Voltage Stability Assessment and Control of Power Systems using Computational Intelligence

Von der Fakultät für Ingenieurwissenschaften,
Abteilung Elektrotechnik und Informationstechnik
der Universität Duisburg-Essen
zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

genehmigte Dissertation

von

Worawat Nakawiro

aus

Bangkok, Thailand

Gutachter: Prof. Dr.-Ing. habil István Erlich
Prof. Dr. Thierry Van Cutsem
Tag der mündlichen Prüfung: 3. Februar 2011

Acknowledgement

This dissertation is the product of my research activities carried out at the institute of electric power system (Elektrische Anlagen und Netze; EAN) at the University of Duisburg-Essen. I would like to express my sincere and profound gratitude to my supervisor, Prof. Dr.-Ing. habil István Erlich.

Prof. Erlich gave me an opportunity to conduct unique research in a highly challenging and interesting area. His thorough and in-depth technical skills help develop my research skills dramatically. Throughout this study period, I have been sponsored to attend several conferences in many parts of the world through his research grant. Moreover, my heartful appreciation goes to my co-supervisor Prof. Dr. Thierry Van Cutsem from the University of Liège, Beligum for serving as the external examiner, suggesting several constructive comments in order to improve the value of my dissertation. His excellent editing skill and insightful knowledge are very helpful to my future research works.

My appreciation is also extended to the other members of my examination board consisting of Prof. Dr.-Ing. habil Peter Jung, Prof. Dr.-Ing. Klaus Solbach and Prof. Dr. rer. nat. Gerd Bacher for their valuable comments and suggestions.

Moreover, I am indebted to all staff members of EAN for contributing a pleasant and inspiring working atmosphere. Special thanks go to Ayman Hoballah, Mohd Zamri Che Wanik, Robert van de Sandt and Swaroop Pappala for their cordial assistances in both official and personal matters. Unforgetably, I am very greatful to Ms. Hannelore Treutler for her professional work as secretary of EAN. I would also like to cordially thank all members of small Thai communities in Duisburg and Krefeld. My appreciation is also extended to other organizations and individuals whose names are not listed here.

I would like also to acknowledge the finanacial support from the German Academic Exchange Service (DAAD) and the University of Duisburg-Essen. Without them, it is merely impossible for me to pursue and complete this doctoral degree in Germany.

Last but not least, my utmost gratitude is delivered to all members of my family; my parents, Mr. Wichai and Mrs. Em-on Nakawiro; my brother Dr. Thanawat and his wife Dr. Daochompoo Nakawiro. Their love and encouragement substantially help me overcome

several difficulties in the past five years. I would like to dedicate this dissertation to all of them.

Worawat Nakwiro
Duisburg, March 2011

Abstract

The primary objective of this dissertation is the utilization of an integrated and effective framework for voltage stability assessment and control based on computational intelligence techniques. A method based on artificial neural network (ANN) was developed to estimate the voltage stability margin (VSM) of a power system in real time and used for initiating appropriate control actions. The developed ANN method should provide accurate estimation for any system condition. A new method for generating training samples for ANN was proposed in this dissertation in order to take correlation of loads at different locations and variation of control settings into consideration.

The next focus of this thesis is the development of a black-box optimization algorithm requiring minimum human intervention. The algorithm has to be capable of handling practical engineering optimization problems with complex cost characteristics, mixed-integer variables and a large number of constraints. An adaptive differential evolution namely JADE is extended in this thesis to consider variation of the population size namely JADE-vPS. The algorithm is featured by a parameter-free penalty approach to handle constraints. The results of benchmark problems for unconstrained optimization are very encouraging. For a voltage stability constrained optimal power flow problem, JADE-vPS outperforms the other counterparts in terms of robustness and quality of the solution.

The final investigation is emphasized on fitness approximation for computationally expensive optimization problems. For some engineering problems, the system states corresponding to a given set of inputs are determined by a time-consuming procedure, such as numerical integration methods. In evolutionary computation, this calculation must be repeated for a huge number of times. This makes the entire process sluggish and might be infeasible for real-time implementation. In this thesis, a few models that use ANN to approximate VSM during the optimization course for determining the optimal control variables of voltage stability constrained optimal reactive power dispatch problems.

Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 Motivation

In recent years, modern power systems have experienced many technical challenges due to increasing complexities in operation and structure of the interconnected power grid. Voltage stability is recognized as one of the major problems in many power systems throughout the world such as the western region (WECC) of the United States in 1996, the Chilean power system in 1997 accounting for a loss of 80% of its total load, the Hellenic system covering the entire Athens and the neighboring area in 2004 [1].

Voltage instability is mainly associated with the inability of the power system to maintain acceptable voltages at all buses in the system under normal conditions and after being subject to disturbances such as gradual load increases or outages of critical lines or generating units. The general characteristic of voltage instability is that the voltage level at different locations slightly changes after the disturbance but abruptly declines near to the collapse point. Therefore, the voltage level itself is not a good indicator. The system operator needs performance indices either in online or offline modes to determine how close the system is to the collapse and what the control actions should be carried out in that event. In offline planning activities, computational speed is generally not a problem. However, for online analysis, real-time or faster-than-real-time tools are of the key interest for monitoring and enhancing stability of the power system.

There are a number of challenges in developing such tools for online operation. First, power systems under the deregulated environment in many parts of the world are operated by several independent transmission operators. Among these operators, only limited number of information is exchanged primarily due to business competitions. This makes a study of the entire system harder than before [2]. Classical methods for stability analysis require complete information of the entire power system. They are also time-consuming for a large system. Second, the system stability is also subject to settings of various control devices which may be located outside the authorized area of control. In classical approaches, some values in the

typical range have to be assigned to these unknown parameters. On the other hand, a computational intelligence (CI) method has no requirement for the complete dataset. It can still perform the similar task based on the information actually available to the TSO.

Besides online stability assessment, different optimization problems are usually conducted in a daily power system operation. The set of optimal control variables that suits the specified objective while maintaining various constraints is determined by an optimization engine. Due to complex characteristics in power systems, many of mathematical programming techniques generally may find great difficulty to handle the problem. Some principal assumptions in convexity and availability of gradient information have to be made. Modern heuristic optimization, a branch of CI, is a promising solution to the aforementioned problem because this method does not take any restriction on the cost characteristics. However, one of the major issues in every heuristic optimization method is that the fitness function has to be repetitively evaluated. This is a serious aspect in some optimization problem where the fitness function is computationally expensive. Therefore, an alternative model has to be developed so that the entire process is speeded up while sacrificing an acceptable degree of error.

1.2 Objectives

The overall objectives of this thesis are to study the applications of CI techniques for assessing and enhancing static voltage stability margin. The specific objectives of the research described hereafter are summarized as follows.

- **Study of power system voltage stability**: There are various techniques for assessing voltage stability of the power system, such as continuation method, optimization method, or performance indices. The study emphasizes comparing performances of different indices and suggesting appropriate selection of the analysis method according to availability of data and computational budget.
- **Development of a robust global search algorithm**: Practical real time optimization problems should be solved by a global search method without considerable transformation and reduction of the cost model. This task is achieved by the evolutionary algorithms (EAs). A powerful EA without control parameters namely JADE-vPS is developed.

- **Voltage constrained optimal power flow**: To obtain a stable and secure operating condition, traditional practices rely on trial-and-error heuristics. These methods are either time consuming or imprecise. This thesis presents a methodology that can guarantee sufficient voltage stability margin for a given operating point. Practical non-convex and discontinuous characteristics of thermal generators are also considered.
- **Approximation of voltage stability margin**: Determination of a large number of some stability indices, such as continuation power flow method is quite time consuming. It is therefore hard to directly incorporate such calculation to optimal power flow (OPF) solved by an EA. Two methods namely neural network and sensitivity approach are introduced to approximate the voltage stability margin.
- **Optimization with the approximated fitness function**: This is an emerging research field in evolutionary computation. For a problem with expensive fitness function, it is sometimes a wise idea that the fitness should be approximated by a faster method, such as neural networks or statistical learning models. Few examples on reactive power optimization with voltage stability constraint are demonstrated.

1.3 Organization of Thesis

Chapter 2 introduces the fundamentals of voltage stability problems. Various analysis techniques are explained in details. Several voltage stability indices proposed in literature are reviewed and categorized to two groups. A case study is set up to demonstrate the effectiveness of each index. Finally, different countermeasures used to prevent the power system from voltage collapse are discussed.

Chapter 3 presents the computational intelligence tools that were applied in various applications. Dimensionality reduction as data pre-processing is summarized. Neural networks and the supervised learning scheme are conceptually described. Different evolutionary algorithms are extensively investigated in this thesis. A new parameter-free algorithm namely JADE-vPS is developed and statistically tested with benchmark problems.

Chapter 4 discusses two techniques to approximate the voltage stability margin: neural network approach and sensitivity approach. The key idea of these methods is to replace the

role of continuation power flow in optimal power flow (OPF) problems. A novel and quite realistic method for generating database of operating states is proposed.

Chapter 5 copes with a voltage stability constrained OPF problem. The voltage stability constraint is controlled in a similar manner to thermal and voltage limits. Simulation results reveal that the developed JADE-vPS algorithm outperforms other algorithms in terms of high-quality and robust results.

Chapter 6 deals with various OPF models as countermeasures against voltage collapse. A completely intelligent OPF without power flow calculation is introduced. A hybrid intelligent (two stage) model is also presented. The cost due to energy loss and the number of control actions of discrete devices are simultaneously minimized in a combined-objective model. A under voltage load shedding problem is formulated to take technical and economical aspects into account.

Finally, the thesis is concluded and future research directions are summarized in Chapter 7.

Chapter 2
Voltage Stability

2.1 Introduction

Voltage stability refers to the ability of a power system to maintain acceptable voltages at all buses both under normal operating conditions and after being subject to contingency conditions [3-6]. A power system enters a state of voltage instability when a disturbance results in a progressive and uncontrollable voltage decline. The primary cause of this problem is the sudden load increase, the outages of major generator and transmission line, or a combination of multiple events [5]. In the recent decades, the consequence of voltage instability has resulted in several widespread power interruptions in many power systems throughout the world [1, 3, 7]. These incidents have caused serious losses in terms of economy and public welfare. Therefore, voltage stability studies have to be incorporated in planning and operating tools of modern power systems.

Voltage stability can be classified as static voltage stability and dynamic voltage stability based on the time frame of incidents [8-11]. In dynamic considerations, exact modeling of equipments such as transformer, induction motors, SVCs, loads, etc is required. The analysis can be further categorized into small signal stability and large disturbance stability problems [3, 6]. In the first problem, differential and algebraic equations (DEAs) are linearized around an equilibrium point and eigen-methods can be applied to determine modes of voltage stability such as [12]. This information is very useful in finding the suitable locations for reactive power compensation devices and in designing controllers such as [13]. On the other hand, DEAs cannot be generally linearized when the disturbance is large. Therefore, the large disturbance voltage stability is handled by numerical integration methods such as [10, 11].

Some of voltage stability analyses involve examination of wide range of system conditions and a number of contingencies. In this case, static analysis approach becomes attractive because the method normally involves only the solution of algebraic equations and is less computationally intensive than the dynamic approach. This method provides a practical measure in terms of the distance to collapse for given contingencies. On the other hand,

dynamic methods such as time-domain simulations are very appropriate to study the mechanism, chronology leading to instability. According to [5], system dynamic associated with voltage stability are usually slow. Therefore, static methods are adequate for examining the existence of equilibrium point for a specified operating condition. . Practical indices describing voltage stability margin (VSM) of power systems are thoroughly reviewed in Chapter 4 of [3]. Some of them are selected and implemented in this thesis as will be discussed later in this chapter.

The chapter is organized as follows. Voltage stability assessment is discussed in details in section 2.2. Power system models used in voltage stability analysis are presented in section 2.2.1. Different techniques which can be applied to analyze the voltage stability problem are explained in section 2.2.2. Various voltage stability indices proposed in literatures are summarized in section 2.2.3. Section 2.3 introduces control measures for preventing the power system from voltage collapse. Finally, some conclusions are drawn in section 2.4.

2.2 Voltage Stability Assessment

The static voltage stability analysis for a given power system state involves the determination of how close the system is to voltage instability. The proximity to instability can be measured by an index preferably defined in terns of physical quantities such as load level, reactive power reserve, etc.

2.2.1 Power system models

The typical quasi-steady state model of a power system considered in voltage stability analysis is generally given by the differential and algebraic equations as follows [1, 3]:

$$\begin{aligned} \dot{\mathbf{x}} &= f(\mathbf{x}, \mathbf{y}, \lambda) \\ 0 &= g(\mathbf{x}, \mathbf{y}, \lambda) \end{aligned} \tag{2.1}$$

where $\mathbf{x}$ is the vector of state variables; $\mathbf{y}$ is the vector of algebraic variables. The variable λ is a parameter or a set of parameters that slowly changes over time so that the power system moves from an equilibrium point to another until reaching the collapse point. To

simplify the power system description, a new vector $\mathbf{z} = [\mathbf{x}, \mathbf{y}]^T$ is defined so that (2.1) can be rewritten as:

$$\begin{bmatrix} \dot{\mathbf{x}} \\ 0 \end{bmatrix} = F(\mathbf{z}, \lambda) \tag{2.2}$$

In static voltage stability analysis, we focus only the operating condition reaching an equilibrium point given by $(\mathbf{z}_0, \lambda_0)$. The system equation as shown in (2.2) becomes $F(\mathbf{z}_0, \lambda_0) = 0$. Based on the singularity assumption, an equilibrium point $(\mathbf{z}_*, \lambda_*)$ which makes $\partial F(\mathbf{z}_*, \lambda_*)/\partial \mathbf{z}$ singular is mathematically defined as the saddle node bifurcation point. Such a bifurcation point is directly associated with the voltage collapse problems. It is well demonstrated in the literature that different control parameters of system components in generation and load sides affect the location of collapse points.

Under a certain assumption made on slow variation of active and reactive power demand as the key driving force of the system to the collapse point, the power flow model yields adequate results comparing to the ones from the complete dynamical model. Also, it is generally of interest to locate the system operating condition where the power flow is unsolvable because most of online decisions are made based on power flow solutions. The power flow model used in this thesis to obtain different voltage stability indices is represented by the typical nonlinear equation of active and reactive power mismatches at the system buses such as:

$$\begin{bmatrix} \Delta P(\mathbf{u}, \lambda) \\ \Delta Q(\mathbf{u}, \lambda) \end{bmatrix} = F(\mathbf{u}, \lambda) = 0 \tag{2.3}$$

where **u** represents a vector of system variables such as voltage magnitudes *U* and voltage angles δ. Besides the two variables earlier, any other variables such as reactive power generation *Q* can be swapped with *U* in **u** when a reactive power limit is reached.

The variable λ is a scalar parameter used to simulate the system load changes that drive the system to voltage collapse in the following way:

$$\begin{aligned} P_{D,i} &= P_{D0,i}(1+k_{P,i}\lambda) \\ Q_{D,i} &= Q_{D0,i}(1+k_{Q,i}\lambda) \\ &= P_{D0,i}\tan(\varphi_i)(1+k_{Q,i}\lambda) \end{aligned} \tag{2.4}$$

where $P_{D,i}$ and $Q_{D,i}$ represent the active and reactive power demand at bus *i*, respectively; $P_{D0,i}$ and $Q_{D0,i}$ are the initial active and reactive power demand before the load changes, respectively; $k_{P,i}$ and $k_{Q,i}$ are constants representing changes (either increase or decrease) in active and reactive power demand at bus *i*, respectively; φ_i is the power factor angle at bus *i*.

For a generator, the active power output of generator *i* should be modified to accommodate the changed power demand according to:

$$P_{G,i} = P_{G0,i}(1+\lambda k_{Gi}) \tag{2.5}$$

where $P_{G0,i}$ is the initial active power generation of bus *i*; $k_{G,i}$ is the constant specifying the rate of change in generation when λ is varied.

2.2.2 Methods of analysis

In the steady-state stability analysis, four analysis techniques are popularly used consisting of direct method, modal analysis, continuation method and optimization method. The methods are so generic that they can be applied to any system.

2.2.2.1. Direct method

This method is also known as the point of collapse method. It directly computes singular bifurcation points of nonlinear systems by solving the following equations:

$$\begin{aligned} &F(\mathbf{z},\lambda) = 0 \\ &\left[\frac{\partial F(\mathbf{z},\lambda)}{\partial \mathbf{z}}\right]^T \mathbf{w} = 0 \\ &\|\mathbf{w}\|_\infty = 1 \end{aligned} \tag{2.6}$$

The system of nonlinear equations shown above involves the power flow equations, the singularity conditions at the collapse point and the non-zero left eigenvector requirement. Methods for solving nonlinear equations such as Newton-Raphson iteration can applied to obtain the collapse point given by $(\mathbf{z}_*, \lambda_*)$. This method converges very fast if the initial operating condition is close to the critical point. Moreover, the approach is very dependent on the good estimate of λ. Comparing to the conventional power flow, the complexity of system of nonlinear equations in the direct method is also double.

2.2.2.2. Continuation method

Voltage profiles shown in the well-known PV and QV curves are of the practical use for determining the proximity to collapse so that operators can take proper preventive control actions to safeguard the system. To achieve the complete voltage profile, successive power flow solutions or continuation methods can be used [14]. The first technique provides reliable and very informative results.

Continuation methods overcomes certain difficulties in the successive power flow method because the complete voltage profile is generated by automatic changes of the loading parameter λ and overcome the singularity problem of the system equations in the vicinity of the voltage stability limit [15, 16].

The strategy used in the continuation method can be depicted in Fig. 2.1 where the system is initially at the equilibrium state $(\mathbf{z}_1, \lambda_1)$. This known equilibrium point is then used to compute the direction vector $\Delta\mathbf{z}_1$ and a change $\Delta\lambda_1$ in the so-called predictor step which thereby generates an initial guess $(\mathbf{z}_1 + \Delta\mathbf{z}_1, \lambda_1 + \Delta\lambda_1)$. Notice that this point may be not an equilibrium state. Therefore, the corrector step is applied to compute a new equilibrium point $(\mathbf{z}_2, \lambda_2)$ which lies on the system profile (bifurcation diagram). There are a number of techniques that can be used to determine the actual equilibrium point $(\mathbf{z}_2, \lambda_2)$. For example, a hyper-plane perpendicular to the tangent vector given by the predictor step can be used to find the correct point on the bifurcation diagram as shown in Fig. 2.1. A parameterization technique may be used in the continuation method to overcome the difficulty in ill-conditioning system equations close to the collapse point.

From the power flow equation given in (2.2), the fundamental equation for sensitivity analysis at an equilibrium point i.e. $(\mathbf{z}_1, \lambda_1)$ abbreviated as the superscript 1 in derivatives can be written as:

$$\left.\frac{\partial F}{\partial z}\right|_1 dz + \left.\frac{\partial F}{\partial \lambda}\right|_1 d\lambda = 0 \Rightarrow \left.\frac{dz}{d\lambda}\right|_1 = -\left[\left.\frac{\partial F}{\partial z}\right|_1\right]^{-1} \left.\frac{\partial F}{\partial \lambda}\right|_1 \tag{2.7}$$

The right hand side of (2.7) represents the tangent vector at the equilibrium point 1. The parameter change can be computed from the normalization of the tangent vector such as:

$$\Delta\lambda_1 = \frac{k}{\left\| dz/d\lambda \big|_1 \right\|} \tag{2.8}$$

where *k* is a positive scalar controlling the step size of the predictor step, named as the step size. Then, the direction vector can be computed by the sensitivity derived in (2.7) as:

$$\Delta z_1 = \Delta\lambda_1 \left.\frac{dz}{d\lambda}\right|_1 \tag{2.9}$$

With the two quantities above, it is now possible to find the next initial guess. One should observe that the step size influences the continuation process. If the step is too large, it would create some convergence problems in the corrector step whereby the small step may well trace the manifold but too many steps are required. Several techniques have been proposed to determine an optimal and adaptive step size.

Followed the predictor step with the initial guess $(\mathbf{z}_1 + \Delta\mathbf{z}_1, \lambda_1 + \Delta\lambda_1)$, the actual point $(\mathbf{z}_2, \lambda_2)$ on the system profile must be computed by solving the following equations for **z** and λ:

$$\begin{aligned} F(\mathbf{z}, \lambda) &= 0 \\ \rho(\mathbf{z}, \lambda) &= 0 \end{aligned} \tag{2.10}$$

The first vector equation correspond to the system-state (in this case power flow) equations. The second scalar equation is a phase condition that ensures non-singularity of the

system Jacobian matrix at the bifurcation point. Two phase conditions have been successfully used in the corrector techniques. The details of these two methods can be obtained from [3].

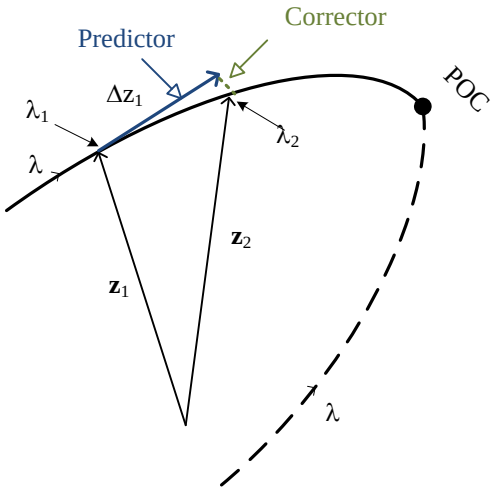


Fig. 2.1 Illustration of continuation method

2.2.2.3. Modal analysis

When the power system nearly reaches the collapse point, the power flow Jabobian matrix approaches the singularity. The singular value decomposition can be used to determine the rank of a matrix which is equal to the number of non-zero singular values. Therefore, the smallest singular value of the Jacobian matrix can be traced up to the voltage collapse point where the value becomes zero [17]. Linearizing the steady state equation $F(\mathbf{z}_0, \lambda_0) = 0$ around an equilibrium point $(\mathbf{z}_0, \lambda_0)$ gives:

$$\Delta F(\mathbf{z}, \lambda) = \mathbf{J}\Delta\mathbf{z} \tag{2.11}$$

$$\begin{bmatrix} \Delta\mathbf{P}(\boldsymbol{\delta}, \mathbf{U}, \lambda) \\ \Delta\mathbf{Q}(\boldsymbol{\delta}, \mathbf{U}, \lambda) \end{bmatrix} = \begin{bmatrix} \frac{\partial P}{\partial \delta} & \frac{\partial P}{\partial U} \\ \frac{\partial Q}{\partial \delta} & \frac{\partial Q}{\partial U} \end{bmatrix}_{(\mathbf{z}_0, \lambda_0)} \begin{bmatrix} \Delta\boldsymbol{\delta} \\ \Delta\mathbf{U} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{J}_2 \\ \mathbf{J}_3 & \mathbf{J}_4 \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{\delta} \\ \Delta\mathbf{U} \end{bmatrix} \tag{2.12}$$

where $\mathbf{J}_1, \mathbf{J}_2, \mathbf{J}_3$ and $\mathbf{J}_4$ are sub-matrices of the power flow Jacobian matrix **J**. For the real $n \times n$ Jacobian matrix, the singular value decomposition is given by:

$$J = R \sum S^T = \sum_{i=1}^{n} r_i \sigma_i s_i^T \tag{2.13}$$

where R and S are $n \times n$ orthornomal matrices whose i^{th} columns are singular vectors r_i and s_i, respectively and $\sum$ is a diagonal matrix of positive real singular value σ_i such that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$. Given the decomposition in (2.13), at an equilibrium point other than the collapse one, (2.12) can be rewritten as:

$$\begin{bmatrix} \Delta\boldsymbol{\delta} \\ \Delta\mathbf{U} \end{bmatrix} = \sum_{i=1}^{n} \sigma_i^{-1} s_i r_i \begin{bmatrix} \Delta\mathbf{P}(\boldsymbol{\delta}, \mathbf{U}, \lambda) \\ \Delta\mathbf{Q}(\boldsymbol{\delta}, \mathbf{U}, \lambda) \end{bmatrix} \tag{2.14}$$

Notice from (2.14) that the minimum singular value is a relative measure of how close the system is to the voltage collapse point. When the system Jacobian is singular (i.e. $\sigma_n \sim 0$), the system state variables presents a large variation even subjected to a small load perturbation. This observation can be applied to the voltage stability analysis because the small incremental load changes may cause dramatic and sometimes uncontrollable voltage drops. Moreover, the left r_n and right s_n singular vectors contain important information related to the critical buses and the most sensitive direction for changes in power injection, respectively.

Because there is a strong relation between reactive power injections Q and bus voltages, the power flow model shown in (2.12) can be converted to consider only the Q-changes as follows:

$$\begin{aligned} \Delta\mathbf{Q}(\boldsymbol{\delta}, \mathbf{U}, \lambda) &= (\mathbf{J}_4 - \mathbf{J}_3 \mathbf{J}_1^{-1} \mathbf{J}_2) \Delta\mathbf{U} \\ &= \mathbf{J}_{QV} \Delta\mathbf{U} \end{aligned} \tag{2.15}$$

where $\mathbf{J}_{QV}$ is called the reduced Jacobian matrix. It is shown in the literature that the singular value of $\mathbf{J}_{QV}$ shows the better profile than the one of the complete **J** [18].

2.2.2.4. Optimization method

The maximum loading margin can be alternatively determined by an optimal power flow problem that seeks to maximize the loading factor while maintain the power flow equations, and other security limits. A theoretical study of equivalency between continuation and optimization methods was presented in [19]. An improved optimization model based on complementarity constraints previously introduced by [20] was further developed to accurately represent the generator voltage control. As initially proposed in [21], the voltage collapse can be obtained by solving the following optimization model:

$$\text{Maximize} \quad \lambda \tag{2.16}$$

$$\text{subject to } F(\mathbf{z}, \lambda) = 0. \tag{2.17}$$

The solution of this problem represents the last stable operating point before the system loses the voltage control ability. This problem can be solved by any optimization method. Most of the works presented so far such as [19, 20, 22-24], [25] rely on interior point methods. However, modern heuristic optimization methods were also applied in [26, 27]. In [27], the objective function is modified from maximizing λ to minimizing the mismatch of (2.6) in order to ensure the singularity condition and the non-zero left eigenvector requirement as follows:

$$\text{Maximize} \quad \Delta_{mis} = \sum \begin{bmatrix} D_{\mathbf{z}} F(\mathbf{z}, \lambda)^T \mathbf{w} \\ \|\mathbf{w}\|_\infty - 1 \end{bmatrix} \tag{2.18}$$

$$\text{subject to } F(\mathbf{z}, \lambda) = 0 \tag{2.19}$$

where Δ_{mis} is the mismatch of the singularity condition and the eigenvector requirement at the collapse point; **w** is the left eigenvector.

