ABSTRACT

Stochastic Dynamic Optimal Power Flow under the Variability of Renewable Energy with Modern Heuristic Optimization Techniques

Wenlei Bai, Ph.D.

Mentor: Kwang Y. Lee, Ph.D.

With the increasing penetration of renewable energy to power systems, such as wind power, more challenges have been brought to system operations due to the intermittent nature of wind. Such influence can be reflected on ancillary services of systems such as frequency control, scheduling and dispatch, and operating reserves. To tackle those challenges, wind power forecast has become an important tool.

Nowadays, forecasters typically have access to information scattered through a huge number of observed wind power time-series data from a large number of wind farms. However traditional multivariate time-series models can only process small number of data and capture only the temporal correlation in wind. In this work we utilized a probabilistic forecast model, dynamic factor model (DFM), to predict wind power. The DFM is able to capture both the spatial and temporal correlation of data, and generate as many scenarios as possible to represent the uncertainty of wind power forecast.

This work also focuses on the optimization of the system integrated with wind power and storage devices over 24 hours. Thus we formulate such problem as a stochastic dynamic optimal power flow (DOPF) problem. The essence of solving stochastic problem is to make a decision that performs well on average under almost all possible scenarios. In all, the objective functions are to optimize the expected value over all scenarios generated by DFM.

Once the stochastic optimization problem is formulated, a proper methodology is required to solve the problem. Static optimal power flow (OPF) is a highly non-linear, mixed-integer, non-convex and non-smooth problem, and traditional techniques such as nonlinear programming, quadratic programming, interior point method simplifies the problem which sacrifices the accuracy of the solution, and fails to consider the non-smooth, non-differentiable and non-convex objective functions. Therefore, to circumvent these downsides we proposed a novel heuristic method called artificial bee colony (ABC) to tackle the static OPF without approximation. In this study, the ABC has been tested on small, medium and large power system for OPF (IEEE-30, IEEE-57 and IEEE-118 buses) and then it was modified and extended to solve a dynamic optimization problem recursively.

Stochastic Dynamic Optimal Power Flow under the Variability of Renewable Energy with Modern Heuristic Optimization Techniques

by

Wenlei Bai, B.S., M.S.

A Dissertation

Approved by the Department of Electrical and Computer Engineering

_____

Kwang Y. Lee, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of

Doctor of Philosophy

Approved by Dissertation Committee

_____

Kwang Y. Lee, Ph.D., Chairperson

_____

Randall Jean, Ph.D.

_____

Robert J. Marks, Ph.D.

_____

Liang Dong, Ph.D.

_____

Qin Sheng, Ph.D.

Accepted by the Graduate School

August 2017

_____

J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABC	Artificial Bee Colony
ACE	Automatic Control Error
ACO	Ant Colony Optimization
ACOPF	Alternating Current Optimal Power Flow
CDF	Cumulative Distribution Function
CFD	Computational Fluid Dynamics
DCOPF	Direct Current Optimal Power Flow
DE	Differential Evolution
DFM	Dynamic Factor Model
DOPF	Dynamic Optimal Power Flow
ED	Economic Dispatch
EGA	Enhanced Genetic Algorithm
ERCOT	Electrical Reliability Council of Texas
ESS	Energy Storage System
FA	Factor Analysis
GA	Genetic Algorithm
GSA	Gravitational Search Algorithm
IABC	Improved Artificial Bee Colony
LDI-PSO	Linearly Decreasing Inertia Weight Particle Swarm Optimization
ISO	Independent System Operator
MAE	Mean Absolute Error
MDE	Modified Differential Evolution
MSE	Mean Square Error
MSFA	Modified Shuffle Frog Leaping Algorithm

NNs	Neural Networks
NWP	Numerical Wind Prediction
OA	Orthogonal Array
OED	Orthogonal Experimental Design
OL	Orthogonal Learning
PCA	Principle Component Analysis
PDF	Probability Density Function
PSO	Particle Swarm Optimization
RMSE	Root Mean Square Error
SCOPF	Security Constrained Optimal Power Flow
SOC	State of Charge
SVMs	Support Vector Machines
UC	Unit Commitment
VAR	Vector Auto Regression
WPF	Wind Power Forecast

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Dr. Kwang Lee, who has been a role model, consistently guiding me for the last five years with patience and wisdom. Not only his diligent and rigorous academic attitude has influenced and inspired me, but also his love for Jesus has touched me. I see the manifestation of Christ in him.

I would also give thanks to all my other committee members: Dr. Liang Dong, Dr. Robert Marks, Dr. Qin Sheng, and Dr. W. Randall Jean for their support and guidance during my proposal and research.

I would like to thank my colleague Guiying Wu, whom I often turn to when I have difficulties. She's always there. Thanks also go to Dr. Duehee Lee who provided me a great support on the dynamic factor model. Without his selfless and patient help the work cannot be done. I'd also like to thank Dr. Ibrahim Eke who was a visiting scholar to Baylor Power & Energy Systems Laboratory for 2013-2014; by collaboration we developed the heuristic method for optimal power flow. Appreciation goes to Dr. Jack Lee, who provided me an internship opportunity at the M.D. Anderson Cancer Center in 2015 summer, and Youjiao Yu who helped me on fundamentals in stochastic processes.

I want to thank my entire family members who are always there for me. Their endless support, understanding and caring carried me through my Ph.D. study.

My dear spiritual brothers in Christ, I'd like to express my great gratitude to you, Yudong Cao, Chenguang Li, Jianhui Lin, Jiancheng Lin, Daniel Wang, Daniel Wong, etc. We are, by His heavenly calling, holy brothers forever! I will miss the

fellowship/prayer/encouragement time we have had and may the anointing spirit who resurrected Jesus from death guides us on the heavenly journey. I also want to thank all the brothers and sisters in Waco Chinese Church who have been helping me in every aspect of life: cooking for me, cleaning for me, giving me rides, and challenging my spiritual growth. I will miss and never forget the time together with you. Missionary Tan Sheng, Rebecca and their son Enoch, I thank you for sharing the Gospel to me. Your pure heart and willingness to sacrifice always inspires me. May God remember and fortify your work.

Last but not least, the work is dedicated to my dear Lord Jesus who is holy, righteous, just, and yet gracious, merciful, caring and loving.

Lord, may you come, Maranatha!

To my dear Lord Jesus,
Who saved me out of darkness.

To my Mom,
Who taught me love and endurance.

To my Dad,
Who taught me to be diligent.

是爱的神，作我牧人，
The God of love, my Shepherd is,

他常喂养，他常施恩；
Grace He doth give, me He doth feed,

他是属我，我是属他，
O He is mine and I am His,

何来需要？何来缺乏？
What else I lack? What else I need?

CHAPTER ONE

Introduction

1.1 Wind Energy Integration

Wind is providing one of the cheapest and cleanest sources of electrical energy. Installed capacity of wind power plants has increased from1.5 GW to 26 GW in 20 years in USA and near $50 billion has been invested on wind energy in this period of time [1]. In 2010 for the first time ever, the installed capacity of new wind power plants in developing countries became more than developed countries [2]. It shows that wind power is not a technology only for developed countries, which cannot be deployed in other places. It also implies that wind power is getting more and more economic, and that in addition to its environmental advantages, more countries install wind units because the cost of these units is constantly falling.

Applications of wind power for producing electrical energy were studied in the literatures at early 19th century [3] [4]. After the oil crisis, interest in the power of the wind re-emerged. Capacity of installed wind power plants increased year after year and more researchers started to work on wind power and its unique issues. Generally, researches on wind power can be categorized into five different categories; i) system operation with high wind penetration issue, focusing on increasing system's reliability by advanced forecasting methods and operational tools, ii) wind turbine issues, concentrating on mechanical aspects of wind power, iii) power generator types, focusing on different types of generators employed to convert the mechanical power into the

electrical, iv) integration issues, seeking best topologies to connect the wind generator to the network using converters/capacitor banks/soft-starters, and v) control methods, focusing on different control strategies including pitch-angle control and converter control.

1.2 Modern Heuristic Techniques

Modern heuristic techniques have been developed and drawn researchers' attention in recently years because of their advantages such as ease to implement, robustness, and efficiency to handle non-convex, non-linear, discontinuous and complex problems [5]. These techniques include evolutionary computation, simulated annealing, Tabu search, particle swarm optimization, and so forth. Reports of applications of these tools have been widely published, including in power system operation area [6]-[8]. This section briefly introduces fundamentals of these techniques.

Evolutionary computation is a major technique in modern heuristic optimization methods. Evolutionary computation is a term to describe any concerns of evolutionary algorithms, which provide practical advantages for difficult optimization problems [5]. Natural evolution is a hypothetical population-based optimization process. Simulating this process on a computer results in stochastic optimization techniques which can often outperform classic methods of optimization. Evolutionary algorithms include evolutionary strategies, evolutionary programming, genetic algorithm, differential evolution, etc. One of the main advantages in evolutionary computation is that it uses a simple concept, as demonstrated in Fig. 1.1. The algorithm consists of initialization, which may be a purely random sampling of possible solutions. Then, those possible solutions are updated before evaluating their "fitness's". The fitness values indicate

"good" or "bad" solutions. Later, better solutions will be chosen. Such process does not require gradient information which may be required of in other conventional algorithms. Over the iterations of random variation and selection, the population can be made to converge [9],[10].

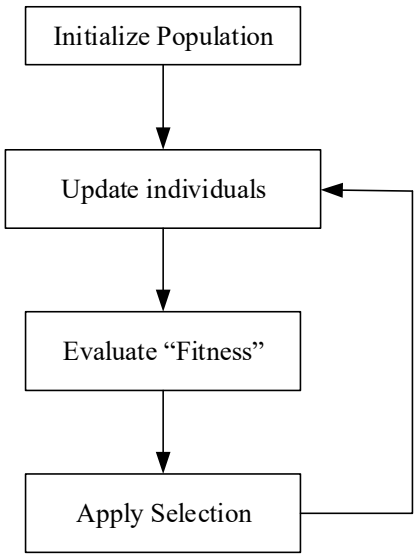

Figure 1.1. Steps of a typical evolutionary algorithm.

Genetic algorithm (GA) is a search algorithm based on the analogy of natural selection and genetics. The features of GA are different in several aspects from other heuristic search algorithms: i) GA is a multipath that searches many peaks in parallel, hence reducing the possibility of trapping in local minima. ii) GA adopts coding and decoding schemes for parameters, which will help the GA to evolve the current state into the next state with minimum computations. iii) GA uses the roulette wheel selection scheme to improve solution. Such scheme enables improved performance with high probability.

Another category of heuristic methods are called swarm intelligence techniques, such as particle swarm optimization (PSO), ant colony optimization (ACO), artificial bee

colony (ABC), etc. Those algorithms mimic the behaviors of natural creatures when they form into a swarm. Colorni, Dorigo, and Maniezzo developed ACO based mainly on the social insect, especially ant metaphor [11]. Each individual exchanges information through pheromones implicitly in ACO. Eberhart and Kennedy developed PSO based on the analogy of swarms of birds and fish schooling [12]. Each particle remembers its best route so far, called *personal best*, and meanwhile recognize neighbors' best route, called *global best*. Then, each particle will be updated based on personal best and global best information till they search the optimum of certain objective functions. The particle updating scheme is defined as:

$$\begin{aligned} v_i^{t+1} &= wv_i^t + c_1 rand_1 \times \left(P_i - x_i^t\right) + c_2 rand_2 \times \left(P_g - x_i^t\right) \\ x_i^{t+1} &= x_i^t + v_i^{t+1} \end{aligned} \tag{1.1}$$

where v_i^t is the velocity of the *i*-th particle at iteration *t*, *w*, c_1 and c_2 are weight coefficients, r*and*$_1$ and *rand*$_2$ are two random numbers between 0 and 1, x_i^t is the current position of particle at iteration *t*, P_i is the best personal solution the *i*-th particle has memorized and P_g is the global best of the whole group. The velocity is updated incorporating the information of personal best and global best. The modification of position is demonstrated in Fig. 1.2, which is based on equation (1.1).

1.3 Operational Challenges in Wind Integrated System

Driven by increasing prices for fossil fuels and concerns about emissions, the wind energy, which is environment-friendly and costless, is rapidly penetrating into current power systems. Such penetration introduces more variability and uncertainty due to the intermittent nature of the wind, which cause significant impacts/challenges in

power system operations [13]. The impacts of wind power in the electric power system depend on a large extent on the level of wind power generation, grid size and types of power generation in the system. The established control methods and available system reserves for dealing with variable demand and supply are adequate for dealing with the additional variability at the wind energy penetration level up to 20 percent, depending on the nature of a specific system. For higher penetration levels, additional changes to operation may be required to accommodate wind energy.

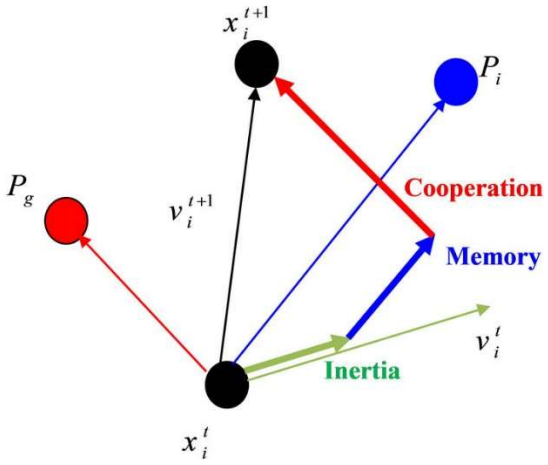

Figure 1.2. Description of velocity and position updates in PSO [12].

1.3.1 Impact of Wind Generation on Automatic Generation Control

With the expansion of installed wind power capacity, corresponding regulating reserve will be required to accommodate both load uncertainty and high wind power variations. Conventionally, the fast-responding automatic generation control (AGC) units will respond to automatic control error (ACE) signal to regulate frequency whenever there is an imbalance between load and generation. The ACE signal arises due to the imbalance of load and generation. The information regarding frequency and tie-line

power flow errors is contained in the ACE signals [14]. Based on the premise that inter-area (different areas) effects are normally weaker than intra-area (same area) effects, all generators within an area are assumed to be synchronized in the same frequency. However such assumption may not be carried to the system where large wind penetration is imposed, because one may observe potentially larger imbalances at locations where installed wind capacity is high. Consequently, frequency is spatially differentiated even in the same area. Such spatial variation requires the allocation of regulating reserve, unbundling of control input down to the generator level. In all, the conventional simple control by AGC will not meet the control requirements. There have been studies on the impact of wind integration on conventional AGC in real-world systems. For instance, the work in [15] shows that the integration of large-scale wind generation requires novel frequency regulation and load following mechanisms for the California Independent System Operator (ISO).

1.3.2 Impact of Wind Generation on Unit Commitment

The Unit Commitment (UC) is the decision process to determine which units should be turned on at which generation level [16]. Since wind power is less predictable than demand, thus the integration of significant wind power requires UC to be carried out more frequently, preferably each time new wind power forecast are available, for example, every 6 hour. It is obvious that the commitment decisions are very sensitive to wind power forecast when there is considerable wind penetration in the system. Neglecting wind power forecast is a decision error that may lead to unnecessary cost. Bad forecast may lead to even worse case than no forecast [17]. One example can illustrate such impact. If an independent system operator (ISO) makes a day-ahead UC considering

no wind power available and this leads to operating certain number of coal-fired power plant. However, there is plenty of wind power available the following day, and the system ought to maximize using wind power according to today's practice. Since the coal-fired power plant cannot reduce the output fast enough, coal-fired units will be turned off and fast-response and yet expensive gas-fired power plants need to be started. The whole process would lead to the operational cost. The discussion above concludes that the main challenge for UC is the forecast uncertainty. A number of possible solutions have been proposed in the literature. For example, in [18] and [19] stochastic UC is formulated under large wind penetration. Stochastic UC requires the UC solution be robust for accommodating low actual wind power when wind was expected high and vice versa. Robustness means that the UC is flexible enough to accommodate wind variations, without emergency commitment or emergency de-commitment.

1.3.3 Impact of Wind Generation on Economic Dispatch

Once the UC has decided which units to turn on, normally one-day ahead, economic dispatch (ED) will take place every 10-15 min to adjust the output of committed units due to the load and wind variations. The complexity in ED arises from various factors: technical, economic, and environmental policies. In [20], the wind forecast in reducing the system-wide emission as well as generation cost is quantified assuming static dispatch. Another key factor that complicates the operation is rate of response of various resources. For example, coal-fired unit is slower than gas-fired turbines. The work in [21] shows that a look-ahead ED approach, which considers wind power dispatchable, reduces the cost than that of static ED, which treats wind power as negative load by explicitly accounting for the rate of response of different types of

generators. Despite the fact that seemingly “free energy” is not fully utilized, the overall system generation cost is reduced. This is because the fact that by smoothing out high inter-temporal changes from the wind generation, the generation from the fast responsive and yet expensive unit is also reduced.

Obviously, should a suitable wind power forecast model be developed, the risk and challenges in system can be significantly reduced. In order to further reduce the negative impacts of wind forecast due to its uncertainty, a proper stochastic optimization dispatch scheme is in urgent need. Driven by the operational challenges due to wind and the advantages of modern heuristic techniques, this work is proposed in the dissertation. The following lists the scope of the dissertation.

1.4 Scope of the Dissertation

The outcome of the proposed research is highlighted as following:

1. Developed a wind power forecasting model, dynamic factor model (DFM), which is able to process large multi-dimensional data with less computational burden. The model adopts the principle component analysis technique.
2. By different dynamic shocks, the DFM is able to generate various scenarios which represent the uncertainty of forecast. Such scenarios are of great interest for stochastic optimization when considering power system with high wind penetration.
3. Adopted the modern heuristic technique, artificial bee colony (ABC), as the methodology to solve optimal power flow problem. Since OPF is highly non-linear and non-convex, ABC is more efficient and flexible to tackle such problem.

4. Improved ABC based on orthogonal learning such that the exploration and exploitation of the original ABC has been improved into good balance. The original ABC is good at exploration and yet not sufficient in exploitation.
5. Proposed a methodology for dynamic OPF over 24 hour horizon. The methodology is built based on the ABC, solving static OPF recursively with additional dynamic constraints.
6. Proposed a stochastic optimization to optimize the expected cost function under all available scenarios. The uncertainty arises from the wind power forecast and is represented by scenarios generated from the DFM.

CHAPTER TWO

Wind Power Forecast and Power System Operations

2.1 Wind Power Forecast Definitioin

The forecasted wind power at time *t*, for a look-ahead time, *t+k*, is denoted by $\hat{P}_{t+k|t}$ meaning that the wind farm is expected to produce such power during the time resolution of the forecast (e.g., 1 hr.) given constant wind speed. Time horizon, *T*, is the total forecast period in the future (e.g., 24 hrs.). It is worthy to mention that the forecasted power $\hat{P}_{t+k|t}$ is the point forecast, whereas another type of forecast is called probabilistic forecast. Details will be discussed in the following content.

A forecasting system is characterized by its time horizon, which is the future time period for which the wind power will be predicted (e.g., the next 1 hour). The forecasting system usually is characterized in accordance with its time horizon – very short term, short term, medium term, or long term. In wind power forecasting (WPF) problem, generally time horizon can be categorized into three groups: 1) *Very short term*, where the time horizon is a few hours. A limit value of 4 hours for this term is proposed in [22]. 2) *Short term*, where the time horizon rises from the very short term to 48 or 72 hr. Application in this horizon is the trading in one day-ahead market. 3) *Medium term*, where it ranges from the short term to 7 days. As the forecasting horizon increases, the forecast error increases as well [23]. Applications of using this horizon in forecasting are the unit commitment (UC) of the conventional generation, and maintenance planning [24].

Advanced forecasting methods are generally divided into two main groups. The first one is called physical approach, and second is statistical approach. The dynamic factor model proposed in the thesis is a time series statistical model.

2.1.1 Physical Approach

The physical approach model takes the physical phenomena of the wind flow and terrain of wind farm into account, plus using the wind-turbine manufacturer's power curve to predict the wind power. The numerical wind prediction (NWP) provides the atmospheric information over certain area [25]. In order to obtain the detailed characterization of the weather variables in wind farms, physical approach has to downscale the wind speed and direction from NWP to the turbine hub's height [26][27]. The main procedure is depicted in Fig. 2.1.

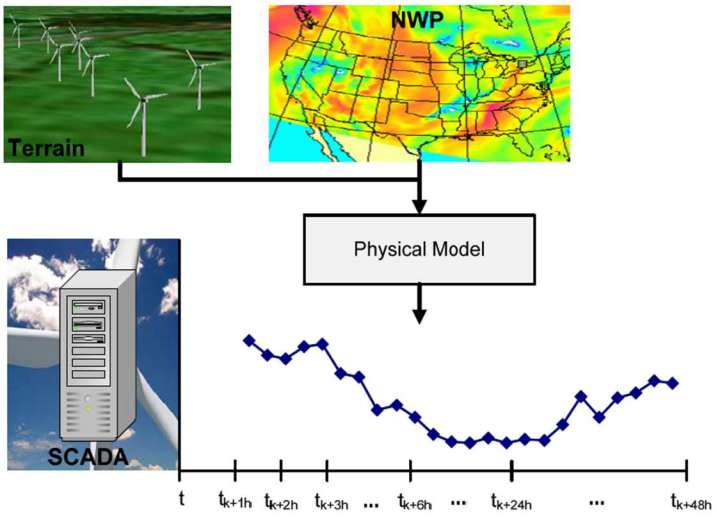

Figure 2.1. Physical Approach Structure [32].

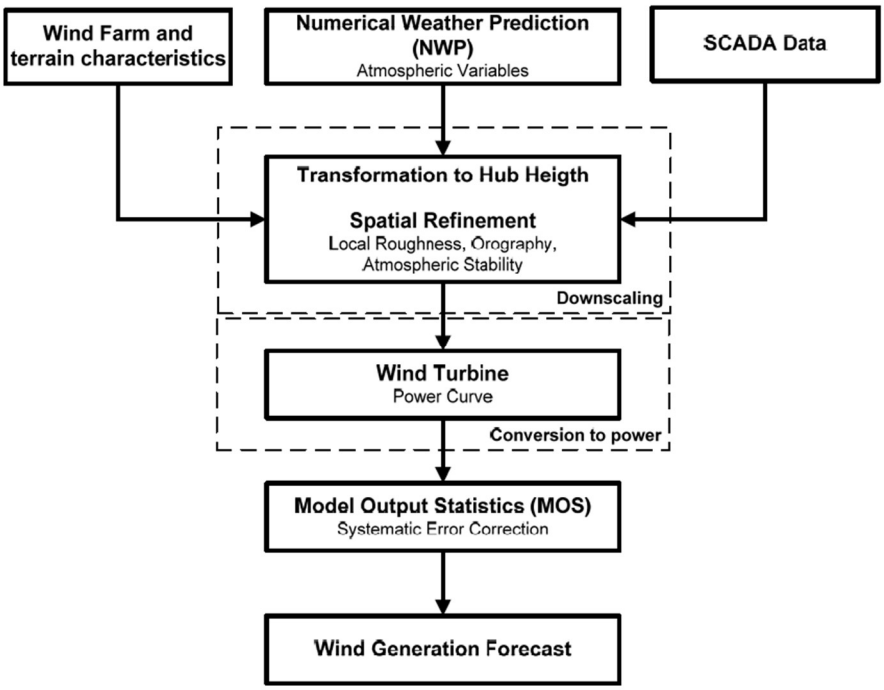

Figure 2.2. Main steps of physical approach for wind forecasting.

As shown in Fig. 2.2, the main two procedures in physical approach are downscaling and conversion to power. There are typical two methods for downscaling: 1) combine the modeling of the wind profile and the geostrophic law to obtain surface winds [26], 2) model a full description of the terrain (local roughness, orography, obstacles etc.) by computational fluid dynamics (CFD) to obtain a prediction of local wind regime [28]. When it comes to converting data into wind power, the manufacturers' power curve is needed.

2.1.2 Statistical Approach

There is an alternative approach, statistical approach, for wind forecasting. Such approach directly transforms the input data into wind generation by only one step, as shown in Fig. 2.3. For instance, the statistical model block in Fig. 2.4 is able to combine

the variables from NWPs such as wind speed, direction, temperature, etc., with on-line measurements from SCADA data such as wind power, speed, direction, etc. By considering these inputs, a direct estimate of regional wind power can be obtained by statistical model.

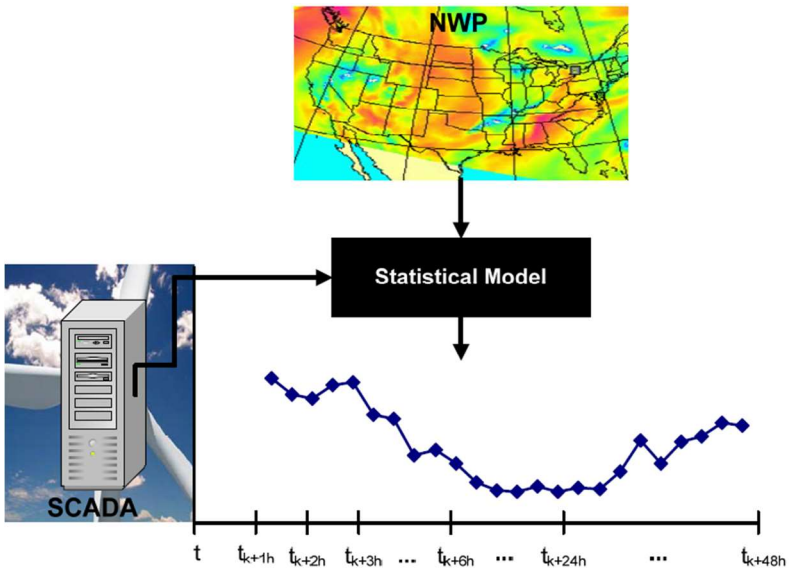

Figure 2.3. Statistical approach structure [32].

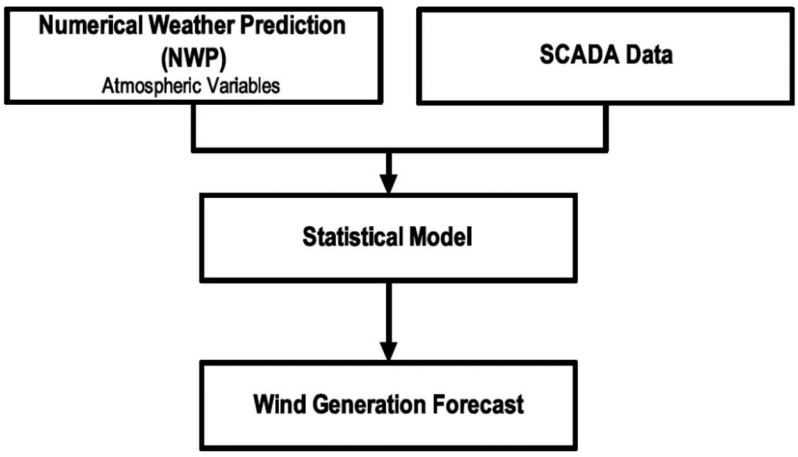

Figure 2.4. Main steps of the statistical forecasting approach.

There are different types of statistical models. For example, artificial intelligence based models, such as Neural Networks (NNs) and Support Vector Machines (SVMs) are considered as nonlinear "black-box" models. There are also models that can be expressed analytically, such as time-series model, and Kernel regression. In all, the statistical model can be expressed mathematically as:

$$\hat{P}_{t+k|t} = f(P_t, P_{t-1}, \cdots, P_{t-n}, x_t, x_{t-1}, \cdots, x_{t-n}, \hat{x}_{t+1|t}, \cdots \hat{x}_{t+k|t}) \tag{2.1}$$

where P is the wind power, x is the explanatory variable, $\hat{p}_{t+k|t}$ is the k-step ahead forecast power, and $\hat{x}_{t+k|t}$ is the k-step ahead forecast explanatory variable. In other words, the statistical model is a function of past values of p and a set of past values and forecasts of the explanatory variables x.

