality PIs, the decision makers can confidently draw up future plans, better manage risks, and maximize their benefits.

2.1.2.2 Difference Between Prediction and Confidence Intervals

When discussing PIs, it is necessary to distinguish PIs from confidence intervals (CIs). It is often assumed that targets can be modeled by [41],

$$t_i = y_i + \epsilon_i \tag{2.5}$$

where t_i is the $i-th$ messured target (totally n target). ϵ_i is the noise, also called error, with a zero expectation. The error term moves the target away from its true regression mean, y_i, towards the measured value, t_i. It is assumed that errors are independently and identically distributed. In practice, an estimation of the true regression mean is obtained using a model, $\hat{y}_i$. According to

this, we have,

$$t_i - \hat{y}_i = [y_i - \hat{y}_i] + \epsilon_i \tag{2.6}$$

CIs deal with the variance of the first term in the right hand side of (2.6). They quantify the uncertainty between the prediction, $\hat{y}_i$, and the true regression, y_i. CIs are based on the estimation of characteristics of the probability distribution $P(y_i \mid \hat{y}_i)$. In contrast, PIs try to quantify the uncertainty associated with the difference between the measured values, t_i, and the predicted values, y_i. This relates to the probability distribution $P(t_i \mid \hat{y}_i)$. Accordingly, PIs will be wider than CIs and will enclose them [41].

If the two terms in (2.6) are statistically independent, the total variance associated to the model outcome will become [41],

$$\sigma_i^2 = \sigma_{\hat{y}_i}^2 + \sigma_{\hat{\epsilon}_i}^2 \tag{2.7}$$

The term $\sigma_{\hat{y}_i}^2$ originates from model misspecification and parameter estimation errors, $\sigma_{\hat{\epsilon}_i}^2$ is the measure of noise variance.

CIs and PIs are two well-known tools for quantifying and representing the uncertainty of predictions. While a CI describes the uncertainty in the prediction of an unknown but fixed value, a PI deals with the uncertainty in the prediction of a future realization of a random variable. By definition, a PI accounts for more sources of uncertainty (model misspecification and noise variance) and is wider than the corresponding CI [43].

2.1.3 Fuzzy Logic

Fuzzy logic [20] was first proposed in 1965 by Lotfi Zadeh with the proposal of fuzzy set theory [44]. Fuzzy logic has been applied to many fields, such as control theory, artificial intelligence, power system, scheduling and optimization, signal analysis for tuning and interpretation etc. Fuzzy logic is a form of multi-valued logic or probabilistic logic; it deals with reasoning that is approximate rather than fixed and exact. In the traditional logic theory, binary sets have only two-valued logic, true or false. On the contrary, fuzzy logic variables may have a truth value that

ranges in degree between 0 and 1, so fuzzy logic can also handle partial truth.

2.1.3.1 Fuzzy Logic Systems

A fuzzy logic system (FLS) is defined as a nonlinear mapping of an input data (feature) vector into a scalar output (the vector output case decomposes into a collection of independent multi-input/single-output systems) [44]. The primary tool for doing this is a list of *IF-THEN* statements. All rules are fired in parallel. A typical FLS is show in Fig. 2.2, which consists of four components: rules, fuzzifier, inference engine, and defuzzifier [44]. A FLS maps crisp inputs into crisp outputs. The process of fuzzy logic is explained as follows: firstly, a crisp set of input data are gathered and converted to a fuzzy set using fuzzy linguistic variables, fuzzy linguistic terms and membership functions. This step is known as fuzzification. Afterwards, an inference is made based on a set of rules. The inference engine of the FLS maps fuzzy input sets into fuzzy output sets. Lastly, the resulting fuzzy output is mapped to a crisp output using the membership functions, in the defuzzification step.

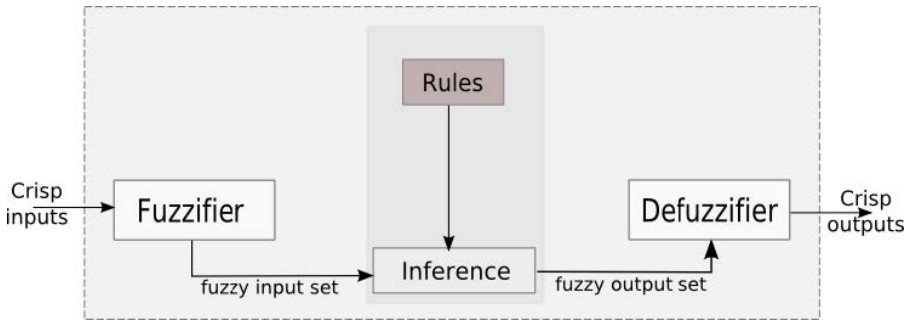

Figure 2.2: A fuzzy logic system and its components.

2.1.3.2 Fuzzy Algorithm Implementation

A fuzzy algorithm can be summarized in the following five steps:

1) *Initialization*. Define the linguistic variables and terms; construct the membership functions and construct the rule base.

2) *Fuzzify the inputs*. Fuzzification is the process of assigning a degree of truth to statements about the input variables (all those statements in the if part, or antecedent, of the rule). The membership functions associated with the input variables determine this degree of truth. Any statement in the antecedent evaluates to a number between 0 and 1.

3) *Fuzzy inference*. Apply implication operator: The implication method is defined as the shaping of the consequent (a fuzzy set) based on the antecedent (a single number). Implication occurs for each rule.

4) *Aggregate output across all rules*. The above operations occur for all rules, and each rule results in a clipped output fuzzy set. Joining all these clipped output fuzzy sets into a single combined output membership function is known as aggregation and it is performed by the aggregation (max) operator.

5) *Defuzzify the aggregated output fuzzy set*. The aggregated membership function needs to be reduced to a crisp value. The defuzzification method returns this value given from the sometimes oddly shaped aggregate.

Linguistic variables [45] are the input or output variables of the system whose values are not numbers but words or sentences in a natural or artificial language. A linguistic variable is generally decomposed into a set of linguistic terms. Take the linguistic values of the temperature as an example, then, $T(t) = \{\text{too-cold, cold, warm, hot, too-hot}\}$ can be the set of decompositions for the linguistic variable temperature. Each member of this decomposition is called a linguistic term and can cover a portion of the overall values of the temperature. Membership functions are used in the fuzzification and defuzzification steps of a FLS, to map the non-fuzzy input values to fuzzy linguistic terms and vice versa. A membership function is used to quantify a linguistic term. There are different forms of membership functions such as triangular, trapezoidal, piecewise linear, Gaussian, or singleton. Fig. 2.3 shows these functions.

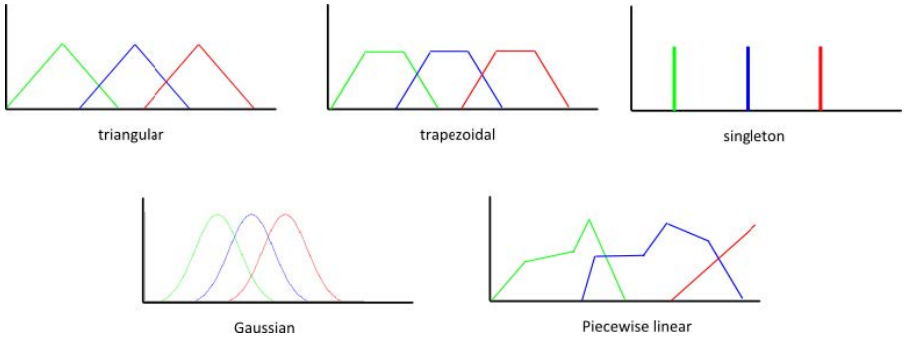

Figure 2.3: Different types of membership functions.

The most common types of membership functions are triangular, trapezoidal, and Gaussian shapes. The type of the membership function can be context dependent and it is generally chosen arbitrarily according to the user experience. Note that, an important characteristic of fuzzy logic is that a numerical value does not have to be fuzzified using only one membership function. In other words, a value can belong to multiple sets at the same time.

Rules may be provided by experts (you may be such a person) or can be extracted from numerical data. In either case, engineering rules are expressed as a collection of *IF-THEN* statements. The rules can also be represented in a matrix form. The evaluations of the fuzzy rules and the combination of the results of the individual rules are performed using fuzzy set operations. The operations on fuzzy sets are different from the operations on non-fuzzy sets. The mostly-used fuzzy set operations are displayed in Table 2.1, where μ_A, μ_B are membership functions for fuzzy sets A and B. After evaluating the result of each rule, these results should be combined to obtain a final result. This process is called inference.

Table 2.1: Fuzzy Set Operations

OR (Union)		**AND (Intersection)**	
MAX	$Max\{\mu_A(x), \mu_B(x)\}$	MIN	$Min\{\mu_A(x), \mu_B(x)\}$
ASUM	$\mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)$	PROD	$\mu_A(x)\mu_B(x)$
BSUM	$Min\{1,\ \mu_A(x) + \mu_B(x)\}$	BDIF	$Max\{0,\ \mu_A(x) + \mu_B(x) - 1\}$

After the inference step, the overall result is a fuzzy value. This result should be defuzzified to obtain a final crisp output. For example, the center of gravity, left/right most maximum algorithms can be used. This is the purpose of the defuzzifier component of a FLS. Defuzzification is performed according to the membership function of the output variable. In a controls application, for example, such a number corresponds to a control action [44].

2.1.4 Stochastic Model

A general description of stochastic programming and the typical procedure of Monte Carlo simulation are reviewed.

2.1.4.1 Stochastic Programming

In the field of mathematical optimization, stochastic programming is a framework for modeling optimization problems that involve uncertainty. Whereas deterministic optimization problems are formulated with known parameters, real world problems almost invariably include some unknown parameters. When the parameters are known only within certain bounds, one approach to tackling such problems is called robust optimization [46]. Here the goal is to find a solution which is feasible for all such data and optimal in some sense. Stochastic programming models are similar in style but take advantage of the fact that probability distributions governing the data are known or can be estimated. The goal here is to find some policy that is feasible for all (or almost all) the possible data instances and maximizes the expectation of some function of the decisions and the random variables. More generally, such models are formulated, solved analytically or numerically, and analyzed in order to provide useful information to a decision-maker [47, 48].

$$\begin{aligned} \text{Minimize} \quad & F_0(x) = \mathbf{E}[f_0(x,\tau)] \\ \text{Subject to} \quad & F_i(x) = \mathbf{E}[f_i(x,\tau)] \leq 0, \quad i = 1, ..., m \end{aligned} \tag{2.8}$$

where the objective and constraint functions $f_i(x,\tau)$ depend on optimization variable x and a random variable τ. The parameter τ may be a random variation in implementation, manufacture

or operation. If $f_i(x, \tau)$ are convex in x for each τ, hence $F_i(x)$ are convex and the stochastic programming problem is convex. $F_i(x)$ have analytical expressions in only a few cases, in other cases the problem is solved approximately. Usually the value of τ is unknown but its distribution is known or can be estimated. The goal here is to choose the decision variable x so that:

- Constraints are satisfied on average, or with high probability;
- The objective is small on average, or with high probability.

2.1.4.2 Monte Carlo Simulation

Monte Carlo simulation is a general method for (approximately) solving stochastic programming problem. Monte Carlo simulation is a type of simulation that relies on repeated random sampling and statistical analysis to compute the results [11]. Monte Carlo methods are often used in computer simulations of physical and mathematical systems. They are used to model phenomena with significant uncertainty in inputs, such as the calculation of risk in business. Monte Carlo methods vary in different applications, but the physical process trends to follow a typical procedure [11]:

1) *Static Model Generation*:

   Every Monte Carlo simulation starts up with developing a deterministic model which closely resembles the real scenario. This is also called the base case, which uses the most likely value of the inputs variables and generates the model outputs.

2) *Input Distribution Identification*:

   In the base case, the deterministic inputs are applied. Since the risks originate from the stochastic nature of the input variables, we try to identify the underlying distributions, if any, which govern the input variables. The historical data are then used to fit the distribution parameters. This distribution fitting is nothing but a nonlinear optimization problem, where the variables are parameters of the distributions.

3) *Random Variable Generation*:

After the input distribution has been identified, a set of random variables are generated from the probability distribution over the domain. That is, to generate N samples (scenarios or realizations) $\tau \in \{\tau_1, \tau_2, ..., \tau_N\}$, with associated probabilities $\pi_j = \textbf{Prob}(\tau = \tau_j)$ (usually $\pi_j = 1/N$). One set of random numbers is one possible representation of the input variables, which will be further used in the deterministic model to provide one set of output values. This process will be repeated times by generating more sets of random numbers, one for each input distribution, and collect different sets of possible output values. This part is the core of Monte Carlo simulation.

4) *Problem Solving*:

Form sample average approximations:

$$\hat{F}_i(x) = \sum_{j=1}^{N} \pi_j f_i(x, \tau_j), \qquad i = 0, ..., m \tag{2.9}$$

These are random variables ($via \ \ \tau_1, ..., \tau_N$) with mean $\mathbf{E}[f_i(x, \tau)] = F_i(x)$.

Now solve finite event problem:

$$\begin{aligned} \text{Minimize} \quad & \hat{F}_0(x) \\ \text{Subject to} \quad & \hat{F}_i(x) \leq 0, \quad i = 1, ..., m \end{aligned} \tag{2.10}$$

Solution $x^\star_{mcs}$ and optimal value $\hat{F}_0(x^\star_{mcs})$ are random variables (hopefully close to $x^\star_{mcs}$ and $p^\star$, optimal value of original problem). Theory [48] says, (with some technical conditions) as $N \to \infty, x^\star_{mcs} \to x^\star$ and $\mathbf{E}[\hat{F}_0(x^\star_{mcs})] \leq p^\star$.

5) *Analysis and Decision Making*:

After a sample of output values have been collected from the simulation, the statistical analysis will be implemented. Some statistical indices such as mean, median, the maximum, minimum, the standard deviation, variance, skewness and kurtosis can be calculated for analysis. The analysis results are then used for decision making.

Take the UC problem with wind power as an example. In the stochastic security constrained unit commitment (SCUC) with high level of wind power penetration, Monte Carlo simulation

is usually used to model the uncertainty of the volatile wind power. The distribution of the wind power generation output can be obtained from a forecasting procedure. The Monte Carlo method is applied to generate a series of scenarios from the distribution. Each scenario represents one possible realization of the wind power output with certain probability. The objective is to minimize the expected cost of overall scenarios [49]. The obtained UC results can then be used to determine the generation scheduling, for example to determine the on/off status of the generators and the output power of each generators.

2.2 Construction of Neural Network-Based Prediction Intervals

Both the traditional PI construction method and the new LUBE method are reviewed. Their disadvantages and advantages are summarized.

2.2.1 Traditional Methods

After an overview of traditional methods, the detailed implementation of delta method is reviewed. The disadvantages of traditional method are provided at the end of this subsection.

2.2.1.1 Overview of Traditional Methods

The delta [50, 51], Bayesian [52], mean-variance estimation [53] and bootstrap [54] techniques are traditional methods to construct NN-based PIs [55]. The cornerstone of the delta technique lies in interpreting NNs as nonlinear regression models and linearizing them based on Taylor's series expansion [56]. The Bayesian technique interprets the NN parameter uncertainty in terms of probability distributions and integrates them to obtain the probability distribution of the target conditional on the observed training set [52]. The mean-variance estimation is implemented through developing two NNs for prediction of mean and variance of targets [54]. The bootstrap technique is essentially a resampling method that its computation requirement is massive. Selection of any of these techniques for constructing PIs depends on problem domain, computation

burden, number of available samples, and analysis purpose.

2.2.1.2 Delta Method

The root theory of delta method is nonlinear regression, which represents the NN as a nonlinear regression model [41, 56]. The method first linearizes the NN model around a set of parameters obtained through minimization of the sum of squared error (SSE) cost function. Then, standard asymptotic theory is applied to the linearized model for constructing PIs [43]. Consider that ω^* is the set of optimal NN parameters that approximates the true regression function, $y_i = f(x_i, \omega^*)$. In a small neighborhood of this set, the NN model can be linearized based on Taylor's series expansion [41, 56],

$$\hat{y}_0 = f(x_0, \omega^*) + g_0^T(\hat{\omega} - \omega^*) \tag{2.11}$$

g_0^T is the NN output gradient against the network parameters, ω^*,

$$g_0^T = \left[\frac{\partial f(x_0, \omega^*)}{\partial \omega_1^*} \;\; \frac{\partial f(x_0, \omega^*)}{\partial \omega_2^*} \;\; ... \frac{\partial f(x_0, \omega^*)}{\partial \omega_p^*}\right] \tag{2.12}$$

where p is the number of NN parameters. In the training process of NN, the NN parameters $\hat{\omega}$ are adjusted through the minimization of the SSE cost function. Under certain regularity conditions, it can be shown that $\hat{\omega}$ is very close to ω^*. Accordingly, we have,

$$t_0 - \hat{y}_0 \approx [y_0 + \epsilon_0] - \left[f(x_0, \omega^*) + g_0^T(\hat{\omega} - \omega^*)\right] = \epsilon_0 - g_0^T(\hat{\omega} - \omega^*) \tag{2.13}$$

where t_0 and ϵ_0 are the same meaning as in (2.5). They are the target values and the noise or error terms respectively. Then the variance is calculated as,

$$var(t_0 - \hat{y}_0) = var(\epsilon_0) + var(g_0^T(\hat{\omega} - \omega^*)) \tag{2.14}$$

Assuming that the error terms are normally distributed$\big(\epsilon \approx N(0, \sigma_\epsilon^2)\big)$, the second term in the right hand side of (2.14) can be expressed as,

$$\sigma_{\hat{y}_0}^2 = \sigma_\epsilon^2 \; g_0^T (J^T \; J)^{-1} g_0 \tag{2.15}$$

J in (2.15) is the Jacobian matrix of the NN model with respect to its parameters computed for the training samples [41],

$$J = \begin{bmatrix} \frac{\partial f(x_1,\hat{\omega})}{\partial \hat{\omega}_1} & \frac{\partial f(x_1,\hat{\omega})}{\partial \hat{\omega}_2} & \dots & \frac{\partial f(x_1,\hat{\omega})}{\partial \hat{\omega}_p} \\ \frac{\partial f(x_2,\hat{\omega})}{\partial \hat{\omega}_1} & \frac{\partial f(x_2,\hat{\omega})}{\partial \hat{\omega}_2} & \dots & \frac{\partial f(x_2,\hat{\omega})}{\partial \hat{\omega}_p} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f(x_n,\hat{\omega})}{\partial \hat{\omega}_1} & \frac{\partial f(x_n,\hat{\omega})}{\partial \hat{\omega}_2} & \dots & \frac{\partial f(x_n,\hat{\omega})}{\partial \hat{\omega}_p} \end{bmatrix} \tag{2.16}$$

By replacing (2.15) in (2.14), the total variance can be expressed as,

$$\sigma_0^2 = \sigma_\epsilon^2 \left(1 + \ g_0^T (J^T \ J)^{-1} g_0\right) \tag{2.17}$$

An unbiased estimate of σ_ϵ^2 can be obtained from,

$$s_\epsilon^2 = \frac{1}{n-1} \sum_{i=1}^{n} (t_i - \hat{y}_i)^2 \tag{2.18}$$

According to this, the $(1-\alpha)\%$ PI for $\hat{y}_i$ is computed as detailed in [50],

$$\hat{y}_0 \pm t_{n-p}^{1-\frac{\alpha}{2}} s_\epsilon \sqrt{1 + g_0^T (J^T \ J)^{-1} g_0} \tag{2.19}$$

where $t_{n-p}^{1-\frac{\alpha}{2}}$ is the $\frac{\alpha}{2}$ quantile of a cumulative t-distribution function with $(n-p)$ degrees of freedom.

In order to minimize the overfitting problem and to improve the generalization power of NN, the Weight Decay Cost Function (WDCF) can be used instead of the SSE cost function. The WDCF tries to keep the magnitude of the NN parameters as small as possible,

$$WDCF = SSE + \lambda \, \omega^T \, \omega \tag{2.20}$$

De Veaux et al. derived the following formula for PI construction for the case that NNs are trained using the WDCF [51],

$$\hat{y}_0 \pm t_{n-p}^{1-\frac{\alpha}{2}} s_\epsilon \sqrt{1 + g_0^T (J^T \ J + \lambda I)^{-1} (J^T J) (J^T \ J + \lambda I)^{-1} g_0} \tag{2.21}$$

Inclusion of λ improves the reliability and quality of PIs, particularly for cases where $J^T J$ is nearly singular. This comes to the end of delta method for constructing NN-based PIs. Computationally, the delta technique is more demanding in its development stage than its application

stage. Due to the special assumptions on error distribution, the quality of the generated PIs is low for noisy data.

2.2.1.3 Disadvantages of Traditional Methods

In spite of advantages of PIs, applications of these methods are still less popular than NN point forecasts. One important reason is that the implementation of these methods is complex. For instance, the delta and Bayesian methods need to calculate the Jacobian matrix and Hessian matrix of the parameters separately [57]. In each iteration, the Jacobian or Hessian matrix needs to be updated, which is very time consuming. Also calculation of derivatives suffers from singularity problems that decrease the reliability of PIs. On the other hand, traditional methods make assumptions about the data distribution. The delta method assumes that the noises are normally distributed and t-distribution is applied [50]. Mean-variance estimation method assumes that NN (predicting the mean) can precisely estimate the true mean of the targets. If this assumption is violated, the NN generalization power is weak and it results in accumulation of uncertainty in the estimation of the target value and leads to a low coverage probability. Bootstrap method assumes that an ensemble of NN models will produce a less biased estimate of the true regression of the targets [54]. The main disadvantage is that the total variance will be underestimated resulting in narrow PIs with a low coverage probability. Implementation difficulties, special assumption about the data distribution, and massive computational requirements hinder widespread applications of these methods for decision-making [58].

2.2.2 Lower Upper Bound Estimation (LUBE) Method

Lower upper bound estimation (LUBE) method is a new proposed method for construction of PIs in [43]. Its concept is different from traditional methods. Obvious advantages are shown compared to traditional methods.

2.2.2.1 Concept of LUBE Method

The basic concept of LUBE method is to adopt a NN with two outputs to directly generate the upper and lower bounds of PIs. The first and second outputs correspond to the upper and lower bounds of PIs separately. The symbolic NN with two outputs for the LUBE method is shown in Fig. 2.4 [58]. The real NN architectures will vary in different applications. For a typical three-

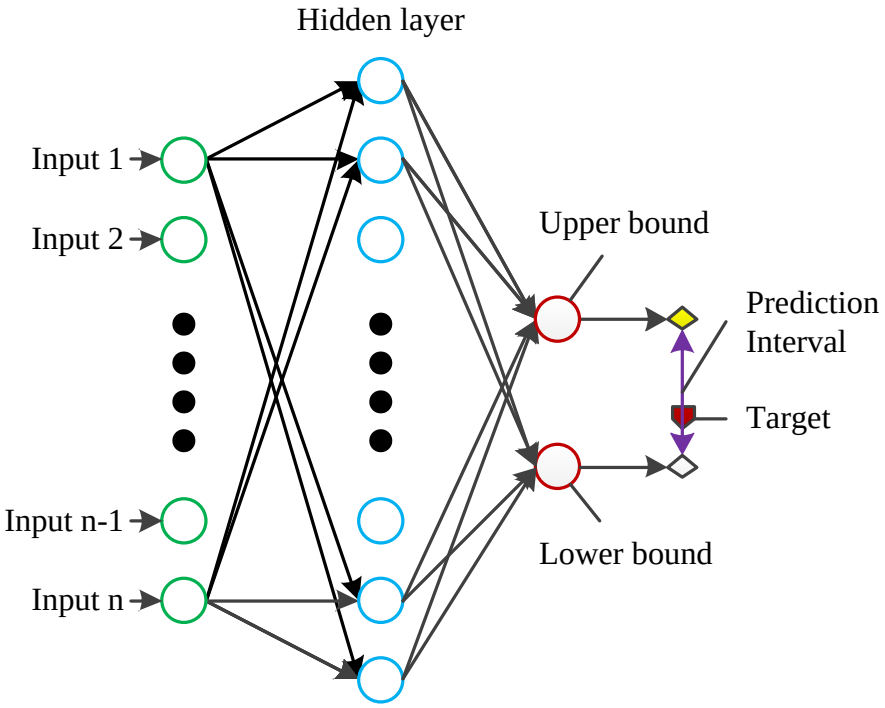

Figure 2.4: NN model for LUBE method to generate upper and lower bounds of PIs

layered NN, the mathematical mapping between the inputs and the outputs is shown in Formula (2.22) [59]:

$$y_i = f_1\left(\sum_{j=1}^{N_h}\left(w_{ij} f_2\left(\sum_{k=1}^{N_i} v_{jk}x_k + b_{vj}\right) + b_{wi}\right)\right), i = 1, 2, ..., N_o \tag{2.22}$$

where y_i is the output of the ith node on the output layer; x_k is the input of the kth node in the input layer; w_{ij} represents the connection weight between nodes in the hidden and output layers; v_{jk} is the connection weight between nodes in the input and hidden layers; and b_{wi} and b_{vj} are bias terms that represent the threshold of the transfer function f_1 and f_2. The number of nodes in the input, hidden and output layers are N_i, N_h and N_o respectively. In the LUBE method,

$N_o = 2$, and y_1, y_2 correspond to the upper and lower bounds of PIs.

2.2.2.2 Advantages of LUBE Method

Traditional methods for construction of PIs attempt to estimate the mean and variance of the target datasets. They construct PIs in two steps [55]:

1) They regress the given dataset to a specified model or function, which is the same as point forecasts;

2) According to the assumed data distribution, the statistical mean and variance values are calculated, if Jacobian or Hessian matrix are needed, they are also calculated at this step. Based on this information, PIs are then constructed.

In contrast, the LUBE method directly generates the upper and lower bounds of PIs in one step. Then PI construction is as simple as point forecasting. Compared with delta, Bayesian and bootstrap methods, LUBE method has the following advantages [58, 60]:

1) Makes no assumptions about the dataset.

2) Is simpler and avoids calculation of derivatives of NN output with respect to its parameters, such as Jacobin or Hessian matrix. There is no singularity problem, thus it increases the reliability of the PI construction.

3) The computational load is much lower than other methods. Once the NN is trained and the optimal weights are fixed, PI construction is as simple as point forecasting.

2.3 Unit Commitment (UC) with Intermittent Renewable Energy Sources

The UC with IRESs is different from traditional UC scheduling. IRESs such as wind and solar power have penetrated into the traditional power systems. If the output power of IRESs is re-

garded as a negative load, the net load is the load minus the output power of IRESs. The power balance constraints is changed to meet the net load balance. The potential risks of the uncertain IRESs need to be addressed.

2.3.1 Overview of the UC Problem Formulation

There are mainly three types of formulation methods for the UC problem [61]. The first type of UC is the traditional one, which is usually formulated to minimize the total operating cost under various constraints. It is also called Security Constrained Unit Commitment (SCUC). In some deregulated markets, some companies change the objective function from cost minimization to profit (revenue-operational cost) maximization. Beside this, they also change the demand constraints from an equality to less than or equal. This UC is called Price-Based Unit Commitment (PBUC) [62], which consists of the second type of UC formulation. With the increase in penetration of renewable energy, the third type of UC is UC of Power System with Renewable Energy Sources and Storage System. These renewable sources include wind power and solar power generation, the battery storage system etc, to name a few. Although these three types of formulation methods have different objective functions, they all can be summarized as an optimization problem with several constraints. Here the deterministic SCUC is illustrated as an example for problem formulation.

The objective of the SCUC problem is to minimize the total generation costs under several constrains. Mathematically, the objective function is formulated as follows [61, 63, 64]:

$$E(X,P) = \sum_{i=1}^{N} \sum_{t=1}^{H} \left[F_i(P_{i,t}) + SU_{i,t}(1 - X_{i,(t-1)}) \right] X_{i,t} \tag{2.23}$$

where $F_i(P_{i,t})$ of thermal generators is usually represented as the quadratic function:

$$F_i(P_{i,t}) = a_i + b_i P_{i,t} + c_i P_{i,t}^2 \tag{2.24}$$

The start up cost is determined according to the continuously off time of generators:

$$SU_{i,t} = \begin{cases} HSU_{i,t}, & if\ T_{i,t}^{off} \leq T_i^{Down} + T_i^{cold}; \\ CSU_{i,t}, & if\ T_{i,t}^{off} > T_i^{Down} + T_i^{cold}. \end{cases} \tag{2.25}$$

Subject to the following constraints:

1) Power balance constraints:

$$\sum_{i=1}^{N} X_{i,t} P_{i,t} = D_t - IRES_t \tag{2.26}$$

2) Spinning reserve constraints:

$$\sum_{i=1}^{N} X_{i,t} P_{i,max} \geq D_t + R_t \tag{2.27}$$

or in another form as:

$$\sum_{i=1}^{N} X_{i,t} \left[P_{i,max} - P_{i,t} \right] \geq R_t \tag{2.28}$$

3) Generation limit constraints:

$$P_{i,min} X_{i,t} \leq P_{i,t} \leq P_{i,max} X_{i,t} \tag{2.29}$$

4) Minimum up time constraints:

$$(T_{i,t}^{on} - T_i^{Up})(X_{i,(t-1)} - X_{i,t}) \geq 0 \tag{2.30}$$

where,

$$T_{i,t}^{on} = (T_{i,(t-1)}^{on} + 1) X_{i,t} \tag{2.31}$$

5) Minimum down time constraints:

$$(T_{i,t}^{off} - T_i^{Down})(X_{i,t} - X_{i,(t-1)}) \geq 0 \tag{2.32}$$

where,

$$T_{i,t}^{off} = (T_{i,(t-1)}^{off} + 1)(1 - X_{i,t}) \tag{2.33}$$

6) Ramp up rate constraints [64]:

$$P_{i,t} - P_{i,(t-1)} \leq UR_i \tag{2.34}$$

7) Ramp down rate constraints [64]:

$$P_{i,(t-1)} - P_{i,t} \leq DR_i \tag{2.35}$$

8) Transmission flow limits from bus k to bus m:

$$-L_{km}^{max} \leq L_{km,t} \leq L_{km}^{max} \tag{2.36}$$

where, L_{km}^{max} is maximum allowable real power flow through transmission line km.

Besides the above constrains, there are various other constrains. These constrains include but not limited to, system emission limit, must run units, must out units, crew constraints, fuel constrains etc [65].

2.3.2 UC Solution Methods

UC solution methods can be classified into three categories: deterministic methods, meta-heuristic methods and hybrid models. Deterministic methods include priority listing (PL) [66], dynamic programming (DP) [66], Lagrangian relaxation (LR) [67], integer and linear programming [68]. Meta-heuristic approaches include expert or fuzzy systems [69, 70], genetic algorithm (GA) [71, 72], evolutionary programming, simulated annealing, tabu search, PSO [73, 74], ant colony optimization and differential evolution [75]. Hybrid models, where one method compensates with another, may have a better performance than individual models. Srinivasan et al. in [76] proposed a PL method-based evolutionary algorithm to solve the UC problem. PL was used as a good initialization of the evolutionary algorithm. In [63], the authors proposed a hybrid LRGA model; GA was applied to evolve the Lagrange multipliers to improve the LR method.

2.3.2.1 Deterministic Methods

Exhaustive enumeration is the earliest method used to solve the UC problem. It enumerates all possible combinations of the generating units and then the combinations that yield the least cost of operation are chosen as the optimal solution. Even though the method was not suitable for a large size electric utility, it was capable of providing an accurate solution [66] .

PL method initially arranges the generating units based on lowest operational cost characteristics. The predetermined order is then used for UC such that the system load is satisfied [65]. PL is the simplest and fastest but achieves poor final solution.

DP holds the Theorem of optimality that "An optimal policy must contain only optimal sub-policies." [65]. It has the advantage of being able to solve problems of a variety of sizes and to be easily modified to model characteristics of specific utilities [66]. But its computation time suffers from the curse of dimensionality, which leads to more mathematical complexity and increase in computation time [61].

LR is based on a dual optimization theory. It solves the UC problem by "relaxing" or temporarily ignoring the coupling constraints and solving the problem as if they did not exist. This is done through the dual optimization procedure. The dual procedure attempts to reach the constrained optimum by maximizing the Lagrangian with respect to the Lagrange multipliers, while minimizing with respect to the other variables in the problem [65]. Fig. 2.5 shows a typical LR method for solving the UC problem [77], where λ and μ are Lagrange multipliers for the equality and inequality constraints respectively.

2.3.2.2 Meta-Heuristic Approaches

Meta-heuristic approaches are computational intelligence methods inspired by natural evolution or social behavior. These methods include expert systems [69], fuzzy systems [70], GA [71, 72], evolutionary programming, simulated annealing, tabu search, PSO [73], ant colony optimization and differential evolution [75].

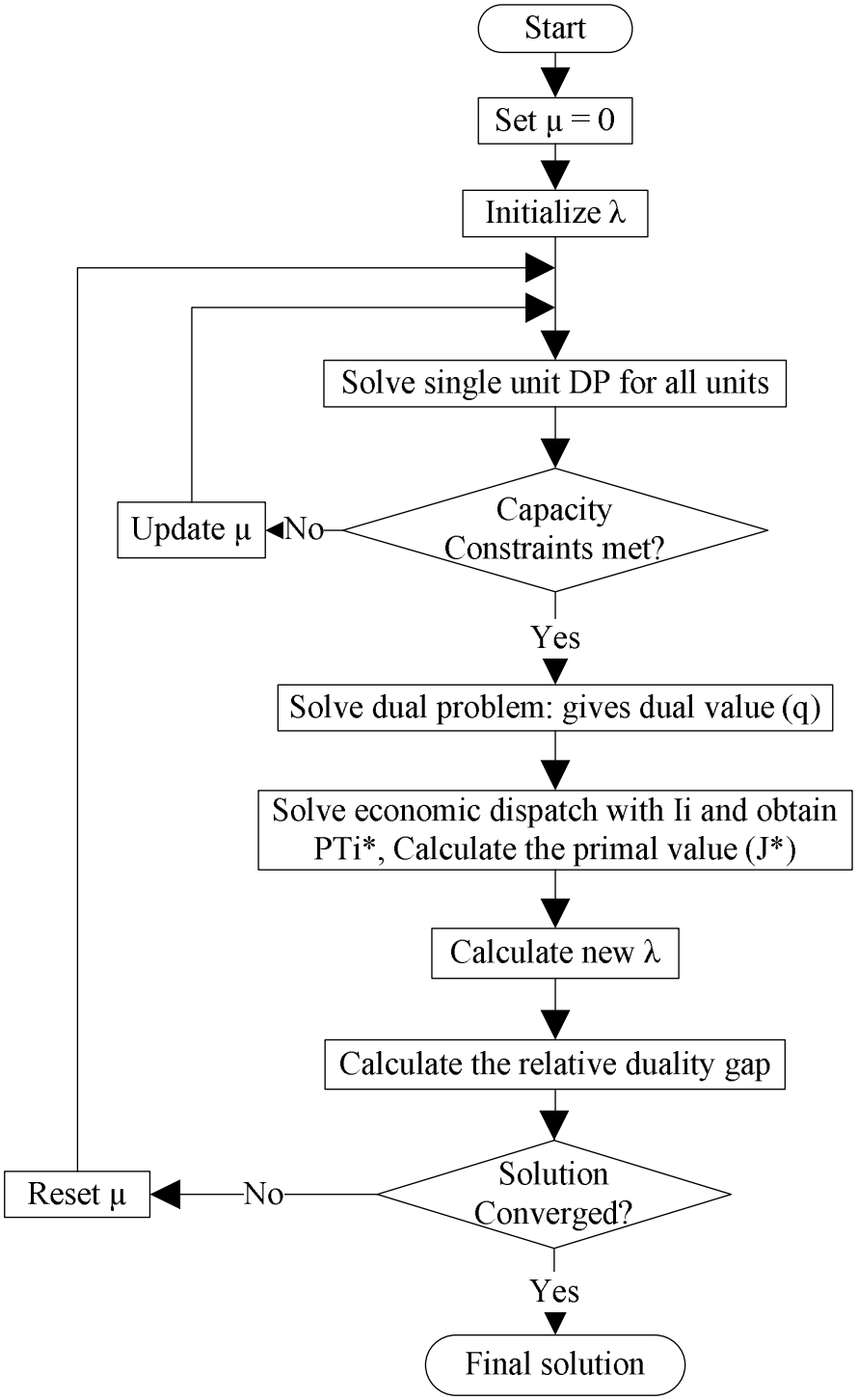

Figure 2.5: LR method to solve UC

GA mimics the natural evolutionary process. It encodes the UC variables as the chromosomes, each population represents one potential solutions of the problem. A fitness function is used as criteria for evaluation of the goodness of the solution. The objective function involves

penalty term to penalize those potential solutions in which the problem constraints are not fulfilled. Selection, crossover and mutation are the three main operators of GA. It has potential of obtaining near-global minimum and the capability of obtaining the solution within short time and the constraints can be easily included [71].