2.2.3 Performance indices

2.2.3.1. Simplified power flow model

Voltage Collapse Proximity Index

The voltage collapse proximity index *VCPI* is derived from the basic power flow equation. The derivation starts from finding the complex power injected to the bus *k* by the *N*-bus power system. Assuming that the complex voltages of the rest *N*-1 buses are known, two unknowns consisting of voltage magnitude U_k and voltage angle δ_k of the bus *k* are left to be solved as follows [28]:

$$f_1(U_k,\delta_k) = |U_k|^2 - \sum_{m=1,m\neq k}^{N} U'_m U_k \cos(\delta_k - \delta'_m) \tag{2.20}$$

$$f_2(U_k,\delta_k) = \sum_{m=1,m\neq k}^{N} U'_m U_k \sin\delta \tag{2.21}$$

where $\underline{U}'_m$ in the above two equations is defined by:

$$\underline{U}'_m = U'_m \angle \delta'_m = \frac{\underline{y}_{km}}{\sum_{i=1,i\neq k}^{N} \underline{y}_{ki}} \underline{U}_m \tag{2.22}$$

where $\underline{y}_{km}$ is the complex element of the admittance matrix at row *k* and column *m*.The collapse of voltage at bus *k* means that there is no solution to the partial derivative (Jacobian) matrix of (2.20) and (2.21). This means that the determinant of Jacobian matrix becomes zero at the voltage collapse point. The *VCPI* of bus *k* can be found from:

$$VCPI_k = \left| 1 - \frac{\sum_{m=1,m\neq k}^{N} \underline{U}'_m}{\underline{U}_k} \right| \quad ,\forall k \in \alpha_D \tag{2.23}$$

It is observed that *VCPI* in (2.23) is characterized by voltage phasor information of the participating buses and the network admittance matrix. The typical range of *VCPI* is between zero and one. When the voltage at bus *k* collapses, the index is one. The global index describing the stability margin of the entire system $VCPI_{sys}$ is given by the maximum value of *VCPI* of all load buses.

$$VCPI_{sys} = \max_{k \in \alpha_D} \{VCPI_k\} \tag{2.24}$$

L-Index

The line index *L* is a simple method to assess the voltage stability of a single transmission line initially proposed by [29] and further studied by [30]. The index uses information available in a normal power flow study. The derivation of *L* index starts from examining a simple two bus system with a generator and a load. The load bus voltage is expressed in terms of the generator bus voltage, the line admittance and the self-admittance connected to the two buses. The condition that the solution of load bus voltage exists was found that it can characterize the proximity of a line to the voltage stability limit. This concept is generalized to any multi-node system. The *L* index of each node *j* can be determined by:

$$L_j = \left| 1 - \sum_{i \in \alpha_G} C_{ij} \frac{\underline{U}_i}{\underline{U}_j} \right| \quad , \forall j \in \alpha_D \tag{2.25}$$

where α_G is the set of generator buses; α_D is the set of load buses; $\underline{U}_j$ is the complex voltage of bus *j* and C_{ij} is the element in row *i* column *j* of the gain matrix **C** determined by:

$$\mathbf{C} = -\left[\mathbf{B}''\right]^{-1} \mathbf{B}_{LG} \tag{2.26}$$

where $\mathbf{B}''$ is the imaginary part of the matrix $\underline{\mathbf{Y}}_{LL}$; $\mathbf{B}_{LG}$ is the imaginary part of the matrix $\underline{\mathbf{Y}}_{LG}$. The matrices $\underline{\mathbf{Y}}_{LL}$ and $\underline{\mathbf{Y}}_{LG}$ are sub-matrices of the admittance matrix. Decomposition techniques are generally applied to determine the matrix inversion in (2.26). The typical range of *L* index is between zero and one. When the voltage at bus *k* collapses, the index is one. The global index describing the stability margin of the entire system L_{sys} is given by the maximum value of L indices of all lines as:

$$L_{sys} = \max_{j \in \alpha_D} \{L_j\} \tag{2.27}$$

2.2.3.2. Local measurement model

Tracking Thevenin parameters

With the phasor measurement technology providing accurate time-synchronized measurements, it is possible now to track system dynamics in real-time. Phasor measurement units (PMUs) are placed in a region of a power system to ensure observability of bus voltages in that region. Voltage stability is one of the applications that PMUs enhance online monitoring. According to [31], PMU-based detection of voltage stability can be categorized into two broad categories depending upon the input data needed. In the first category, the methods [32, 33] assume few or no information exchanges among different locations. The methods of the second category [28, 34] require the observability of the entire region prone to voltage stability. In this thesis, we follow the first category.

Consider a load bus *k* having a load demand of $\underline{S}_k = P_k + jQ_k$ connected to the rest of power system as shown in Fig. 2.2(a). This part of the system can be represented by the Thevenin equivalent circuit as shown in Fig. 2.2(b). At an operating condition, two Thevenin parameters in a complex form consisting of voltage $\underline{U}_{TH}$ and impedance $\underline{Z}_{TH}$ must be determined by solving the voltage equation given as:

$$\underline{U}_{TH} = \underline{U}_k + \underline{Z}_{TH}\underline{I}_k \tag{2.28}$$

where $\underline{U}_k$ is the voltage at bus *k*; $\underline{I}_k$ is the current flowing to the load bus *k*. When the load bus *k* is considered as the reference (the voltage angle becomes zero), the load current $\underline{I}_k$ can be expressed by:

$$\underline{I}_k = \frac{P_k - jQ_k}{U_k} \tag{2.29}$$

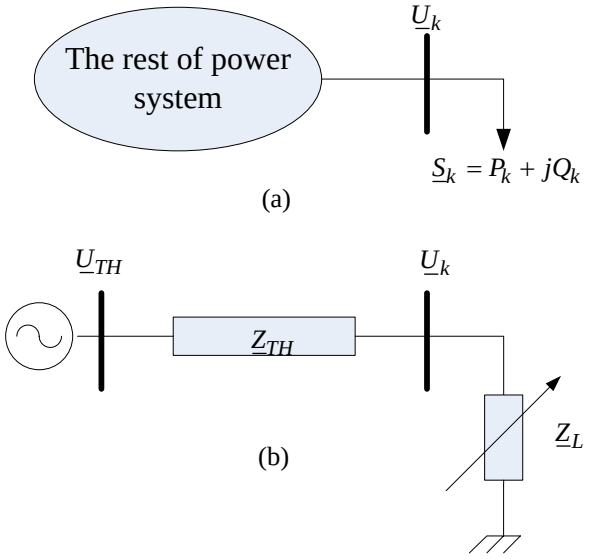


Fig. 2.2 Representation of a load bus *k* in a general power system

The voltage equation (2.28) can be transformed as follows:

$$U_{THr} + jU_{THm} = U_k + (R_{TH} + jX_{TH})\underline{I}_k \tag{2.30}$$

where U_{THr} and U_{THm} are real and imaginary part of $\underline{U}_{TH}$, respectively; R_{TH} and X_{TH} are resistive and reactive parts of $\underline{Z}_{TH}$, respectively . It is obvious from (2.30) that there are four unknowns to be solved. Therefore, at least two voltage and current measurements of bus *k* at time *t* are required to solve such an equation. An additional subscript *j* is introduced to the voltage and current to indicate the number of measurement as $\underline{U}_{k,j}$ and $\underline{I}_{k,j}$, respectively. Rewriting (2.30) in a matrix format with two set of measured quantities, (2.30) is converted to [32]:

$$\begin{bmatrix} 1 & 0 & -I_{kr,1} & I_{km,1} \\ 0 & 1 & -I_{km,1} & -I_{kr,1} \\ 1 & 0 & -I_{kr,2} & I_{km,2} \\ 0 & 1 & -I_{km,2} & -I_{kr,2} \end{bmatrix} \begin{bmatrix} U_{THr} \\ U_{THm} \\ R_{TH} \\ X_{TH} \end{bmatrix} = \begin{bmatrix} U_{kr,1} \\ U_{km,1} \\ U_{kr,2} \\ U_{km,2} \end{bmatrix} \tag{2.31}$$

where $I_{kr,j}$ and $I_{km,j}$ are real and imaginary part of $\underline{I}_{k,j}$, respectively; $U_{kr,j}$ and $U_{km,j}$ are real and imaginary part of $\underline{U}_{k,j}$,respectively.

Now, the four unknowns representing real and imaginary part of the two Thevenin parameters can be determined from (2.31). In practical power system operation, the measured data may contain some noise or error which could make Thevenin parameter estimates inaccurate. In such a case, more than two sets of measurement quantities collected from past loading conditions should be used and the least square curve fitting can be applied to solve for the Thevenin parameters as follow [32]:

$$\mathbf{A^T A x = A^T b} \tag{2.32}$$

where $\mathbf{A^T}$ denotes the transpose of **A**.

In the steady state time frame, the voltage collapse occurs when the system transfer limit is exceeded. Therefore, the estimated value of maximum loading condition can indicate the voltage stability margin. Observe the circuit in Fig. 2.2(b), the magnitude of load apparent power at bus *k* can be expressed as [35, 36]:

$$S_k = \frac{U_{TH}^2 Z_L}{Z_{TH}^2 + Z_L^2 + 2Z_{TH}Z_L\cos(\beta - \varphi)} \tag{2.33}$$

where $\underline{Z}_{TH} = Z_{TH}\angle\beta$ is the Thevenin impedance and $\underline{Z}_L = Z_L\angle\varphi$ is the load impedance. For a given load power factor, the maximum apparent power that can be transferred to the load occurs when the condition is satisfied.

$$\frac{\partial S_k}{\partial Z_L} = 0 \tag{2.34}$$

The solution of (2.34) gives the condition of maximum power transfer when:

$$Z_L = Z_{TH} \Rightarrow Y_L = Y_{TH} \tag{2.35}$$

where Y_L and Y_{TH} are the load impedance and Thevenin equivalent impedance, respectively. The maximum apparent power that can be drawn from the system to the load bus *k* can be determined by substituting (2.35) to (2.33) as:

$$S_{k,\max} = \frac{U_{TH}^2}{2Z_{TH}(1+2\cos(\beta-\varphi))} \tag{2.36}$$

Note that the Thevenin parameters vary with the system operating condition. Several indicators can be derived based on the Thevenin parameters to assess the distance to voltage collapse of a load bus. Based on (2.35), the voltage stability margin in terms of admittance denoted as VSM_y can be defined as [35]:

$$VSM_y = \frac{Y_{TH} - Y_L}{Y_{TH}} \tag{2.37}$$

The value of VSM_y falls between zero and one. When the voltage collapse occurs, VSM_y becomes zero. The VSM_y indicator is not so intuitive because it is not expressed in a directly measurable quantity. Therefore, it is more desirable to express the indicator in terms of load apparent power.

Power Transfer Stability Index

The power transfer stability index *PTSI* represents the ratio between the load apparent power and the maximum loadability. At the voltage collapse point, such a ratio is equal to one, that is [37]:

$$PTSI_k = \frac{S_k}{S_{k,\max}} \tag{2.38}$$

Substituting (2.36) into (2.38), such a margin can be expressed as:

$$PTSI_k = \frac{2S_L Z_{TH}(1+2\cos(\beta-\phi))}{U_{TH}^2} \quad , \forall k \in \alpha_D \tag{2.39}$$

where α_D is the set of load buses. The typical range of *PTSI* is between zero and one. When the system reaches the voltage collapse point, the *PTSI* value reaches one. The global index

describing the stability margin of the entire system $PTSI_{sys}$ is given by the maximum value of *PTSI* of all load buses as:

$$PTSI_{sys} = \max_{k \in \alpha_D} \{PTSI_k\} \tag{2.40}$$

Power based Voltage Stability Margin

Because the voltage collapse occurs when the apparent power drawn from the system is equal to the maximum loading capability, the power-based voltage stability margin *PVSM* may be defined as [35]:

$$PVSM_k = \frac{S_{k,\max} - S_k}{S_{k,\max}} \tag{2.41}$$

Substituting (2.33) and (2.36) into (2.41), the *PVSM* can be expressed as:

$$PVSM_k = \frac{(Z_L - Z_{TH})^2}{Z_{TH}^2 + Z_L + 2Z_{TH}Z_L \cos(\beta - \phi)} \quad , \forall k \in \alpha_D \tag{2.42}$$

The typical range of *PVSM* is between zero and one. In contrast to *PTSI*, the value of *PVSM* reaches zero when the voltage collapse occurs. The global index describing the stability margin of the entire system $PVSM_{sys}$ is given by the minimum value of *PVSM* of all load buses

$$PVSM_{sys} = \min_{k \in \alpha_D} \{PVSM_k\} \tag{2.43}$$

2.2.3.3. Simulation results

A number of indices presented in this chapter are conducted to assess voltage stability of the IEEE-14 bus test system the single line diagram of which is depicted in Fig. 2.3. The system consists of five synchronous generators connected to buses 1,2,3,6 and 8. Among five generators, generator 1 is chosen as the slack bus and generators 3, 6 and 8 are synchronous condensers supplying only reactive power. Nine load buses are not directly attached to any

generator while two loads at buses 2 and 3 are connected to a generator. The base load condition of this system is 362.62 MW and 113.96 MVar.

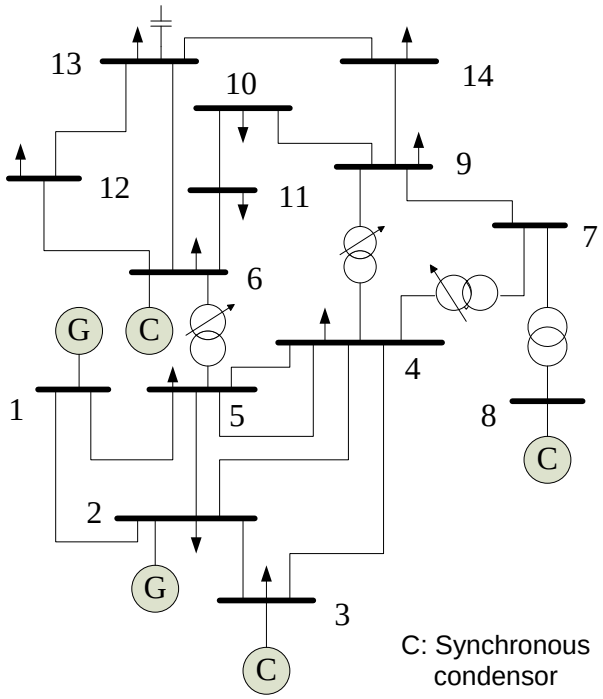


Fig. 2.3 Single line diagram of the IEEE 14-bus system

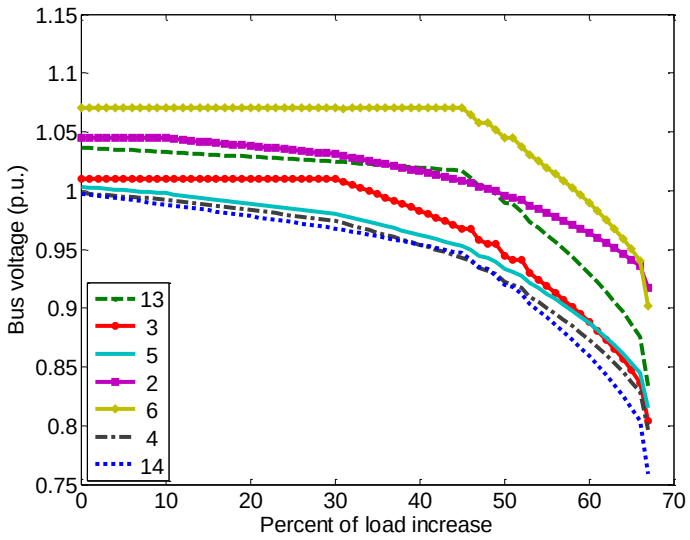


Fig. 2.4 Bus voltages at different loading conditions

The loading scenario considered in this example is uniform and simultaneous load increases at every load bus according to (2.4) for both active and reactive power demands. For all load buses, the load increment coefficients $k_{P,i}$ and $k_{Q,i}$ are fixed to 0.01 (1%). The rate of change in generation $k_{G,i}$ as shown in (2.5) is equally set to 0.01 for all generators. The loading parameter λ is running from zero (the base-load condition) until the maximum loading condition $\lambda^{\max}$ is met where the power flow solution does not converge. In this simulation, the collapse occurs at 168% of the base loading level.

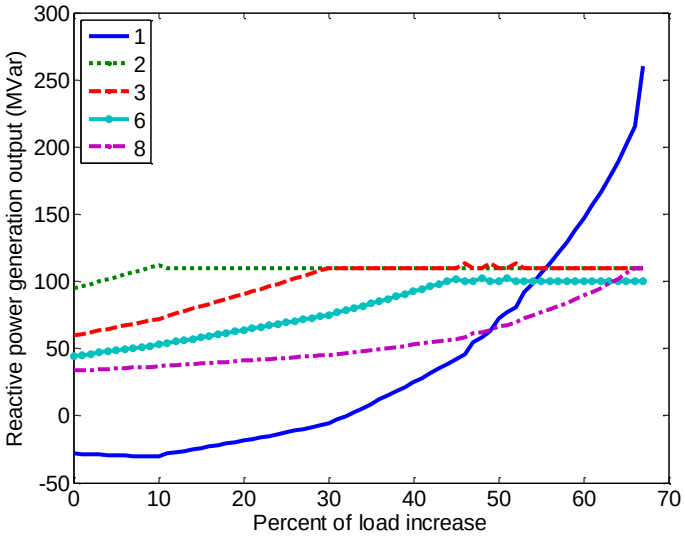


Fig. 2.5 Reactive power generation outputs at different loading conditions

During the course of load increase, a power flow routine is conducted to determine the corresponding power system state. In this situation, the voltage profiles at selected load buses are shown in Fig. 2.4. Observe that the voltage profiles of buses 2, 3 and 6 with generating units attached remain constant up to a certain loading level before starting to decrease. This characteristic is due to the loss of voltage control capability of a generator. The relationship between the reactive power generations and load increases are depicted in Fig. 2.5. When the reactive power output of a generator reaches the limit, reactive power generation of that generator has to be fixed at its maximum capacity and the bus type is converted from a generator (PV) bus to a load (PQ) bus. It is assumed in this example that the slack bus (at bus 1) is represented by a very large generator with unlimited reactive power generation.

Observe from the voltage profile shown in Fig. 2.4 that bus voltages are slightly changed in response to the load variation and abruptly diminished near to the collapse point. This is an undesirable characteristic of an indicator for voltage stability. A good indicator should be easily monitored as system parameters change. The indicator profile should also have a predictable and smooth shape. Minimum singular value (MSV) of the power flow Jacobian matrix and its sub-matrices are among various indicators. The relationship between the variation of MSV of the complete Jacobian FJ different sub-matrices namely $\mathbf{J}_1$, $\mathbf{J}_4$ and $\mathbf{J}_{QV}$ (see (2.15)) and the loading scenario can be shown in Fig. 2.6.

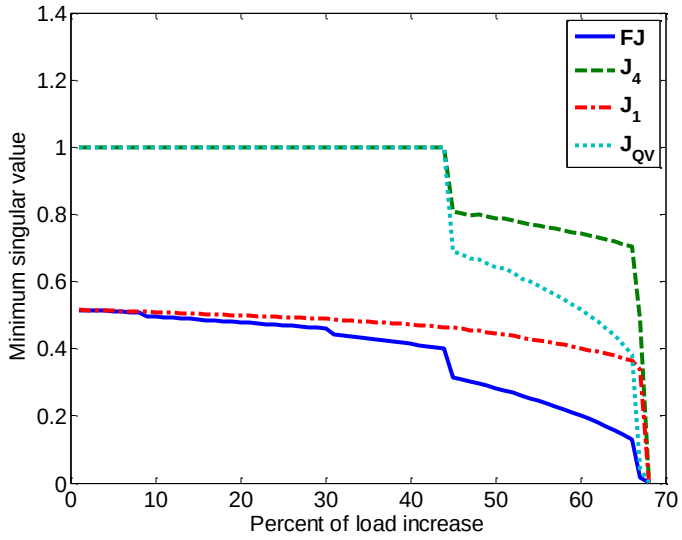


Fig. 2.6 Minimum singular value at various loading levels

It is quite obviously seen from Fig. 2.6 that MSV of $\mathbf{J}_1$ is quite insensitive to the load change pattern. Therefore, it is not a good indicator. This is not surprising because the primarily cause of voltage instability is the lack of reactive power. $\mathbf{J}_4$ and $\mathbf{J}_{QV}$ provide sensitivity information between reactive power injection and voltage at buses (Q-V sensitivity) by assuming the weak coupling between reactive power and voltage angle where considering MSV of $\mathbf{J}_4$ alone would completely ignore such a coupling. In some practical large power system, MSV of the reduced $\mathbf{J}_{QV}$ may be of more interest because it could reduce computing burden instead of handling the complete **FJ**. It should be observed that there are some sudden declines of MSV of **FJ** because of the change of bus types. In mathematical

interpretations, this bus type conversion alters the list of state variables and the dimension of power flow Jacobian matrix.

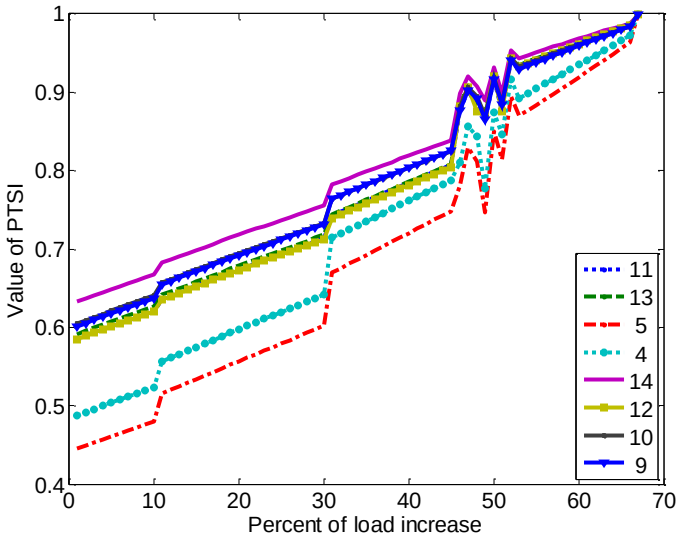


Fig. 2.7 PTSI of load buses at various system loading levels

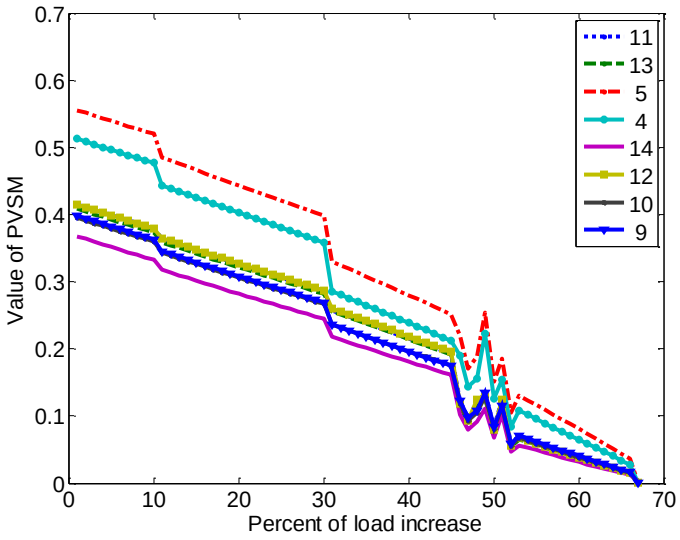


Fig. 2.8 PVSM of load buses at various system loading levels

Notice that MSV is a measure for degradation of power flow equations. Therefore, this value represents the stability of the entire power system. On the other hand, some indicators such as L-index, PTSI and PVSM provide stability information at local buses. Based on the

technique for tracking Thevenin parameters discussed in section 2.2.3.2, a few indicators can be derived. Among them, the variations of PTSI and PVSM corresponding to the loading scenario discussed earlier are shown in Fig. 2.7 and Fig. 2.8, respectively. An important merit of these indicators is that they rely only on the information available at local load buses. Network topology is not required for this method. It is clearly depicted that PTSI and PVSM of all load buses reach the critical value of one and zero, respectively, at the maximum loading point of 68%. The most critical bus in this operating condition is bus 14 where PTSI is the greatest and PVSM is the lowest compared to the ones at the other load buses. The conversion of bus status is also reflected by the sudden changes in the PTSI and PVSM profiles. As discussed earlier in section 2.2.3.2 that at least two measurements of bus voltage magnitude and bus voltage angle are required to solve (2.31), five different measurements are used in this example. A small random perturbation is introduced to each system loading condition so that (2.31) is numerically solvable. It can be further noticed that at high loading conditions after the bus type change of generator 6 (see Fig. 2.5), PTSI and PVSM mildly fluctuate due to the greater difficulty of estimating Thevenin parameters near to the collapse point. However, the two indicators are able to correctly indicate the proximity to voltage collapse.

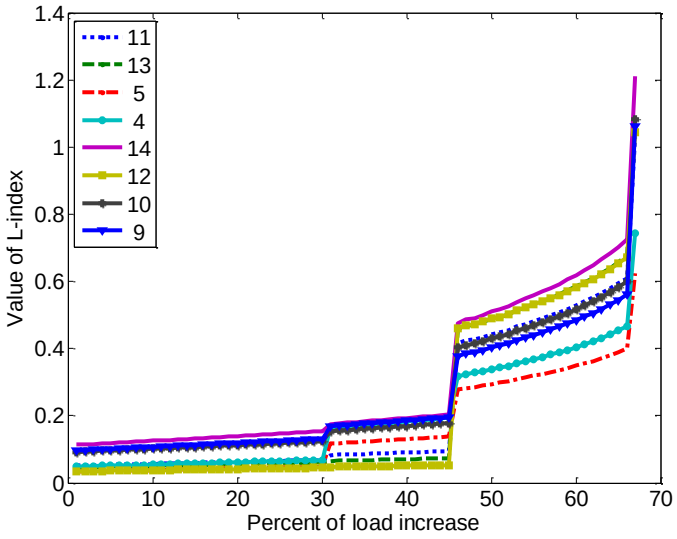


Fig. 2.9 L-index of load buses at various system loading levels

The last indicator presented in this example is the famous L-index widely used by many researchers. This indicator is a representative of indicators which are derived from the simplified power flow equations based on certain assumptions. Therefore, the information required to compute these indicators are generally similar to those for a power flow calculation. L-index needs the complete admittance matrix bus voltages in a complex form. The critical value of L-index is one as discussed earlier. Fig. 2.9 shows the response of L-index to the load variation scenario. Nonlinearity of the indicator profile also exists in this case. Bus 14 has the greatest value of L-index at nearly every loading level. This observation conforms to the identification of the most critical bus made by the other indicators.

2.3 Preventing voltage collapse

Based on the discussion regarding design and operating measures in preventing voltage collapse in a power system given in [5], this section summarizes some aspects relating to the studies developed in this thesis. Recent developments in each method are further elaborated.

2.3.1 Reactive power and voltage control

The power system should be operated with an adequate voltage stability margin. If the system operates with the margin less than required in the operating guidelines, the appropriate scheduling of reactive power sources and voltage profiles in the system is the first tool that the operator should use to enhance the voltage stability. The problem is essentially defined by an optimization problem whose objective can be either to minimize the total active power transmission losses or to minimize the voltage deviation at load buses. The required stability margin is generally constrained to ensure the stable system operation. Control devices that can be used to achieve this task comprise reactive compensating devices, transformer tap changers and reactive power output or terminal voltage of a generator.

2.3.1.1. Reactive compensation devices

Reactive power compensating devices can improve voltage stability of a power system. For example, shunt capacitors are by far the most inexpensive means to providing reactive power and voltage support in load premises. With the appropriate size and location of the devices, they can free up the reactive spinning reserve in generators and prevent voltage collapse in

many situations. However, these devices have a number of limitations, for instance when in a highly-compensated system the voltage control may become poor. Series capacitors are self-regulating devices providing effects similar to shortening long transmission lines.

The advent of flexible AC transmission systems (FACTS) technology gives a new possibility of controlling power flow without the need to reschedule generation and network topology. There are a number of devices in this category used in power systems for different purposes. Among these, unified power flow controller (UPFC) is the most advanced FACTS controller providing significant flexibility in injecting a controlled series and shunt compensation. Given various merits provided by these devices in improving the system performance including voltage stability margin [38], a number of technical issues such as coordination of FACTS with the other devices have to be taken into consideration. Beside this, the devices are relatively expensive. Therefore, studies should be carried out to assess the investment returns and benefits.

2.3.1.2. Control of transformer tap changers

Transformer tap changers can be either locally or centrally controlled to mitigate the risk of voltage collapse. A simple method is to block tap changing when the voltage sags and unblock when it recovers. This is a common control scheme used in several utilities.

2.3.2 Under-voltage load shedding

In extreme and unplanned situations, it may be necessary to use under-voltage load shedding schemes analogous to under-frequency load shedding. Load shedding is a low-cost means of preventing wide spread system collapse and sometimes is effective method to drive operation of an unstable system back into a stable operating zone. This control measure is generally applied as the last resort when the system conditions and contingencies leading to voltage instability are of low probability but would caused widespread consequences.

2.4 Summary

This chapter presents fundamental concepts of voltage stability, analysis techniques, indices and control measures for preventing voltage collapses. Various indices proposed in literature are categorized into two main groups: the simplified model and the local measurement model.

The simulation results showed that the indices in these two groups have nearly identical ability to assess voltage stability of the power system. The major difference is only the information required to determine the index. Preventive and corrective control measures that can safeguard the power system from the collapse are also presented.

Chapter 3
Computational intelligence tools

3.1 Introduction

This chapter discusses implementations of the computational intelligence (CI) techniques applied in this thesis for voltage stability assessment and control problems. The chosen CI methods consists of two main categories namely neural networks (NNs) and evolutionary algorithms (EAs). Two types of NNs namely multi-layered feedforward NNs (FFNNs) and learning vector quantization (LVQ) are selected for different regression tasks and classification tasks, respectively. LVQ, sometimes refers to a Kohonen network, is the supervised counterpart of the self-organizing map (SOM) for classification problems developed by Kohonen [39].