The dynamic factor model (DFM) presented in this work is a multivariate time-series model which takes into account the multiple wind farms as input. Such model considers both the spatial and temporal correlation of input data which provides more information of the data and thus results in more practical and precise forecast [29]. The details of DFM is found in Chapter three.

2.2 Evaluation of Forecasts

The WPFs are inherently uncertain because no forecast model can be perfect. Thus the evaluation of the forecasts become essential. Practitioners should not only be able to assess the performance of the forecasts but also to understand what factors influence the prediction uncertainty. Evaluation of the quality of forecasting methods is conducted by comparing wind power predictions directly with the actual corresponding observations. This section describes a framework of forecast evaluations.

2.2.1 Standard Error Measures

The prediction error at a given time $t+k$ is defined as the difference between the forecast and observation:

$$e_{t+k|t} = P_{t+k} - \hat{P}_{t+k|t} \tag{2.2}$$

where $e_{t+k|t}$ is the forecast error corresponding to time $t+k$ for the prediction made at time t. The forecast error is often normalized by the wind farm installed capacity in order to compare results with other wind farms regardless of the difference in wind power capacity:

$$e_{t+k|t} = \frac{e_{t+k|t}}{P_{cap}} = \frac{1}{P_{cap}}\left[P_{t+k} - \hat{P}_{t+k|t}\right] \tag{2.3}$$

where P_{cap} is the wind power capacity. Any prediction error can be decomposed into two components: systematic error and random error. The systematic error should be zero for a perfect model while random error can be modelled as a series of independent random variables of Gaussian distribution [30].

There are several error measures defined as followings to assess the quality of forecasting methods. The bias error, $BIAS_k$, which corresponds to an estimate of the systematic error that is provided by the mean error over the whole evaluation period:

$$BIAS_k = \frac{1}{N}\sum_{t=1}^{N} e_{t+k|t} \tag{2.4}$$

where N is the number of prediction errors used for method evaluation. The $BIAS_k$ provides an indication of whether the method tends to overestimate or underestimate the forecasted variable. However it is very unlikely that a forecasting method with a zero $BIAS_k$ will result in perfect predictions because the $BIAS_k$ could cancel out positive and

negative error values along the prediction horizon. Thus the mean square error (MSE) is defined to identify the contribution of both positive and negative errors to a forecasting method:

$$MSE_k = \frac{1}{N}\sum_{t=1}^{N} e_{t+k|t}^2 \tag{2.5}$$

where N is the number of prediction errors used for method evaluation. There are two other basic criteria for performance evaluation: the mean absolute error (MAE) and the root mean square error (RMSE):

$$MAE_k = \frac{1}{N}\sum_{t=1}^{N} \left| e_{t+k|t} \right| \tag{2.6}$$

$$RMSE_k = \sqrt{MSE_k} = \sqrt{\frac{\sum_{t=1}^{N} e_{t+k|t}^2}{N}} \tag{2.7}$$

Similar to the MAE, both systematic and random errors affect the RMSE criterion.

2.2.2 Comparison of Forecasting Methods

Different forecasting methods are usually made comparison by various criterion because one method may work best at certain criterion while less effective at other criterions. The performance of forecasting methods not only depends on the variance of errors but on the evaluation period because certain methods are designed specifically for different prediction horizons. To compare the performances of various methods, the "skill score" which is defined as the improvement relative to the reference is an important indication [31]:

$$IMP_k = \frac{EvC_k^{ref} - EvC_k}{EvC_k^{ref}} \times 100\% \tag{2.8}$$

where the *EvC* is the selected evaluation criterion, such as MAE, RMSE, or their normalized version, etc. Positive value indicates the advanced approach is better than the reference model and negative values imply that advanced approach performs worse.

2.3 Uncertainty in WPF

Short-term forecasting tools that are currently widely in use can provide single-valued point forecasts. However, the main drawback of point forecasts is that no information is provided on the dispersion of observations around the predicted value. Additional information on the uncertainty associated with future wind power predictions is required. Recent research has focused on associating uncertainty estimates with point forecasts to become probabilistic forecasts or scenarios of wind power.

Probabilistic forecasting consists of estimating the future uncertainty of wind power that can be expressed as a probability measure. The forecasted wind power can be described as random variables in probability density function (PDF), cumulative distribution function (CDF), moments of distributions (mean, variance, skewness), or a set of percentiles. In all, the PDF are the most fundamental representation for uncertainty because other forms can be deduced from PDF. For example, let f_{t+k} be the PDF of P_{t+k} (wind power for look-ahead time $t+k$) and let F_{t+k} be the consequent CDF. Since F_{t+k} is a monotone increasing function, the percentile q_{t+k}^{α} with proportion $\alpha \in [0,1]$ of the random variable P_{t+k} is uniquely defined as the value x, such that $prob(P_{t+k} < x) = \alpha$, or in another form is defined as $q_{t+k}^{\alpha} = F_{t+k}^{-1}(\alpha)$, and $\hat{q}_{t+k}^{\alpha}$ is denoted as an estimate of percentile q_{t+k}^{α} at time step t for look-ahead time $t+k$.

Percentiles are then used to build intervals that provide a range of possible values where the observed value is expected to lie within under certain confidence level. This type of probabilistic forecast is called interval forecasts [32]. The interval forecast is defined by its lower and upper bounds, which are two forecasted quantiles. Figure 2.5 gives an example of interval forecasts with different intervals.

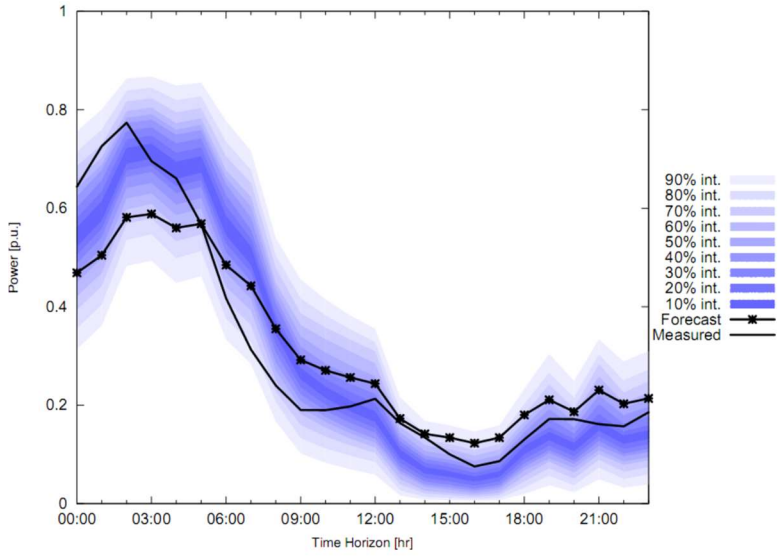

Figure 2.5. Interval forecasts [32].

The common representation is to center the intervals on the median. Thus, the probabilities are symmetric around the median. However, the distances are not symmetric. For example, if the first quantile is 500 MW with $\alpha = 35\%$, and the second is 1,800 MW with $\alpha = 65\%$ around a median of 800 MW, the corresponding interval is [500, 1,800], with a coverage rate of 30% and an amplitude of 1300 MW. The distance to each side is not the same: 300 MW for first quantile bound and 1000 for the other quantile bound.

Note that since the error distribution in WPF is skewed and heavy-tailed, the forecasted distribution of wind power output might also be asymmetric. Therefore, it is

uncommon to center the intervals on the mean or on the point forecasts because, in asymmetrical distributions, the mean and the median may be very different. In this situation, the point forecast may not lie inside an interval with low coverage rate.

The probabilistic forecasts are produced for each look-ahead time independently. Therefore, they are not able to develop the forecast uncertainty over the forecasted time series due to the fact that the probabilistic forecasts do not consider the temporal correlation in the forecasted time series. However such temporal correlation over the time horizon is a valuable information in time-dependent decision-making problems such as the unit commitment under large wind power penetration. A scenario generation method described in [34] is used to generate a number of wind power scenarios that provide information on forecasting uncertainty through the set of look-ahead times. It is worthy to mention that there is no standard or systematic way to evaluate probabilistic wind power forecast because an individual probabilistic forecast cannot be determined as incorrect. For example, a probabilistic forecast states that the expected power generation for a given horizon is between 1 and 1.6 MW with 40% probability, and the actual wind power is 0.8 MW. Since the probabilistic forecast only covers 40% of the cases, we cannot say the forecast is incorrect. Evaluation schemes can be defined uniquely from case to case based on applications. For example reference [33] have evaluated the probabilistic forecast models based on one-day ahead economic dispatch or power system operating cost.

CHAPTER THREE

Dynamic Factor Model for Wind Forecast

3.1 Dynamic Factor Model

With the increasing penetration of renewable energy to power system, more challenges have been brought to system operations because of the intermittent nature of such energy [35]. Wind energy has great impact on system's stability and reliability because the wind power is highly uncertain and unpredictable. Such influence can be reflected on ancillary services of systems such as frequency control, scheduling and dispatch, and operating reserves. There are two major approaches in wind power forecast in terms of output: point forecast and probabilistic forecast. The point forecast gives a single value for future wind power, which usually is an estimate of conditional mean of wind power. However probabilistic forecast is of more interest because it provides a complete characterization of the conditional distribution of future wind power [36],[37]. In other words, it considers the wind power's volatility. Extensive researches on probabilistic forecast have been conducted due to the fact that such forecast can provide the information of uncertainty from the forecast, which is of importance for system operators to make decisions [38]. The uncertainty representation is in the form of quantiles, interval forecasts, probability density function (pdf), and scenarios, etc.

Reference [37] gives a comprehensive review of the methodologies for wind power forecast which include physical models, statistical models and combined models. Traditional time-series model (statistical model) and autoregressive model (AR), captures

the temporal correlation in wind. The model predicts the future wind power as the linear combination of current and past data plus a white noise error [39]. Advanced statistical model, space-time model, is proposed in [40]. The model considers the terrain, wind direction, wind speed as the input data to the model and the quality of forecast is assessed by economic dispatch model simulating the northwest region of America and such forecast model implies cost saving compared with AR and persistence models. However few of the aforementioned works have focused on the stochastic process with large number of multivariate, even though the spatial correlation among wind farms has placed an important factor on power system reliability. For example, wind farms with positively correlated output power and sharing the same transmission lines might cause congestion which leads to curtailment [41].

As discussed above, scenario synthesis has evolved from point values to a single time-series; however the correlation among time-series has not been systematically implemented. In this work we proposed a novel multivariate time series model, DFM, to address the correlation among wind farms in nearby area where wind power is affected by similar weather condition and the correlation inherited in time domain. Thus it is also a spatio-temporal model. The model is a multivariate stochastic process because it includes 96 wind farms time-series data as input. Factor analysis (FA) is a statistical tool to reduce data dimension by describing the observed correlated variables in terms of the unobserved/latent variables, called factors. The observed variables are modeled as linear combinations of the latent factors. DFM utilizes factor analysis concept to reduce the data dimension such that the computational burden of regular multivariate time-series analysis can be reduced significantly. Thus the observed data can be decomposed by the

multiplication of factor loadings and dynamic factors [42]-[44]. In addition, by DFM various forecast scenarios/realizations can be generated to represent the uncertainty because the model is essentially driven by uncorrelated dynamic shocks.

In all, the advantages of DFM are 1) an arbitrary number of wind power scenarios can be generated; 2) all scenarios have the similar spatial and temporal correlation, statistical characteristics as the actual observations do; 3) computational burden can be reduced significantly since by factor analysis much less dimensional variables are required than full vector autoregressive model (VAR).

3.1.1 Data Description

Wind power observation data of total $N = 96$ wind farms from EROCT were used. The hourly data were from January 1st to March 31th in 2013 for total $T = 24 \times 90$ hours. Total wind power capacity is 10,407 MW. Thus the observation data set A is a $N \times T$ matrix as shown in Fig 3.1. The forecast period is the following day of the historical data, which is April 1st, 2013.

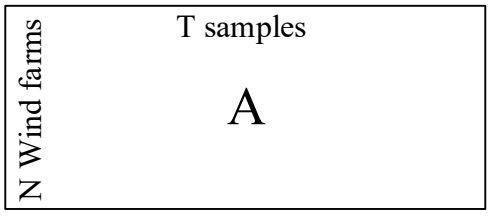

Figure 3.1. Observation data set A.

The original data is preprocessed by the followings: 1) subtract the average value of every wind farm's data from the original data; 2) divide the residuals by the standard deviation of every wind farm; 3) extract the periodic pattern of data. Steps 1 and 2 are the

standardizing process and step 3 is the process of removing seasonality. The periodic component is estimated by averaging data for the same hour every day: the observation set $A \in \Re^{96\times(24\bullet 90)}$ is reshaped into $Y \in \Re^{96\times 24\times 90}$ first, which is a three dimensional matrix (90 number of 96×24 matrix), then the mean of each hour over 90 days is stacked into $Z \in \Re^{96\times 24}$, and finally the diurnal pattern was subtracted from hourly data from the observations. Note that after synthesizing scenarios using residuals, the reversed process will be applied in order to obtain final scenarios. The process is defined in the following equations:

$$\begin{aligned} A &= P + X_R \\ X_R &= \chi + \xi \end{aligned} \tag{3.1}$$

Where A is the observation data set and is decomposed into components: periodic data P, common component $\chi \in \Re^{N\times T}$ and idiosyncratic component (unpredictable data) $\xi \in \Re^{N\times T}$; X_R is the residual signals after subtracting periodic data from the observation data A. Then X is normalized and standardized by

$$X = \frac{X_R - mean(X_R)}{std(X_R)} \tag{3.2}$$

3.1.2 Derivation of DFM

Factor analysis model has been applied widely in energy, load and economic forecasting areas [45]. Such model is used to find the unobserved correlated latent variables (factors) out of observations. Factor analysis (FA) is a statistical tool to reduce data dimension. The following example illustrates the FA process of reducing the observed variables of school courses by latent variables.

Five students were collected as samples in Table 3.1, and their course grades on four different subjects were observed variables. Some students are good at arithmetic courses, and some are good at linguistic courses. The arithmetic ability and linguistic ability are not observed; however we can extract such information.

Table 3.1. Sample data of four subjects

Student Number	Grade			
	Math	Literature	Spanish	Physics
1	100	44	37	93
2	80	72	71	79
3	50	50	50	50
4	20	76	83	27
5	10	82	91	19

Table 3.2. Ability loading

Subject	Ability Loading	
	Arithmetic	Linguistic
Math	1	0
Literature	0.2	0.8
Spanish	0.1	0.9
Physics	0.9	0.1

Table 3.3. Unobserved Variables

Student	Ability	
	Arithmetic	Linguistic
1	100	30
2	80	70
3	50	50
4	20	90
5	10	100

As shown in Tables 3.2 and 3.3, the arithmetic ability and linguistic ability are loaded differently with respect to course. In all, the observed data grade can be found by the multiplication of ability loading and ability matrix, which are latent variables:

$$\text{Grade}^{T}_{(4\times5)} = \text{Ability Loading}_{(4\times2)} \times \text{Ability}^{T}_{(2\times5)}$$

In this work, the latent variables can be wind speed, air density, wind direction, etc. Evolved from static factor model, the dynamic model not only considers the spatial correlation (among various wind farms) but the temporal correlation as well. "Dynamic" means that there is time lag in factor loadings. Same as for time-series model, the assumption for DFM is that the data is wide-sense stationary, meaning that the auto-covariance function is only dependent on the time-lag and the whole stochastic process has constant expected value and variance.

After aforementioned preprocessing of observation data, the residual data matrix $X \in \Re^{N\times T}$ can be decomposed into two components, the common component $\chi \in \Re^{N\times T}$ and the idiosyncratic component $\xi \in \Re^{N\times T}$ [46], where N is the total number of wind farms over T hours. The column vector $X_t = \{x_{1t}, x_{2t}, \cdots, x_{Nt}\}$ represents the wind power of N wind farms at time t, where $t = 1, \cdots, T$:

$$X_t = \chi_t + \xi_t \tag{3.3}$$

The goal is to synthesize the common component by dynamic factor model. It is assumed that $\chi \in \Re^{N\times T}$ can be decomposed into the multiplication of factor loading $A(L)$, an $N\times Q\times M$ three dimensional polynomial matrix which can be seen as M number of $N\times Q$ matrices, and the dynamic factor $f \in \Re^{Q\times T}$ as:

$$\chi_t = A(L) f_t = A_0 + \cdots + A_M f_{t-M} \tag{3.4}$$

where S is the lag/lead steps, $A(L)$ is a notation for polynomial matrix: $A(L) = A_0 L^0 + \cdots + A_M L^M$, and L^M is the delay operator meaning that f_t is delayed by M, $\boldsymbol{f}_{t-M}$. Since the derivation starts from the residuals $\boldsymbol{X} \in \Re^{N\times T}$, it is reasonable to assume

that dynamic factors can be obtained through filter $B(L)$, an $Q\times N\times M$ three dimensional polynomial matrix similar as $A(L)$:

$$f_t = B(L)X_t \tag{3.5}$$

Furthermore, dynamic factors f is a time series structure; thus vector autoregressive (VAR) model can be used to model f:

$$\varepsilon_t = C(L) = C_0 f_t - \cdots - C_R f_{t-R} \tag{3.6}$$

where ε_t is the column vector of white noise $\varepsilon \in \Re^{Q\times T}$ and $C(L)$ is the coefficients Q×Q×R three dimensional polynomial matrix: $C(L) = C_0 L^0 + \cdots + C_R L^R$. Another way to look at equation (3.6) is that dynamic factors f is driven by noise $\boldsymbol{\varepsilon}$; however the noise term has the same spatial correlation as in the actual observation data, while the goal here is to formulate a series of uncorrelated noise to drive f so that as many scenarios can be synthesized. Guided by this idea, we extract the correlation structure from $\boldsymbol{\varepsilon}$ by Cholesky decomposition [41]:

$$\varepsilon_{\mathrm{t}} = H\delta_{\mathrm{t}} \tag{3.7}$$

where $\delta \in \Re^{Q\times T}$ is a series of spatially and temporally uncorrelated white Gaussian noise with zero mean and variance one. H is the $Q\times Q$ matrix that preserves the correlation structure. There is one additional assumption in DFM: χ and ξ are uncorrelated as:

$$E[\chi_j \xi_k^T] = 0 \tag{3.8}$$

where $j,k = 1,\cdots,N$.

Combing equations (3.4) (3.5) (3.6) and (3.7), the dynamic factor is represented as:

$$\chi_t = A(\mathrm{L})\left[C(\mathrm{L})\right]^{-1} H\delta_t \tag{3.9}$$

which leads to the comprehensive form of DFM. Then the question of how to estimate the polynomial matrices $A(L)$, $B(L)$, and $C(L)$ arises naturally, and the procedure is discussed in the following. In all, the DFM can be summarized in Figs. 3.2 and 3.3.

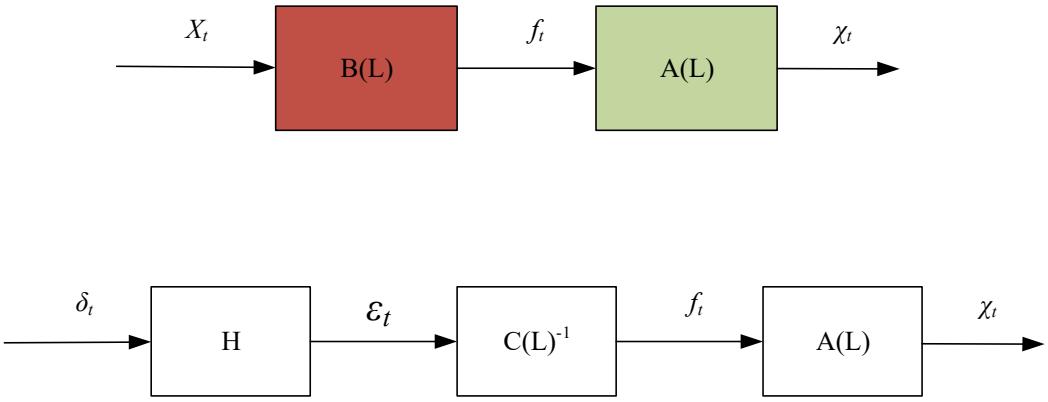

Figure 3.2. Block diagram for DFM (Top: estimating A(L) and B(L); Bottom: generating scenarios).

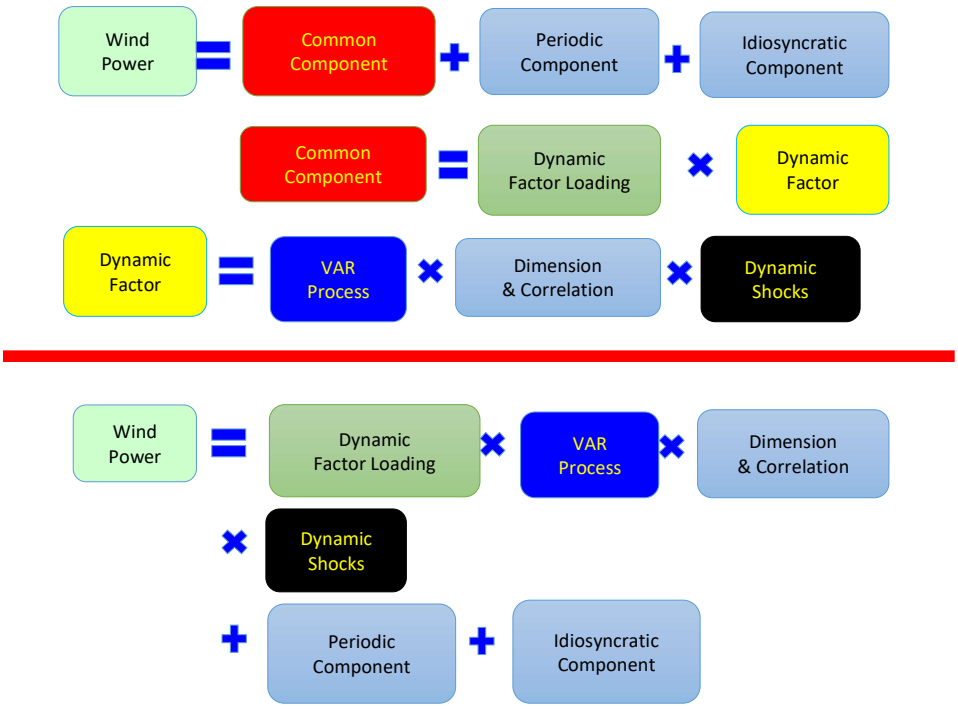

Figure 3.3. DFM decomposition components.

The common component of wind power consists of deterministic components (factor loadings and VAR process) and stochastic component (random dynamic shocks).

3.1.3 Estimation Parameters for DFM

Previous section has raised a question of how to estimate $A(L)$, $B(L)$ and $C(L)$, before beginning the derivation, several fundamental concepts need to be introduced.

Auto-correlation (in stationary time series) is defined as:

$$R_{X1,X1}(k) = \frac{\mathrm{E}\{(X_1[t]-\mu)(X_1[t-k]-\mu)\}}{\sigma^2} \tag{3.10}$$

where $R_{X1,X1}$ is the auto-correlation function of one random sequence X_1. E is the expected value operator and k is the time lag. Note that since after normalization and standardization the data has zero mean, $\mu = 0$ and variance $\sigma^2 = 1$, thus the auto-correlation function is equivalent with auto-covariance function, which is defined as:

$$R_{X1,X1}(k) = \mathrm{E}\{X_1[t]X_1[t-k]\} \tag{3.11}$$

The cross-covariance is defined as:

$$\begin{gathered} R_{X1,X2}(k) = E\{X_1[t+k]X_2[t]\} \\ \Downarrow \\ \hat{R}_{x1,x2}(k) = \frac{\sum_{t=1}^{T}(x_1[t+k]x_2[t])}{T-|k|+1} \qquad \text{unbiased estimation form} \end{gathered} \tag{3.12}$$

where $R_{X1,X2}(k)$ is the cross-covariance function of two random sequences X_1 and X_2 with k lag, E is the expected value, k is the time lag, and x_1 and x_2 are the one realization (observation) of X_1 and X_2 random sequences respectively. Note that we can only estimate such random sequences because in practice, only a finite segment of the realization of the infinite-length random process is available. Unless otherwise,

correlation and covariance functions are treated the same in this work because after normalization and standardization the data has zero mean, $\mu = 0$ and variance $\sigma^2 = 1$.

Since there are total N signals (wind farms), and the $N \times N$ covariance matrix at lagged k is defined as:

$$\Sigma_X(k) = \begin{bmatrix} R_{1,1}(k) & R_{1,2}(k) & \cdots & R_{1,N}(k) \\ R_{2,1}(k) & R_{2,2}(k) & \cdots & R_{2,N}(k) \\ \vdots & \vdots & \ddots & \vdots \\ R_{N,1}(k) & R_{N,2}(k) & \dots & R_{N,N}(k) \end{bmatrix} \quad k = 1, \cdots, M \tag{3.13}$$

where $R_{x1,x2}$ is the covariance function between wind farms. It is assumed that common components and idiosyncratic components are uncorrelated ($\mathrm{E}[\chi\xi^{\mathrm{T}}] = 0$), which leads to:

$$\begin{aligned} \Sigma_X(k) &= E\left\{X[t]X^T[t-k]\right\} = E\left\{\left(\chi[t] + \xi[t]\right)\left(\chi[t-k] + \xi[t-k]\right)^T\right\} \\ &= E\left\{\chi[t]\chi^T[t-k]\right\} + E\left\{\xi[t]\xi^T[t-k]\right\} \\ &= \Sigma_\chi(k) + \Sigma_\xi(k) \end{aligned} \tag{3.14}$$

where $\Sigma_X(k)$ is the auto/cross - covariance function of data matrix X at time lag k. Similarly, $\Sigma_\chi(k)$ and $\Sigma_\xi(k)$ are the auto/cross - covariance function of common component and idiosyncratic component, respectively.

In order to reduce the dimension of observed variables, we adopt the concept of principle component analysis (PCA) to transfer data into orthogonal basis sets such that the largest possible variance can be represented by less dimensional orthogonal basis. The covariance matrix Σ_X is a symmetric matrix and by conducing PCA, equation (3.13) can be decomposed into two parts:

$$\Sigma_X(k) = V^1(k)\Omega^1(k)V^1(k)^T + V^2(k)\Omega^2(k)V^2(k)^T = \Sigma_\chi(k) + \Sigma_\xi(k) \tag{3.15}$$

where Ω^1(Q×Q)is the diagonal matrix whose diagonal elements are top Q large eigenvalues of $\Sigma(k)$ and Ω^2, (N-Q) ×(N-Q) matrix, is the diagonal matrix whose diagonal is the rest of eigenvalues (N-Q), and V is the corresponding eigenvectors and V^T is its transpose. In all, the Q number of top large eigenvalues and their corresponding eigenvectors of the covariance matrix in X at the given lag are converted to the covariance matrix of common components in χ, and the rest of the eigenvalues and their corresponding eigenvectors are converted to the covariance of idiosyncratic components in ξ.