PSO is inspired by a swarm of birds to find food. Each individual in PSO flies in the search space with a velocity which is dynamically adjusted according to its own flying experience and its companions' flying experience. Personal best (p_{best}) denotes the best performance of the individual particle, while global best (g_{best}) is the best performance in the whole swarm. Through the velocity and position update, each particle shares the information efficiently throughout the whole swarm. Attention on PSO in UC problem is increased because of its better convergence and small simulation time [58, 73].

2.3.2.3 Hybrid Models

As mentioned above, each method has its own advantages and disadvantages. Some researchers develop hybrid models to further improve the solution of the UC problem. In the hybrid models, one method compensates with another, they may have better performance than individual models.

Srinivasan, D et al. in [76] proposed a PL-based evolutionary algorithm to solve the UC problem. PL was used as a good initialization of the evolutionary algorithm. The authors further studied the effect of the number of PL solution in the initial population. Results showed that best results were obtained with only 1 PL solution in the initial population, remaining solutions being generated randomly. In [63, 77], the authors proposed a hybrid LRGA model; GA was applied to evolve the Lagrangian multipliers to improve the LR method. The LRGA method consisted of a two-stage cycle. The first stage was to search for the constrained minimum of Lagrangian function under constant Lagrangian multipliers by two-state dynamic programming. The second stage was to maximize the Lagrangian function with respect to Lagrangian multipliers adjusted by GA [63, 77].

2.4 Incorporating Wind Generation Forecast Uncertainties into UC

Wu et al. [49] applied two methods to consider the uncertainty of wind power generation; comparisons between the two methods are also implemented in this paper. The two methods are stochastic SCUC and interval optimization approaches. In stochastic SCUC, multiple scenarios are generated using Monte Carlo method to simulate the possible realization of wind power generation uncertainties. In addition to the base case operation cost, the scenario-based approach minimizes expected costs of corrective actions. Instead of sampling scenarios, the interval optimization uses confidence intervals in terms of upper and lower bounds to represent the uncertainty spectrum, and derives optimistic and pessimistic solutions for satisfying the system security requirement [49].

Wang et al. [78] also incorporated wind power generation uncertainty using stochastic SCUC. The wind power uncertainty in this paper emphasizes on the aspects of intermittency and volatility. Monte Carlo method is employed to simulate possible wind power volatility scenarios. The results for the six-bus system and the IEEE 118-bus system show that the physical limitations of units such as ramping are crucial for accommodating the volatility of wind power generation.

Ortega-Vazquez et al. [79] estimated the spinning reserve requirements in systems with significant wind power generation penetration. Wind power generation is viewed as a negative load, the uncertainty on this generation increases the uncertainty on the net demand (i.e., the system wide demand minus the wind power generation). Net demand takes into account the load forecast errors and wind power generation uncertainty together. The results show that, contrary to what is commonly believed, an increased wind power penetration does not necessarily require larger amounts of spinning reserve.

Botterud et al. [80, 81] presented a new model for optimal trading of wind power in day-ahead electricity markets under uncertainty in wind power and prices. A wind power producer is facing three main uncertainties when bidding into the day-ahead market: the day-ahead locational

marginal prices (LMP), the real-time LMP, and the real-time delivery of wind power. A stochastic model is developed for the bidding strategy.

Sturt et al. [82] used stochastic UC with rolling planning for simulation of wind-integrated power systems. There are three main contributions. First, they present a new formulation of the stochastic UC problem. Second, the proposed formulation uses a quantile-based scenario tree structure. Third, the authors compare the performance of various tree topologies in year-long simulations of a large system.

Matos et al. [34] considered various uncertainties when setting operating reserves. These uncertainties come from load, conventional generation and wind generation uncertainties etc. Load uncertainty is represented as a Gaussian distribution model, conventional generation as capacity outage probability table. Wind generation uncertainties take into account both forecast errors and wind turbines' unplanned outages. The forecast errors can be represented in three ways. A nonparametric probabilistic forecast can be represented by quantiles, intervals or PDFs. The other two representations take the form of risk indices of the forecasts, and scenarios incorporating temporal or spatial interdependence structure of prediction errors.

Jiang et al. [83] studied the robust UC with wind power and pumped storage hydro. In the beginning, the authors presented some problems of stochastic SCUC to accommodate wind power uncertainty. For example, problem size and computational requirement increase with the number of scenarios; wind power ramp event cannot be represented in stochastic UC etc. Instead, the robust UC represents the wind power uncertainties in confidence intervals. The algorithm can automatically simulate the worst scenario where the wind power output changes between upper and lower bounds and has the greatest impact on system operations. Robust UC also does not require the distribution of the wind power output. Benders' decomposition algorithm is used to obtain a robust unit commitment solution. Interested reader can find more about robust UC from the relevant studies [84–88].

In spite of the advantages of the above references, the wind power scenarios in the stochastic

model are mainly generated from an assumed probability distribution, either a normal distribution [78, 79] or a Weibull probability distribution [49]. The robust UC needs to predefine the uncertainty set and find the worst-case scenario [83, 84, 86]. A typical PI consists of three components: an upper bound, a lower bound and the coverage probability. It is difficult to directly apply the PIs for computational purpose due to this multivalued problem. There is a gap between the PIs and the scenarios in the stochastic model. Previous studies mainly focus on one or two aspects of the uncertainties, such as the load and wind power forecast uncertainties. However, other system uncertainties still exist such as the solar power forecast uncertainty, generator outage uncertainty, etc. An integration framework is therefore needed to integrate different sources of uncertainty together for decision making and risk assessment.

Chapter 3

Construction of Neural Network-Based Prediction Intervals for Forecast Uncertainty Modeling

3.1 Introduction

Since neural networks (NNs) were first proposed by McCulloch and Pitts in 1943, they have been successfully applied to different areas. Multilayer feedforward NNs are theoretically universal approximators [89]. Due to the strong approximation capacity and learning ability, NNs are suitable for prediction and regression problems. In literatures there are numerous applications, such examples include, but not limited to, electric power systems [90–92], transportation systems [93], pattern recognition [94–96], and financial price forecasting [97].

As mentioned in Subsection 2.1.2, PIs are powerful tools to quantify the uncertainties associated with point forecasts. Compared to point forecasts, PIs are more reliable and can provide more information during forecasting for decision making. Traditional methods for construction of NN-based PIs are delta [50, 51], Bayesian [52], mean-variance estimation [53] and bootstrap [54] techniques. Subsection 2.2.1 has summarized some drawbacks of traditional methods,

such as implementation difficulties, special assumption about the data distribution, and massive computational requirements.

Khosravi et al. in [43] proposed a new method for construction of PIs called lower upper bound estimation (LUBE) method. LUBE method applies a NN with two outputs to directly generate the upper and lower bounds of PIs. This method makes no assumptions about data distribution and avoids calculation of derivatives such as Jacobian and Hessian matrix. Therefore, it is easy for implementation and fast in generating PIs. Comparative results reported in [43] reveal that the LUBE method is simpler, faster, and more reliable than traditional methods for PI construction. For high quality PIs in real-world applications, a higher coverage probability and narrower width are always expected. After introducing two quantitative PI evaluation indices, specially, a new width evaluation index that is suitable for NN training is proposed. A comprehensive measure to evaluate both coverage probability and width of PIs is also developed. This measure was successfully applied to different areas, such as electrical load forecasting [56], travel time prediction [42] and industrial systems [98].

The measurement cost function is nonlinear, complex, discontinuous and nondifferentiable. Traditional derivative-based algorithms cannot be applied for its minimization. In [43], simulated annealing algorithm is applied to minimize the cost function. Obtained results can be significantly improved in case of using more powerful optimization methods such as the PSO algorithm. Classical NN training algorithm such as backpropagation, is gradient-based and may be trapped into local optima. Compare to other evolutionary methods such as GA, the advantage of PSO is its algorithmic simplicity. PSO is more powerful for parameter optimization (real number), especially for optimization of NN connection weights in numerous literatures. The mutation operator, which helps to achieve diversity in evolution algorithm, is also integrated into PSO to improve the exploratory capabilities and help to jump out of local optima.

The main contributions of the chapter are listed below. 1) A new PI width evaluation index, which is suitable for NN training, is proposed to help improve the quality of PIs. 2) PSO

associated with mutation operator is firstly integrated into the LUBE method called the PSO-based LUBE method. This PSO associated with mutation operator has a very strong searching ability. 3) Demonstrated results from six case studies indicate that the quality of PIs has been significantly improved compared with the same case studies of two journal papers [41,43]. 4) PI construction time of the PSO-based LUBE method is also much shorter than both the delta [41] and LUBE [43] method.

The rest of this chapter is organized as follows. Section 3.2 describes indices used for PI evaluation. Section 3.3 introduces the proposed PSO-based LUBE method. Case studies, results and discussions are presented in Section 3.4 and Section 3.5 respectively. Finally, Section 3.6 concludes the whole chapter with some remarks of future work in this domain.

3.2 PI Evaluation Indices

Before discussing the evaluation indices for PIs, the evaluation indices for point forecasts are reviewed. Two frequently used evaluation indices are mean square errors (MSE) and mean absolute percentage errors (MAPE). Both MAPE and MSE are used as the cost function in the learning algorithms such as back propagation (BP) for training NN parameters. These indices are also calculated for examining performance of models for test samples. Likewise, PIs, which estimate the lower and upper bounds for targets with a prescribed confidence level ((1-α)%), need to be evaluated.

In literatures, two indices have been introduced for evaluation of PI qualities. These are based on coverage probability and width of intervals.

3.2.1 PICP

Prediction interval coverage probability (PICP) is considered as the fundamental feature of the PIs. It measures which percentage of targets is covered by intervals. The definition of PICP is as

follows [56]:

$$PICP = \frac{1}{n}\sum_{i=1}^{n} c_i \tag{3.1}$$

where n is the number of samples, $c_i = 1$ if the target $y_i \in [L_i, U_i]$, otherwise $c_i = 0$. L_i and U_i are the lower and upper bound of the target y_i. To have valid PIs, PICP should be at least equal to or greater than the nominal confidence level of PIs. Otherwise, PIs are invalid and should be discarded. The ideal case for PICP is $PICP = 100\%$, which means all targets lie within PIs.

3.2.2 PINAW and PINRW

If the width of intervals is large enough, the requirement for high PICP can be easily satisfied. But on the other hand, too wide intervals convey little information about the targets, which is of no use for decision making. According to this, it is important to evaluate PIs based on their widths as well. In literatures, prediction interval normalized average width (PINAW) has been introduced [98]. PINAW is the other aspect for higher quality of PIs, which has an overall evaluation of the width of PIs,

$$PINAW = \frac{1}{nR}\sum_{i=1}^{n}(U_i - L_i) \tag{3.2}$$

where R is the range of the underlying targets. Originally, only prediction interval average width (PIAW) is defined. The purpose of normalization is to compare the average widths for datasets of different ranges. PINAW is the percentage presentation of PIAW to the range of the underlying targets.

Inspired by the successful performances of MSE for training, a new width evaluation index for PIs, called prediction interval normalized root-mean-square width (PINRW), is proposed for NN training in this chapter:

$$PINRW = \frac{1}{R}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(U_i - L_i)^2} \tag{3.3}$$

L_i and U_i and R are the same as in PINAW. Unlike PINAW presents the normalized average width of PIs, the PINRW is the normalized root-mean-square width of PIs. PINAW and PINRW

correspond to the 1-norm and 2-norm of the width of PIs. The format of PINAW is similar to MAPE for point forecasting, and it gives equal weights to all widths of PIs. In contrast, the new index PINRW is functionally similar to MSE, and magnifies wider intervals. In practice, experiment results show that PINRW trends to obtain better results than PINAW, just as MSE is a much better cost function for training than MAPE. Next in this chapter PINRW is used for NN training and PINAW for testing.

3.2.3 Comprehensive Evaluation of PIs

Larger PICP and narrower PINRW are essential properties of higher quality PIs. These properties are conflicting, as requesting a greater PICP will always result in wider PINRW, and narrowing PINRW may lead to an unsatisfactorily low PICP. From the optimization perspective, this is a two-objective problem. For the purpose of simplification and comprehensive comparisons for different PIs, the primary multi-objective is transformed into a single-objective one using cost function.

One successful PI-based cost function is developed in [98], which has a comprehensive balance between both the PICP and PINAW. Here the PINAW is substituted by PINRW in the training process. This coverage width-based criterion (CWC) cost function is defined as:

$$CWC = PINRW(1 + \gamma(PICP)e^{-\eta(PICP-\mu)}) \tag{3.4}$$

where $\gamma(PICP) = 1$ for training. For test samples, $\gamma(PICP)$ is a step function which entirely depends on PICP:

$$\gamma(PICP) = \begin{cases} 0, & PICP \geq \mu; \\ 1, & PICP < \mu. \end{cases} \tag{3.5}$$

η and μ are two constant hyper-parameters that determine the penalty term in case of unsatisfying the required condition. μ is the nominal confidence level associated with PIs and can be set to $(1 - \alpha)$. η exponentially magnifies the difference between the PICP and μ. It is always set to a large value to give a heavy penalty of violation of PICP to the given μ.

At the beginning of the training process, PICP is always very low, the cost function CWC gives an extremely heavy penalty of unsatisfying the assigned μ. Due to $\gamma(PICP)e^{-\eta(PICP-\mu)}$ is a decreasing function of $(PICP - \mu)$, the penalty term exponentially decreases along with the increase of PICP. When the PICP reaches around μ, there is a balance between the PINRW and PICP. CWC provides a comprehensive evaluation of both evaluation indices and tries to find a trade off between informativeness and validity of PIs. This conservative approach for PICP leaves a slack for the test samples [43].

While $\gamma(PICP)$ is a step function for testing samples. This is because, as mentioned in Subsection *2.1*, PICP determines the validity of PIs. Once the PICP is no less than the nominal confidence level ((1-α)%), these PIs are considered as equally valid. Then the comparison between the two CWCs reasonably pays more attention to the narrower width (PINAW). When evaluating PIs for testing, $\gamma(PICP) = 0$ in the step function (5) gives equal measurements of PICP if it is no less than assigned μ; otherwise the corresponding penalty will be accounted by CWC through setting $\gamma(PICP) = 1$.

3.3 PSO-based LUBE Method

The proposed PI construction method is called the PSO-based LUBE method. The detailed implementation of the proposed method is provided. Determination of some important experimental parameters such as the optimal NN structure is demonstrated.

3.3.1 LUBE Method

A detailed description of LUBE method [43] has been introduced in Subsection 2.2.2. The basic concept of LUBE method is to adopt a NN with two outputs to directly generate the upper and lower bounds of PIs. Traditional methods for construction of PIs attempt to estimate the mean and variance of the target datasets. They construct PIs in two steps, first do point forecasts, second the mean and variance values of point forecast errors are applied for PI construction. In

contrast, the LUBE method directly generates the upper and lower bounds of PIs in one step. Then PI construction is as simple as point forecasting. Interesting readers can find more about LUBE method in Subsection 2.2.2.

3.3.2 Optimal NN Structure

Quality of PI is sensitive to the structure of NNs. The question is how to determine layers and neurons per layer. Too small or large NNs have a low generalization power and usually suffer under-fitting or over-fitting problems. The determination of the NN structure is key to the successful construction of high quality PIs. How to determine this optimality is still an open question. For the case of point forecasting, several methods such as cascade correlation [99], hybrid evolutionary neural network construction [100], network pruning [101–103], and k-fold cross validation [104] have been discussed in literatures.

In this chapter, fully connected feed-forward three-layered NNs are chosen and the number of neurons in the hidden layer is changed from 1 to 20. A k-fold cross validation method is applied to determine the optimal NN structure. This method relies on a k-fold cross correlation. To keep the training and test set independent, the k-fold cross validation is implemented on the training set. The whole training set is divided equally into k complementary folds, out of which, usually k-1 are used for training the candidate NNs, and the remaining fold is used for validation [104]. To avoid biased sampling, each structure of NN is trained and validated for k times using k different sub-training and validation datasets. A 5-fold cross validation is implemented here. Traditionally error-based measures such as MAPE and MSE are used for selecting the optimal structure of NNs. The focus of this chapter is on construction of high quality PIs. Therefore, it is more reasonable to determine the optimal structure of NNs directly using PI-based evaluation indices, such as PICP, PINAW and CWC. Determination of the optimal NN structure needs to balance between the network complexity, generalization and learning capacity of NNs.

3.3.3 Implementation of the Proposed PSO-Based LUBE Method

The flow chart of the proposed PSO-based LUBE method is shown in Fig. 3.1.

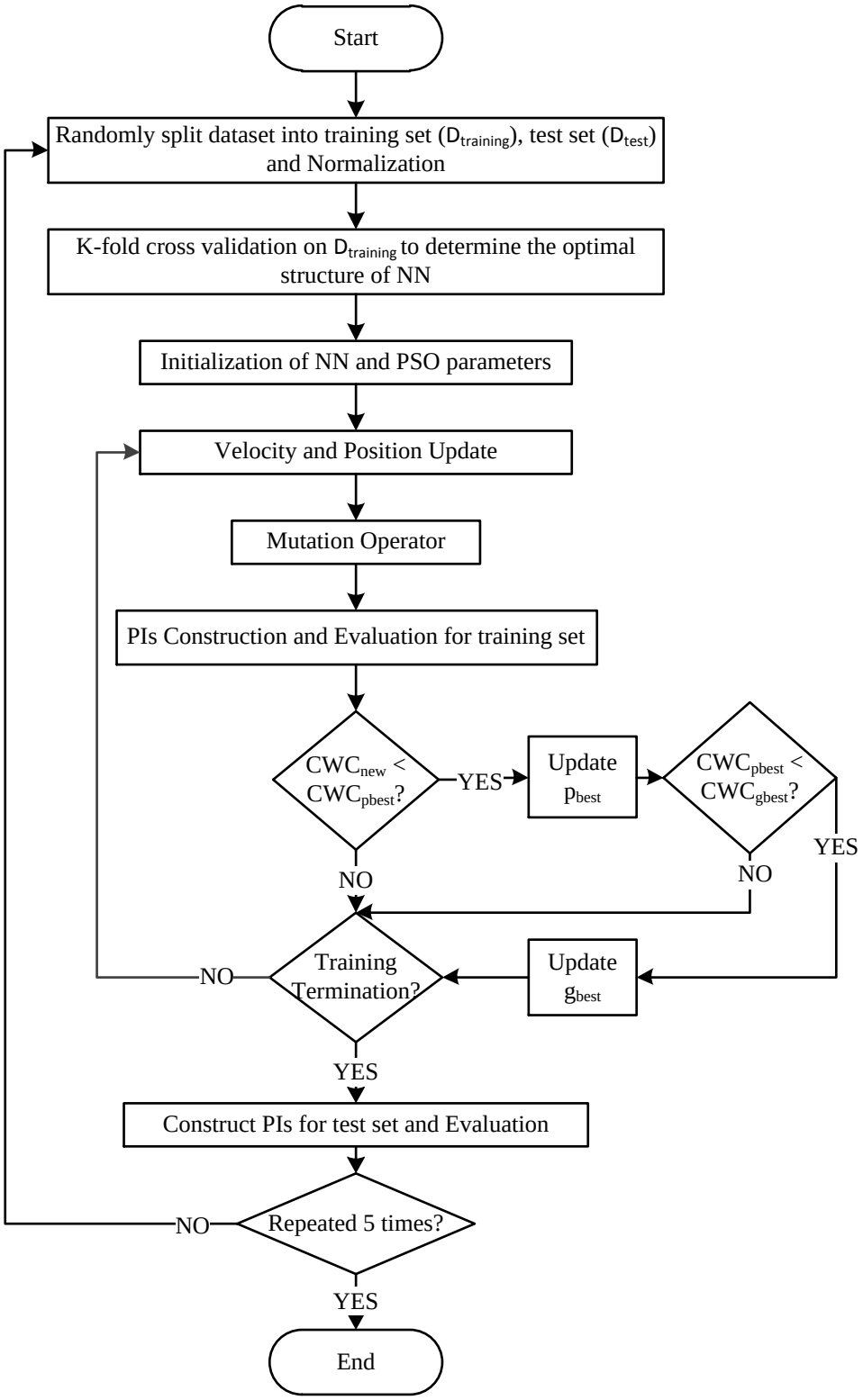

Figure 3.1: PSO-based LUBE method for construction and evaluation of PIs

3.3.3.1 Data Splitting

In each case study, the dataset is randomly split into training set and test set. 70% of the whole dataset is used for training and 30% for testing. After splitting, the training dataset is normalized to [-1, 1], and the same settings are applied to test set for normalization.

3.3.3.2 Determination of the Optimal NN Structure

The training set ($D_{training}$) is further split into sub-training sets and validation sets using the 5-fold cross validation method. Median value of CWC is used to determine the optimal structure of NNs.

3.3.3.3 Initialization

This consists of NNs and PSO parameter initialization. NN weights initialization is very important for this algorithm. A bad initialization may never converge to a good result. After investigation [105] and experiments of several methods such as zero symmetric and nonsymmetric random initialization, fixed initial weights and Nguyen-Widrow (NW) method, the results show that the worst performance is zero nonsymmetric random initialization. The reason may be that the input datasets are normalized to [-1, 1] which is symmetric. On the other hand, NW method always obtains best results and is more stable. NW method chooses initial weights in order to distribute the active region of each neuron in the layer approximately evenly across the layer's input space [106]. Thus NW method is applied in this algorithm for NN initialization.

PSO parameter initialization includes particle position and velocity initialization. Since particle position is coded in the NN weights, so position initialization has been completed in NN initialization. Particle velocity is randomly initialized with zero symmetric numbers.

3.3.3.4 Velocity and Position Update

Velocity and position update are two important operations in PSO algorithm. The searching information of personal and global best particle is conveyed by the velocity and position update. As mentioned in the relevant literatures [107, 108], the classic formulas for velocity and position update are shown respectively,

$$\begin{aligned} v_n(t+1) = W v_n(t) \quad &+ \quad C_1\, rand()(p_{best,n} - x_n(t)) \\ &+ \quad C_2\, rand()(g_{best,n} - x_n(t)) \end{aligned} \tag{3.6}$$

$$x_n(t+1) = x_n(t) + v_n(t+1) \tag{3.7}$$

where v_n is the particle velocity in the nth dimension, rand() is a random number between 0 and 1, W is a scaling factor, and C_1 and C_2 are scaling factors that determine the relative "*pull*" of p_{best} and g_{best} [109]. Also the limitations of the velocity (V_{max}) and position (X_{max}) are set in update. If the absolute value of velocity is greater than V_{max}, then the velocity is constrained in V_{max}. This is the same for position update.

3.3.3.5 Mutation Operator

Mutation operator has been successfully applied to genetic algorithm (GA). Cross-over and mutation are two main operators in GA. Some studies also introduce mutation operator into PSO and obtain better results [110]. In this chapter, mutation operator is integrated into PSO in order to increase the search capacity and avoid being trapped in local optima. When Gaussian mutation is added to each connection weight after the position update, the mean of Gaussian distribution equals to that parameter value and standard deviation is 10% of that parameter value. The mutation rate exponentially decreases as the optimization continues.

3.3.3.6 PI Construction and Evaluation for Training Set

After the update of the connection weights, new PIs are constructed using the LUBE method. The network architecture is similar to the one shown in Fig. 2.4. $\gamma(PICP)$ in the cost function for training is set to 1. After construction of PIs, PI evaluation indices PICP and PINRW are calculated and further the value of cost function CWC is calculated using Equation (3.4).

3.3.3.7 Update p_{best} and g_{best} particle

The criterion for p_{best} and g_{best} update is the value of cost function CWC. p_{best} denotes the personal best value of each particle, if the new CWC is smaller than CWC of p_{best}, then p_{best} for this particle is substituted by the new corresponding connection weights. Meanwhile, g_{best} is the best value of the whole swarm, if the new CWC of p_{best} is smaller than CWC of g_{best}, then the g_{best} is updated.

3.3.3.8 Training Termination

The training process terminates if the maximum number of iterations is reached or only small improvement is made in certain number of iterations, otherwise it returns to step *3.3.3.4*. The maximum number of iterations can be different in case studies, which is set according to the decreasing and converging process of CWC. Usually the PSO-based LUBE method is converged.

3.3.3.9 Test and Evaluation

After training the g_{best} value is used to construct the PIs for test samples. As mentioned in the second section, for test samples, PINAW is chosen instead of PINRW and $\gamma(PICP)$ is a step function defined in Equation (3.5). The evaluation indices PICP, PINAW and CWC are measured and recorded. The whole process is repeated for five times.

3.4 Case Studies

Six case studies are used to examine performance of the PSO-based LUBE method. The datasets and methodology used for six case studies are introduced.

3.4.1 Datasets

The first two cases are synthetic with heterogeneous and homogenous noise. The other four cases are from real world systems. Table 3.1 summarizes these case studies. These cases are as follows:

Table 3.1: Datasets for Case Studies

Case study	Target	Samples	Attributes
1	1-D function with heterogeneous noise (Ding10)	500	1
2	5-D function with homogenous noise (HAS)	300	5
3	Dry bulb temperature (DBT)	867	3
4	Plasma beta-carotene (PBC)	315	12
5	T70	272	2
6	T90	272	2

1) Ding10 is a one-dimensional synthetic mathematical function, $f(x) = x^2 + \sin(x) + 2 + \epsilon$, where x is randomly generated in $[-10, 10]$, and ϵ is the added noise with a heterogeneous distribution.

2) HAS is a five-dimensional synthetic mathematical function, $f(x_1, x_2, x_3, x_4, x_5) = 0.0647(12 + 3x_1 - 3.5x_2^2 + 7.2x_3^3)(1 + \cos(4\pi x_4))(1 + 0.8\sin(3\pi x_5))$. Unlike case study 1, the added noise is normally distributed with a constant variance.

3) Case study 3, dry bulb temperature (DBT) comes from an industrial dryer sampled every ten seconds. Three inputs are used for estimating the output of dry bulb temperature.

4) Data in case study 4 comes from a medical study, which contains 315 observations on 14

variables. This study tries to investigate the relationship between personal characteristics, dietary factors, and plasma beta-carotene.

5) T70 comes from a real baggage handling system, which is frequently affected by probabilistic events. The target is to forecast the travel time for 70% of each flight bags (T70).

6) T90 is similar to T70. It represents the travel time for 90% of each flight bags (T90). The level of uncertainty for T90 is higher than T70.

3.4.2 Methodology Used for Case Studies

Two synthetic and four real-world case studies can stand for different types of datasets and validate the PSO-based LUBE method. In each case study, a 5-fold cross validation method is implemented to determine the optimal structure of NNs. The implementation is very similar to Figure 2. However, there are two small differences. The first difference is the data splitting. It is the training set that is split into five folds using the cross validation method. The second difference is that the NN structure is fixed for each run, and N_h is changed from 1 to 20 ($1 \leq N_h \leq 20$). Each candidate NN structure is trained and validated 5 times using 5 different sub-training and validation datasets, the median value of CWC is chosen to determine the optimal NN structure.

Case study 3 of DBT is taken as an example to illustrate the 5-fold cross validation method. Fig. 3.2 shows the median value of CWCs in percentage terms corresponding to the number of neurons in the hidden layer (N_h). As shown in Fig. 3.2 for case study 3, $N_h = 4, 5, 16$ are the best three options. NNs with $N_h = 4$ and $N_h = 5$ have the top two smallest median value of CWCs. NN with $N_h = 16$ is obviously more complex and has more parameters, which is difficult for training. Median CWC of $N_h = 4$ is also smaller than median CWC of $N_h = 5$. Balancing between the network complexity, generalization and regression capacity of NNs, $N_h = 4$ is chosen as the optimal NN structure.

After the determination of the optimal NN structure, the LUBE method is used to construct PIs and PSO is applied to minimize the cost function through optimizing the connection weights

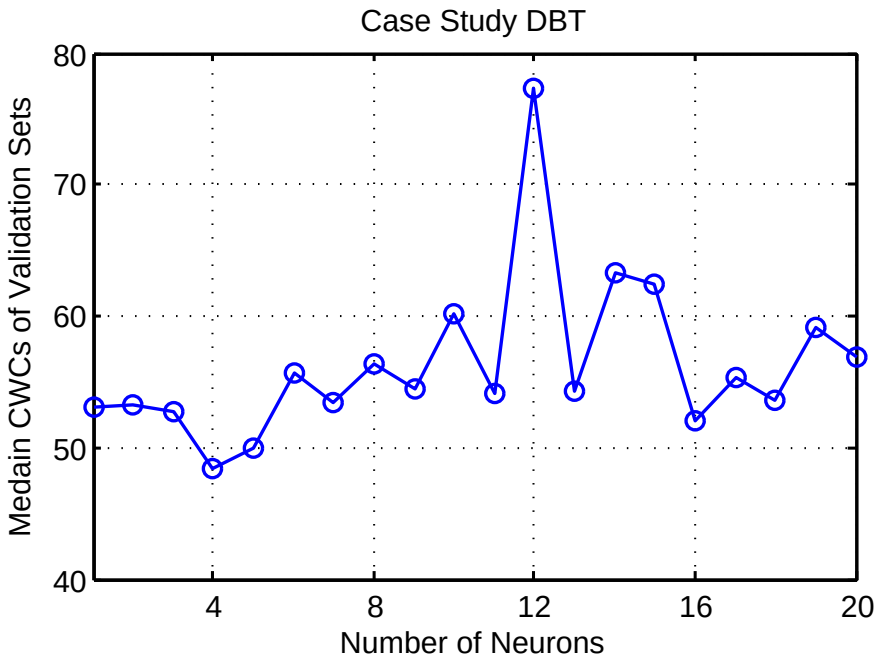

Figure 3.2: Median CWC vs. NN structure for DBT

of NNs. When the maximum number of iterations is reached, the training process terminates. The obtained g_{best} is then used to construct the PIs and PICP, PINAW, CWC are used to evaluate the quality of PIs for test samples.

Table 3.2: Parameters for PSO and CWC

	Parameter	**Numerical value**
PSO	C_1	1.22
	C_2	1.49
	W_{max}	0.9
	W_{min}	0.7
CWC	α	0.1
	μ	0.90
	η	80

Parameters of PSO can change for different case studies in order to control the convergence speed and the searching process. The typical parameter sets of PSO and the CWC cost function are shown in Table 3.2. W_{max} and W_{min} are the maximum and minimum value of the

inertia weight W for previous velocity, which linearly decreases and plays an important role in controlling the PSO convergence.

3.5 Experimental Results and Discussions

For construction of NN-based PIs, representation of results consists of mainly three parts, the training process, test results and discussions on computation time. As shown in Fig. 3.3, the training process shows the convergence behavior of the algorithm. Through training, the value of cost function becomes smaller and smaller. The test results are based on the constructed PIs for test samples. Through testing, the performance of the trained NN is validated. In this section, the training process, the test results and discussions on computation time are presented.

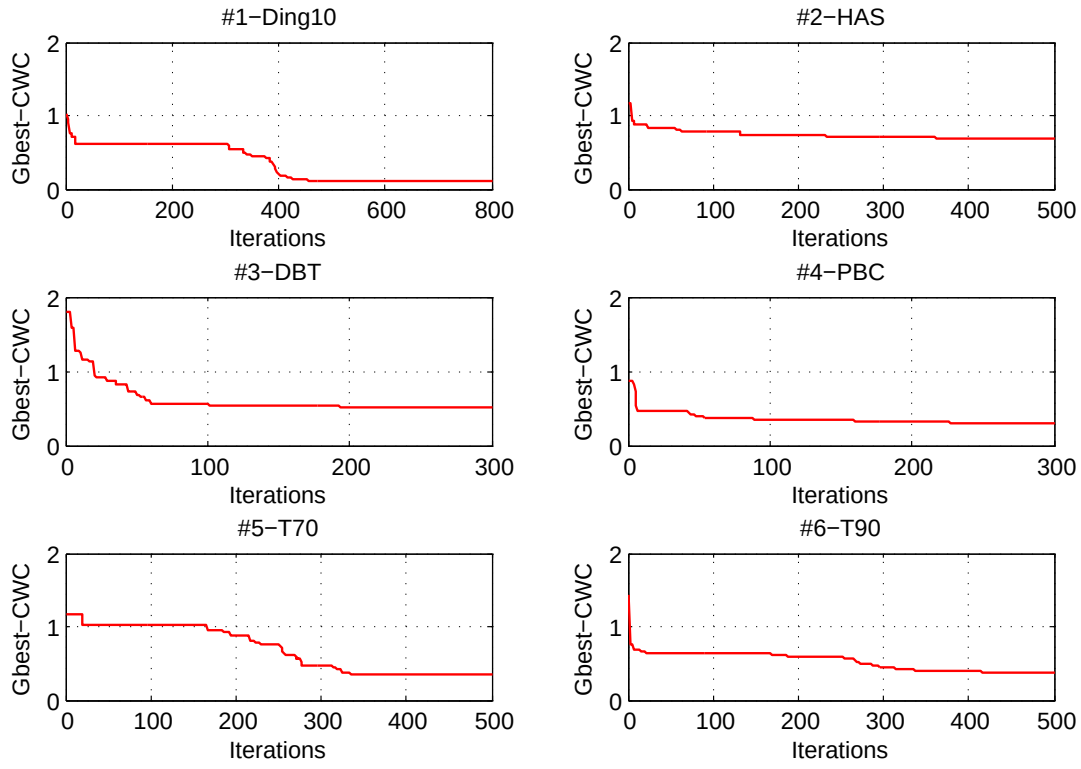

Figure 3.3: CWC of the g_{best} particle in each generation of the PSO algorithm.

3.5.1 Training Process

In order to clearly present the result of the six case studies, the training process for minimizing CWC cost function is illustrated in Fig. 3.3. The figure shows the CWC of the g_{best} particle in each generation of the PSO algorithm. As can be seen from the figure, the CWC value of the g_{best} particle for each case study converges to a good result. CWC decreases rapidly at the beginning of the minimization process. As the optimization proceeds, CWC gradually plateaus and takes its optimal value. For some case studies, such as case study 2, 3 and 4, CWC drops sharply at the beginning and after about 100 iterations it decreases slowly until the end. But for case study 1, 5 and 6, not only at the beginning a sharp drop occurs, but also a significant drop happens at the middle. For each case study, the algorithm converges and obtains a good result, which indicates the strong searching capacity of PSO combined with the mutation operator.

3.5.2 Test Results

Fig. 3.4 shows the constructed PIs for test samples of six case studies. For better visualization, only PIs for the first sixty samples are shown. For each case study the constructed PIs cover the targets in a great percentage, which implies that the PICP indices for the test samples are very satisfactory. For all case studies we have $PICP >= 90\%$. This clearly indicates that the cost function CWC and PSO training algorithm can construct high quality PIs. On the other hand, PINAWs vary from one case study to another. They can be very narrow, such as the PINAWs in case study 1, 4 and 5; but PINAWs in case study 2 and 3 are much wider. This indicates that under certain PICP, the widths for PINAWs are determined by the level of uncertainty of the datasets. Higher level of uncertainty results in wider PINAW. In contrast, a lower level of uncertainty leads to a smaller PINAW. Thus the PSO-based LUBE method can generate high quality PIs for different types of datasets.

In order to validate the stability of the algorithm, and present quantitative and convincing results, each case study is repeated for 5 times. The median values of PICP, PINAW, CWC

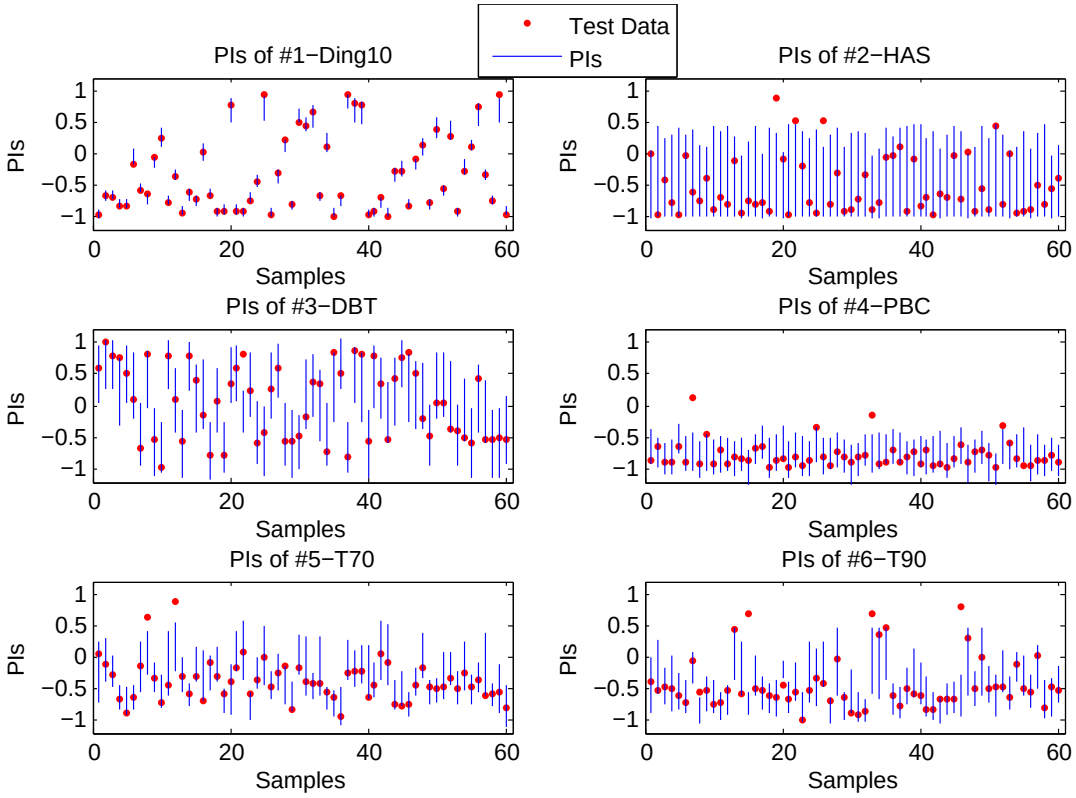

Figure 3.4: The constructed PIs of the 6 case studies.

instead of the best ones are displayed in Table 3.3 for six case studies. Although $\gamma(PICP) = 0$ for test samples, for some case studies such as case study 1, 2 and 4, PINAW is not equal to CWC. This is due to the fact that shown values are the median of CWC.