In this thesis, FFNNs are principally devoted to estimate voltage stability margins (VSMs) of power systems without the need to conduct a complete simulation. LVQ as applied in Chapter 6 is responsible for indicating voltage stability of the current operating condition . An unstable operating condition will be further analyzed for appropriate control measures in the framework of optimization problems.

With the recent development in evolutionary computation, the search performance of EAs has been enhanced. They provide a very robust and powerful solution to highly complex optimization problems. This thesis investigates applications of a few well-known EAs to handle power system operation problems with major emphasis given on the control measures to prevent voltage collapse. Those EAs applied in this thesis consist of genetic algorithm (GA), ant colony optimization (ACO) and differential evolution (DE). In DE, special interests were given to a recent variant called JADE [40]. Parameters of JADE are adapted based on the knowledge gained during the course of optimization. However, a control parameter namely population size still needs to be tuned and specified by the user. In this thesis, the original JADE was improved by allowing the population size to automatically change in the manner similar to the other control parameters. The modified JADE method is named JADE-vPS standing for JADE with variable population size.

This rest of this chapter is organized as follows. Dimensionality reduction methods as the pre-processing are discussed in Section 3.2. Essential concepts of NNs are summarized in Section 3.3. Evolutionary algorithms are presented in details in Section 3.4. This section also gives brief introduction to backgrounds and different methods of constraint handling. Finally, Section 3.5 concludes the chapter.

3.2 Dimensionality reduction

The number of information describing an operating state directly grows in proportion to the sheering size of a power system. For an interconnected system, the complete state information is generally too large to be effectively handled by any CI method. The problem of curse of dimensionality is very well-known in this respect. It is also advisable that the number of input information should be reduced to a number reasonably manageable by CI approaches to enhance generalization capability and speeding up the learning process. In statistics, the term dimensionality reduction refers to the process of reducing the number of variables under consideration and can be classified into "feature selection" and "feature extraction". In CI terminology, "feature" is a quantity used to characterize input information which is in our case an operating state. The idea of dimensionality reduction techniques can be depicted as in Fig. 3.1 (adapted from [41]).

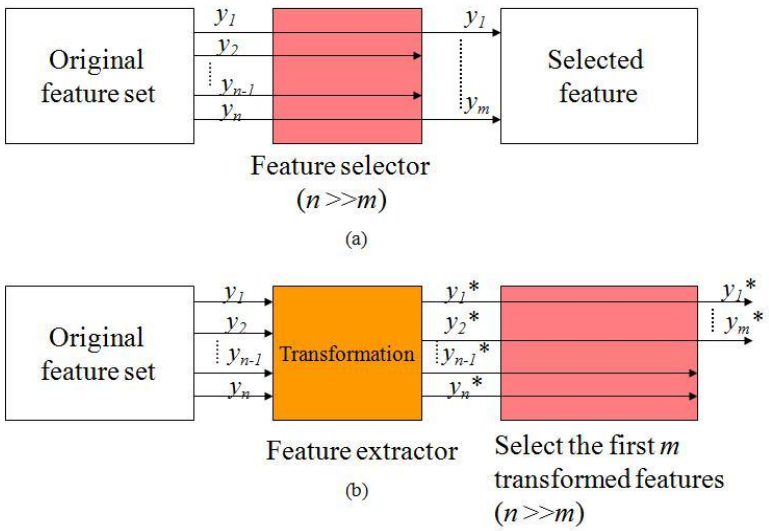


Fig. 3.1 A conceptual illustration of dimensionality reduction (a) feature selection (b) feature extraction

In feature selection, a part of the original feature set is selected to describe the input information whereby the remainder is discarded. Fig. 3.1(a) shows that by some means set up in the feature selector, m variables $[y_1,y_2,\ldots,y_m]$ are chosen to equivalently represent the original data set with n variables. On the other hand, the original feature set is transformed into a new space $\left[y_1^*, y_2^*,..., y_n^*\right]$ in feature extraction. Then, the first m features $\left[y_1^*, y_2^*,..., y_m^*\right]$ that preserve much of the relevant information are selected.

3.2.1 Feature selection

Feature selection is the technique to select a subset of the inputs and to discard the rest. From the theoretical viewpoint, the optimal subset of features can be regarded as a supervised learning problem requiring exhaustive searches of all possible subsets of features [42]. This method is impractical when the number of features is large. Feature selection is particularly helpful if the input features are highly correlated. This means that the same information is repeated in several variables.

Two categories can be classified among feature selection algorithms: feature ranking and subset selection. In feature ranking, features are ranked by a pre-defined score. The features that do not achieve an adequate score are eliminated. Canonical correlation and mutual information are two common scores used to describe the relationship between two features or two sets of features. Linear discriminant analysis, such as Fisher's discriminant [43] can also be used to determine separability of two feature classes. Then, a sequential ranking method is carried out based on the computed Fisher's values to construct the set of suitable features. In subset selection methods, a subset of features is evaluated for suitability. Clustering can be considered as a member of this category.

The k-mean algorithm is a well-known and powerful clustering method applicable to most situations. This is measured by the Euclidean distance between any two vectors. Then, each point is assigned to the cluster whose center is nearest. Given a set of observations $\left(\mathbf{x}_1, \mathbf{x}_2,..., \mathbf{x}_n\right)$ where each observation is in a D-dimensional space. The k-mean clustering assigns n observed vectors into $k<N$ clusters denoted by $\left\{\mathbf{S}_1, \mathbf{S}_2,..., \mathbf{S}_k\right\}$ so that the within-cluster sum of squares is minimized according to:

$$\min_{\mathbf{S}} \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in \mathbf{S}_i} \left\| \mathbf{x}_j - \mathbf{c}_i \right\|^2 \tag{3.1}$$

where $\mathbf{c}_i$ is the center of the cluster $\mathbf{S}_i$ defined by the mean of points in $\mathbf{S}_i$ given by

$$\mathbf{c}_i = \frac{1}{m} \sum_{\substack{j=1 \\ \mathbf{x}_j \in S_i}}^{m} \mathbf{x}_j \tag{3.2}$$

where m is the number of observations in the cluster $\mathbf{S}_i$. Given an initial mean (center) of k clusters, the iterative procedures are repeated and the mean of each iteration is computed. The algorithm is deemed converged if the means have no changes.

3.2.2 Feature extraction

Principal component analysis (PCA) also known as Karhunen-Loéve transformation is one of the widely used techniques for reducing the number of input features. Beside the dimensionality reduction, the primary goal of PCA is to preserve as much of the relevant information as possible. The PCA procedures entirely rely on the input data itself without considering any target data. Therefore, it can be regarded as an unsupervised learning method. Given an input vector $\mathbf{x}$ in a D dimensional space $\left(x_1, x_2, ..., x_D\right)$, PCA maps such a vector to a new vector $\mathbf{z}$ in a M dimensional space $\left(z_1, z_2, ..., z_M\right)$,where $M<D$. Without any loss of generality, an input vector $\mathbf{x}$ can be expressed by a linear combination of D orthonormal vector $\mathbf{u}_i$ according to:

$$\mathbf{x} = \sum_{i=1}^{D} z_i \mathbf{u}_i \tag{3.3}$$

Suppose that a subset of $M<D$ basis vectors $\mathbf{u}_i$ are retained for z_i and the rest coefficients are replaced by constants b_i, an input vector $\mathbf{x}$ is approximated by:

$$\tilde{\mathbf{x}} = \sum_{i=1}^{M} z_i \mathbf{u}_i + \sum_{i=M+1}^{D} b_i \mathbf{u}_i \tag{3.4}$$

Consider the entire dataset with N patterns, the approximation error for an input vector n=1,…,N introduced by the dimensionality reduction is governed by:

$$\mathbf{x}^n - \tilde{\mathbf{x}}^n = \sum_{i=M+1}^{D} (z_i^n - b_i)\mathbf{u}_i \tag{3.5}$$

For the best approximation, the total sum of the square errors over the whole dataset should be minimized. Therefore, we minimize the following error function.

$$E = \frac{1}{2}\sum_{n=1}^{N} \left\| \mathbf{x}^n - \tilde{\mathbf{x}}^n \right\|^2 = \frac{1}{2}\sum_{n=1}^{N}\sum_{i=M+1}^{D} (z_i^n - b_i)^2 \tag{3.6}$$

After following the optimization procedure, it can be shown [42] that the minimum occurs when the basis vectors satisfy

$$\mathbf{C}\mathbf{u}_i = \lambda_i \mathbf{u}_i \tag{3.7}$$

where **C** is the covariance matrix of the vector set $\{\mathbf{x}^n\}$. It should be observed that (3.7) is the eigen-characteristic of the covariance matrix. Therefore, a PCA algorithm starts by computing the covariance matrix and its eigenvectors and eigenvalues. Then, the eigenvectors corresponding to the M largest eigenvalues are used to project the input vectors $\{\mathbf{x}^n\}$ onto a new M dimensional space.

The algorithm discussed here is based on a linear dimensionality reduction with the sum-of-square criterion. It is possible to consider other covariance measures and population entropy. Detailed discussions can be found from [44].

3.3 Neural networks

In machine learning, artificial neural network is the electrical analogue of the biological nervous system. For short, the term "neural network" (NN) will be used throughout the thesis. This section gives a brief background of NN and outlines few useful NN configurations. The concepts of supervised learning are also discussed.

As suggested by its name, a NN consists of various connected neurons. Input signals are transmitted to selected neurons through weighted factors. Fig. 3.2 shows a typical model for a

single artificial neuron where there are n inputs x_1 to x_n activating this neuron. At each neuron j, the intermediate output u is computed by the linear combination (dot product) between the input values and the connection weights. Furthermore, a neuron can receive an additional input called bias b which does not come from any other neuron. This bias helps improve the network generalization in some cases in which all inputs are zero and the target is non-zero. Because the domain of u is mathematically infinite, an activation function is required to map such an infinite range to a pre-specified range. Five activation functions are commonly employed in NNs consisting of [45]: (a) the linear function, (b) the step function, (c) the ramp function, (d) the sigmoid function and (e) the Gaussian function. Each activation function is illustrated in Fig. 3.3. In summary, the final output of a neuron is computed by:

$$y_j = f(u_j + b_j) \tag{3.8}$$

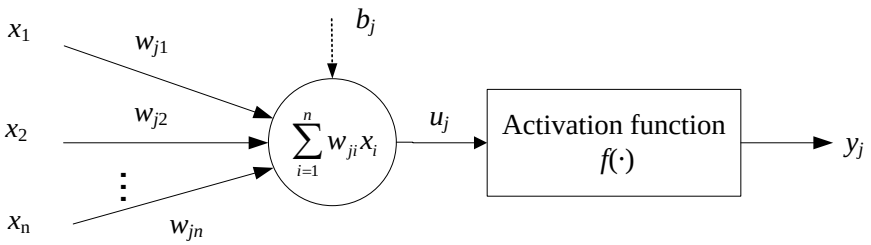


Fig. 3.2 A typical single neuron model

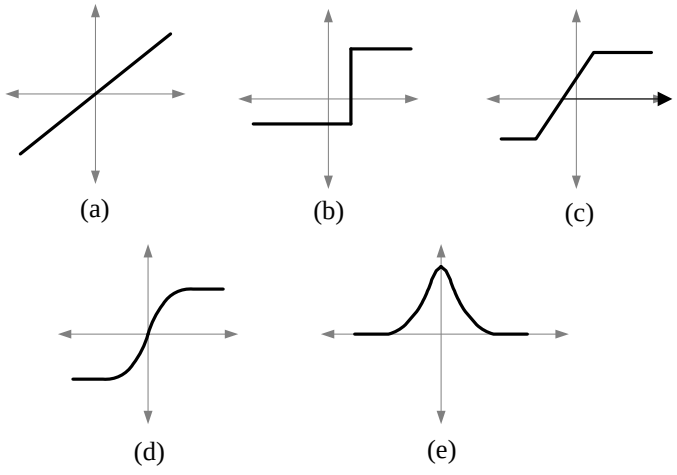
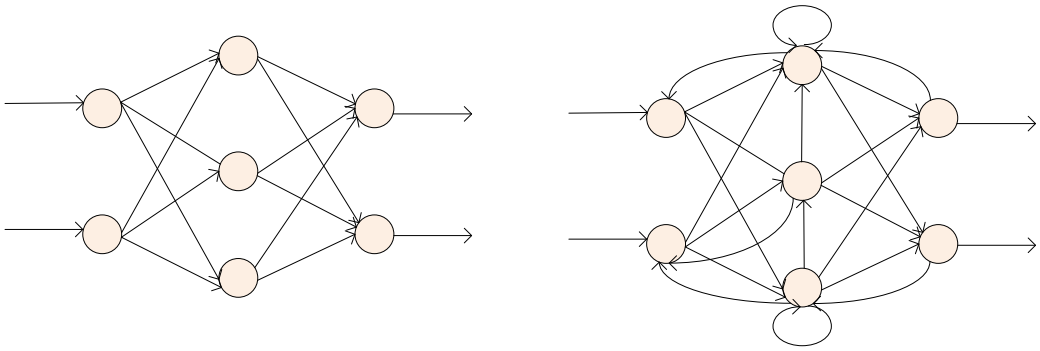


Fig. 3.3 Five most common activation functions

Neurons in a NN can be connected in different topological configurations. The best network configuration practically depends on the problem to be solved. The multilayer feed-forward network is commonly used in nonlinear mapping problems. A feed-forward network can be configured with a number of layers in which there are a number of neurons in each layer. Input signals flow in a forward manner and there is no feedback from one neuron to the other as shown in Fig. 3.4(a). This type of networks is generally adequate for most of regression and classification problems. However, one of its disadvantages is the ability to handle temporal information. This problem can be overcome using a recurrent network in which there exist feedbacks among different neurons either in the same layer of in the previous one as depicted in Fig. 3.4(b). This type of network is well suited for predicting time series such as wind power forecasting in [46] or equivalencing of dynamic systems such as load modeling in voltage stability assessment in [47].



(a) (b)

Fig. 3.4 Configurations of neural networks
(a) a feed-forward network (b) a recurrent network

Once the appropriate network configuration is defined, a training session is required so that the connection weight values can be adapted. This is the indispensable process for every NN in order to perform the given task correctly. Adaptation of weights can be accomplished by four different methods namely supervised learning, unsupervised learning, competitive learning and reinforcement learning. The training method used in this thesis is the supervised learning in which relevant concepts are illustrated in Fig. 3.5 [48].

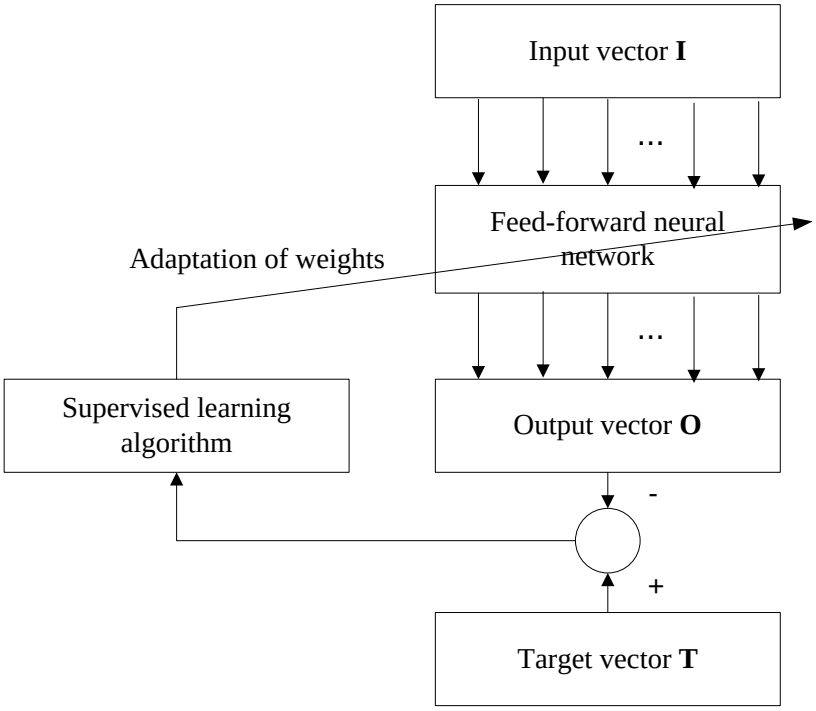


Fig. 3.5 A supervised learning scheme

Given an input vector **I**, the corresponding output vector **O** is computed based on the input **I** and weights. At the initial stage, weights are randomly initialized. Compared to the known target vector **T**, the error vector **E** = **T**-**O** is computed. Then, the supervised learning algorithm will adjust the network weights. The back-propagation algorithm is the popular method for network parameter adaptation. In NN training, an epoch is the presentation of the entire training set to the NN. The training session can be terminated when the number of epochs reaches the pre-specified maximum number or when the estimation error in an independent data set called the validation set starts to increases. The later intends to prevent the network from the over-fitting problem.

3.4 Evolutionary algorithms

3.4.1 Overview

In computational intelligence, an evolutionary algorithm (EA) is a subset of evolutionary computation (EC) which is a population-based meta-heuristic optimization method [49]. The

search mechanism of an EA is inspired by natural processes such as biological evolution, reproduction, mutation, crossover, or selection. Candidate solutions of an optimization problem are represented by individuals in a population. The quality of a solution is estimated by the fitness function. The population is repetitively evolved through evolutionary operators. These operators can be varied in different EAs.

EC optimization methods differ from traditional optimization methods in a number of aspects. One of the major differences is the requirement for derivatives information. In many real-world optimization problems, it is hard or even not possible to calculate derivatives. The other merit of the EC paradigm is that this approach reduces the chance of getting stuck to local optimum by allowing the population to explore the search space simultaneously through effective operators over several trials [50].

Regardless of the algorithm implemented in the EC paradigm, EAs often follow a similar procedure as listed below [45].

1. Initialize the population.
2. Calculate the fitness of each individual in the population.
3. Reproduce selected individual to create a new population.
4. Perform evolutionary operations, such as crossover and mutation on the population and select parent individuals for the next generation
5. Loop to step 2 until the termination criterion is met.

EAs have been successfully implemented in diverse engineering areas including electric power systems. Three EAs applied in this thesis will be presented in section 3.4.2. The penalized fitness function adopted to handle constraints is discussed in section 3.4.3.1.

3.4.2 Algorithms

This section gives essential and compact summary of the EAs applied in this thesis in the first three subsections. Finally, the constraint handling method based on the self-adaptive penalty scheme is discussed.

3.4.2.1. Genetic algorithm

Genetic algorithm (GA) is a search method in the EC paradigm reflecting the primitive process of biological evolution introduced by Holland [51]. A population of GA represents

the set of candidate solutions represented by different individuals. An individual is called a chromosome due to the analogy to genetic evolution of organisms.

The early version of GA treats real numbers in the binary form. Encoding and decoding are required to convert continuous variables to binary variables and vice versa. Each element of an individual can be encoded in any specified number of bits. A unique discrete value is then assigned to each sub-range. The difference between the actual function value and the quantization level is known as the quantization error. The number of bits directly influences the quantization error and ultimately the performance of GA [52].

In continuous search spaces, the binary encoding scheme may become sluggish if the number of variables is large and a high precision is required. Moreover, in that condition the binary GA also suffers from the memory storage problem because the size of chromosome is getting much larger. The points discussed earlier motivate the development of continuous GAs where a variable is represented by a single floating-point number instead of various binary integers such as [53, 54].

3.4.2.2. Ant colony optimization

Ant colony optimization (ACO) is the algorithm inspired by the foraging behavior of real ants and was initially proposed to solve combinatorial optimization problems [55]. Recently, there have been great interests in extending ACO to handle problems in the real space [56]. Among these techniques, the one proposed by Socha and Dorigo in [57] namely ACO_R, demonstrates better performance than other ant-inspired algorithms in dealing with various test problems in unconstrained optimization. Therefore, we adopted this ACO_R with some modification for constrained optimization as the optimization engine of various tasks developed in this thesis.

Conventional ACO initially developed to deal with combinatorial optimization problems is not appropriate to handle continuous variables. This is because it maps the entire search space of every dimension into a discretized graph namely the pheromone trail with a definite number of potential solutions. An ant selects the new direction to move from one step to the other using a selection method such as the roulette wheel or tournament selection based on a probabilistic criteria defined as a function of pheromone over each trail.

On the other hand, ACO_R defines the entire search domain of each dimension by a continuous probability distribution function (PDF). Multiple promising search regions are modeled by the Gaussian kernel PDF defined as a weighted sum of k single PDFs as follows:

$$G^i(x) = \sum_{l=1}^{k} \omega_l g_l^i(x) = \sum_{l=1}^{k} \omega_l \frac{1}{\sigma_l^i \sqrt{2\pi}} e^{-(x-\mu_l^i)^2/2(\sigma_l^i)^2} \tag{3.9}$$

where k is the number of single Gaussian PDF at the construction step i; ω, μ, and σ are vectors of size k defining the weights, means and standard deviations associated with every individual Gaussian PDF at the construction step i.

For each ant, a new variable value can be developed at each construction step by a random sampling technique of a given PDF based on mean μ and standard deviation σ. Since the global optimization in a continuous domain involves an indefinite number of candidate solutions, the pheromone trail concept of conventional ACO is no longer applicable. Therefore, ACO_R stores the knowledge gained from previous searches in the tabular format namely the archive $\mathbf{X}_T$.

	Candidate solution						Fitness	Probability	Feasibility
$\mathbf{x}_1$	x_1^1	x_1^2	. . .	x_1^i	. . .	x_1^D	$f(\mathbf{x}_1)$	p_1	1
$\mathbf{x}_2$	x_2^1	x_2^2	. . .	x_2^i	. . .	x_2^D	$f(\mathbf{x}_2)$	p_2	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$\mathbf{x}_j$	x_j^1	x_j^2	. . .	x_j^i	. . .	x_j^D	$f(\mathbf{x}_j)$	p_j	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$\mathbf{x}_n$	x_n^1	x_n^2	. . .	x_n^i	. . .	x_n^D	$f(\mathbf{x}_n)$	p_n	0

Fig. 3.6 Data structure of the solution archive

Fig. 3.6 shows the data structure of the solution archive $\mathbf{X}_T$ designed to handle constrained optimization problems. The archive stores the set of good solutions that the algorithm has discovered so far. The first part of $\mathbf{X}_T$ contains the set of candidate solutions $\mathbf{x}_j$, j=1,2,…,n where $\boldsymbol{x}_j \in \Re^D$; D is the problem dimension. The next column of $\mathbf{X}_T$ stores the corresponding fitness value of j[th] candidate solution $f(\boldsymbol{x}_j)$. The probability of selecting the j[th] solution as a

mean p_i is recorded in the next column. The last column holds the binary indicating feasibility status of the the ith solution (1 if feasible and 0 if infeasible).

As mentioned earlier that ACO_R applies the Gaussian kernel to generate a new solution, it is however not simple to find the inverse cumulative distribution function of an arbitrary distribution function. Therefore, ACO_R applies an alternative sampling technique in order to increase the implementation flexibility in two steps.

In the first step, a single component of the kernel is probabilistically selected for each ant. The weight ω_j of the solution j is the Gaussian PDF value with mean of 1 and standard deviation of qn. It is computed according to:

$$\omega_j = \frac{1}{qn\sqrt{2\pi}} e^{-(j-1)^2/2q^2n^2} \tag{3.10}$$

where q is a parameter of the algorithm and n is the size of solution archive. When q is small, the solutions with lower ranks in the archive have very strong influences in guiding new search directions whereby a larger q allows the wider search diversification over the entire space. For each archive solution of rank j in $\mathbf{X}_T$, the corresponding probability is calculated by:

$$p_j = \frac{\omega_j}{\sum_{r=1}^{n} \omega_r} \; ; \forall j = 1, 2, ..., n \text{ and } p_j \in \mathbf{p} \tag{3.11}$$

where $\mathbf{p}$ is the vector containing probability of selection of all archived solutions.

Then, to generate an ant k of the descent ant population, the Roulette wheel selection method is applied to randomly select which candidate solution of $\mathbf{X}_T$ should be set as the vector of mean values expressed by:

$$\boldsymbol{\mu}_k = \mathbf{x}_j \; ; \forall k = 1, 2, ..., m \text{ and } j = \text{Roulette}(\mathbf{p}) \tag{3.12}$$

where n is the problem dimension; m is the size of ant population; Roulette($\mathbf{p}$) is the Roulette selection function with $\mathbf{p}$ as the input and returning the selected rank. For each ant k, the standard deviation σ for every construction step i is calculated from the average distance from the chosen solution $\mathbf{s}_j$ to the other solutions in $\mathbf{X}_T$ according to:

$$\sigma_k^i = \xi \sum_{e=1}^{n} \frac{\left| x_e^i - x_j^i \right|}{n-1};\ \forall k = 1,2,...,m \text{ and } \forall i = 1,2,...,D \tag{3.13}$$

where ξ is the pheromone evaporation coefficient; and n is the problem dimension. Based on determined mean and standard deviation, the new position x_a of an ant k is generated by:

$$x_{a,k}^i = x_j^i + \sigma_k^i N(0,1)\ ;\ \forall k = 1,2,...,m \text{ and } \forall i = 1,2,...,D \tag{3.14}$$

After the complete generation of all ants, the better ant solutions replace the archived solutions based on the following rules.

- Any feasible solution is preferred to any infeasible solution;
- Between two feasible solutions, the one with the better objective value is preferred;
- Between two infeasible solutions, the one having smaller fitness value (smaller constraint violation) is preferred.

All solutions in $\mathbf{X}_T$ are then ranked after the update according to their feasibility status and fitness value based on a descending order of the solution quality. This means that the solution with the lower rank is better than the one with the higher one.

3.4.2.3. Differential evolution

Differential evolution (DE) initially proposed in [58] is a population based EA gaining significant interests in global optimization problems during the recent years due to the simplicity in implementation, the efficiency and robustness of the search process. DE works quite similarly to other EAs except the scheme of generating a new trial vector. A new individual of DE moves to a new one in the search space by adding a resulting vector of difference between two other random vectors randomly chosen from the population. A weight namely mutation factor is introduced to control the magnitude of the vector of difference. The resulting (mutant) vector replaces the vector with which it is compared only if the new vector yields the better fitness value. DE is inherently a simple to use and powerful optimization method. Moreover, it is possible to implement DE in parallel computation.

Besides those merits outlined earlier, the performance of DE is quite influenced by the setting of three important control parameters namely the mutation factor, the crossover rate

and the population size. Although there are some guidelines for appropriate parameter settings such as [59-61], the interaction between the parameter setting and the search performance is quite complicated and perhaps not completely well understood. The tedious trial-and-error method is generally used to tune the control parameters which are most of the cases problem-specific.

In recent years, there have been keen interests in the area of parameter adaptation where the value of control parameters is dynamically adjusted without prior knowledge given by the users [60, 62, 63]. JADE is a decent variant of DE recently proposed by Zhang and Sanderson [40]. Parameters of the algorithm are automatically changed during the course of optimization. JADE relies on a simple mechanism developed to accumulate the knowledge about the set of parameters that produce the better trial vectors and to adapt the control parameters in the remaining generations with the intention to improve the solution quality.

This thesis focuses on the application of JADE to optimization tasks in power systems problems. Various case studies in economic load dispatch were conducted to verify our implementation. Major steps of JADE algorithm are summarized as follows.

Solution encoding and initialization

Before DE starts the search, the component *j* or an individual *i* is initialized (denoted as the generation 0) by:

$$x_{i,j}^{0} = x_j^{\min} + rand \cdot (x_j^{\max} - x_j^{\min}) \qquad (3.15)$$

$$\forall i \in \{1,2,...,NP\} \text{ and } \forall j \in \{1,2,...,D\}$$

where $x_j^{\min}$ and $x_j^{\max}$ are minimum and maximum of variable *j*, respectively; *rand* is a uniformly distributed random number between 0 and 1; *NP* is the population size.