In DFM, the goal is to minimize the variance of idiosyncratic component, in others words, to minimize the summation of diagonals of the covariance matrix of idiosyncratic component, which is the trace of the covariance matrix defined by:

$$\begin{aligned} & tr\{\mathrm{E}[(X - A(L)B(L)X)(X - A(L)B(L)X)^T]\} \\ & \Updownarrow \\ & tr[(I_N - A(L)B(L))\Sigma_X(I_N - A(L)B(L))^T] \\ & \Updownarrow \\ & tr[(\sqrt{\Sigma_X} - A(L)B(L)\sqrt{\Sigma_X})(\sqrt{\Sigma_X} - A(L)B(L)\sqrt{\Sigma_X})^T] \end{aligned} \tag{3.16}$$

By the Courant-Fischer Theorem [42],

$$A(L)B(L)\sqrt{\Sigma_X} = V^1\sqrt{\Omega^1}V^{1T} \tag{3.17}$$

Recall that $\sum_X = V^1\Omega^1V^{1T} + V^2\Omega^2V^{2T}$ by Principle Component Analysis (PCA), thus equation (3.17) is further expand to

$$\begin{aligned} A(L)B(L) &= V^1\sqrt{\Omega^1}V^{1T}\sqrt{\Sigma_X{}^{-1}} \\ &= V^1\sqrt{\Omega^1}V^{1T}\left(V^1\sqrt{\Omega^{1-1}}V^{1T} + V^2\sqrt{\Omega^{2-1}}V^{2T}\right) \\ &= V^1V^{1T} \end{aligned} \tag{3.18}$$

From (3.18), $A(L)$ and $B(L)$ are found by: A(L) = $V^1(L)$, B(L) = $V^1(L)^T$, where $V^1(L) \in \Re^{N \times Q \times M}$ represents M numbers of $N \times Q$ matrices. Note that each column vector in matrices V^1 and V^2 are orthogonal basis (e.g., $V^1(:,1) \times V^2(:,1) = 0$) due to the property of eigen-decomposition of a symmetric matrix.

The Courant-Fischer Theorem is the fundamental to find $A(L)$ and $B(L)$ which is stated as: Let Γ be an $N \times N$ symmetric matrix with eigenvalues λ_i in descending order, where i is from 1,…, N. The $N \times N$ matrix A with rank (A) = Q that minimizes the top Q eigenvalues of

$$(\Gamma - A)(\Gamma - A)^* \tag{3.19}$$

to zero is given by

$$A = V^1 \Omega^1 V^{1*} \tag{3.20}$$

where Ω^1 denotes the diagonal matrix containing the top Q large eigenvalues of Γ and V^1 is the corresponding eigenvectors, and V^* means the conjugate transpose of matrix V.

For the vector auto regressive model (VAR) process:

$$f_t = C_1 f_{t-1} + C_2 f_{t-2} + \cdots + C_R f_{t-R} + \varepsilon_t \tag{3.21}$$

where the coefficient matrix $C(L)$ is estimated by Yule-Walker equation [47]. After multiplying shifted f_t on both side and calculating the covariance matrix by taking the expected value, we obtain:

$$\begin{aligned} \Sigma_1 &= C_1 \Sigma_0 + C_2 \Sigma_1 + \cdots + C_R \Sigma_{R-1} + E(\varepsilon_t f_{t-1}) \\ \Sigma_2 &= C_1 \Sigma_1 + C_2 \Sigma_0 + \cdots + C_R \Sigma_{R-2} + E(\varepsilon_t f_{t-2}) \\ &\vdots \\ \Sigma_R &= C_1 \Sigma_{R-1} + C_2 \Sigma_{R-2} + \cdots + C_R \Sigma_0 + E(\varepsilon_t f_{t-R}) \end{aligned} \tag{3.22}$$

where Σ_R is the covariance matrix of R lag. Since f_t is uncorrelated with ε_t, the expected value of them are zero. Then the equation is rearranged as:

$$\begin{bmatrix} \Sigma_1 \\ \Sigma_2 \\ \vdots \\ \Sigma_R \end{bmatrix} = \begin{bmatrix} \Sigma_0 & \Sigma_1 & \cdots & \Sigma_{R-1} \\ \Sigma_1 & \Sigma_0 & \cdots & \Sigma_{R-2} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{R-1} & \Sigma_{R-2} & \cdots & \Sigma_0 \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_R \end{bmatrix} \tag{3.23}$$

Therefore, $C(L)$ can be calculated by multiplying the inverse of the matrix in the right-hand side to the matrix in the left-hand side.

3.2 Verification of DFM

In all, the procedures of synthesizing scenarios can be summarized as following: (a) $\hat{\chi}$, $\hat{\xi}$, and $\hat{f}$ are estimated from data set X (the 'hat' denotes for the estimated value), (b) since dynamic factors f can be modeled as VAR process, the noise term in VAR model can be estimated as $\hat{\varepsilon}$ and meanwhile, the correlation structure matrix H can be estimated by Cholesky decomposition in [41], (c) Scenarios of $\hat{\chi}$ can be synthesized by giving dynamic shocks δ which are uncorrelated noises into the model as denoted by (3.7), (d) $\hat{X}$ can be finally synthesized by adding $\hat{\chi}$ with $\hat{\xi}$, and the estimation of original data can be obtained by adding seasonality and reversing the normalization process.

3.2.1 Verification in Time and Frequency Domain

Figures 3.4 and 3.5 show the comparison between actual wind power and its common component $\hat{\chi}$ at 75th and 81th wind farm site. From two figures it is observed that the common component can catch the trend of actual wind power (similar ramp, maximum, minimum and overall shape with the actual measurements).

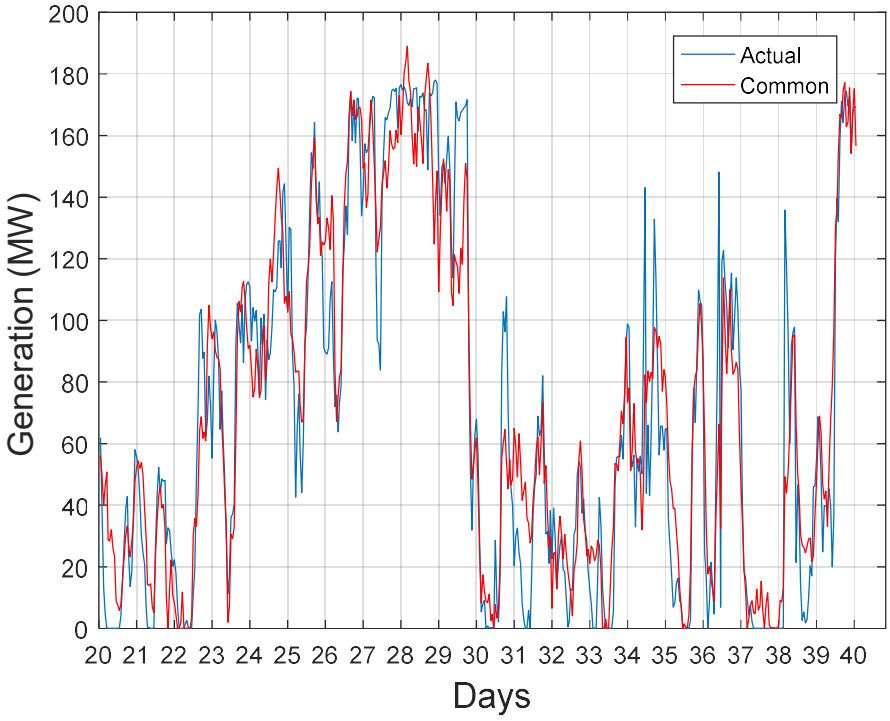

Figure 3.4. Actual wind power and common component of wind power at #75 wind farm.

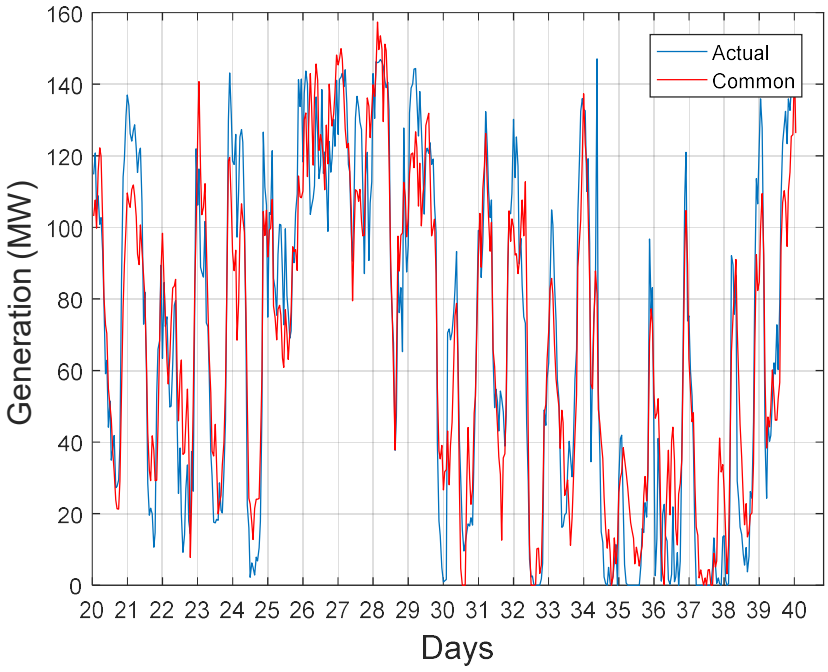

Figure 3.5. Actual wind power and scenario of wind power at #81 wind farm.

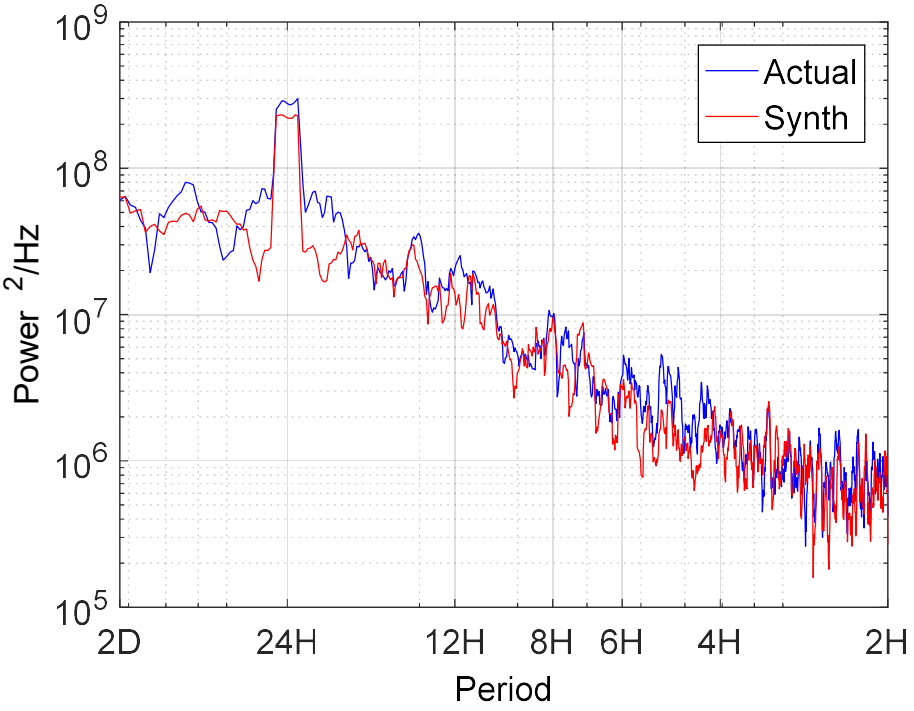

Figure 3.6. PSD of Actual wind power and synthesized wind power at #81 wind farm.

Figure 3.6 also verifies the model by PSD plot in frequency domain. Note that the x-axis is 'period', the reciprocal of 'frequency'. It is found the function with frequency corresponding to 24 hour period contributes the most in #81 wind farm observation data. Similarly the synthesized PSD agrees with the actual wind data.

As described earlier, one of the advantages is that the DFM can capture the co-movement of synthesized scenarios; in other words, the synthesized scenarios of different wind farms have similar correlations as the actual ones do. Figures 3.7 and 3.8 demonstrate the actual wind power correlation between #9 and #10 wind farms with the correlation coefficient 0.9215, and the synthesized scenarios have the correlation coefficient of 0.9000.

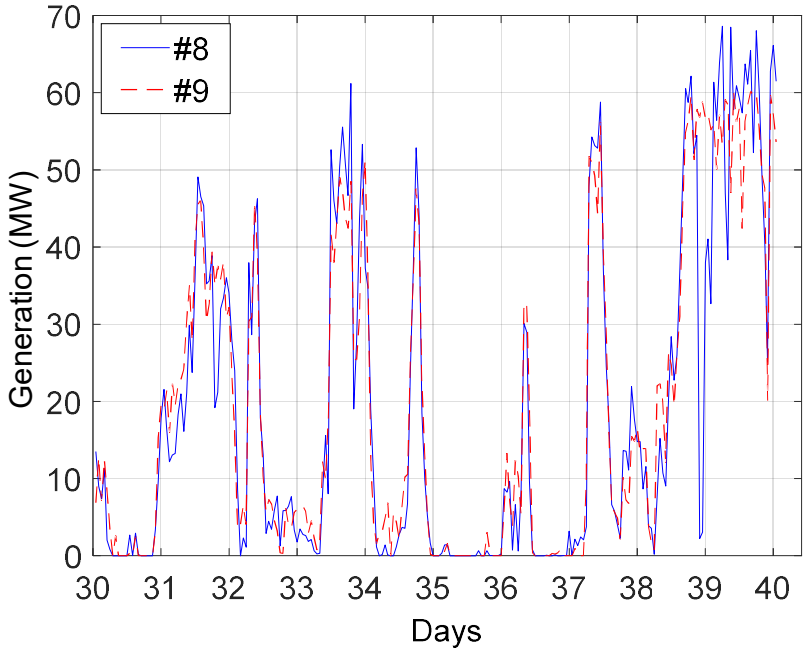

Figure 3.7. Correlation of wind power at #8 and #9 wind farms with correlation coefficient of 0.9215.

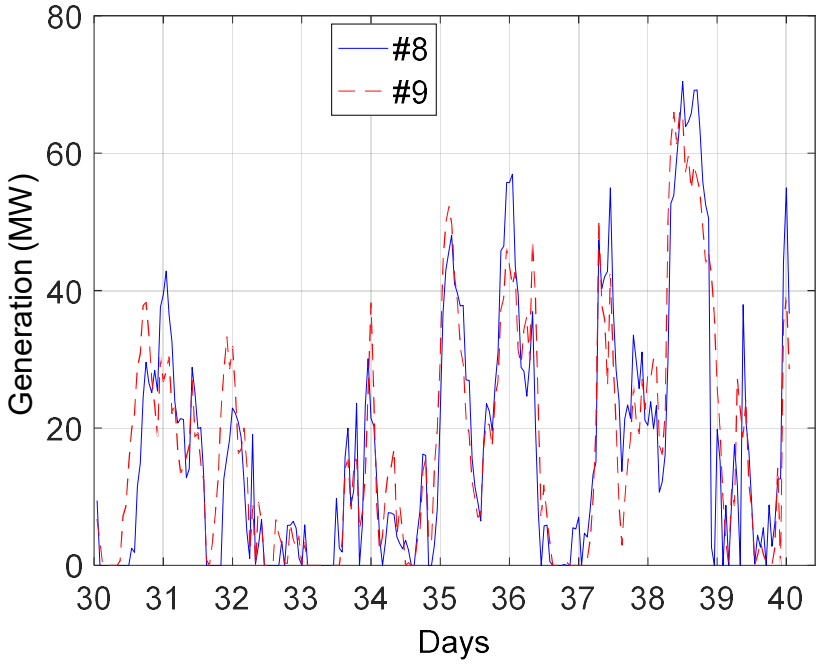

Figure 3.8. Scenarios correlation of wind power at #8 and #9 wind farms with correlation coefficient of 0.9.

3.2.2 Forecast by DFM

Forecast by time-series model means to calculate the expected value by the model, in other words, considering the expected value of noise as zero and such forecast provides minimum square error [48]:

$$\chi_{t+k}^{t} = A(L)\left[C(L)\right]^{-1} H\delta_{t+k}^{t} \tag{3.24}$$

where k is the future steps past the end of the observed series. The superscript t is to be read as "given data up to time t". Therefore in order to have the minimum square error forecast over next 24 hours, columns in dynamic shock, $\delta_{t+1},\ldots,\delta_{t+24}$ are zeros which is illustrated by Fig. 3.9.

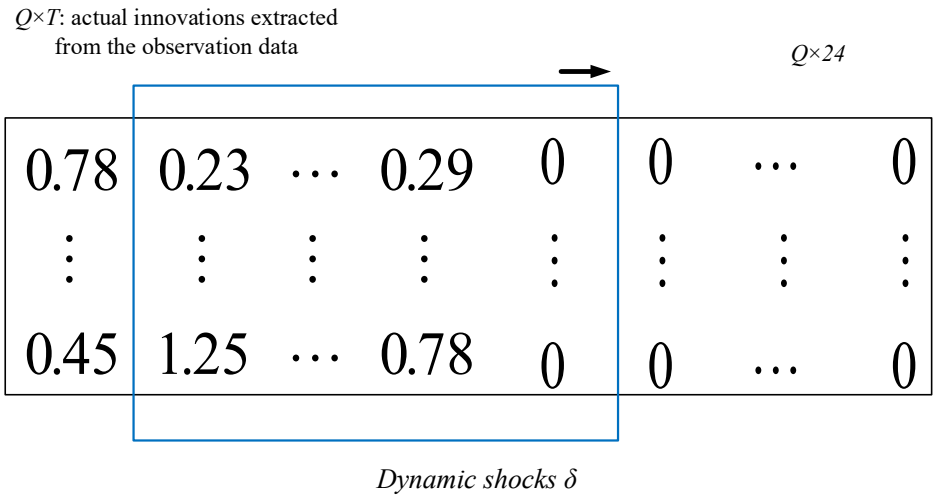

Figure 3.9. Dynamic shock illustration in forecasting.

The actual innovation extracted from the observation data is $Q\times T$, and is augmented by future innovation (zeros) of $Q\times 24$. The filter $A(L)\left[C(L)\right]^{-1} H$, a moving window, is indicated in blue box and applied to the innovation to obtain the common component χ_{t+k}^{t}.

The forecast result is provided in Figures 3.10 and 3.11. There are 15 forecasted future scenarios plotted in blue lines, the minimum mean square error (MSE) forecast plotted in red (dynamic shock columns are zero), and the actual wind power plotted in black.

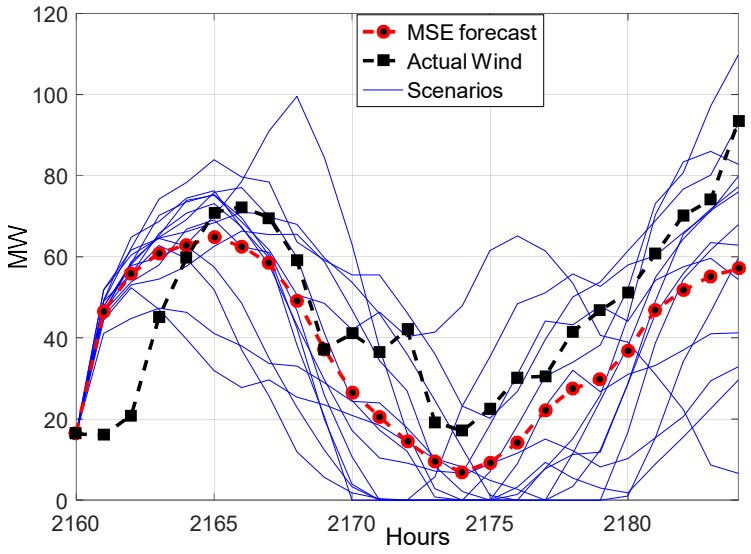

Figure 3.10. Forecast at #29 wind farm.

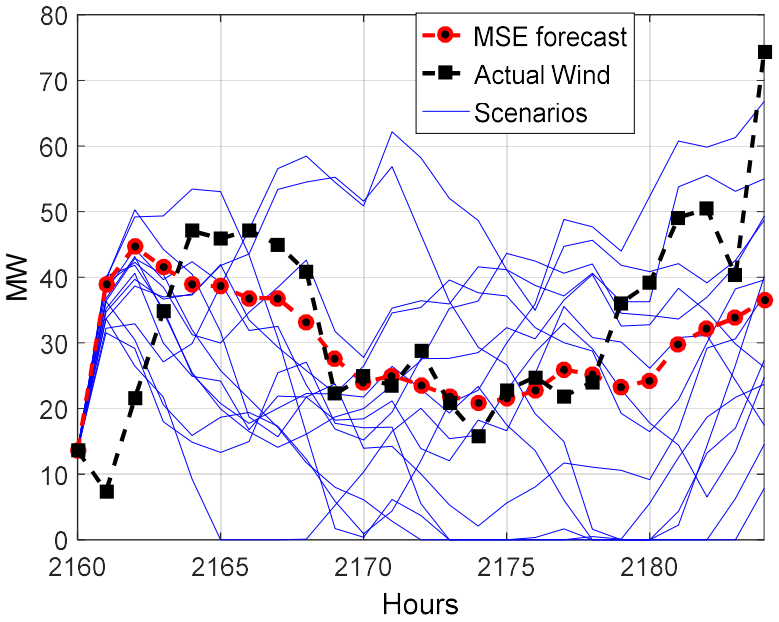

Figure 3.11. Forecast at #32 wind farm.

The forecast is able to capture the trend of actual wind power, and the generated scenarios are used by stochastic programming to make decisions concerning the penetration of intermittent wind power.

CHAPTER FOUR

Heuristic Method on Power System Application

4.1 Introduction

Modern heuristic optimization techniques have been developed in the past decades to facilitate solving optimization problems which were previously difficult or impossible to solve in power systems area [5]. Those techniques include evolutionary programming, simulated annealing, Tabu search, particle swarm optimization, etc. The applications in power system range over security assessment, generation and maintenance scheduling, economic dispatch, optimal power flow, transmission network expansion planning, generation expansion and reactive power planning, distribution system optimization, power plant and power system controls, etc. In this chapter a novel heuristic technique called artificial bee colony (ABC) is introduced with its application to optimal power flow.

4.1.1 Optimization in Power System Operation

Power system optimization has evolved with developments in computing and optimization theory in decades. As early as in the first half of the 20th century, the optimal power flow was 'solved' by empirical methods, rules of thumb and primitive tools such as analog network analyzers [49]. With the development of computing ability of computers, the optimal power flow problem was first formulated by Carpentier in 1962 [89] and has proven to be a very difficult problem to solve.

There are generally three types when it comes to power system operation optimization problems in the literature: power flow (load flow), economic dispatch, and optimal power flow. Here a brief description of the differences/commons is given among these problems. Table 4.1 lists the major characteristics of power system operation problems. The power flow focuses on the generation; load and transmission network models and solves a nonlinear mathematical problem. However, the solution might not be optimal or physically feasible under certain constraints. For instance, the power flow models do not consider generator reactive power and transmission line limits. Historically economic dispatch (ED) has been the major tool for system operation and planning of the power systems; however the control variables for ED are only real power and the electrical network is solely represented by single equality constraint, the power balance equation. The economic dispatch fails to consider the power flow constraints in network.

The optimal power flow (OPF) solves for the optimal solution under a specified objective function subject to the power flow constraints, generator limits, transmission lines' thermal capacity, switching equipment limits, etc. From Table 4.1, there are various sub-problems stemming from the three general types in order to meet specific situations. Such as alternating current OPF (ACOPF), direct current OPF (DCOPF), security constrained OPF (SCOPF), etc. Note that OPF can be a fundamental tool for power system operation, and based on such tool various modified versions can be developed for specific purposes. For instance, nowadays with high penetration of renewable energy and storage devices, system operators are developing dispatch methods to meet the needs and the essence of such methods are still OPF problem.

Table 4.1. Major Types of Power System Problems

General problem type	Problem name	Voltage angle constraint?	Bus voltage constraint?	Transmission constraint?	Assumption	Generator costs?	Contingency constraints?
OPF	ACOPF	Y	Y	Y	----	Y	N
OPF	DCOPF	N	N	Y	V is constant	Y	N
OPF	Security Constrained ED (SCED)	Y	N	Y	V is constant	Y	Y
PF	Power Flow	N	Y	N	----	N	N
ED	ED	N	N	N	No transmission constraints	Y	N
OPF	Security Constrained OPF (SCOPF)	Y	Depends	Y	Depends	Y	Y

System operators and planners consider the OPF is one of the most essential problems in power systems because of the detailed controls over power system which result in a significant reduction in the cost. The OPF was first proposed as early as to 1960s by Carpentier and has grown into a powerful tool for system operation and planning [50]-[52]. The OPF problem, under the objective function of minimizing cost, takes into account the constraints of AC load flow at each node, transmission line capacity, voltage limits, etc. and controls real, reactive power, voltage, transformer taps, shunt compensators which lead to significant reduction in the cost. In other words, the aim of OPF is to optimize an objective function representing the total cost of generation, power losses, voltage stability, and/or other relevant information, while satisfying the system constraints [53].

Classical methods such as linear programming [54], quadratic programming [55] and interior point method [56]-[57], rely on theoretical assumptions of convexity and are

very sensitive to starting points. Therefore these methods can be easily trapped into a local optimum or diverge and only fit for limited types of objective functions. The OPF becomes a highly non-linear, non-convex and large dimensional optimization problem after incorporating non-smooth, non-convex, non-linear, and non-differential objective functions and constraints, which is difficult, if not impossible, to solve by classical methods. Therefore researchers have been focused on developing more efficient and robust methods to handle OPF problems without simplifying the system. For this purpose, meta-heuristic methods such as genetic algorithm [58], evolutionary programming [59], Tabu search [60], particle swarm optimization [61] have been placed attention on solving the OPF problems. Reference [61] proposed an improved PSO to tackle the problem considering the valve point effect on the regular quadratic fuel cost function.