It can be seen from the table for all case studies, the assigned confidence level (90%) can be satisfied. The median values of PICP are all greater than 90%, which means the PSO-based LUBE method can generate PIs with a high PICP. Meanwhile, the median PINAWs are also compared with the delta [41] and LUBE [43] method respectively. For the parameters of cost function CWC, $\gamma(PICP)$ (a step function) and μ are the same. The penalty term η in this chapter is 80 which is larger than 50 in [41, 43]. Although heavier penalties are applied, the obtained CWCs for all case studies are significantly smaller than the results of the same case studies reported in [41, 43]. In spite of the median PINAW in case study 1 of delta method [41] is smaller than the median PINAW of our proposed method, but its PICP is unsatisfied which

Table 3.3: PI Evaluation Indices and Construction Time for Test Samples

Comparisons	Case studies	PICP(%)	PINAW(%)	CWC(%)	Time(ms)
The Proposed Method	#1-Ding10	92.67	11.49	15.05	0.0693
	#2-HAS	91.11	61.81	63.46	0.0276
	#3-DBT	91.15	49.24	49.24	0.0895
	#4-PBC	93.62	27.30	27.49	0.0503
	#5-T70	91.46	37.69	37.69	0.0778
	#6-T90	90.24	38.77	38.77	0.0508
LUBE [43] and Delta [41] Method	#1-Ding10 [41]	89.50	6.32	20.49	1460
	#2-HAS [43]	90.00	72.22	89.27	3
	#3-DBT [43]	94.62	69.39	69.39	4
	#4-PBC [43]	91.15	33.91	33.91	3
	#5-T70 [41]	91.67	46.44	64.01	980
	#6-T90 [43]	91.46	41.51	41.51	3

is less than 90%. The proposed method provides an obvious better CWC than delta method. The percentage improvements $(100\% * (compared\ result - new\ result)\ /\ compared\ result)$ of CWCs are 26.55%, 28.91%, 29.04%, 18.93%, 41.12% and 6.60% respectively for six case studies.

Therefore, through minimization of cost function for each case study, the PSO-based LUBE method can construct high quality PIs with a narrower width and a satisfactory coverage probability.

3.5.3 Discussion on Computation Time

Computation time is one of the key characteristics for algorithm design. For construction of NN-based PIs, computation time consists of two parts: the off-line training and on-line computation. The time for off-line is regarded as less important because we will have enough time to train a NN properly. The NN can be trained and parameters can be adjusted for enough times, then the one with the best results will be used in real practice. But the on-line computation time becomes

crucial if it is used for on-line applications. For example, case study 6, T90, comes from a real baggage handling system, thousands of flight bags at peak hours need to be scheduled in a short time, algorithms with long computation time will be discarded.

In order to evaluate the computation time of the proposed PSO-based LUBE method, Table 3.3 also summarizes average PI construction time for test samples. The hardware configuration of the simulation computer is as follows: Intel(R) Core(TM)2 Duo CPU E8500 @ 3.16GHz, and 4 GB of RAM. Under this hardware configuration, for all case studies, PI construction time of test samples is less than 0.1ms. This is much faster than both the results reported in [41, 43]. As can be seen from Table 3.3, traditional method such as delta method is very time consuming. This is because Jacobin matrix needs to be calculated and updated in each iteration. According to the same case studies reported in [43], PI construction requirement of traditional methods is at least 10 times more than LUBE methods. Compared with traditional methods, LUBE method is straightforward and fast to construct PIs. The reason for faster construction of PIs than [43] is that matrix operation is applied to our proposed method, all the PIs for test samples can be obtained at one time.

The above result comparisons strongly indicate that the PSO-based LUBE method can construct higher quality PIs in a very short time and can be more widely used in real-world applications.

3.6 Conclusions

A recently introduced PI construction method named LUBE method is applied and extended in this chapter. To quantitatively evaluate the quality of PIs, PI evaluation indices for both coverage probability and width are introduced. Specially, a new width evaluation index PINRW, which is suitable for NN training, is proposed. Compared with PINAW which gives equal weights to all intervals, PINRW magnifies the wider PIs. Case study results show that the new index can obtain averagely narrower PIs with a satisfactory coverage probability. The cost function

CWC is developed to comprehensively balance between the PICP and PINRW. PSO associated with mutation operator is integrated into the LUBE method called PSO-based LUBE method. The proposed method is then applied to construct PIs and minimize the cost function through optimizing the connection weights of NNs.

Six case studies, two synthetic and four real-world ones, are implemented to validate the PSO-based LUBE optimization method. The 5-fold cross validation method is used to determine the optimal structure of NNs. Demonstrated results show that not only the quality of PIs has considerable improvements, but also the computation time is much shorter than traditional methods. What's more, the proposed method can construct PIs in an easier and faster manner without any assumption about the data distribution, it can be more generally and widely applicable in practice. The minimization process of CWC of the g_{best} particle in each generation clearly indicates that the cost function can always converge to a sufficient small CWC. With the minimization of CWC, both the high PICP and narrow PINAW can be reached, which has constructed high quality PIs. In conclusion, the PSO-based LUBE method is very efficient in constructing high quality PIs in a short time.

Chapter 4

Uncertainty Handling Using Neural Network-Based Prediction Intervals for Electrical Load and Wind Power Forecasting

4.1 Introduction

With power systems growth and the penetration of renewable energy sources, the system complexity and uncertainty levels have significantly increased. The load and the renewable energy forecasting processes have become even more complex, and more accurate forecasts are required for the management of power systems. Long-term forecasts of the peak electricity demand are needed for capacity planning and maintenance scheduling [111]. Medium-term forecasts are used for maintenance planning, fuel scheduling, and hydro reservoir management [112]. Short-term load forecasting (STLF) is a fundamental and vital factor in day-to-day operations, unit commitment (UC) and scheduling functions, evaluation of net interchange, and system security

analysis [112]. The STLF and the renewable energy forecasting are required for the control and scheduling of power systems and affect the system reliability and fuel consumptions.

Forecasting models that are popularly applied to electrical load and the renewable energy forecasting can be classified into three categories. 1) statistical models, such as aurore-gressive (AR), AR integrated moving average (ARIMA) and exponential smoothing (ES) models [111,113,114]; 2) artificial intelligence models, such as neural networks (NNs) [90,115–118], fuzzy logic systems (FLSs) [119–121], expert systems etc.; and 3) hybrid models as neuro-fuzzy systems [122, 123], to name a few. In [111], different forecasting methods including ARIMA modeling, periodic AR modeling, double seasonality of Holt-Winters ES, and principal component analysis are considered for STLF. Electricity demands from 10 European countries are used as case studies to compare these methods. In [114] two time series models are proposed, namely, the multiplicative decomposition model and the seasonal ARIMA model. These two models are implemented and compared for STLF using Singapore datasets. Hippert et al. in [90] provide a comprehensive review of STLF using NN models. Different forecasting strategies, the iterative, multi-model, and single-model multivariate forecasting are investigated. Issues, such as NN designing, implementation and validation are also covered. In [116], three techniques called error output, resampling and multilinear regression are applied to STLF for constructing confidence intervals using NN models. In [118], the second-generation wavelets are combined with recurrent NNs to improve the accuracy of solar radiation prediction. In [4], a hybrid NN and enhanced PSO are used as the forecast engine for wind power forecast. Focusing on feature selection, an irrelevancy filter and a redundancy filter are applied to select the set of candidate inputs of NN.

Like NNs, FLSs are also universal approximators. Recently, FLSs have a quick development and are popularly applied to forecasting applications [123]. In [122], the adaptive neuro-fuzzy inference system (ANFIS) model is applied for very short-term wind forecasting using Tasmania datasets. In [119], an interval type-2 FLS (IT2 FLS) is applied to STLF to handle uncertainties. The output of an IT2 FLS is an interval (the type reduced set composed of the

left and right end points), but is not a prediction interval (PI). A PI has a prescribed probability associated with it called the confidence level. The output obtained from an IT2 FLS does not have this feature and is a simple interval. An IT2 FLS model has built-in features (e.g., membership functions with uncertain mean and variance) for handling uncertainties and minimizing their effects on the quality of output. However, it still lacks to assign a confidence level to its type reduced set (interval).

Electrical power systems are evolving from today's centralized bulk systems to more decentralized systems [1]. Renewable energy sources, such as the wind and solar power generations, with their advantages of being cheaper, more flexible, and environmentally friendly, become the key to a sustainable energy supply infrastructure. But penetrations of these increase the level of uncertainty in power systems. More advanced methods for accurate load and wind power forecasts under various uncertainties become urgent for smart grid applications. In [2], wind generation is considered as a negative load. This further increases the complexity of load forecasting. Most of the applications on STLF and wind power forecasting are point forecasts. However, point forecasts cannot properly handle the uncertainties associated with datasets [42, 56].

In this chapter, STLF and wind power forecasting are implemented using NN-based PIs. PIs are excellent tools for quantification of uncertainties associated with point forecasts and predictions [5, 55]. By definition, a PI is an estimate of an interval in which a future observation will fall, with a certain probability $((1-\alpha)\%)$, given what has already been observed [19]. Typically, a PI consists of a lower and an upper bound, and the confidence level that the targets will lie within the two bounds. For point forecasts, only one predicted point is provided for one target value. Point forecasts provide only the prediction error but tell nothing about the probability for correct predictions. PIs not only provide a range in which targets are highly likely to be covered, but also have an indication of their accuracy called the coverage probability.

Delta, Bayesian and bootstrap are three traditional methods commonly used for construction of NN-based PIs [42, 124]. In spite of advantages of PIs, applications of these methods are

still less popular than NN point forecasts. Implementation difficulties, special assumptions about the data distribution, and massive computational requirements [50, 54, 57] hinder widespread applications of these methods for decision-making. To overcome these problems, a new method called lower upper bound estimation (LUBE) method for PI construction was proposed by Khosravi et al in [43]. The LUBE method makes no assumption about data distribution, and avoids calculation of matrices such as the Jacobian and Hessian matrix. As we know, wind power is intermittent and very volatile in nature. Thus assumptions about the data distributions seem problematic and in doubt. That is where we can use the non-parametric LUBE method. Comparative results reported in [43] reveal that the LUBE method is simpler, faster, and more reliable than traditional methods.

From the decision-making point of view, larger coverage probability and narrower width are always expected. But these two aspects of PIs are conflicting with each other. This optimization problem can be formulated and solved in different ways. In this chapter, after introducing the evaluation indices of PIs, three problem formulation methods are summarized and developed. Obviously, the primary problem is a multi-objective optimization problem for larger coverage probability and narrower width. Previously, the primary problem has been translated into a single objective one using cost functions [5, 42, 55, 56, 98]. In this chapter, a new problem formulation method is proposed. The nominal coverage probability is considered as a hard constraint. Our only objective is to minimize the width of PIs. The new constrained single-objective problem formulation is closer to the primary problem and has fewer parameters than cost functions.

To solve the new constrained single-objective problem, traditional derivative-based algorithms, such as gradient descent methods, cannot be applied. These methods also run the risk of being trapped into local optima. Therefore, more intelligent and powerful optimization methods are needed. To optimize NN parameters, PSO [4, 125], genetic algorithm (GA) [126], and simplified swarm optimization [127] have been introduced in literatures. In this chapter, PSO which is powerful for parameter optimization is employed to solve the problem. The mutation operator,

which helps achieve diversity in GA, is also integrated into PSO to improve the exploratory capabilities and help jump out of local optima. The objective of using PSO has mainly two aspects. For one thing, PSO is used to solve the newly formulated constrained single-objective problem. That is, to handle the constraints and optimize the objective. For another thing, PSO with mutation operator is used as the training algorithm through optimizing the connection weights of NN models. Datasets from electrical load demands and wind power generations are used to validate this method. For the purpose of comparisons, ARIMA, ES and naive models are also built using the same datasets. Comparative results show that the PSO-based LUBE method can construct higher quality PIs for load and wind power forecasting applications.

The main contributions of this chapter are as follows: 1) A new problem formulation method for PI construction is proposed. The primary multi-objective problem is formulated and solved as a constrained single-objective problem. 2) A new PI width evaluation index, which is suitable for training NN models, is proposed. 3) PSO associated with mutation operator is firstly integrated into the LUBE method called the PSO-based LUBE method. This PSO associated with mutation operator has a very strong searching capability. 4) Different types of prediction tasks including electrical load and wind power generation forecasts are implemented and compared together. 5) The obtained results from three case studies indicate that the quality of PIs has been significantly improved compared with ARIMA, ES and naive models. 6) Implementation of the proposed method is straightforward and much easier; the PI construction time is also much shorter than traditional methods.

The rest of this chapter is organized as follows. Section 4.2 introduces evaluation indices of PIs. Three problem formulation methods are provided in Section 4.3. The proposed PSO-based LUBE method is described in Section 4.4. Case studies, results and discussions are implemented in Section 4.5 and Section 4.6 separately. Finally, Section 4.7 concludes this chapter and provides guidelines for future work.

4.2 Evaluation Indices of PIs

In literature, various methods have been applied to evaluate the performance of point forecasts. Out of these methods, the most popular ones are error-based measures, such as mean square errors (MSE) and mean absolute percentage errors (MAPE). Likewise, quality of PIs needs to be quantitatively evaluated. In this section, evaluation indices for both the coverage probability and width of PIs are firstly introduced. Specially a new index for width evaluation is proposed. The new index is suitable for training NN models. Finally, a cost function for the comprehensive evaluation of PIs is developed.

4.2.1 PI Coverage Probability

Usually the coverage probability (or confidence level) is considered as the key feature of PIs. PI coverage probability (PICP) indicates in which probability target values will be covered by the upper and lower bounds. A larger PICP means more targets lie within the constructed PIs and vice versa. PICP is defined as [56]:

$$PICP = \frac{1}{N}\sum_{i=1}^{N}\epsilon_i \tag{4.1}$$

where, N is the number of samples and ϵ_i is a boolean variable which indicates the coverage behavior of PIs. If the target value y_i is covered between the lower bound L_i and upper bound U_i, $\epsilon_i = 1$; otherwise $\epsilon_i = 0$. Mathematically, ϵ_i is defined as follows:

$$\epsilon_i = \begin{cases} 1, & if\ y_i \in [L_i, U_i]; \\ 0, & if\ y_i \notin [L_i, U_i]. \end{cases} \tag{4.2}$$

To have valid PIs, PICP should be no less than the nominal confidence level of PIs. Otherwise, PIs are invalid and should be discarded. The ideal case for PICP is $PICP = 100\%$, which means all targets are covered by PIs.

4.2.2 PI Normalized Average Width and Normalized Root-Mean-Square Width

Quality of PIs is often evaluated by PICP and discussion about the width of PIs is either ignored or vaguely presented [41]. If the upper and lower bounds of PIs are chosen as extreme values of the targets (maximum and minimum values), a high PICP (even 100% PICP) can be easily achieved. The argument here is that too wide PIs convey little information and are of no use for decision-making. Width of PIs determines their informativeness. In literature, a quantitative measure of the width is defined as PI normalized average width (PINAW) [98]:

$$PINAW = \frac{1}{NR}\sum_{i=1}^{N}(U_i - L_i) \tag{4.3}$$

where R is the range of the underlying targets (maximum minus minimum). The purpose of using R is to normalize the PI average width in percentage. In this way, PINAW can be used for objective comparisons, regardless of techniques used for their construction or the magnitudes of the underlying targets.

The format of PINAW is similar to the MAPE used for point forecasts. It gives equal weights to each width of PIs. It is MSE that is frequently used to train NN models instead of MAPE. This is due to the fact that MSE magnifies bigger forecasting errors and results in better training performances. Inspired by this, a new width evaluation index for PIs, called PI normalized root-mean-square width (PINRW), is developed for training in this chapter:

$$PINRW = \frac{1}{R}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(U_i - L_i)^2} \tag{4.4}$$

The new index PINRW is functionally similar to MSE, and magnifies wider intervals. In practice, experiment results show that PINRW trends to obtain higher quality PIs than PINAW, just as MSE is a much better cost function for training than MAPE. Next in this chapter, PINRW is used for training NN models and PINAW for testing.

4.2.3 Coverage Width-Based Criterion

PICP and PINAW (or PINRW) assess only one aspect of PIs individually. Focusing on only one side of PIs may lead to misleading results. In practice, a measure is required to simultaneously address both aspects and comprehensively evaluate the overall quality of PIs. An interesting index, called coverage width-based criterion (CWC), is proposed in [98]:

$$CWC = PINAW(1 + \gamma(PICP)e^{-\eta(PICP-\mu)}) \tag{4.5}$$

where $\gamma(PICP) = 1$ for training. μ and η are two controlling parameters. The nominal confidence level ($(1-\alpha)\%$) can be used as a guidance for choosing μ. It stands for the preassigned PICP that must be satisfied. η is a hyper-parameter that magnifies the difference between the PICP and μ. If the preassigned PICP is not satisfied, CWC exponentially penalizes on this term. When the PICP reaches around μ, there is a balance between the PINRW (for training) and PICP. CWC provides a comprehensive assessment for both evaluation indices. It tries to find a tradeoff between informativeness (PINAW & PINRW) and validity (PICP) of PIs.

If the preassigned PICP is satisfied, the comparison between the two CWCs reasonably pays more attention to the narrower PINAW. Thus for test samples, $\gamma(PICP)$ is a step function whose value is determined upon the satisfaction of PICP:

$$\gamma(PICP) = \begin{cases} 0, & PICP \geq \mu; \\ 1, & PICP < \mu. \end{cases} \tag{4.6}$$

When evaluating test PIs, if PICP is no less than the assigned μ, $\gamma(PICP) = 0$ gives equal measurements of PICP. Otherwise, $\gamma(PICP) = 1$ and the corresponding penalty will be accounted by CWC.

4.2.4 Winkler SCORE

Winkler in [128] proposed a scoring rule for interval estimation. This score $S_i^{(\alpha)}$ rewards narrow PI as well as provides the penalty term if the target value y_i is not covered by the constructed PI.

$$S_i^{(\alpha)} = \begin{cases} -2\alpha\delta_i^{(\alpha)} - 4(L_i^{(\alpha)} - y_i), & y_i < L_i^{(\alpha)}; \\ -2\alpha\delta_i^{(\alpha)}, & y_i \in I_i^{(\alpha)}; \\ -2\alpha\delta_i^{(\alpha)} - 4(y_i - U_i^{(\alpha)}), & y_i > U_i^{(\alpha)}. \end{cases} \tag{4.7}$$

where $\delta_i^{(\alpha)}$ is the size of the interval forecast $I_i^{(\alpha)}$:

$$\delta_i^{(\alpha)} = U_i^{(\alpha)} - L_i^{(\alpha)}. \tag{4.8}$$

Then, the average score is derived to address the overall performance:

$$\overline{S_i^{(\alpha)}} = \frac{1}{m}\sum_{i=1}^{m} S_i^{(\alpha)}. \tag{4.9}$$

A detailed comparison between Winkler score and CWC for the comprehensive evaluation of PIs can be found from one of our publications in [129]. Both the advantages and disadvantages of CWC and SCORE are discussed and compared together for uncertainty handling of electrical load forecasting applications [129]. In the following of this chapter, CWC is chosen as the comprehensive evaluation index of PIs.

4.3 Problem Formulation for Construction of NN-Based PIs

Three different problem formulations are summarized and compared together. They are the primary multi-objective problem formulation, the single-objective formulation and constrained single-objective formulation.

4.3.1 Primary Problem Formulation

From the optimization point of view, higher PICP and narrower PINAW are two objectives for high quality PIs. Thus the primary problem can be modeled as a multi-objective problem. However, these objectives are conflicting with each other as improving one objective will decrease

another. If it is solved by some multi-objective methods, the nondominated solutions will appear on the Pareto front.

$$\begin{aligned} \textbf{Objectives}: \quad & Finding\ optimal\ weights\ \omega^*\ to: \\ & \textbf{Maximize}: PICP(\omega); \\ & \textbf{Minimize}: PINAW(\omega). \end{aligned} \tag{4.10}$$

$$\begin{aligned} \textbf{Constraints}: \quad & PINAW(\omega) > 0; \\ & 0 \leq PICP(\omega) \leq 100\%. \end{aligned} \tag{4.11}$$

4.3.2 Single-Objective Problem Formulation

In our previous work [42, 55, 56, 98], if the nominal PICP (μ) was preassigned, the primary problem was transformed into a single-objective problem. This is realized through using CWC as a cost function defined in Equation (3.4).

$$\begin{aligned} \textbf{Objective}: \quad & Finding\ optimal\ weights\ \omega^*\ to: \\ & \textbf{Minimize}: CWC(\omega). \end{aligned} \tag{4.12}$$

CWC cost function provides a comprehensive evaluation on both the PICP and PINAW (or PINRW for training). At the beginning of the training process, PICP is usually very low, and then CWC gives a heavily exponential penalty on this term. As the training continues, PICP becomes higher and higher, the penalty for the unsatisfying PICP exponentially decreases. Once PICP is near the preassigned coverage probability (μ), there is a balance between the validity (PICP) and the informativeness (PINAW & PINRW) of PIs.

4.3.3 Constrained Single-Objective Problem Formulation

In this chapter, a new problem formulation method is proposed. Since PICP is usually considered as the fundamental feature and determines the validity of PIs, it is reasonably to be regarded as the hard constraint. This means, the preassigned PICP must be satisfied for valid PIs. Under this

hard constraint, the remaining objective is to minimize PINAW. In this way, the primary problem is successfully represented as a constrained single-objective problem:

$$\textbf{Objective}: \quad Finding\ optimal\ weights\ \omega^*\ to: \tag{4.13}$$

$$\textbf{Minimize}: PINAW(\omega).$$

$$\textbf{Constraints}: \quad PINAW(\omega) > 0; \tag{4.14}$$

$$\mu \leq PICP(\omega) \leq 100\%.$$

where μ is the nominal confidence level which can be set to $(1-\alpha)\%$. As shown in the model, this new formulation is a single-objective problem with three constraints. Our only objective is to minimize the average width of PIs. For the three constraints, out of which, $\mu \leq PICP(\omega)$ is the hard constraint that must be met. If proper controls are given in the process of calculation, other two constraints can be satisfied automatically.

Compared to the cost function method, there are two obvious advantages of this problem formulation. It has fewer parameters and is closer to the primary problem. Taking the above CWC cost function as an example, parameters such as μ, η and $\gamma(PICP)$ need to be assigned in advance. The performance of the cost function is also sensitive to these parameters. They need to be fine tuned carefully.

To solve the constrained single-objective problem using PSO, the criteria for replacing one particle $\vec{m}$ with another particle $\vec{n}$ are as follows [108, 130]:

- Particle $\vec{n}$ is feasible and particle $\vec{m}$ is not;
- Both particles are feasible or they have the same satisfaction of constraints, but $\vec{n}$ yields a better objective function value.
- Both particles are infeasible, but $\vec{n}$ results in the lower sum of constraint violations.

4.4 PSO-based LUBE Method for STLF

With the new problem formulation, the improved PSO-based LUBE method is applied to STLF and wind power forecasting. The detailed implementation is introduced below.

4.4.1 LUBE Method

Traditional methods for construction of PIs suffer from various problems. For examples, the delta method makes assumption on data and residual distributions [41]. In the process of PI construction, the derivatives also need to be calculated. Jacobian matrix and Hessian matrix are required by the delta and Bayesian method respectively [41]. This may result in singularity problems then decrease the reliability of PIs. Calculations for Jacobian matrix and Hessian matrix also significantly increase the computation time. Thus the complexity of traditional methods hinders widespread applications of PIs.

A new method called the lower upper bound estimation (LUBE) method was proposed in [43] to construct NN-based PIs. LUBE method adopts a NN with two outputs to directly construct PIs in one step without any assumption about the data distribution. The two outputs of NN correspond to the lower and upper bounds of PIs. This design format is similar to point forecasts, the process of PI construction is easy and straightforward. But the functions are totally different. LUBE method constructs PIs in just one step, thus is simpler and faster to be implemented. Interesting readers can find more about LUBE method in Subsection 2.2.2

4.4.2 Implementation of the PSO-based LUBE Method for STLF

The flow chart of the proposed PSO-based LUBE method for STLF is shown in Fig. 4.1.

4.4.2.1 Data Splitting

The whole dataset is split into three sets: training set, validation set, and test set. The training set is used to adjust the connection weights of NNs. The validation set is applied to determine

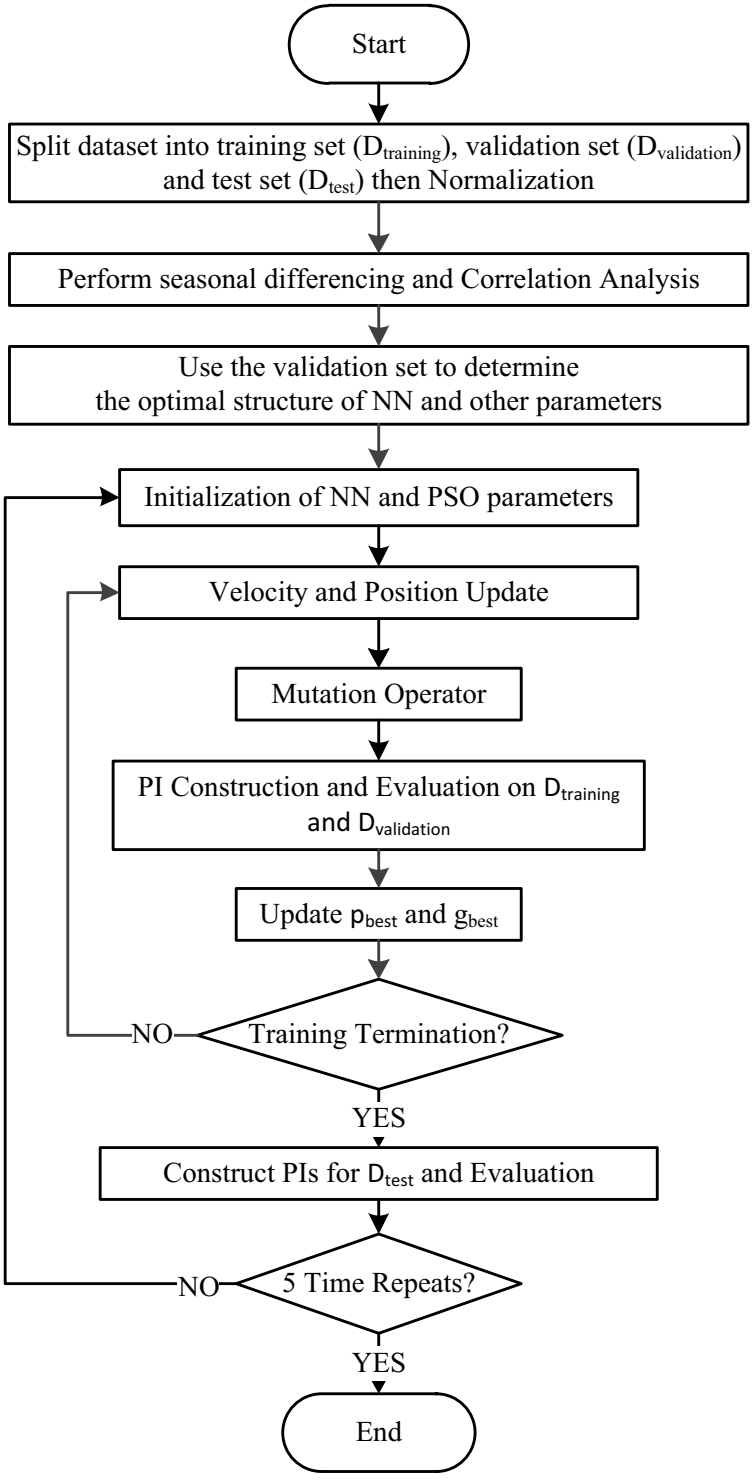

Figure 4.1: PSO-based LUBE method for construction and evaluation of PIs for STLF

the optimal NN structure and other undetermined parameters. The test set will evaluate the final performance of the algorithm. After splitting, the training and validation datasets are normalized to [-1, 1], and the same settings are applied to test set for normalization.

4.4.2.2 First Seasonal Difference and Correlation Analysis

The purpose of differencing is to make a time series stationary. It is particularly effective for a seasonal time series. Correlation analysis is applied to help choose the inputs of NNs. The detailed implementation of this part can be found in Section V.

4.4.2.3 Determination of the Optimal NN Structure

For each candidate NN structure, the NN is trained and validated using the training and validation sets for five times. Median values of PINAWs (with satisfied PICPs $\geq \mu$) are used to determine the optimal structure of NNs. Similarly, this method is also applicable to other undetermined parameters.

4.4.2.4 Initialization

NN weight and PSO parameter initializations become crucial for the proposed algorithm. The initialization process directly influences the quality of PIs and repeatability of the algorithm. Several NN weight initialization methods have been investigated [105] and compared in advance. The performances of fixed initial weights, zero symmetric and nonsymmetric random initialization and Nguyen-Widrow (NW) method are examined by a list of experiments. Comparative results show that zero nonsymmetric random initialization has the worst results. One possible explanation is that the input datasets are normalized to [-1, 1] which is symmetric. On the other hand, NW method repeatedly obtains best and stablest results. NW method chooses initial weights in order to distribute the active region of each neuron in the layer approximately evenly across the layer's input space [106]. Thus NW method is chosen in this algorithm for NN weight initialization.

PSO parameter initialization consists of particle position and velocity initialization. Since NN connection weights are represented as the position of particles, so position initialization has been completed in weight initialization. Particle velocity is randomly initialized with zero

symmetric numbers.

4.4.2.5 Velocity and Position Update

Velocity and position update are the core of the PSO algorithm. The particles will exchange their "*findings*" with each other in the update process. In this way, the information will be exchanged efficiently throughout the whole swarm. The classic formulas for velocity and position update [108] are shown below:

$$\begin{aligned} v_n(t+1) = Wv_n(t) &+ C_1\, rand()(p_{best,n} - x_n(t)) \\ &+ C_2\, rand()(g_{best,n} - x_n(t)) \end{aligned} \tag{4.15}$$

$$x_n(t+1) = x_n(t) + v_n(t+1) \tag{4.16}$$

where v_n is the particle velocity in the nth dimension, rand() is a random number between 0 and 1, W is a scaling factor, and C_1 and C_2 are scaling factors that determine the relative "pull" of p_{best} and g_{best} [109]. Besides the two updates, the ranges for velocity and position are limited to V_{max} and X_{max} separately.

4.4.2.6 Mutation Operator

Selection, cross-over and mutation are three main operators in GA. Mutation operator, which helps achieve diversity in GA, is integrated into PSO. This integration strongly enhances the searching capacity and avoids being trapped into local optima. In flow chart shown in Fig. 4.1, Gaussian mutation is added to each connection weight after the position update. The mean and standard deviation of Gaussian distribution are the weight value and 10% of that weight value respectively. The mutation rate exponentially decreases as the optimization continues.

4.4.2.7 PI Construction and Evaluation for Training

The validation set has been applied to determine the optimal NN structure. In this step, the training and validation sets are combined together to train the NN. After update of the NN connection

weights, LUBE method is then applied to construct new PIs. PI assessment indices (PICP and PINRW) are calculated.

4.4.2.8 Update p_{best} and g_{best} particle

p_{best} is the personal best value of each particle and g_{best} denotes the best value of the whole swarm. Compared with the cost function method, the constrained single-objective optimization is different. When updating the p_{best} and g_{best}, the criterion introduced in Section III is applied. The feasibility and the objective function will be considered together. While only CWC is considered for cost function method.

4.4.2.9 Training Termination

The training termination criterion can be set as the reach of the maximum number of iterations or few improvements made in certain number of iterations. Otherwise, the training process continues and returns to Step 4.4.2.5.

4.4.2.10 Test and Evaluation

Once the training process terminates, the g_{best} value is chosen to generate PIs for the test set. PICP and PINAW instead of PINRW are calculated and recorded. For the comprehensive evaluation purpose, CWC is also calculated. $\gamma(PICP)$ is a step function defined in Equation (4.6) for testing. The whole process is repeated five times. Results in each run and the median values are reported.

4.5 Electrical Load and Wind Power Forecasting in DPSs

Three case studies are investigated to validate the improved PSO-based LUBE method. Two cases are real datasets for electrical load forecasting and one case is for wind power forecasting.

4.5.1 Datasets

The demand datasets are real electrical load data from Singapore (SG) and New South Wales (NSW) (Australia). These two different areas stand for two different types of load profiles. Load pattern of Singapore is obvious more fixed than NSW's due to the climate and regional reasons. Singapore has a tropical rainforest climate with no distinctive seasons, uniform temperature and pressure. Temperatures of this city country usually range from 23 to 32 °C (73 to 90 °F). On the other hand, NSW is more influenced by the seasonal factors thus with more complex load patterns.

The chosen time periods are from Jan, 2007 to Dec, 2011, with 48 load points in each day. These five-year load datasets are further divided into three sets: training set, validation set and test set. The periods used for training, validation and testing are the first three years, the fourth year and fifth year respectively. The typical weekly curves of SG and NSW load, from Monday to Sunday, are shown in Fig. 4.2. One-week ahead load forecasting using NN-based PIs will be implemented and compared for both case studies.

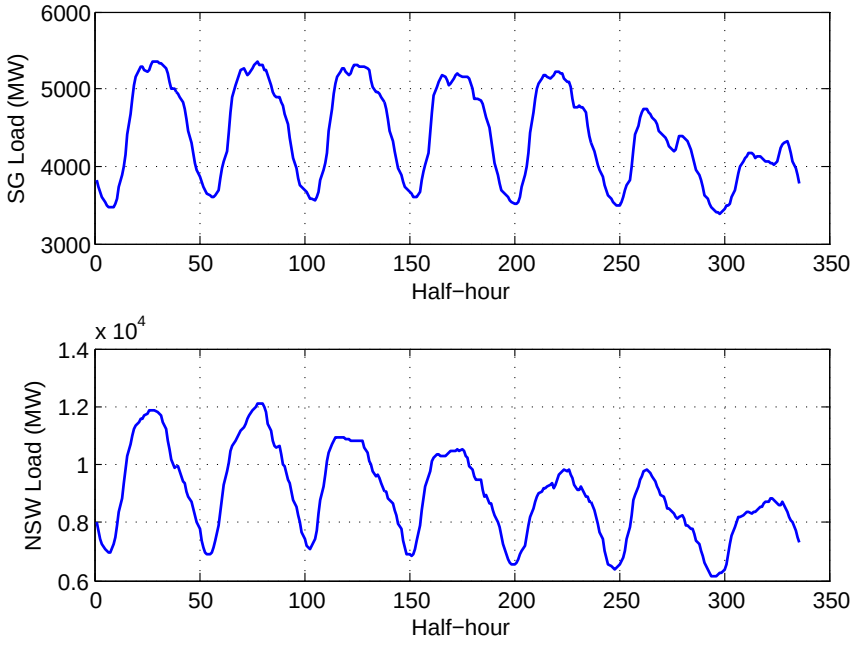

Figure 4.2: Typical weekly load of SG and NSW (22nd to 28th Jan 2007)

The wind power generation datasets are from Capital Wind Farm (Captl_WF). The Captl_WF is located in NSW, around 30 kilometers north east of Canberra, just southeast of Lake George and north of Bungendore. The wind farm was completed in 2009 and cost around A$220 million. The wind farm was built as part of the Kurnell Desalination Plant project to offset the power usage of the desalination plant. The total capacity is 140 MW. The original 5-minute interval datasets are combined into one-hour interval ones using average values. Some missing points are filled by the neighborhood values. The whole year of 2010 is chosen, out of which, the first six months are used for training, the following three months for validation and last three months for testing. Since one-day (24-hour) ahead wind power generation forecasting is commonly used in the UC and economic dispatches (EDs), one-day ahead PIs for wind power generation is implemented in this chapter.