Mutation

For each individual *i* at the generation *G* $\mathbf{x}_i^G$ (also known as a target vector), mutation is applied to produce the corresponding mutant vector $\mathbf{v}_i^G = \left(v_{i,1}^G, v_{i,2}^G, ..., v_{i,D}^G\right)$. The basic mutation strategy is the "DE/rand/1" as follow:

$$\mathbf{v}_i^G = \mathbf{x}_{r0}^G + F_i(\mathbf{x}_{r1}^G - \mathbf{x}_{r2}^G) \tag{3.16}$$

where $r0 \neq r1 \neq r2 \neq i$ are distinct integers uniformly chosen from the set $\{1,2,...,NP\}$; F_i is the mutation strategy for the individual *i* usually in the range (0,1).. The parent vector at the generation *G* is randomly chosen as $\mathbf{x}_{r0}^G$ and updated by the difference vector $\mathbf{x}_{r1}^G - \mathbf{x}_{r2}^G$. This strategy provides great diversification to the search process because the parent vector is randomly chosen from the population. However, an effective EA should balance the search capabilities between "diversification" and "intensification". This is achieved by the "DE/current-to-best/1" mutation strategy as follows:

$$\mathbf{v}_i^G = \mathbf{x}_i^G + F_i(\mathbf{x}_{best}^G - \mathbf{x}_i^G) + F_i(\mathbf{x}_{r1}^G - \mathbf{x}_{r2}^G) \tag{3.17}$$

where $r1 \neq r2 \neq i$ are distinct integers uniformly chosen between 1 and *NP*. For each individual *i*, the mutation scheme updates the new position by adding two differential vectors to the current position at each generation *G*. The first one is the vector pointing from the current position $\mathbf{x}_i^G$ to the global best $\mathbf{x}_{best}^G$ while the second one is the difference vector between any two arbitrary points. This strategy helps accelerate the convergence because of the additional information given by the global best vector. The improved version of this strategy is called "DE/current-to-*p*best/1"

$$\mathbf{v}_i^G = \mathbf{x}_i^G + F_i(\mathbf{x}_{best,p}^G - \mathbf{x}_i^G) + F_i(\mathbf{x}_{r1}^G - \tilde{\mathbf{x}}_{r2}^G) \tag{3.18}$$

where $\mathbf{x}_{best,p}^G$ is randomly chosen from the top best 100*p* % of the current population denoted as X_P at generation *G* with probability $p \in [0,1]$. $\tilde{\mathbf{x}}_{r2}^G$ is a vector randomly chosen from the set of $X_P \cup X_A$ where X_A is the solution archive storing the local best solutions that the algorithm has discovered so far.

Crossover

After the mutation, a trial vector $\mathbf{u}_i^G = \left(u_{i,1}^G, u_{i,2}^G, ..., u_{i,D}^G\right)$ is generated by mixing the target vector and the mutant vector through the binomial crossover as:

$$u_{i,j}^{G} = \begin{cases} v_{i,j}^{G} & \text{if } rand \le CR_i \text{ or } j = j_{rand} \\ x_{i,j}^{G} & \text{otherwise} \end{cases} \tag{3.19}$$

where $CR_i \in [0,1]$ corresponds to the crossover rate of the vector *i*. The crossover rate roughly determines the fraction of vector component that are altered by the mutation process. j_{rand} is an integer randomly chosen between 1 and *D*. This is to ensure that at least one of the vector components comes from the mutation process.

Selection

The better vector between the parent $\mathbf{x}_i^G$ and the trial $\mathbf{u}_i^G$ is selected to be a member of population in the next generation *G*+1 according to their fitness values $f(\cdot)$. For a minimization problem, the selected vector is given by:

$$\mathbf{x}_i^{G+1} = \begin{cases} \mathbf{u}_i^G & \text{if } f(\mathbf{u}_i^G) < f(\mathbf{x}_i^G) \\ \mathbf{x}_i^G & \text{otherwise} \end{cases}. \tag{3.20}$$

A trail vector $\mathbf{u}_i^G$ is selected to be a parent vector of the next generation $\mathbf{x}_i^{G+1}$ only if it yields the better fitness value, otherwise the old target vector $\mathbf{x}_i^G$ is retained.

Control parameter adaptation

The crossover rate CR_i^G and the mutation factor F_i^G associated with a target vector $\mathbf{x}_i^G$ are defined as random variables as follows:

$$CR_i^G = randn(\mu_{CR}^G, 0.1) \tag{3.21}$$

$$F_i^G = randc(\mu_F^G, 0.1) \tag{3.22}$$

where $randn(a,b)$ is a random number based on the normal distribution with mean *a* and standard deviation *b*; $randc(a,b)$ is a random number based on the Cauchy distribution with location *a* and scale *b*. μ_{CR}^G and μ_F^G are the mean value of crossover rates and mutation factors corresponding to successful individuals at generation *G*, respectively. An individual is

successful at any generation if it is selected as a parent vector according to. Upon the completion of evolutionary operations to all individuals, the mean values of *CR* and *F* for the next generation *G*+1 are updated based on:

$$\mu_{CR}^{G+1} = (1-c)\mu_{CR}^{G} + c \cdot mean_A(\mathbf{s}_{CR}) \tag{3.23}$$

$$\mu_{F}^{G+1} = (1-c)\mu_{F}^{G} + c \cdot mean_L(\mathbf{s}_{F}) \tag{3.24}$$

where c is a constant generally set between 0.01 and 0.2. The vectors $\mathbf{s}_{CR}$ and $\mathbf{s}_F$ store crossover rates and mutation factors corresponding to successful individuals. The operator $mean_A(\cdot)$ and $mean_L(\cdot)$ are used to compute the arithmetic and Lehmer means of a vector. For any vector $\mathbf{x}$ with n elements, the Lehmer mean is a nonlinear moving average with a parameter p as follows:

$$L_p(\mathbf{x}) = \frac{\sum_{k=1}^{n} x_k^p}{\sum_{k=1}^{n} x_k^{p-1}}. \tag{3.25}$$

Adapting population size

Parameter adaptation can be brought into the population level so that the number of individuals can be dynamically updated over generations. Some examples of these methods are a rule-based adaptation in [64] and self-adaptive methods in [65, 66]. These strategies enable the EA to adapt itself to any class of the problem without the need for the user to specify the value of control parameters. If well designed, they may also bring more diversity to the population which is generally helpful in many optimization problems especially in multi-modal problems. So far, those techniques have not shown distinct performances in improving the solution quality. They have not been applied to any practical optimization problems.

Therefore, we developed the so-called JADE-vPS in this thesis to investigate those gaps. The mechanism to change the control parameters *F* and *CR* are kept unchanged from the original JADE. The strategy to change the population size is extended by coding a variable called *NPn* to the individual level as shown in the last column in Fig. 3.7. The variable *NPn* denotes the normalized number of individuals participating in the evolutionary operation at the current generation. After de-normalization, *NPn* becomes *NP* representing the number of

individuals to be evolved at the current generation. Moreover, *PS* represents the population size at the current generation. At the first generation, *NP* is essentially equal to *PS*.

$\mathbf{x}_1^G$	F_1^G	CR_1^G	NPn_1^G
$\mathbf{x}_2^G$	F_2^G	CR_2^G	NPn_2^G
$\mathbf{x}_3^G$	F_3^G	CR_3^G	NPn_3^G
⋮	⋮	⋮	⋮
$\mathbf{x}_{PS}^G$	F_{PS}^G	CR_{PS}^G	NPn_{PS}^G

Fig. 3.7 Encoding scheme in JADE-vPS

At the beginning, *PS* can be set to a fixed number according to one of the guidelines discussed earlier. Because *PS* can be automatically adjusted during the optimization process, *PS* has to be bounded in the range given by $[PS^{\min}, PS^{\max}]$ representing the minimum and maximum allowable size of the population. In the current version, $PS^{\min}$ and $PS^{\max}$ are set to $0.5 \cdot PS^{ini}$ and $2 \cdot PS^{ini}$, respectively where PS^{ini} is the initial population size. PS^{ini} is set to $3 \cdot D$ where *D* is the problem dimension. For each target vector *i*, the normalized number of individuals is defined by:

$$NPn_i^G = randn(\mu_{NPn}^G, 0.1) \tag{3.26}$$

where μ_{NPn}^G is the mean value of all *NPn* variables. When the algorithm starts, μ_{NPn} is initialized to 0.5. The mean value of *NPn* for the next generation is updated by:

$$\mu_{NPn}^{G+1} = (1-c_1)\mu_{NPn}^G + c_1 \cdot mean_A(\mathbf{s}_{NPn}) \tag{3.27}$$

where c_1 is a small constant (in this study set to 0.01); $\mathbf{s}_{NPn}$ is the vector storing the normalized number of individuals corresponding to the successful candidates. Then, the number of individual that the EA will evolve in the next generation denoted as NP^{G+1} is determined by:

$$NP^{G+1} = PS^{\min} + round\{\mu_{NPn}^{G+1}(PS^{\max} - PS^{\min})\} \tag{3.28}$$

where $round\{\cdot\}$ is the round-up function. At this point, there exists a possibility that NP^{G+1} is greater than PS^G. The question is now how to handle the marginal (NP^{G+1}- PS^G) individuals. There are two simple alternatives to this issue. First, the existing population is sorted and the best individuals are copied to fulfill the requirement. The other solution is to insert new individuals which can be randomly close to the local best solutions to intensify the search process or in the other hyper-spheres excluding the local bests. The latter strategy in some cases is beneficial to prevent the population trapping to a local optimum. This thesis follows the first mechanism in adding local bests. By doing this, as discussed earlier the search process could be biased toward a not truly optimal direction. At the moment, we have restricted the number of local bests PS^{added} that can be inserted to the population by offsetting (NP^{G+1}- PS^G) according to:

$$PS^{added} = \begin{cases} \min\{(NP^{G+1} - PS^G), 0.05 \cdot PS^{ini}\} & \text{if } NP^{G+1} > PS^G \\ 0 & \text{otherwise} \end{cases} \tag{3.29}$$

Hence the population size of the next generation becomes:

$$PS^{G+1} = PS^G + PS^{added} \tag{3.30}$$

Up to this point, the other important question may arise. Because *NP* changes during the optimization process, then which individuals should be updated by the evolutionary operations? Each individual in the population has equal chance of being evolved by the algorithm. The mutation scheme of JADE-vPS is identical to the one of JADE as shown in (3.18) except the definition of the individual pointer *i*. Here in JADE-vPS, a vector storing random permutation of integers between 1 to *PS* is generated. Let us denote this *PS*-element vector $\mathbf{r}_{pi}$. Each element of $\mathbf{r}_{pi}$ is a pointer to the individual on which JADE-vPS should be operated. Then, only the first *NP* elements are kept and the rest are trimmed out. Now the individual pointer *i* can be redefined by:

$$i = \mathbf{r}_{pi}(k) \; ; \; k = 1, 2, ..., NP \tag{3.31}$$

The condition of random pointers to two difference vectors $r1 \neq r2 \neq i$ remains unchanged except the two distinct integers are now uniformly chosen from the range between 1 and *PS*.

Performance of JADE-vPS

JADE-vPS is developed to handle large-scale optimization problems including power systems. At the same time, the algorithm is parameter-less and very user-friendly. To validate computational performances of the proposed algorithm, a set of selected standard test problems established in a special session of the 2005 IEEE Congress on Evolutionary Computation (CEC'05) is employed in this section. These test problems provide fair comparisons for real-parameter optimization [67]. Special operators and working principles are designed to favor the search capability to some test problems [68]. An example of these operators is the neighborhood competition operator in [69]. An EA with this operator often yields very good results in the problems the optimal solution of which lies at the center of origin. Moreover, in practical applications interchanging variables is generally not acceptable because they can be of different physical meanings. CEC'05 test problems are designed to offset these influences by shifting the global optimum to an arbitrary point and rotating the landscape of search space.

In this section, ten functions of the CEC'05 test set as listed Table 3.1are selected to represent various characteristics of optimization problems including unimodal, multimodal and expanded functions. Here, an expanded function *EF* is defined by a combination of 2-D starting function $F(x,y)$ as follows:

$$EF(x_1, x_2, ..., x_D) = F(x_1, x_2) + F(x_2, x_3) + ... + F(x_{D-1}, x_D) \tag{3.32}$$

All ten functions are used to verify the performance of JADE-vPS w/Arch and compare with five reference algorithms. The reference algorithms comprise:

1. Conventional DE with the "current-to-best" mutation scheme (DE/ current-to-best/1);
2. Developed JADE-vPS without the solution archive (JADE-vPS w/o Arch);
3. Original JADE with the solution archive (JADE w/Arch);
4. Original JADE without the solution archive (JADE w/o Arch);
5. Self-adaptive DE (SADE) [62].

For all algorithms and all test functions, the following parameters are used in all experiments: 1) the population size is set to 100; 2) the maximum number of fitness evaluations *FE* is 200000. If the population size is fixed for every generation, *FE* is the

multiplication between the population size and the maximum number of generations. The parameters F_i and CR_i are set to 0.7 and 0.2 for all individuals in the "DE/ current-to-best/1" for all test functions. The problem dimension D is set to 30 for all problems.

Table 3.1 Description of test problems

Function	Problem	Range	$F(\mathbf{x}^*)$
Unimodal functions			
$F1$	Shifted Sphere Function	$[-100,100]^D$	-450
$F2$	Shifted Schwefel's Problem 1.2	$[-100,100]^D$	-450
$F4$	Shifted Schwefel's Problem 1.2 with Noise in Fitness	$[-100,100]^D$	-450
$F5$	Schwefel's Problem 2.6 with Global Optimum on Bounds	$[-100,100]^D$	-310
Multimodal functions			
$F6$	Shifted Rosenbrock's Function	$[-100,100]^D$	390
$F8$	Shifted Rotated Ackley's Function with Global Optimum on Bounds	$[-32,32]^D$	-140
$F10$	Shifted Rotated Rastrigin's Function	$[-5,5]^D$	-330
$F12$	Schwefel's Problem 2.13	$[-\pi,\pi]^D$	-460
Expanded functions			
$F13$	Expanded Extended Griewank's plus Rosenbrock's Function ($F8F2$)	$[-3,1]^D$	-130
$F14$	Shifted Rotated Expanded Scaffer's $F6$	$[-100,100]^D$	-300

D: Problem dimension $\mathbf{x}^*$: Optimal solution

Evolutionary algorithms are stochastic by nature. This means that the results are, in some sense, random and uncertain. This randomness is due to the use of a random number generator as a part of most evolutionary operators. Therefore, to obtain statistically significant results, 25 independent trials were repeated for each test case. At every pre-specified number of fitness evaluations, we record the absolute error defined by:

$$e = \left| F(\mathbf{x}^*) - F(\mathbf{x}^{g,best}) \right| \tag{3.33}$$

where $\mathbf{x}^*$ is the known global optimum; $\mathbf{x}^{g,best}$ is the best solution that the algorithm has found so far; F is the objective function.

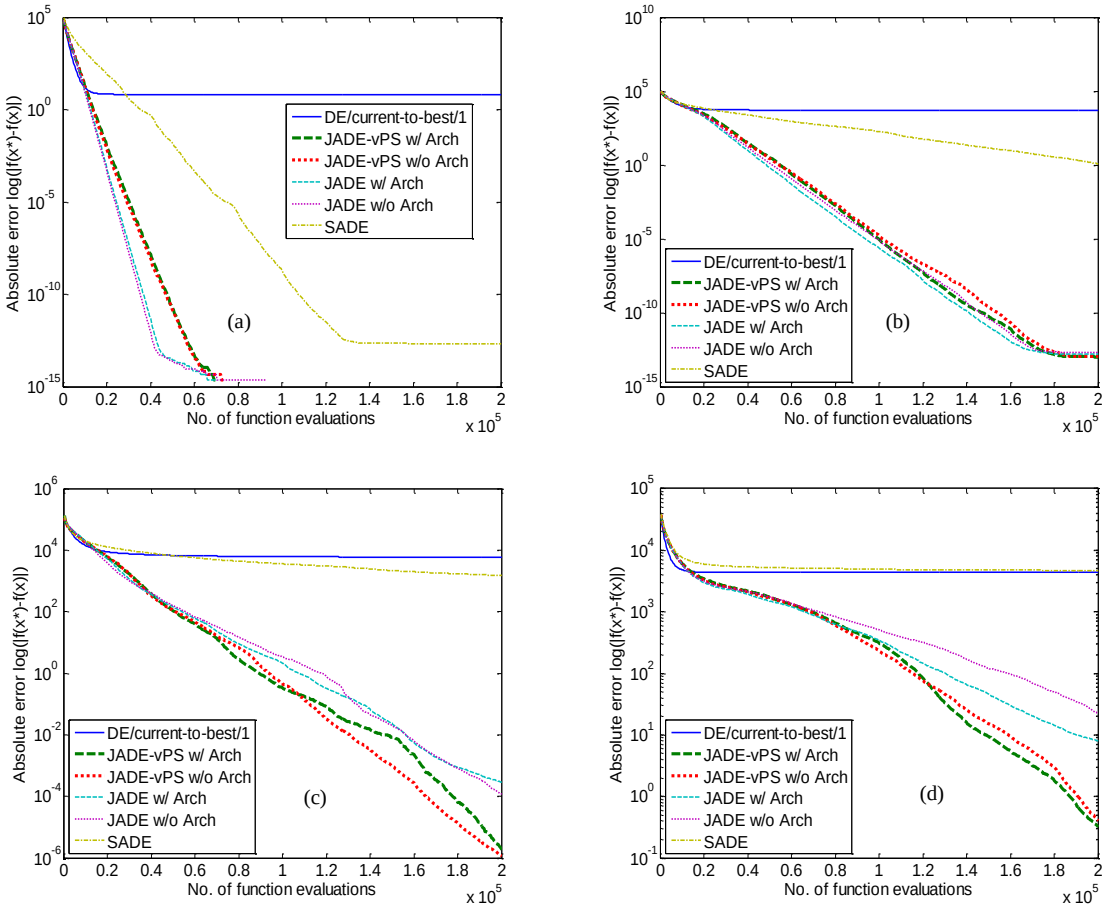


Fig. 3.8 Convergence characteristics in unimodal functions (a) *F*1 (b) *F*2 (c) *F*4 (d) *F*5

For each test problem and each algorithm, statistical values of the absolute error e consisting of average, standard deviation, best and worst results are summarized in Table 3.2 and Table 3.3. The average of sample results is the commonly used and perhaps simplest measure of EA performance. It is quite evident that JADE-vPS outperforms the other algorithms in terms of convergence and accuracy of the final results. For unimodal problems as shown in Fig. 3.8, JADE-vPS w/ Arch and JADE-vPS w/o Arch show the ability to

enhance the search diversity (exploration) in the noisy function *F*4 and the problem with optimum on bounds *F*5.

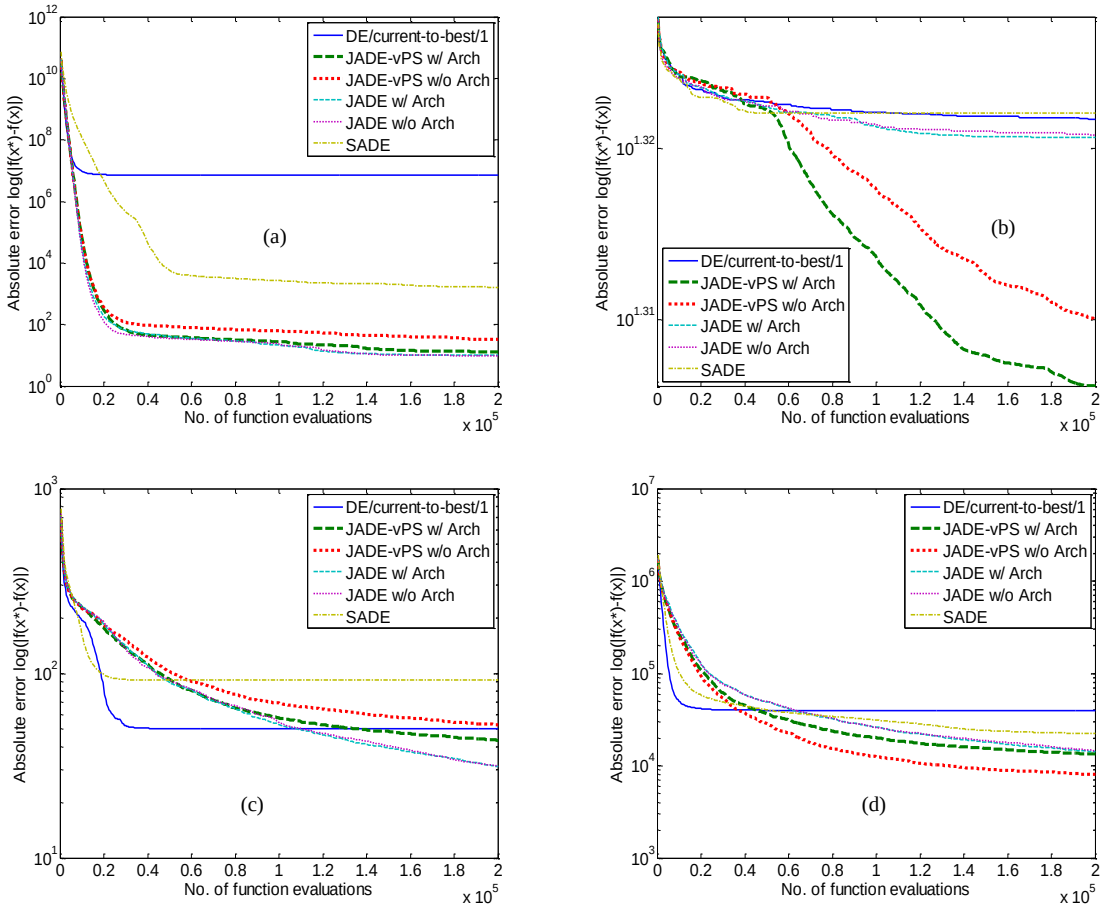


Fig. 3.9 Convergence characteristics in multimodal functions (a) *F*6 (b) *F*8 (c) *F*10 (d) *F*12

For multimodal problems, it is also interesting to see that JADE-vPS w/Arch can escape from the pre-mature convergence in the problem *F*8 as shown in Fig. 3.9(b). The function *F*8 is very difficult because there are several local optimums and the global optimum lies on the search space bounds. Even though the original JADEs reach better results in *F*10, the differences compared to JADE-vPSs are not very significant. At the moment, the results of JADE-vPS are still not outstanding in expanded functions as depicted in Fig. 3.10. More research is required to improve the search performance.

However, comparing the performance of EAs based on the average of result samples is in general insufficient. This is because a few good or bad solutions can greatly distort the

average value. Therefore, we apply the Mann-Whitney (or Wilcoxon) significant test [70] to rigorously compare the performance of EAs. The Mann-Whitney test is a non-parametric version of student t-test. It is more preferable than the t-test because there is no assumption of normality of the distribution.

The one-tailed Mann-Whitney test with the test statistic U is used in this example. The null and alternative hypotheses are:

$$\begin{aligned} H_0 &: P(fr_p \le fr_i) = 0.5 \\ H_1 &: P(fr_p \le fr_i) < 0.5 \text{ (or)} \\ &\quad P(fr_p \le fr_i) > 0.5 \end{aligned} \tag{3.34}$$

where $P(fr_p \le fr_i)$ is the probability that a result of the proposed algorithm fr_p is less (better) than or equal to a result of the i^{th} reference algorithm fr_i. If the null hypothesis H_0 is accepted, it means that the performance of the two algorithms is statistically similar. If H_0 is rejected, there are two possibilities that either the proposed algorithm statistically perform better or on the other way round. The decision on the alternative hypothesis H_1 can be made based on the p-values for left-tailed test and right-tailed test. The significant level α is chosen to be 0.05.

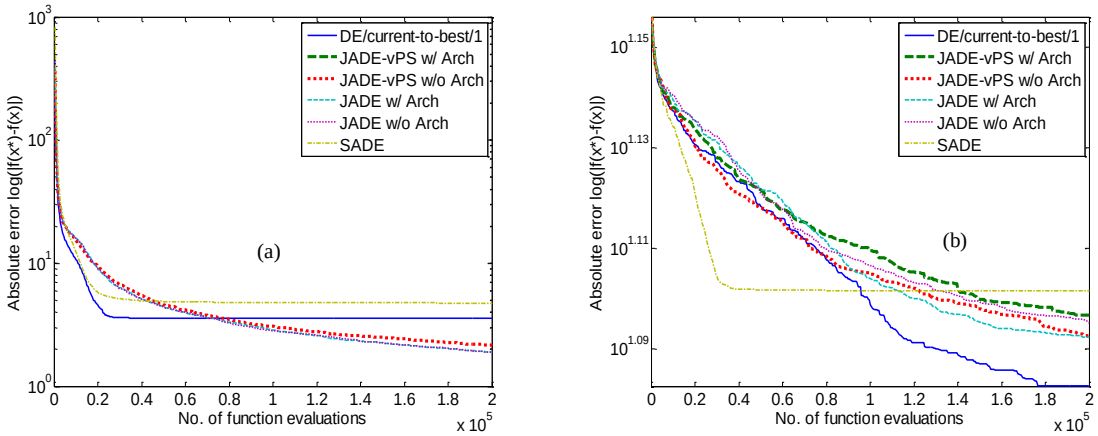


Fig. 3.10 Convergence characteristics in expanded functions (a) *F*13 (b) *F*14

Table 3.2 Statistics of final results from 25 independent trials

Method	Minimum						Maximum					
	M1	M2	M3	M4	M5	M6	M1	M2	M3	M4	M5	M6
*F*1	1.81E-03	0.00E+00	0.00E+00	0.00E+00	0.00E+00	5.68E-14	3.07E+01	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.14E-12
*F*2	1.57E+03	5.68E-14	5.68E-14	5.68E-14	5.68E-14	3.18E-02	1.27E+04	1.71E-13	2.84E-13	2.84E-13	3.98E-13	8.31E+00
*F*4	1.65E+03	5.89E-11	1.74E-09	1.97E-07	1.09E-06	1.01E+02	1.29E+04	2.73E-05	1.51E-05	4.89E-03	8.57E-04	4.50E+03
*F*5	2.84E+03	5.35E-04	1.31E-04	1.11E-02	4.32E-02	2.66E+03	7.09E+03	3.71E+00	3.99E+00	7.55E+01	2.09E+02	7.38E+03
*F*6	1.36E+04	5.68E-14	1.14E-13	5.68E-14	5.68E-14	4.73E+00	5.94E+07	1.11E+02	4.62E+02	1.25E+02	1.10E+02	1.29E+04
*F*8	2.08E+01	2.00E+01	2.00E+01	2.00E+01	2.08E+01	2.08E+01	2.10E+01	2.10E+01	2.10E+01	2.11E+01	2.10E+01	2.11E+01
*F*10	2.89E+01	2.19E+01	2.57E+01	2.16E+01	2.26E+01	5.57E+01	8.97E+01	6.59E+01	8.76E+01	4.32E+01	3.79E+01	1.78E+02
*F*12	8.43E+03	9.95E+02	1.73E+01	8.21E+03	7.58E+03	7.27E+02	8.94E+04	2.97E+04	3.19E+04	2.52E+04	2.34E+04	1.37E+05
*F*13	1.90E+00	1.69E+00	1.83E+00	1.61E+00	1.72E+00	2.74E+00	5.79E+00	2.67E+00	2.52E+00	2.15E+00	2.11E+00	7.57E+00
*F*14	1.07E+01	1.16E+01	1.18E+01	1.14E+01	1.18E+01	1.18E+01	1.31E+01	1.32E+01	1.30E+01	1.28E+01	1.28E+01	1.33E+01

M1: DE/ current-to-best/1 M2: JADE-vPS w/Arch M3: JADE-vPS w/oArch
M4: JADE-w/Arch M5: JADE-w/o Arch M6: SADE

Table 3.3 Statistics of final results from 25 independent trials (continued)

Method	Average						Standard deviation					
	M1	M2	M3	M4	M5	M6	M1	M2	M3	M4	M5	M6
*F*1	6.67E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.16E-13	9.23E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.24E-13
*F*2	5.23E+03	1.05E-13	1.07E-13	1.57E-13	1.93E-13	1.24E+00	2.51E+03	4.24E-14	5.76E-14	5.26E-14	8.37E-14	1.95E+00
*F*4	5.82E+03	1.88E-06	1.15E-06	2.92E-04	1.19E-04	1.50E+03	2.70E+03	5.52E-06	3.08E-06	9.78E-04	2.14E-04	1.45E+03
*F*5	4.36E+03	3.23E-01	4.09E-01	7.94E+00	2.22E+01	4.66E+03	1.08E+03	8.00E-01	9.33E-01	1.75E+01	4.78E+01	1.18E+03
*F*6	7.27E+06	1.32E+01	3.37E+01	1.03E+01	9.79E+00	1.63E+03	1.41E+07	2.78E+01	9.99E+01	2.58E+01	2.71E+01	3.45E+03
*F*8	2.10E+01	2.02E+01	2.04E+01	2.09E+01	2.09E+01	2.10E+01	5.30E-02	3.46E-01	4.27E-01	1.98E-01	5.99E-02	7.76E-02
*F*10	5.01E+01	4.33E+01	5.28E+01	3.12E+01	3.14E+01	9.18E+01	1.60E+01	1.26E+01	1.60E+01	4.72E+00	5.08E+00	2.96E+01
*F*12	3.97E+04	1.34E+04	8.01E+03	1.43E+04	1.46E+04	2.22E+04	1.93E+04	7.78E+03	6.79E+03	3.91E+03	3.93E+03	2.75E+04
*F*13	3.58E+00	2.14E+00	2.15E+00	1.89E+00	1.89E+00	4.74E+00	1.02E+00	2.38E-01	1.88E-01	1.40E-01	9.91E-02	1.11E+00
*F*14	1.21E+01	1.25E+01	1.24E+01	1.24E+01	1.25E+01	1.26E+01	7.33E-01	4.23E-01	3.54E-01	3.42E-01	2.30E-01	4.11E-01

M1: DE/ current-to-best/1 M2: JADE-vPS w/Arch M3: JADE-vPS w/oArch
M4: JADE-w/Arch M5: JADE-w/o Arch M6: SADE

Table 3.4 Statistical comparisons among different algorithms

Problem	M2<M1	M2<M3	M2<M4	M2<M5	M2<M6
*F*1	/	=	=	=	/
*F*2	/	x	/	/	/
*F*4	/	/	/	/	/
*F*5	/	x	/	/	/
*F*6	/	x	/	x	/
*F*8	/	x	/	/	/
*F*10	=	/	x	x	/
*F*12	/	x	=	=	=
*F*13	/	=	x	x	/
*F*14	=	=	=	=	=

M1: DE/ current-to-best/1 M2: JADE-vPS w/Arch M3: JADE-vPS w/oArch
M4: JADE-w/Arch M5: JADE-w/o Arch M6: SADE

/ : M2 performs statistically better than the corresponding algorithm.
x: The corresponding algorithm is statistically better than M2.
=: There is no statistical difference between the two algorithms.