4.1.2 Optimal Power Flow Problem

Mathematically, solving an OPF problem is equivalent to finding a set of optimal decision/control vectors that minimizes an objective function under several constraints. The OPF problem to be considered is formulated as follows [53]:

$$\begin{aligned} Min \quad & f(x,u) = 0 \\ s.t. \quad & g(x,u) = 0 \\ & h(x,u) \le 0 \end{aligned} \tag{4.1}$$

where

u : decision/control vectors,

x : state vectors,

f : objective functions,

g : equality functions,

h : inequality functions,

The vector u includes generator real power P_G except at the slack bus, generator bus voltage V_G, transformer tap T, and shunt compensator Q_C at selected buses. The vector x includes real power P_{G1} at the slack bus, voltage V_L at load buses, reactive power Q_G at generator buses, and transmission line loadings S_L. There are four objective functions chosen in the study: quadratic cost function, quadratic cost function with valve-point effect, power loss, and voltage stability. They are respectively listed in following:

$$f_1 = a_i P_{Gi}^2 + b_i P_{Gi} + c_i \tag{4.2}$$

$$f_2 = a_i P_{Gi}^2 + b_i P_{Gi} + c_i + \left| d_i \sin\left(e_i \left(P_{Gi,\min} - P_{Gi}\right)\right)\right| \tag{4.3}$$

$$f_3 = \sum_{k=1}^{N_l} \frac{r_k}{r_k^2 + x_k^2}\left[V_i^2 + V_j^2 - 2V_i V_j \cos\left(\delta_i - \delta_j\right)\right] \quad \forall i, \forall j \tag{4.4}$$

$$f_4 = a_i P_{Gi}^2 + b_i P_{Gi} + c_i + \omega \sum_{i=1}^{N_{pq}} \left|V_i - 1\right| \tag{4.5}$$

The equality constraints g from (4.1) are the AC power flow balance equations at each bus representing that the power flowing into that specific bus is equal to the power flowing out, and the equations are defined as:

$$\begin{aligned} P_i &= V_i \sum_{j=1}^{N} V_j Y_{ij} \cos\left(\delta_i - \delta_j - \theta_{ij}\right) \\ Q_i &= V_i \sum_{j=1}^{N} V_j Y_{ij} \sin\left(\delta_i - \delta_j - \theta_{ij}\right) \quad \forall i, \forall j \end{aligned} \tag{4.6}$$

Inequality constraints h in (4.1) are listed as generator limits, tap position of transformers, shunt capacitor constraints, and security constraints on load bus voltage and transmission line flows.

$$\begin{aligned} &P_{Gi,\max} \le P_{Gi} \le P_{Gi,\max} \\ &Q_{Gi,\max} \le Q_{Gi} \le Q_{Gi,\max} \\ &V_{Gi,\max} \le V_{Gi} \le V_{Gi,\max} \quad i \in N_G \end{aligned} \tag{4.7}$$

$$TP_{i,\min} \le TP_i \le TP_{i,\max} \qquad i \in N_T \tag{4.8}$$

$$Q_{ci,\min} \le Q_{ci} \le Q_{ci,\max} \qquad i \in N_c \tag{4.9}$$

$$\begin{aligned} &V_{Li,\min} \le V_{Li} \le V_{Li,\max} \qquad i \in N_{pq} \\ &S_{Li} \le S_{Li,\max} \qquad i \in N_l \end{aligned} \tag{4.10}$$

where

a_i, b_i, c_i, d_i, e_i, : fuel cost coefficients of the i-th unit,

P_{Gi} : real power of the i-th unit,

V_i : voltage magnitude at bus i,

r_k, x_k : the resistance and reactance of the transmission line k that links bus i and j,

V_i, V_j : voltages at bus i and j,

δ_i, δ_j : angles at bus i and j,

ω : the weighting factor,

N_{pq} : the number of PQ buses,

N_l : the total number of transmission lines,

N_G : the number of generators,

N_T : the number of tap-changing transformers,

Y_{ij}, θ_{ij} : the Y-bus admittance matrix elements,

$P_{Gi,min}$, $P_{Gi,max}$: the minimum/maximum real power limits of generating unit i,

$Q_{Gi,min}$, $Q_{Gi,max}$: the minimum/maximum reactive power limits of generating unit i,

$V_{Gi,min}$, $V_{Gi,max}$: the minimum/maximum voltage limits of generating unit i,

$TP_{i,min}$, $TP_{i,max}$: the limits of transformers,

$Q_{ci,min}$, $Q_{ci,max}$: the limits of shunt capacitors,

$V_{Li,min}$, $V_{Li,max}$: the limits of load bus voltage,

$S_{Li,max}$: the maximum line flow of transmission line i,

It is worth to mention that the control variables (real power generation of PV buses, voltage at all generator buses, transformer tap settings, and shunt compensators) are randomly initialized within the feasible domain, while a penalty function is introduced in order to ensure that the dependent/state variables are in the feasible domain as well. In other words, penalty function is utilized to handle the inequality constraints. The penalty cost function is defined as:

$$p(x_i) = \begin{Bmatrix} (x_i - x_{i,\max})^2 \\ (x_{i,\min} - x_i)^2 \\ 0 \end{Bmatrix} \quad \begin{matrix} if \quad x_i > x_{i,\max} \\ if \quad x_i < x_{i,\min} \\ if \quad x_{i,\min} \le x_i \le x_{i,\max} \end{matrix} \tag{4.11}$$

where

$p(x_i)$: the penalty function of dependent variable x_i at bus i,

The penalty cost increases with a quadratic form when dependent variables are exceeding the limits and the cost is zero if the constraints are not violated. For example, if one of the PQ bus voltage exceeds the limit, certain amount of penalty will be added, which leads to the increase of total cost and eventually this solution will be abandoned. Thus the augmented objective function by adding the penalty function of the slack bus, reactive power generation, *PQ* bus voltage and transmission line capacity is described as:

$$F = f + C_p p(P_{G1}) + C_q \sum_{i=1}^{N_G} p(Q_{Gi}) + C_v \sum_{i=1}^{N_{pq}} p(V_{Li}) + C_s \sum_{i=1}^{N_l} p(S_{Li}) \tag{4.12}$$

where f is the original fuel cost function (f_1, f_2, f_3, or f_4 in (4.2)-(4.5)), C_p, C_q, C_v and C_s respectively denote the penalty factors of real power generation of slack bus, reactive power output of the generator buses, and *PQ* bus voltage and transmission line capacity, i.e.,

f : the original fuel cost function (f_1, f_2, f_3 or f_4 in this study),

C_p: penalty factors of real power generation of slack bus,

C_q: penalty factors of reactive power output of the generator buses,

C_v: penalty factors of *PQ* bus voltage,

C_s: penalty factors of transmission line capacity.

4.2 Artificial Bee Colony

Artificial bee colony (ABC) is a population-based search procedure inspired from the intelligent behavior of honeybees [62]. It is as simple as particle swarm optimization (PSO) and differential evolution (DE) algorithms, and uses only common control parameters such as colony size and maximum cycle number. There are three types of bees in the ABC system: employed bee, onlooker bee and scout bee. The aim of all bees is to find the best food source (possible solution) with highest nectar (fitness value); in other words, artificial bees fly in a multidimensional search space to find the global optimal. Employed bees search for food sources based on their memory and the information gathered on food sources is shared with onlooker bees. Onlooker bees tend to choose good sources with higher nectar and further explore new food sources around the

selected food sources. Scout bees abandon old food sources and randomly start a new source in a way to avoid local minimum.

As the ABC has proven its robust, efficient and simple characteristics, it has been widely implemented in solving a range of optimization problems in recent years such as job shop scheduling and machine timetabling problems [63]. The ABC was also implemented to tune the PI controller parameters in microgrid power electronics control [64]. In reference [65], authors applied the ABC or improved ABC in power system problems such as OPF, modified OPF which integrated wind power and storage devices. Continuing this section the detailed procedures of ABC algorithm is presented.

4.2.1 Employed Bee Phase

Before entering into employed bee phase, initial possible solutions are generated from the search space. After the initialization, the search process will be carried out in a repeated cycles by those three types of bees. At initialization each vector solution $X_i = \{X_{i,1},X_{i,2},\ldots,X_{i,D}\}$ is generated randomly within the limits of the control variables as follows:

$$X_{i,j} = X_{i,j\_\min} + rand(0,1)\times(X_{i,j\_\max} - X_{i,j\_\min}) \tag{4.13}$$

where

SN: the number of employed bees and onlooker bees, *i* is from 1 to *SN*,

D: the number of control variables, *j* is a random number from 1 to *D*,

X_{i,j_min} : the lower bounds for dimension *j*,

X_{i,j_max} : the upper bounds for dimension *j*,

rand(0,1): a uniformly distributed random number in (0,1).

In the employed bee phase, employed bees update the current solution based on neighborhood information and then evaluate the nectar (fitness) of the new food source. The update equation is defined as:

$$V_{i,j} = X_{i,j} + \Phi_{i,j} \times (X_{i,j} - X_{k,j}) \tag{4.14}$$

where

k : is an integer different from i, uniformly chosen from the range [1, SN],

$\Phi_{i,j}$: a random number from [-1,1],

If the new source is better (higher nectar) than the old one, employed bee will memorize the new source and disregard the old. Otherwise the old one will be remained. Such scheme is simply known as greedy selection.

4.2.2 Onlooker Bee Phase

The onlooker bee phase starts when the food source information was shared from employed bees. In nature, onlooker bees tend to select the food source with higher nectar. The nectar information has been shared by employed bees. To mimic such phenomenon, the roulette wheel selection scheme [5] is used in the onlooker bee phase, which ensures good food source will have a higher probability to be selected, and then onlooker bee will update those solutions. The roulette wheel selection scheme is defined in the following:

$$P_i = \frac{\hat{fit}_i}{\sum_{j=1}^{SN} \hat{fit}_j} \tag{4.15}$$

where

$\hat{fit}_i$: the fitness value associated with solution i,

P_i : the probability associated with solution i,

The onlooker bee updates the selected solution using equation (4.14) as the employed bees do and memorize the solution by greedy selection. This process will continue until every onlooker bee finishes its search.

4.2.3 Scout Bee Phase

After a predefined number of searching cycles, food sources become exhausted (inactive solution) if their quality could not be improved anymore, then an employed bee will become a scout bee to start a random direction to search for new food source. This process is to avoid local optima. In the original ABC, only one scout is allowed to occur in each cycle [62]. After finding a new food source, the scout bee will turn itself back to employed bee. Note that the random search by scout is also performed by equation (4.13) same as in the initialization stage.

4.2.4 ABC for the OPF Problem

As mentioned in previous sections, the OPF problem is to find the optimal decision variables so that the objective function can be optimized. The control/decision vector u consists of:

$$u=[P_G; V_G; T; Q_C] \tag{4.16}$$

where

P_G : real power output at PV buses,

V_G : bus voltage at PV and slack buses,

T : transformer tap settings,

Q_C : shunt compensators settings,

The state vector x consists of

$$x=[P_{G1};\ V_L;\ Q_G;\ S_L] \tag{4.17}$$

where

P_{G1} : real power output at slack bus,

V_L: PQ bus voltage,

Q_G : generator reactive power output,

S_L : transmission line loadings,

It is necessary to clarify for those readers who are new to power system that there are three types of buses in power system: slack bus, PV bus and PQ bus. Slack bus is to balance the real and reactive power in the system while performing load flow calculations, it is also known as reference bus. The PV bus is the node where real power P and voltage magnitude V are specified, it is also known as generator bus. The PQ bus is the node where real and reactive power is specified, known as load bus.

There are four types of objective functions (f_1, f_2, f_3 and f_4) in this study as defined in previous sections. The fitness value of one solution can be evaluated using the following equation [62]:

$$fit_i = \begin{cases} \dfrac{\alpha}{1+f} & \text{if } f \geq 0 \\ 1+abs(f) & \text{if } f < 0 \end{cases} \tag{4.18}$$

where

α : constant,

f: objective function of (4.12).

It can be easily seen that for the cost function f greater than zero (OPF problems), to minimize cost function is to maximize the fitness value.

The framework of ABC algorithm is summarized as follows:

Step 1) Initialization:

1.1) Randomly generate SN points in the search space as feasible solution X_i by (4.13).

1.2) Run AC power flow and evaluate the fitness equation (4.18).

Step 2) For all employed bees ($i = 1,\ldots,SN$):

2.1) Update a candidate solution V_i by (4.14).

2.2) Run power flow and evaluate the fitness function, and calculate the probability p associated with its fitness by (4.15).

2.3) Choose a solution (from X_i and V_i) with better fitness value.

Step 3) For all onlooker bees (will only be executed under certain probability p):

3.1) Update a new candidate solution V_i by (4.14).

3.2) Run AC power flow and evaluate the fitness function by (4.18).

3.3) Choose a solution (from X_i and V_i) with better fitness function.

Step 4) For all scout bees (they will be executed only after the maximum trial m). Note that the maximum trial m is a predefined number that if a certain food cannot be improved by an employed bee after m times, the employed bee will become a scout bee.

4.1) Replace X_i with a new random solution X_i by (4.13).

4.2) Run AC power flow and evaluate the fitness function by (4.18).

After initialization, the algorithm repeats the search processes of employed bees, onlooker bees, and scout bees by a predefined cycle.

4.3 Case Studies by ABC

All simulations were performed on a computer with 3.4 GHz Intel core i7 Processor and 8 GB RAM. Power flow was calculated by Newton-Raphson method in MATPOWER package [66].

4.3.1 Test System Description

All four test cases were performed on IEEE-30 bus system. The configuration of the system is shown in Fig. 4.1. Reference [67] gives the data of IEEE 30-bus test system, and control variable limits. There are total 24 control variables which consist of real power generation at five PV buses and voltage magnitude of all six generator buses, nine shunt compensator controls for injecting reactive power and four transformer tap controls. There are six generators and buses 10, 12, 15, 17, 20, 21, 23, 24 and 29 are equipped with shunt compensators. In addition, lines 4-12, 6-9, 6-10, and 28-27 are equipped with tap-changing transformers as shown in Fig. 4.1. The system is at 100 MVA base with active power demand of 2.834 p.u. and reactive power demand of 1.262 p.u. The quadratic cost fuel cost coefficients were taken from [51].

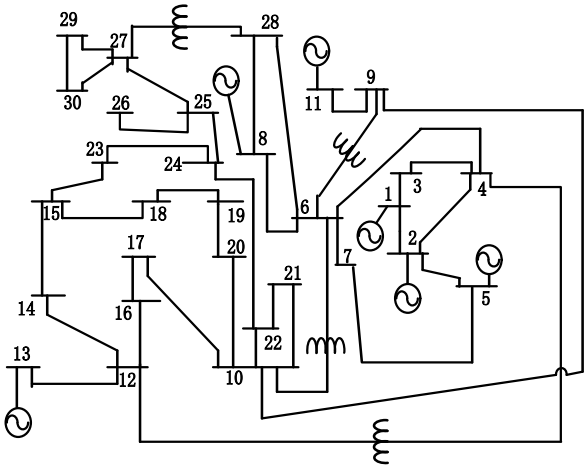

Figure 4.1. IEEE-30 bus system.

4.3.2 ABC Parameters for OPF

There are several parameters to be pre-determined in the ABC algorithm. For instance, the number of colony size is 200, in other words, the size of employed bees and onlooker bees is 100 for each. The number of food sources is 100. The trial limit is 100; meaning that a food source could not be improved after 100 updating attempts and then the food source will be abandoned by its employed bee. The number of cycles for foraging (stop criteria) is 400 iterations and there are total 24 parameters to be optimized as listed in Table 4.2.

Table 4.2. Parameters for ABC algorithm

Parameters	Values
Colony Size	200
Food Number	100
Limited Trials	100
Maximum Cycles	400
Parameters to be Optimized	24

4.3.3 Case 1: Quadratic Cost Function

Case 1 is the standard OPF problem with quadratic cost function. The objective of this case is to minimize total generator fuel cost (4.2). Simulation was run 30 times in order to conduct statistical analysis. The minimum total cost from ABC is 799.904 $/h, with the maximum 801.518 $/h, the average 800.944 $/h, and standard deviation 0.162. Results from other methods such as gravitational search algorithm (GSA), linearly decreasing inertia weight particle swarm optimization (LDI-PSO), enhanced genetic algorithm (EGA), modified differential evolution (MDE), and modified shuffle-frog leaping algorithm (MSFA) [58], [68]-[71] were made comparison to the results from

ABC. The comparison including execution time is given in Table 4.3. Fig. 4.2 shows the convergence properties of ABC.

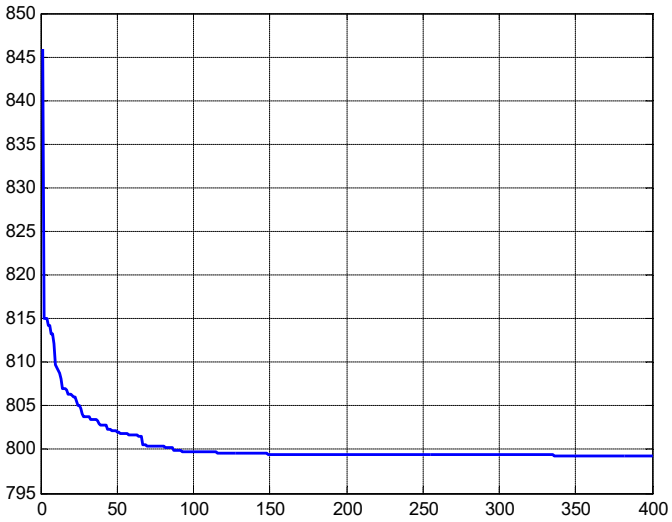

Figure 4.2. Convergence characteristics of ABC method in Case 1.

Table 4.3. Comparison of fuel cost in Case 1

Method	Min. ($/h)	Avg. ($/h)	Max. ($/h)	Std. Dev. (σ)	t (s)
ABC	799.904	800.944	801.518	0.162	39.8
GSA [68]	805.175	812.194	827.459	N/A	10.8
LDI-PSO [69]	800.734	801.557	803.869	N/A	N/A
EGA [58]	802.060	N/A	802.140	N/A	N/A
MDE [70]	802.376	802.382	802.404	N/A	23.3
MSFLA [71]	802.287	802.414	802.509	N/A	N/A

Note that the computational time can be affected by several variables, such as the performance of computer, the complexity of algorithms, efficiency of code, etc. Although the ABC algorithm used in the study did not outperform some of other algorithms, we argue that the computational performance can be improved by using advanced computers

or parallel computing methods. Therefore computational time comparison is not the focal point in the study.

4.3.4 Case 2: Quadratic Cost Function with Valve-Point Effect

In a real power plant, steam is controlled by valves to enter the turbine through separate nozzle groups. The best efficiency is achieved when each nozzle group operates at full output [72]. Therefore in order to achieve highest possible efficiency for given output, valves are opened in sequence and this results in a rippled cost curve as in Fig. 4.3. The objective function is given by equation (4.3). Table 4.4 shows the comparison with other methods.

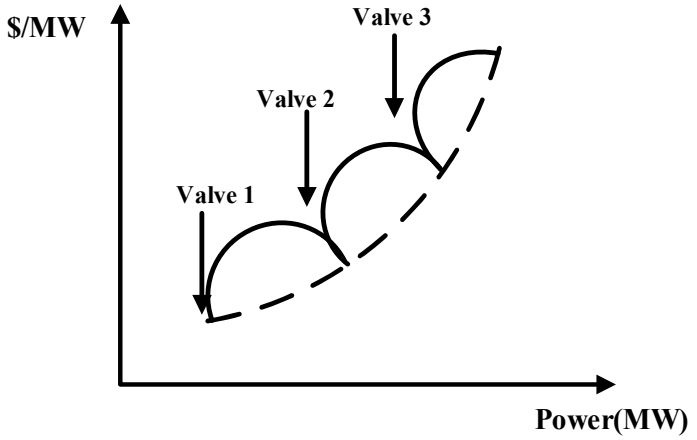

Figure 4.3. Effect of valve-point loading on a quadratic cost function.

Table 4.4. Comparison of case 2 fuel cost

Method	Min. ($/h)	Avg. ($/h)	Max. ($/h)	Std. Dev. (σ)	t (s)
ABC	923.436	924.124	924.894	0.562	39.8
GSA [68]	929.724	930.925	932.049	N/A	9.83
MDE [70]	930.793	942.501	954.073	N/A	N/A

From Table 4.4 the minimum total cost from ABC is 923.436$/h, with the maximum 924.894$/h, the average 924.124$/h, and standard deviation 0.562, and the convergence property is shown in Fig. 4.4.

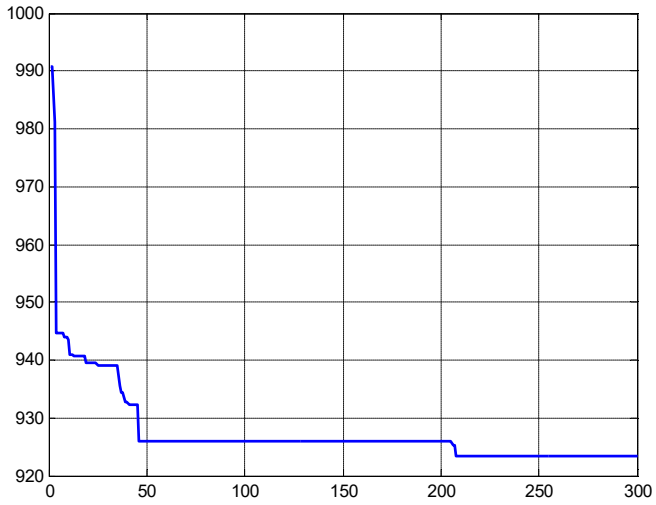

Figure 4.4. Convergence characteristics of ABC method in Case 2.

4.3.5 Case 3: Minimization of Power Loss

Real power loss is due to the power flowing through transmission lines which consist of resistance and reactance. It is apparent that minimizing real power loss is one of the major concerns for system operation. The objective function is given by (4.4). The convergence property is shown in Fig. 4.5 and Table 4.5 shows the comparison of case 3.

Table 4.5. Comparison of power loss in Case 3

Method	Min. (MW/h)	Avg. (MW/h)	Max. (MW/h)	Std. Dev. (σ)	t (s)
ABC	3.096	3.112	3.177	0.036	70.8
HS [73]	N/A	2.967	N/A	N/A	N/A
EGA [58]	N/A	3.201	N/A	N/A	N/A

From Table 4.5, the minimum total power loss from ABC is 3.096 MW/h, with the maximum 3.177 MW/h, the average 3.112 MW/h, and standard deviation 0.036. However the power loss found by harmony search from reference [73], according to reference [65], is not a feasible solution because the authors in [65] verified the power flow based on the optimal decision variables, there were bus voltage violations at all load buses except bus 7.

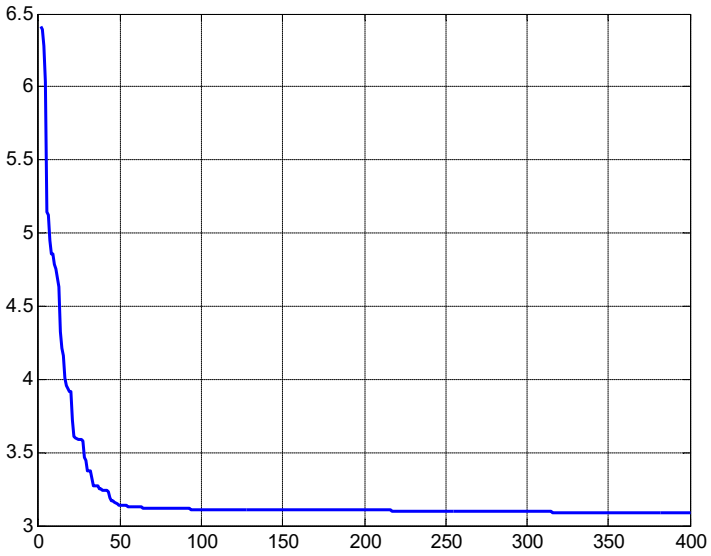

Figure 4.5. Convergence characteristics of ABC method in Case 3.

4.3.6 Case 4: Voltage Profile Improvement

In power system operation, minimizing total cost is usually not the only objective considered, and other issue such as minimizing voltage derivation is of great importance. Thus the improvement of voltage profile is also investigated. The objective function for minimizing all *PQ* bus voltage *V* deviating from 1.0 p.u is described by (4.5). Fig. 4.6 shows the convergence property.

After 30 times simulations, the minimum total cost found by ABC is 955.227 $/h, with the maximum 958.147 $/h, the average 956.824 $/h, and standard deviation 0.512.

Since this is not a standard comparison case, no reference from other methods with the same system parameters can be found. Fig. 4.7 compares the PQ bus voltage profiles with the case of minimizing basic quadratic cost function (Case 1).

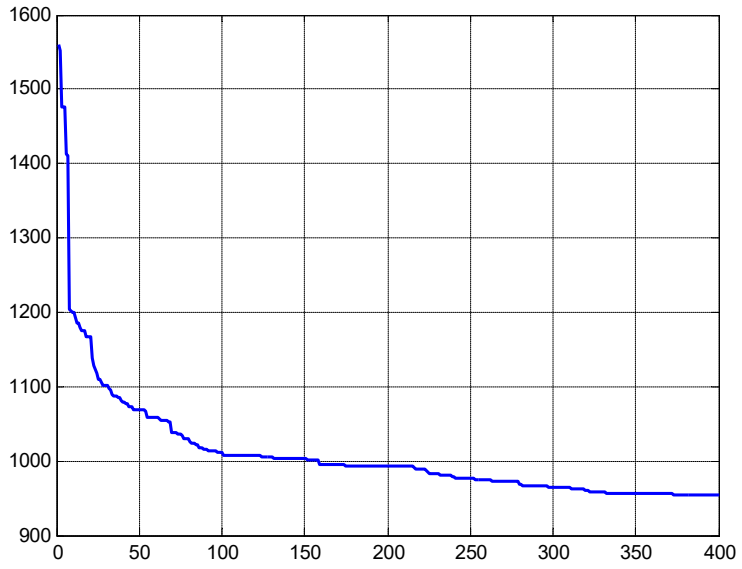

Figure 4.6. Convergence characteristics of ABC method in Case 4.

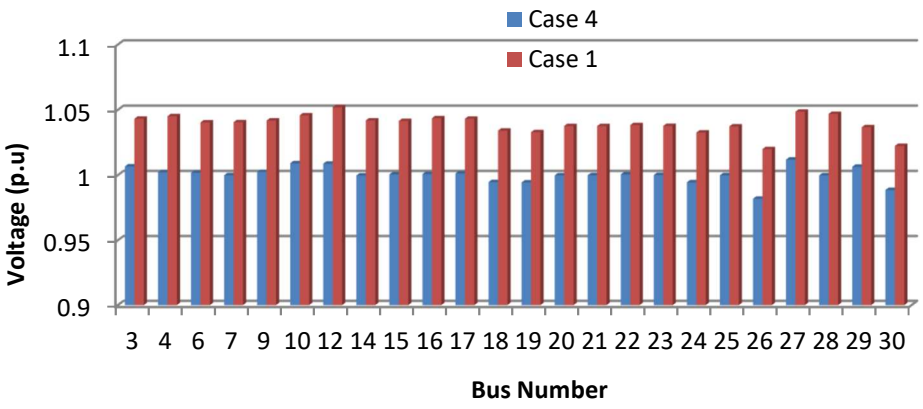

Figure 4.7. Voltage profiles for Case 1 and Case 4.

From Fig. 4.7 it is obvious that by considering voltage improvement, the voltages of PQ buses stay close to 1 p.u.; on the other hand the voltage deviates from 1 p.u but within feasible limits when voltage improvement is not considered.

This section introduced the artificial bee colony (ABC) algorithm for handling non-linear, non-convex optimal power flow problems. In this study the ABC has been proven its efficiency and robustness over four test cases of OPF problems. The comparison results have shown that ABC outperforms other heuristic methods in terms of finding better solutions with fewer costs. In addition, with the help of high speed computers or parallel computing algorithm, the computational burden can be further reduced, which enable the ABC algorithm to promise as a useful tool for OPF problems.

4.3.7 Case 5: Large System OPF by ABC

Previous cases were tested on modified IEEE 30-bus system and in order to test the robustness of ABC on large systems, IEEE 118-bus system was adopted as a test system. Such system contains 54 generator buses, 9 shunt compensators and 9 transformers. The diagram of IEEE 118-bus is shown in Fig. 4.8. Similarly, simulation were run 30 times for this case study, and Table 4.6 lists the simulation results. The average cost is 130,321 $/hr, with maximum 130,410 $/hr and minimum 130,210 $/hr; standard deviation 90.5 $/hr and it takes 4037.5 seconds.