4.5.2 Correlation Analysis

In time series analysis, identification of models usually relies on correlation analysis. For ARIMA models, $ARIMA(p, d, q)$, the autocorrelation function (ACF) and partial ACF (PACF) are used to determine the q and p orders. ACF and PACF are useful tools to analyze the correlation between the forecast values and the historical datasets. Inspired by the successful applications of ACF and PACF, we apply them to determine the input values of NN that are most related to the forecast values. However, in time series analysis, only for a stationary time series can make ACF and PACF more sense. A time series that is seasonal or has varied mean and variance values is absolutely nonstationary.

Unfortunately, the load datasets are nonstationary because they are seasonal time series with daily and weekly patterns. Thus ACF and PACF make little sense on the original datasets. How to remove the seasonality of datasets becomes important for ACF and PACF analysis. Differencing is a common way to make a time series stationary. Since one-week ahead load forecasting is

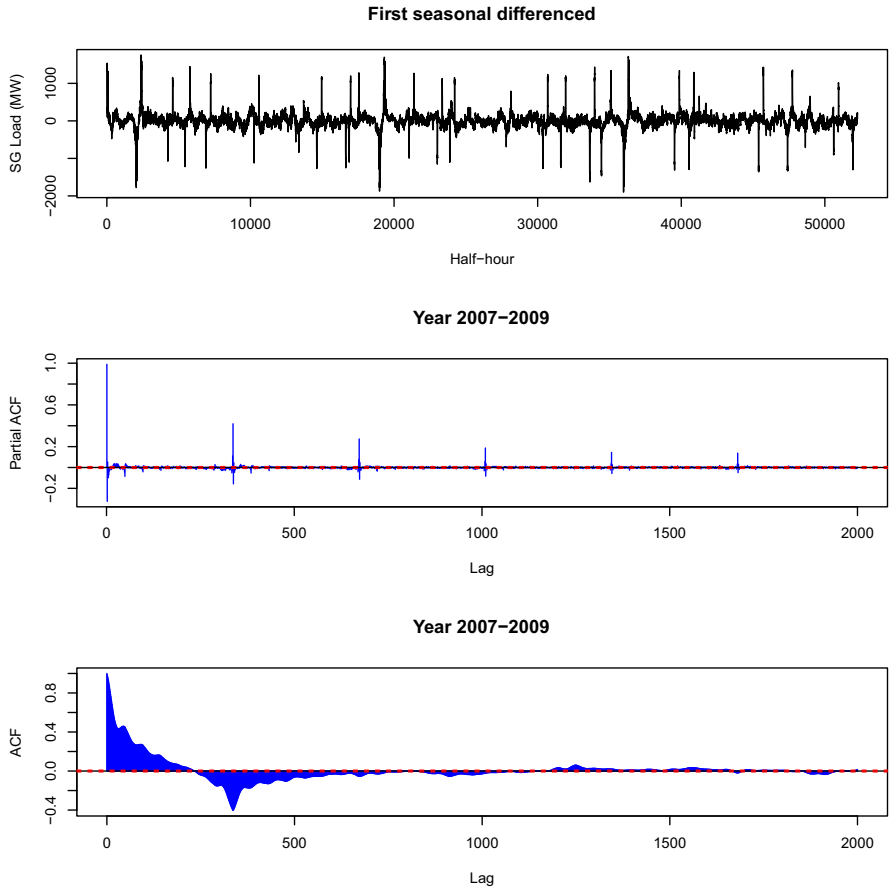

Figure 4.3: ACF and PACF analysis of first seasonal differenced SG load from Year 2007 to 2009

studied in this chapter, first seasonal difference [131] is conducted here:

$$y(t) \;=\; y(t_0) \;-\; y(t_0 - T) \tag{4.17}$$

where T is the cycle of a time series, here T is the weekly cycle: $T = 48 * 7 = 336$. $y(t_0)$ and $y(t)$ are the time series before and after the first seasonal difference. ACF and PACF analysis of first seasonal differenced SG and NSW load from Year 2007 to 2009 are implemented. The example of SG is shown in Fig. 4.3.

The implementations of wind power generation PIs are similar to the short-term load forecasting as shown in the flow chart of Fig. 4.1. The main difference is that no differencing is conducted. Wind power is the intermittent and volatile renewable resource. It fluctuates from time to time, so it has no obvious daily and weekly patterns. Then no seasonal differencing is conducted on wind power generation forecasts. The PIs are constructed on the original datasets.

Furthermore, some limitations are set to the upper and lower bounds of Captl_WF. The upper limitation is set to the capacity of 140MW, and the lower limitation is 0.

4.5.3 Inputs to the Neural Network

How to determine the inputs of NN for STLF is still an open question. Previously they were usually determined based on experience or a priori knowledge about the behavior of the system. A rather intuitive guess is that, there must be homologous instants in the past to the current instant, either the same moment yesterday (24 hours ago) or the same instant one week ago, two weeks ago and so on [132].

In this chapter, the inputs of NN are chosen based on the ACF and PACF analysis, as well as some empirical guidelines from references [91, 132–134]. The first input of NN for load forecasting is the day mark. The day mark in each week noted as $\{1, 2, 3, 4, 5, 6, 7\}$, then it is normalized and added to the input set of NN. The day mark can distinguish different daily load patterns in each week. Lagged (one-week ahead, $lags > T$) values of the larger absolute ACF and PACF are also considered. Based on the trial and error method, there are totally 16 inputs for NN load forecasting models. We also leave half an hour time slack for power system weekly ahead planning, such as the UC scheduling and EDs. The number of inputs for NN wind power generation forecasting is chosen as 24 without day marks. In the following of the chapter, training of NN for load forecasting is based on the datasets after first seasonal difference. When calculating the assessment indices of PIs, the datasets are transformed back into the original load. While the PIs of wind power generation forecasting are constructed on the original datasets.

4.6 Results and Discussions

Representations of results for NN-based PIs mainly consist of three parts: the training process, test results and discussions on quality of PIs. The training process shows the convergence behavior of PICP and PINRW for the g_{best} particle. It implies how they change through the optimiza-

tion process. The final performance of the proposed method is examined by test set. Test results are cast in the form of figures and tables. Finally, the qualities of PIs including the repeatability and computation time are further discussed.

4.6.1 Determination of NN Structure and Parameters

Fully connected feed-forward NNs with two hidden layers are chosen in three case studies. The activation functions in the hidden and output layers are tansig and purelin respectively. The quality of PIs is sensitive to the structure of NNs. Too small NNs have a poor learning capacity and too large NNs have a low generalization power. They also suffer from under-fitting and over-fitting problems. How to determine the number of neurons in the two hidden layers becomes crucial for constructing high quality PIs. In literatures, several methods such as network pruning [101–103], cascade correlation [99] and hybrid evolutionary NN construction [100] have been applied to determine this optimality. In our previous work [55], a k-fold cross validation was applied to address this problem. Because in load and wind power forecasting the time series is in sequence, the k-fold cross validation cannot be directly implemented here. But the idea is very similar to cross validation. The number of neurons in the two hidden layers (n_1 and n_2) vary from 1 to 10, $1 \leq n_1, n_2 \leq 10$. Thus there are totally $10 * 10 = 100$ candidate NN structures. Each candidate NN is trained and validated for 5 times using the training and validation set. Then the median PINAW of the validation set is used as the criterion for a better structure. If the hard constraint of $\mu \leq PICP(\omega)$ has not been met, then the object of PINAW will be arbitrarily set to a very big value. Thus this candidate NN structure will be discarded.

Determination of the NN structure needs to balance between the network complexity, generalization, and learning capacity of NNs. Fig. 4.4 shows the NN structure versus the median PINAW of the validation set of NSW load. The lowest point in this 3D-plot is chosen as the NN structure. As shown in Fig. 4.4, the number of neurons in the hidden layers are $n_1 = 8$ and $n_2 = 1$ for the lowest point. So the chosen NN structure is $16 - 8 - 1 - 2$ for NSW load. Due to

the space limitation, the similar plots for SG load and Captl_WF are omitted, the corresponding NN structure is $16-5-1-2$ and $24-8-5-2$.

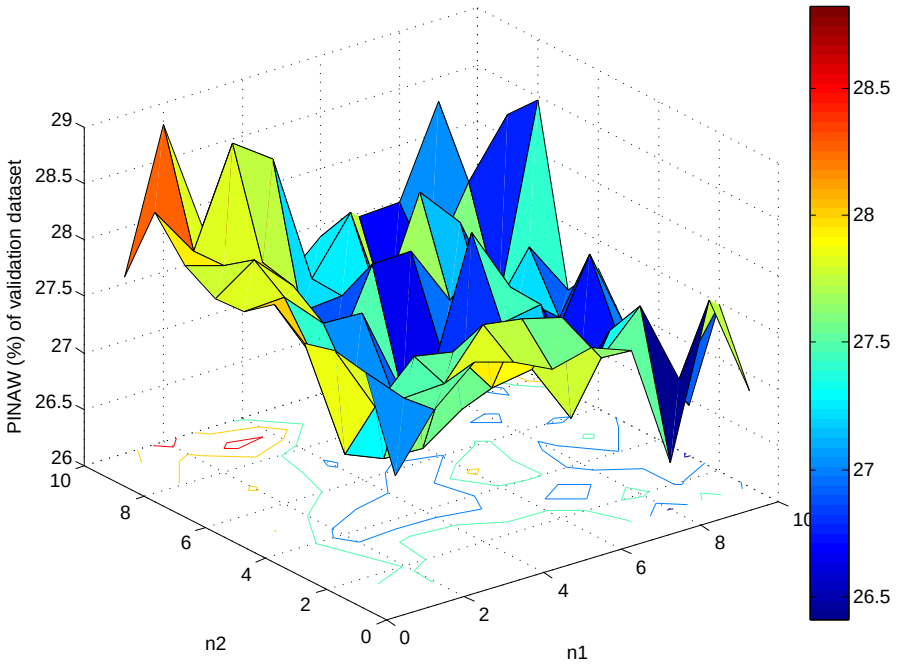

Figure 4.4: Optimal NN structure of NSW load data

Besides the optimal NN structure, the validation set can also be used to determine other parameters in the algorithm. Actually the cost function CWC is not necessary needed to solve the constrained single-objective optimization problem. For the purpose of comprehensive evaluation, CWC is also applied for testing. Table 4.1 shows the typical parameters of three case studies for PSO and CWC. W_{max} and W_{min} are the maximum and minimum value of the inertia weight W for previous velocity. W plays an important role in controlling the PSO convergence and it linearly decreases as iterations increase.

The nominal coverage probability $\mu_{nominal}$ is 90%. When training NNs, the μ_{train} is set to 91%–93% according to performances of the validation set. Usually μ_{train} is 1–3% greater than $\mu_{nominal}$. This conservation leaves a slack for the test set. In this way, the nominal coverage probability will be easily reached for testing.

Table 4.1: Experiment Parameters in STLF

	Parameter	Numerical value
PSO	C_1	1.22
	C_2	1.49
	W_{max}	1
	W_{min}	0.7
CWC	$\mu_{nominal}$	90%
	μ_{train}	91%–93%
	η	50

4.6.2 Training Process

Fig. 4.5, 4.6 and 4.7 show the PICP and PINRW of g_{best} particle in iterations for SG, NSW load and Captl_WF separately. The population size is 80 for SG and NSW load, and 200 for Captl_WF; the numbers of iterations are all set to 2000.

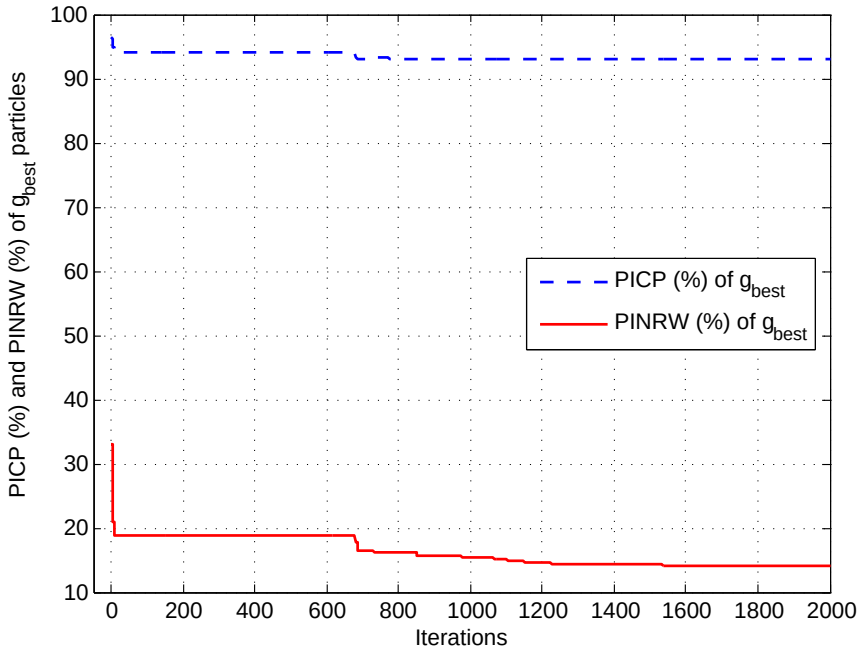

Figure 4.5: PICP and PINRW of g_{best} particle during training for SG load data

As shown in Fig. 4.5, 4.6 and 4.7, the training processes for all case studies simply converge. The PICP of g_{best} particle for load forecasting only has a rapid drop at the beginning.

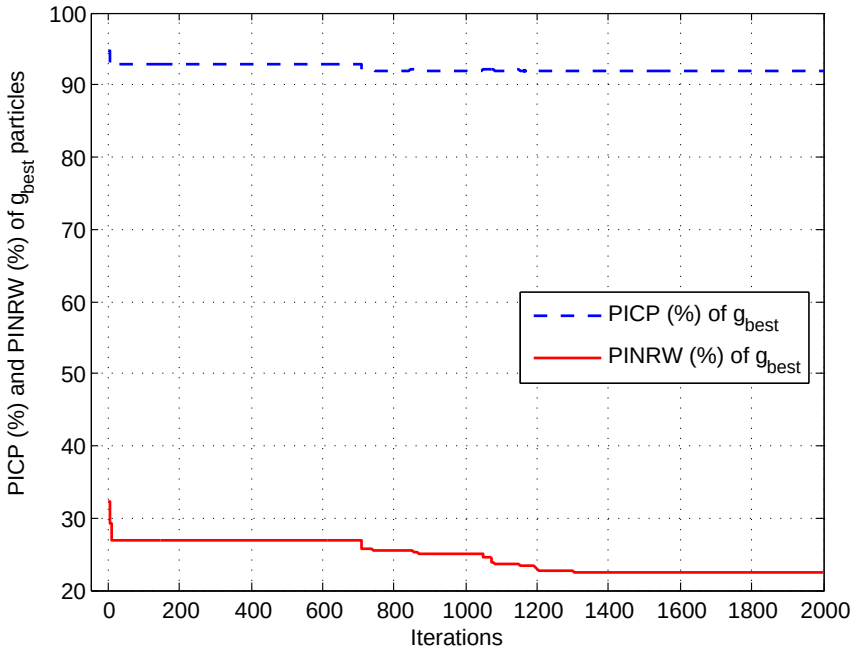

Figure 4.6: PICP and PINRW of g_{best} particle during training for NSW load data

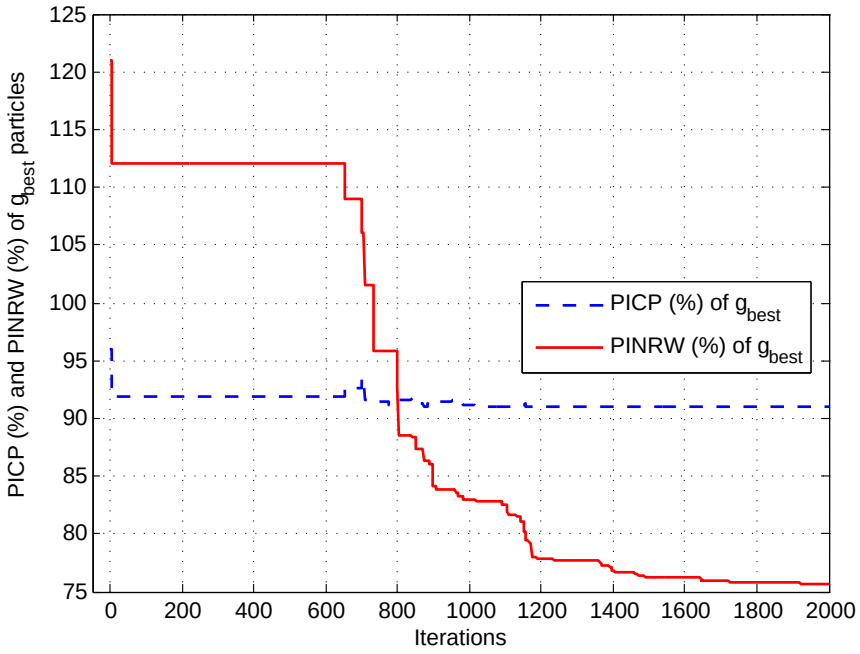

Figure 4.7: PICP and PINRW of g_{best} particle during training for Captl_WF wind data

After the first few iterations, PICP makes little changes and quickly reaches to the preassigned coverage probability μ_{train}. For the g_{best} PICP of Captl_WF, not only a rapid drop happens at the beginning, but also it has a small perturbation in the middle. This implies the algorithm pays

much more attention to the hard constraint at first. The particles with satisfied PICP have the priority to survive and are chosen as the g_{best} particle. It just meets our design of the problem formulation. The PINRW of g_{best} particle decreases sharply at the beginning. This means once the hard constraint of μ_{train} is satisfied, the algorithm quickly shifts the emphasis on smaller objective function of PINRW. As the optimization proceeds, PINRW gradually plateaus. Even from 100 to 600 iterations, PINRW minimization makes little improvement. But after about 600 iterations, PINRW continues to reduce step by step. Finally PINRW takes its optimal value and converges to a good solution. This implies the strong searching capacity of PSO combined with the mutation operator.

4.6.3 Test Results

For unbiased assessments, the whole year of 2011 is used for testing. So there are totally $365 * 48 = 17520$ test samples for load forecasting. The number of test samples for Captl_WF is $92 * 24 = 2208$ (last three months of 2010). For better visualization, test result figures for only one week are shown in Fig. 4.8, 4.9 and 4.10. The numerical results of the whole test set are shown later in Table 4.2.

As can be seen from Fig. 4.8, 4.9 and 4.10, the constructed PIs cover the real test samples in a great percentage. The real test samples (blue line) lie within the constructed lower and upper bounds (pink and red lines) in most cases. For SG and NSW load forecasts, the shapes of the three lines are very similar to each other. But for Captl_WF, the upper bounds and the real test data have strong fluctuations, which indicate the high uncertainties of wind power. The lower bounds of Captl_WF mostly drop to zeros. That is because in the training and validation sets, wind power outputs unexpectedly drop to zeros frequently. The percentage of zero values is 30.42% in the whole Captl_WF dataset. This is called the intermittence of wind power. In order to cover the real test data with a high PICP ($\geq 90\%$), the lower bounds have to be set down as zeros frequently.

Table 4.2: Evaluation on the Quality of PIs Between the Proposed Method and Benchmark Models for Test Sets

PSO-Based	**SG Load**			**NSW Load**			**Captl_WF**		
LUBE Method	**PICP(%)**	**PINAW(%)**	**CWC(%)**	**PICP(%)**	**PINAW(%)**	**CWC(%)**	**PICP(%)**	**PINAW(%)**	**CWC(%)**
1	91.75	15.95	15.95	90.40	23.65	23.65	90.40	71.57	71.57
2	90.98	16.10	16.10	90.34	23.45	23.45	92.30	75.78	75.78
3	91.20	15.85	15.85	90.20	23.50	23.50	92.16	74.07	74.07
4	91.33	16.08	16.08	90.70	26.52	26.52	91.80	72.43	72.43
5	91.20	16.05	16.05	90.00	23.41	23.41	90.67	72.57	72.57
Median	91.20	16.05	16.05	90.34	23.50	23.50	91.80	72.57	72.57
ARIMA	91.16	18.66	18.66	88.80	20.79	58.74	91.44	75.58	75.58
ES	91.38	18.73	18.73	88.75	22.02	63.15	92.71	78.45	78.45
Naive	90.14	22.24	22.24	88.65	24.24	71.92	88.81	74.43	209.16
Time(ms)		12.30			14.56			4.27	

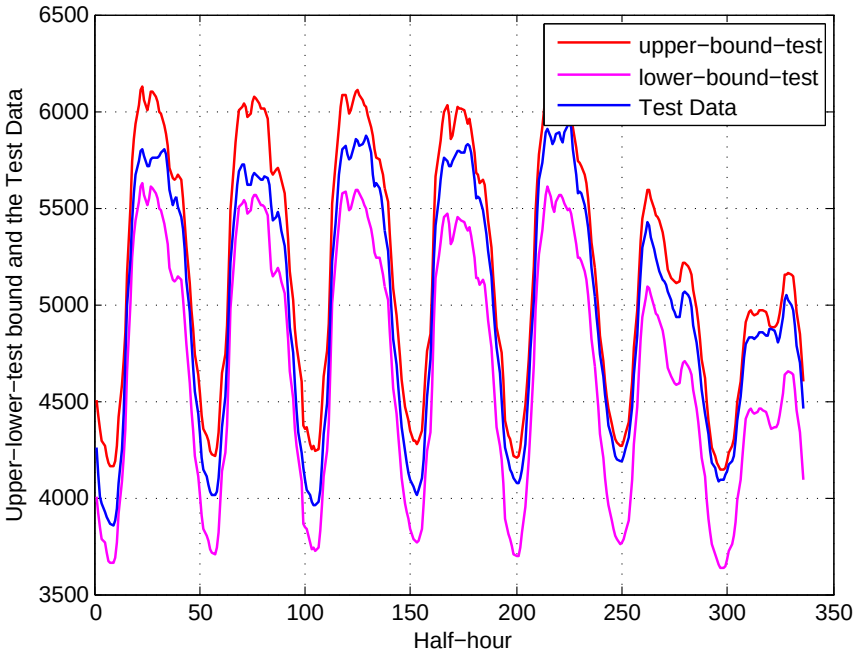

Figure 4.8: SG weekly load and PIs for testing (21-27 Mar. 2011)

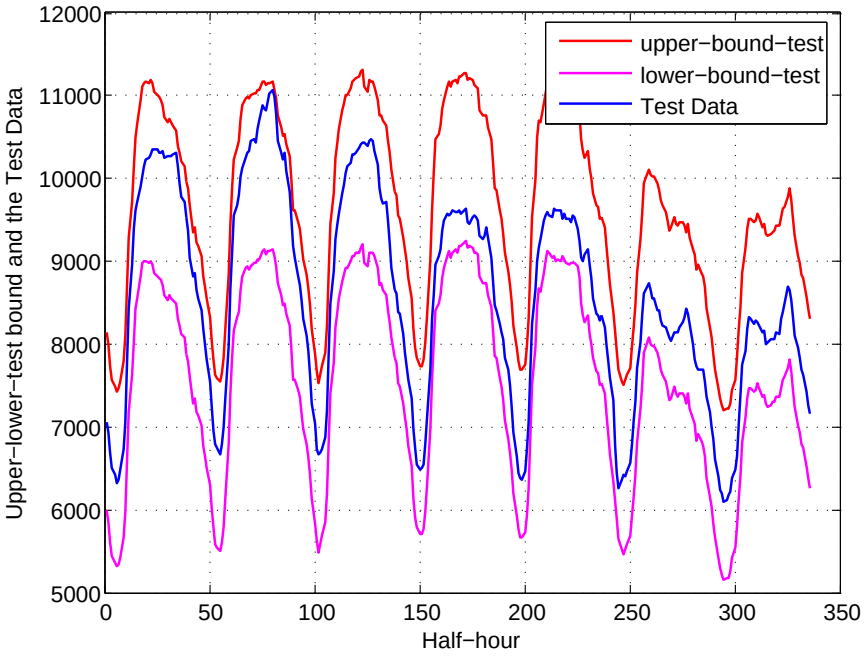

Figure 4.9: NSW weekly load and PIs for testing (21-27 Mar. 2011)

High PI coverages imply that the PICP indices for the test samples are very satisfactory using the proposed method. In this way, the validity of PIs has been confirmed. On the other hand, the width of PIs for SG load is different from NSW load. They can be very narrow, as

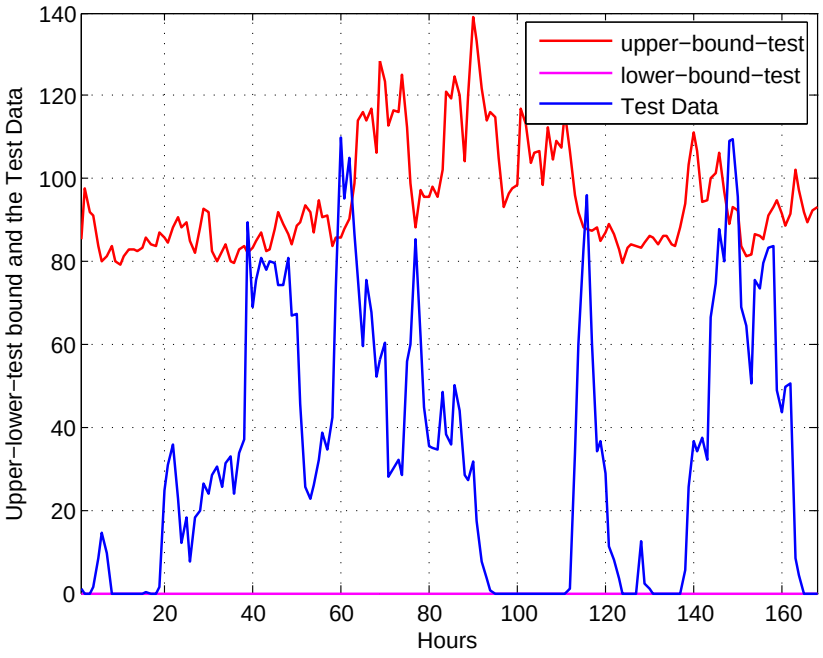

Figure 4.10: Captl_WF weekly generation and PIs for testing (1-7 Oct. 2010)

the pink and red lines of the lower upper bounds are tight with each other. It is wider for NSW load. This is because, as mentioned before, load pattern of NSW is more irregular than SG due to the climate and regional reasons. Moreover, for all SG and NSW loads in Fig. 4.8 and 4.9, PIs become wider during the daily maximum and minimum load demand periods. These are periods with the maximum level of uncertainty in operation of the energy system. These uncertainties are appropriately reflected in the width of PIs.

Furthermore, widths of PIs for Captl_WF are much larger than load forecasts. The widths of PIs are determined by the uncertainty level of the datasets. Under the preassigned PICP, a lower level of uncertainty results in narrower PIs and vice versa. Thus the PSO-based LUBE method can construct high quality PIs for datasets under different levels of uncertainty. It can handle different types of prediction tasks.

4.6.4 Discussions on Quality of PIs

In order to validate the repeatability of the algorithm, and provide quantitative and convincing results, each case study is repeated 5 times. Results in each run as well as the median values of PICP, PINAW and CWC instead of the best ones are displayed in Table 4.2. PI construction time for test samples is also reported. Based on Table 4.2, we can have the following conclusions:

1) Demonstrated results imply the strong repeatability and stability of the proposed algorithm. For five time replicates, the results show a high consistency for PICP, PINAW and CWC. The standard deviations of CWCs for three case studies are 0.1045, 1.3525 and 1.6602. The obtained results have definitely small variations in different runs.

2) For all runs and case studies, the preassigned PICP (90%) can be satisfied. This means that the constructed PIs cover the target values with a high probability. It clearly shows that the hard constraint ($\mu \leq PICP(\omega)$) in the problem formulation is successfully met.

3) The median value of PINAW for NSW load is 23.50%, that is obviously larger than 16.05% of SG load. The average widths of PIs are different for two case studies. Under certain PICP, widths of PIs rely on the uncertainty level of the datasets. NSW load has a higher level of uncertainty thus its PIs are wider than PIs of SG load. What's more, the widths of PIs for wind power generation are obviously lager than the load forecasting. Although the forecast horizon is one-day ahead, PINAWs of PIs are still much larger than one-week ahead load forecasting. This strongly indicates the high uncertainty of wind power.

4) PI construction time is one of the key characteristics for algorithm design. This is especially true for on-line applications. Under a hardware configuration of Intel(R) Core(TM)2 Duo CPU E8500 3.16GHz, and 4 GB of RAM, the average PI construction time for the test set is 12.30ms, 14.56ms and 4.27ms separately. This is very fast and as simple as point forecasts.

4.6.5 Comparison of Results with Benchmark Models

For unbiased comparisons, ARIMA, ES and naive models are used as three benchmarks for one-week ahead load and one-day ahead wind power generation PIs using the same datasets. Two methods, iterative multi-step and direct one-step ARIMA models are tried. For the iterative multi-step method, a multiplicative seasonal ARIMA model [135] is built. Double seasonal intra-day and intra-week cycles are considered in this method. The iterative multi-step method works well on point forecasts. But the widths of PIs increase very quickly with the increasing of time steps. PINAWs of PIs for one-week ahead are even larger than 100%. These are too large and are of no use for comparisons.

Multi-step models have the risk of running accumulative errors. Usually, ARIMA models have a better performance for one-step ahead forecasting. The concept of direct one-step method is to resample the original half-hour interval datasets into the new one-week interval time series. One step of the new time series is one week. Thus the one-step ahead forecasting on the new time series using ARIMA models can directly construct one-week ahead PIs. For example, to forecast load point $y\{t\}$, the chosen time series is:

$$\begin{aligned} & y\{t-LT\},\ y\{t-(L-1)T\}, y\{t-(L-2)T\} \\ & ,\ .....\ ,\ y\{t-3T\},\ y\{t-2T\},\ y\{t-T\} \end{aligned} \tag{4.18}$$

$T = 48 * 7 = 336$, is the weekly cycle for load forecasting. L is the length of "look-back" weeks. One year contains 52 weeks, if the "look-back" length is 4 years, then $L = 52 * 4 = 208$. PI of $y\{t\}$ is then constructed based on the resampled time series. For the wind power generation forecasts, the forecast horizon is one-day ahead, so the resampling cycle in Equation (4.18) is $T = 24$ for 24 hours, and the "look-back" length L is nine months (approximately $L = 9 * 30 = 270$ days).

The above direct one-step method is conducted on the ARIMA, ES and naive models to construct 90% PIs. The naive model is similar to the persistence model of point forecasts, which

states that the variable's future value will be the same as the last one measured. The simulations of the three benchmark models are implemented in a statistical software named R [136]. The numerical results of the three models are listed in Table 4.2. As can be seen from the table, the proposed PSO-based LUBE method outperforms the ARIMA, ES and naive models. The quality of PIs has been significantly improved. Since CWC has a comprehensive evaluation on both aspects of PIs, the following discussions on the improvements are mainly focused on CWC. The percentage improvement is defined as:

$$\frac{Compared\,result\ -\ New\,result}{Compared\,result}\ *\ 100\% \tag{4.19}$$

Table 4.3: CWC Percentage Improvements to Benchmark Models

Percentage Improvements		**ARIMA**	**ES**	**Naive**
SG Load	**Median**	13.99%	14.31%	27.83%
	Best	15.06%	15.38%	28.73%
NSW Load	**Median**	59.99%	62.79%	67.32%
	Best	60.15%	62.93%	67.45%
Captl_WF	**Median**	3.98%	7.50%	65.30%
	Best	5.31%	8.77%	65.78%

For the three case studies, percentage improvements of the median and best CWCs compared with ARIMA, ES and naive models are listed in Table 4.3. For the NSW load, all three benchmarks fail to construct valid PIs with satisfied PICPs. Although PINAWs of the first two models are narrower than the proposed method, their PICPs are all unsatisfied and lower than the preassigned value of 90%. Thus CWCs put a penalty on the violation of the PICP hard constraint. This penalty term is also added to naive model of Captl_WF. In all the four methods, the proposed method obtains the best results, while the naive models have the worst results for three case studies. Further, the proposed method uses only one NN model for one test set. But then, the three benchmark models apply multiple forecast models. The number of models is equal to

the number of test samples.

What is more, the LUBE method outperforms traditional methods on both the quality of PIs and the computation time. These advantages have been verified in our previous work reported in [43, 55]. LUBE method can construct PIs with satisfied PICP and narrower PINAW than traditional methods. On the other hand, PI construction requirement of traditional methods is at least 10 times more than LUBE methods [43]. Thus the proposed method can construct higher quality PIs in a shorter time for load and wind power generation forecasting applications.

4.7 Conclusions

STLF and the renewable energy forecasting are of great importance for control and scheduling of smart grids. The uncertainty of power systems increases due to the penetration of renewable energy such as wind and solar power. To overcome the deficiencies of point forecasts to handle uncertainty, this chapter implements the short-term load and wind power forecasting using NN-based PIs. PIs are excellent tools for quantification of uncertainties associated with point forecasts and predictions. Traditional methods for PI construction suffer from various problems. A newly proposed method called LUBE method is adopted and further developed to construct PIs. The primary multi-objective problem is successfully transformed into a constrained single-objective problem. Advantages of this new problem formulation are closer to the original problem and having fewer parameters than the cost function. PSO with a strong searching ability for parameter adjustment is integrated with the mutation operator. With the enhanced searching capacity, PSO is then used to solve the new problem and train the NN models. Correlation analysis is applied to help choose the inputs of NN models. Comparative results for both load and wind power generation datasets show that not only the high PICP and narrow PINAW are obtained, but also the PI construction time remains short. Quality of PIs is significantly improved compared with ARIMA, ES and naive models. In conclusion, the proposed PSO-based LUBE method constructs higher quality PIs for different types of prediction tasks in a short time.

Chapter 5

Incorporating Wind Power Interval Forecast Uncertainties into Stochastic Unit Commitment for Decision Making

5.1 Introduction

Due to the random nature of weather, intermittent power generation sources, such as wind and solar systems, have great uncertainties. With sudden weather changes, the output power of a wind farm can drop most part of its power or even drop to zero. High wind power penetration will have significant impacts on power system operation economics, stability, security, and reliability due to fast fluctuation and unpredictable characteristics of wind speed [137]. Unit commitment (UC) and economic dispatch (ED) are essential and basic tasks in power system daily operations. UC problems incorporated with wind generation uncertainties become a more important and challenge task than before in smart grid applications [2, 3, 138].

UC problem is usually formulated to minimize the total generation costs under some operating constraints. This formulation is also called the security constrained unit commitment (SCUC) problem. UC solution methods can be classified into three categories: deterministic

methods, meta-heuristic methods and hybrid models. Deterministic methods include priority listing (PL) [66], dynamic programming [66], Lagrangian relaxation (LR) [67], integer and linear programming. Meta-heuristic approaches include expert systems [69], fuzzy logic [70], genetic algorithm (GA) [71, 72], evolutionary programming, simulated annealing, tabu search, particle swarm optimization (PSO) [73], ant colony optimization and differential evolution (DE) [75]. Hybrid models, where one method compensates with another, may have a better performance than individual models. Srinivasan et al. in [76] proposed the PL method-based evolutionary algorithm to solve the UC problem. PL was used as a good initialization of the evolutionary algorithm. In [63], the authors proposed a hybrid LRGA model; GA was applied to evolve the Lagrange multipliers to improve the LR method.

Conventional wind power forecasting produces a value, or the conditional expectation (point forecasts) of wind power output at a time point in the future [36]. However, currently wide-used point forecasts cannot provide additional quantitative information on the uncertainty associated with wind power generation [139, 140]. From the decision making and risk assessment point of view, probabilistic forecasts of wind power, such as prediction intervals (PIs), quantiles or scenarios, are optimal inputs [36].

For point forecasts (cannot represent uncertainty), there is only one forecast value at each time point. Power system operators can straightforwardly use this single value for decision making. But for PIs, even a single level PI consists of three components: an upper bound, a lower bound and the corresponding confidence level [19]. It is hard to directly apply the two bounds to make a decision. Therefore, the linkage between the PIs and decision making becomes urgent for smart grids [55, 141]. In this chapter, a computational framework is proposed to build this important linkage through a proposed scenario generation method and the stochastic modeling.