Table 3.4 gives statistical comparisons based on empirical experiments between JADE-vPS (M2) and the five reference algorithms at α=0.05. The first column compares the performance of M1 and M2. It is obvious that M2 performs better than M1 in most of the problems. In the second column, M2 does not show significant superiority to M3. Note that the difference between the two algorithms is the solution archive. Therefore, one may conclude that the archive of local best solutions is not helpful in obtaining the better solutions. Comparing M2 and M4 where the only difference is the adaptation of population size, M2 outperforms M4 in a number of problems especially in unimodal functions. The similar observation can be made in the case between M2 and M5. Therefore, it can conclude that the proposed scheme for population size adaptation can help improve the search performance. The last column exhibits the superior search capability of the proposed M2 to an adaptive DE variant M6 in most of the test problems.

3.4.3 Constraint handling

The success of EA in solving a complex optimization problem greatly depends on the ability to treat both equality and inequality constraints. Therefore, one of the major components for

any EA applied to a constrained optimization problem is the fitness function. Fitness functions are used to assign a qualitative measure to individuals in the population. To properly handle constraints in an EA, the fitness function should be carefully designed such that it can help guide the evolutionary process to the promising and feasible search space. The issue of how to deal with infeasible individuals in the EC paradigm is quite complicated and still far from trivial. There are a number of researchers in the evolutionary computation community devoting to this area. This section briefly discusses classification of constraint handling techniques in nonlinear programming problems. The penalization scheme commonly used in many EA applications is further elaborated.

A general nonlinear programming problem is to find the decision variable **x** in the *D*-dimensional real space so as to:

$$\text{minimize } f(\mathbf{x}) \text{ ,} \mathbf{x} = (x_1, x_2, ..., x_n) \in R^D \tag{3.35}$$

subject to

$$\begin{aligned} & g_i(\mathbf{x}) \le 0 \text{ , } i = 1, 2, ..., q \\ & h_j(\mathbf{x}) = 0 \text{ , } j = q+1, q+2, ..., q+m \end{aligned} \tag{3.36}$$

where *q* and *m* are the number of inequality and equality constraints, respectively. A condition of optimality is $\mathbf{x} \in \mathbb{F} \subseteq \mathbb{S}$ where $\mathbb{F}$ is a feasible part of the search space $\mathbb{S} \subseteq R^n$. The search space in each dimension is normally bounded as:

$$x_{i,\min} \le x_i \le x_{i,\max} \text{ , } i = 1, 2, ..., D \tag{3.37}$$

An inequality constraint is called an active constraint if $g_i(\mathbf{x}) = 0$. Thereby, all equality constraints are active constraints. In general, an EA recognizes feasibility of an individual only when the value of constraint function is negative. Therefore, equality constraints should be transformed to inequality ones by introducing the tolerance value ε as follows:

$$\left| h_j(\mathbf{x}) \right| - \varepsilon \le 0 . \tag{3.38}$$

The major issue in constraint handling techniques is how to handle infeasible individuals properly. According to [71], those methods can be categorized as follows:

a) Rejection strategy

The infeasible individuals are discarded from the population. This method has serious limitations if the ratio between the sizes of $\mathbb{F}$ and $\mathbb{S}$ is small and the initial population

consists of only infeasible solutions. Comprehensive research has revealed that infeasible solutions are not always undesirable. They provide useful information and should not be simply discarded from the population.

b) Repair strategy

A problem-dependent mechanism can be used to repair an infeasible solution and transform it to a feasible one. This strategy can help the optimization algorithm solve the problem in a reasonable time and sometimes even provide a better result. In this framework, constraint satisfaction methods [72] can be incorporated in a normal EA routine. However, no definitive conclusion can be made so far and it is still an open-end question requiring further research.

c) Multi-objective optimization strategy

An alternative constraint handling method is to formulate a multi-objective optimization problem with the objectives consisting of the original objective function f and constraint violation measures g_j. Any EA for multi-objective optimization can be used in the context of constrained optimization.

d) Penalizing strategy

As discussed earlier that the rejection strategy may give very harmful effects to the search process and a special operator in the repair strategy may be not simply developed and generally problem dependent, the most common method to treat infeasible individuals in EAs is the penalizing scheme. A suitable model and parameters of fitness function can influence the solution quality. Reviews on essential backgrounds and the penalty methods used in this thesis are summarized in the following subsection.

3.4.3.1. Penalty functions

A penalty function transforms a constrained optimization problem to an unconstrained one by one of the two ways. First, it is called the additive form as follows [73]:

$$eval(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \text{ is feasible} \\ f(\mathbf{x}) + p(\mathbf{x}) & \text{otherwise} \end{cases} \tag{3.39}$$

where $p(\mathbf{x})$ represents the penalty term for an individual $\mathbf{x}$. If there is no violation, p($\mathbf{x}$) is zero and a positive value otherwise. The overall EA objective function becomes $eval(\mathbf{x})$. The second method is the multiplicative form as:

$$eval(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } \mathbf{x} \text{ is feasible} \\ f(\mathbf{x})\,p(\mathbf{x}) & \text{otherwise} \end{cases} \tag{3.40}$$

If there is no violation, $p(\mathbf{x})$ is one and a positive value otherwise. In most EAs, the additive form is more popular than the multiplicative one.

Static penalties

In static penalty methods, penalty parameters are independent of the generation number. An example of these methods is proposed in [74]. In this approach, the users can describe different levels of violation and thereby assign different penalty coefficients of each constraint. The fitness evaluation function can be described by:

$$eval(\mathbf{x}) = f(\mathbf{x}) + \sum_{j=1}^{q+m} R_{ij} \max[0, g_j(\mathbf{x})]^{\beta} \tag{3.41}$$

where R_{ij} indicates the penalty coefficient of the constraint j at the violation level i; β is the order of penalty term usually set to 1 or 2. Equality constraints are transformed to inequality ones by (3.38). The evident disadvantage of this method is the large number of parameters that must be set by the user. For example, if there are 5 constraints and 4 levels of violation, there are 45 parameters to be set. It was demonstrated in [75] also that the solution quality is very sensitive to the value of these parameters.

Adaptive penalties

To eliminate the parameter tuning process, the self-learning penalty function proposed in [76] is applied in this thesis. The elegant feature of this technique is that it is parameter-less and capable of adjusting the penalized fitness function at different stages of the search process. The fitness evaluation function is computed by:

$$eval(\mathbf{x}) = d(\mathbf{x}) + p(\mathbf{x}) \tag{3.42}$$

where $d(\mathbf{x})$ and p(x) represents the distance term and the penalty term of an individual $\mathbf{x}$, respectively. The distance value can be determined by:

$$d(\mathbf{x}) = \begin{cases} v'(\mathbf{x}) & \text{if } r_f = 0 \\ \sqrt{f'(\mathbf{x})^2 + v'(\mathbf{x})^2} & \text{otherwise} \end{cases} \tag{3.43}$$

where r_f is the ratio of the feasible individuals in the population. The normalized objective value of a vector **x** is computed by:

$$f'(\mathbf{x}) = \frac{f(\mathbf{x}) - f^{\min}}{f^{\max} - f^{\min}} \tag{3.44}$$

where $f(\mathbf{x})$ is the corresponding objective value of the given vector **x**; $f^{\min}$ and $f^{\max}$ are minimum and maximum objective values for all individuals in the population. Based on the definition in (3.36), the total constraint violation for each individual **x** is determined by:

$$v(\mathbf{x}) = \sum_{i=1}^{q+m} \max(0, g_i(\mathbf{x})) \tag{3.45}$$

Note that all equality constraints are transformed to ineqaultiy ones by using (3.38). The total constraint violation is normalized within [0,1] by

$$v'(\mathbf{x}) = \frac{v(\mathbf{x}) - v^{\min}}{v^{\max} - v^{\min}} \tag{3.46}$$

where $v^{\min}$ and $v^{\max}$ are minimum and maximum of the total constraint violation for all individuals in the population.

The penalty term $p(\mathbf{x})$ is set up to ensure that an infeasible individual will receive the appropriate fitness value at different stages of the search process in terms of the feasibility and objective value. For example, when the population lacks feasible individuals, infeasible individuals with lower constraint violations should receive relatively less penalty irrespective to the objective value. On the other hand when there are reasonable feasible individuals, infeasible individuals with lower objective value are assigned lower penalty. This idea can be mathematically expressed by:

$$p(\mathbf{x}) = (1 - r_f) X(\mathbf{x}) + r_f Y(\mathbf{x}) \tag{3.47}$$

where

$$X(\mathbf{x}) = \begin{cases} 0 & \text{if } r_f = 0 \\ v'(\mathbf{x}) & \text{otherwise} \end{cases} \tag{3.48}$$

$$Y(\mathbf{x}) = \begin{cases} 0 & \text{if } \mathbf{x} \text{ is feasible} \\ f'(\mathbf{x}) & \text{if } \mathbf{x} \text{ is infeasible} \end{cases} \tag{3.49}$$

3.5 Summary

Advanced developments in computation intelligence (CI) over the past few decades have enhanced solving many complex problems more conveniently and more effectively than before. This chapter presents CI tools and relevant techniques that have been developed and applied in this thesis. Pre-processing is an important process that can have significant effect in the generalization performance. After normalizing the raw input data to a certain range, dimensionality reduction schemes can be applied to transform the original and perhaps large inputs to the new and more meaningful ones. Feature selection and feature extraction are major members of this paradigm.

Neural network (NN) is a powerful tool which maps the set of input features into the considered outputs. Once properly trained, NN may perform the designated task closely to the actual system.

Evolutionary computation is the primary concentration of this chapter. Three evolutionary algorithms; genetic algorithm, ant colony optimization and differential evolution have been applied as the optimization tool in various problems in this thesis. A new DE variant called JADE-vPS is developed and extensively addressed in this chapter. Various test problems for unconstrained optimization are used to verify the performance of the algorithm. The investigation based on non-parametric statistical test reveals that the JADE-vPS performs better than the other versions of DE. The convergence of the new algorithm is also improved. Apart from the improvement in search capabilities, the proposed algorithm is parameter-less. Therefore, it eliminates the need for parameter tuning.

Finally, theoretical backgrounds of constraint handling are reviewed. Different categories of constraint handling methods are presented.

Chapter 4
Estimation of voltage stability margin

4.1 Introduction

Power systems in many parts of the world have been operated closer to technical limits than before majorly due to restrictions in generation and transmission grid expansions. It is therefore desirable to optimally utilize the existing network without jeopardizing the system security. An online tool for estimating the proximity to various stability limits under normal and contingency conditions is one of the important components of dynamic security assessment (DSA) software. For voltage stability, several techniques as per discussed in Chapter 2 have been used to analyze voltage stability margin (VSM) of power systems. Among those approaches, the maximum loadability margin (MLM) is the most popular index. MLM is defined by the distance in terms of power from the current operating condition to the voltage collapse point. Continuation power flow is the common method to determine MLM. However, CPF has two distinct drawbacks namely the high computational burden and the information of load increase directions. The first issue is principally because multiple power flow calculations are required during the continuation process. Secondly, the direction of load increase and generation increases is not readily available on the real-time basis.

To overcome the aforementioned difficulties, several techniques have been introduced for fast and accurate estimation of the system stability margin. Neural network (NN) as presented in section 3.3 is one of powerful nonlinear mapping tools. One of pioneering attempts NN for dynamic security assessment was presented in [77]. In online stability assessment, NN has been successfully applied to transient stability [78, 79], small signal stability [80] and voltage stability [81-93].

This chapter presents major steps for implementing a NN-based MLM estimator. An offline training session can be set up to determine appropriate weighs in each layer of NN. A number of the selected input features and the corresponding target(s) are essential for training a NN. Operational data and several network calculations are maintained in energy

management system and available in real practice. For research purposes, a similar group of data can be also simulated. Few methods were applied and developed in this thesis to generate databases for developing MLM estimators.

The remaining of this chapter is organized as followings. Section 4.2 discusses all relevant aspects for developing a neural network to estimate the stability margin. The method for estimating voltage stability margin of the power system by sensitivity analysis is presented in Section 4.3. Simulation results of the two methods of estimation are discussed in Section 4.4. Finally, the chapter is summarized in Section 4.5.

4.2 Neural network approach

4.2.1 General concepts

A stability index including voltage stability is a highly nonlinear function of the system operating condition defined by the vector of system variables **x** before the disturbance occurs. In most cases, the mapping function $f(\mathbf{x})$ cannot be analytically expressed because power system models usually involves a large number of coupled nonlinear differential and algebraic equations as discussed in Chapter 2. Therefore, machine learning techniques can play a significant role in approximating the stability index. In this thesis, the focus is given on neural networks applied to estimate VSM. Such a problem can be described as follows.

An operating condition of a power system can be characterized by different combinations of power system variables. Pre-selection based on engineering judgment is a very important process in order to select the variables relevant to the problem at hand. The size of input vector can be reduced by a dimensionality reduction technique as discussed in Chapter 3. Then, continuation power flow or any other analysis method which were presented in Chapter 2 can be used to determine the stability index corresponding to the given operating point.

The entire database is divided into two data sets. The first set is to train the estimating function $\hat{f}(\mathbf{x})$ while the second set is to validate the estimated function. The term "generalization" refers to the ability of the developed NN in predicting the corresponding target of an unforeseen input pattern accurately. The training session is accepted if the pre-defined performance measure is within a tolerable range. Otherwise, some actions have to be done such as changing training algorithms and/or network configurations, or using different input features. However, setting the appropriate threshold level for performance measures can

play a significant role in the system performance and practically requires in-depth expertise of that particular technical field.

4.2.2 Performance metrics

After implementing a computational intelligence (CI) system, such as a neural network, one of the important concerns is how well the CI system performs. A performance metric is a general term to measure between the activities of an engineering system and the performance. There are general issues in determining the performance of CI implementation. For NNs, these issues may include specifying the size and the number of iterations of the training dataset. Also, selection of the training and testing datasets may also influence the network performance. In this section, some of those crucial issues in NN implementations are discussed. Comprehensive presentation of this topic can be found in Chapter 10 of [45].

4.2.2.1. Selecting standard measures

The objective of the back-propagation algorithm is to adapt the network weights so as to minimize the average sum-square error defined as:

$$e_t = 0.5\sum_{k=1}^{m}\sum_{j=1}^{q}\left(y_{kj} - \hat{y}_{kj}\right)^2 \tag{4.1}$$

where m is the number of patterns; q is the number of neurons in the output layer; y_{kj} is the target output j for the pattern k that NN is supposed to accurately predict and $\hat{y}_{kj}$ is the result that NN actually gives for the corresponding y_{kj}. Even though the total error is a part of a back-propagation algorithm, it has very little physical meaning. To better understand a system's performance, it is desired to have a measure at a more intuitive level. An example of those measures is the root mean square error defined by:

$$e_{rms} = \sqrt{\frac{1}{mq}\sum_{k=1}^{m}\sum_{j=1}^{q}\left(y_{kj} - \hat{y}_{kj}\right)^2} \tag{4.2}$$

For a classification problem, the performance measure is quite straightforward because we only need to show disagreement between a pattern's actual class and the obtained result. A problem of two-group classification of power system security is taken as an example here.

Three measures can be used to assess the performance of the implemented security classifier namely; success rate *SR*, false alarm rate *FAR* and missed alarm rate *MAR* defined as follows [94]:

$$SR = \frac{\text{No. of operating states correctly classified}}{\text{Total no. of operating states}} \tag{4.3}$$

$$FAR = \frac{\text{No.of SOSs classifed as ISOSs}}{\text{No. of SOSs}} \tag{4.4}$$

$$MAR = \frac{\text{No.of ISOSs classifed as SOSs}}{\text{No. of ISOSs}} \tag{4.5}$$

where SOS stands for a secure operating state and ISOS stands for an insecure operating state.

4.2.2.2. Partitioning the patterns

Testing a developed NN with the same input patterns as the network is generally not accepted. A portion of patterns is often reserved using random selection for testing. It is frequently a good idea to rotate the training and testing cases through all available data. This is intentionally to examine the generalization performance of the implemented network. When training a NN by back-propagation algorithms, one of the important concerns is the number of patterns for each classification and the numerical distribution of patterns. To our knowledge, there are still no empirical rules for allocating training patterns. Some ideas have been discussed in CI research communities. For a classification problem, some people believe that the same number of patterns should be used for each classification. To some other people, this idea is counter-intuitive. They believe that the numerical distribution of patterns should reflect the probability distribution of the class. This means, for example, that if a class appears 20% of the time. Then, 20% of the training cases should contain this class. From the author's experience, this issue is very crucial in a classification task.

4.2.2.3. Cross validation

Cross validation is a technique to estimate how well a predictive model such as NN will perform with an unforeseen dataset. The key objective of cross validation is to prevent the network from over fittings to the training patterns. The common procedures for cross

validation starts from portioning the training set into *S* distinct segments. Then, the data from *S*-1 remaining segments is used to train the network. The trained network is test using the remaining segment and the error function is evaluated. The process is repeated for all possible *S* combinations of the segment. The test errors are averaged over all *S* results. The disadvantage of this method is clearly the large amount of computing time because of repetitive trainings. Cross validation can be alternatively used in early stopping which is a form of network regularization. After each sweep of the new training set, the network performance is evaluated with the validation dataset. When the performance stops improving, the training halts. In other words, the validation dataset is used to decide when to stop the training. This simple method is very common to NN implementations and often produces good generalization.

4.2.3 Database generation

To train a machine learning method for approximating the stability index, a large number of input patterns (in this case operating states) are required. Historical data and results of unusual operating cases are used for training. However, the available data are usually inadequate for this application. Thus, some methodologies have been developed to simulate a number of operating states. Monte Carlo simulation is a technique to obtain statistical parameters of historical power system behaviors. Based on these parameters, new operating states resembling these statistical characteristics can be generated. This method is very technically attractive because it gives very realistic operating states. However, an initial set of historical operating states are still required in Monte Carlo simulations. In academic research communities, it is quite hard for us to obtain this actual data. Therefore, some techniques have been applied to generate a database of operating states based on pre-specified and assumed statistical characteristics. In this section, an example of traditional method used in literatures is reviewed. Afterwards, a new method developed in this thesis will be discussed.

4.2.3.1. Traditional methods

A simple idea to simulate a database of operating states is to introduce random perturbation to a base-operating state. Most of the works in literature assumed that there is no correlation

among load demands at different locations. For an operating state *k*, an example of such methods can be described as follows:

$$P_d^i(k) = P_{d0}^i \Delta L(k) + 2(0.5 - \varepsilon_{pd}^i(k)) P_{d0}^i \Delta P_d^i \tag{4.6}$$

$$Q_d^i(k) = Q_{d0}^i \Delta L(k) + 2(0.5 - \varepsilon_{qd}^i(k)) Q_{d0}^i \Delta Q_d^i \tag{4.7}$$

where $P_d^i(k)$ and $Q_d^i(k)$ are real and reactive power demands at the bus *i*, respectively; P_{d0}^i and Q_{d0}^i are base-case real and reactive power demands at the bus *i*, respectively; $\Delta L(k)$ is a random change of the base-case power demand (i.e. varied between 0.8 and 1.3); ΔP_d^i and ΔQ_d^i are maximum allowable change of active and reactive power at the bus *i*, respectively; $\varepsilon_{pd}^i(k)$ and $\varepsilon_{qd}^i(k)$ are uniformly uncorrelated random numbers in the range [0,1] for active and reactive power perturbation at the bus *i*, respectively. The idea of the two above formulas is to randomly perturb the base-case demand by either reducing (ΔL <0) or increasing (ΔL >0) the demand level. Then, the last term of (4.6) and (4.7) levitates the power demand level if $\varepsilon_{pd}^i(k)$ and $\varepsilon_{qd}^i(k)$ are greater than 0.5. Otherwise, the power demand level will be reduced.

Similarly, the active power generation and the generator voltage set point at a given generator bus *i* for an operating state *k* can be randomly generated around the base case as follows:

$$P_g^i(k) = P_{g0}^i \Delta P(k) + 2(0.5 - \varepsilon_{pg}^i(k)) P_{g0}^i \Delta P_g^i \tag{4.8}$$

$$U_g^i(k) = U_{g0}^i \Delta U(k) + 2(0.5 - \varepsilon_{ug}^i(k)) U_{g0}^i \Delta U_g^i \tag{4.9}$$

where P_{g0}^i and U_{g0}^i are base-case active power generation and voltage set point at the generator bus *i*, respectively; $\Delta P(k)$ and $\Delta U(k)$ are random changes of the base-case active power generation and voltage set point; $\varepsilon_{pg}^i(k)$ and $\varepsilon_{ug}^i(k)$ are uniformly uncorrelated random numbers in the range [0,1] for active power generation and voltage set point at the bus *i*, respectively; ΔP_g^i and ΔU_g^i are maximum allowable change of active power generation and voltage set point at the bus *i*, respectively

4.2.3.2. The proposed method

The tradition method recently discussed generates a new operating state by introducing small and random perturbation to the base operating state. This method has two major shortcomings. First, the power demand at all bus locations of a new sample is uncorrelated. Second, variation of control parameters, such as generator reactive power outputs, transformer tap positions, settings of compensating devices is not taken into consideration. Therefore, the operating states in the dataset produced by the method may be clogged around the base-condition. The machine learning trained by this dataset may perform very well but only within the vicinity of that operating point.

In this thesis, we developed a new dataset generation method that spans over the entire daily operation and incorporates various operational aspects. Active and reactive power demands are considered to be statistically correlated. Multivariate statistics is used to model this relation. Moreover, the system parameters associated with reactive power control are defined by the uniform distribution within the corresponding ranges.

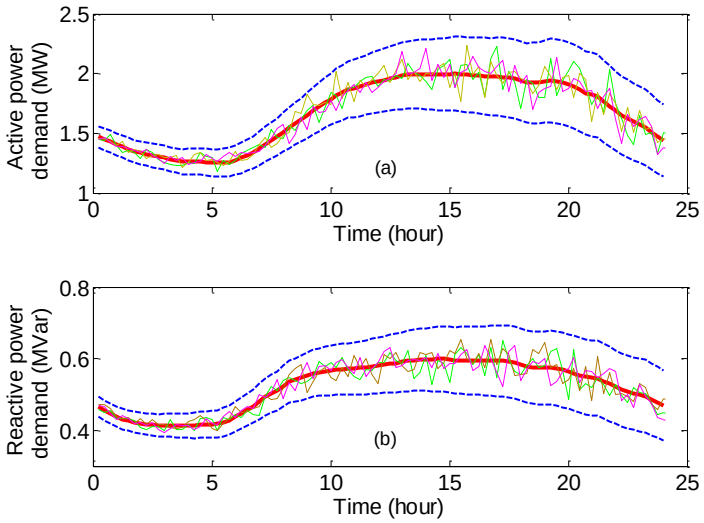


Fig. 4.1 Daily active (a) and reactive (b) power demand profiles with probable operating regions

To define several operating points of a power system, a daily load profile with forecast error is considered. The forecast error is assumed to be linearly increasing from 2% to 5% from the beginning toward the end of the day. This is due to the cumulative forecast error

from the preceding time steps. The nominal load profile with upper and lower bands of forecast error define the probable operating region in the daily operation as shown in Fig. 4.1. The red bold line is the nominal profile while the dotted upper and lower lines correspond to maximum and minimum boundaries, respectively. The total demand variation at every time step *t* is modeled by a multivariate normal distribution function. There are two major inputs for this distribution function namely the mean values of random variables and the covariance matrix.

For predictive purposes, statistical properties of a complex system can be derived from analyses of the observed or measured data. Unfortunately, the data is neither always available nor complete. One of the solutions to this problem is to randomly create a valid correlation matrix defined in terms of the physical units of the problem under investigation (in our case MW and MVar).

In multivariate statistics, a correlation matrix is valid only if it is a symmetric positive-semidefinite matrix whose (*i*,*j*) element is the correlation coefficient between any two random variables. In this thesis, we have used the technique in [95] to compute the correlation matrix nearest to validity. The valid correlation matrix is then converted to the covariance matrix $\mathbf{\Sigma}$ based on an additional input vector containing expected standard deviation of independent variables [96].

A two-step method is developed in this thesis in order to construct a large number of operating conditions. To better illustrate the ideas, the load profiles shown in Fig. 4.1 are taken as the example.

The sampling time used in this study is 15 minutes. Therefore, there are 96 time steps for a day. Note that at the beginning we have only forecast active PF_j and reactive QF_j power demands at each time step *j*. The new matrix of *n*-sample **XS** can be drawn from a multivariate normal population of *p* dimension according to:

$$\mathbf{XS} \sim N_p(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) \; ; \; \forall j \in \{1, 2, ..., 96\} \tag{4.10}$$

where $\boldsymbol{\mu}_j$ is the vector of mean values given by $\boldsymbol{\mu} = \left[PF_j, QF_j \right]$ (p=2) and $\boldsymbol{\Sigma}_j$ is the covariance matrix with 2×2 dimension at the time step *j*. Let suppose if this process is

repeated three times, it is possible to generate n= 288 (96×3) samples of the total active and reactive power demand.

Denote $\mathbf{XS}_i = [PT_i, QT_i]$ the i^{th} row vector of the n-sample matrix where PT_i and QT_i are the total active and reactive power demand of the operating point i, respectively. Then, the second step will allocate PT_i and QT_i to all the load (PQ) buses by using a random distribution with a multivariate probability density function.