Table 4.6. Results for Case 5

Method	Min. (MW/h)	Avg. (MW/h)	Max. (MW/h)	Std. Dev. (σ)	t (s)
ABC	130,210	130,321	130,410	90.5	4037.5

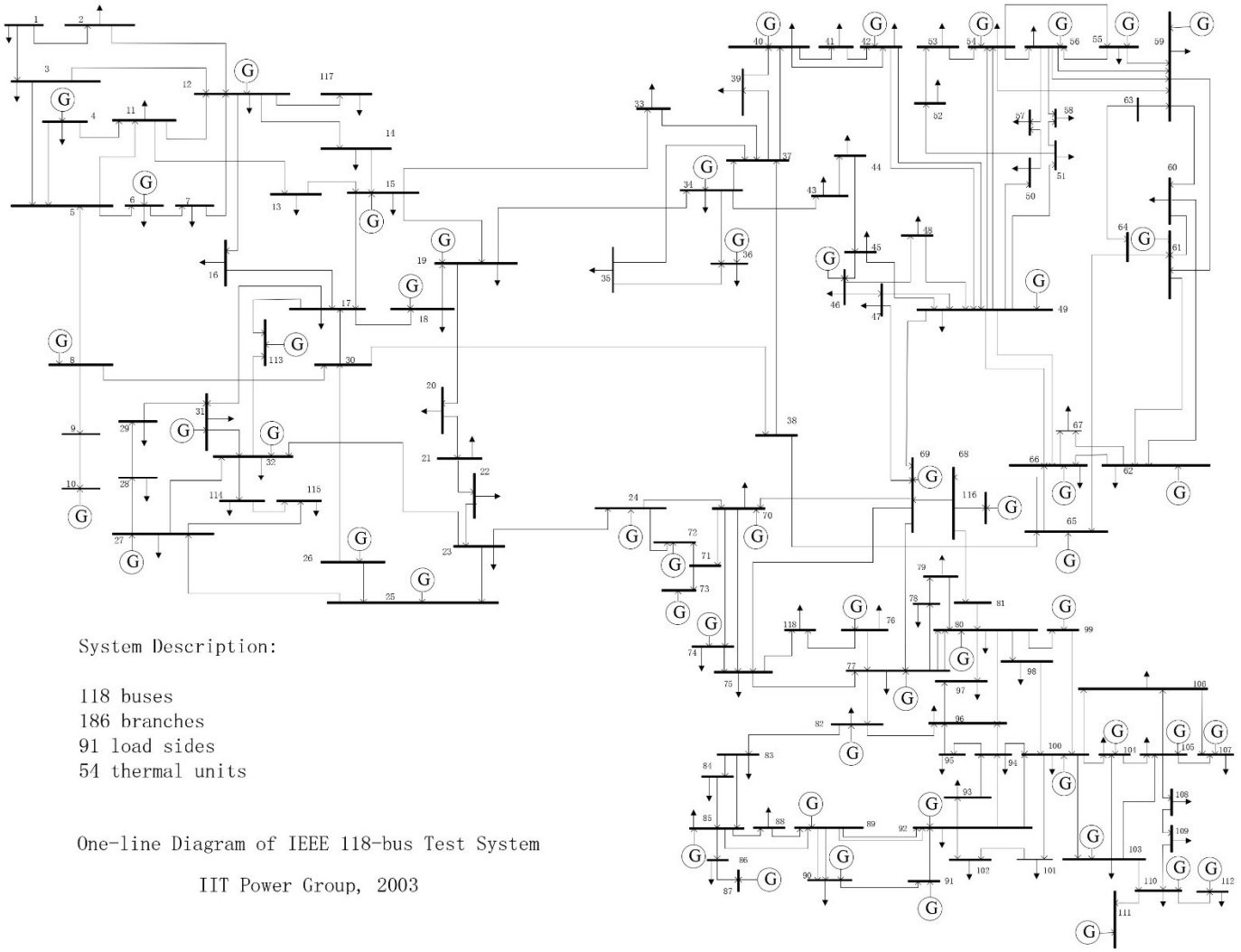

Fig. 4.8. IEEE 118-bus Test System [88].

4.4 Improved Artificial Bee Colony Based on Orthogonal Learning

There are two critical issues on modern heuristic optimization techniques, that is, exploration and exploitation, where the former is the capability of creating diverse population in search space, and the latter is the ability to make the best decision given current information [74]. In reality, however, the two aspects are contradictory to each other and therefore a well balanced approach needs to be found. The search process of ABC performs well for exploration; however, it performs poorly for exploitation which leads to poor convergence [75]. In order to enhance the ability of exploitation, inspired by differential evolution (DE) researchers proposed a search mechanism which utilizes the information of current best solution. In other words, onlooker bees only search around the best solution of the previous iteration according to a predefined probability [76]-[77]. Reference [78] improved the initialization phase in that the chaotic system was utilized, and modified the search mechanism using the information of current best solution.

The search equation of the original ABC randomly selects a dimension of the solution vector and performs mutation with the same dimension of another solution vector. Here the dimension refers to the number of control variables in a solution vector. For example, if the solution vector consists of 24 control variables, it is interpreted as 24 dimensions in such solution vector. However, this search scheme falls short of effectiveness, because one solution vector may contain useful information on some dimensions while the other solution may contain good information on its other dimensions. In other words, merely concentrating on a specific dimension of the solution will be likely to lose other useful information for solution improvement. Therefore, in order to update the solution considering all the information of each dimension from two

candidate solutions, inspired by the orthogonal experimental design (OED) we propose an orthogonal learning (OL) technique to obtain better exploitation. The OED is utilized to determine the best combination out of two vectors via a relatively small number of experimental tests instead of exhaustive trials [79]-[80]. The OL strategy is implemented with the help of OED, and details of such strategy will be described later.

Thus far, the application of ABC based on orthogonal learning on power system operation problems has not been documented in the literature yet. Here we first propose this method to handle the OPF problem. The performance was tested on modified IEEE 30- and 118-bus test systems and comparative analysis was conducted with other methods.

4.4.1. Orthogonal Learning Strategy

The OL method in ABC is the strategy which is analogous to orthogonal experimental design (OED) in order to obtain the best candidate solution with few searching combinations. The OED was first introduced by R. A. Fisher in the 1920's to study the effect of multi-variables/factors to the experimental output. As a powerful statistical tool, the OED was utilized to discover how much rain, water, fertilizer, sunshine, etc., are required to produce the best crop [81]. To illustrate the concept of OED the following simple chemical reaction experiment is considered as shown in Table 4.7 [80].

In this experiment, there are three *factors*: temperature (A), amount of Oxygen (B) and percentage of water (C) determining a chemical conversion rate. In addition, each factor contains three *levels*. For instance, the water can be 5%, 6%, or 7%. Thus there are $3^3 = 27$ total number of combinations that need to be experimented to find the best

conversion rate. However, with the help of OED, the best combination can be predicted by only testing few representative combinations, thus reducing total testing cost. Following describes the definition of orthogonal array and factor analysis, which leads to a comprehensive understanding of OL.

Table 4.7. Chemical reaction experiment [80]

	Factors		
	A	B	C
Levels	Temp. °C	Oxygen (cm^3)	Water (%)
1 *L*1	80	90	5
2 *L*2	85	120	6
3 *L*3	90	150	7

1) Orthogonal array: First we use '$L_N(s^k)$' to denote an array with s levels per factor for k factors, and L and N respectively represent an array and the total number of combinations. For example, in the chemical reaction experiment given in Table 4.6 we define an array $L_9(3^3)$ with 3 factors, 3 levels per factor, and 9 combinations,

$$L_9(3^3) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 3 & 3 \\ 2 & 1 & 2 \\ 2 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 3 \\ 3 & 2 & 1 \\ 3 & 3 & 2 \end{bmatrix} \tag{4.19}$$

An $N \times k$ array A is defined as an orthogonal array (OA) which has *index* λ with *strength* t on $0 \le t \le k$ when each $N \times t$ sub-array of A contains all the combinations of t-tuple exactly λ times as a row [82]. Equation (4.19) gives an example of a 9×3 OA with

strength 2 and index 1 (tuples (1,1) (1,2) (1,3) (2,1) (2,2) (2,3) (3,1) (3,2) (3,3) appear in any two columns one time). Note that an array with strength 3 and index 1 yields the full 27 combinations of triplets.

An OA is a predefined table for the OED method to work on. As mentioned earlier, the benefit of utilizing OED is to obtain the best combination by conducting only few experiments. The total nine experiments specified by the $L_9(3^3)$ are presented in Table 4.8. For instance, the first row is [1 1 1], which means that the factors A (Temperature), B (Oxygen), and C (Water) are all designed to the first level (80°, 90cm^3, and 5%, respectively). The last column shows the results of the experiment for each combination.

Table 4.8. Best combination levels By OED

Comb.	A: Temp. (°C)	B: Oxygen (cm^3)	C: Water (%)	Results (reaction rate)
*Cb*1	(1) 80	(1) 90	(1) 5	$f_1 = 31$
*Cb*2	(1) 80	(2) 120	(2) 6	$f_2 = 54$
*Cb*3	(1) 80	(3) 150	(3) 7	$f_3 = 38$
*Cb*4	(2) 85	(1) 90	(2) 6	$f_4 = 53$
*Cb*5	(2) 85	(2) 120	(3) 7	$f_5 = 49$
*Cb*6	(2) 85	(3) 150	(1) 5	$f_6 = 42$
*Cb*7	(3) 90	(1) 90	(3) 7	$f_7 = 57$
*Cb*8	(3) 90	(2) 120	(1) 5	$f_8 = 62$
*Cb*9	(3) 90	(3) 150	(2) 6	$f_9 = 64$
levels	Factor Analysis			
*L*1	H_{A1}= $(f_1+f_2+f_3)/3=41$	H_{B1}= $(f_1+f_4+f_7)/3=47$	H_{C1}= $(f_1+f_6+f_8)/3=45$	
*L*2	H_{A2}= $(f_4+f_5+f_6)/3=48$	H_{B2}= $(f_2+f_5+f_8)/3=55$	H_{C2}= $(f_2+f_4+f_9)/3=57$	
*L*3	H_{A3}= $(f_7+f_8+f_9)/3=61$	H_{B3}= $(f_3+f_6+f_9)/3=48$	H_{C3}= $(f_3+f_5+f_7)/3=48$	
OED Results	*A*3	*B*2	*C*2	

2) Factor analysis: Factor analysis (FA) is to evaluate the effects of each factor on the experimental results in order to determine the best combination of levels. With all N cases of experimental results of OA known, The FA is conducted to determine the best combination. The process of FA is described as:

To determine the effect of each level for each factor, H_{ks} is evaluated as the average effect of level s (s = 1, 2, 3) for the k-th factor ($k = A, B, C$),

$$H_{ks} = \frac{\sum_{n=1}^{9} f_n \times z_{nks}}{\sum_{n=1}^{9} z_{nks}} \tag{4.20}$$

where f_n is the experimental result of the n-th (n=1,2…,9) combination, z_{nks} is 1 if in the n-th (n=1,2…,9) combination, the level of the k-th factor ($k = A, B, C$) is s (s = 1, 2, 3), otherwise is 0. For instance if we want to evaluate the effect of level 1 in factor B ($B1$), by inspection from the 3-rd column of Table 4.6 we find that combinations $Cb1$, $Cb4$ and $Cb7$ involve all the experiments of level 1 for factor B, with the corresponding experimental results f_1 = 31, f_4 = 53 and f_7 = 57, and the average effect H_{B1} = 47. After computing the effect of all levels for each factor, the most effective level for each factor can be determined by selecting the highest quantity of H_{ks} for each factor. The FA results can be found in Table 4.8 and the details of FA is explained in [79]-[80]. From Table 4.7, the best combination determined by FA is ($A3$, $B2$, $C2$). Note that this combination (90°C, 120cm^3, 6%) is not one of the nine tested combinations. The OL will be implemented in the ABC algorithm in order to obtain the best candidate solution efficiently with few searching combinations by the analogy of OED.

3) Improved ABC with orthogonal learning

As mentioned previously, the original ABC has poor efficiency on exploitation, and to overcome these issues, the OL strategy is proposed to find an efficient candidate solution. Considering the aforementioned discussion, the process of implementing OL into ABC is described below. First, a transmission vector T_j is formed:

$$\begin{aligned} T_j &= V_k + rand(0,1) \times (V_{best} - V_k) \\ k &\neq j \in [1, SN] \end{aligned} \tag{4.21}$$

where V_{best} is the best individual which has the best fitness value in current iteration, V_k is one of the SN feasible solutions different than solution V_j. The best candidate solution V_s is formed by combining the information of T_j and V_j; in other words, OL is applied to predict the best candidate solution by combining T_j and V_j with few tests as the analogy of the OED experiment in current iteration. It is worthy to mention that if OL applies to every pair of T_j and V_j, at each iteration the number of function evaluations is $SN \times (N+1)$, where N is the number of total combinations by the OA formed based on T_j and V_j. However the original ABC only has SN function evaluations, therefore it brings too much computational burden to implement OL on every pair of T_j and V_j. Hence OL is applied only to one pair selected randomly at each iteration. The overall structure of the IABC algorithm is given below:

Step 1) Initialization:

1.1) Initialize SN solutions $X_{1\ldots,SN}$ which satisfy the constraints of control variables by (4.13) randomly.

1.2) Perform Load Flow to compute the fitness values.

Step 2) Randomly select an index s from $\{1, \ldots, SN\}$.

Step 3) In *employed bees* phase ($j = 1, \ldots, SN$):

3.1) Generate a candidate solution V_j by (4.14) if $j \neq s$.

3.2) Perform Load Flow to compute the fitness.

3.3) Construct a candidate solution V_s by implementing OL if $j = s$ (Only apply OL once at each iteration to save computational cost).

3.4) Perform Load Flow to compute the fitness.

3.5) Choose a solution (from X_j and V_j) with better fitness function.

Step 4) In *onlooker bees* phase:

4.1) Choose a solution with high fitness value (roulette wheel selection scheme)

4.2) Update a new candidate solution by V_j (4.14).

4.3) Perform Load flow to compute the fitness.

4.4) Choose a solution (from X_j and V_j) with better fitness function.

4.4) The best food source is memorized.

Step 5) For all *scout bees* (execute after maximum trails):

5.1) Replace X_j with a new randomly produced solution X_j by (4.13).

5.2) Perform Load Flow to compute the fitness values.

In the employed bees phase (Step 3), the solution will be chosen between X_j and V_j whichever has the higher fitness values. Next in Step 4, onlooker phase is executed. Onlookers will select a food source under a certain probability and generates a candidate solution. The solution is selected between X_j and V_j based on their fitness as well. Finally in Step 5, the scout process is executed. After initialization, the algorithm repeats until a stop criterion is met.

An individual employed bee is randomly chosen to use OL strategy to generate a candidate solution, while other employed bees employ equation (4.14) to generate a candidate solution. The idea of adopting OL is that we want to formulate a solution vector by combining the good information of every dimension of two solution vectors. Instead of conducting exhaustive tests, OL is implemented to predict the best combination of dimensions based on two candidate solutions.

4.4.2 Optimal Power Flow Based on IABC

As described in the previous section, an orthogonal array (OA) is a predefined table for the OED method to work on. An $N \times k$ array A is defined as an orthogonal array which has *index* λ with *strength* t on $0 \leq t \leq k$ when each $N \times t$ sub-array of A contains all the combinations of t-tuple exactly λ times as a row. We use '$L_N(s^k)$' to denote an array with s levels per factor for k factors, and L and N respectively represent an array and the total number of combinations. In OPF case, 'factor' stands for the 'control variable' and 'level' means the values of such control variable. For example, since there are total 24 control variables in the IEEE-30 bus test case, a 2-level and 24-factor OA is needed, denoted by '$L_{32}(2^{24})$'. The reason why '2-level' OA is required is because that the value of each factor is determined by either transmission vector T_j or solution V_j. The following paragraph describes the procedures of implementing IABC to OPF.

To satisfy the definition of OA given previously, the 2-level and 24-factor OA is generated to be a 32×24 array with '1' or '2' entries. The procedures of generating 2-level OA can be found in appendix *A*. Appendix *B* gives the full $L_{32}(2^{24})$ OA structure. To illustrate how the OL is applied to ABC, which is the Step 3.3 in the IABC procedure, equation (4.21) gives the partial structure of OA.

As described in Step 3.3, now there are two candidate solutions V_j and T_j available, and we want to predict the best solution V_s based on V_j and T_j. First we map the values of those two solutions in OA: whenever the entry is '1', the corresponding value in V_j is chosen, and whenever the entry is '2', the corresponding value in T_j is chosen as shown in Table 4.9. Then with the help of *Factor Analysis,* OED can predict the best combination by conducting only few experiments. Factor analysis (FA) is to evaluate the effects of each factor on the experimental results in order to determine the best combination of levels. The construction of the predicted solution V_s is summarized as follows:

1) Generate a 2-level OA $L_N(2^k)$, with $N = 2^{\lceil log_2(D+1) \rceil}$, where N denotes for the total number of combinations for an OA and D is the dimension of the problem. ('⌈ ⌉' is the ceiling bracket, meaning round the number to the integer closer to ∞). The procedure to generate a 2-level OA is given in Appendix A. The reason why 2-level OA is developed is because there are only two candidate solutions (one is the current solution V_j chosen for OL, and the other one is the transmission vector T_j) used for OL. Thus by choosing either level, the values from vector V_j or T_j will be used to combine the best solution.
2) Fill the OA $L_N(2^k)$ by the information of T_j and V_j. The OA is a 2 level, denoted by '1' and '2' and 24 factors (control variables) OA, and in such OA the value of T_j is chosen when the entry of OA is '1', and that of V_j is chosen otherwise as shown in Table 4.9.
3) Obtain N test solutions Z_n ($1 \leq n \leq N$) with the corresponding value of T_j (4.14) and V_j according to a 2-level OA $L_N(2^k)$.

4) Evaluate every test solution Z_n ($1 \le n \le N$), $f(Z_n)$, and record the best solution Z_b according to fitness values.

5) For each factor conduct FA to obtain the best level.

6) With the best levels determined in Step (4), predict the best combination solution Z_p, and evaluate Z_p.

7) If Z_p has better fitness value than Z_b, it is adopted as the candidate solution vector V_j.

$$OA(32\times 24)=\left[\begin{array}{cccc:ccc:cccc}
1 & 1 & 1 & 1 & \cdots & \cdots & 1 & 1 & 1 & 1\\
1 & 1 & 1 & 1 & \cdots & \cdots & 2 & 2 & 2 & 2\\
1 & 1 & 1 & 1 & \cdots & \cdots & 1 & 1 & 1 & 2\\
1 & 1 & 1 & 1 & \cdots & \cdots & 2 & 2 & 2 & 1\\
\hdashline
\vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots\\
\hdashline
2 & 1 & 2 & 2 & \cdots & \cdots & 1 & 2 & 1 & 2\\
2 & 1 & 2 & 2 & \cdots & \cdots & 2 & 1 & 2 & 1\\
2 & 2 & 1 & 1 & \cdots & \cdots & 2 & 2 & 1 & 1\\
2 & 2 & 1 & 1 & \cdots & \cdots & 1 & 1 & 2 & 2\\
\hdashline
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{array}\right] \tag{4.22}$$

As shown in Table 4.9, the number in the parenthesis 1 or 2 represents that the corresponding control variables from V_j or T_j is chosen. In other words, whenever the entry is '1', the corresponding value in V_j is chosen, and whenever the entry is '2', the corresponding value in T_j is chosen. The control variables of real power output, PV bus voltages, transformer taps, and reactive power compensators are denoted by 'P2, P5, etc.' With all N cases of experimental results of OA known, The FA is conducted to determine the best combination. In IEEE 30-bus OPF example, FA is illustrated in Table 4.10. The FA decides which level of each factor should be chosen, in other words, by FA, the value of each control variable can be determined based on V_j and T_j to construct V_s.

Table 4.9. Orthogonal array construction

	Factors									Fitness Value
Comb.	P2	P5	P8	P11	…	Qc21	Qc23	Qc24	Qc29	
1	(1) 49	(1) 21	(1) 21	(1) 12	…	(1) 5	(1) 3	(1) 5	(1) 3	f1=110
2	(1) 49	(1) 21	(1) 21	(1) 12	…	(2) 1	(2) 3	(2) 5	(2) 4	f2=50
3	(1) 49	(1) 21	(1) 21	(1) 12	…	(1) 5	(1) 3	(1) 5	(2) 4	f3=60
4	(1) 49	(1) 21	(1) 21	(1) 12	…	(2) 1	(2) 3	(2) 5	(1) 3	f4=70
⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮
23	(2) 65	(1) 21	(2) 16	(2) 17	…	(1) 5	(2) 3	(1) 5	(2) 4	f23=70
24	(2) 65	(1) 21	(2) 16	(2) 17	…	(2) 1	(1) 3	(2) 5	(1) 3	f24=70
25	(2) 65	(2) 38	(1) 21	(1) 12	…	(2) 1	(2) 3	(1) 5	(1) 3	f25=80
26	(2) 65	(2) 38	(1) 21	(1) 12	…	(1) 5	(1) 3	(2) 5	(2) 4	f26=90
⋮	⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮

Table 4.10. Factor analysis illustration

Level:	Factor Analysis:								
1	(f1+f2+ f3+f4)/4 =72.5	(f1+f2+ f3+f4+f23+ f24)/6= 71.7	(f1+f2+ f3+f4+f25+f 26)/6= **76.7**	(f1+f2+ f3+f4+f25+ f26)/6= **76.7**	…	(f1+f3+f23+f 26)/4 = **82.5**	(f1+f3+f24+f 26)/4 = **82.5**	(f1+f3+f23+f 25)/4 = **80**	(f1+f4+ f24+f25)/4= **82.5**
2	(f23+f24+f 25+f26)/4= **77.5**	(f25+f26)/2 = **85**	(f23+f24)/2= 70	(f23+f24)/2 = 70	…	(f2+f4+f24+f 25)/4 = 67.5	(f2+f4+f23+f 25)/4 = 67.5	(f2+f4+f24+f 26)/4 = 70	(f2+f3+f23+f 26)/4= 67.5
The predicted solution:	(2) 65	(2) 38	(1) 21	(1) 21	…	(1) 5	(1) 3	(1) 5	(1) 3

4.4.3 Case Studies Based on IABC

The following paragraph presents different cases to illustrate the performance of IABC.

1) Case 1: Minimizing Fuel Cost for IEEE 30-bus System

Case 1 is the standard OPF problem with quadratic cost function. Section 4.3.1 describes the IEEE 30-bus test system including the control variables; fuel costs coefficients, load demand, etc.

The objective in this case is to minimize the total generator fuel cost (4.2). Simulation was run 30 times in order to conduct statistical analysis. The minimum total cost from IABC is 799.321 $/h, with the maximum 799.322 $/h, the average 799.321 $/h, and zero standard deviation. Results from other methods such as basic ABC, gravitational search algorithm (GSA), linearly decreasing inertia weight particle swarm optimization (LDI-PSO), enhanced genetic algorithm (EGA), modified differential evolution (MDE), and modified shuffle-frog leaping algorithm (MSFA) were made comparison to the results from IABC. The comparison including execution time is given in Table 4.11. Fig. 4.9 shows the convergence properties of ABC, and IABC algorithms.

Table 4.11 shows that the IABC approach found the minimum solution of 799.321 $/h, less than all other methods in the literature, and faster convergence of IABC was demonstrated in Fig. 9. It is worth mentioning that for the standard quadratic fuel cost function, the improvement seems not significant because the cost function is not complex enough to show the improvement.

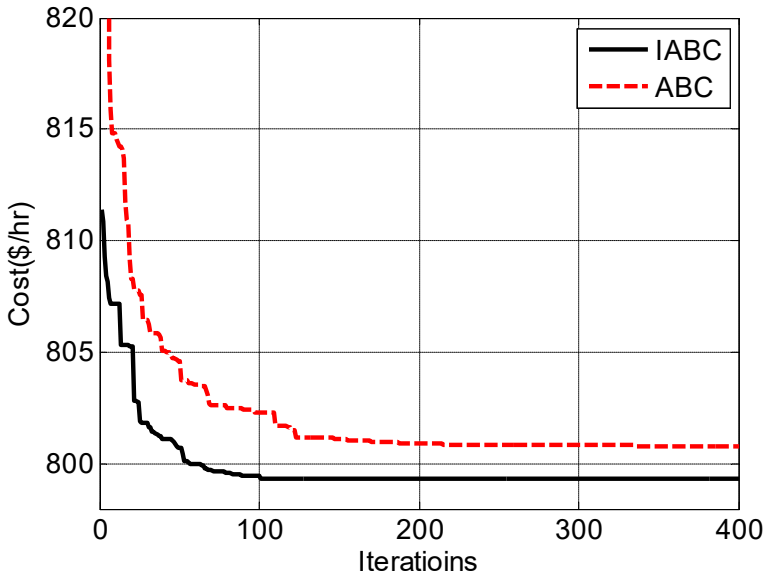

Figure 4.9. Convergence performance in case 1for IEEE-30 bus system.

Table 4.11. Comparison for fuel cost minimization in IEEE 30-bus system

Method	Fuel cost ($/h)				
	Min	Avg.	Max	Std. Dev. (σ)	t(s)
IABC	799.321	799.321	799.322	0.000	56.8
ABC	800.834	800.944	801.518	0.162	39.8
GSA [68]	805.175	812.194	827.459	N/A	10.8
LDI-PSO [69]	800.734	801.557	803.869	N/A	N/A
EGA [83]	802.060	N/A	802.140	N/A	N/A
MDE [70]	802.376	802.382	802.404	N/A	23.3
MSFLA [71]	802.287	802.414	802.509	N/A	N/A

2) Case 2: Fuel Cost with Valve-Point Effect

In this case, bus 1 and bus 2 have units with the fuel cost function with vale-point effect (4.3). Simulation was run 30 times again to obtain statistical results. The minimum total fuel cost from IABC is 918.167 $/h, the average is 919.567 $/h, the maximum is 921.458 $/h and with the standard deviation of 0.662. Table 4.11 gives the comparison with the results obtained from other methods. Note that since this case is not a typical bench mark problem, less studies are found to make comparison. The same issue applies

to the remaining two cases. Fig. 4.10 gives the convergence characteristics from the basic ABC and IABC method.

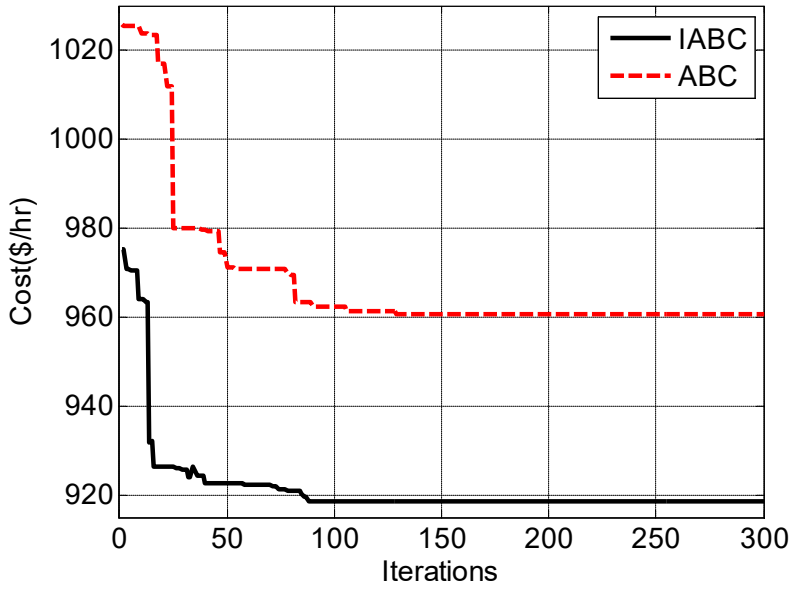

Figure. 4.10. Convergence performance in case 2 for IEEE-30 bus system.

Table 4.12. Comparison for valve-point loading effect in IEEE 30-bus system

METHOD	Fuel cost ($/h)				
	Min	Avg.	Max	Std. Dev. (σ)	t(s)
IABC	918.167	919.567	921.458	0.662	96.2
ABC	945.450	960.565	973.599	8.547	74.6
GSA [68]	929.724	930.925	932.049	N/A	N/A
MDE [70]	930.793	942.501	954.073	N/A	N/A

As shown in Table 4.12, the IABC approach found the minimum solution of 918.167 $/h, less than all other methods in the literature, and better convergence property is shown in Fig. 4.10.

3) Case 3: Loss Minimization for IEEE 30-bus System

The objective is to minimize the total real power loss as defined in (4.2). The control and state variables are identical with the previous two cases and the fuel cost function for this case is in the regular quadratic form. The results are compared with original ABC and EGA from reference [83]. The convergence property and comparison can be found in Fig. 4.11 and Table 4.13. As seen from Fig. 4.11, IABC outperforms ABC by obtaining smaller power loss. However, the convergence rates are similar.