In order to incorporate the wind power forecast uncertainties into the UC problems, some research work has been done in references. Wu et al. [49] implemented and compared the stochastic SCUC and interval optimization approaches. In the stochastic SCUC, the scenarios were gener-

ated with a presumption that the wind speed uncertainty follows a Weibull probability distribution function with the autocorrelation factor and diurnal pattern. The interval optimization used confidence intervals in terms of upper and lower bounds to represent the uncertainty spectrum, and derived optimistic and pessimistic solutions [49]. Wang et al. in [78] emphasized on the aspects of intermittence and volatility of wind power in SCUC. Wind power was assumed to follow a normal distribution and Monte Carlo simulation was used to generate scenarios subject to this normal distribution.

Ortega-Vazquez et al. [79] estimated the spinning reserve (SR) requirements in systems with significant wind power generation penetration. Wind power generation was viewed as a negative load. The net demand forecast error was generated using a Gaussian cumulative probability distribution. Sturt et al. [82] used stochastic unit commitment (SUC) with rolling planning for simulation of wind-integrated power systems. It presented a new formulation of the SUC problem and used a quantile-based scenario tree structure. The performances of various tree topologies in year-long simulations of a large system were compared. Jiang et al. [83] studied the robust UC with wind power and pumped storage hydro. The robust UC [84, 86] represented the wind power uncertainties in a predetermined uncertainty set containing the worst-case scenario. Benders' decomposition algorithm was used to obtain a robust UC solution.

In spite of the advantages of the above references, the wind power scenarios in the stochastic model are mainly generated from an assumed probability distribution, either a normal distribution [78, 79] or a Weibull probability distribution [49]. The robust UC needs to predefine the uncertainty set and find the worst-case scenario [83, 84, 86]. To generate wind power scenarios, in [142, 143] the complex covariance matrix needs to be calculated based on a multivariate Gaussian distribution assumption. The main contributions of this chapter are summarized below:

1) A new scenario generation method is proposed to capture the uncertainty of the wind power forecasting. The proposed method can be easily implemented and avoids making strong assumptions on the wind power probabilistic distributions.

2) A nonparametric lower upper bound estimation (LUBE) is used to generate neural network (NN)-based PIs.

3) Meta-heuristic GA is used to solve the stochastic SCUC problem. The obtained total cost is among the best one from the relevant literature.

4) Five deterministic and four stochastic case studies are implemented. These different UC strategies are further discussed and compared together.

5) Different reserve strategies are investigated. The scheduled reserve and real time ED reserve are compared together.

6) The computational framework in this chapter overcomes the multivalued problem when the PIs are applied for decision making in power system operations. Some guidelines are also provided for system operators from the risk assessment perspective.

The rest of this chapter is organized as follows. Wind power forecasting using NN-based PIs are introduced in Section 5.2. Section 5.3 proposes a new scenario generation method to handle the interval forecast uncertainties. Problem formulation of the stochastic SCUC problem is described in Section 5.4. Section 5.5 presents the GA-based solution method for the stochastic SCUC problem. Case studies are introduced in Section 5.6. Simulation results and discussions are provided in Section 5.7. Finally, Section 5.8 concludes the chapter with some remarks for further study in this domain.

5.2 PIs for Wind Power Forecasting

PIs are powerful tools to handle uncertainties associated with point forecasts. In our previous work [60], a new PSO-based LUBE method was proposed to construct NN-based PIs for electrical load and wind power forecasting. The basic idea of LUBE method is to use a NN with two outputs to directly generate the upper and lower bounds of PIs [43, 60]. The first output of NN is

the upper bound of a PI and the second output corresponds to the lower bound.

Delta, Bayesian and bootstrap are traditional methods to construct prediction intervals (PIs) based on NNs [41, 43]. Compared to traditional methods, the key advantages of the LUBE method for PI constructions are as follows:

1) It is simpler and constructs higher quality PIs in just one step. Traditional methods first do point forecasting and then construct PIs.

2) As known to all, wind power is volatile and intermittent in nature. Thus, assumptions about the data distributions seem problematic and in doubt. The LUBE method is a nonparametric method and no assumption on data distribution is made. Traditional methods always consider a parametric distribution (e.g., Normal) for data and then try to find its parameters for construction of intervals [41].

3) Its computational burden for PI construction is significantly lower than alternative methods [43]. This is due to the fact that the developed NN directly generates PIs. Alternative methods need to linearize NN models or calculate complex matrices such as the Jacobian and Hessian matrices [43].

In practice, PIs with a high coverage probability and narrow width are expected for decision making. However, these two aspects of PIs contradict with each other. For example, increasing the coverage probability will also widen the PIs while squeezing the PIs may lead to a lower coverage probability. This multi-objective problem can be formulated and solved in different ways. In [43,58], this problem is modeled as a single-objective problem through a comprehensive cost function. In [60], the primary multi-objective problem is transformed into a constrained single-objective problem. The preassigned coverage probability is treated as a hard constraint and the only objective is to minimize the width of PIs. Compared with the cost function method, advantages of this new formulation are: (i) it is closer to the primary problem; and (ii) it has fewer parameters and avoids choosing the format of cost functions. The extra parameters used in

the cost function are not required in the new formulation.

In [60], PSO, which is powerful for parameter optimization, is employed to solve the problem. The mutation operator, which helps to achieve diversity in GA, is also integrated into PSO to improve the exploratory capabilities and help in jumping out of local optima. The objective of using PSO has mainly two aspects. For one thing, PSO is used to solve the newly formulated constrained single-objective problem. That is, to handle the constraints and optimize the objective. For another thing, PSO with mutation operator is used as the training algorithm through optimizing the connection weights [144] of NN models. To solve the constrained single-objective problem using PSO, the criteria for replacing one particle $\vec{a}$ with another particle $\vec{b}$ are as follows [108, 130]:

1) Particle $\vec{b}$ is feasible and particle $\vec{a}$ is not;

2) Both particles are feasible or they have the same satisfaction of constraints, but $\vec{b}$ yields a better objective function value.

3) Both particles are infeasible, but $\vec{b}$ results in the lower sum of constraint violations.

The wind power data used in this chapter is from Alberta wind farm, Canada in 2012 [145]. The data set is preprocessed and has an hourly resolution. It is well known that wind power is intermittent and performs like a stochastic process. Different from electrical demand, wind power has no obvious daily or weekly patterns. In this chapter, the autocorrelation function and partial autocorrelation function analyses [136] are applied and help to choose the inputs of NN. Based on the correlation analysis of the 2012 wind power data and the trial and error method, the 24 points in the previous day are chosen as the inputs of NN. Although the PSO-based LUBE method is used to forecast a list of PIs of wind power in this chapter, the methodology can be extended into other forecasting tasks for uncertainty modeling, such as the electrical load and solar power forecasting etc. To make the chapter self-contained, a brief introduction of the method for wind power interval forecast is provided. More details about the method can be found in [60].

5.3 Scenario Generation from the Wind Power PIs

The uncertain wind power is firstly captured by a list of PIs. Then a new scenario generation method is proposed to generate scenarios from PIs. These scenarios are further incorporated into the stochastic model for decision making.

5.3.1 Forecast Uncertainty Handling Using a List of PIs

In [60] only one level PIs, the 90% PIs, are constructed using the PSO-based LUBE method. For most decision-making processes, such as power system operation, a single level PI is not sufficient for making an optimal decision for a given lead time. Theoretically, any type of probabilistic distributions can be represented by a sufficient number of its PIs or quantiles [142]. Different from our previous work, a list of PIs are generated ranging from 5% to 95% with an increment of 5% for uncertainty handling. PIs of 19 levels are constructed for one-day ahead wind power generation forecasting using the PSO-based LUBE method as introduced in [60]. These PIs are applied to estimate the empirical cumulative distribution function (ECDF) of unknown probabilistic distributions of the wind power generation outputs.

5.3.2 Scenario Generation for Uncertainty Representation for Decision Making

The wind power uncertainty mainly comes from the NN-structured PIs with different confidence levels. However, the main difficulty here is how to apply these PIs to mathematical models for decision making. As mentioned before, compared to point forecasts (cannot represent uncertainty) PI suffers the multivalued problem for decision making, not to mention the fact that we have PIs of 19 different levels here. Therefore, a new scenario generation is proposed to properly represent the uncertainties included in the NN-structured PIs. The wind forecast uncertainties are represented as scenarios. Further, these generated scenarios can be mathematically involved into the stochastic model for computational purpose for decision making. The following steps show the implementation of the proposed scenario generation method.

1) **A List of Wind power PIs:** Apply the PSO-based LUBE method to make forecasts for a list of PIs with continuous levels of 5% to 95%. An example from ARIMA is shown in Fig. 5.1. The wind power datasets are from a wind farm of Alberta, Canada [145]. The forecasting periods are the 24 hours on Dec. 2, 2012. All the predicted and measured wind power values are normalized by the capacity of the wind farm.

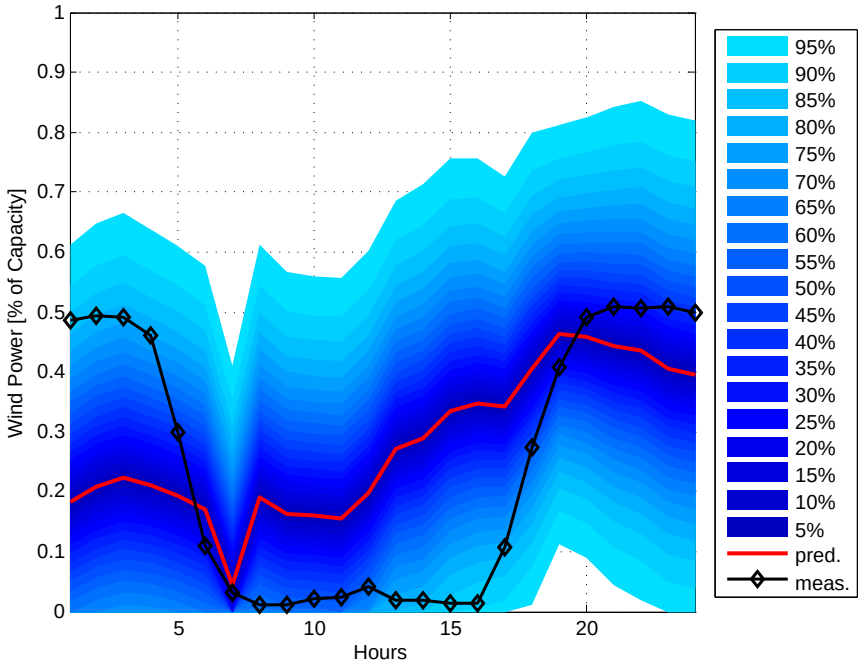

Figure 5.1: A list of PIs for one-day ahead wind power forecasting.

2) **Decomposing PIs into quantiles:** If the forecast errors are assumed to be symmetrically distributed, then each PI can be uniquely decomposed into two quantiles. The $(1-\alpha)\%$ PIs consists of two bounds, i.e. the lower and upper bounds. They correspond to the $(\alpha/2)\%$ and $(1-\alpha/2)\%$ quantiles respectively. For example, the 90% PI is decomposed into 5% and 95% quantiles of wind power and the 10% PI is decomposed into 45% and 55% quantiles.

3) **Obtaining discrete points on ECDF:** For each hour, every quantile value corresponds to one point on the ECDF curve. If the upper and lower bounds of a 90% PI at $hour\ t$ are 0.92 and 0.1 (per unit value of wind power), then two points on the ECDF curve are obtained. They are point A (0.95, 0.92) and point B (0.05, 0.1). The number of PIs are 19, so the number of

quantiles is $19 * 2 = 38$ for each hour. Besides the 38 points, two more points, point (0, 0) and point (1, 1), are also on the ECDF curve. These two points are the two ends of the ECDF curve, which mean the probabilities are 0 and 1 when the wind power outputs is less than zero and the nominal capacity. In sum, totally $38 + 2 = 40$ points are obtained on the ECDF curve.

4) **Fitting the ECDF curve:** Once the 40 points are obtained, the ECDF curve can be fitted from these points. Here piecewise cubic Hermite interpolation is used to fit the ECDF. The fitted ECDF curve passes through all the known 40 points. Fig. 5.2 shows the fitted ECDF curve of Hour 1 on Dec. 2, 2012 for the Alberta wind farm.

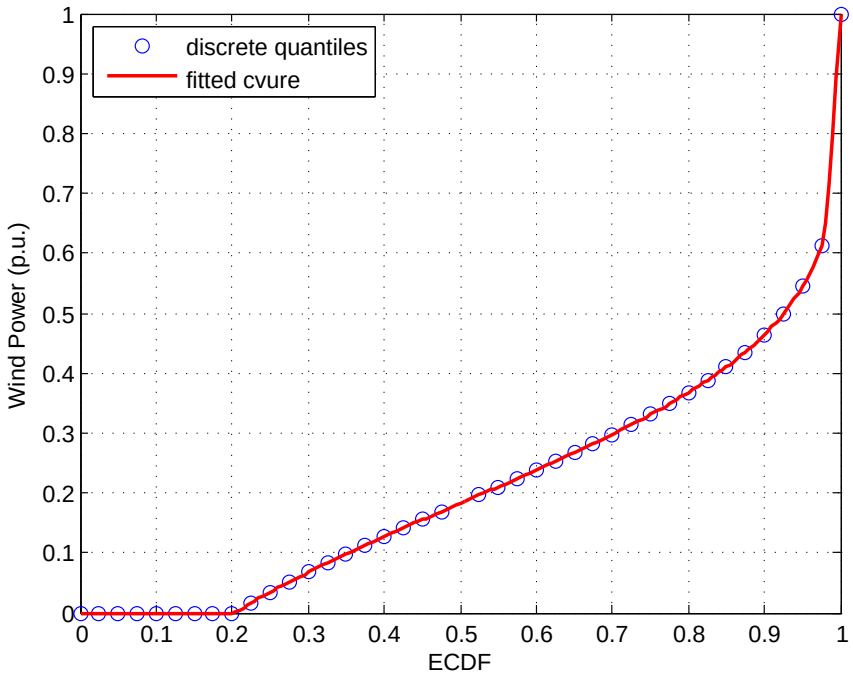

Figure 5.2: The fitted ECDF curve of Hour 1 on Dec. 2, 2012.

5) **Monte Carlo simulation to generate scenarios:** After the ECDF curve fitting, Monte Carlo simulation [11] is applied to generate scenarios from the ECDF. That is, for each scenario, a random number between 0 and 1 is uniformly generated. The corresponding wind power on the ECDF curve (per unit) is the wind power output at this scenario. This process is repeated 24 times for the 24 hours of the day. The ECDF on each hour is fitted individually and

the corresponding scenarios are generated from the ECDF of this hour. Fig. 5.3 shows the generated 50 scenarios for 24 hours on Dec. 2, 2012 for the Alberta wind farm.

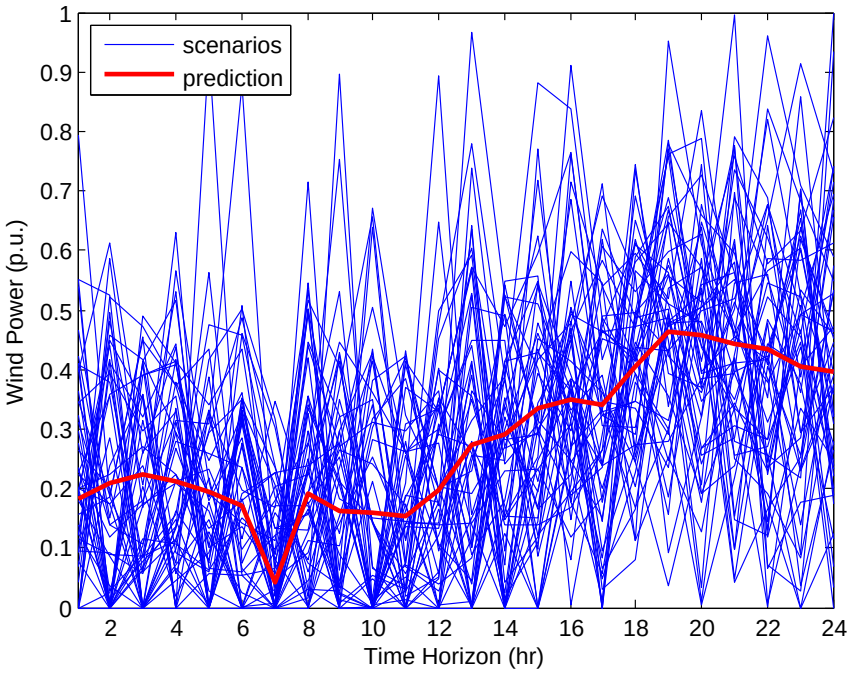

Figure 5.3: The generated 50 wind power scenarios for 24 hours on Dec. 2, 2012.

5.3.3 Advantages and Significance

Advantages of the proposed scenario generation method are as follows:

1) It can be easily implemented. The core parts of the proposed method are the ECDF curve fitting and Monte Carlo simulation. No complex computation such as the covariance matrix [142, 143] is needed.

2) It avoids making strong assumptions on the wind power probabilistic distributions. The only assumption in the scenario generation is that the forecast errors are symmetrically distributed. In this way, one level PI can be uniquely decomposed into two quantiles. Under this assumption, the 90% PI is decomposed into the 5% and 95% quantiles rather than the 6% and 96% or 4% and 94% quantiles. This assumption is a relaxation of the specific probabilistic distributions. Whereas, the normal distribution [78, 79] or a Weibull probability distribution [49]

is assumed to capture the volatility of the wind power. A multivariate Gaussian distribution needs to be assumed before the recursive estimation of the covariance matrix in [142] and [143].

The proposed method also builds an important bridge between the PIs and the scenarios used in the stochastic model. As mentioned before, power system operators need to deal with the multivalued problem when the PIs are applied to make a decision. Moreover, a list of PIs has much more values than the single level PI. These values cannot be directly used for computational purpose. The purpose of the proposed method is to generate the scenarios from a list of PIs. These generated scenarios are further computationally involved into the stochastic models for decision making and risk assessment.

5.4 Formulation of the Stochastic SCUC Problem

As shown in last section, wind power forecast uncertainties have been modeled as the generated scenarios. In this section, the intermittent and fluctuating wind power is further incorporated into the formulation of the stochastic SCUC problem. In the mathematical formulation, the W_t^s is the wind generation at time t in scenario s. As shown in Fig. 5.3, the generated 50 scenarios for 24 hours on the operating day are provided. These lines stand for the stochastic realizations of W_t^s for next day wind power generation outputs. W_t^s is then incorporated into a thermal generating UC and ED system through the power balance constraints in Equation (2.26). It further influences the expected generation costs during the UC and ED scheduling. In this way, W_t^s establishes a mathematical link between the generated scenarios and the formulation of the stochastic SCUC problem for uncertainty handling of wind power forecasting. A detailed mathematical problem formulation is shown below.

The objective of the stochastic SCUC problem is to minimize the expected generation costs under several constrains. Mathematically, the objective function is formulated as follows [63,

138, 146]:

$$E(X,P) = \min \sum_{s=1}^{S} p_s \Big[\sum_{i=1}^{N} \sum_{t=1}^{H} F_i(P_{i,t}^s) X_{i,t} + \sum_{t=1}^{H} (C_{ens} * ENS_t^s + C_{rns} * RNS_t^s) \Big] + \sum_{i=1}^{N} \sum_{t=1}^{H} \left[SU_{i,t}(1 - X_{i,(t-1)}) \right] X_{i,t} \tag{5.1}$$

where $F_i(P_{i,t}^s)$ of thermal generators is usually represented as a quadratic function:

$$F_i(P_{i,t}^s) = a_i + b_i P_{i,t}^s + c_i (P_{i,t}^s)^2 \tag{5.2}$$

Energy not served (ENS) and the reserve not served (RNS) will be at the cost of C_{ens} and C_{rns} in the objective function. The start up cost is determined according to the continuously off time of generators:

$$SU_{i,t} = \begin{cases} HSU_{i,t}, & if\ T_{i,t}^{off} \leq T_i^{Down} + T_i^{cold}; \\ CSU_{i,t}, & if\ T_{i,t}^{off} > T_i^{Down} + T_i^{cold}. \end{cases} \tag{5.3}$$

The following constraints are also defined:

1) Power balance constraints:

$$\sum_{i=1}^{N} X_{i,t} P_{i,t}^s + W_t^s = D_t - ENS_t^s \tag{5.4}$$

2) Spinning reserve constraints:

$$\sum_{i=1}^{N} X_{i,t} \left[P_{i,max} - P_{i,t}^s \right] \geq R_t^s - RNS_t^s \tag{5.5}$$

3) Generation limit constraints:

$$P_{i,min} X_{i,t} \leq P_{i,t}^s \leq P_{i,max} X_{i,t} \tag{5.6}$$

4) Minimum up time constraints:

$$(T_{i,t}^{on} - T_i^{Up})(X_{i,(t-1)} - X_{i,t}) \geq 0 \tag{5.7}$$

where,

$$T_{i,t}^{on} = (T_{i,(t-1)}^{on} + 1) X_{i,t} \tag{5.8}$$

5) Minimum down time constraints:

$$(T_{i,t}^{off} - T_i^{Down})(X_{i,t} - X_{i,(t-1)}) \geq 0 \tag{5.9}$$

where,

$$T_{i,t}^{off} = (T_{i,(t-1)}^{off} + 1)(1 - X_{i,t}) \tag{5.10}$$

5.5 GA-Based Solution Method for the Stochastic SCUC

To solve the stochastic SCUC problem, a GA-based solution method is proposed. As mentioned in Chapter 4, compare to other evolutionary methods such as GA, PSO is powerful for parameter optimization (real numbers), especially for optimization of NN connection weights. However, in Chapter 5 the formulated optimization problem is a UC problem. It is used to determine the on/off status of generators. The chromosome is presented as binary strings (0 and 1 integer numbers). Compared to PSO, GA is more suitable and powerful for optimization of binary problems. The flow chart of the GA-based solution method is provided in Fig. 5.4.

5.5.1 Binary Representation of UC Solutions

Each chromosome of the GA represents a potential solution of the UC problem. The chromosome is a row vector, whose length is the number of scheduled hours multiplied by the number of generators. A bit string of a N units H hour UC chromosome is shown in Fig. 5.5.

5.5.2 Initialization

Initialization is crucial to the repeatability and success rate of the GA-based solution method for the stochastic SCUC problem. In this chapter, two initialization methods are applied. One third of the populations are initialized using the PL method and the remaining ones are randomly initialized. Before the PL, full-load average cost (FLAC) of each unit is calculated.

$$FLAC(i) = \frac{F_i(P_{i,max})}{P_{i,max}} = \frac{a_i}{P_{i,max}} + b_i + c_i * P_{i,max} \tag{5.11}$$

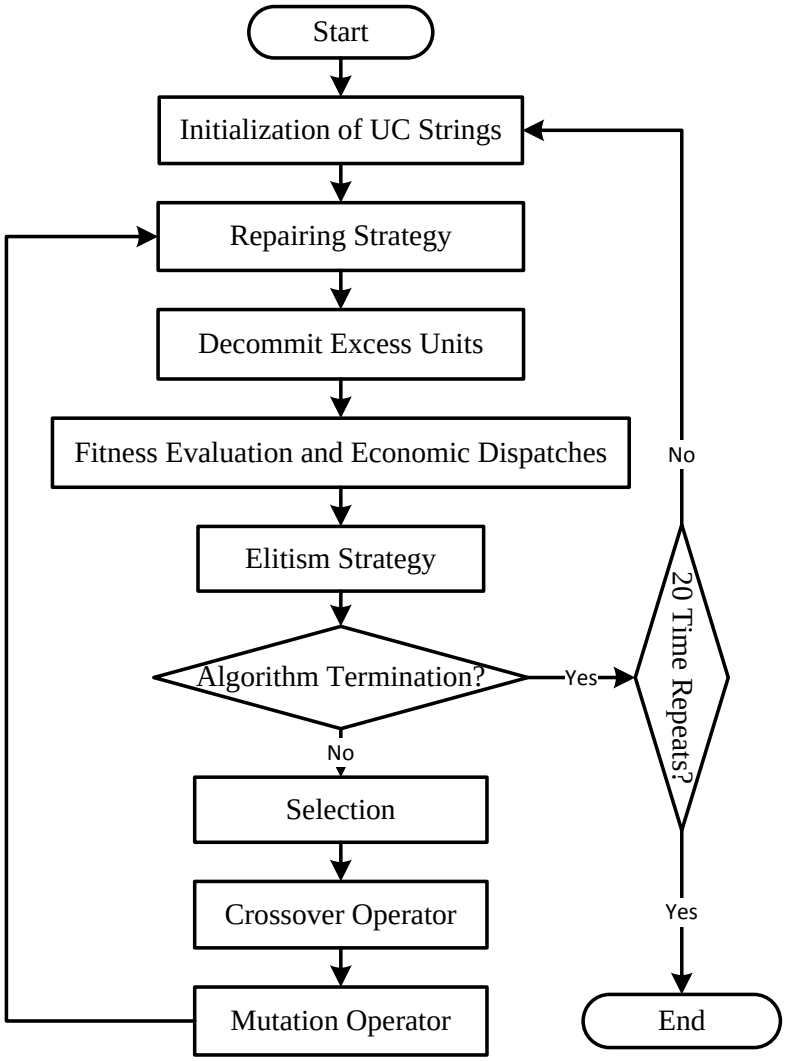

Figure 5.4: Flow chart of the GA-based solution method.

Hour 1	Hour 2	.....	Hour H
1100011101	1101111101 ...	1000110110 ...	1110001011
Unit 1 ... N	Unit 1 ... N	.....	Unit 1 ... N

Figure 5.5: Binary representation of chromosomes.

Net load is also predefined as in Equation (5.12):

$$load_{net}(t) = load(t) - wind\ power(t) \tag{5.12}$$

In PL method, the units are committed in a ascending order of FLAC to meet the net load. For each chromosome, one wind power scenario is chosen to calculate the net load. So the net load is changing with wind power scenarios in different chromosome initializations.

5.5.3 Repairing Strategy for the Minimum Up and Down Time Constraints

During the initialization, the minimum up and down time constraints of generators are not considered. Thus the minimum up and down time constraints may be violated for some chromosomes. Generally speaking, there are two methods to deal with these constraint violations. One method is to add a penalty term to the objective function for the violation. So the offspring with a satisfied constraint have the priority to survive. Another is to repair the chromosomes if the minimum up and down time constraints are not met. After the repairing, the minimum up and down time constraints can be satisfied for every chromosome. In this chapter, a similar repairing strategy [75, 147] for the minimum up and down time constraints is implemented.

5.5.4 Decommitment of the Excess Units

Repairing the minimum up and down time constraints can lead to excessive generations and spinning reserves at some operation hours. This is not desirable from the economic operation perspectives and extra costs may be caused. A heuristic search algorithm in [75, 147] based on a PL is used to decommit some units one by one, in descending order of their FLAC as given by Equation (5.11), until the spinning reserve constraint is just satisfied at any time instant. Please note that before decommitting the excess unit, the minimum up and down time constraints must be checked in advanced. If turning off the generator will make the minimum up and down time constraints unmet, then this generator will not be shut down.

5.5.5 Fitness Evaluation and Economic Dispatches

For every chromosome, the UC decision is fixed. With a fixed UC decision, the λ-iteration [65] is used to solve the ED problem for every wind power scenario. The power balance constraint as in Equation (2.26) can be satisfied through the λ-iteration method. If the ENS or the RNS of spinning reserve requirement as in Equation (2.27) happens, the not served parts will be at the cost of C_{ens} and C_{rns} separately. If the sum of the minimum power of the on units is larger than

the net load, a heavy penalty term is added to the expected cost. The fitness function is defined in Equation (5.13) and is positively oriented, i.e., the greater, the better. A is a system-dependent constant used to avoid getting too small fitness values. Its magnitude should be of the order of the systems maximum operating cost over the scheduling time periods or larger [148].

$$Fitness = \frac{A}{E(X,P)} \tag{5.13}$$

5.5.6 Elitism Strategy

Elitism is a useful strategy frequently used in GA. The idea is to reserve the best one or more populations in previous generation and directly copy them to the next generation without modifications. In this way, the best fitness values will not decrease with iterations. In this chapter, the two fittest solutions are reserved first, and the two worst solutions after the selection are replaced by the two elitism members.

5.5.7 Selection

q-tournament selection is implemented in the GA. The most used value for q is 2. The larger value the q is, the higher the selection pressure becomes in the population. A large q means the whole population will be dominated by the members with high fitness values. There needs a balance to maintain the selection pressure and the diversity of the population.

5.5.8 Crossover Operator

The crossover operator is a two-point crossover. Under a certain crossover probability, two-point crossover calls for two points to be selected on the parent chromosome strings. Everything between the two points is swapped between the parent binary strings, rendering two child chromosomes.

5.5.9 Mutation Operator

Adaptive mutation is chosen. The mutation rate exponentially decreases with the increase of the number of generations. For each mutation, a random number is generated. If the rand number is smaller than the mutation rate, the value on this bit is flipped; otherwise, the value remains the same as before.

5.5.10 Termination

The termination criterion can be set as the reach of the maximum number of iterations or few improvements are made in a certain number of generations.

5.6 Case Studies

Five deterministic and four stochastic cases are implemented and compared. Reserve strategies with and without wind power forecast are investigated.

5.6.1 Datasets

The modified UC test system [63] incorporated with wind power generation uncertainties is used as case studies in this chapter. Please see the appendix for the detailed datasets of this test system. The 24-hour ahead load forecasting is shown in Table 1. Table 2 and 3 show the unit data for generators. The total installed capacity of the thermal generators is 1662 MW, which is 10.8% higher than the peak load (1500MW at Hour 12). The wind power generation uncertainties are considered as scenarios as shown in Fig. 5.3. The capacity of the penetrated wind farm is assumed to be 200 MW. The basic spinning reserve requirement is considered as 10% of the system load. The test systems can be extended up to 100 units by duplicating the units and scaling the load and wind profiles based on the original system.

5.6.2 Wind Power Generation Uncertainties

The wind power generation uncertainties are modeled by a stochastic process using the generated scenarios (W_t^s). Firstly, a list of PIs are constructed using the PSO-based LUBE method. They are further decomposed as quantiles and used to fit an ECDF curve. Then Monte Carlo simulation method is applied to generate the scenarios from the ECDF curve. A detailed description of the scenario generation method can be found from Section III. With the uncertain wind power, the UC decision of the stochastic SCUC problem is solved by the GA in this study.

5.6.3 Deterministic and Stochastic Cases

To investigate the effects of different UC strategies on the economic costs and the ED reserves, five deterministic and four stochastic cases are implemented and compared. Deterministic cases set different forecasts of wind power generations, such as no forecast, point forecast, perfect forecast, the positive 80% quantile and the conservative 20% quantile forecast of wind power. A 10% of load forecast reserve is considered in all deterministic cases. While, stochastic cases set different ED reserve requirements for scenarios. These reserve requirements are explained in detail in next subsection. Table 5.1 summarizes these cases of different UC strategies.

5.6.4 Investigations on Different Reserve Strategies

In order to accommodate wind power uncertainty, several reserve strategies are investigated. The actual reserve requirements need to be set by the power system operators for their preferences. Usually there is a compromise between the economic cost and risks. In S1, a basic reserve requirement of 10% of the load is supposed to accommodate system uncertainties such as the load and wind forecast errors and generator outages. Additional reserve requirement can be added to compensate the extra wind uncertainties not captured by the generated scenarios and the potential emergencies. Thus additional reserves are added to the basic level according to different reserve strategies. These strategies can be considered with and without wind power

Table 5.1: Deterministic and Stochastic UC Strategies

Cases	Descriptions	UC	Forecasts
D1	Det.UC w/no wind	Det.	No
D2	Det.UC w/point forecast	Det.	Point
D3	Det.UC w/perfect forecast	Det.	Perfect
D4	Det.UC w/80% quantile forecast	Det.	Quantile
D5	Det.UC w/20% quantile forecast	Det.	Quantile
S1	Stoch.UC w/regular reserve (10% of load forecast)	Stoch.	Scenarios
S2	Stoch.UC w/additional reserve (additional 5% of load forecast)	Stoch.	Scenarios
S3	Stoch.UC w/additional reserve (50% of point forecast)	Stoch.	Scenarios
S4	Stoch.UC w/additional reserve (point forecast - 10% quantile forecast)	Stoch.	Scenarios

forecast, such as additional 5% of load forecast (S2), additional 50% of point forecast (S3) and additional (point forecast - 10% quantile forecast) reserves (S4). In S2, total 15% of load is considered as reserve. In S3, only half of point forecast is considered to be reliable since an additional reserve of 50% of point forecast is added. In S4, additional wind reserve requirement is set by deducting the 10% quantile forecast from the point forecast. Because the realized wind power will most likely be higher than the 10% quantile forecast (90% of the time). Although we set different reserve strategies at the UC scheduling stage, in the real time ED stage the regular reserve (10% of load forecast) is used for the real operations of the power systems.

5.7 Simulation Results and Discussions

Simulation results of deterministic and stochastic cases are provided and compared together. Generation costs as well as the available reserves are discussed.

5.7.1 Deterministic Cases of Wind Power

The deterministic case can be considered as a special case of the stochastic SCUC model. This special case contains only one scenario with deterministic wind power in each hour. If no forecast is obtained, no wind power is penetrated in the thermal units system. This is a classical study implemented in many references [63,67,75,147]. The deterministic case of no wind power (D1) can simply set the number of scenarios to one and the wind power to zero. Other deterministic cases of point (D2), perfect (D3) and quantile forecast (D4 & D5) of wind power can be conducted in a similar way as D1.

The experimental parameters are listed in Table 5.2. Where, pop_size is the population size, and N_PL is the number of populations using PL initialization method. The maximum number of iterations is 300, the crossover probability is 0.7, and q in the q-tournament is set to 2. The number of elitism members is 2, and β is the initial mutation probability which exponentially decreases with the increasing iterations. The system-dependent constant A is chosen as 600,000. The cost of reserve curtailment is $1,100/MWh, and the cost of unserved energy is $3,500/MWh.

Table 5.2: Parameters for GA

Parameter	**Numerical value**
pop_size	64
N_PL	20
max_iter	200 or 300
$crossover\_prob$	0.7
q	2
$N\_elitism$	2
β	0.2
A	600000
C_{ens}	3500
C_{rns}	1100

Table 5.3: The UC Solutions for the 10-Unit 24-Hour Stochastic SCUC Problem

Hour	Unit Number									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	1	1	0	0	0	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0	0
3	1	1	0	0	1	0	0	0	0	0
4	1	1	0	0	1	0	0	0	0	0
5	1	1	0	1	1	0	0	0	0	0
6	1	1	1	1	1	0	0	0	0	0
7	1	1	1	1	1	0	0	0	0	0
8	1	1	1	1	1	0	0	0	0	0
9	1	1	1	1	1	1	1	0	0	0
10	1	1	1	1	1	1	1	1	0	0
11	1	1	1	1	1	1	1	1	1	0
12	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	0	0
14	1	1	1	1	1	1	1	0	0	0
15	1	1	1	1	1	0	0	0	0	0
16	1	1	1	1	1	0	0	0	0	0
17	1	1	1	1	1	0	0	0	0	0
18	1	1	1	1	1	0	0	0	0	0
19	1	1	1	1	1	0	0	0	0	0
20	1	1	1	1	1	1	1	1	0	0
21	1	1	1	1	1	1	1	0	0	0
22	1	1	0	0	1	1	1	0	0	0
23	1	1	0	0	0	1	0	0	0	0
24	1	1	0	0	0	0	0	0	0	0

The best total cost solved for D1 is $563937.687 and the UC decision is given in Table 5.3. In order to valid the efficiency of the proposed GA to solve the UC problem, the scheduled UC cost of D1 is compared to related results using the same dataset. These benchmark methods include GA [71], LR [71], IBPSO [73], LRPSO [73], DBDE [75], DE [147], ES-EPSO [69], ELR [67] and LRGA [63]. As is clearly shown in Table 5.4, the obtained total cost is among the best one from the relevant literature. If no wind power is considered, this is a classical test system, and the lowest value in Table 5.4 is already the best solution for this test system. Therefore, the effectiveness of the proposed GA-based solution method has been validated.

Table 5.4: Comparison of the best total generation costs without wind power (D1).

Methods	**Generation Cost ($)**
GA [71]	565,825
LR [71]	565,825
IBPSO [73]	563,999
LRPSO [73]	565,870
DBDE [75]	563,977
DE [147]	563,938
ES-EPSO [69]	563,938
ELR [67]	563,977
LRGA [63]	564,800
Proposed GA	563,937.687

Both the scheduled UC generation cost and the real time ED cost are listed in Table 5.5. The scheduled UC generation cost is based on the day-ahead UC decision and the different wind power forecasts. The same UC decision is used in the real time ED cost, however, the real wind power next day (perfect forecast) is applied to real time ED. Table 5.5 lists the two kinds of costs for the deterministic and stochastic cases. The results are based on the best results of the repeated 20 runs using the proposed algorithm.