The dimension of random variables p is now equal to the number of load buses denoted by m. A two step transformation technique is used to simulate dependent random variables. First, random vectors are randomly chosen from multivariate normal distribution according to (4.10) with a zero-mean vector and a random covariance matrix. Each independent variable is assumed to have random variation between -1 to 1. The valid covariance matrix is computed as discussed earlier. Observe that most of samples drawn from a normal distribution lie within $\pm 3\sigma$ where σ is the standard deviation. In the second step of the transformation, the cumulative normal distribution function (CDF) is applied to **XS** with the parameters including mean of zero and standard deviation of one as follows:

$$\mathbf{XK} = CDF(\mathbf{Xs}, 0, 1) \tag{4.11}$$

where **XK** is a $n \times m$ matrix whose elements randomly lie in [0,1]. Since our target is to find the vector of random distribution factors, therefore it is desirable to calculate such a vector the element of which is in [0,1]. The sum of all elements should also be exactly one. Then, each row i of **XK** denoted by $\mathbf{XK}_i$ is normalized according to:

$$\mathbf{XK}'_i = \mathbf{XK}_i \Big/ \sum_{j=1}^{m} XK_{i,j} \tag{4.12}$$

where $\mathbf{XK}'$ is the normalized matrix of **XK** such that $\sum_{j=1}^{m} XK'_{i,j} = 1 \; ; \forall i$ and $XK'_{i,j} > 0 \; ; \forall i$ and $\forall j$. Given the distribution vector, the load demand at bus j of operating condition (pattern) i can be determined from:

$$PD_{i,j} = PT_i \cdot XK'_{i,j} \tag{4.13}$$

$$QD_{i,j} = QT_i \cdot XK'_{i,j} \tag{4.14}$$

To visualize the simulated correlation of power demand among load buses, few scatter plots of active power demand P_D of two buses at different correlation coefficients *R* are shown in Fig. 4.2. Linear regression is shown as a red line in each sub figure. It can be observed from Fig. 4.2 (a) and (b) that P_D at bus 5 has positive correlation with P_D at bus 15 while negative correlation exists between P_D at buses 5 and 12. There is relatively less correlation in case of Fig. 4.2 (c) and (d).

Beside power demand adjustments discussed above, random changes of reactive power control variables are also considered in this study. This variation is included in the dataset generation method because of two reasons. Firstly, these control variables are continuously varied in practical operating situations. Secondly, the developed NN will be used to approximate VSM in an optimal power flow (OPF) model which will be presented in Chapter 6. The decision variables of that OPF are varied during the ACO course of optimization. Therefore, the knowledge of such variations should be inherited in the dataset.

Shunt capacitors and transformer tap positions are of discrete characteristics. The two variables are assumed to be of uniform distribution. Therefore, a discrete variable *j* for an operating state *i* can be generated according to:

$$XD_{i,j} = XD_{j,\min} + round((XD_{j,\max} - XD_{j,\min}) \cdot rand) \tag{4.15}$$

where $XD_{j,\min}$ and $XD_{j,\max}$ are minimum and maximum boundaries of the discrete variable *j*, respectively; *rand* is a uniform random variable in [0,1] and the function *round*(·) is used to round to the nearest integer.

Generator reactive power and shunt compensating devices are modeled as continuous variables. Therefore, a variable *j* of an operating state *i* is given by:

$$XC_{i,j} = XC_{j,\min} + (XC_{j,\max} - XC_{j,\min}) \cdot rand \tag{4.16}$$

where $XC_{j,\min}$ and $XC_{j,\max}$ are minimum and maximum boundaries of the continuous variable *j*, respectively.

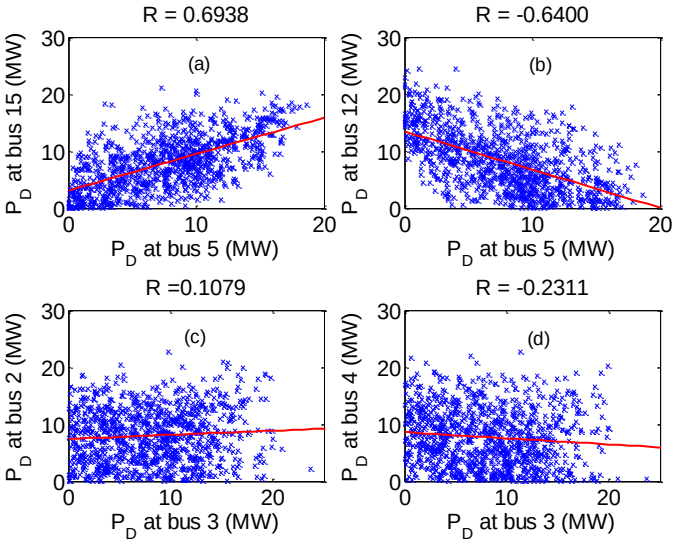


Fig. 4.2 Correlation of load demands

For all defined operating conditions, power flow is solved to ensure that it is a feasible one. The unsolvable power flow cases are simply rejected. Voltage stability of a power system can be expressed by an index as discussed in Chapter 2. In this chapter, the system loadability margin λ is the chosen stability index.

4.3 Sensitivity approach

The objective of this section is to discuss another alternative approach beside NN to estimate VSM. In voltage stability analysis, it is generally not adequate to merely obtain the critical point. It is also important to know how the change of system conditions will influence such a critical point. Sensitivity analysis is a key instrument to obtain the information about how different system parameters affect the stability index. However, this sensitivity information is valid only in the vicinity of the point of linearization. To correctly analyze the system sensitivity, the system equations should be linearized.

Recalled the system DAE model introduced in Chapter 2, it can be parameterized by the scalar λ and the vector of parameters $\mathbf{p}$ as follows:

$$\begin{aligned} \dot{\mathbf{x}} &= f(\mathbf{x}(\lambda(\mathbf{p}),\mathbf{p}),\mathbf{y}(\lambda(\mathbf{p}),\mathbf{p}),\lambda(\mathbf{p}),\mathbf{p}) \\ 0 &= g(\mathbf{x}(\lambda(\mathbf{p}),\mathbf{p}),\mathbf{y}(\lambda(\mathbf{p}),\mathbf{p}),\lambda(\mathbf{p}),\mathbf{p}) \end{aligned} \tag{4.17}$$

At the voltage collapse (saddle node bifurcation) point which is also an equilibrium point, the partial derivatives with respect to the parameter vector **p** of the above two equations can be shown as:

$$\underbrace{\begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}}_{\mathbf{A_{sys}}} \begin{pmatrix} \frac{\partial x}{\partial \lambda}\frac{\partial \lambda}{\partial p} + \frac{\partial x}{\partial p} \\ \frac{\partial y}{\partial \lambda}\frac{\partial \lambda}{\partial p} + \frac{\partial y}{\partial p} \end{pmatrix} + \begin{pmatrix} f_\lambda \\ g_\lambda \end{pmatrix}\frac{\partial \lambda}{\partial p} + \begin{pmatrix} f_p \\ g_p \end{pmatrix} = 0 \tag{4.18}$$

Multiply (4.18) by the left eigenvector $\left(v_f^T, v_g^T\right)$ corresponding to the zero eigenvalue of $\mathbf{A_{sys}}$. The first term of (4.18) vanishes and becomes:

$$\left(v_f^T, v_g^T\right)\begin{pmatrix} f_\lambda \\ g_\lambda \end{pmatrix}\frac{\partial \lambda}{\partial p} + \left(v_f^T, v_g^T\right)\begin{pmatrix} f_p \\ g_p \end{pmatrix} = 0 \tag{4.19}$$

Therefore, the sensitivity of λ with respect to parameter changes is:

$$\frac{\partial \lambda}{\partial p} = \frac{-\left(v_f^T, v_g^T\right)\begin{pmatrix} f_p \\ g_p \end{pmatrix}}{\left(v_f^T, v_g^T\right)\begin{pmatrix} f_\lambda \\ g_\lambda \end{pmatrix}} \tag{4.20}$$

To compute the above sensitivity, it is necessary to determine matrices of partial derivatives f_p, g_p and f_λ, g_λ. Given the loading scenario given by (2.4) and the generation scenario in (2.5), the load parameterization equations can be written in vector notation as follows:

$$\mathbf{s}(\lambda) = \mathbf{s_0} + \lambda \mathbf{k} \tag{4.21}$$

where $\mathbf{s}(\lambda)$, $\mathbf{s_0}$ and $\mathbf{k}$ are defined by:

$$\begin{aligned} \mathbf{s}(\lambda) &= [\ldots P_{D,i}(\lambda) \ \ldots \ Q_{D,i}(\lambda)]^T \\ \mathbf{s_0} &= [\ldots P_{D0,i} \ \ldots \ Q_{D0,i}]^T \\ \mathbf{k} &= [\ldots k_{P,i} \ \ldots \ k_{Q,i}]^T \end{aligned} \tag{4.22}$$

Using the above notations and the chain rule, the partial derivatives f_λ, g_λ can be written as:

$$\begin{pmatrix} f_\lambda \\ g_\lambda \end{pmatrix} = \begin{pmatrix} f_s \\ g_s \end{pmatrix} \mathbf{k} \tag{4.23}$$

As discussed in chapter static voltage stability analysis involves only algebraic equations of the DAE model. Therefore, the sensitivity formula in (4.20) can be rewritten as:

$$\frac{\partial \lambda}{\partial p} = \frac{-v_g^T g_p}{v_g^T g_s \mathbf{k}} \tag{4.24}$$

Various sensitivity matrices f_p, g_p with respect to parameter changes such as network admittance, governor, exciter and loading pattern are summarized in [1]. The sensitivities with respect to base case active and reactive power demands can be respectively shown as:

$$\begin{aligned} \frac{\partial \Delta P_i}{\partial P_{D0,i}} &= -(1 + \lambda k_{P,i}) \\ \frac{\partial \Delta Q_i}{\partial Q_{D0,i}} &= -(1 + \lambda k_{Q,i}) \end{aligned} \tag{4.25}$$

In modern power system operation, coordinated controls are often used to yield better performance index. Notice that the sensitivity in (4.24) is a first-order estimation. Therefore, it can be linearly superimposed when *k* parameters are varied. Therefore, the stability margin variation can be found from:

$$\Delta\lambda = \frac{\partial \lambda}{\partial p_1}\Delta p_1 + \frac{\partial \lambda}{\partial p_2}\Delta p_2 + \ldots + \frac{\partial \lambda}{\partial p_k}\Delta p_k \tag{4.26}$$

4.4 Simulation results

The IEEE 30 bus is the test system for all simulations presented in this chapter. There are 6 generation buses, 21 load buses, 4 transformers and 41 transmission lines in this test system. Shunt reactive power elements are connected to buses 15,16,17,18,20,22,23,25 and 30. The test system, parameters and initial bus data are given in [97]. The generator reactive power limits are listed in Table 4.1 whereby the limits of shunt reactive power sources are given in

Table 4.2. The single line diagram is shown in Fig. 4.3. The PSAT software package [98] is used for all power system simulations in this chapter. MATLAB neural network toolbox is applied for the neural network VSM estimation.

Table 4.1 Generator reactive power limits

Generator reactive power limit (p.u.)					
Bus no.	2	5	8	11	13
$q_{Gi}^{\min}$	-0.4	-0.4	-0.1	-0.06	-0.06
$q_{Gi}^{\max}$	0.5	0.4	0.4	0.24	0.24

Table 4.2 Shunt reactive power source limits

Reactive power source limits (p.u.)									
Bus no.	22	15	16	17	18	20	23	25	30
$q_i^{\min}$	-0.1	0	0	0	0	0	0	0	0
$q_i^{\max}$	0	0.2	0.15	0.1	0.2	0.1	0.3	0.25	0.2

4.4.1 Neural network approach

There are various types of neural networks appropriate for different applications. Among them, feed-forward neural network (FFNN) is a very suitable tool for function approximation problems. FFNN learns and gains the knowledge through the training session. Back propagation is a method to teach the FFNN to perform a given task. The weights are adjusted over iterations until the minimum error is obtained. A three-layer FFNN is developed for VSM estimation in this example. The augmented vector of input features and the corresponding target value for an operating point (pattern) *i* is shown as:

$$\mathbf{x}_i = \left[sf_{1,i}, sf_{2,i}, ..., sf_{nl,i}, u_{1,i}, u_{2,i}, ..., u_{nb,i}, \lambda_i \right] , \forall i = 1, 2, ..., n \tag{4.27}$$

where $sf_{k,i}$ is the MVA power flow over line k for pattern i; $u_{j,i}$ is the voltage at bus j for pattern i; nl is the number of lines; nb is the total number of buses excluding the slack ;λ_i is

the VSM of pattern *i* and the target of FFNN estimation and *n* is the total number of input patterns. This is a convenient and reasonable choice of system variables for VSM approximation because bus voltage magnitude can reasonably describe the influences of changing power demand levels and variation of control variable settings. Line flow is also a nonlinear function of bus voltages and line admittances. Therefore, this quantity should be able to reflect network topological changes.

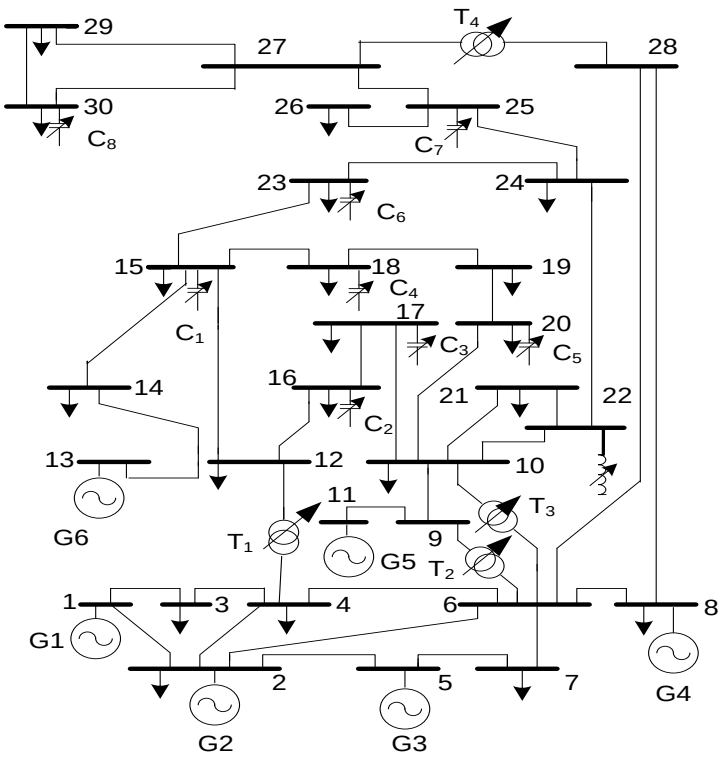


Fig. 4.3 Single line diagram of IEEE 30-bus test system

Principal component analysis is applied to extract the useful input features and improve the learning capability of FFNN. The size of the original feature according to (4.27) is 70 (41+29). The features whose contribution to the total variation (represented by eigenvalues) is less than 0.001 are removed. At the end of this process, the number of transformed features is 29. The number of hidden neurons is set to 20. The training session uses 15,000 operating conditions generated offline based on the procedures discussed in section 4.2.3.2.

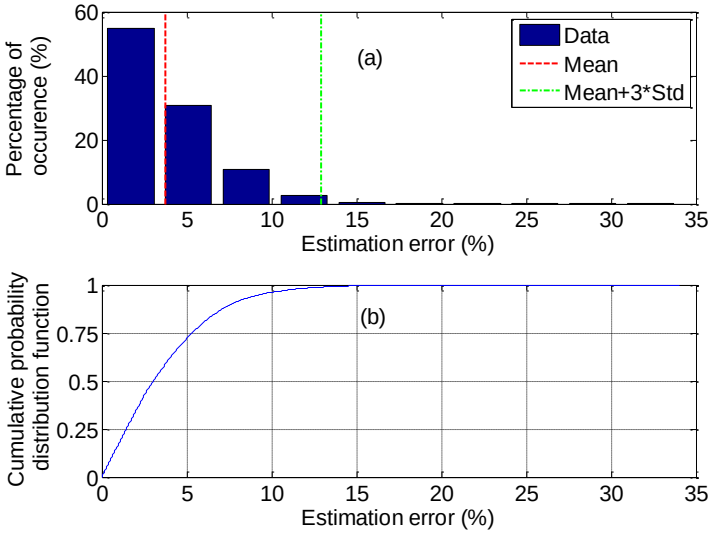


Fig. 4.4 Statistics of FFNN testing (a) histogram of 5000 testing conditions (b) cumulative probability distribution of estimation error

After the proper training, VSMs of new 5000 operating conditionswhich have never been presented during the training phase were estimated by the FFNN. The estimation error is defined by the percentage of absolute difference between estimated and actual values relative to the actual one. Histogram showing the distribution of estimation error is shown in Fig. 4.4(a). Mean and standard deviation of estimation errors are computed. Vertical two lines in Fig. 4.4(a) indicate the mean and three times the standard deviation (3σ) based on the 5000 testing samples. If estimation errors are assumed to be normally distributed, the latter quantity can be set as the threshold value for an outlier analysis. A sample with an estimation error greater than 3σ is labeled an outlier.

From the statistical analysis, the average error is 3.72%., the standard deviation is 3.07% and the maximum error is 34.02%. According to the criterion for an outlier discussed above, there are 59 outliers accounting for 1.18% (59/5000) of the entire sample group. The cumulative probability distribution function (CPDF) of estimation errors is plotted as shown in Fig. 4.4(b). It can be observed that accurate results can be obtained with reasonably high probability. For example, the probability of estimation error of FFNN is less than or equal to 5% is 0.7171. The large estimation errors may result from very versatile operating conditions for the traning set.

An additional test for VSM estimation by FFNN has been also carried out. This experiment aims at verifying if the developed FFNN still performs well under the condition of fixed power demand level but varying control variable settings. This is precisely the condition that FFNN will have to perform when incorporated in the ACO process. An hourly demand profile for a day was randomly generated. The demand and generation profiles at each hour are fixed. At every time step, 100 different samples are generated based on the rule that only control variables are allowed to be changed within the respective limits. The corresponding VSM is calculated for each operating conditions. Estimation errors as defined earlier are calculated and box and whisker plots at each hour are shown in Fig. 4.5. For each plot, the box has lines at lower quartile, median and upper quartile. The extreme values are extended by the whiskers within the range of 1.5 times of the interquartile range. The values beyond this range are considered outliers. It is shown that the median of errors throughout the day is less than 8%. Although there are outliers at each hour, the number of those points is relatively small.

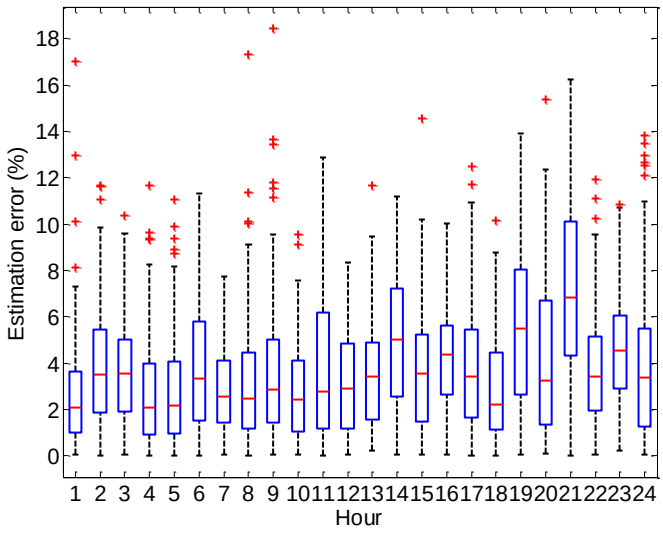


Fig. 4.5 Box-plot of testing for variation of control settings

4.4.2 Sensitivity approach

The sensitivity approach was used to approximate VSM in an optimization task for under-voltage load shedding which was developed in this thesis (see Chapter 6 for more details).

The new value of VSM denoted by λ is determined by adding the influence of parameter changes to the old value of VSM given by λ_0 according to:

$$\lambda = \lambda_0 + \sum_{i=1}^{N} \frac{\partial \lambda}{\partial p_{di}} \Delta p_{di} + \sum_{i=1}^{N} \frac{\partial \lambda}{\partial q_{di}} \Delta q_{di} \tag{4.28}$$

where $\frac{\partial \lambda}{\partial p_{di}}$ and $\frac{\partial \lambda}{\partial q_{di}}$ are sensitivities of VSM with respect to the active and reactive load shedding at bus *i*, respectively (calculated by (4.24)); Δp_{di} and Δq_{di} are the total amount of active and reactive power curtailed at bus *i*.

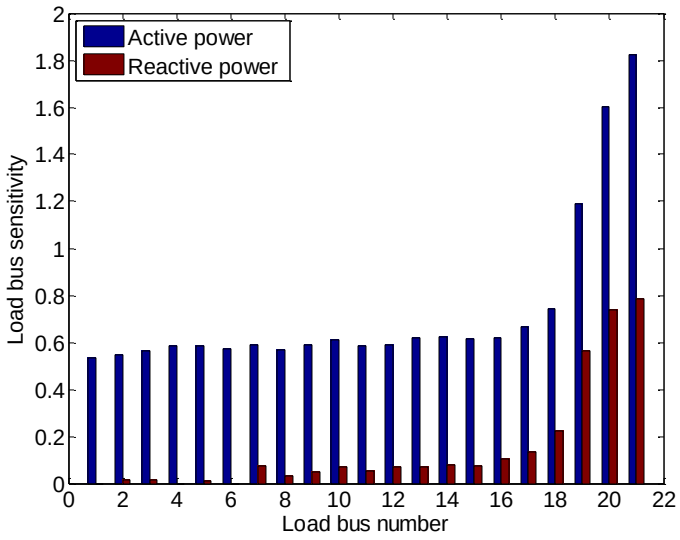


Fig. 4.6 Sensitivity of load buses

As mentioned earlier that the IEEE 30-bus test system has 21 load buses, the two aforementioned sensitivities calculated at each load bus calculated at an operating point close to the voltage collapse (saddle node bifurcation) point are depicted as in Fig. 4.6. It is obvious that buses with relatively high sensitivity are effective in enhancing voltage stability of the power system. Therefore, five buses with the highest sensitivities (load buses number 17-21 corresponding to buses 23, 24,26,29 and 30, respectively) should be considered for the load shedding program. Pre-screening load buses is very advantageous because in operational viewpoints load reduction at these effective buses will better improve the stability margin. From the computational viewpoint, this process can reduce the number of decision variables

of the load shedding problem. Two examples showing the results of VSM approximation are extracted from that work and presented in this chapter.

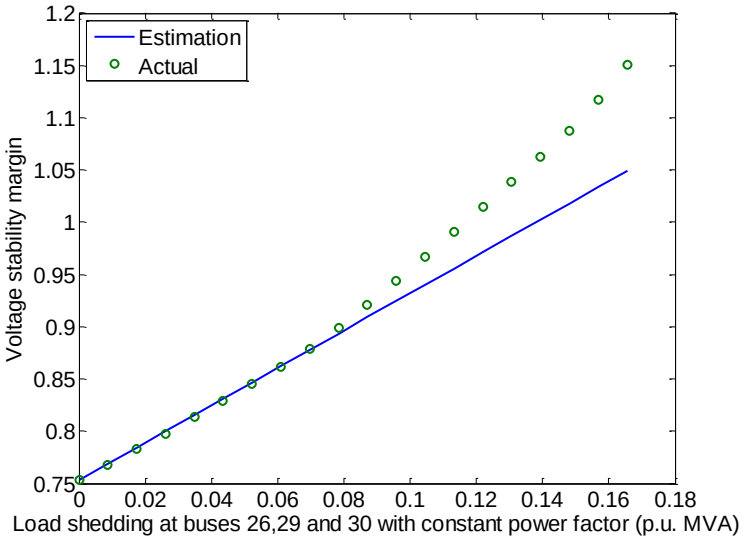


Fig. 4.7 VSM versus load shedding

It is observed from Fig. 4.6 that the last three load buses (numbers 19-21) representing buses 26, 29 and 30, respectively have relatively the greatest sensitivities. The major reason for these critical locations is that these buses are geographically located far from the generating units and lacks of reactive power supports. If the loads at these locations are shed, the power transfer margin will consequently be increased. The stability margin increases almost linearly in the operating region close to the base condition as shown in Fig. 4.7. However, when more loads are shed, the margin deviates from the linear relationship and show more nonlinearity. The linear estimate becomes more erroneous in this situation. The accuracy of estimation could be improved by regularly updating the sensitivity values.

The second example simulates the situation in which the sensitivity method is embedded in optimization of the load shedding problem. The amount of load shed at buses 23, 24,26,29 and 30 is randomly chosen between minimum and maximum allowable limits $\Delta p_{di,min}$ and $\Delta p_{di,max}$ as shown in Table 6.11 of Chapter 6, respectively. The estimated and actual values of VSM are plotted in Fig. 4.8(a) whereby the corresponding estimation errors are given in Fig. 4.8(b). It is revealed that the sensitivity method is capable of giving the accurate estimate of VSM which is one of the important constraints of the optimal load shedding problem.

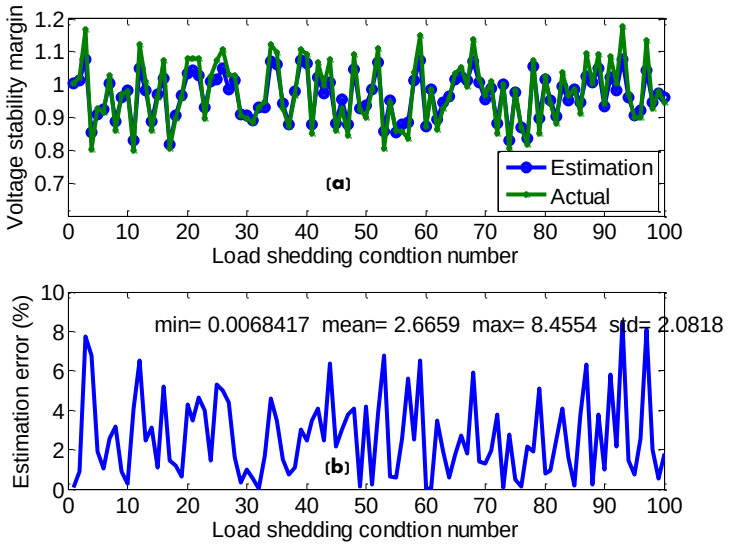


Fig. 4.8 Testing of VSM estimation by sensitivity method

4.5 Summary

This chapter presents two techniques to estimate voltage stability margin (VSM) of the power system namely the neural network approach and the sensitivity approach. A new method for generating realistic operating conditions is proposed. Correlation of the power demand among different locations is modeled by the multivariate normal distribution. Control variables in the power system are assumed to statistically independent and modeled by the uniform distribution within their respective operating limits. In the first method, a feed-forward neural network was developed to learn the relationship between the selected set of input features and the corresponding VSM calculated by the continuation power flow (CPF) offline. The influence of changes in control parameters to VSM can be approximated by the sensitivity formula as shown in the second approach. Accurate estimation of VSM has been observed for the operating conditions close to the base loading point. The greater degree of nonlinearity has been observed for the operating point far away from the base one.

Chapter 5
Voltage Stability Constrained Optimal Power Flow

5.1 Introduction

One of the important tasks in operating an interconnected power system is to schedule the active and reactive power outputs of each power plant in such a way that the operating cost is minimized and constraints are within the respective limits. These limits consist of the capacity of equipments such as generators and compensating devices, the security limits and the stability limits. This problem is generally known as the optimal power flow (OPF). The set of optimal decision variables is sought out by solving a constrained optimization problem. The objective of this problem can be economic costs, system security or any other objectives. In general, the system quantities that must be maintained within their limits are defined as the dependent variables, such as load bus voltages and generator reactive power outputs. Since the early stage [99], many researchers have studied the OPF with different objectives and various methods including computational intelligence [100].

Conventionally, optimal generation scheduling was carried out by a simplified OPF problem namely economic dispatch (ED). Active power output of generating units is the only decision variable. Power transmission losses are approximated by loss formulas. In the past few decades, researchers have immensely investigated several techniques to solve ED [101, 102]. However, the significant shortcoming of ED is inability to capture constraints of dependent variables due to simplified power flow equations. Thanks to fast processing speed of computers, OPF is a standard practice to determine optimal control variables from global criterion.

Besides general objective functions of OPF in planning and operating stages, the current practice of operators for testing stability of the power system is to adopt the optimal solution obtained from the OPF and evaluate the system stability performance for all credible contingencies. If the system is unstable, the OPF solution is modified by heuristic rules derived from engineering experience and judgment to achieve minimum rescheduling costs. This process is usually time consuming and also may sacrifice the optimality. Therefore, the

stability limit should be properly incorporated in the automated OPF software [103, 104]. Moreover, generation scheduling is one of the effective control measures for improving different stability problems [105]. Therefore, voltage stability constrained OPF (VSCOPF) is presented in this chapter. The objective of VSCOPF is to minimize instantaneous operating costs while maintaining security and voltage stability constraints within permissible ranges. Practical characteristics of thermal generating units were also taken into consideration. Given those non-convex and discontinuous cost characteristics, differential evolution (DE) algorithms discussed in Chapter 3 were applied as optimization tools.