Table 4.13. Comparison for Total power loss in IEEE 30-bus

METHOD	Total real power loss (MW)				
	Min	Avg.	Max	Stand. Dev. (σ)	t(s)
IABC	3.084	3.086	3.100	0.003	104.2
ABC	3.206	3.212	3.227	0.006	70.8
EGA [12]	N/A	3.201	N/A	N/A	N/A

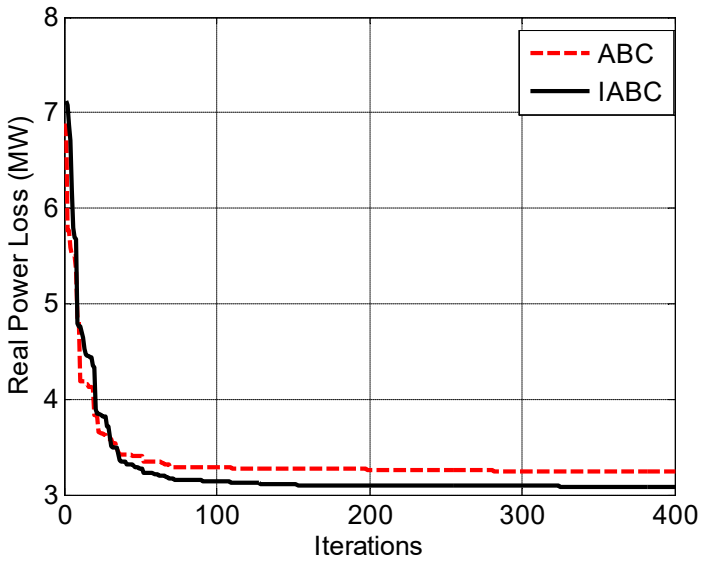

Fig. 4.11. Convergence performance in case 3 for IEEE-30 bus system.

The minimum total real power loss from IABC is 3.084 MW, the average is 3.086 MW, the maximum is 3.100 MW and with the standard deviation of 0.003.

4) Case 4: Minimizing Fuel Cost for IEEE 118-bus System

In order to test the effectiveness and robustness of IABC on a large-scale power system, IEEE 118-bus test system is adopted. For this case, there are total 130 control variables including 9 transformer tap controls, 14 shunt compensator controls, 53 real power output controls and voltage magnitude control of all 54 generators buses. Note that one of the generator bus is slack bus and thus is not considered as control/decision variable. Details of IEEE 118-bus data can be found in [84]. The minimal cost found is 129,862 $/h. The IABC was compared with regular ABC in Table 4.13. Figure 4.12 gives the convergence property.

Table 4.14. Comparison for case 4

METHOD	Fuel cost ($/h)				
	Min	Avg.	Max	Std. Dev.(σ)	t(s)
IABC	129,862	129,895	129,941	40.8	4157.8
ABC	130,210	130,321	130,410	90.5	4037.5

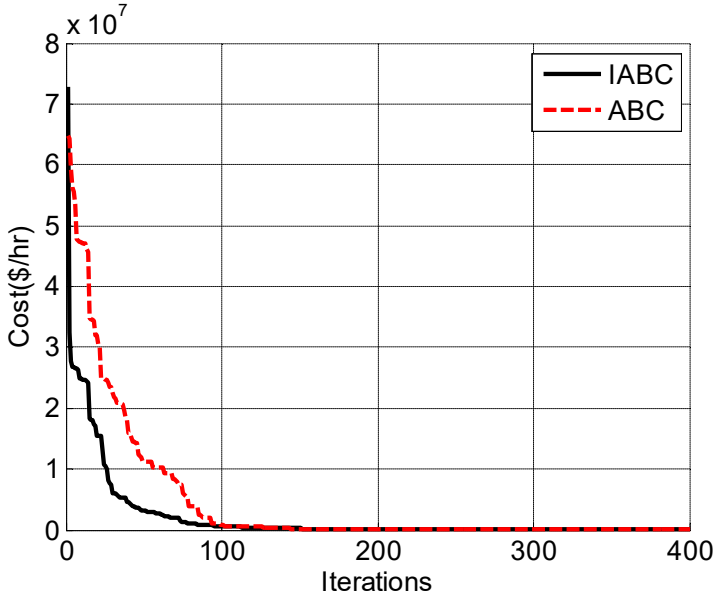

Figure 4.12. Convergence performance in case 4 for IEEE-118 bus system.

From Table 4.14, a better feasible solution can be found compared to the basic ABC. It is worth mentioning that the IABC takes longer execution time than other methods, because OL is implemented at each iteration to conduct the deep search.

4.4.4 Statistical Analysis

In order to draw convincing conclusions, statistical analysis over all cases were conducted. Fig. 4.13 presents the box plot for IABC and ABC algorithms of Case 4 study. Table 4.15 gives the one-tail paired *t*-test results.

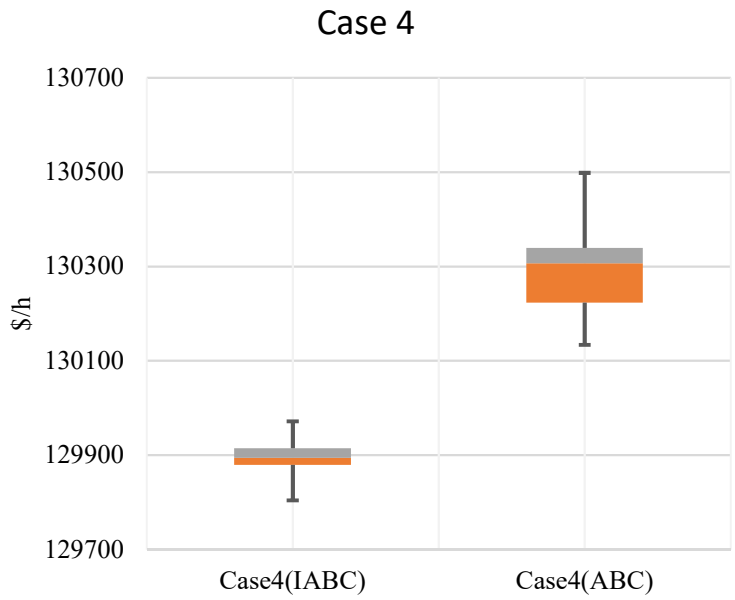

Fig. 4.13. Box plot for Case 4.

The box plot for Case 4 showed the 1st quartile, median, 3rd quartile, minimum and maximum values out of 30 simulation runs and it is obvious that the results from IABC is more consistent (smaller deviation) and better fuel cost is found. Therefore we conclude the effectiveness and robustness of IABC. Similar performance of Cases 1 through 3 is also observed and due to space the rest of box plots are not listed.

Table 4.15. Paired statistical *t*-test for IABC and ABC

Cases	ABC		IABC		
	Best	Avg.	Best	Avg.	P-value
1	800.834	800.944	799.321	799.321	5.20e-33
2	945.450	960.565	918.167	919.567	3.81e-18
3	3.206	3.212	3.084	3.086	5.79e-41
4	130,210	130,321	129,862	129,895	6.26e-23

Table 4.15 gives the *t*-test results and it is seen that the IABC outperforms ABC in all four cases at 0.05 confidence level in terms of the total generation cost and power losses. The null hypothesis (H0) is defined as that there is no differences between two algorithms and the alternative hypothesis (H1) is that the performance of IABC is better than the original ABC. Since all the p-values are smaller than 0.05, we can draw conclusion that there is significant difference between two algorithms; in other words, H0 is rejected and H1 is accepted.

It is worth to point out that it takes longer time for running IABC, because on employed bee deep search on solution space was performed in order to find promising solution. Meanwhile the computation time are not well reported for other techniques in the literature, thus making the full comparison on computation time impossible. In order to reduce computational cost, parallel computing has been placed much attention; however it is out of the scope of this work and can be a future work.

This section formulates the OPF problem and describes three objective functions for case studies. It is shown that OPF is a non-smooth, non-convex and mixed-integer optimization problem because of the same properties inherited from the objective functions and constraints. A wide range of heuristic techniques have been applied to this problem, and the balance of the exploration and exploitation ability has always been of

great importance. In order to find more promising solution, an improved ABC (IABC) optimization technique is developed based on the orthogonal learning to improve the exploitation of the basic ABC. On the employed phase, orthogonal learning is implemented to predict the best combination of two solution vectors based on limited trials instead of exhaustive trials to conduct deep search in the solution space.

In order to verify the effectiveness of proposed algorithm, IABC and basic ABC were tested on modified IEEE systems (30 and 118 bus). The results were compared with other modern heuristic methods and were able to find better feasible solutions. Different case studies and statistical analysis have demonstrated that the IABC is effective, accurate and robust with better optimization performance. In addition, IABC can be applied to large-scale power systems.

CHAPTER FIVE

Stochastic Dynamic Optimal Power Flow

5.1 Stochastic Dynamic Optimal Power Flow Formulation

In essence, stochastic optimization is to perform well under all possible scenarios, and thus the objective function of such problem is to minimize the expected value. This chapters introduces the formulation of stochastic dynamic optimal power flow (DOPF) integrated with wind and storage devices. There are typical two approaches to model wind energy in power system operation. First one is to model the wind energy as control/dispatchable variable. In other words, wind energy can be dispatched according to reference. However, since naturally wind power is not easy to dispatch, penalties on overestimation and underestimation of wind will be imposed on the objective function. Second method is to model wind power as negative *PQ* buses, provided that wind power forecast is known. Such scheme maximizes utilizing wind energy.

The work adopts the second approach treating wind power, and yet we still give discussion regarding wind power as control variable in the first subsection. Then scenario based stochastic model is introduced. Wind power forecast scenarios generated by DFM will be provided for stochastic DOPF. The original ABC is modified to tackle the dynamic optimization. Case studies regarding various cost functions have also been presented.

5.1.1 Dispatchable Wind Power

The wind energy integration model is introduced by defining wind power as a random variable and an additional penalty costs due to the unknown future wind power. The model was developed from the perspectives of system operators (SOs). The SO can own conventional generators or both conventional and wind generators, or neither of them. In this case, the SO does not own wind farms. Since the wind power is uncertain, overestimation and underestimation factors need to be considered in the model. The overestimation is when the available wind generation is less than the scheduled reference. Reserve power will be purchased from other sources to compromise the insufficiency, and otherwise load will be shed. Those activities lead to incremental cost of the SO.

Underestimation occurs when the available wind generation is greater than the scheduled reference, the SO needs to buy extra power which they have not expected from wind farms and deal with the extra power. Note that if the SO owns wind farms, the cost for underestimation penalty will not exist. The SO will usually sell the extra wind power to adjacent power grids by re-dispatching. If none of the above methods can be implemented, the excess energy can be dumped through dummy load resistors. In all, these activities can be modeled by an overestimation and underestimation penalty cost functions, and augmented to the generation cost while supporting load demands and complying with system constraints. It is worthy to mention that the constraints in economic dispatch is the minimum and maximum generator outputs and power balance, while in OPF there are constraints on generator outputs, voltage limits, transmission line capacity, transformers, power balance, etc. Thus, the following objective function is developed:

$$\sum_{i}^{M} f_i(P_i) + \sum_{i}^{N} C_{p,\omega,i}(W_{i,av} - \omega_i) + \sum_{i}^{N} C_{r,\omega,i}(\omega_i - W_{i,av})$$
$$s.t. \quad 0 \le \omega_i \le \omega_{r,i} \tag{5.1}$$

The objective function consists of the three aspects:

1) Fuel cost for thermal plants is represented in the first term; f_i is the fuel cost function depending on power p_i at the i-th unit.
2) The penalty for underestimating wind power is accounted for in the second term, where $W_{i,av}$ is a value in the range of $0 \le W_{i,av} \le \omega_{r,i,}$; ω_i and $\omega_{r,i}$ are respectively the scheduled wind power and rated power on the i-th wind generator, and $W_{i,av}$ is a random variable with probabilities varying with a given probability density function (PDF), where Weibull PDF is considered in this study.
3) The penalty for overestimating wind power is accounted for in the third term, where C_p and C_r are penalty cost functions, respectively, for underestimating and overestimating wind power.

We assume that the underestimation penalty cost and the overestimation reserve cost have linear relationships with the gap between the actual and scheduled wind generation [85]. Then the penalty and reserve cost functions, respectively, can be calculated as:

$$C_{p,\omega,i}(W_{i,av} - \omega_i) = k_{p,i}(W_{i,av} - \omega_i) = k_{p,i}\int_{\omega_i}^{\omega_{r,i}} (\omega - \omega_i) f_W(\omega) d\omega$$
$$C_{r,\omega,i}(\omega_i - W_{i,av}) = k_{r,i}(\omega_i - W_{i,av}) = k_{r,i}\int_{0}^{\omega_i} (\omega_i - \omega) f_W(\omega) d\omega \tag{5.2}$$

where k_p and k_r are the cost coefficients for penalty and reserve, respectively, and f_W is the PDF of wind power. Note that the unit of cost coefficients is '$/h•MW'. They were

used to reflect the additional cost, and demonstrate the relationship between optimal output and the coefficients.

If we want to assess both the reserve and penalty costs numerically, the PDF for the output of wind power needs to be known. In general, the PDF of wind speed is in compliance with Weibull distribution. The following paragraph derives wind power PDF $f_W(\omega)$. The PDF of wind speed is considered as Weibull distribution:

$$f(v) = \left(\frac{k}{c}\right)\left(\frac{v}{c}\right)^{k-1}\left(e^{-(v/c)^k}\right) \quad 0 < v < \infty \tag{5.3}$$

where v is wind speed, $f(v)$ is the PDF function of wind speed random variable. k and c are two factor parameters. Figs. 5.1 and 5.2 give the Weibull PDF functions for k = 1 and 2, respectively, with c = 10, 15, and 20.

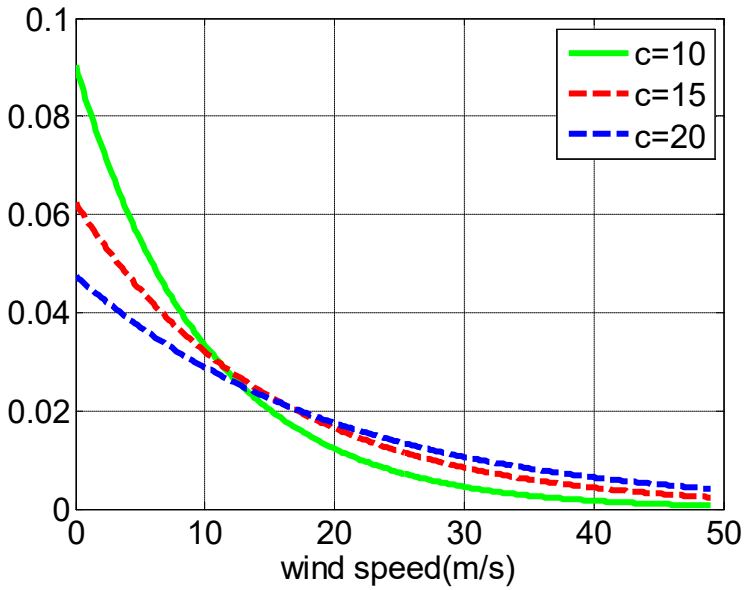

Figure. 5.1. Weibull PDF with k = 1.

The output power of wind farms is also random and can be obtained through a transformation from wind speed. Wind turbine power is related to wind speed as:

$$\begin{aligned} &\omega = 0, \quad v < v_n \quad and \quad v > v_o \\ &\omega = \omega_r \frac{(v - v_0)}{(v_r - v_n)}, \quad v_n \le v \le v_r \\ &\omega = \omega_r, \quad v_r \le v \le v_o \end{aligned} \tag{5.4}$$

where v_r is the rated wind speed, and v_n and v_o are cut-in and cut-out speeds, respectively. With a given wind speed Weibull distribution, references [85] give the details of transforming from wind speed to wind power distribution, and the transformation of two random variables are as shown below:

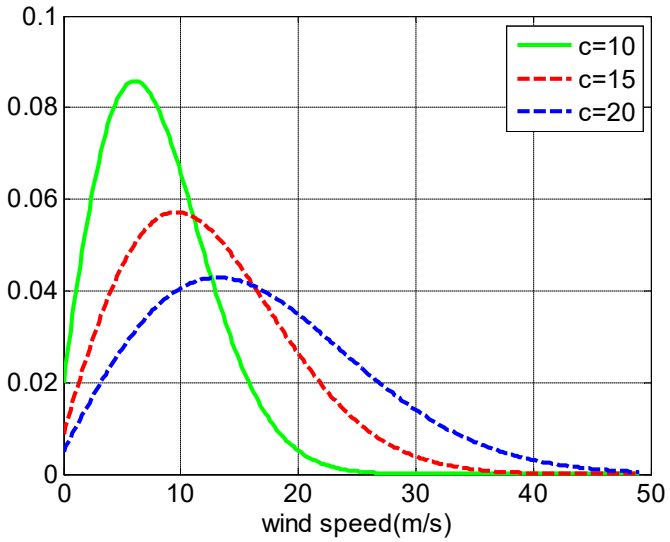

Figure. 5.2. Weibull PDF with $k = 2$.

$$\begin{aligned} &\omega = g(v) \\ &g : R \to R \\ &f_W(\omega) = f_V(v) \left| \frac{d}{d\omega} g^{-1}(\omega) \right| \end{aligned} \tag{5.5}$$

where v and ω are respectively wind speed and power random variables, g is the function that maps v to ω. Given g, the wind speed PDF $f_V(v)$ can be transformed to the wind power PDF $f_W(\omega)$ by (5.5). It is worth mentioning that the wind speed PDF $f_V(v)$ can be

obtained by historical meteorological data of a specific site, and we assume that it can be used for determining the expected value of wind speed and wind power in (5.2). The expected wind power is considered as the predicted available power.

Fig. 5.3 demonstrates the wind power PDF with normalized output of wind power corresponding to the given wind speed PDF with $k = 2$ and $c = 10$, 15, and 20.

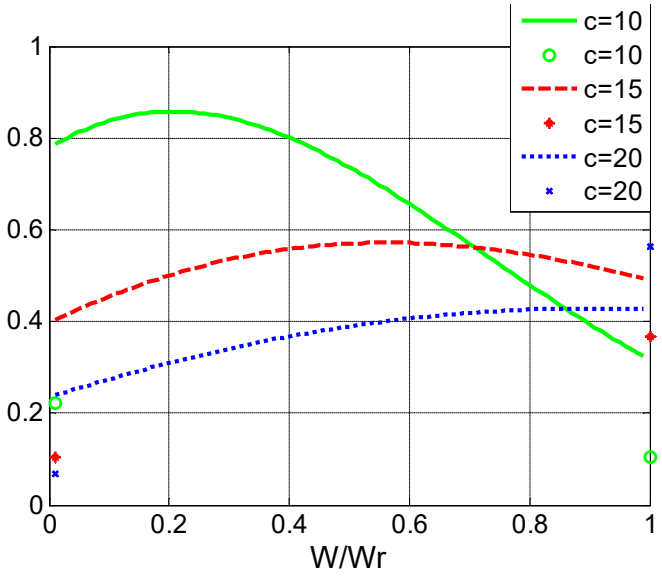

Figure. 5.3. Wind power PDF with $k = 2$ (discrete at 0 and 1; continuous between 0 and 1).

Note that the PDF of the wind power output consists of continuous random variable and discrete random variable (at 0 and 1).

5.1.2 Scenario Based Stochastic Model

A thorough description of wind power uncertainty would be by multiple random variables, and each one corresponds to a probability density function (PDF) (the dimension is the number of time steps within the prediction horizon). An approximate representation of this PDF can be obtained with a set of scenarios, sampled from the PDF representing the historical (observed) error distribution. Every wind power forecast

scenario can be treated as negative PQ buses in the sense of consuming negative power from the system (supplying power). One common way to represent the uncertainty of wind power is to assume that the wind power is subject to normal distribution $N(\mu, \sigma^2)$ with expected value μ as forecasted wind and σ as the forecasting error. Then the Monte Carlo simulation [34] is utilized to generate a large set of scenarios subject to normal distribution and each scenario is assigned a probability which quantifies the likelihood of that scenario. In the case of wind power forecast, a scenario is a time-series or a particular sequence of values representing an assumed possible realization of wind power along some period [36]. An example of 10 scenarios wind power prediction is shown in Fig. 5.4. The predicted expected value μ was obtained from Electrical Reliability Council of Texas (ERCOT) [86] on the date of Dec 21th 2014 by scaling to 1/10 of the original real power.

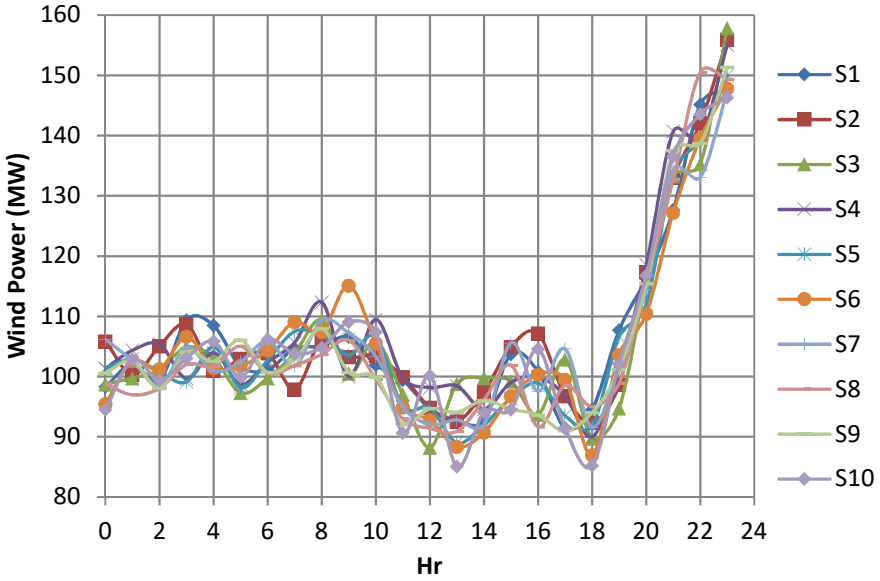

Figure. 5.4. Scenarios of wind power output over 24 hr.

As described in Chapter three, the scenarios generated by Monto Carlo simulation is not able to catch the co-movement of different wind farms. In other words, the scenarios cannot represent the spatial correlations. By using dynamic factor model, scenarios are synthesized by dynamic innovations, which are Gaussian white noises. Those scenarios capture the co-movement of wind farms as demonstrated in Chapter three. Fig. 5.5 gives the forecast scenarios by DFM.

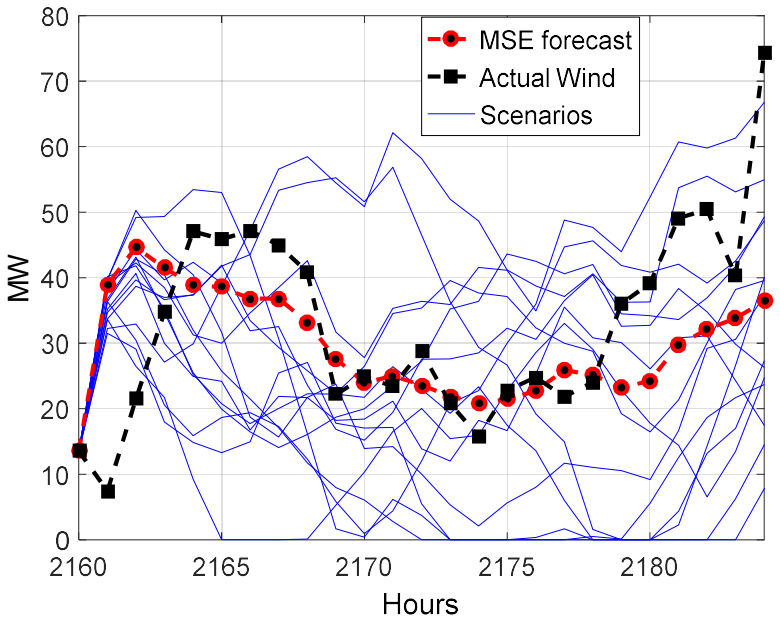

Figure. 5.5. Scenarios of wind power output by DFM.

The following content introduces the dynamic optimal power flow (DOPF), which is the optimization over a time period, energy storage device model, and the formulation of stochastic DOPF ubder given wind power forecast scenarios.

5.1.3 Traditional Dynamic Optimal Power Flow

Static/traditional OPF is to find the optimal solution at specific time; however, independent system operators (ISO) often need to make operational schedule plan one-

day ahead. Therefore dynamic OPF becomes the heart of ISO to conduct economically efficient and reliable operation on power grid. Dynamic OPF is an extension of OPF over a time horizon under dynamic constraints, which breaks the timeline into t steps (24 hours in this work) and considers OPF at each time step. The goal of traditional DOPF is to choose a set of control variables in order to minimize objective functions. In general the traditional dynamic OPF is defined as:

$$\text{Min } f(x,u) \tag{5.6}$$

$$\text{s.t. } g(x,u) = 0 \tag{5.7}$$

$$h(x,u) \le 0 \tag{5.8}$$

where f is the objective function, g is the equality constraints which represents nonlinear AC power flow equations, and h is the system inequality constraints. The vector u is the vector of independent control variables and it includes generator active power outputs P_G at PV bus, bus voltages V_G at PV and slack buses, transformer tap settings T, and shunt VAR compensators Q_C. There are three types of objective functions chosen for the study: minimization of total quadratic fuel cost, fuel cost with valve-point effect, and minimization of total power losses. The total generation cost is defined as:

$$FC = \sum_{t=1}^{24} \sum_{i=1}^{N_G} \left(a_i P_{Git}^2 + b_i P_{Git} + c_i \right) \tag{5.9}$$

The voltage improvement objective function is formed as:

$$FCV = \sum_{t=1}^{24} \sum_{i=1}^{N_G} \left[a_i P_{Git}^2 + b_i P_{Git} + c_i + \left| d_i \sin\left(e_i \left(P_{Git,\min} - P_{Git} \right) \right) \right| \right] \tag{5.10}$$

The total power loss is formed as:

$$FL=\sum_{t=1}^{24}\sum_{k=1}^{N_l}\left\{\frac{r_k}{r_k^2+x_k^2}\left[V_{it}^2+V_{jt}^2-2V_{it}V_{jt}\cos\left(\delta_{it}-\delta_{jt}\right)\right]\right\} \tag{5.11}$$

where a_i b_i c_i are the fuel cost coefficients of the i-th generating unit, N_G is the number of generators, P_{Git} is the power generated at bus i at hour t, V_{it} is the voltage at bus i at hour t, N_{pq} is the number of PQ buses, w is the weighting factor, N_l is the number of transmission lines, r_k and x_k are the resistance and reactance of the transmission line k that connects buses i and j, δ_{it} and δ_{jt} are the voltage angles at bus i and j, respectively, and t is from 1 to 24, representing one-day period. The equality constraint is the power flow equation at each bus, defined as:

$$\begin{aligned} P_{it} &= V_{it}\sum_{j=1}^{N}V_{jt}Y_{ij}\cos\left(\delta_{it}-\delta_{jt}-\theta_{ij}\right) \\ Q_{it} &= V_{it}\sum_{j=1}^{N}V_{jt}Y_{ij}\sin\left(\delta_{it}-\delta_{jt}-\theta_{ij}\right) \quad \forall t,\forall i,\forall j \end{aligned} \tag{5.12}$$

Inequality constraints are listed as generator limits, ramp rate constraints of generators, tap position of transformers, shunt capacitor constraints, security constraints, load bus voltage and transmission line flows.

$$\begin{aligned} &P_{Gi,\max}\le P_{Git}\le P_{Gi,\max} \\ &Q_{Gi,\max}\le Q_{Git}\le Q_{Gi,\max} \\ &V_{Gi,\max}\le V_{Git}\le V_{Gi,\max} \quad i\in N_G,\forall t \end{aligned} \tag{5.13}$$

$$-DR_i\le P_{Gi(t+1)}-P_{Git}\le UR_i \quad i\in N_G,\forall t \tag{5.14}$$

$$TP_{i,\min}\le TP_{it}\le TP_{i,\max} \quad i\in N_T,\forall t \tag{5.15}$$

$$Q_{ci,\min}\le Q_{cit}\le Q_{ci,\max} \quad i\in N_c,\forall t \tag{5.16}$$

$$\begin{aligned} &V_{Li,\min}\le V_{Lit}\le V_{Li,\max} \quad i\in N_{pq},\forall t \\ &S_{Lit}\le S_{Li,\max} \quad i\in N_l,\forall t \end{aligned} \tag{5.17}$$

where N is the total number of buses; P_{it} and Q_{it} are the injected real and reactive power of bus i at time t; Y_{ij} and θ_{ij} are the Y-bus admittance matrix elements; the minimum/maximum real, reactive power and voltage limits of unit i are denoted by $P_{Gi,min}$, $P_{Gi,max}$, $Q_{Gi,min}$, $Q_{Gi,max}$, $V_{Gi,min}$, and $V_{Gi,max}$; UR_i and DR_i are the ramp up/down limits of unit i, N_T is the number of tap transformers; $TP_{i,min}$, $TP_{i,max}$, $Q_{ci,min}$, $Q_{ci,max}$, $V_{Li,min}$ and $V_{Li,max}$ are the limits of transformers, shunt capacitors, and load bus voltage, respectively; and $S_{Li,max}$ is the maximum line flow of transmission line i.