For the scheduled UC costs as shown in Table 5.5, the 80% quantile forecast case (D4) obtains the least cost in all deterministic cases. The largest cost corresponds to the no forecast case of D1. This is obvious, since the cost of wind generation is regarded as zero in this study. Then the larger proportion of wind power penetration will lead to a less generation cost and vice versa. The order of the generation costs from largest to smallest is $D1 > D5 > D3 > D2 > D4$. D3 and D2 are the perfect and point forecast cases. Their wind power values are moderate so the generation costs also stay at middle levels. D5 obtains the second largest cost because it uses a conservative 20% quantile forecast of wind power. The analysis of the available reserves of the deterministic cases will be discussed later together with the stochastic cases.

Table 5.5: Deterministic and Stochastic UC and ED Costs

Cases	Scheduled UC Costs ($)	Value	Real Time ED Costs ($)
D1	563937.687	Single	542496.106
D2	532707.667	Single	612254.542
D3	538556.091	Single	538556.091
D4	514111.980	Single	873267.106
D5	552152.144	Single	556400.825
S1	538104.445	Expected	542496.106
S2	574268.924	Expected	548266.135
S3	543801.510	Expected	546078.819
S4	550554.515	Expected	547896.795

For the real time ED cost, the perfect forecast (D3) definitely has the lowest cost. This can be clearly seen from the last column of Table 5.5. However, for the optimistic 80% quantile forecast of D4, the RNS occurs (shown in detail later in Subsection 5.7.3). In contrast to the least scheduled UC cost, D4 obtains the largest real time ED cost due to the high penalty from RNS. This is the same reason for the higher costs of D2 and even D5, they both have the RNS. Since RNS does not occur in D1 (no wind case), it obtains the second lowest cost and is only more than the perfect forecast.

5.7.2 Stochastic SCUC with Wind Power Scenarios

The parameters used in GA and the stochastic SCUC problem are shown in Table 5.2. The only difference is that the maximum number of iterations is 200 instead of 300 in the stochastic SCUC cases to save the computation time. The number of scenarios is 50 as shown in Fig. 5.3.

Fig. 5.6 shows the best and average fitness values during optimization for S1 with a 10% reserve requirement. Due to the elitism strategy, the best fitness continuously increases with the iterations. As can be seen from Fig. 5.6, the initial best fitness is relatively high (lager than

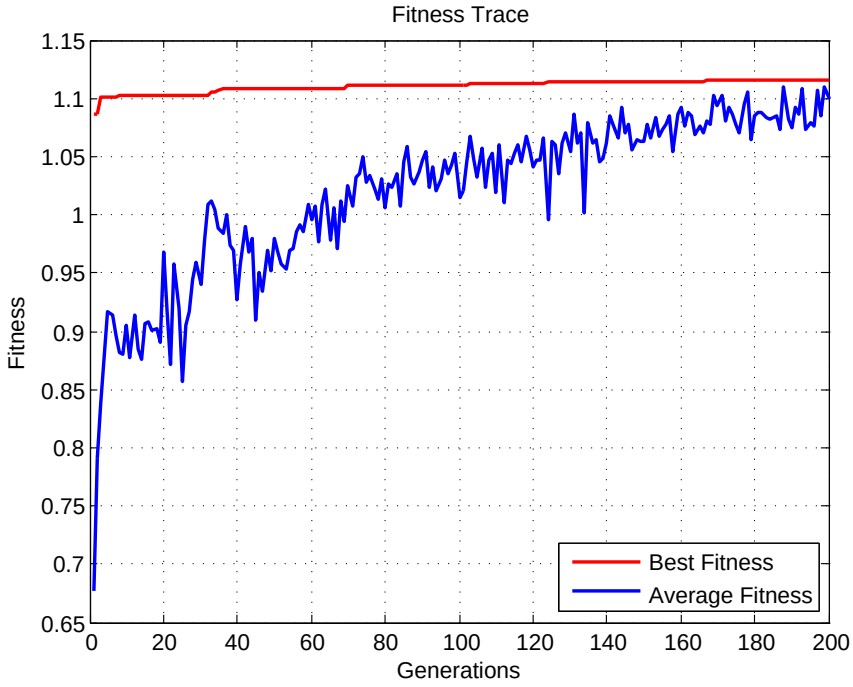

Figure 5.6: Fitness trend during optimization for S1.

1.05), this is due to the good initialization using the PL method. This indicates the high quality of seeds used at the start of the algorithm, which will lead to a high success rate for optimal or near optimal solutions.

The system scheduled expected cost of S1 is \$538104.445 for the stochastic SCUC. The UC decision is the same as in Table 5.3. Compared with the no wind power case (D1), the expected cost reduction due to the penetration of wind power generation is $\$563937.687 - \$538104.445 = \$25833.242$. Different from the deterministic cases, the scheduled UC costs and the real time ED costs of S1-S4 remain at the same level. The higher the reserve requirement is, the more the generation cost will be. S1 has only the regular reserve requirement (10% of load), it has the lowest values for the two kinds of costs. The order of the scheduled UC costs and real time ED costs from the lowest to the highest is $S1 < S3 < S4 < S2$. The system costs of S2 are the largest due to a highest reserve requirement (totally 15% of load). In S2, most generators need to be turned on to maintain the highest spinning reserve, which can be clearly seen from Fig. 5.7. At every hour, the number of units turned on of S2 is largest. Moreover, due to the high penalty

of RNS on some scenarios, the difference between the scheduled UC cost and the real time ED cost is also the largest in S2 out of all stochastic cases.

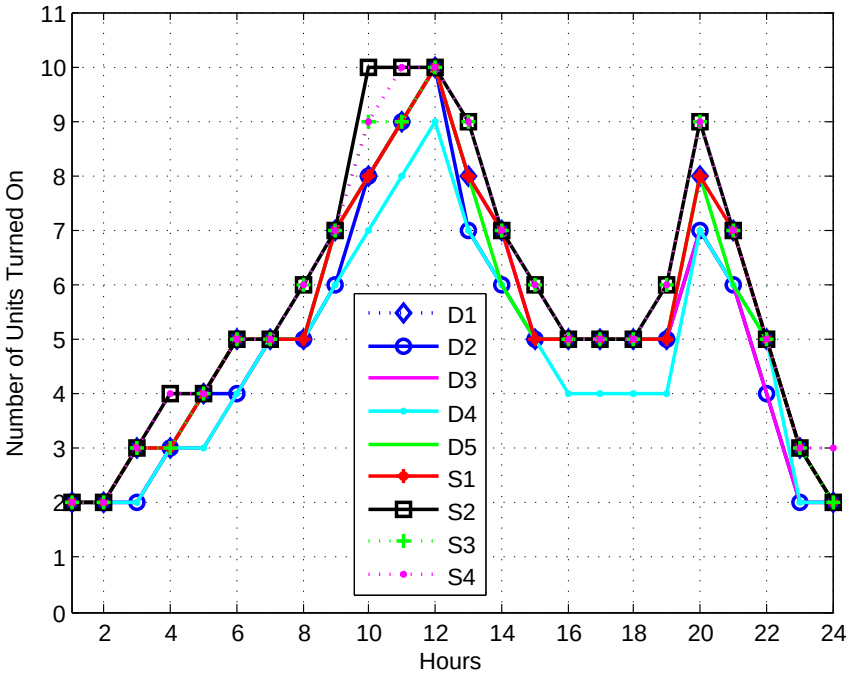

Figure 5.7: The number of units turned on.

5.7.3 Discussion on Available Reserves

Two kinds of available reserves, the scheduled reserve and the real time ED reserve, are investigated and compared together for different UC strategies. The scheduled available reserve is based on the day-ahead UC decision and forecasted wind power generation. The same UC decision but the real wind power next day is applied to real time ED. The purpose of doing so is to show the difference between the planned and realized reserves, and further to investigate the potential risks for different UC strategies.

5.7.3.1 Number of Units Turned on

It can be observed from Fig. 5.7 that the stochastic cases of S2 and S4 have the most generators turned on in each hour. S2 has the most on-line generators to maintain a high reserve requirement

of 15%. D4 turns least generators on using a positive 80% quantile forecast of wind power. On the other hand, D5 has to turn on more generators due to a conservative 20% quantile forecast.

5.7.3.2 Scheduled UC Reserves

As shown in Fig. 5.8 and 5.9, almost all the UC strategies can meet the scheduled reserve requirement. The objective function in Equation (2.23) gives heavy penalties on the ENS and

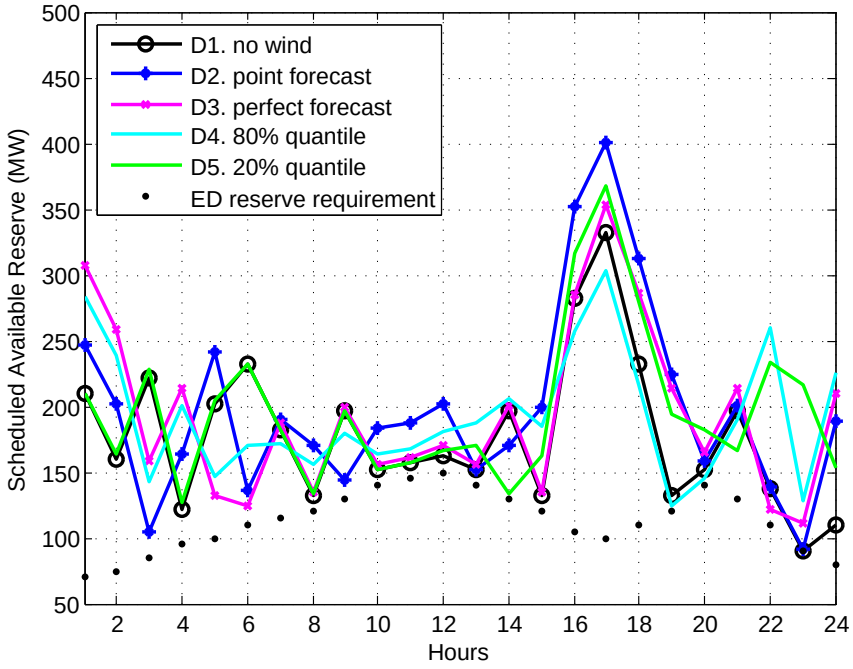

Figure 5.8: Scheduled available reserve of Det. cases.

RNS, so all deterministic and stochastic cases schedule enough generators on-line to avoid ENS and RNS occurring. In Fig. 5.8, all the deterministic cases of D1-5 can meet the scheduled 10% reserve requirement. In Fig. 5.9, all the stochastic cases of S1-S4 can meet the scheduled reserve requirement respectively. However, there is only one exception of S2 (15% reserve of load) at Hour 12. The RNS of 18.68MW (expected value) occurs at the peak load, which is shown in the pink rectangle in Fig. 5.9. This is due to the fact that the total installed capacity of the thermal generators is only 10.8% higher than the peak load. It also implies that wind power is not reliable to contribute to the system reserve. Therefore, the scheduled UC cost of S2 is the highest among

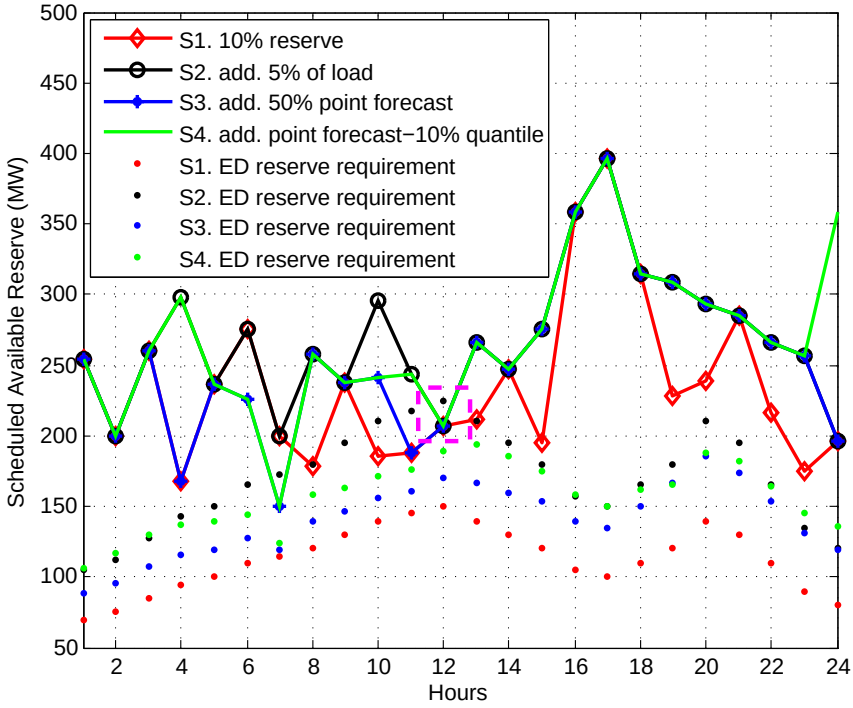

Figure 5.9: Scheduled available reserve of Stoch. cases.

all the stochastic cases as indicated in Table 5.5.

5.7.3.3 Real Time ED Reserves

There are some differences between the planned and the realized reserves with the uncertain wind power. As per Fig. 5.10, RNS still occurs in three real time ED cases. D4, D2 and D5 have RNS during Hours 8-20, although the ENS does not happen. D4 has the most hours (8-15, 19-20) of RNS due to the optimistically deterministic forecast (80% quantile) of wind power. As a result of high uncertainties, wind power will be under or over estimated. It is especially risky if the forecasted wind power is highly over estimated. According to Fig. 5.1, the real wind power is obviously lower than the point forecasts during Hours 6-19. This is the direct cause for RNS. If the penetration level of wind power is higher in this system, even ENS may happen. This is risky for the operation of the system and needs to be avoided. Another interesting finding is that most of the RNS occurs at the peak load hours such as Hours 9-14 and 19-21. The overall reserve at the peak load hours is obviously lower than non-peak load hours. This implies that the peak load

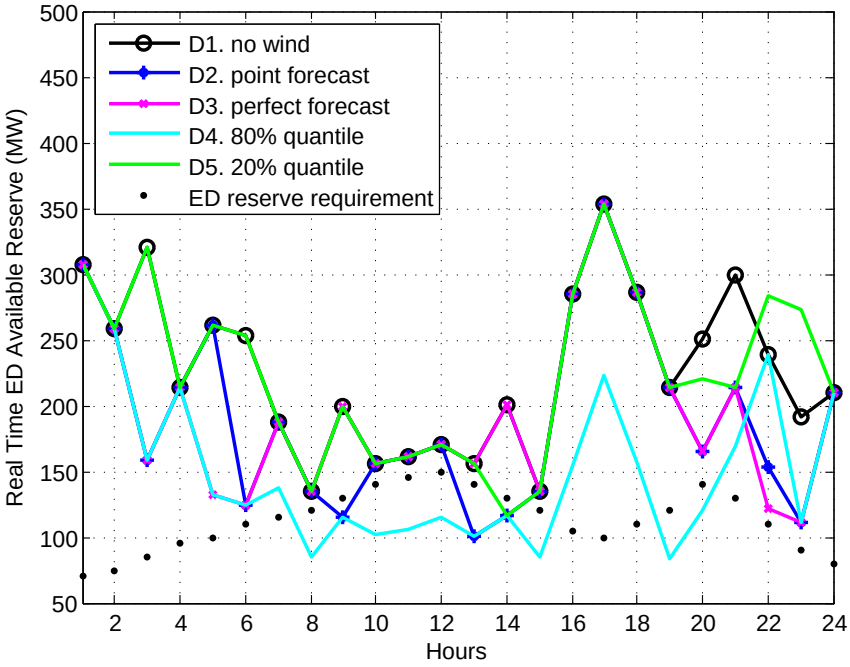

Figure 5.10: Real time ED available reserve of Det. cases.

hours run higher level of risk (uncertainties) for system operations.

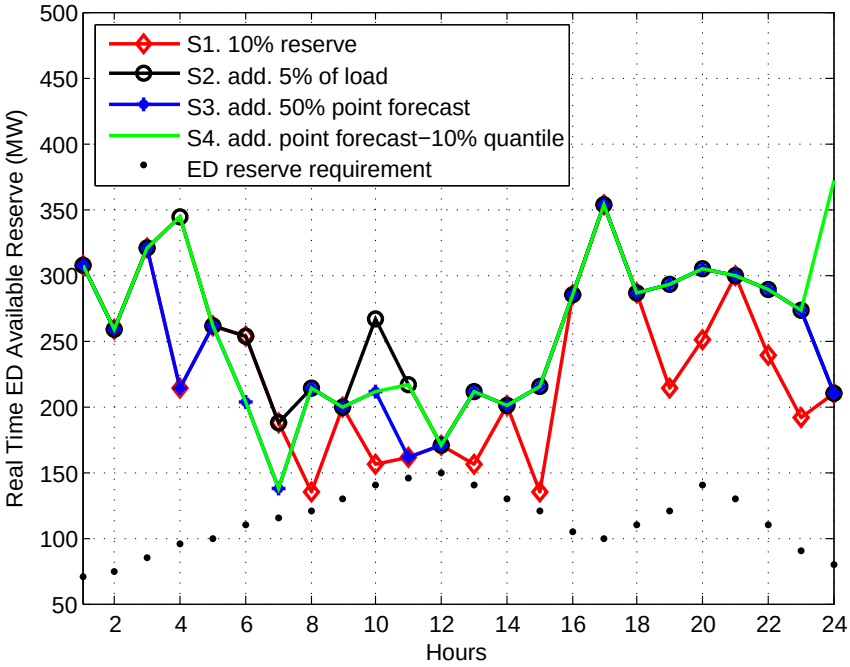

Figure 5.11: Real time ED available reserve of Stoch. cases.

On the contrary, all the four stochastic cases (with different reserve strategies) can always

meet the reserve requirement in real time ED in Fig. 5.11. S2 and S4 have overall higher reserve levels than S1 and S3. This is because the higher reserve requirements of S2 and S4. All the four lines of S1-S4 are always above the 10% reserve requirement in Fig. 5.11. However, the RNS occurs in three out of five deterministic case studies. It needs to be noticed that the conservative 20% quantile forecast in D5 has RNS in Hour 14. This case tells us that even if the conservative deterministic forecast is applied, the obtained UC scheduling still runs the risk and is unreliable in system operations. Thus the stochastic model is more robust than deterministic ones.

5.8 Conclusions

A nonparametric and recently proposed PSO-based LUBE method is used to construct NN-based PIs. Wind power forecast uncertainties are represented by a list of PIs with the coverage probabilities ranging from 5% to 95%. A new scenario generation method is proposed. It can be easily implemented and avoids making strong assumptions on the wind power probabilistic distributions. This method also builds an important bridge between the PIs and the scenarios used in the stochastic model. Wind generation uncertainties are considered as a stochastic process through the generated scenarios. Further a stochastic SCUC model is built and the wind scenarios are incorporated into this model. Five deterministic and four stochastic case studies with uncertain wind power are implemented. GA is used to solve the stochastic SCUC problem. The generation costs for different UC strategies are discussed and compared together. Specially two available reserves, the scheduled and real time ED reserves, are discussed. Results indicate that the stochastic model is more robust than deterministic ones. The peak load hours run higher level of risk (RNS and ENS) than non-peak hours for system operations. What is more, there are some differences between the planned and real time ED reserves. The potential risks need to be considered by system operators.

Chapter 6

A Computational Framework for Uncertainty Integration with Intermittent Renewable Energy Sources

6.1 Introduction

Renewable energy resources such as wind and solar power have several major benefits including low economic costs and zero environmental footprints. That is why their deployment has sustained a high growth rate in many countries worldwide [4]. However, the level of uncertainties associated with operation of intermittent renewable energy sources (IRESs) is high. In addition, uncertainties do exist in other parts and components of power systems including but not limited to consumers (load demands), generators (shutdowns), and transmission lines (faults and leakages).

Uncertainties in forecasting the output of IRESs such as wind and solar generation, as well as system loads, are not incorporated into existing energy management systems (EMSs) and tools

used for generation commitment, dispatch, and market operation. With the growing penetration of intermittent resources, these uncertainties could result in significant unexpected load following and dispatch problems and pose serious risks to control and operation performance characteristics as well as the grid reliability. Without quantifying prevailing risks, system operators have limited means to assess the likelihood of occurrence and the magnitude of problems to mitigate imposed adverse impacts [149]. As per these, there is a real business need to develop and deploy a computational framework to integrate these uncertainties together and mitigate the potential risks for system operations.

In [150], a fuzzy-optimization approach was proposed for solving the generation scheduling problem with consideration of wind and solar energy systems. Hourly load, available water, wind speed and solar radiation forecast errors were taken into account using fuzzy sets. A methodology to set operating reserve considering load forecast uncertainty, conventional generation outages and wind power forecast uncertainty was described in [34]. Makarov in [2] proposed an approach to integrate the uncertainty model with an existing EMS. Different levels of integration such as passive, active and proactive are presented. However, only load and wind power forecasting uncertainties are considered and the solar power uncertainty is not considered in simulations. A stochastic programming framework [151] was built as a multi-objective problem. Different sources of uncertainties were considered for optimal operation of micro-grids. In [152], intermittent wind units were integrated into a GENCO's generation assets and the GENCO's hourly wind generation schedule was coordinated with that of natural gas units and hydro units for maximizing the GENCO's payoff.

Previous studies mainly focus on one or two aspects of the uncertainties, such as the load and/or wind power forecast uncertainties [49, 78, 79, 83, 84]. It is very important to address this problem comprehensively by including all sources of uncertainty (load, wind and solar power generation, forced generator outages, etc.). In [138, 146], the influence of wind power forecast uncertainties on generation costs and different reserves have been investigated. In order to fo-

cus on wind power uncertainties, other uncertainties such as load and forced generator outages were temporarily ignored. In this paper, a computational framework is proposed to integrate the various uncertainties together. A preliminary research has been implemented to investigate the framework for uncertainty integration.

The purpose of this research is multi-fold. 1) Previous studies have not considered adequate uncertainties as part of the scheduling. 2) This integration framework quantifies all grid uncertainties and then integrates them into the decision-making and risk assessment. 3) Several uncertainty modeling methods are proposed to address uncertainties from different sources. 4) A new scenario generation method is proposed to generate scenarios for IRESs such as wind and solar power. 5) The integration framework is validated on the stochastic unit commitment (UC) scheduling problem. 6) Generation costs as well as the available reserves of different UC strategies are investigated and compared together.

The rest of this chapter is organized as follows. The proposed framework for uncertainty integration in DPSs is introduced in Section 6.2. Uncertainty representation, problem formulation of the stochastic SCUC problem and the solution method are described in Section 6.3. Section 6.4 describes the datasets and simulation cases. Simulation results and discussions are provided in Section 6.5. Finally, Section 6.6 concludes the chapter with some remarks for further study in this domain.

6.2 Proposed Computational Framework for Uncertainty Integration in DPSs with IRESs

The Section 1.4 in Chapter 1 has provided a comprehensive summary of uncertainty in DPSs with IRES. As shown in Fig. 1.2, DPSs operations are mainly subject to four types of uncertainties: generation uncertainties, transmission uncertainties, distribution uncertainties and departures from forecasts. In this chapter, a computational framework for uncertainty integration

in DPSs with IRESs is proposed. The integrated uncertainties include load forecast uncertainty, wind and solar power forecast uncertainty and forced generator outages etc. Integrating all of the uncertainties in DPSs is out of the scope of this study. Fig. 6.1 shows the proposed computational framework for uncertainty integration in DPSs with IRESs.

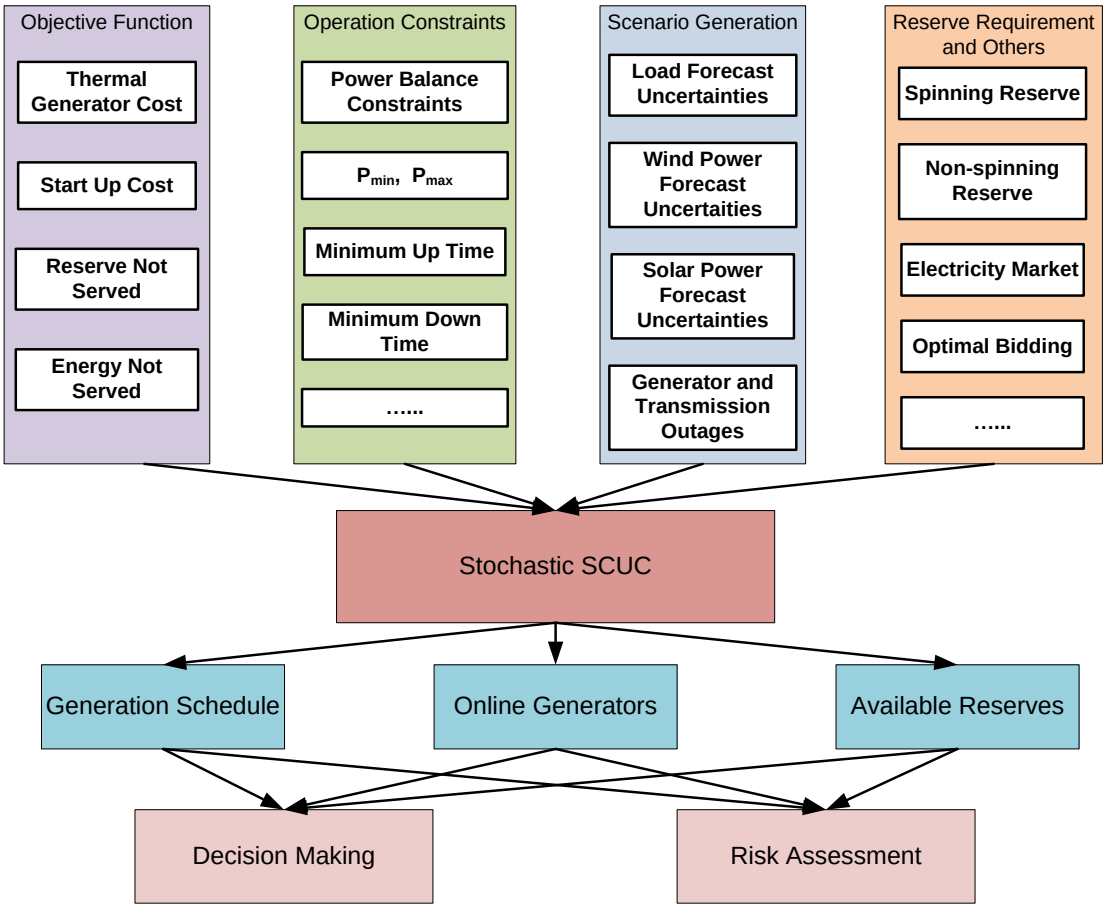

Figure 6.1: The computational framework for uncertainty integration in DPSs with IRESs.

6.2.1 Inputs of the Integration Framework

Objective setting and data gathering are the first stage of the proposed framework. The inputs of the integration framework consist of mainly four parts: the objective function, the operation constraints, scenario generation, reserve requirement and other input data.

The objective function includes the thermal generator cost, the start up cost and penalty terms from the reserve and energy not served. So the objective is to minimize the cost and

mitigate the operation risk. The objective function is subject to several constraints such as the power balance constraint, the generating limitations of thermal generators, the minimum up/down time constraints and reserve requirement. The load forecast uncertainty, wind and solar power forecast uncertainty and forced generator outages are represented as the generated scenarios. The detailed scenario generation will be introduced in next section. Reserve requirement includes the spinning and non-spinning reserve requirements. Other input information such as the electricity market and optimal bidding of generation (thermal, wind and solar) can also be covered.

6.2.2 Stochastic SCUC Modeling

After setting the objective function and gathering the input data, a stochastic stochastic security-constrained unit commitment (SCUC) model is built. The stochastic SCUC model tries to cover as many as major uncertainty sources in DPSs. In particular, forecasting uncertainties for load demand and solar and wind power generation are considered. These uncertainties are modeled as a stochastic process and presented as scenarios. Other uncertainties can be integrated in a similar way. The objective of stochastic SCUC is to minimize the expected generation costs under several constraints for several scenarios. Therefore, it evaluates the overall costs considering all scenarios and their probabilities (p_s). The detailed problem formulation of the stochastic SCUC can be found in Subsection 6.3.2.

6.2.3 Outputs of the Integration Framework

The main purpose of UC is to determine the on/off status and output power of each generator. After solving the stochastic SCUC and economic dispatches, the outputs are obtained. These are the generation schedule, the online units and the available reserve. These outputs are important evidence for decision making and risk assessment in the operation, economics and reliability of DPSs.

6.2.4 Decision Making and Risk Assessment

The decision making and risk assessment are based on the outputs of the stochastic SCUC. The power system operators can schedule the generation next day and set the different reserve strategies according to their preferences. That is to say, the system operators need to decide which generators are turned on next day, and to decide how much power each turned on generator will supply. On the other hand, system operators should maintain a certain level of spinning and non-spinning reserves. The potential risks such as the not-served reserve and the not-served energy need to be handled by the system operator as well.

6.3 Stochastic Model for Uncertainty Integration: Load, Wind, Solar, and Generator Outage Uncertainties

In this section, uncertainty representation for different sources of uncertainty is introduced. The problem formulation and the GA-based solution method are then provided.

6.3.1 Uncertainty Representation

Load forecast uncertainty is assumed to follow a normal distribution. Wind and solar forecast uncertainties are represented as scenarios using the proposed scenario generation method in Chapter 5. Generator outages uncertainty is modeled as discrete scenarios.

6.3.1.1 Load Uncertainty Representation

Load uncertainties are much less than uncertainties associated with IRESs. This is due to the fact that aggregated load demands are fairly smooth and predictable. However, the MW value of load forecasting error is still high. A common method is to model the load uncertainty as the normal distribution [27,34]. In this chapter, load forecast errors are assumed to follow the normal distribution. In addition to the point forecast, standard deviations in each hour are also provided.

Load uncertainty is then represented as stochastic scenarios which are generated using the point forecasts and the corresponding standard deviations (SD).

Table 6.1: Load SD Data for 24 Hours

Hour	Load SD (MW)	Hour	Load SD (MW)	Hour	Load SD (MW)
1	19.221	**9**	25.027	**17**	26.920
2	17.783	**10**	26.514	**18**	31.295
3	18.576	**11**	26.783	**19**	28.534
4	17.819	**12**	36.938	**20**	21.120
5	17.247	**13**	23.636	**21**	20.366
6	17.106	**14**	25.720	**22**	22.101
7	19.942	**15**	25.096	**23**	20.389
8	19.934	**16**	24.428	**24**	20.367

The demand point forecasts are from a classic UC test system as reported in [67,71]. Hourly point forecasts of load demand and their standard deviation [153] are shown in Table 1 (in the appendix) and Table 6.1 respectively. The load scenarios are then generated from this normal distribution at each hour, which are shown in Fig. 6.2. The mean absolute percentage error (MAPE) of the generated load scenarios is around 2%-3%, which is in the common range of day-ahead load forecasting.

6.3.1.2 Wind Uncertainty Representation

As mentioned in Chapter 5, wind power forecasting uncertainty is represented as a series of scenarios. This has been realized by two steps as shown in Section 5.2 and Section 5.3. Firstly, a list of PIs for wind power forecasting is implemented with confidence levels ranging from 5% to 95% (5% increment). Then wind power scenarios are generated by a newly proposed scenario generation method. To avoid the repetition, wind uncertainty representation is omitted in this

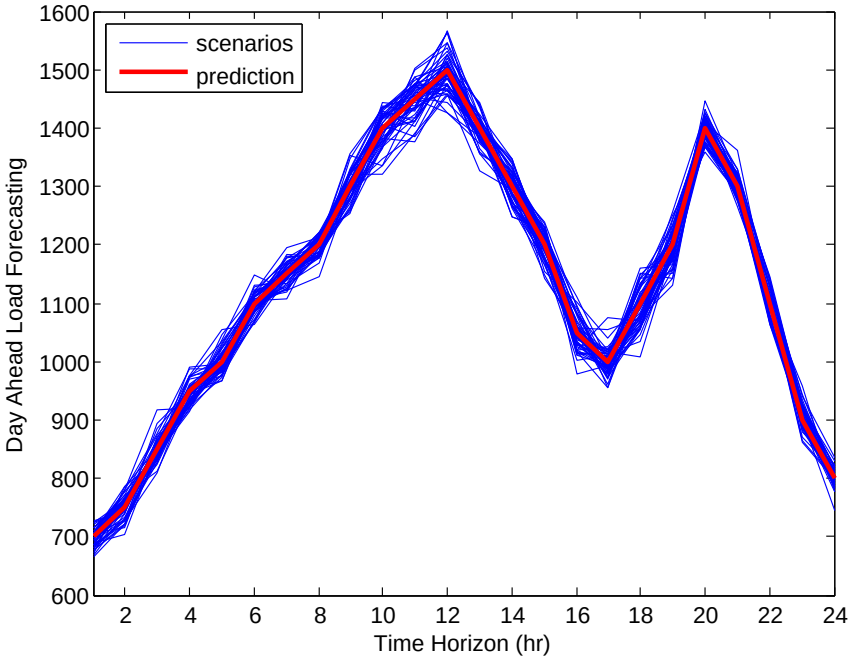

Figure 6.2: The generated 50 load scenarios for 24 hours.

chapter. A detailed implementation can be found from Section 5.2 and Section 5.3 in Chapter 5.

6.3.1.3 Solar Uncertainty Representation

The same uncertainty representation method of wind power is applied to solar power as well. Since wind and solar power are all IRESs, the proposed uncertainty modeling methodology can be used as the common modeling method for IRESs.

1) **A list of PIs for solar power forecasting:** as shown in Section 5.3 in Chapter 5, a list of wind power PIs are generated. The method is similarly applied to solar power. A list of solar power PIs are generated ranging from 5% to 95% with an increment of 5% for uncertainty handling. PIs of 19 levels are constructed for one-day ahead solar power generation forecasting. These PIs can be constructed by our previously proposed PSO-based lower upper bound estimation (LUBE) method [43, 60] or ARIMA models [136]. An example from ARIMA is shown in Fig. 6.3. The solar power datasets are from Solar Radiation Monitoring Laboratory station: Ashland, University of Oregon in USA [154]. The forecasting periods are the 24 hours on

July. 16, 2013. All the predicted and measured solar power values are normalized by the capacity of the PV modules at Ashland station (15KW).

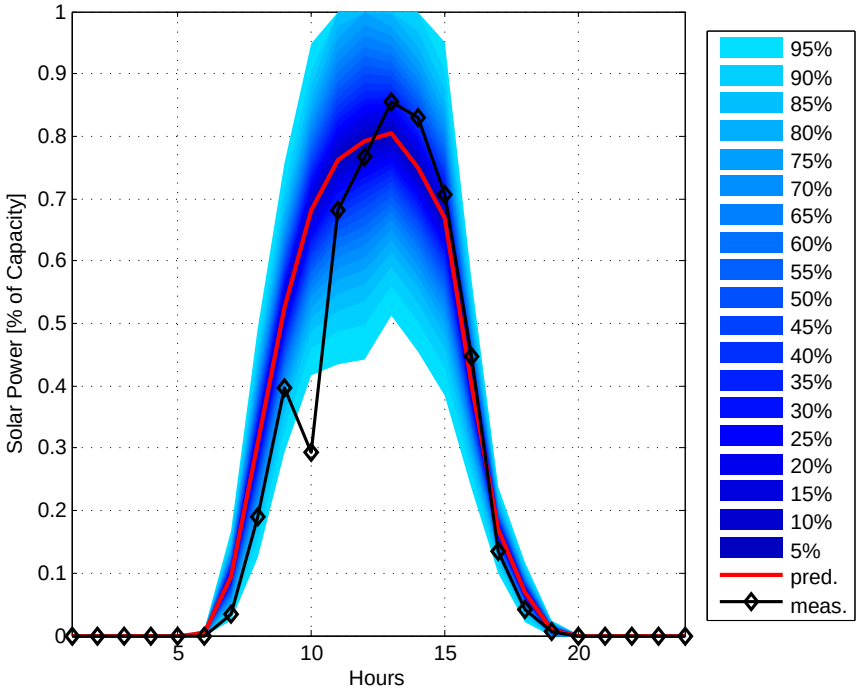

Figure 6.3: A list of PIs for day-ahead solar power forecasting for Ashland PV modules on July 16, 2013.