The rest of this chapter is organized as follows. The general OPF formulation is given in Section 5.2. Constraint handling methods are summarized in Section 5.3. In Section 5.4, the objective and constraints of VSCOPF are listed. Simulation results obtained from different DE algorithms are compared. The chapter is summarized in Section 5.5.

5.2 Optimal power flow

OPF is usually conducted at each operation planning period (i.e. 15 minutes). This OPF solution is optimal for that particular time instant. However, inter-temporal constraints such as ramp-rate limits exist in thermal generating units. To capture these restrictions, dynamic OPF (DOPF) such as [106] should be carried out. Given power of computing facilities, it becomes more promising to handle DOPF within the allowable planning period. Since this thesis focuses on the developed EA method, this chapter will not deal with DOPF. The optimal solution for a planning period t is determined.

Two cases can be treated in OPF: optimal active and reactive power flow (objective = instantaneous operating costs; solution = exact economic dispatch) and optimal reactive power flow (objective = total system losses; solution = minimum losses) [99]. Formerly, these two problems were analyzed separately. This chapter introduces an alternative formulation for VSCOPF that can combine the two optimization tasks and solved in a single run. Control variables for both active and reactive dispatch problems are considered. In this thesis, the term OPF refers to exact economic dispatch with total fuel cost minimization as the objective whereas optimal reactive power dispatch (ORPD) means the reactive and voltage control problem with loss minimization as the objective. The latter problem will be

investigated in Chapter 6. Several models are proposed to speed up the computing process in order to suit online applications.

5.2.1 Problem formulation

OPF is essentially a nonlinear constrained optimization. Following the framework of nonlinear programming, OPF can be mathematically formulated with the objective function and constraints as discussed below.

5.2.1.1. Objective functions

The objective function in conventional vertically integrated system is to minimize the total operating cost of generation by assuming that the associated generating units belong to a single entity. On the other hand, the objective function is changed to maximize the social wealthfare in the deregulated environment in which there exist many entities competing in business. This chapter follows the first paradigm. Since operating costs mainly associated with fuel costs, the objective of OPF here is to minimize the total fuel cost of power generation as follows:

$$\min \sum_{i=1}^{n} FC_i(P_i) \tag{5.1}$$

where FC_i is the fuel cost function of generator i; n is the number of generators, P_i is the power output of generator i. The simplified fuel cost function of each generator can be represented as:

$$FC(P_i) = a_i P_i^2 + b_i P_i + c_i \tag{5.2}$$

where a_i, b_i , c_i are cost coefficients of generator i, obtained by heat-characteristic tests.

The valve opening process of multivalve steam turbines produces ripples in the generator heat rate characteristic leading to the higher degree of nonlinearity in the fuel cost function. To realistically capture this phenomenon, the cost function (5.2) should be replaced by the following function to consider the valve-point effect [107]:

$$FC_i(P_i) = a_i P_i^2 + b_i P_i + c_i + \left| e_i \sin(f_i (P_i^{\min} - P_i)) \right| \tag{5.3}$$

where e_i and f_i are cost coefficients of generator i reflecting the valve-point effect. $P_i^{\min}$ is the minimum capacity of generator i. Fig. 5.1 shows the cost characteristic of four generators with the valve-point effect. The cost coefficients of these generators are taken from [108].

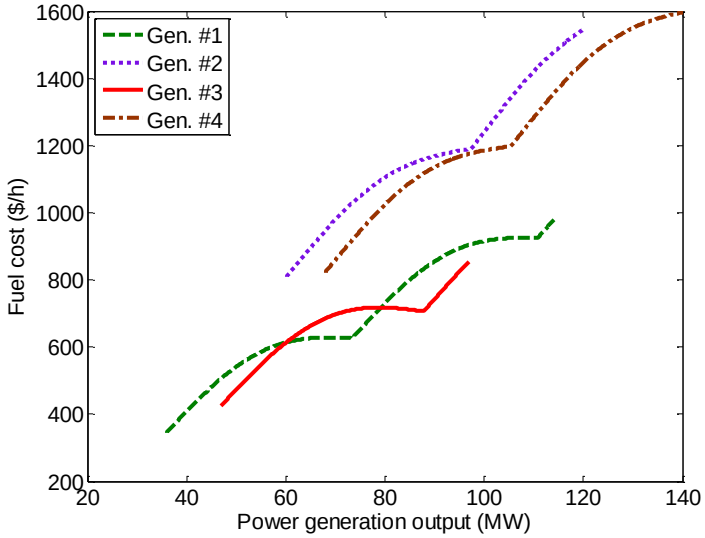


Fig. 5.1 Cost characteristics considering the valve-point effect

Because the generating units can be supplied by multiple-fuel sources such as coal, natural gas, or oil, a practical ED has to determine both the most economic fuel type and generation output [109]. The fuel cost function of a generating unit with such characteristics can be modeled as several piecewise cost function. With the valve-point effect, a cost function of unit i can be described as:

$$FC_i(P_i) = \begin{cases} FC_{i1}(P_i) & \text{fuel 1} & P_i^{\min} \le P_i \le P_{i1}^{\min} \\ FC_{i2}(P_i) & \text{fuel 2} & P_{i2}^{\min} \le P_i \le P_{i2}^{\max} \\ \vdots & \vdots & \vdots \\ FC_{ik}(P_i) & \text{fuel } k & P_{ik}^{\min} \le P_i \le P_i^{\max} \end{cases} \tag{5.4}$$

where the fuel cost function of generator i for fuel type k is defined by:

$$FC_{ik}(P_i) = a_{ik}P_i^2 + b_{ik}P_i + c_{ik} + \left| e_{ik} \sin(f_{ik}(P_{ik}^{\min} - P_i)) \right| \tag{5.5}$$

where $a_{ik}, b_{ik}, c_{ik}, e_{ik}$ and f_{ik} are the cost coefficients of generator i for fuel type k, $P_{ik}^{\min}$ is the minimum power output of generator i using fuel type k. Fig. 5.2 shows the cost characteristic of three generators with multi-fuel options. The cost coefficients of these generators are taken from [109].

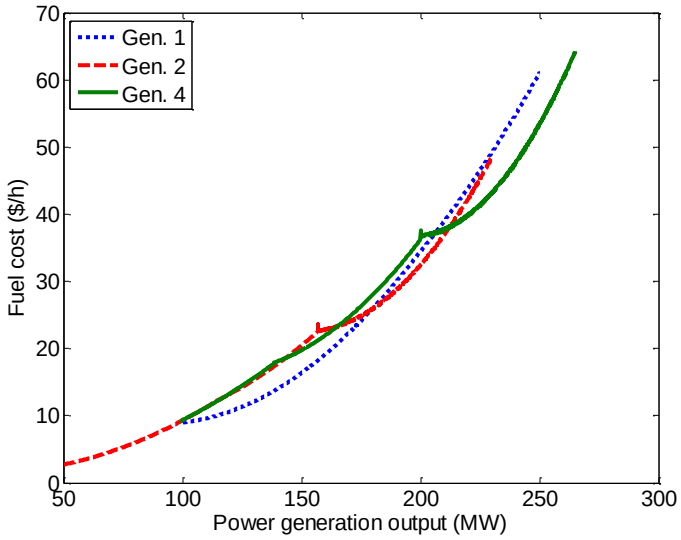


Fig. 5.2 Cost characteristics considering multi-fuel options

5.2.1.2. Constraints

Typically, power flow equations are used to represent active and reactive power in OPF rather than an approximated formula in ED. This allows OPF to consider security constraints including voltage levels and transmission power flows. Moreover, active power restrictions of each generating unit have to be incorporated as inequality constraints of OPF.

- **Power flow equations**

Given the complex bus voltage at bus i written by $U_i e^{j\theta_i}$, the power balance equations which include the active and reactive power balance equations for each load bus are described as follows:

$$0 = P_{Gi} - P_{Di} - U_i \sum_{j=1}^{N} U_j Y_{ij} \cos(\theta_i - \theta_j - \delta_{ij}) \tag{5.6}$$

$$0 = Q_{Gi} - Q_{Di} - U_i \sum_{j=1}^{N} U_j Y_{ij} \sin(\theta_i - \theta_j - \delta_{ij}) \tag{5.7}$$

where P_{Gi} and Q_{Gi} are active and reactive power generation at bus i, respectively; P_{Di} and Q_{Di} are active and reactive power load at bus i, respectively; Y_{ij} is the admittance matrix corresponding to the ith row jth column and δ_{ij} is the phase angle of the admittance between the ith and jth buses; N is the total number of buses.

- **Ramp-rate limits**

The actual range of all online units during each dispatch period t is restricted by the corresponding ramp rate limits according to [110]:

$$P_{i,t}^{\min} \le P_i \le P_{i,t}^{\max} \tag{5.8}$$

where $P_{i,t}^{\min}$ and $P_{i,t}^{\max}$ are minimum and maximum generation limits at dispatch period t determined by:

$$P_{i,t}^{\min} = \max\{P_i^{\min}, P_i^0 - DR_i\} \tag{5.9}$$

$$P_{i,t}^{\max} = \min\{P_i^{\max}, P_i^0 + UR_i\} \tag{5.10}$$

where DR_i and UR_i are down ramp and up ramp limits of generator i, respectively. $P_i^{\max}$ is maximum capacity of generator i.

- **Prohibited operating zones**

In some cases, a thermal unit cannot be operated in the entire capacity range due to some physical limitations. This is because of faults in the electrical machine itself or in the associated auxiliaries units such as boilers, feed pumps, etc. Furthermore, a typical thermal unit may have a steam valve in operation, or a vibration in a shaft bearing which may result

in interference [111]. Therefore, for units with prohibited zones additional constraints have to be satisfied:

$$P_i \in \begin{cases} P_{i,t}^{\min} \le P_i \le PZ_{i,1}^{\mathrm{l}} & \\ PZ_{i,k-1}^{\mathrm{u}} \le P_i \le P_{i,k}^{\mathrm{l}} & k = 2,3,...,npz_i \\ PZ_{i,pz_i}^{\mathrm{u}} \le P_i \le P_{i,t}^{\max} & i \in \mathbf{vpz} \end{cases} \tag{5.11}$$

where $PZ_{i,k}^{\mathrm{l}}$ and $PZ_{i,k}^{\mathrm{u}}$ are lower and upper bounds of prohibited zone k of unit i, respectively , npz_i is the number of prohibited zones of unit i: **vpz** is the set of units with prohibited operating zones. Fig. 5.3 shows the cost characteristics of two generators with prohibited operating zones. The cost coefficients of these generators are taken from [112].

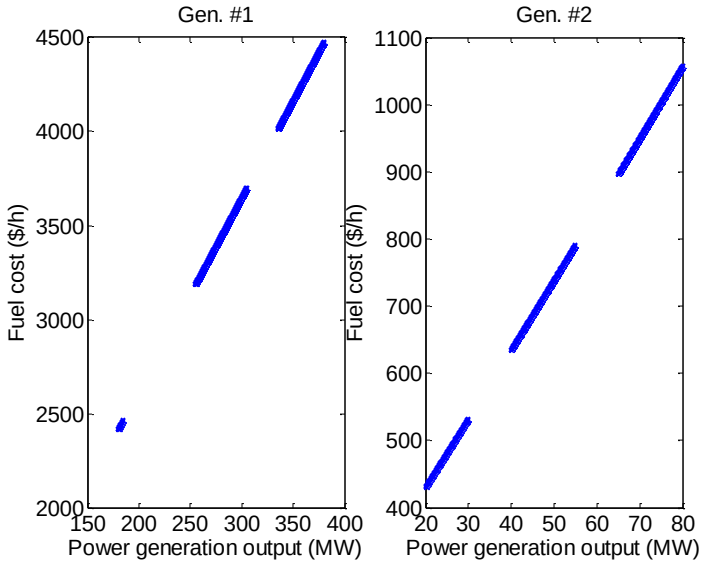


Fig. 5.3 Cost characteristics with prohibited operating zones

5.3 Constraint handling methods

5.3.1 Treatment of inequality constraints

The power generation outputs have to be within its feasible operating boundaries. This constraint becomes even harder to maintain in the case of generators with prohibited operating zones because of the existence disjoint search spaces. Therefore, for all individuals in the population after crossover every element of power generation vector is checked and adjusted according to:

$$P_i = \begin{cases} P_i & \text{if } P_{i,t}^{\min} \le P_i \le P_{i,t}^{\max} \\ P_{i,t}^{\min} + rand \cdot (P_{i,t}^{\max} - P_{i,t}^{\min}) & \text{otherwise} \end{cases} \quad (5.12)$$

$$\forall i \in \{1,2,\ldots,n\}$$

Note that the minimum and maximum power generations should be adjusted to consider the ramp-rate constraints given by (5.9)-(5.10). For the units with prohibited operating zones, the power generation outputs should not be inside any of their respective prohibited zones (PZ). Therefore, if any generator *j* with PZ whose power output lies in the prohibited zone *k* the corresponding power output should be adjusted according to:

$$P_j = \begin{cases} PZ_{j,k-1}^{\mathrm{u}} + rand \cdot (PZ_{j,k}^{\mathrm{l}} - PZ_{j,k-1}^{\mathrm{u}}) & \text{if } d_{\mathrm{l}} < d_{\mathrm{u}} \\ PZ_{j,k}^{\mathrm{u}} + rand \cdot (PZ_{j,k+1}^{\mathrm{l}} - PZ_{j,k}^{\mathrm{u}}) & \text{otherwise} \end{cases} \quad (5.13)$$

$$d_{\mathrm{l}} = \left| P_j - PZ_{j,k}^{\mathrm{l}} \right| \quad (5.14)$$

$$d_{\mathrm{u}} = \left| P_j - PZ_{j,k}^{\mathrm{u}} \right| \quad (5.15)$$

$$\forall j \in \mathbf{vpz}$$

where d_{l} and d_{u} are the absolute distances from the power generation inside the prohibited zone *k* to the corresponding lower and upper bounds, respectively.

5.3.2 Fitness function

The adaptive penalized fitness function discussed in Section 3.4.3.1 is used to treat state (dependent) variables of OPF.

5.4 Simulation results

With the serious threat of voltage collapse in modern power systems as discussed in the previous chapters, the system stability should be evaluated in the generation dispatch problem. Given non-convex and non-continuous characteristics of generating units, evolutionary computation is a highly potential optimization engine.

This example presents an example of the voltage stability constrained optimal power flow (VSCOPF) problem solved by different variants of differential evolution (DE). Simulation was done on the IEEE-30 bus test system with six generating units.

The objective of VSCOPF is to minimize the total fuel cost consisting of individual non-smooth cost functions by regulating the active power outputs of all committed generators while maintaining operation, security and stability constraints. The VSCOPF problem is therefore expressed as follows:

$$\textbf{Minimize } F_T = \sum_{i=1}^{n} HRC_i \cdot HR_i(P_i)$$

Subject to:

1. Prohibited operating zone limits

   See (5.11)

2. Generator reactive power (Q_G) limits

   $Q_{Gi}^{\min} \le Q_{Gi} \le Q_{Gi}^{\max}$ for all generators

3. Transformer tap position (a) limits

   $a_i^{\min} \le a_i \le a_i^{\max}$ for all transformers

4. Bus voltage (U) limits

   $U_i^{\min} \le U_i \le U_i^{\max}$ for all buses

5. Line flow (S) limits

   $\left|S_i\right| \le \left|S_i^{\max}\right|$ for all transmission lines

6. Voltage stability margin (λ) limit

   $\lambda \ge \lambda^{\min}$

(5.16)

where F_T is the total fuel cost at the considered period. In this formulation, the reactive power output of all generators and transformer tap positions are considered as the decision variables in addition to the active power output. Generators are modeled as load buses with negative power generation. Therefore, generator bus voltages become dependent variables in addition to the magnitude of apparent power flows across transmission lines, load bus voltages and the voltage stability margin of the entire system. In summary, the limits of the decision variables constitute items 1 to 3 of (5.16) whereby items 4 to 6 of the same equation are the constraints of the dependent variables. Voltage stability margin is calculated by the direct method. PSAT software package [98] is used as the simulation tool. It should be noted that the proposed formulation allows us to analyze active and reactive power dispatch

problems at the same time by adding the cost of energy losses to the objective function in (5.16).

The valve point effect was considered in the fuel cost characteristic as a recurring rectified sinusoidal contribution to the quadratic function for every generator. Therefore, the heat consumption of the generating unit i in terms of MBtu per per-unit of power generated is shown as follows:

$$HR_i(P_i) = ah_i P_i^2 + bh_i P_i + ch_i + \left| eh_i \sin(fh_i(P_i^{\min} - P_i)) \right| \tag{5.17}$$

where ah_i, bh_i, ch_i, eh_i and fh_i are coefficients of the heat rate characteristic of the generator i. Therefore, the fuel cost characteristic can be shown as:

$$FC_i(P_i) = HRC_i \cdot HR_i(P_i) \tag{5.18}$$

where HRCi is the fuel cost for generating a unit of heat in $/MBtu. The parameters of the heat rate characteristic of each unit are given in Table 5.1. Every generator is assumed to have prohibited operating zones whose intervals are listed in Table 5.2 along with the operating boundary.

Table 5.1 Heat characteristics of generators

Coefficients	Generating unit					
	1	2	3	4	5	6
ah_i (MBtu)	0	0	0	0	0	0
bh_i (Mbtu/puW)	200	175	100	325	300	300
ch_i (Mbtu/puW^2)	37.5	175	625	83.4	250	250
eh_i (MBtu)	15	10	10	5	5	5
fh_i (rad/puW)	6.283	8.976	14.784	20.944	25.133	18.48
HRC_i ($/MBtu)	10	10	20	20	30	30

Table 5.2 Generator limits and boundaries of prohibited operating zones

Unit	$[P_i^{min},P_i^{max}]$	Prohibited operating zone		
		1	2	3
1	[0.5,2.5]	[1,1.25]	[1.75,2]	[2.1,2,25]
2	[0.2,1.6]	[0.35,0.4]	[0.5,0.75]	[1.1,1.25]
3	[0.15,1]	[0.3,0.5]	[0.7,0.85]	-
4	[0.1,0.7]	[0.2,0.35]	[0.5,0.6]	-
5	[0.1,0.6]	[0.2,0.35]	[0.5,0.6]	-
6	[0.12,0.8]	[0.3,0.4]	[0.5,0.6]	

$[P_i^{min},P_i^{max}]$: Min. and Max. limit of generator *i*

To verify the performance of the JADE-vPS developed in this thesis in handling the VSCOPF problem, different DE algorithms presented in Chapter 3 were applied as the reference algorithms in order to compare the results. Statistical values computed from ten independent runs are given in Table 5.3. The testing algorithms consist of:

1. DE1 is the classical DE with the "DE/rand/1" mutation strategy (see (3.16));
2. DE2 is the classical DE with the "DE/curren-to-best/1" mutation strategy (see (3.17));
3. JADE-PS*x* is the adaptive JADE with the population size *x* and the external solution archive;
4. JADE-vPS is the modified variant of JADE with the variable population size and the external solution archive;
5. SaDE is the self-adaptive DE proposed in [62].

All testing algorithms terminate when the number of function (power flow) evaluations reaches 5000. The statistical values listed in Table 5.3 clearly show that the proposed JADE-vPS in comparison to the other algorithms is very robust to the random search nature. Its standard deviation and maximum of the final results are the least. The principal difference between the proposed JADE-vPS and the original JADE is the population size. As mentioned in Chapter 3 , the population size of JADE-vPS is bounded between 0.5 and 2 times of the initial population size. In this example, the initial population size PS^{ini} is set to 50 for all algorithms. To fairly compare the two algorithms, the population size of the classical JADE is

set to 25 ($0.5 \cdot PS^{ini}$) and 100 ($2 \cdot PS^{ini}$). For this particular problem, it is quite surprising to see that the smaller population size results in the better search performance as reflected in the least minimum and average costs among the JADE with 25, 50 and 100 population size.

Table 5.3 Statistical results of different algorithms from 10 independent trials

Algorithm	Statistical values			
	Min.	Avg.	Max.	Std.
DE1	6111.03	8425.89	11329.32	1590.57
DE2	6206.01	6516.50	6874.96	185.58
JADE-PS25	**5957.69**	**6058.50**	6424.77	135.24
JADE-PS50	6025.43	6156.95	6263.32	76.27
JADE-PS100	6541.07	6829.51	7025.58	176.94
JADE-vPS	6029.66	6110.61	**6167.73**	**48.41**
SaDE	6425.22	7089.61	8192.98	697.57

Min.: Minimum Avg.: Average
Max: Maximum Std.: Standard deviation

The average convergence characteristics among all algorithms as shown in Fig. 5.4 confirm this observation. It is obvious that the JADE-PS25 converges at the fastest rate and arriving to the least average cost. The proposed JADE-vPS performs nearly identical to the JADE-PS50. On the other hand, the convergence of the JADE-PS100 is far behind its counterparts and even worse than the DE2 where there is no parameter adaptation involved. The algorithms with the worst convergence property are DE1 and SaDE.

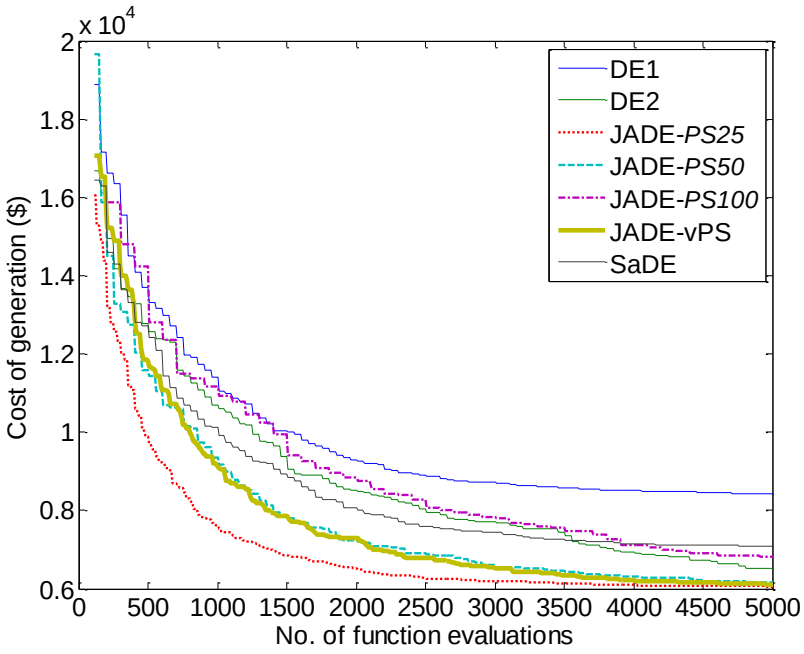


Fig. 5.4 Convergence characteristics of all algorithms

During the optimization course, the decision variables corresponding to the global best individual are recorded. Because the prohibited operating zones are considered for all generating units, any feasible generation schedule should not lie in one of these zones. Based on the best result obtained by JADE-vPS, Fig. 5.5 shows the variation of active power output P_G of two generators (no. 2 and 5) versus the number of function evaluations. For both generators, the method for handling the prohibited operating zone constraint in (5.13) works well as there is no P_G falling inside any prohibited zone. Because the active power output is directly related to the objective function, therefore P_G changes very rapidly at the beginning of the search and settles to an optimal or suboptimal value when the algorithm converges.

The transformer tap position a and the generator reactive power output Q_G are also decision variables in the VSCOPF problem. Even these variables do not directly influence the objective function, they are very crucial for regulating the security and stability constraints. All transformers are assumed to be equipped with the on-load tap changer (OLTC) with 13 steps (±6% of the nominal). The variation of a of two selected transformers is shown in Fig. 5.6(a) whereby the one of Q_G of the generators 2 and 5 is depicted in Fig. 5.6(b). The latter is treated as of the continuous characteristic. The response of voltage stability margin λ during the optimization process is shown in Fig. 5.7. It can be observed that the power system has a

relatively larger stability margin at the beginning of the optimization process. However, it reduces in response to the minimum cost measure. Therefore, the stability margin should be appropriately included in any OPF problems.

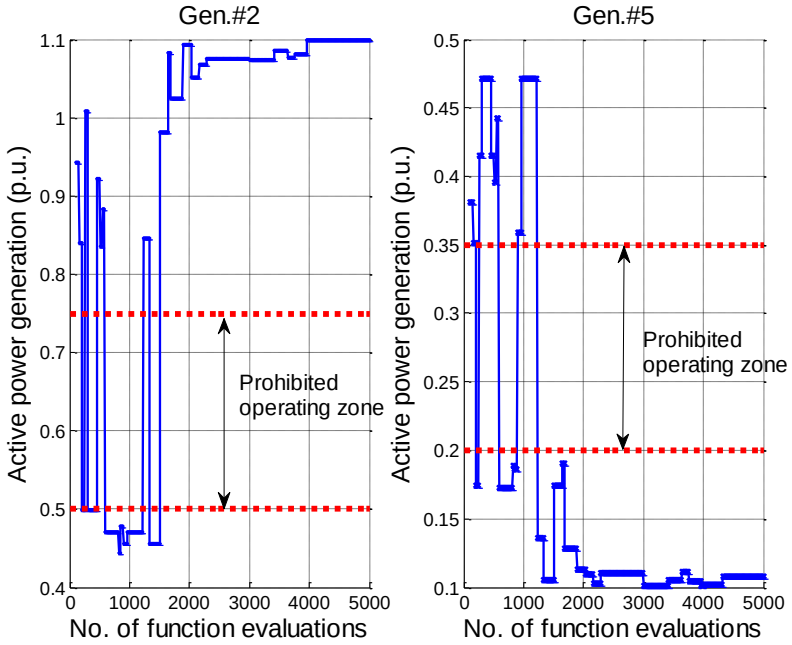


Fig. 5.5 Variation of active power output of two generators

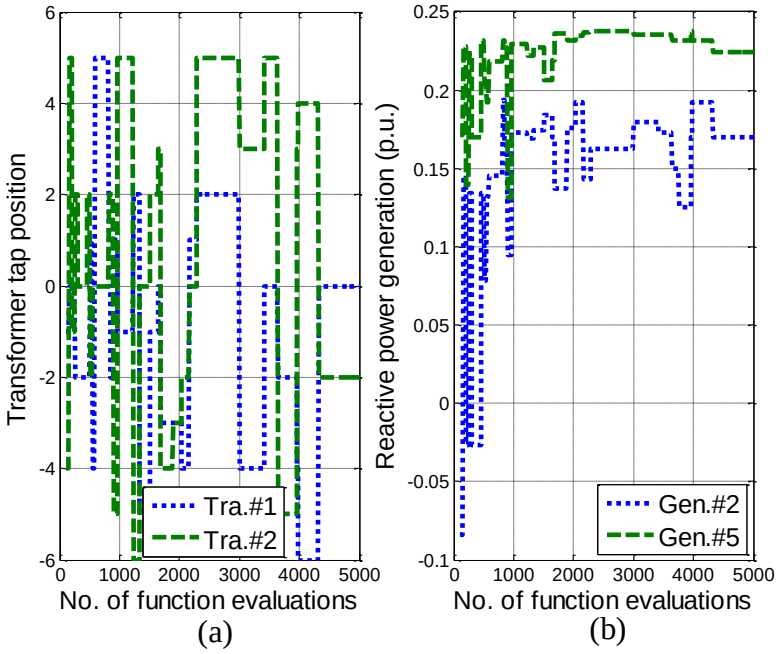


Fig. 5.6 Variation of (a) transformer tap positions (b) reactive power outputs

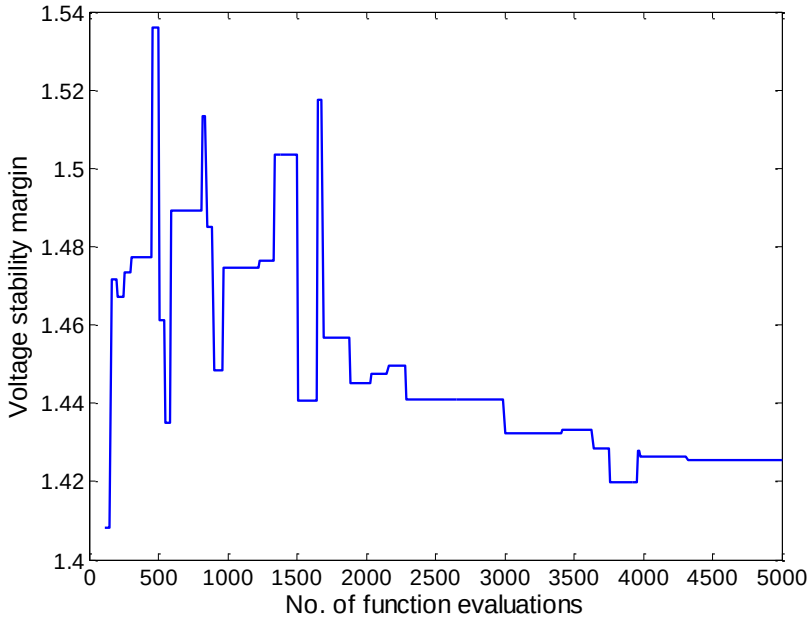


Fig. 5.7 Variation of voltage stability margin

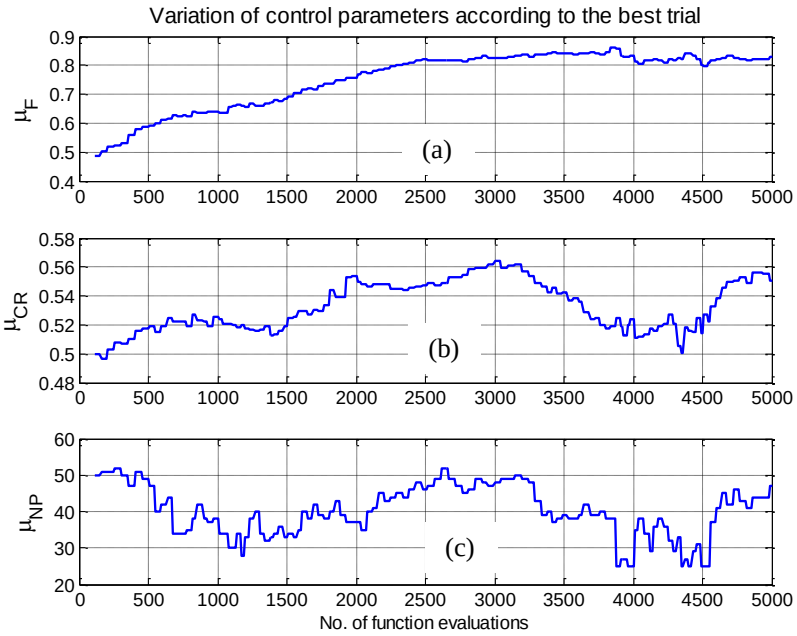


Fig. 5.8 Variation of JADE-vPS control parameters (from the best trial)

The key reason of the JADE-vPS success is the parameter adaptation scheme. The evolution of μ_F, μ_{CR} and μ_{NP} (see section 3.4.2.3 for explanations) corresponding to the global

best individual is plotted in Fig. 5.8. The CPU time for all testing algorithms is nearly identical because the most computationally demanding part is power-flow and stability margin calculations. The average CPU time is approximately 3 minutes on the computing platform: Pentium Core 2 Quad with CPU 2.83 GHz and RAM 3.25 GB. For large and practical power systems, this computing duration is not feasible. The next chapter will propose alterative solutions to reduce the CPU time by employing fitness approximation techniques.