5.1.4 Energy Storage System Modeling

Energy storage system (ESS) has the ability to absorb energy and dispatch energy around the network. Thus in this work, ESS are modeled as generators with the ability to inject positive or negative power onto network and state of charge (SOC) is adopted as the variable to keep track of the stored energy. It is assumed that the SOC is periodic during 24 hours, that is, the initial SOC level of each day is the same.

On those nodes installed with ESS, we define the power charged/discharged from ESS is P_{ESS}, which can be decomposed into:

$$P_{ESS} = P_{ESS}^{dis} + P_{ESS}^{cha} \tag{5.18}$$

where P_{ESS}^{dis} is the 'discharging generator' with positive power from the perspective of grid, and P_{ESS}^{cha} is the 'charging generator' with negative power. The model is shown in Fig. 5.6.

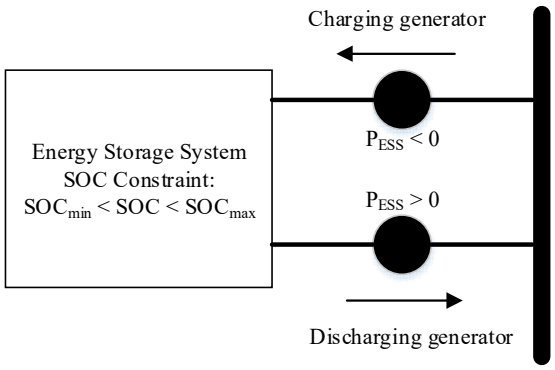

Figure. 5.6. ESS model.

Therefore the SOC at time *t* during each period considering the power flow of P_{ESS}^{dis} and P_{ESS}^{cha} is defined as:

$$SOC_{ESS}(t) = SOC_{ESS}(0) - \frac{\xi_{cha}\Delta t}{E_{ESS}^{cap}} \sum_{t'}^{t} P_{ESS}^{cha}(t') - \frac{\Delta t}{\xi_{dis} E_{ESS}^{cap}} \sum_{t'}^{t} P_{ESS}^{dis}(t') \tag{5.19}$$

As mentioned earlier, it is expected for ESS that the initial SOC(0) is the same with the final SOC(23) of one day. Where E_{ESS}^{cap} is the energy capacity of ESS, and ξ_{cha} and ξ_{dis} are the charging and discharging efficiencies of ESS, respectively. The SOC is constrained within the limits:

$$SOC_{ESS}^{\min} \le SOC_{ESS}(t) \le SOC_{ESS}^{\max} \quad \forall t \tag{5.20}$$

5.1.5 Stochastic Dynamic Optimal Power Flow Model

When considering the uncertainty of wind power forecasting error, the objective functions become the minimization of the expected value of total generation cost, voltage profile improvement or power loss among all generated scenarios. For example, the objective function of minimizing total expected fuel cost is

$$FC = \sum_{s1=1}^{15} prob^{s1} \left[\sum_{t=1}^{24} \sum_{i=1}^{N_G} c_i \left(P_{Git}^{s1} \right) + \sum_{t=1}^{24} \sum_{i=1}^{N_S} c_k \left(P_{Skt}^{s1} \right) \right] \tag{5.21}$$

where s_1 is the scenarios of wind power, $prob^{s1}$ is the corresponding probabilities of each scenario, c_i and c_k are the respective cost functions of generators and ESSs, and P_{Git}^{s1} and P_{Skt}^{s1} are the respective power of the *i*-th generator and *k*-th ESS at time *t* under scenarios. Note that the production cost of the renewable energy is negligible. The equality constraints will remain the same as (5.12) representing the power balance at each node, and all the inequality constraints will be reserved in the form that each scenario is considered. For example, the limits of load voltage:

$$V_{Li,\min} \le V_{Lit}^{s1} \le V_{Li,\max} \quad i \in N_{pq}, \forall t, \forall s_1 \tag{5.22}$$

The additional constraints are applied on the wind and ESS output power which is denoted as:

$$P_{i,\min} \le P_{it}^{s1} \le P_{i,\max} \quad \forall t, \forall s_1 \tag{5.23}$$

where P_{it}^{s1} is the wind or ESS output power under individual scenario at time *t*.

5.2 Modified ABC for Dynamic OPF

Dynamic OPF is to optimize an objective function over a time period while considering constraints such as generating units ramp-rate constraints, power balance, transmission line limits, etc. The following introduces the methodology for solving Dynamic OPF.

As mentioned previously, the original ABC, however, is designed for static optimization, while we modified the process to tackle the dynamic problem. The optimization at each hour is based on the optimization information of previous hours while satisfying the ramp constraints. For example, once the OPF at 00:00 am is solved, the solution found at 00:00 am is considered as known parameters to solve the

optimization at 1:00 am. In such case the control variables can be reduced significantly. In other words, the optimal solution was found recursively. The main structure of ABC maintains, and yet the solution found by time t is then saved as the known information for the next time t+1 as illustrated in the following:

For time t:

Step 1) Initialization:

1.1) Randomly generate SN points in the search space as feasible solutions X_i by (4.13).

1.2) Run Load Flow and evaluate the objective function by (4.18).

Step 2) From all employed bees (i = 1, …, SN):

2.1) Generate a candidate solution V_i by (4.14).

2.2) Run Load Flow and evaluate the objective function by (4.18).

2.3) Choose a solution (from X_i and V_i) with better fitness function.

Step 3) For all onlooker bees (only 'good' solutions will be executed under certain probability p):

3.1) Generate a new candidate solution by V_i (4.14).

3.2) Run Load Flow and evaluate the objective function by (4.18).

3.3) Choose a solution (from X_i and V_i) with better fitness function.

Step 4) For all scout bees (only the solution which has not been updated after a predefined maximum number of trails will be executed):

4.1) Replace X_i with a new randomly produced solution X_i by (4.13).

4.2) Run Load Flow and evaluate the objective function by (4.18).

Save the optimal solution at time t, and use it as known information to find optimal solution at t+1.

The modification enables the ABC process much fewer number of control variables at each hour compared with that of static optimization, where a large vector containing control variables for 24 hours is formed. If the original ABC was used, the control vector, which consists of real power of all generators and voltage magnitudes, transformer taps, and shunt capacitors over 24 hours, is in significantly large size. Such method has much computational burden and takes 48hrs while still can't find feasible solutions. However, by modified ABC, the solution was found in 20mins.

5.3 Case Studies and Analysis

The following paragraphs illustrate different case studies which varies from fuel cost minimization to the integration of energy storage system.

5.3.1 Case 1: Quadratic Fuel Cost Minimization

Case 1 is the standard OPF problem with quadratic cost function. The data of IEEE 30-bus test system, and control variable limits and fuel cost coefficients are the same as in Section 4.3. There are total 15 control variables which consist of real power generation at five PV buses, voltage magnitude of all six generator buses and tap-settings of four transformers. The objective in this case is to minimize the total generator fuel cost in equation (5.9).

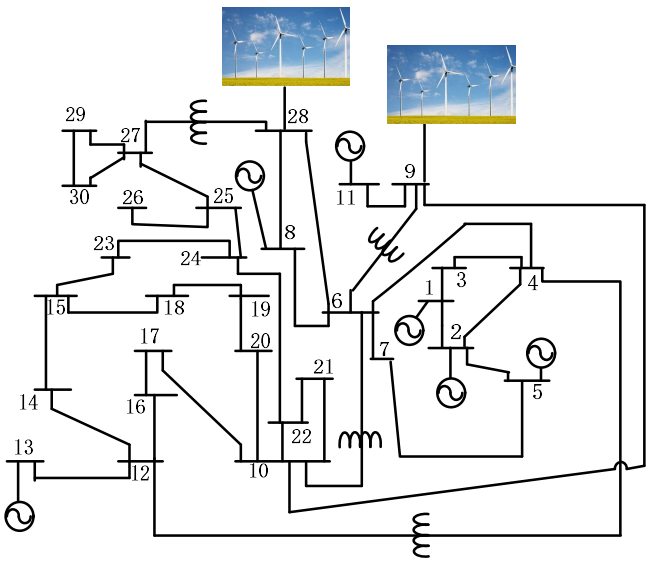

Figure.5.7. Modified IEEE-30 bus system.

As shown in Fig. 5.7, buses 28 and 9 were modified as negative PQ buses with wind farms. Table 5.1 gives the generator data including the ramp up/down constraints and the optimal dispatch is shown in Fig. 5.8.

Table 5.1 Generator data

Buses	Pmax (MW)	Pmin (MW)	Qmax (Mvar)	Qmin (Mvar)	Ramp up (MW)	Ramp down (MW)
1	250	50	200	-40	60	60
2	80	20	200	-40	40	40
5	50	15	160	-30	30	30
8	35	10	120	-30	30	30
11	30	10	100	-20	15	15
13	40	12	120	-30	15	15

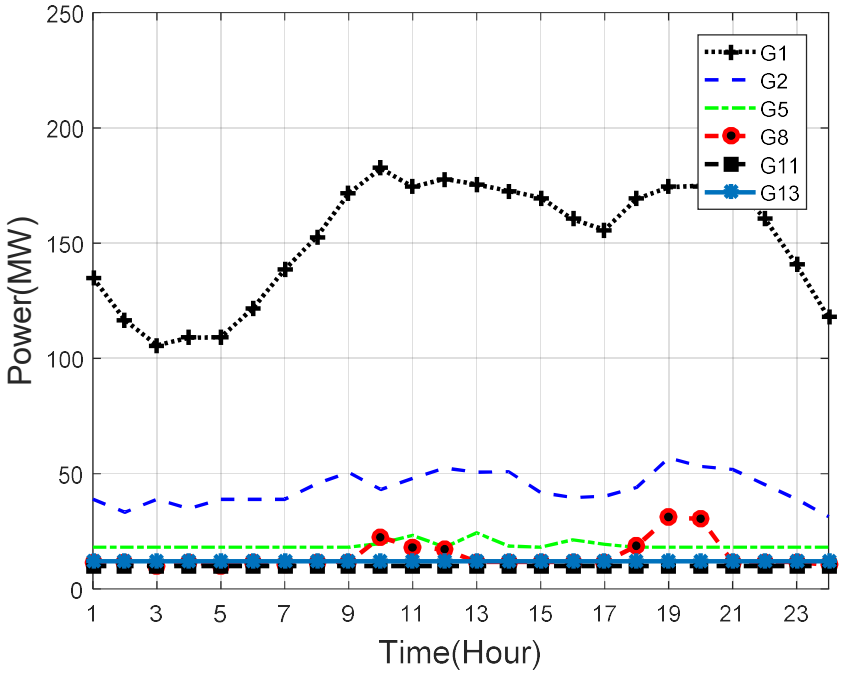

Figure 5.8. Power dispatch of six generators in Case 1.

Load profile are obtained from Electric Reliability Council of Texas (ERCOT) [86] on the date of Dec 21st 2014. The dispatch scheme considers all the scenarios of wind forecast and as verified with Table 5.1, the ramp constraints were all complied. It is noticed that Generator 1 plays the major role in supplying and shrinking power due to its lower cost coefficients. The total generating cost is $15,929.

5.3.2 Case 2: Quadratic Fuel Cost with Valve-Point Effect Minimization

This case aims to minimize the cost function defined in equation (5.10). The control and state variables are identical with the previous case and the optimal dispatch is shown in Fig. 5.9. The total operating cost is $20,553. It can be seen that such cost function imposes additional cost because of the valve-point effect.

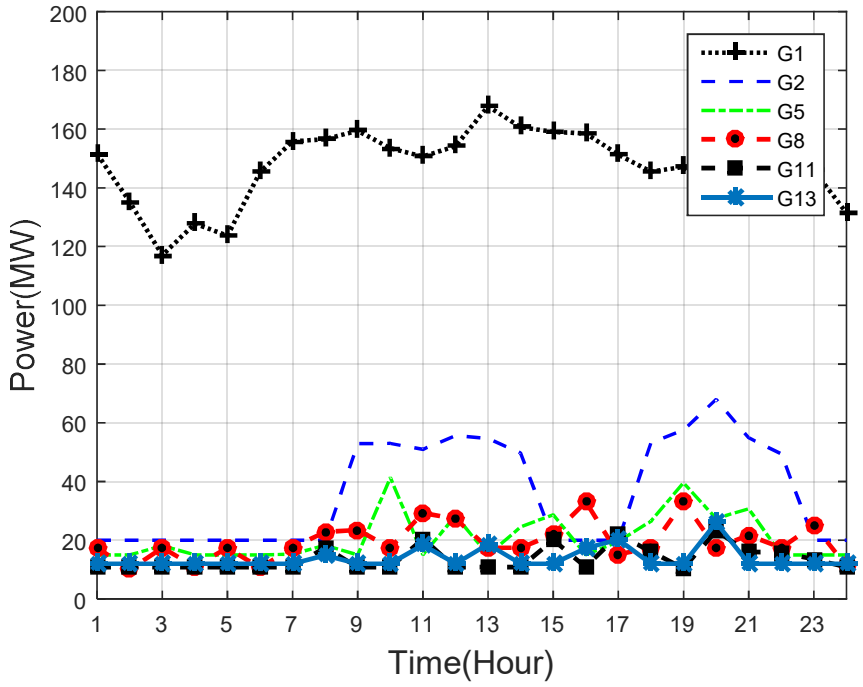

Figure 5.9. Power dispatch of six generators in Case 2.

5.3.3 Case 3: Loss Minimization

The objective is to minimize the total real power loss as defined in equation (5.11). The control and state variables are identical with the previous case and the optimal dispatch is shown in Fig. 5.10. The minimum total real power loss is from 99.98 MW. It is obvious that with different control objective, the optimal dispatch has different forms. For example generator at bus 5 always outputs its maximum power as opposed to the control scheme in Case 1. Due to the limit of space, all control variables including PV bus voltages are not shown, and yet they are within the limits between 0.95 to 1.05 p.u.

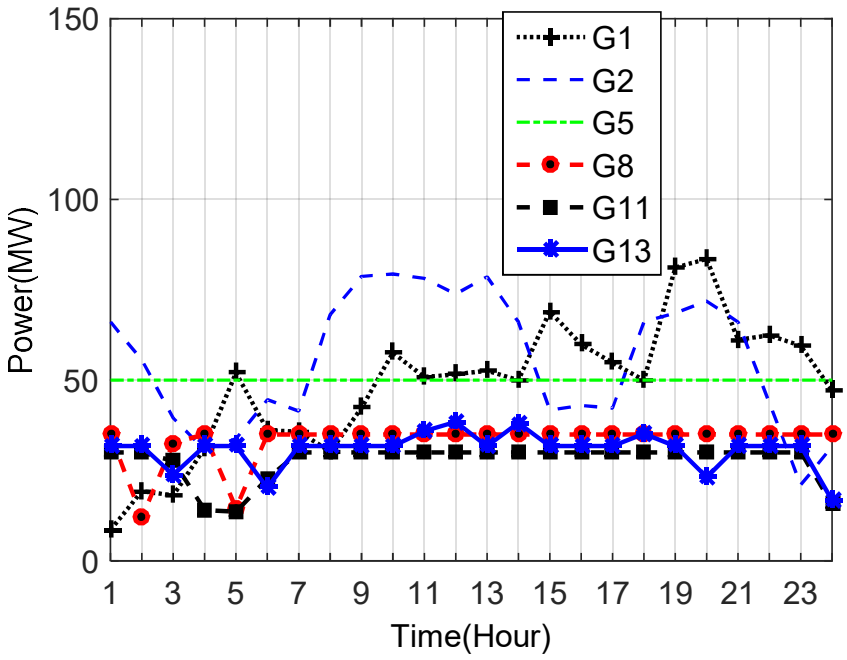

Figure 5.10. Power dispatch of six generators in Case 3.

5.3.4 Case 4: Impact of the Energy Storage System

In this case, the system is modified with three generation buses, which are buses 1, 2 and 5, to see how ESS may affect the generation profile. ESS is installed at bus 13. We investigate cases by considering constant fuel cost and dynamic fuel cost.

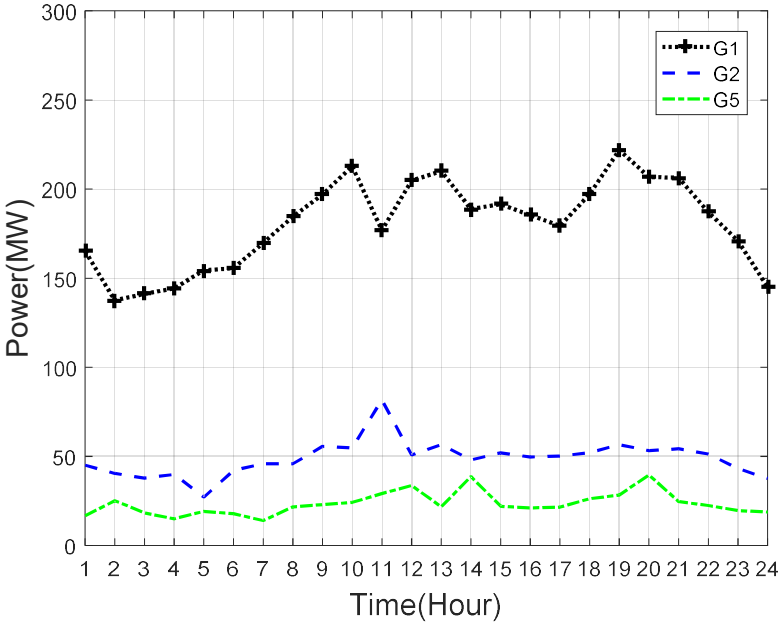

Figure 5.11. Power dispatch of 3 generators without ESS under constant fuel cost.

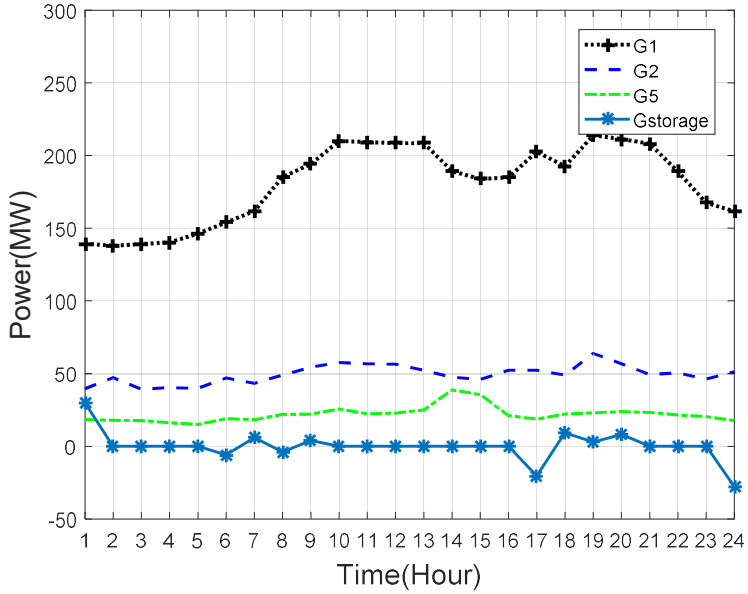

Figure 5.12. Power dispatch of 3 generators with one ESS at bus 13 under constant fuel cost.

Figure 5.11 shows the power dispatch of 3 generators without ESS installed under constant fuel cost over 24 hours as a base case for comparison. It is found that small ramps occurred on three generators. The total fuel cost is $16,203. Figure 5.12 depicts the

power dispatch of 3 generators with one ESS installed at bus 13 under constant fuel cost over 24 hours. Negative power of storage means to receive power from the grid. The total fuel cost is $16,105. Another observation is that power generation profiles of buses 2 and 5 were smoothened due to the installation of ESS.

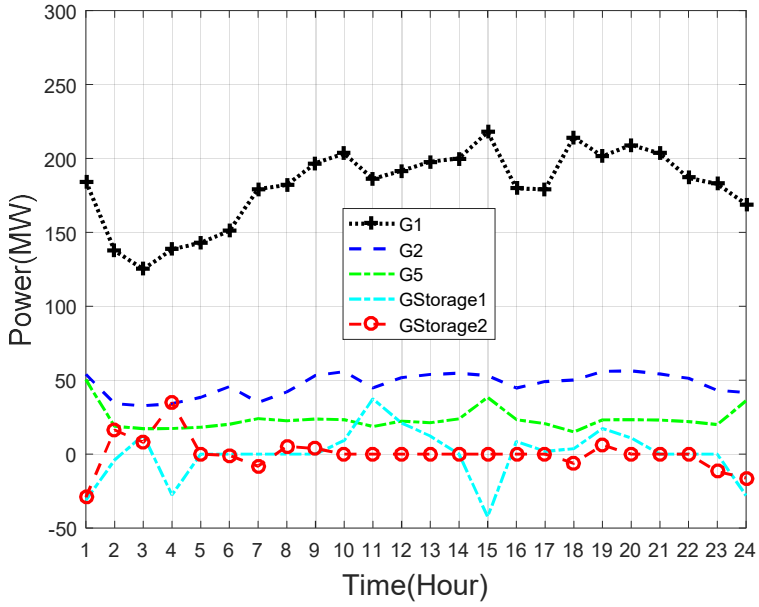

Figure 5.13. Power dispatch of 3 generators with two ESSs at buses 11 and 13 under constant fuel cost.

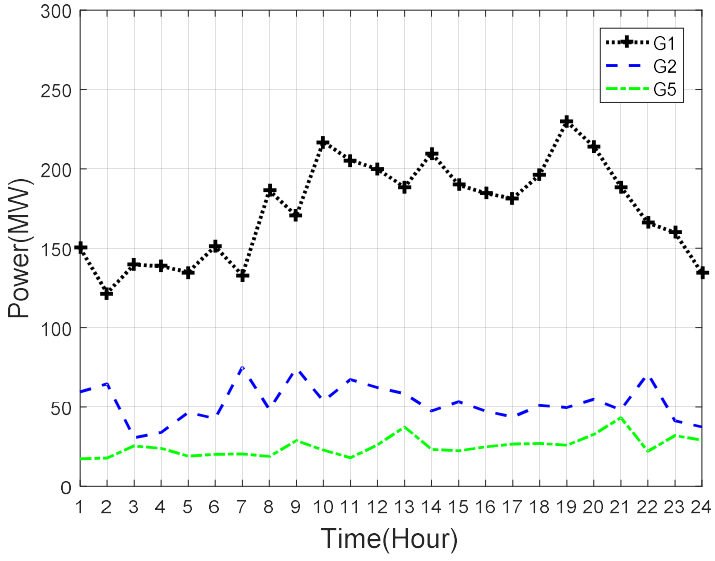

Figure 5.14. Power dispatch of 3 generators without ESS under dynamic fuel cost.

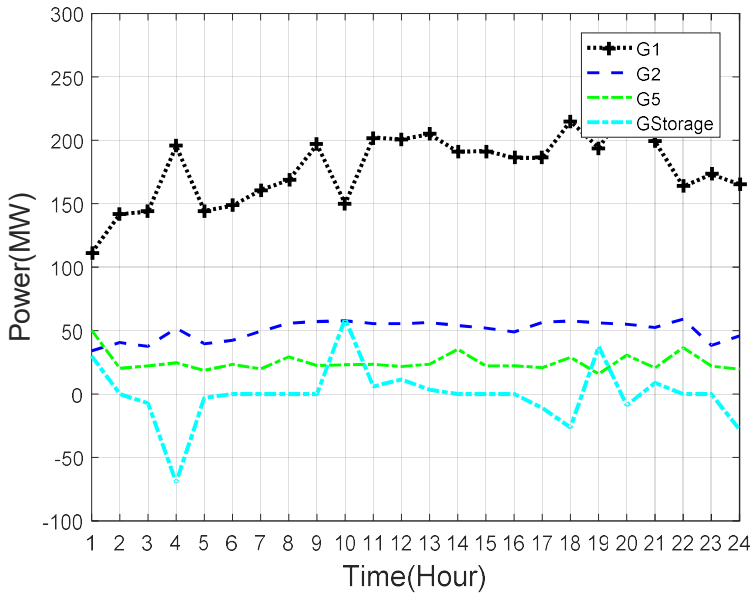

Figure 5.15. Power dispatch of 3 generators with one ESS at bus 13 under dynamic fuel cost.

Figure 5.13 depicts the power dispatch of 3 generators with two ESSs installed at buses 11 and 13 under constant fuel cost over 24 hours. Negative power of storage means to receive power from grid. The total fuel cost is $16,085, which is less than the previous two cases where no ESS and only one ESS were installed. It is also observed that power generation profiles of buses 2 and 5 were smoothened due to the installation of two ESSs.

In order to further investigate the impact of ESS, we tested the system by considering the dynamic fuel cost. In other words, generation fuel cost varies in accordance with the load. Figure 5.14 gives the generation profiles and again we inspect small ramps on profiles. The total cost is $35,363. Figure 5.15 shows the generation profiles of three generators and one ESS installed at bus 13 under dynamic fuel cost. By inspecting the profile of ESS, we found that whenever the fuel price is low, ESS charges power (negative power) and whenever fuel price increases, it discharges power to grid.

The load profile is shown in Figure 5.16. The total cost is $34,346. It is also observed that by installing ESS, generation profiles of buses 2 and 5 have been smoothened.