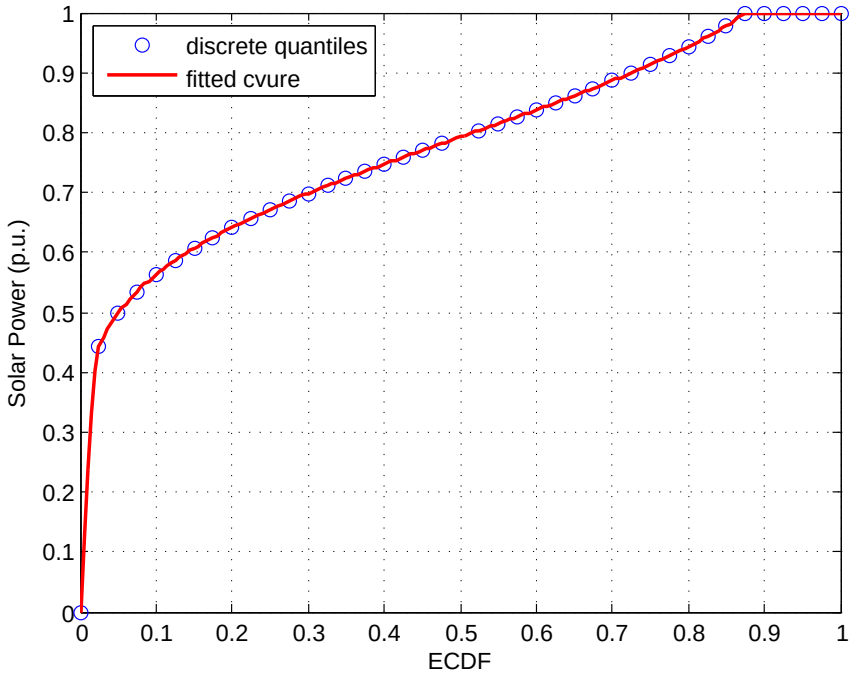

Figure 6.4: The fitted solar power ECDF curve of Hour 12 on July 16, 2013.

2) **The fitted ECDF curve for solar power:** once a list of PIs for solar power forecasting are obtained, these PIs are then decomposed into quantiles. For each hour, every quantile value corresponds to one point on the ECDF curve. Like the wind power in Section 5.3, totally $19 * 2 + 2 = 40$ points are used to fit the ECDF curve. Fig. 6.4 shows the fitted ECDF curve of Hour 12 on July 16, 2013 for the Ashland solar station. The ECDF will be used to generated solar power scenarios for the stochastic model.

3) **The generated 50 scenarios for solar power:** After fitting the ECDF curve, Monte Carlo simulation [11] is applied to generate scenarios from the ECDF. That is, for each scenario, a random number between 0 and 1 is uniformly generated. The corresponding solar power on the ECDF curve (per unit) is the solar power output at this scenario. This process is repeated 24 times for the 24 hours of the day. The ECDF on each hour is fitted individually and the corresponding scenarios are generated from the ECDF of this hour. Fig. 6.5 shows the generated 50 scenarios for 24 hours on July 16, 2013 for the Ashland solar station.

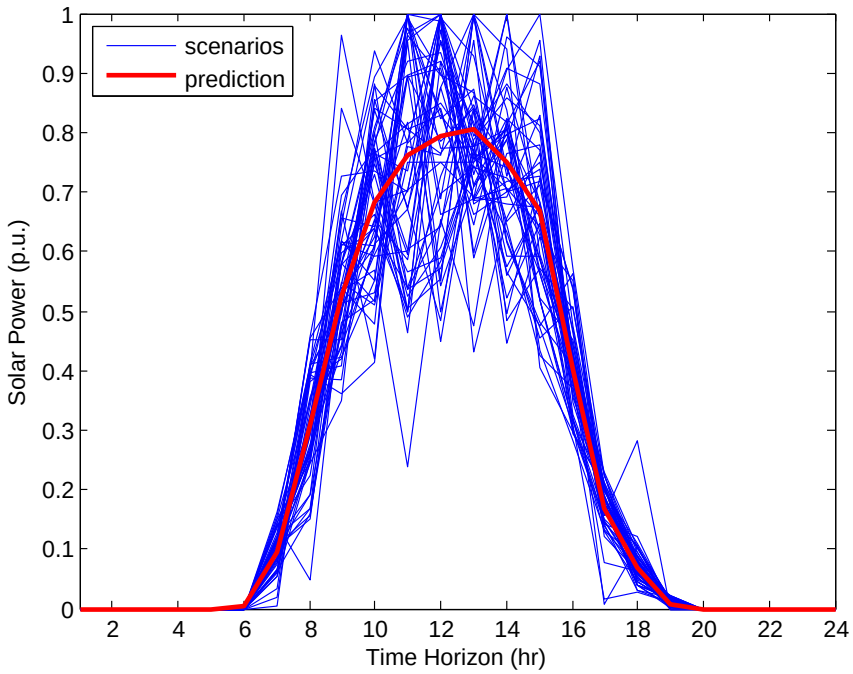

Figure 6.5: The generated 50 solar scenarios for 24 hours on July 16, 2013.

6.3.1.4 Generator Outage Uncertainty Representation

Conventionally generator outage is considered as a discrete probability distribution and represented by the capacity outage probability table (COPT) [34, 65]. COPT gives the probability of occurrence for each possible outage capacity level. In order to simplify the integration framework and be compatible with the net load concept, the similar method in [155] is implemented to represent the generator outage uncertainty. The implementation steps are as follows:

1) 5 out of 10 generators are considered for generator outages. That is, Generators 6-10 (intermediate and peaking units [156]) in the 10-unit 24-hour test systems are considered for the outage events (Table 3).

2) Generator outage is modeled by creating a scenario which has demand increases equal to the unavailable generator's capacity. These demand increases occur during the periods in which the generator is expected to be down [155]. Therefore, if one generator outage occurs, the net load during the outage periods is added by the unavailable generator's capacity. If two or more outages occur together for some hours, then the unavailable generators' capacities are added together to the net load for these hours.

3) The assumption for the equivalent outage duration is 8 hours per day.

4) The assumed forced outage rate is 2%. Please note that the outage rate and duration can be adjusted in the real operation according to the historical outage statistics for different generators.

According to the above steps, the generated scenarios for generator outages are shown in Table 6.2. These scenarios are generated for the 50 scenarios of the Generators 6-10 in Table 3.

6.3.2 Problem Formulation

The problem formulation of the proposed uncertainty integration framework is similar to the formulation of the stochastic SCUC problem in Section 5.4 in Chapter 5. The main difference is

Table 6.2: The Generated Scenarios for Generator Outages

Scenario	Outage Generator No.	Outage Hours
7	8	6-13
16	6	10-17
16	9	7-14
28	10	2-9
44	6	16-23

the power balance constraint, which is shown in Equation (6.5). Compared to the only wind case in Section 5.4, the load, wind and solar power forecasting uncertainties and the generator outages are all considered together in the integration framework. Then the net load has a new definition:

$$load_{net}(t) \; = \; load(t) \; - \; wind\, power(t) \; - \; solar\, power(t) \tag{6.1}$$

The objective of the stochastic SCUC problem is to minimize the expected generation costs under several constrains.The objective function consists of four parts: the thermal generator costs, the start up cost, and the penalty terms from the energy not served (ENS) and the reserve not served (RNS). Mathematically, the objective function is formulated as follows [63, 138]:

$$\begin{aligned} &E(X,P) = \\ &\min \sum_{s=1}^{S} p_s \Big[\sum_{i=1}^{N} \sum_{t=1}^{H} F_i(P_{i,t}^s) X_{i,t} + \sum_{t=1}^{H} (C_{ens} * ENS_t^s \\ &+ C_{rns} * RNS_t^s) \Big] + \sum_{i=1}^{N} \sum_{t=1}^{H} \left[SU_{i,t}(1 - X_{i,(t-1)}) \right] X_{i,t} \end{aligned} \tag{6.2}$$

where thermal generator cost $F_i(P_{i,t}^s)$ is usually represented as a quadratic function:

$$F_i(P_{i,t}^s) = a_i + b_i P_{i,t}^s + c_i (P_{i,t}^s)^2 \tag{6.3}$$

RNS and ENS will be at the cost of C_{ens} and C_{rns} in the objective function. The start up

cost is determined according to the continuously off time of generators:

$$SU_{i,t} = \begin{cases} HSU_{i,t}, & if\ T_{i,t}^{off} \leq T_i^{Down} + T_i^{cold}; \\ CSU_{i,t}, & if\ T_{i,t}^{off} > T_i^{Down} + T_i^{cold}. \end{cases} \tag{6.4}$$

The following constraints are also defined:

1) Power balance constraints:

$$\sum_{i=1}^{N} X_{i,t} P_{i,t}^{s} + W_t^s + PV_t^s = D_t^s + FO_t^s - ENS_t^s \tag{6.5}$$

The D_t^s, W_t^s and PV_t^s are the load, wind and solar power outputs at time t in scenario s respectively. FO_t^s and ENS_t^s are the forced outage of generators and ENS at time t in scenario s.

2) Spinning reserve constraints:

$$\sum_{i=1}^{N} X_{i,t} \left[P_{i,max} - P_{i,t}^{s} \right] \geq R_t^s - RNS_t^s \tag{6.6}$$

3) Generation limit constraints:

$$P_{i,min} X_{i,t} \leq P_{i,t}^{s} \leq P_{i,max} X_{i,t} \tag{6.7}$$

4) Minimum up time constraints:

$$(T_{i,t}^{on} - T_i^{Up})(X_{i,(t-1)} - X_{i,t}) \geq 0 \tag{6.8}$$

where,

$$T_{i,t}^{on} = (T_{i,(t-1)}^{on} + 1) X_{i,t} \tag{6.9}$$

5) Minimum down time constraints:

$$(T_{i,t}^{off} - T_i^{Down})(X_{i,t} - X_{i,(t-1)}) \geq 0 \tag{6.10}$$

where,

$$T_{i,t}^{off} = (T_{i,(t-1)}^{off} + 1)(1 - X_{i,t}) \tag{6.11}$$

6.3.3 GA-Based Solution Method

The proposed integration framework of the stochastic SCUC problem is solved using genetic algorithm (GA). The solution procedure has been introduced in Section 5.5 in Chapter 5. Each UC solution is represented by the binary chromosome. The binary values of 0 and 1 stand for the off and on status of each generator at one hour. So the length of the binary string equals to the number of generators multiplied by the number of scheduled hours.

As indicated in Equation (6.1), both the uncertain load, wind and solar power are involved in the net load. The generator outages also cause demand increases which equal to the unavailable generators' capacity. These are not the same to the only wind case in Chapter 5. These changes need to be reflected by the fitness evaluation and economic dispatches in the GA-based solution method. That is, the λ iteration is updated to satisfy the new power balance constraint as shown in Equation (6.5). When the fitness function is evaluated, all the modeled uncertainties, such as the load, wind, solar and generator outage uncertainties are calculated. These uncertainties have influences on thermal generator costs, the start up costs and the penalty terms from the ENS and RNS. To make this chapter self-contained, a brief description of the GA-based solution method is provided. The detailed implementation of the solution method can be found from Section 5.5.

6.4 Datasets and Simulation Cases

The modified test system in Chapter 5 is used in this chapter. The load forecast and unit data are shown in Table 1-3 respectively. In Chapter 5, only wind uncertainty is considered. But in the integration framework, load, wind, solar and generator outages uncertainties are considered together and represented as scenarios in the stochastic SCUC model. The capacity of the penetrated wind and solar power are assumed to be 200 MW and 100MW respectively. Therefore, the total capacity of the IRESs is 300 MW, which is 1/5 of the peak load.

As the five deterministic and four stochastic cases shown in Table 5.1 in Chapter 5, the same

concept is applied to the integration framework. However, the wind and solar power forecasting are summed up together in the five deterministic and four stochastic cases. Taking the point forecast case (D2) as an example, the point forecast value in the integration framework is the summation of the wind and solar power point forecasts. The summation of the wind and solar power forecasting are also applied for setting different reserve strategies in the four stochastic cases. For each simulation case, the algorithm is run for 10 repeated times and the best results are reported.

6.5 Simulation Results and Discussions

Results of the deterministic and stochastic cases and the available reserves are provided and discussed in this section. The experimental parameters are listed in Table 6.3. Where, pop_size is

Table 6.3: Parameters Used in the Integration Framework

Parameter	**Numerical value**
pop_size	64
N_PL	20
max_iter	200 or 300
$crossover\_prob$	0.7
q	2
$N\_elitism$	2
β	0.2
A	600000
C_{ens}	3500
C_{rns}	1100

the population size, and N_PL is the number of populations using PL initialization method. The maximum numbers of iterations are 300 and 200 for deterministic and stochastic cases respectively. The crossover probability is 0.7, and the q in the q-tournament is set to 2. The number of elitism members is 2, and β is the initial mutation probability which exponentially decreases

with the increasing iterations. The system-dependent constant A is chosen as 600,000. The cost of reserve curtailment is $1,100/MWh, and the cost of unserved energy is $3,500/MWh.

6.5.1 Deterministic Cases of Wind and Solar Power

D1 is the no forecast case, this means the penetrated wind and solar power is 0. In Chapter 5, to demonstrate the effectiveness and efficiency of the proposed GA-based solution method, Section 5.7 has compared the results with several solution methods. These benchmark methods include GA [71], LR [71], IBPSO [73], LRPSO [73], DBDE [75], DE [147], ES-EPSO [69], ELR [67] and LRGA [63]. The comparison indicates that the solution quality of D1 is among the best one from the relevant literature.

Table 6.4: Deterministic and Stochastic UC and ED Costs in the Integration Framework

Cases	**Scheduled UC Costs ($)**	**Value**	**Real Time ED Costs ($)**
D1	563937.687	Single	531586.520
D2	516074.046	Single	646367.876
D3	523810.166	Single	523810.166
D4	495996.705	Single	924968.863
D5	538420.600	Single	536318.881
S1	528092.051	Expected	531626.164
S2	539477.652	Expected	538873.934
S3	535792.825	Expected	537290.562
S4	536678.455	Expected	538374.976

For the scheduled UC cost in Table 6.4, the overall UC cost is less due to the solar power penetration compared to the only wind case in Chapter 5. The larger wind and solar power penetration will lead to a less (unit) generation cost. Case D1 has no wind and solar power penetration, so the generation cost of D1 is the largest. D5 uses the conservative 20% quantile forecast of wind and solar power, and then D5 obtains the second largest costs. D4 uses the

optimistic 80% quantile forecast of wind and solar power, and then D4 obtains the least costs. For the case D2 (point forecast) and D3 (perfect forecast), the point forecast of wind and solar power is overall more than the perfect forecast, so cost of D2 is less than the cost of D3.

For the real time ED cost, the perfect forecast (D3) definitely has the lowest cost. This can be clearly seen from the last column of Table 6.4. However, for the optimistic 80% quantile forecast of D4, the RNS occurs. In contrast to the least scheduled UC cost, D4 obtains the largest real time ED cost due to the high penalty from RNS. This is the same reason for the higher costs of D2 and even D5, they both have the RNS. Compared to the only wind case of D4 and D2 in Chapter 5, the penalty terms are much larger since more RNSs occur. Since RNS does not occur in D1 (no wind case), it obtains the second lowest cost and is only more than the perfect forecast.

6.5.2 Stochastic SCUC with Wind and Solar Power Scenarios

The simulation parameters of the stochastic SCUC are provided in Table 6.3. Fig. 6.6 shows the fitness trends of the best and average fitness values for S4. It is obvious that both the best and average fitness curves have an increasing trend. For one thing, the best fitness has a good starting point (> 1.05). This is due to the good initialization using PL method. For another thing, the best fitness continuously increases with iterations due to the utilization of the elitism strategy. The best fitness has a very sharp increase at the first 20 iteration. Then it continues to increase step by step and finally coverages to a high value. The fitness trend during optimization implies the efficiency of the proposed GA-based solution method.

Both the scheduled UC cost and real time ED cost of stochastic cases are displayed in Table 6.4. Just as the deterministic cases, due to the penetration of solar power, the overall real time ED costs of stochastic cases are less than the only wind case in Chapter 5. As can be seen from Table 6.4, the reserve requirements have a significant impact on the two kinds of generation costs. Different from the deterministic cases, the order of the scheduled UC cost and real time ED cost is the same. That is, generation costs from the lowest to the highest is $S1 < S3 < S4 < S2$.

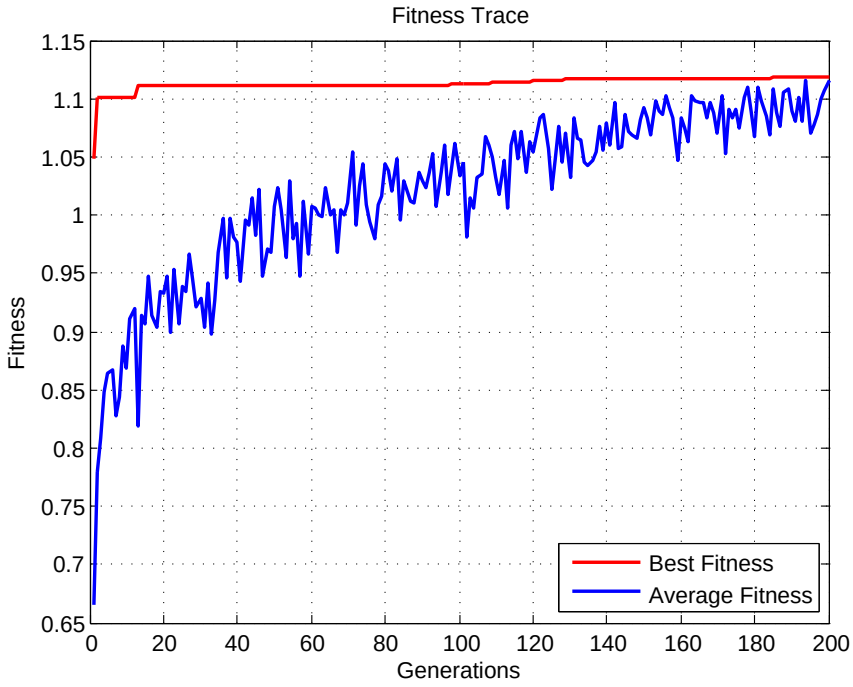

Figure 6.6: Fitness trend during optimization for S4.

This conformity implies the robustness of the stochastic model. In all stochastic cases, the high penalty from RNS and ENS is little. The larger the reserve requirement, the higher the generation cost. S1 is the regular reserve case, the level of the spinning reserve requirement is the lowest. Therefore, S1 has the lowest cost out of the four stochastic cases. It is the opposite situation for S2. Since S2 has the overall highest reserve level, it obtains the highest cost accordingly. As can be clearly seen from Fig. 6.7, most generators need to be turned on to maintain the highest spinning reserve for S2.

6.5.3 Discussion on Available Reserves

Two kinds of available reserves, the scheduled UC reserve and the real time ED reserve, are investigated and compared together for different UC strategies. The main difference of the two kinds of reserve is the usage of the forecast or the real generation outputs of IRESs (wind and solar). The scheduled and real time available reserves are both based on the day-ahead UC decision. The forecast wind and solar power generations are used to calculate the scheduled UC

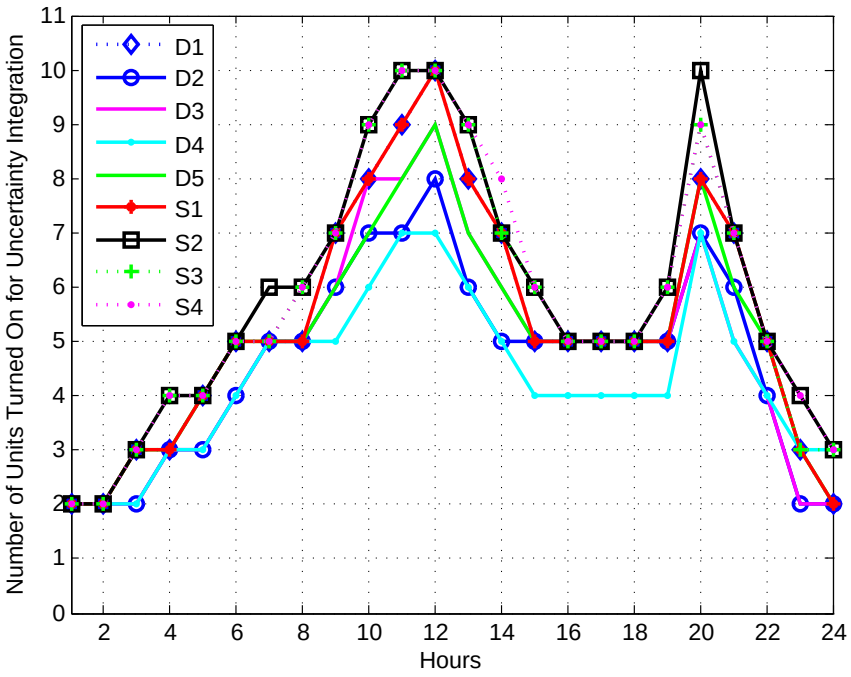

Figure 6.7: The number of units turned on for uncertainty integration.

reserve. However, the real time ED reserve uses the real wind and solar power next day. The purpose of doing so is to show the difference between the planned and realized reserves, and further to investigate the potential risks for different UC strategies.

6.5.3.1 Number of Units Turned on

Fig. 6.7 shows the number of units turned on for the nine simulation cases. As can be seen from Fig. 6.7, for the deterministic cases, the penetration level of renewable energies (wind and solar) has a significant impact on the generator scheduling. The reason behind this is that the less penetration of wind and solar power needs more generator being turned on to maintain the power balance constraint. Then D1 has the most generators turned on and D4 the least out of five deterministic cases. Moreover, for the stochastic cases, the reserve requirement determines the units turned on. The same load, wind, solar and generator outage scenarios are applied in all stochastic cases. Therefore, more generators need to be turned on to maintain a higher level of reserve. In this way, S2 and S1 show the outer and inner edges in all stochastic cases in Fig. 6.7.

6.5.3.2 Scheduled UC Reserves

Fig. 6.8 and 6.9 show the scheduled UC reserves for the deterministic and stochastic cases separately. In all cases, the preassigned reserve requirements can be fully satisfied. This is due to the high penalty terms from RNS and ENS. Then the scheduled UC reserves have to satisfy the different reserve requirements in all cases.

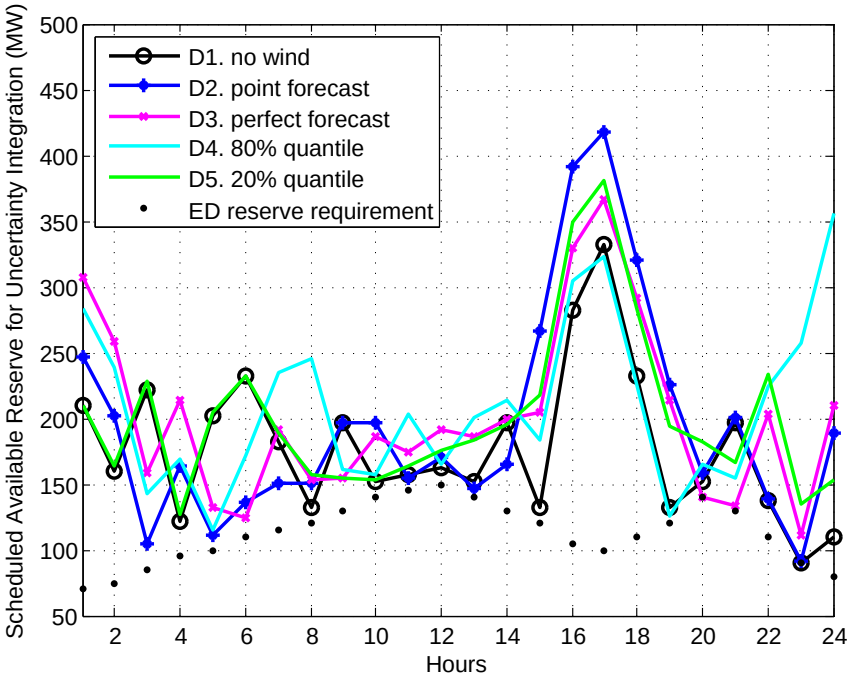

Figure 6.8: Scheduled available reserve of Det. cases for uncertainty integration.

For the deterministic cases, all cases have the same reserve requirement of 10% of load forecasting. As clearly shown in Fig. 6.8, all reserve lines of D1-5 are above the black dotted lines (10% of load) throughout the scheduled 24 hours. An interesting finding is the distribution of the reserve margins during the 24 hours. The least reserve margins happen during Hour 8-13, which are the peak load hours in the day scheduled. This implies the higher level of risk in system operations during these peak load hours. It can be considered as a guideline for system operators for more cautious decision making and risk assessments.

For the stochastic cases, although different reserve strategies are applied and the reserve

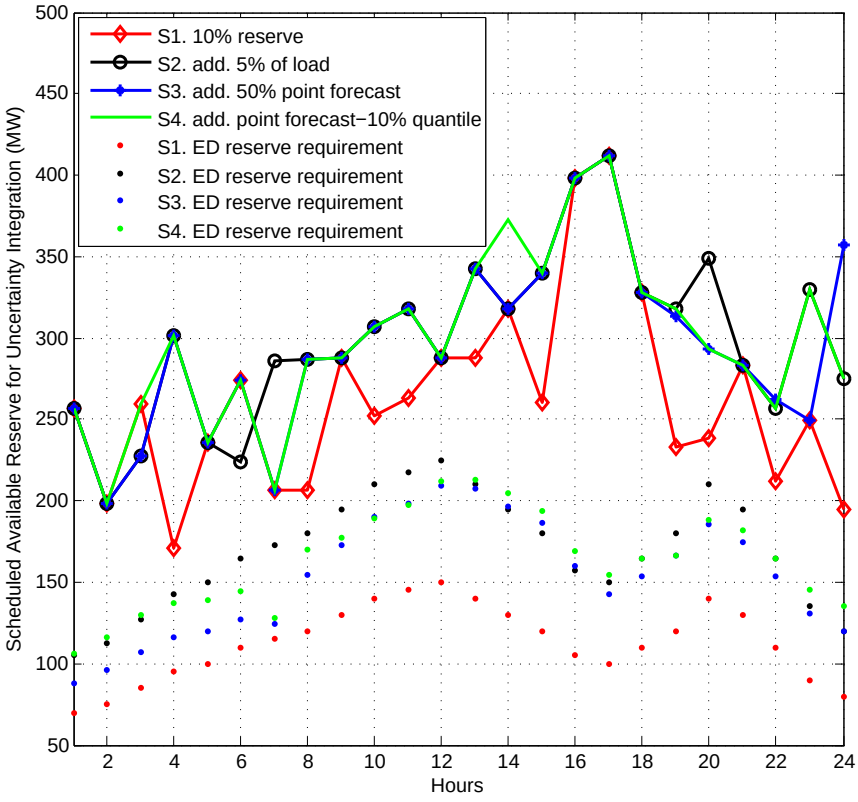

Figure 6.9: Scheduled available reserve of Stoch. cases for uncertainty integration.

requirements are different, all the preassigned reserve requirements can be met. This has been clearly indicated in Fig. 6.9. The dotted lines show the different reserve requirements for S1-S4. For each case, the same color is applied. The solid line stands for the scheduled available reserve, and the dotted line with the same color is the corresponding reserve requirement. S2 has the overall highest scheduled available reserve (black line) while S1 has the lowest (red line).

6.5.3.3 Real Time ED Reserves

Although different reserve strategies are applied to the day-ahead UC scheduling stage, the 10% reserve requirement is set in the real time ED stage to mimic the real operations of the test system. As shown in Fig. 6.10 and 6.11, there are some differences between the planned and the realized reserves with the uncertain IRESs. As per Fig. 6.10, although the ENS does not happen, RNS still occurs in three real time ED cases. D2 has RNS during Hours 8-14, D4 has RNS during Hours 9-20. Even there is RNS at Hour 10 for D5 (conservative 20% quantile forecast of wind and

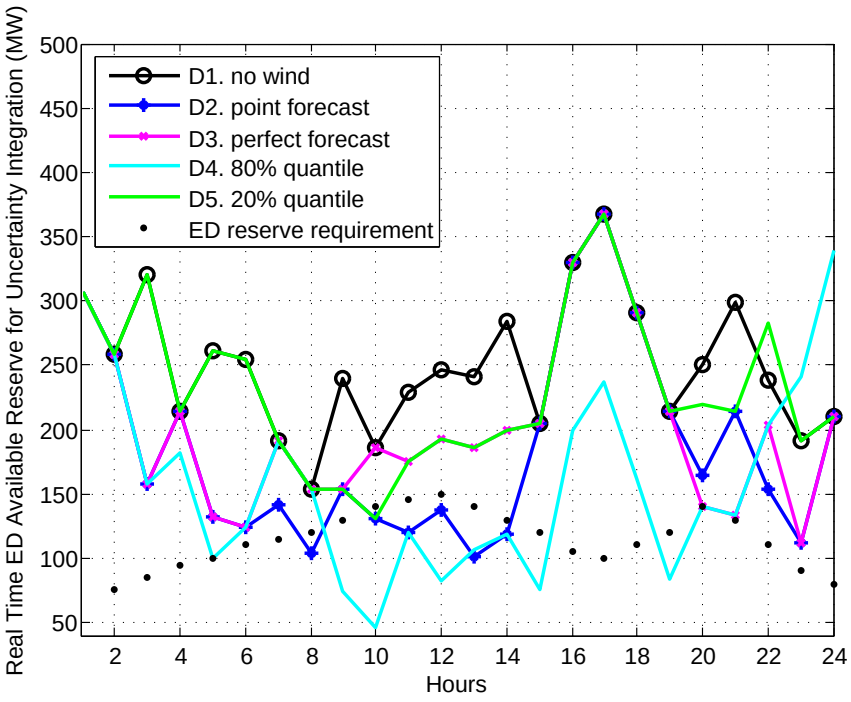

Figure 6.10: Real time ED available reserve of Det. cases for uncertainty integration.

solar power). Compared with the RNS in Chapter 5, RNSs of uncertainty integration of D2 and D4 are obvious larger. It implies the superimposed effect of uncertainties in the computational framework for uncertainty integration.

The case of D5 tells us that even if the conservative deterministic forecast is applied, the obtained UC scheduling still runs the risk and is unreliable in system operations. As can be seen from Fig. 6.8, the peak load hours (Hour 8-13) have the least reserve margins for the scheduled UC reserve. It is not surprising that RNS happens during these peak load hours for the real time ED reserve. As a result of high uncertainties, IRESs may be under or over estimated. It is especially risky if the forecasted IRESs are highly over estimated. Since the high uncertainties of IRESs, the involved risk is also much higher during these peak load hours. The potential risks have been clearly reflected by the RNSs as shown in Fig. 6.10.

In contrary to deterministic cases, all the four stochastic cases (with different reserve strategies) can always meet the reserve requirement in real time ED in Fig. 6.11. Out of the four stochastic cases, there is no RNS and ENS. Uncertainties from load forecasting, IRESs and gen-

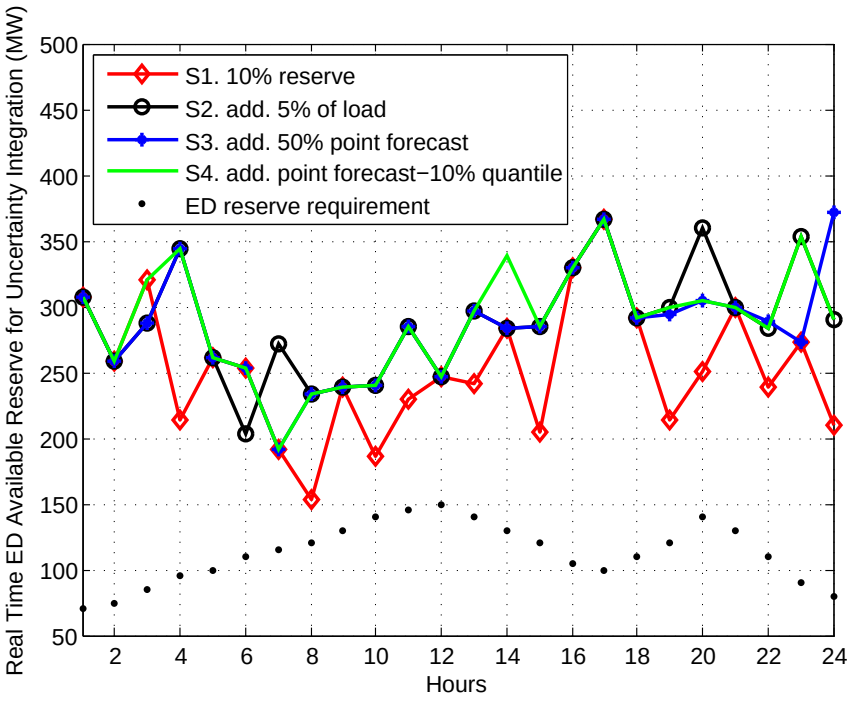

Figure 6.11: Real time ED available reserve of Stoch. cases for uncertainty integration.

erator outages are well represented by the generated scenarios in the stochastic model. The obtained results from the stochastic model are significantly more robust than the deterministic model. Deterministic model considers only one scenario of the uncertain IRESs (wind and solar power). The uncertainties of IRESs cannot be adequately represented in the deterministic model. However, the stochastic model captures the wind and solar power uncertainties in different scenarios. Different cases of IRES realizations, as well as the load and generator outage uncertainties, have been considered. Therefore, stochastic model is more robust than the deterministic ones. Just like the scheduled UC reserve in Fig. 6.9, higher reserve requirement will lead to a higher reserve margin in the real time ED reserve. Correspondingly, the cost is also higher for lower risk (higher reserve margin). The trade-off and the reserve strategies can be determined by the system operators for their preferences during decision making.

6.6 Conclusions

It is very important to address the uncertainty problem comprehensively by including different sources of uncertainty. A comprehensive computational framework for uncertainty integration in DPSs with IRESs has been proposed in this chapter. In contrast to previous studies, the load, wind and solar power forecast uncertainties as well as the uncertainty from generator outages are included in the proposed framework. A new scenario generation method is proposed to model the uncertainty of IRESs such as wind and solar power forecasting. Then these uncertainties are all represented as scenarios and incorporated into a stochastic SCUC model, and a heuristic GA is utilized to solve this stochastic SCUC problem. Five deterministic and four stochastic case studies have been investigated and compared together. Different UC and reserve strategies have been demonstrated. According to the results, the overall costs are lower than the only wind case in Chapter 5 due to the penetration of solar power. The RNSs in the deterministic cases of the integration framework are also larger than the only wind case. It implies the superimposed effect of uncertainties in the computational framework for uncertainty integration. Comparative results show the power systems run higher level of risk during peak load hours. The stochastic model is more robust than the deterministic ones.

Chapter 7

Conclusions and Future Work

In this final chapter, major findings and contributions of this research are summarized. Their implications are discussed and areas for future work are suggested.

7.1 Overall Conclusions

This study explored uncertainty modeling methods applied in distributed power systems (DPSs). The uncertainty modeling methods for forecasting and incorporating forecast uncertainty into unit commitment (UC) in DPSs for decision making were developed. After the literature review on uncertainty modeling and DPSs, a new and efficient uncertainty modeling method for forecasting using the neural network (NN)-based prediction intervals (PIs) was proposed. The PSO-based LUBE method was then extended and applied to electrical load and wind power interval forecasting. Sufficient case studies from different areas have been implemented. The forecasting results clearly show the advantages of the proposed PSO-based LUBE method in terms of quality of PIs and computation time.

Wind power forecast uncertainty was further incorporated into the stochastic UC scheduling and economic dispatches (EDs) for decision making and risk assessment. Finally, a computational framework was proposed for uncertainty integration in DPSs. Load, wind and solar power forecast and the generator outage uncertainties were integrated together. Five deterministic and

four stochastic case studies were investigated from the perspective of generation costs and available reserves. Simulation results prove that the stochastic model outperforms the deterministic ones. There are some differences between the scheduled reserves and real time ED reserves. Power systems run higher level of risk during peak load hours and these potential risks need to be considered by system operators during decision making.

7.2 Main Contributions

Penetration of intermittent renewable energy sources (IRESs) such as wind and solar power into traditional power systems has significant impacts on system operation, economics, reliability and security. Forecasting and UC scheduling with uncertain IRESs have become a lot more challenging and important than before in daily operations of DPSs. The main contributions of this research consist of the following aspects:

- A Comprehensive Overview of Uncertainty Modeling in DPSs

    1) This was the first systematic overview of uncertainty modeling approaches in DPSs and some application case studies were implemented. Uncertainty representations of different sources of uncertainty associated with IRESs were comprehensively investigated.

    2) A comprehensive investigation on forecasting, UC and renewable energy integration was conducted from the uncertainty quantification point of view. i) The uncertainty modeling methods for forecasting and UC problems in DPSs were reviewed. ii) The traditional and new LUBE methods for construction of NN-based PIs were introduced and compared. iii) UC with IRESs, i.e., the UC problem formulations and different solution methods were investigated. iv) Incorporation of wind power forecast uncertainties into UC and other DPS applications were reviewed and the drawback of current study was summarized.

- Advanced Uncertainty Handling Methods for Forecasting

  1) A recently introduced PI construction method named LUBE method was applied and extended. A new width evaluation index PINRW, which is suitable for NN training, was proposed. Further, a new cost function was developed for the comprehensive evaluation of PIs based on their width and coverage probability. PSO with mutation operator was used for minimizing the cost function and adjusting NN parameters in the LUBE method. The proposed PSO-based LUBE method showed significant improvements in the quality of results and speed.