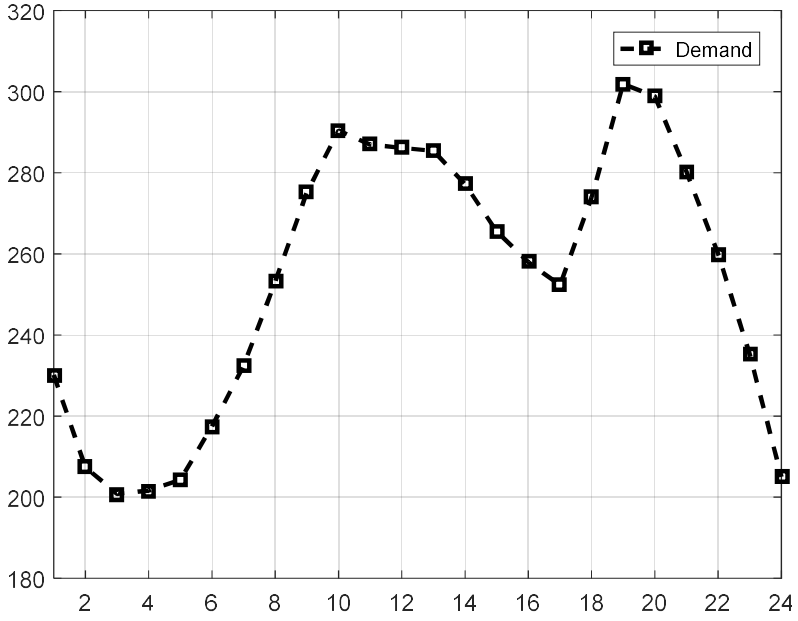

Figure 5.16. Load Profile.

Figure 5.17 shows the generation profiles with two ESSs installed at bus 11 and 13 under dynamic fuel cost. By inspection, we see that the generation profiles on buses 2 and 5 have been smoothened even further. The total fuel cost is $34,024. Therefore by inspection, we conclude that ESS can mitigate the total operational cost and with more ESS installed, smoothen certain generation profiles. In other words, due to the flexibility of charging and discharging, ESS is able to mitigate risk of ramp effect which may be caused by the integration of renewable energy. If considering dynamic fuel cost, ESS is capable of shifting power in the sense that it would store energy when the fuel price is low, and dispatch energy when the price is high in order to save the cost.

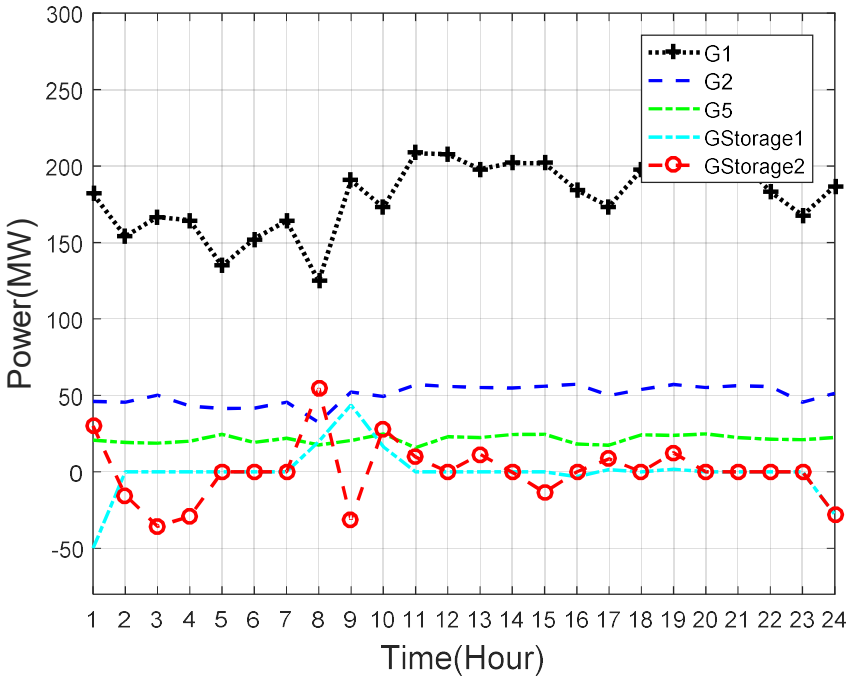

Figure 5.17. Power dispatch of 3 generators with two ESSs at buses 11 and 13 under dynamic fuel cost.

CHAPTER SIX

Conclusion and Future Studies

6.1 Summary

In summary, this work has focused on the power system operational challenges with high wind power penetration. In order to accommodate the uncertainty and variability of wind power, the dynamic factor model (DFM) is proposed to forecast wind energy by considering both temporal and spatial correlation of data. One of the advantages of DFM is that temporal and spatial correlated scenarios can be generated for the stochastic dynamic optimal power flow, which aims to optimize the system operation on various objectives, such as minimizing operating cost, power loss, etc. Due to the high non-linear, non-convex, discontinuity properties of the complex dynamic optimal power flow problem, modern heuristic optimization techniques have stood out because of the ability to solve such problem and easy implementation. We improved artificial bee colony (ABC) based on orthogonal learning to improve the balance of exploration and exploitation ability. Promising results are found compared to other methods listed in the literature in terms of finding the lower cost efficiently and robustly. Lastly the original ABC is modified to fit for dynamic optimization problem. Several case studies including minimizing quadratic cost, quadratic cost with valve-point effect, and power loss, and impacts of ESS, have been used to test the stochastic optimization.

6.2 Conclusions

The following are the conclusions of this dissertation:

1) The dynamic factor model (DFM) can successfully capture the spatial and temporal correlation among 96 wind farm data, which is more practical, compared to traditional vector autoregressive models.

2) The DFM is able to reduce the computational burden significantly, owing to the dimension reduction technique and principle component analysis (PCA). In addition, DFM can generate forecast scenarios which represent the uncertainty of forecasting.

3) Modern heuristic techniques are good tools for solving complex non-linear and non-convex problems. Promising results can be found because such techniques do not approximate original problems.

4) A good balance between exploration and exploitation needs to be established in heuristic methods. The original ABC, strong for exploration, falls short in exploitation. The improved ABC based on orthogonal learning has been proposed to enhance the exploitation and promising results are found by IABC.

5) Heuristic technique, ABC, has been widely applied for static optimization, and in this study, we extend its application to dynamic optimization where the problem is solved recursively.

6) By installing wind power and storage devices, the operating cost can be significantly reduced. The ESS is able to mitigate risk of ramping effect which may be caused by the integration of renewable energy.

7) If considering dynamic fuel price, the ESS is capable of shifting power in the sense that it would store energy when the fuel price is low and dispatch energy when the price is high in order to save the cost

6.3 Future Studies

The followings can be continued in future studies base on the results provided in this work:

1) The comparison on deterministic optimization and stochastic optimization on power system integrated with wind can be a future work. Such comparison includes the economic effects and system reliability. Since economic results are greatly affected by ancillary services, the objective function needs to include the ancillary services cost.

2) There are typically two approaches of operating the system considering wind energy: one is to treat wind energy as negative load and the other is to treat the wind power as dispatchable resources such that the exceeding wind power can be placed in storage devices. The comparison between those two approaches is a future research topic.

3) Detailed model of storage devices are to be developed for accurate control. For example, since storage devices have their lifetime, the degradation cost can be added in the model.

4) Economics in the power market is always of interest. The stochastic dynamic OPF developed in this work can be a fundamental tool for power market analysis. For example, to determine financial transmission right (FTR) over entities is

essentially a complex optimization problem with various constraints in wind integrated power system.

APPENDIX

APPENDIX

A. Construction of 2-level and D-factors OA[87]:

OA = Generate_OA (*D*)

{ $n := 2^{\lceil log_2(D+1) \rceil}$;

For i := 1 **to** n **do**

For j := 1 **to** D **do**

level := 0;

k := j;

mask := n/2;

While k > 0 **do**

If ($k$mod2) **and** (bitand(i-1, *mask*) ≠ 0) **then**

level := (*level* + 1) mod 2;

$k := \lfloor k/2 \rfloor$;

mask := *mask*/2;

$OA[i][j]$:= *level* + 1;}

where the '⌈ ⌉' is the ceiling bracket, meaning round the number to the integer closer to ∞, '⌊ ⌋' is the floor bracket, meaning round the number to the integer closer to 0. 'bitand(integ1, integ2)' returns the bit-wise AND of values of integer1 and integer2. For example:

a = 60 (60 = 0011 1100), b = 13 (13 = 0000 1101);

bitand(a,b) returns 12 (12 = 0000 1100).

B. $L_{32}(2^{24})$ *OA*:

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2
1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	2
1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1
1	1	1	2	2	2	2	1	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2	1
1	1	1	2	2	2	2	1	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1	2
1	1	1	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	2	2	2	2	2
1	1	1	2	2	2	2	2	2	2	2	1	1	1	1	2	2	2	2	1	1	1	1	1
1	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1
1	2	2	1	1	2	2	1	1	2	2	1	1	2	2	2	2	1	1	2	2	1	1	2
1	2	2	1	1	2	2	2	2	1	1	2	2	1	1	1	1	2	2	1	1	2	2	2
1	2	2	1	1	2	2	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1	1	1
1	2	2	2	2	1	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2	1	1	1
1	2	2	2	2	1	1	1	1	2	2	2	2	1	1	2	2	1	1	1	1	2	2	2
1	2	2	2	2	1	1	2	2	1	1	1	1	2	2	1	1	2	2	2	2	1	1	2
1	2	2	2	2	1	1	2	2	1	1	1	1	2	2	2	2	1	1	1	1	2	2	1
2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1
2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	2	1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	2	1	2	1	2	1	2	1	1	2	1	2	1	2	1	2	2
2	1	2	1	2	1	2	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	1
2	1	2	2	1	2	1	1	2	1	2	2	1	2	1	1	2	1	2	2	1	2	1	1
2	1	2	2	1	2	1	1	2	1	2	2	1	2	1	2	1	2	1	1	2	1	2	2
2	1	2	2	1	2	1	2	1	2	1	1	2	1	2	1	2	1	2	2	1	2	1	2
2	1	2	2	1	2	1	2	1	2	1	1	2	1	2	2	1	2	1	1	2	1	2	1
2	2	1	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1	1
2	2	1	1	2	2	1	1	2	2	1	1	2	2	1	2	1	1	2	2	1	1	2	2
2	2	1	1	2	2	1	2	1	1	2	2	1	1	2	1	2	2	1	1	2	2	1	2
2	2	1	1	2	2	1	2	1	1	2	2	1	1	2	2	1	1	2	2	1	1	2	1
2	2	1	2	1	1	2	1	2	2	1	2	1	1	2	1	2	2	1	2	1	1	2	1
2	2	1	2	1	1	2	1	2	2	1	2	1	1	2	2	1	1	2	1	2	2	1	2
2	2	1	2	1	1	2	2	1	1	2	1	2	2	1	1	2	2	1	2	1	1	2	2
2	2	1	2	1	1	2	2	1	1	2	1	2	2	1	2	1	1	2	1	2	2	1	1

BIBLIOGRAPHY

[1]. R. Wiser, L. Berkeley, and M. Bolinger, *2008 Wind Technologies Market Report*, US Department of Energy, 2008.

[2]. *Global Wind Report - Annual market update 2010*, Global Wind Energy Council (GWEC) 2010.

[3]. A. B. Mallinson, "Justifiable Small Power Plants," *Journal of the Institution of Electrical Engineers,* Vol. 63, No. 345, pp. 896-901, 1925.

[4]. W. Stirling, "Wind-Driven Generators," *Journal of the Institution of Electrical Engineers,* Vol. 61, No. 323, pp. 1096-1099, 1923.

[5]. K. Y. Lee and M. A. EI-Sharkawi (Editors), *Modern Heuristic Optimization Techniques with Applications to Power Systems*, IEEE Press, ISBN: 978-0-471-45711-4, Wiley: New York, 2008.

[6]. H. D. Chiang and C. Chu, "A systematic search method for obtaining multiple local optimal solutions of nonlinear programming problems," *IEEE Trans Circuits Systems*, vol. 43, pp. 99-106, 1996.

[7]. M. R. Adaryani, and A. Karami, "Artificial bee colony algorithm for solving multi-objective optimal power flow problem," *Elect. Power and Energy Syst.*, vol. 53, pp. 219-230, 2013.

[8]. J. Y. Kim, K. J. Mun, H. S. Kim, and J. H. Park, "Optimal power system operation using parallel processing system and PSO algorithm," *Elect. Power and Energy Syst.*, vol. 33, pp. 1457-1461, Aug. 2011.

[9]. D. B. Fogel, "Asymptotic convergence properties of genetic algorithms and evolutionary programming: analysis and experiments," *Cybern. Syst*. vol. 25, pp. 389-407.

[10]. G. Rudolph, "Convergence analysis of canonical genetic algorithms," *IEEE Trans. Neural Networks* vol.5, pp. 96-101, 1994.

[11]. A. Colorni, M. Dorigo, and V. Maniezzo, "Distributed optimization by ant colonies," *Proceedings of first European Conference on Artificial Life*, Cambridge, MA: MIT Press, pp. 134-142, 1991.

[12]. J. Kennedy and R. Eberhart, "Particle swarm optimization," *Proceedings of IEEE International Conference on Neural Networks*, Perth, Australia: IEEE Press, vol. 4, pp. 1942-1948.

[13]. W. Bai and K. Y. Lee, "Modified optimal power flow on storage devices and wind power integrated system," *IEEE Power Engineering Society General Meeting*, Boston, MA, pp.1-5. Jul. 2016.

[14]. O. I. Elgerd and C. E. Fosha, "Optimum mega," *IEEE Trans. Power Apparatus Syst.*, vol. PAS-89, no.4, pp. 556-563, Apr. 1970.

[15]. Y. V. Makarov, C. Loutan, J. Ma, and P. de Mello, "Operational impacts of wind generation on California power systems," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1039-1050, May 2009.

[16]. F. C. Schweppe, *Spot Pricing of Electricity*. New York: Springer-Verlag, 1988.

[17]. B. C. Ummels, M. Gibescu, E. Pelgrum, W. L. Kling, and A. J. Brand, "Impacts of wind power on thermal generation unit commitment and dispatch," *IEEE Trans. Energy Conv*. vol. 22, no.1, pp.44-51, Mar. 2007.

[18]. A. Tuohy, P. Meibom, E. Denny, and M. O'Malley, "Unit commitment for systems with significant wind penetration," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 592-601, May 2009.

[19]. F. Bouffard and F. D. Galiana, "Stochastic security for operations planning with significant wind power generation," *IEEE Trans. Power Syst*. vol. 23, no. 2, pp. 306-316, May 2008.

[20]. E. Denny and M. O'Malley, "Wind generation, power system operation, and emissions reduction," *IEEE Trans. Power Syst*., vol. 21, no. 1, pp. 341-347, Feb. 2006.

[21]. L. Xie and M. D. Ilic, "Model predictive economic/environment dispatch of power systems with intermittent resources," in *Proc. IEEE Power Energy Soc. General Meeting*, Calgary, AB, Canada, Jul. 2009.

[22]. J. Usaola, O. Ravelo, G. Gonzalez, F. Soto, M.C. Davila, and B. Diaz-Guerra, "Benefits for wind energy in electricity markets from using short term wind power prediction tools: a simulation study," *Wind Engineering*, vol. 28, no. 1, pp. 119-127, 2004.

[23]. C. Moerhlen, *Uncertainty in wind energy forecasting*, Ph.D. dissertation, University Cork, Ireland, May 2004.

[24]. J. Wang, M. Shahidehpour, and Z. Li, "Security-constrained unit commitment with volatile wind power generation," *IEEE Trans. Power System*, vol. 23, no. 3, pp. 1319-1327.

[25]. G. Giebel, G. Kariniotakis, and R. Brownsword, *State of the art on short-term wind power prediction,* ANEMOS Deliverable Report D1.1, 2003.

[26]. L. Landberg, "Short-term prediction of the power production from wind farms," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 80, pp. 207-220, 1998.

[27]. M. Lange and U. Focken, *Physical Approach to Short-Term Wind Power Prediction*, Berlin: Springer, 15 Dec. 2005.

[28]. M. Magnusson and L. Wern, "Wind energy predictions using CFD and HIRLAM forecast," in *proceeding of the European Wind Energy Conference* EWEC'01, Copenhagen, Denmark, pp. 861-863, June 2001.

[29]. W. Bai, D. Lee, and K. Y. Lee, "Stochastic dynamic optimal power flow integrated with wind energy," *IFAC Workshop on Control of Transmission and Distribution Smart Grids*, Vol. 49, no. 27, pp. 129-134, Prague, Czech republic Oct. 2016.

[30]. R. Jursa and K. Rohrig, "Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models," *International Journal of Forecasting*, vol. 24, pp. 694-709, 2008.

[31]. P. Pinson, *Estimation of the uncertainty in wind power forecasting*, PhD. Dissertation, Ecole des Mines de Paris, Paris, France, 2006.

[32]. *Wind Power Forecasting: State-of-the-art 2009*, Argonne National Lab. Tech. Rep., Lemont, IL, USA, 2009.

[33]. X. Zhu, M. G. Genton, Y. Gu, and L. Xie, "Space-time wind speed forecasting for improved power system dispatch," *TEST*, vol. 23, pp. 1- 25

[34]. P. Pinson, H. Madsen, H. A. Nielsen, G. Papaefthymiou, and B. Klockl, "From probabilistic forecast to statistical scenarios of short-term wind power production," *Wind Energy* Vol. 12 No. 1, pp. 51-62, 2009.

[35]. W. Bai, I. Eke, and K. Y. Lee, "Heuristic optimization for wind energy integrated optimal power flow," *IEEE Power & Engineering Society General Meeting*, pp. 1-5, Jul. 2015.

[36]. A. Botterud, Z. Zhou, and J. Wang, *Use of Wind Power Forecasting in operational decisions*, Argonne National Laboratory Report, Lemont, IL, USA, 2011.

[37]. C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, and G. Conzelmann, *Wind power forecasting state-of-the-art 2009*, Argonne National Laboratory Report, 2009.

[38]. X. Zhu, and M. G. Genton, "Short-term wind speed forecasting for power system operations," *International Statistical Review*, Vol. 80, No. 1, pp. 2-23, 2012.

[39]. Z. Huang, and Z. S. Chalabi, "Use of time-series analysis to model and forecast wind speed," *Journal of Wind Engineering and Industrial Aerodynamics*, Vol. 56, pp. 311-322, 1995.

[40]. L. Xie, Y. Gu, X. Zhu, and M. G. Genton, "Short-term spatio-temporal wind power forecast in robust look-ahead power system dispatch," *IEEE Trans. Smart Grid*, Vol. 5, No. 1, pp. 511-520, 2014.

[41]. J. Morales, R. Minguez, and A. Conejo, "A methodology to generate statistically dependent wind speed scenarios," *Applied Energy*, Vol. 87, No. 3, pp.843-855, 2010.

[42]. D. Lee, J. Lee and R. Baldick, "Wind power scenario generation for stochastic wind power generation and transmission expansion," *IEEE Power & Engineering Society General Meeting*, pp. 1-5, Jul. 2014.

[43]. D. R. Brillinger, *Time Series*: *Data Analysis and Theory*. San Francisco, USA: SIAM, 1981.

[44]. M. Forni and L. Gambetti, "The dynamic effects of monetary policy: A structural factor model approach," *Journal of Monetary Economics*, vol. 57, no. 2, pp. 203-216, 2010.

[45]. C. V. Nieuwenhuyze, "A generalized dynamic factor model for the Belgian economy-useful business cycle indicators and GDP growth forecasts," *National Bank of Belgium Working Paper*, no. 80, 2006.

[46]. M. Forni, M. Hallin, L. Marco, and L. Reichlin, "The generalized dynamic factor model: One-sided estimation and forecasting," *Journal of the American Statistical Association*, vol. 100, no. 471, pp.830-840, 2005.

[47]. G. Eshel, *The Yule Walker Equation for the AR Coefficients*, available: http://www.stat.wharton.upenn.edu/~steele/Courses/956/ResourceDetails/YWSourceFiles/YW-Eshel.pdf

[48]. R. H. Shumway, and D. S. Stoffer, *Time Series Analysis and its Applications with R examples*, ISSN:1431-875X, Springer, 2011.

[49]. M. B. Cain, R. P. O'Neill, and A. Castillo, "History of optimal power flow and formulations," *Federal Energy Regulatory Commission Tech. Report*., 2012.

[50]. J. Carpentier, "Contribution to the economic dispatch problem," *Bull Soc Francaise Elect.*, Vol. 3, No. 8, 1962, pp. 431-447 (in French).

[51]. K. Y. Lee, Y. M. Park, and J. Ortiz, "A united approach to optimal real and reactive power dispatch," *IEEE Trans. Power Apparatus Syst.* Vol. 104, May 1985, pp. 1147-53.

[52]. D. Gan, R. J. Thomas, and R. D. Zimerman, "Stability-constrained optimal power flow," *IEEE Trans. Power Syst*. Vol. 15, No.2, May 2000, pp. 535-540.

[53]. W. Bai, I. Eke, and K. Y. Lee, "Improved artificial bee colony based on orthogonal learning for optimal power flow," *18th Intelligent Systems Application to Power Systems Conference*, Sep. 2015, pp. 1-6.

[54]. R. Palomin, and V. H. Quintana, "Sparse reactive power scheduling by a penalty function – linear programming technique," *IEEE Trans. Power Syst.*, Vol. 1, No. 3, 1986, pp. 31-39.

[55]. R. C. Burchett, H. H. Happ and D. R. Vierath, "Quadratically convergent optimal power flow," *IEEE Trans. Power Apparatus Syst.*, Vol. 103, No. 11, 1984, pp. 3267-3275.

[56]. H. Wei, H. Sasaki, J. Kubokawa, and R. Yokoyama, "An interior point nonlinear programming for optimal power flow problems with a novel data structure," *IEEE Trans. Power Syst.*, Vol. 13, No. 3, 2002, pp. 870-877.

[57]. X. Yan, and V. H. Quintana, "Improving an interior-point-based OPF by dynamic adjustments of step sizes and tolerances," *IEEE Trans. Power Syst.,* Vol. 14, No. 2, 1999, pp. 709-717.

[58]. M. S. Kumari, and S. Maheswarapu, "Enhanced Genetic Algorithm based computation technique for multi-objective optimal power flow solution," *Electrical Power and Energy Syst.* Vol. 32, pp. 736-742, 2010.

[59]. Y. R. Sood, "Evolutionary programming based optimal power flow and its validation for deregulated power system analysis," *Int. J. Electr. Power Energy. Syst.*, 2007, Vol. 29, No.1, 2007.

[60]. M. A. Abido, "Optimal power flow using Tabu search algorithm," *Electric Power Component and Syst.*, vol. 30, pp. 469-483, 2002.

[61]. J. B. Park, Y. W. Jeong, J. R . Shin, and K. Y. Lee, "An improved particle swarm optimization for nonconvex economic dispatch problems," *IEEE Trans. Power Systems*, vol. 25, no. 1, pp. 156-166, Feb. 2010.

[62]. D. Karaboga, *An idea based on honey bee swarm for numerical optimization*, Erciyes Univ. Kayseri, Turkey, Tech. Rep.- TR06, 2005.

[63]. C. W. Fong, H. Asmuni, and B. McCollum, "A hybrid swarm-based approach to university timetabling," *IEEE Trans. Evolutionary Comp*., Vol. 19, No. 6, pp. 870-884, 2015.

[64]. W. Bai, M. R. Abedi, and K. Y. Lee, "Distributed generation system control strategies with PV and fuel cell in microgrid operation," *Control Engineering Practice*, Vol. 53, pp. 184-193, 2016.

[65]. M. R. Adaryani, and A. Karami, "Artificial bee colony algorithm for solving multi-objective optimal power flow problem," *Elect. Power and Energy Syst.*, Vol. 53, pp. 219-230, 2013.

[66]. MATPOWER. Available: http://www.pserc.cornell.edu/matpower/

[67]. Y. M. Park and D. T. Sun, "Optimal power flow based upon P-Q decomposition," *IEEE Trans. Power Apparatus Syst.*, vol. PAS-101, pp. 397-405, Feb. 1982.

[68]. S. Duman, U. Güvenç, Y. Sönmez and N. Yörükeren, "Optimal power flow using gravitational search algorithm," *Energy Conversion and Management*, vol. 59, pp. 86-95, Jul. 2012.

[69]. A. Khorsandi, S. H. Hosseinian, and A. Ghazanfari, "Modified artificial bee colony algorithm based on fuzzy multi-objective technique for optimal power flow problem," *Electric Power Syst. Research*, vol. 95, pp. 206-213, 2013.

[70]. S. Sayah and K. H. Zehar, "Modified differential evolution algorithm for optimal power flow with non-smooth cost function," *Energy Convers. Manage.*, vol. 49, pp. 3036-42, Nov. 2008.

[71]. T. Niknam, M. R. Narimani, M. Jabbari, and A. R. Malekpour, "A modified shuffle frog leaping algorithm for multi-objective optimal power flow," *Energy*, vol. 36, pp. 6420-32, Nov. 2011.

[72]. G. L Decker and A. D. Brooks, "Valve point loading turbines," *Electrical Engineering*, vol. 77, issue 6, pp. 501. Aug. 1958.

[73]. S. Sivasubramani and K. S. Swarup, "Multi-objective search algorithm for optimal power flow problem," *Electrical Power and Energy Sysyt*. Vol. 33, pp. 745-752, 2011.

[74]. M. Crepinsek, S. H. Liu, and M. Mernik, "Exploration and exploitation in evolutionary algorithms: A survey," *ACM Computing Surveys* vol. 45, issue 3, pp35-68, Jun. 2013.

[75]. W. F. Gao, S. Y. Liu, and L. L. Huang, "A novel artificial bee colony algorithm based on modified search equation and orthogonal learning," *IEEE Trans. Cybern.* vol. 43, no. 3, Jun. 2013.

[76]. W. Gao, and S. Liu, "A modified artificial bee colony algorithm," *Computer & Operations Research*, vol. 39, pp.687-697, 2012.

[77]. W. Gao S. Liu, and L. Huang, "A global best artificial bee colony algorithm for global optimization," *Journal of Comp. and Applied Math.*, vol. 236, pp.2741-2753, 2012.

[78]. W. Gao and S. Liu, "Improved artificial bee colony algorithm for global optimization," *Information Processing Letters*, vol. 111, issue 17, pp.871-882, Sep. 2011.

[79]. Q. F. Zhang and Y. W. Leung, "An orthogonal genetic algorithm for multimedia multicast routing," *IEEE Trans. Evol. Comput.,* vol. 3, no. 1, pp. 53-62, Apr. 1999.

[80]. Z. H. Zhan, J. Zhang, Y. Li, and Y. H. Shi, "Orthogonal learning particle swarm optimization," *IEEE Trans. Evol. Comput.,* vol. 15, no. 6, pp. 832-847, Dec. 2011.

[81]. R. K. Roy, *Design of Experiments using The Taguchi Approach*, ISBN: 9780471361015, Wiley, 2011.

[82]. A. S. Hedayat, N. J. A. Sloane, and J. Stufken, *Orthogonal Arrays: Theory and Applications,* ISNB: 987-0387987668, Springer: New York, 1999.

[83]. A. G. Bakirtzis, P. N. Biskas, C. E. Zoumas, and V. Petridis, "Optimal power flow by enhanced genetic algorithm," *IEEE Trans. Power Syst.* vol. 17, pp. 229-36, May 2002.

[84]. Power System Test Case Archive. Available: https://www.ee.washington.edu/research/pstca/pf118/pg_tca118bus.htm

[85]. J. Hetzer, D. C. Yu, and K. Bhattarai, "An economic dispatch model incorporating wind power," *IEEE Trans. Energy Convers*., vol. 23, no. 2, pp. 603-611 Jun. 2008.

[86]. Electric Reliability Council of Texas. Available: http://www.ercot.com/

[87]. S.-Y. Ho, L.-S. Shu, and J.-H. Chen, “Intelligent evolutionary algorithms for large parameter optimization problems,” *IEEE Trans. Evol. Comput*., vol. 8, no. 6, pp. 522–541, Dec. 2004

[88]. One-line diagram of IEEE 118-bus test system, IIT Power Group, 2003. Available: http://al-roomi.org/power-flow/118-bus-system

[89]. J. Carpentier, “Contribution to the economic dispatch problem,” *Bull Soc Fmnc Efeci*, vol. 8, no. 3, pp. 431-447, 1962.