  2) A new problem formulation for PI construction was proposed. The primary multi-objective problem was successfully transformed into a constrained single-objective problem. Advantages of this new formulation are its closeness to the original problem and require fewer parameters than cost functions. Correlation analysis was applied to help choose the inputs of NN models. Comparative results for both load and wind power generation datasets suggest that not only were high PICP and narrow PINAW obtained, but also the PI construction time remained short. Quality of PIs was significantly improved when compared to ARIMA, exponential smoothing and naive models.

- Incorporation of IRES Forecast Uncertainties into Stochastic UC for Decision Making and Risk Assessment

  1) A new scenario generation method was proposed. This method built an important bridge between the PIs and the scenarios used in the stochastic model. Wind generation uncertainties were considered as a stochastic process through the generated scenarios. A security-constrained UC (SCUC) model with uncertain wind power was built and genetic algorithm (GA) was proposed to solve the stochastic SCUC problem. Five deterministic and four stochastic case studies were implemented. Two

kinds of generation costs and available reserves for different UC strategies were discussed and compared together.

2) A computational framework for uncertainty integration in DPSs with IRESs was proposed. Load forecast uncertainty was assumed to follow a normal distribution. Wind and solar power forecast uncertainties were represented by a list of PIs and the proposed scenario generation method was applied to generate scenarios from PIs. The uncertainty of generator outages was modeled as the discrete scenarios. Load, wind, solar and generator outage uncertainty scenarios were further incorporated into a stochastic SCUC model. Deterministic and stochastic cases were implemented to investigate the impacts of generation costs and reserves on system operations from the decision making and risk assessment point of view.

7.3 Future Work

The PSO-based LUBE method for constructing PIs would be a useful tool to improve the existing uncertainty handling method for load and wind power forecasting in DPSs. The proposed scenario generation method and the integration framework have provided valuable insight into the decision making and risk assessment for system operators under uncertainties. Further improvements and interesting areas for future research are as follows:

- The PSO-based LUBE method can be extended to other forecast tasks involving uncertainties, such as electricity price forecasting and transportation system uncertainty modeling.

- In order to reserve the uncertainties in the original demand datasets, we did not have special considerations during holidays and weekends for short-term load forecasting. Electrical loads usually have different patterns and higher uncertainties during holidays and weekends. Further improvements should be made by smoothing out the holiday datasets through average, separating the weekends from weekdays and applying enhanced input

selection methods.

- The proposed scenario generation method and the stochastic models may be applied to other statistical applications for the representation of a stochastic process.

- In SCUC test systems, the focus was on incorporating the wind power forecast uncertainties into stochastic UC for decision-making rather than building a very complex test system. Thus some constraints in UC were omitted, such as the transmission line constraint and the ramping rate constraint of generators. A direct extension of this work is to consider more constraints in UC and use bigger test systems. For example, the number of units can be extended up to 100. New test cases, such as the 26-unit or 52-unit systems, would be further investigated. Some reliability and risk assessment indices, e.g., standard deviation of the scenario cost, will also be considered.

- The proposed PSO-based LUBE method and scenario generation method make little assumptions about data distributions. Therefore, they can be generally and easily conducted. It would also be interesting to extend the methodologies to other areas in smart grid applications for uncertainty modeling, such as renewable integrations, optimal bidding in electricity market and optimal sitting and sizing of distributed generations.

- In the proposed computational framework for uncertainty integration, the uncertainty representation methods for different uncertainty sources could be further improved. For example, the capacity outage probability table and the Markov model can be applied to represent the uncertainty of generator outages. The proposed computational framework would be shifted to other applications involving different kinds of uncertainties, such as the smart grid and transportation system for uncertainty quantification and integration.

Bibliography

[1] Y. Atwa, E. El-Saadany, M. Salama, and R. Seethapathy, "Optimal renewable resources mix for distribution system energy loss minimization," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 360–370, Feb. 2010.

[2] Y. Makarov, P. 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 Transactions on Sustainable Energy*, vol. 2, no. 4, pp. 433–442, Oct 2011.

[3] B. Ummels, M. Gibescu, E. Pelgrum, W. Kling, and A. Brand, "Impacts of wind power on thermal generation unit commitment and dispatch," *IEEE Transactions on Energy Conversion*, vol. 22, no. 1, pp. 44–51, Mar. 2007.

[4] N. Amjady, F. Keynia, and H. Zareipour, "Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle swarm optimization," *IEEE Transactions on Sustainable Energy*, vol. 2, no. 3, pp. 265–276, July 2011.

[5] H. Quan, D. Srinivasan, A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of neural network-based prediction intervals for short-term electrical load forecasting," in *2013 IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG)*, April 2013, pp. 66–72.

[6] B. M. Ayyub and G. J. Klir, *Uncertainty modeling and analysis in engineering and the sciences*. Chapman & Hall/CRC Press, 2006.

[7] N. Attoh-Okine and B. M. Ayyub, *Applied Research in Uncertainty Modeling and Analysis*, ser. International Series in Intelligent Technologies. Springer, 2005.

[8] F. Alvarado, Y. Hu, and R. Adapa, "Uncertainty in power system modeling and computation," in *IEEE International Conference on Systems, Man and Cybernetics*, 1992, pp. 754–760.

[9] M. C. Kennedy and A. O'Hagan, "Bayesian calibration of computer models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 3, pp. 425–464, 2001.

[10] S. H. Lee and W. Chen, "A comparative study of uncertainty propagation methods for black-box type functions," in *ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers, 2007, pp. 1275–1284.

[11] S. Raychaudhuri, "Introduction to Monte Carlo simulation," in *Simulation Conference, 2008. WSC 2008. Winter*. IEEE, 2008, pp. 91–100.

[12] P. J. Smith, M. Shafi, and H. Gao, "Quick simulation: A review of importance sampling techniques in communications systems," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 4, pp. 597–613, 1997.

[13] C. G. Bucher, "Adaptive samplingan iterative fast monte carlo procedure," *Structural Safety*, vol. 5, no. 2, pp. 119–126, 1988.

[14] B. Lallemand, G. Plessis, T. Tison, and P. Level, "Neumann expansion for fuzzy finite element analysis," *Engineering computations*, vol. 16, no. 5, pp. 572–583, 1999.

[15] K. Sepahvand, S. Marburg, and H.-J. Hardtke, "Uncertainty quantification in stochastic systems using polynomial chaos expansion," *International Journal of Applied Mechanics*, vol. 2, no. 02, pp. 305–353, 2010.

[16] X. Du and W. Chen, "A most probable point-based method for efficient uncertainty analysis," *Journal of Design and Manufacturing Automation*, vol. 4, no. 1, pp. 47–66, 2001.

[17] F. Xiong, K. Guo, and W. Zhou, "Uncertainty propagation techniques in probabilistic design of multilevel systems," in *2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE)*. IEEE, 2011, pp. 874–878.

[18] H. Xu and S. Rahman, "A generalized dimension-reduction method for multidimensional integration in stochastic mechanics," *International Journal for Numerical Methods in Engineering*, vol. 61, no. 12, pp. 1992–2019, 2004.

[19] C. Chatfield, "Calculating interval forecasts," *Journal of Business and Economic Statistics*, vol. 11, no. 2, pp. 121–135, 1993.

[20] L. A. Zadeh, "Fuzzy logic," *Computer*, vol. 21, no. 4, pp. 83–93, 1988.

[21] T. Ackermann, G. Andersson, and L. Söder, "Distributed generation: a definition," *Electric power systems research*, vol. 57, no. 3, pp. 195–204, 2001.

[22] T. Logenthiran, "Multi-Agent System for Control and Management of Distributed Power Systems," Ph.D. dissertation, National University of Singapore, Singapore, Mar. 2012.

[23] S. Luo, "A review of distributed power systems part I: DC distributed power system," *IEEE Aerospace and Electronic Systems Magazine*, vol. 20, no. 8, pp. 5–16, Aug. 2005.

[24] T. Logenthiran, D. Srinivasan, A. Khambadkone, and H. Aung, "Multi-Agent System (MAS) for short-term generation scheduling of a microgrid," in *2010 IEEE International Conference on Sustainable Energy Technologies (ICSET)*, Dec. 2010, pp. 1–6.

[25] F. P. Sioshansi, *Smart grid: integrating renewable, distributed & efficient energy*. Academic Press, 2011.

[26] P. Ruiz, C. Philbrick, E. Zak, K. Cheung, and P. Sauer, "Uncertainty management in the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 642–651, May 2009.

[27] R. Billinton and D. Huang, "Effects of load forecast uncertainty on bulk electric system reliability evaluation," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 418–425, May 2008.

[28] J. Valenzuela, M. Mazumdar, and A. Kapoor, "Influence of temperature and load forecast uncertainty on estimates of power generation production costs," *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 668–674, May 2000.

[29] G. Toh and H. Gooi, "Incorporating forecast uncertainties into EENS for wind turbine studies," *Electric Power Systems Research*, vol. 81, no. 2, pp. 430–439, 2011.

[30] R. Bo and F. Li, "Probabilistic LMP forecasting considering load uncertainty," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1279–1289, Aug 2009.

[31] D. Koval and A. Chowdhury, "Outage data concepts for generation and transmission equipment," in *Power Engineering Society General Meeting, 2006. IEEE*, 2006, pp. 1–7.

[32] L. Wu, M. Shahidehpour, and Y. Fu, "Security-constrained generation and transmission outage scheduling with uncertainties," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1674–1685, Aug 2010.

[33] J. H. Roh, M. Shahidehpour, and L. Wu, "Market-based generation and transmission planning with uncertainties," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1587–1598, Aug 2009.

[34] M. Matos and R. Bessa, "Setting the operating reserve using probabilistic wind power forecasts," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 594–603, May 2011.

[35] K. Zou, A. Agalgaonkar, K. Muttaqi, and S. Perera, "Distribution system planning with incorporating DG reactive capability and system uncertainties," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 1, pp. 112–123, Jan 2012.

[36] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.

[37] Z. Liu, F. Wen, and G. Ledwich, "Optimal siting and sizing of distributed generators in distribution systems considering uncertainties," *IEEE Transactions on Power Delivery*, vol. 26, no. 4, pp. 2541–2551, Oct 2011.

[38] M. Khalid and A. Savkin, "A method for short-term wind power prediction with multiple observation points," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 579–586, May 2012.

[39] H. M. Merrill and A. J. Wood, "Risk and uncertainty in power system planning," *International Journal of Electrical Power & Energy Systems*, vol. 13, no. 2, pp. 81–90, 1991.

[40] H. Tijms, *Understanding probability*. Cambridge University Press, 2012.

[41] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Comprehensive review of neural network-based prediction intervals and new advances," *IEEE Transactions on Neural Networks*, vol. 22, no. 9, pp. 1341–1356, Sep. 2011.

[42] A. Khosravi, E. Mazloumi, S. Nahavandi, D. Creighton, and J. van Lint, "Prediction intervals to account for uncertainties in travel time prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 537–547, Jun. 2011.

[43] A. Khosravi, S. Nahavandi, D. Creighton, and A. F. Atiya, "Lower upper bound estimation method for construction of neural network-based prediction intervals," *IEEE Transactions on Neural Networks*, vol. 22, no. 3, pp. 337–346, Mar. 2011.

[44] J. Mendel, "Fuzzy logic systems for engineering: a tutorial," *Proceedings of the IEEE*, vol. 83, no. 3, pp. 345–377, Mar 1995.

[45] L. A. Zadeh, "Fuzzy logic = computing with words," *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 2, pp. 103–111, May 1996.

[46] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*. Princeton University Press, 2009.

[47] A. Shapiro and A. Philpott, "A tutorial on stochastic programming," *Manuscript. Available at www2. isye. gatech. edu/˜ ashapiro/publications. html*, 2007.

[48] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on stochastic programming: modeling and theory*. SIAM, 2009, vol. 9.

[49] L. Wu, M. Shahidehpour, and Z. Li, "Comparison of scenario-based and interval optimization approaches to stochastic SCUC," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 913–921, May 2012.

[50] J. T. G. Hwang and A. A. Ding, "Prediction intervals for artificial neural networks," *Journal of the American Statistical Association*, vol. 92, no. 438, pp. 748–757, 1997.

[51] R. D. Veaux, J. Schumi, J. Schweinsberg, D. Shellington, and L. Ungar, "Prediction intervals for neural networks via nonlinear regression," *Technometrics*, vol. 40, no. 4, pp. 273–282, 1998.

[52] C. M. Bishop, *Neural Networks for Pattern Recognition*. New York, USA: Oxford University Press, Inc., 1995.

[53] D. A. Nix and A. S. Weigend, "Estimating the mean and variance of the target probability distribution," in *IEEE International Conference on Neural Networks*, vol. 1, 1994, pp. 55–60.

[54] T. Heskes, "Practical confidence and prediction intervals," in *Advances in Neural Information Processing Systems 9*. MIT press, 1997, pp. 176–182.

[55] H. Quan, D. Srinivasan, and A. Khosravi, "Construction of neural network-based prediction intervals using particle swarm optimization," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*, June 2012, pp. 647–653.

[56] A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of optimal prediction intervals for load forecasting problems," *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1496–1503, Aug. 2010.

[57] G. Chryssolouris, M. Lee, and A. Ramsey, "Confidence interval prediction for neural network models," *IEEE Transactions on Neural Networks*, vol. 7, no. 1, pp. 229–232, Jan. 1996.

[58] H. Quan, D. Srinivasan, and A. Khosravi, "Particle swarm optimization for construction of neural network-based prediction intervals," *Neurocomputing*, vol. 127, pp. 172–180, 2014.

[59] C.-H. Chen, J.-C. Wu, and J.-H. Chen, "Prediction of flutter derivatives by artificial neural networks," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, no. 10-11, pp. 1925 – 1937, 2008.

[60] H. Quan, D. Srinivasan, and A. Khosravi, "Short-term load and wind power forecasting using neural network-based prediction intervals," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 2, pp. 303–315, Feb. 2014.

[61] T. Logenthiran and D. Srinivasan, "Formulation of unit commitment (UC) problems and analysis of available methodologies used for solving the problems," in *2010 IEEE International Conference on Sustainable Energy Technologies (ICSET)*, Dec 2010, pp. 1–6.

[62] T. Li and M. Shahidehpour, "Price-based unit commitment: a case of Lagrangian relaxation versus mixed integer programming," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 2015–2025, Nov 2005.

[63] C.-P. Cheng, C.-W. Liu, and C.-C. Liu, "Unit commitment by Lagrangian relaxation and genetic algorithms," *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 707–714, May 2000.

[64] C. Wang and S. M. Shahidehpour, "Effects of ramp-rate limits on unit commitment and economic dispatch," *IEEE Transactions on Power Systems*, vol. 8, no. 3, pp. 1341–1350, Aug 1993.

[65] A. J. Wood and B. F. Wollenberg, *Power generation, operation, and control*. John Wiley & Sons, Inc., 1996.

[66] N. Padhy, "Unit commitment-a bibliographical survey," *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 1196–1205, 2004.

[67] W. Ongsakul and N. Petcharaks, "Unit commitment by enhanced adaptive Lagrangian relaxation," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 620–628, 2004.

[68] P. Xiong, "Unit Commitment under Uncertainty," Ph.D. dissertation, National University of Singapore, Singapore, Aug. 2012.

[69] P.-H. Chen, "Two-level hierarchical approach to unit commitment using expert system and elite PSO," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 780–789, 2012.

[70] B. Venkatesh, P. Yu, H. Gooi, and D. Choling, "Fuzzy MILP unit commitment incorporating wind generators," *IEEE Transactions on Power Systems*, vol. 23, no. 4, pp. 1738–1746, 2008.

[71] S. Kazarlis, A. Bakirtzis, and V. Petridis, "A genetic algorithm solution to the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 11, no. 1, pp. 83–92, 1996.

[72] K. Swarup and S. Yamashiro, "Unit commitment solution methodology using genetic algorithm," *IEEE Transactions on Power Systems*, vol. 17, no. 1, pp. 87–91, 2002.

[73] T. Logenthiran and D. Srinivasan, "Particle swarm optimization for unit commitment problem," in *2010 IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2010, pp. 642–647.

[74] X. Li and C. Jiang, "Short-term operation model and risk management for wind power penetrated system in electricity market," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 932–939, May 2011.

[75] X. Yuan, A. Su, H. Nie, Y. Yuan, and L. Wang, "Application of enhanced discrete differential evolution approach to unit commitment problem," *Energy Conversion and Management*, vol. 50, no. 9, pp. 2449–2456, 2009.

[76] D. Srinivasan and J. Chazelas, "A priority list-based evolutionary algorithm to solve large scale unit commitment problem," in *International Conference on Power System Technology, 2004. PowerCon 2004.*, vol. 2, 2004, pp. 1746–1751.

[77] T. Logenthiran and D. Srinivasan, "Short term generation scheduling of a Microgrid," in *TENCON 2009-2009 IEEE Region 10 Conference*, Jan 2009, pp. 1–6.

[78] J. Wang, M. Shahidehpour, and Z. Li, "Security-constrained unit commitment with volatile wind power generation," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1319–1327, Aug 2008.

[79] M. A. Ortega-Vazquez and D. S. Kirschen, "Estimating the spinning reserve requirements in systems with significant wind power generation penetration," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 114–124, Feb 2009.

[80] A. Botterud, Z. Zhou, J. Wang, R. J. Bessa, H. Keko, J. Sumaili, and V. Miranda, "Wind power trading under uncertainty in LMP markets," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 894–903, May 2012.

[81] A. Botterud, J. Wang, R. Bessa, H. Keko, and V. Miranda, "Risk management and optimal bidding for a wind power producer," in *2010 IEEE Power and Energy Society General Meeting*, July 2010, pp. 1–8.

[82] A. Sturt and G. Strbac, "Efficient stochastic scheduling for simulation of wind-integrated power systems," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 323–334, Feb 2012.

[83] R. Jiang, J. Wang, and Y. Guan, "Robust unit commitment with wind power and pumped storage hydro," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 800–810, May 2012.

[84] D. Bertsimas, E. Litvinov, X. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 52–63, Feb 2013.

[85] C. Zhao and Y. Guan, "Unified stochastic and robust unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3353–3361, Aug 2013.

[86] C. Zhao, J. Wang, J.-P. Watson, and Y. Guan, "Multi-stage robust unit commitment considering wind and demand response uncertainties," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2708–2717, Aug 2013.

[87] R. Jiang, J. Wang, M. Zhang, and Y. Guan, "Two-stage minimax regret robust unit commitment," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2271–2282, Aug 2013.

[88] W. Wu, J. Chen, B. Zhang, and H. Sun, "A robust wind power optimization method for look-ahead power dispatch," *IEEE Transactions on Sustainable Energy*, vol. 5, no. 2, pp. 507–515, April 2014.

[89] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

[90] H. Hippert, C. Pedreira, and R. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44–55, Feb 2001.

[91] V. Yadav and D. Srinivasan, "A SOM-based hybrid linear-neural model for short-term load forecasting," *Neurocomputing*, vol. 74, no. 17, pp. 2874–2885, 2011.

[92] K. Methaprayoon, C. Yingvivatanapong, W.-J. Lee, and J. Liao, "An integration of ann wind power estimation into unit commitment considering the forecasting uncertainty," *IEEE Transactions on Industry Applications*, vol. 43, no. 6, pp. 1441–1448, Nov.-Dec. 2007.

[93] D. Srinivasan, M. C. Choy, and R. Cheu, "Neural networks for real-time traffic signal control," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 3, pp. 261–272, Sep. 2006.

[94] B. K. Panigrahi, P. K. Dash, and J. B. V. Reddy, "Hybrid signal processing and machine intelligence techniques for detection, quantification and classification of power quality disturbances," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 3, pp. 442–454, 2009.

[95] A. Jain, R. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, Jan 2000.

[96] R. C. and Muresan, "Pattern recognition using pulse-coupled neural networks and discrete fourier transforms," *Neurocomputing*, vol. 51, no. 0, pp. 487 – 493, 2003.

[97] C. Rodriguez and G. Anders, "Energy price forecasting in the ontario competitive power system market," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 366–374, Feb. 2004.

[98] A. Khosravi, S. Nahavandi, and D. Creighton, "Prediction interval construction and optimization for adaptive neurofuzzy inference systems," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 5, pp. 983–988, Oct. 2011.

[99] S. F. Christian and C. Lebiere, "The cascade-correlation learning architecture," in *Advances in Neural Information Processing Systems 2*. Morgan Kaufmann, 1990, pp. 524–532.

[100] L. Ma and K. Khorasani, "New training strategies for constructive neural networks with application to regression problems," *Neural Networks*, vol. 17, no. 4, pp. 589–609, 2004.

[101] A. and Luchetta, "Automatic generation of the optimum threshold for parameter weighted pruning in multiple heterogeneous output neural networks," *Neurocomputing*, vol. 71, no. 16-18, pp. 3553–3560, 2008.

[102] P. L. Narasimha, W. H. Delashmit, M. T. Manry, J. Li, and F. Maldonado, "An integrated growing-pruning method for feedforward network training," *Neurocomputing*, vol. 71, no. 13-15, pp. 2831–2847, 2008.

[103] E. Ricci and R. Perfetti, "Improved pruning strategy for radial basis function networks with dynamic decay adjustment," *Neurocomputing*, vol. 69, no. 13-15, pp. 1728–1732, 2006.

[104] J. Wright and M. Manic, "Neural network architecture selection analysis with application to cryptography location," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2010, pp. 1–6.

[105] A. Pavelka and A. Proch, "Algorithms for initialization of neural network weights random numbers in matlab," *Control Engineering*, vol. 2, pp. 453–459, 2004.

[106] D. Nguyen and B. Widrow, "Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights," in *The 1990 International Joint Conference on Neural Networks (IJCNN)*, vol. 3, Jun. 1990, pp. 21–26.

[107] B. K. Panigrahi, V. R. Pandi, and S. Das, "Adaptive particle swarm optimization approach for static and dynamic economic load dispatch," *Energy Conversion and Management*, vol. 49, no. 6, pp. 1407–1415, 2008.

[108] G. Pulido and C. Coello, "A constraint-handling mechanism for particle swarm optimization," in *IEEE Congress on Evolutionary Computation, CEC 2004.*, vol. 2, Jun. 2004, pp. 1396–1403.

[109] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the Sixth International Symposium on Micro Machine and Human Science (MHS '95.)*, Oct. 1995, pp. 39–43.

[110] C. Coello, G. Pulido, and M. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 256–279, Jun. 2004.

[111] J. Taylor and P. McSharry, "Short-term load forecasting methods: An evaluation based on European data," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2213–2219, Nov. 2007.

[112] D. Srinivasan and M. Lee, "Survey of hybrid fuzzy neural approaches to electric load forecasting," in *IEEE International Conference on Systems, Man and Cybernetics, 1995. Intelligent Systems for the 21st Century.*, vol. 5, Oct. 1995, pp. 4004–4008.

[113] A. Conejo, M. Plazas, R. Espinola, and A. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1035–1042, May 2005.

[114] J. Deng and P. Jirutitijaroen, "Short-term load forecasting using time series analysis: A case study for singapore," in *2010 IEEE Conference on Cybernetics and Intelligent Systems (CIS)*, June 2010, pp. 231–236.

[115] R. Sood, I. Koprinska, and V. Agelidis, "Electricity load forecasting based on autocorrelation analysis," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*, July 2010, pp. 1–8.

[116] A. da Silva and L. Moulin, "Confidence intervals for neural network based short-term load forecasting," *IEEE Transactions on Power Systems*, vol. 15, no. 4, pp. 1191–1196, Nov. 2000.

[117] Q. Yu, H. Tang, K. Tan, and H. Li, "Rapid feedforward computation by temporal encoding and learning with spiking neurons," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 10, pp. 1539–1552, Oct 2013.

[118] G. Capizzi, C. Napoli, and F. Bonanno, "Innovative second-generation wavelets construction with recurrent neural networks for solar radiation forecasting," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 11, pp. 1805–1815, 2012.

[119] A. Khosravi, S. Nahavandi, D. Creighton, and D. Srinivasan, "Interval type-2 fuzzy logic systems for load forecasting: A comparative study," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1274–1282, 2012.

[120] D. Hidalgo, P. Melin, and O. Castillo, "An optimization method for designing type-2 fuzzy inference systems based on the footprint of uncertainty using genetic algorithms," *Expert Systems with Applications*, vol. 39, no. 4, pp. 4590–4598, 2012.

[121] J. R. Castro, O. Castillo, P. Melin, O. Mendoza, and A. Rodríguez-Díaz, "An interval type-2 fuzzy neural network for chaotic time series prediction with cross-validation and akaike test," in *Soft Computing for Intelligent Control and Mobile Robotics*. Springer, 2011, pp. 269–285.

[122] C. Potter and M. Negnevitsky, "Very short-term wind forecasting for Tasmanian power generation," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 965–972, May 2006.

[123] P. Melin, J. Soto, O. Castillo, and J. Soria, "A new approach for time series prediction using ensembles of ANFIS models," *Expert Systems with Applications*, vol. 39, no. 3, pp. 3494–3506, 2012.

[124] C. Sheng, J. Zhao, W. Wang, and H. Leung, "Prediction intervals for a noisy nonlinear time series based on a bootstrapping reservoir computing network ensemble," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 7, pp. 1036–1048, 2013.

[125] F. Valdez, P. Melin, and O. Castillo, "Evolutionary method combining particle swarm optimisation and genetic algorithms using fuzzy logic for parameter adaptation and aggregation: the case neural network optimisation for face recognition," *International Journal of Artificial Intelligence and Soft Computing (IJAISC)*, vol. 2, no. 1-2, pp. 77–102, 2010.

[126] M. Pulido, P. Melin, and O. Castillo, "Genetic optimization of ensemble neural networks for complex time series prediction," in *The 2011 International Joint Conference on Neural Networks (IJCNN)*, 2011, pp. 202–206.

[127] W.-C. Yeh, "New parameter-free simplified swarm optimization for artificial neural network training and its application in the prediction of time series," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 4, pp. 661–665, 2013.

[128] R. L. Winkler, "A decision-theoretic approach to interval estimation," *Journal of the American Statistical Association*, vol. 67, no. 337, pp. 187–191, 1972.

[129] H. Quan, D. Srinivasan, and A. Khosravi, "Uncertainty handling using neural network-based prediction intervals for electrical load forecasting," *Energy*, vol. 73, pp. 916–925, Aug. 2014.

[130] K. Zielinski and R. Laur, "Constrained single-objective optimization using particle swarm optimization," in *IEEE Congress on Evolutionary Computation, CEC 2006.*, 2006, pp. 443–450.

[131] B. Bowerman, R. O'Connell, and A. Koehler, *Forecasting, Time Series, and Regression: An Applied Approach*, ser. Duxbury applied series. Thomson Brooks/Cole, 2005.

[132] D. Srinivasan, Z. Guofan, A. Khosravi, S. Nahavandi, and D. Creighton, "Hybrid neural-evolutionary model for electricity price forecasting," in *The 2011 International Joint Conference on Neural Networks (IJCNN)*, Jul. 31-Aug. 5 2011, pp. 3164–3169.

[133] S.-T. Chen, D. Yu, and A. Moghaddamjo, "Weather sensitive short-term load forecasting using nonfully connected artificial neural network," *IEEE Transactions on Power Systems*, vol. 7, no. 3, pp. 1098–1105, Aug. 1992.

[134] S. Fan and R. Hyndman, "Short-term load forecasting based on a semi-parametric additive model," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 134–141, Feb. 2012.

[135] D. Yang, P. Jirutitijaroen, and W. M. Walsh, "Hourly solar irradiance time series forecasting using cloud cover index," *Solar Energy*, vol. 86, no. 12, pp. 3531–3543, 2012.

[136] R. Shumway and D. Stoffer, *Time Series Analysis and Its Applications: With R Examples*, ser. Springer Texts in Statistics. Springer, 2010.

[137] Y. Ding, P. Wang, L. Goel, P. C. Loh, and Q. Wu, "Long-term reserve expansion of power systems with high wind power penetration using universal generating function methods," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 766–774, May 2011.

[138] J. Wang, A. Botterud, V. Miranda, C. Monteiro, and G. Sheble, "Impact of wind power forecasting on unit commitment and dispatch," in *8th Int. Workshop on Large-Scale Integration of Wind Power into Power Systems, Bremen, Germany*, 2009.

[139] A. Khosravi and S. Nahavandi, "Combined nonparametric prediction intervals for wind power generation," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 4, pp. 849–856, Oct. 2013.

[140] A. Khosravi, E. Mazloumi, S. Nahavandi, D. Creighton, and J. V. Lint, "A genetic algorithm-based method for improving quality of travel time prediction intervals," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 6, pp. 1364–1376, 2011.

[141] Q. Sun, J. Zhang, Y. Zhao, and D. Ma, "Power distribution self-healing system based on intuitionistic uncertainty-rough sets," *Applied Mathematics & Information Sciences*, vol. 7, no. 1L, 2013.

[142] 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, 2009.

[143] X.-Y. Ma, Y.-Z. Sun, and H.-L. Fang, "Scenario generation of wind power based on statistical uncertainty and variability," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 4, pp. 894–904, Oct 2013.

[144] Q. Sun, Y. Yu, Y. Luo, and X. Liu, "Application of BFNN in power flow calculation in smart distribution grid," *Neurocomputing*, vol. 125, pp. 148–152, 2014.

[145] "Wind Power / AIL Data," 2014, [Online; accessed 28-May-2014]. [Online]. Available: http://www.aeso.ca/gridoperations/20544.html

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

[147] D. Datta and S. Dutta, "A binary-real-coded differential evolution for unit commitment problem," *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 517–524, 2012.

[148] G. Xiong, X. Liu, and T. Hashiyama, "Stochastic unit commitment problem considering risk constraints and its improved GA-based solution method," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 8, no. 5, pp. 463–469, 2013.

[149] Y. V. Makarov, Z. Huang, P. Etingov, J. Ma, R. T. Guttromson, K. Subbarao, and B. Chakrabarti, *Incorporating wind generation and load forecast uncertainties into power grid operations*. Pacific Northwest National Laboratory, 2010.

[150] R.-H. Liang and J.-H. Liao, "A fuzzy-optimization approach for generation scheduling with wind and solar energy systems," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1665–1674, 2007.

[151] T. Niknam, R. Azizipanah-Abarghooee, and M. R. Narimani, "An efficient scenario-based stochastic programming framework for multi-objective optimal micro-grid operation," *Applied Energy*, vol. 99, pp. 455–470, 2012.

[152] C. Sahin, M. Shahidehpour, and I. Erkmen, "Generation risk assessment in volatile conditions with wind, hydro, and natural gas units," *Applied Energy*, vol. 96, pp. 4–11, 2012.

[153] K. Y. Lee, Y. T. Cha, and J. H. Park, "Short-term load forecasting using an artificial neural network," *IEEE Transactions on Power Systems*, vol. 7, no. 1, pp. 124–132, 1992.

[154] "UO Solar Radiation Monitoring Laboratory," 2014, [Online; accessed 21-June-2014]. [Online]. Available: http://solardat.uoregon.edu/index.html

[155] S. Takriti, J. R. Birge, and E. Long, "A stochastic model for the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 11, no. 3, pp. 1497–1508, 1996.

[156] A. Calsetta, P. Albrecht, V. Cook, R. Ringlee, and J. Whooley, "A four-state model for estimate of outage risk for units in peaking service," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-91, no. 2, pp. 618–627, March 1972.

Appendix

The Load Data and Unit Data for the UC Test System

Table 1: Load Data for 24 Hours

Hour	**Load (MW)**	**Hour**	**Load (MW)**	**Hour**	**Load (MW)**
1	700	**9**	1300	**17**	1000
2	750	**10**	1400	**18**	1100
3	850	**11**	1450	**19**	1200
4	950	**12**	1500	**20**	1400
5	1000	**13**	1400	**21**	1300
6	1100	**14**	1300	**22**	1100
7	1150	**15**	1000	**23**	900
8	1200	**16**	1050	**24**	800

Table 2: Unit Data 1 for the 10-Unit System

	Unit 1	**Unit 2**	**Unit 3**	**Unit 4**	**Unit 5**
$P_{i,max}$**(MW)**	455	455	130	130	162
$P_{i,min}$**(MW)**	150	150	20	20	25
a_i**($/h)**	1000	970	700	680	450
b_i**($/MWh)**	16.19	17.26	16.60	16.50	19.70
$c_i(\$/MW^2h)$	0.00048	0.00031	0.00200	0.00211	0.00398
T_i^{Up}**(h)**	8	8	5	5	6
T_i^{Down}**(h)**	8	8	5	5	6
HSU($)	4500	5000	550	560	900
CSU($)	9000	10000	1100	1120	1800
T_i^{cold}**(h)**	5	5	4	4	4
$T_{i,0}$**(h)**	8	8	-5	-5	-6
FLAC	18.61	19.53	22.24	22.01	23.12

Table 3: Unit Data 2 for the 10-Unit System

	Unit 6	**Unit 7**	**Unit 8**	**Unit 9**	**Unit 10**
$P_{i,max}$**(MW)**	80	85	55	55	55
$P_{i,min}$**(MW)**	20	25	10	10	10
a_i**($/h)**	370	480	660	665	670
b_i**($/MWh)**	22.26	27.74	25.92	27.27	27.79
$c_i(\$/MW^2h)$	0.00712	0.00079	0.00413	0.00222	0.00173
T_i^{Up}**(h)**	3	3	1	1	1
T_i^{Down}**(h)**	3	3	1	1	1
HSU($)	170	260	30	30	30
CSU($)	340	520	60	60	60
T_i^{cold}**(h)**	2	2	0	0	0
$T_{i,0}$**(h)**	-3	-3	-1	-1	-1
FLAC	27.45	33.45	38.14	39.48	40.06

Lists of Publications

The publications that was published, accepted, and submitted during the course of my research are listed as follows:

Journal Publications:

[1] H. Quan, D. Srinivasan, and A. Khosravi, "Short-term load and wind power forecasting using neural network-based prediction intervals," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 2, pp. 303-315, Feb. 2014.

[2] H. Quan, D. Srinivasan, and A. Khosravi, "Particle swarm optimization for construction of neural network-based prediction intervals," *Neurocomputing*, vol. 127, pp. 172-180, Mar. 2014.

[3] H. Quan, D. Srinivasan, and A. Khosravi, "Uncertainty handling using neural network-based prediction intervals for electrical load forecasting," *Energy*, vol. 73, pp. 916-925, Aug. 2014.

[4] H. Quan, D. Srinivasan, and A. Khosravi, "Incorporating wind power forecast uncertainties into stochastic unit commitment using neural network-based prediction intervals," *IEEE Transactions on Neural Networks and Learning Systems* (Accepted),
DOI: 10.1109/TNNLS.2014.2376696, Nov. 2014.

[5] H. Quan, D. Srinivasan, and A. Khosravi, "A computational framework for uncertainty integration in stochastic unit commitment with intermittent renewable energy resources," submitted to *Applied Energy* (Under Review), Aug. 2014.

[6] H. Quan, D. Srinivasan, and A. Khosravi, "Integration of renewable generation uncertainties into stochastic unit commitment considering reserve and risk: a comparative study," submitted to *IEEE Transactions on Industrial Informatics* (Under Review), Dec. 2014.

Conference Publications:

[1] H. Quan, D. Srinivasan, and A. Khosravi, "Construction of neural network-based prediction intervals using particle swarm optimization," *in The 2012 International Joint Conference on Neural Networks (IJCNN)*, Jun. 2012, pp. 647-653.

[2] H. Quan, D. Srinivasan, A. Khosravi, S. Nahavandi, and D. Creighton, "Construction of neural network-based prediction intervals for short-term electrical load forecasting," *in 2013 IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG)*, Apr. 2013, pp. 66-72.

[3] H. Quan, D. Srinivasan, and A. Khosravi, "Incorporating wind power interval forecast uncertainties into stochastic unit commitment problems for decision making," to be submitted.

Workshop Presentations:

[1] H. Quan, D. Srinivasan, and A. Khosravi, "Construction of neural network-based prediction intervals using particle swarm optimization," *in The 2nd NUS ECE Graduate Student Symposium*, Singapore, 10-11 May. 2012.

[2] H. Quan, D. Srinivasan, and A. Khosravi, "Construction of neural network-based prediction intervals for short-term electrical load forecasting," *in The 3rd NUS ECE Graduate Student Symposium*, Singapore, 27-28 Feb. 2013.

[3] H. Quan, D. Srinivasan, and A. Khosravi, "Incorporating wind power forecast uncertainties into stochastic unit commitment using neural network-based prediction intervals," *in The 4th NUS ECE Graduate Student Symposium*, Singapore, 22-23 May. 2014.