5.5 Summary

Applications of evolutionary algorithms to optimal generation scheduling problems were presented in this chapter to meet the stressed operation of modern power systems. An example of VSCOPF with practical non-convex and discontinuous cost characteristics was developed in this chapter. Various DE algorithms were applied to solve the problem. The JADE-vPS algorithm demonstrates the outstanding robustness and effectiveness compared to the others.

Chapter 6
Countermeasures against voltage instability

6.1 Introduction

In the previous chapters, voltage stability and several methods of analysis were reviewed. Two techniques namely for estimating voltage stability margin (VSM) were also developed. VSM is one of the important performance indices for the system operator. Besides that, proper control coordination of various reactive power sources is among operational routines of the system operator. This process is not only to improve the steady state operating condition of a power system but also to ensure more voltage stability.

Because one of the major causes of voltage collapse is associated with reactive power, therefore voltage stability criterion should be addressed in reactive power and voltage control. The optimal power flow (OPF) dealing with optimization of reactive power sources is alternatively known as the optimal reactive power dispatch (ORPD). Control variables in ORPD normally consist of transformer tap positions, generator set points (either reactive power injection or voltage), and reactive power compensations. As a countermeasure to prevent a power system from voltage collapse, VSM should be incorporated in ORPD. This chapter primarily deals with various voltage constrained optimal reactive power dispatch (VSCORPD) problems.

ORPD is a mixed-integer nonlinear programming (MINLP) problem with continuous and discrete decision variables. Several methods have been applied to solve different ORPD problems. However, they are generally not efficient in dealing with large-scale and non-convex optimization problems. In contrast to mathematical programming, evolutionary computing has no restriction on cost characteristics. Moreover, compared to conventional gradient-based methods, evolutionary algorithms (EAs) are much simpler to be implemented and barely needs significant changes when dealing with different optimization tasks. Several EAs have been successfully implemented in various power system problems. This chapter presents the well-known genetic algorithm (GA) and the recent variant of ant colony optimization (ACO) applied to ORPD problems. Since ORPD decides the appropriate control

settings for an operating state of the next planning step, it can be considered as a preventive control scheme. In some situations, adjusting reactive power sources may not be adequate in providing a sufficient margin to collapse. Load shedding will be conducted generally as the last resort in counteracting the problem or steering the system away from the unstable operating zone to the stable one [30, 113-115]. In this sense, load shedding can be deemed a corrective control measure. ACO was applied to an under-voltage load shedding problem formulated by taking both technical and economic aspects into account.

In the EC paradigm, fitness of every individual in the population has to be evaluated in each generation. Therefore, a large number of fitness evaluations are required. In many applications, each single fitness evaluation is very time consuming. This is one of the motivations in the research area of fitness function approximation in EC [116]. This thesis applies feed-forward neural networks for this purpose.

The rest of this chapter is organized as follows. Key concepts of fitness function approximation are summarized in Section 6.2. Different preventive and corrective control schemes enhancing voltage security are discussed in Section 6.3. Finally, the chapter is concluded in Section 6.4.

6.2 Fitness function approximation

In many engineering optimization problems, computing the fitness function is very expensive. The analytical fitness function is also sometimes not available. Therefore, fitness functions are often approximated based on the data gathered from experiments or simulations. This fitness approximation is generally known as the meta-model. In this model, the fitness evaluation function becomes [117]:

$$eval(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if the original fitness model is used} \\ f(\mathbf{x}) + e(\mathbf{x}) & \text{if a meta-model is used} \end{cases} \tag{6.1}$$

where $e(\mathbf{x})$ is the approximation error of the meta-model. The challenge in dealing with a meta-model is how to manage the error properly. The right balance between the actual but expensive fitness evaluation and the cheap but erroneous one is very crucial to the success of a meta-model. This issue is regarded as the model management or evolution control. In this scheme, the original fitness function is used to evaluate some individuals in every generation

or all individuals in some generations. The individual whose quality is assessed by the original fitness function is called the controlled individual. Similarly, the generation in which all the individuals are evaluated by the original fitness function is called the controlled generation. Model management in EA can be classified into three categories according to [118] based on the viewpoint of evolution control.

- **No evolution control:** The approximate model is assumed to be very exact and the actual fitness evaluation is not used at all.
- **Fixed evolution control:** In this category, the approximate model (AM) and the original function (OF) are used together. Two levels of approximation are possible in the evolution control; individual-based or generation-based. In individual-based control, controlled individuals are evaluated by the OF whereby the fitness of the rest individuals is determined by the AM. Two strategies can be implemented to select the controlled individuals; the best strategy or the random strategy. At every generation, the population is sorted based on the fitness values given by the AM such that the first rank individual represents the best one. In the best strategy, the best individual is chosen. On the other hand, individuals are randomly selected in the random strategy (see Fig. 6.1 for graphical illustration of a minimization problem). Thereafter, the fitness of the controlled individuals is reevaluated by the OF as highlighted in yellow.

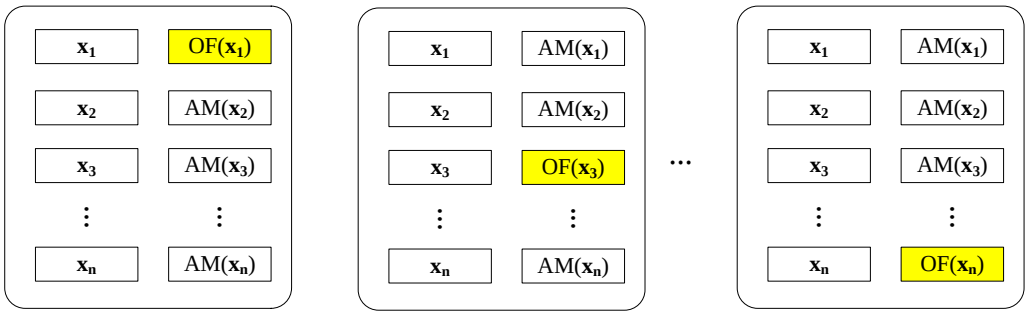


Fig. 6.1 Random selection of the controlled individual

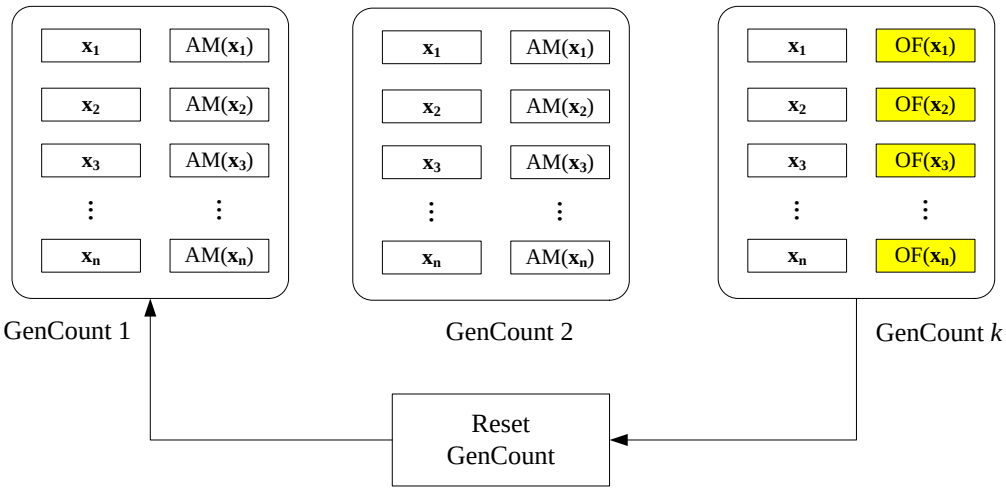


Fig. 6.2 Generation-based evolution control

Generation-based evolution control can be used when the EA with AM converges. Alternatively, the evolution control can be repeated at every fixed number of generations as shown in Fig. 6.2. In this example, all individuals are evaluated by the OF at every *k* generations. An important concern of the methods presented here is how to properly set the frequency of evolution control.

- **Adaptive evolution control:** The frequency of evolution control is adapted based on the accuracy of the AM.

6.3 Control schemes

6.3.1 Optimal reactive power dispatch

In literature, there are two general objective functions for ORPD consisting of:

- Minimization of active power losses in the transmission system;
- Minimization of voltage deviations from the desired voltage levels.

In this thesis, the objective of all VSCORPD problems is to minimize power transmission losses while maintain operational, security and stability limits within the allowable limits. In contrast to some researchers [23, 38, 119], we believe that it is not the objective to operate a power system with maximum VSM. The only requirement is to guarantee sufficient VSM for

normal and severe contingency conditions [120-123]. We disagree to some research works that VSM is not incorporated in the problem formulation because there is no guarantee that the VSM after optimization is greater than the threshold. In this case, the control variables recommended by ORPD with no proper VSM considerations may deteriorate the system stability especially in highly stressed operating conditions where reactive power supports are inadequate. A VSCORPD problem can be formulated as:

Minimize $$f(\mathbf{x}) = P_{loss}(\mathbf{x},\mathbf{d}) \tag{6.2}$$

Subject to:

a) Generator bus voltage limits

$$u_{Gi}^{\min} \le u_{Gi} \le u_{Gi}^{\max} \qquad \forall i \in \mathbf{s}_{PV} \tag{6.3}$$

b) Shunt compensator limits

$$q_{Ci}^{\min} \le q_{Ci} \le q_{Ci}^{\max} \qquad \forall i \in \mathbf{s}_{QC} \tag{6.4}$$

c) Transformer tap setting limits

$$a_i^{\min} \le a_i \le a_i^{\max} \qquad \forall i \in \mathbf{s}_{T} \tag{6.5}$$

d) Load bus voltage limits

$$u_{Li}^{\min} \le u_{Li} \le u_{Li}^{\max} \qquad \forall i \in \mathbf{s}_{PV} \tag{6.6}$$

e) Line flow limits

$$s_{Fi} \le s_{Fi}^{\max} \qquad \forall i \in \mathbf{s}_{L} \tag{6.7}$$

f) Voltage security margin limit

$$VSM \ge VSM^{\min} \tag{6.8}$$

where P_{loss} is the total active power losses in the transmission system; $\mathbf{s}_{PV}$ is the set of generator (PV) buses; $\mathbf{s}_{QC}$ is the set of shunt compensators; $\mathbf{s}_{T}$ is the set of transformers; $\mathbf{s}_{PQ}$ is the set of load (PQ) buses; $\mathbf{s}_{L}$ is the set of transmission lines. The vector **x** contains control variables listed in (6.3)-(6.5) where (6.3) is treated as continuous variables and (6.4) to **(6.5)** are treated as discrete variables. The vector **d** describes the operating condition of the power system represented by active and reactive power demand and active power generation.

This section presents three implementations of different VSORPD problems. Mathematical formulation and simulation results are discussed in each sub section below.

6.3.1.1. A power flow less model

ORPD is a sub problem of optimal power flow (OPF). From the EC viewpoint, the quality of an individual is assessed by calculating power flow to determine the objective value (in this case P_{loss}) and feasibility of corresponding dependent variables. Therefore, over the generations, a large number of power flow calculations are inevitably required. For some OPF applications, such as VSORPD or transient stability constrained OPF, an expensive calculation, such as time-domain simulations or continuation power flow, needs to be carried out in every fitness calculation. The first model presented here investigates feasibilities of a complete meta-model in which the fitness is solely estimated by several NNs. The actual fitness function is not used in this model. The validity should be verified by the accuracy of the optimal solution of the developed model compared with the traditional OPF model.

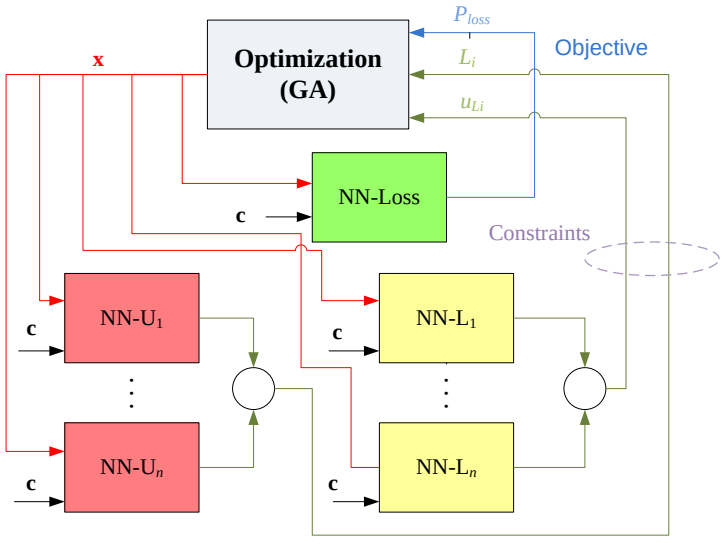


Fig. 6.3 Implementation concept of the power flow less model

The proposed model is conceptually depicted in Fig. 6.3. A group of NNs (denoted by NN-U_1 to NN-U_n) are dedicated to approximate load bus voltages. For large systems, multiple compact NNs with a reasonable number of outputs are generally recommended rather than a single huge one with many outputs. This helps improves the learning capability and saving

the training time. Line (L) indices at every bus are also evaluated by a group of NNs denoted by NN-L_1 to NN-L_n. A NN namely NN-Loss is trained to approximate the total active power transmission loss. Every network uses the same set of input features consisting of two parts namely control variables **x** and constant variables **c**.

Here, we have adopted a binary genetic algorithm (GA) as the optimization engine. Constraints are handled by the static penalty scheme. Note that other powerful optimization algorithms can replace this GA without any major change. The proposed meta-model is termed GA-NN where the conventional model is termed GA-ORPD. Notice that the GA-NN in Fig. 6.3 did not consider the line flow limits shown in (6.7) because we have found some difficulties to accurately predict this quantity based on the given input features. However, the line flow can be computed based on the predicted bus voltages and the system bus admittance matrix. Therefore, the GA-ORPD formulation can be written as:

$$\begin{gathered}\text{Minimize (6.2)} \\ \text{Subject to (6.3)-(6.6) and (6.8)}\end{gathered} \tag{6.9}$$

A dataset of 8000 different operating states was generated based on the technique discussed in Section 4.2.3.1. The training set constitutes the first 6000 patterns whereby the rest 2000 patterns are reserved for testing. For the prediction of voltages and line indices at load buses, four networks (n=4 see Fig. 6.3) are used with six outputs for each network. Generalization capability of the developed networks is examined by giving new input patterns and thereby computing the corresponding estimation error. Given unforeseen 250 operating states, Fig. 6.4 shows the scatter plot between the target value of a bus voltage and the estimation given by a NN-V in the upper figure whereby the corresponding estimation errors are computed and depicted in the lower figure. The similar plotting pattern is carried out for NN-L and NN-Loss as shown in Fig. 6.5 and Fig. 6.6, respectively. These figures clearly demonstrate the developed NNs can effectively and accurately estimate the target values with very small errors. The errors of P_{loss} estimation are relatively the greatest among these three examples with the maximum error roughly about 1% and the average error of 0.3%.

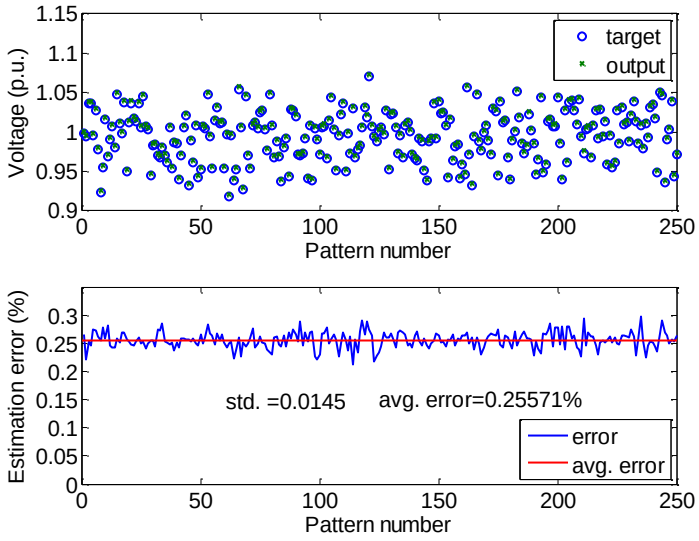


Fig. 6.4 Generalization of NN-V

The operating state number 6900 which is in the testing dataset and was not presented during the NN training phase is used to examine the effectiveness of GA-NN. P_{loss} at the initial state is 11.4151 MW as shown in the red straight line of Fig. 6.7. By adopting the GA-NN to adjust the control variables, the estimated P_{loss} at the final state (after optimization) given by NN-Loss becomes 4.8004 MW. This accounts for 56.16% P_{loss} reduction. However, it must be noted that the final value of P_{loss} is not the actual one due to the existence of NN estimation errors. Therefore, the exact P_{loss} value of the final operating state can be determined by setting system control parameters equal to the optimal control settings given by the GA-NN and running power flow at that demand level. By doing this, the final P_{loss} is 5.0048 MW. The error of estimation in this case is corrected to 4.0828 % which is acceptable in practical applications.

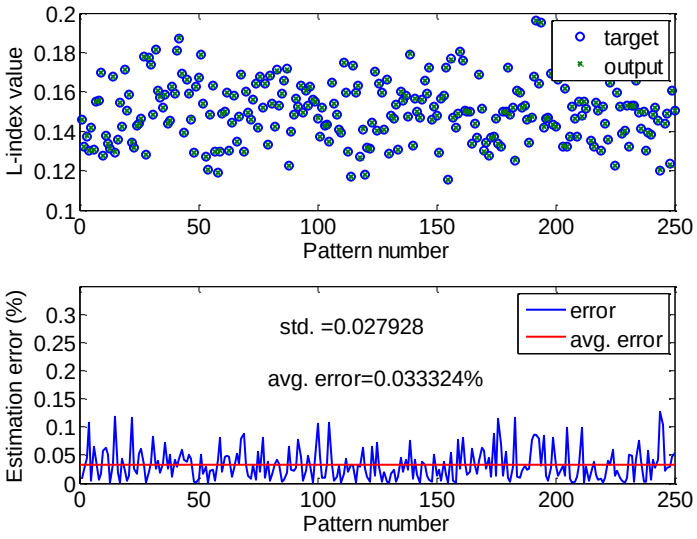


Fig. 6.5 Generalization of NN-L

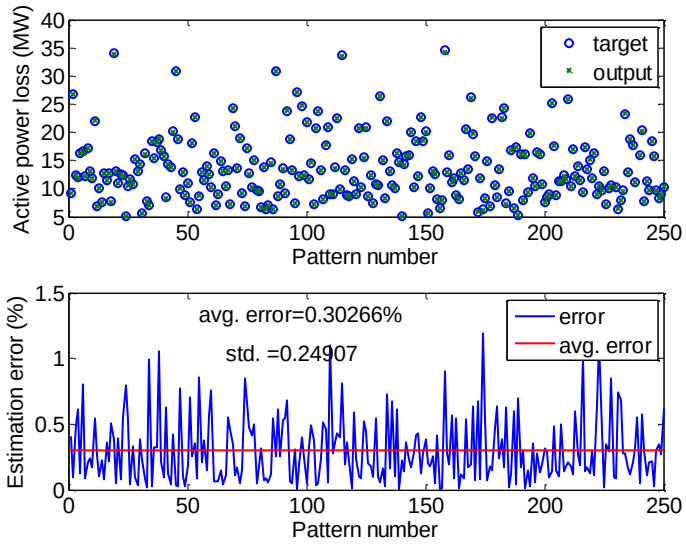


Fig. 6.6 Generalization of NN-Loss

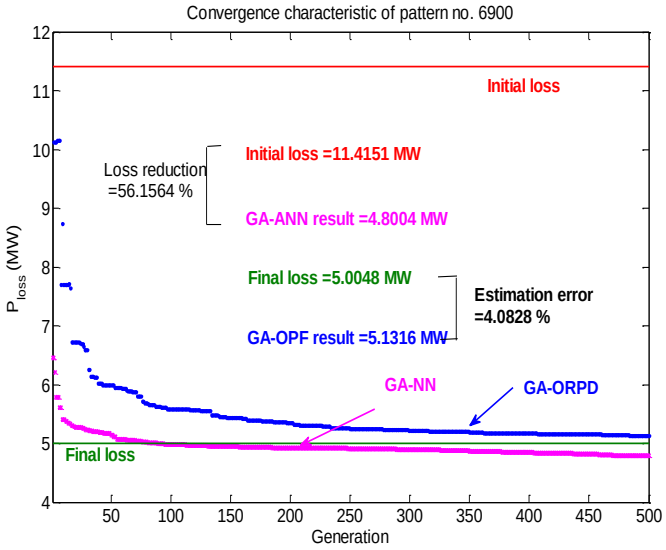


Fig. 6.7 Comparison of convergence properties between GA-NN and GA-ORPD

Comparing the GA-ANN with the conventional GA-based OPF (GA-OPF), the results of both methods are tabulated in Table 6.1. The value of P_{loss} and system voltage stability indicator L^{max} are shown in Table 6.2. It is clear that both methods result in a very close solution. Voltage stability is also improved as reflected in reduction of L^{max}. Considering the computational aspect, the proposed method is much faster than the GA-OPF (approximately 5 times faster).

Besides the computational aspect, it should be further observed that the proposed GA-ANN method requires fewer input information to perform OPF calculation. In the conventional OPF formulation, all system data, such as network parameters, generation and load level and transformer tap setting need to known as inputs for power flow program. On the other hand, GA-ANN needs only few of them to perform the same task because a representative of system quantities sharing statistical similarity determined by k-mean cluster method is used as the inputs.

Table 6.1 Optimal control variables

	GA-ANN	GA-OPF
q_{c1}	-0.03	-0.05
q_{c2}	0.18	0.11
q_{c3}	0.0075	0.0075
q_{c4}	0.065	0.05
q_{c5}	0.05	0.05
q_{c6}	0.025	0.06
q_{c7}	0.09	0.15
q_{c8}	0.175	0.0375
q_{c9}	0.02	0.04

	GA-ANN	GA-OPF
a_1	0.93	1.05
a_2	0.93	0.96
a_3	0.97	0.99
a_4	1.01	0.99
u_{G1}	1.0486	1.0026
u_{G2}	1.0432	0.9997
u_{G3}	1.0188	0.9737
u_{G4}	1.0389	0.9878
u_{G5}	1.1	1.0434
u_{G6}	1.1	1.004

Table 6.2 Simulation results: P_{loss} ,L^{max} and CPU time

	Initial	GA-ANN	GA-OPF
P_{loss} (MW)	11.4151	5.0048	5.1316
L^{max}	0.1775	0.1546	0.1529
Time (s)	-	78.92188	438.2813

6.3.1.2. A two stage model

This section presents a two-stage design of online voltage security assessment and control tool. It is well accepted that dynamic voltage stability is the most accurate index. However, the computing procedure is quite time-consuming and modeling simplifications are normally assumed in many cases. To overcome these limitations, intelligent system modules are currently implemented as a complementary part in addition to conventional simulation tools. The method discussed here utilizes the historical information and results from previous simulations to develop an intelligent classfier capable of identifying voltage security of a power system rapidly. If the current operating state is identified insecure, the appropriate

control settings are suggested to enhance both security and stability levels. The ideas of this model can be conceptually depicted as shown in Fig. 6.8.

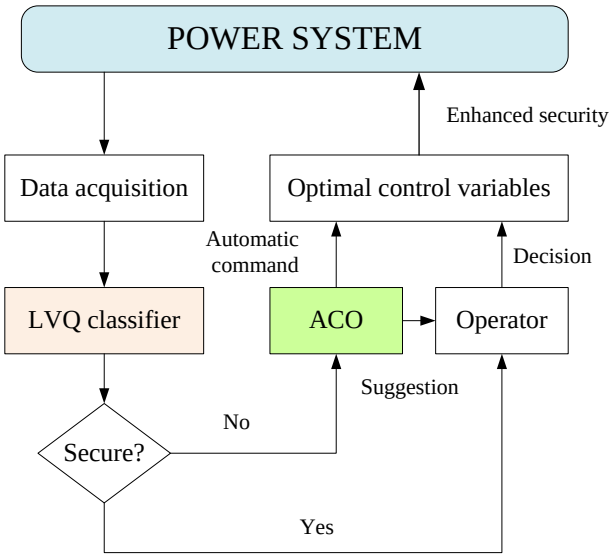


Fig. 6.8 Conceptual diagram of the two-stage model

The model shown here uses two types of NNs for different purposes. Learning vector quantization (LVQ) is applied to classify the voltage security level of the current operating state into two classes namely “secure” and “insecure”. The second NN is used to approximate VSM in the VSCORPD solved by ACO. This network works in the regression mode to substitute the continuation power flow in determining the VSM. Both NNs share the same set of input patterns *i* described by:

$$\mathbf{xn}_i = \left[s_{Fi,1}, s_{Fi,2}, ..., s_{Fi,m}, x_{i,1}, x_{i,2} ..., x_{i,n} \right] \tag{6.10}$$

where $s_{Fi,j}$ is the MVA flow over the selected transmission line j of the operating state i; $x_{i,j}$ is the selected ORPD control variable j of the operating state i; m is the number of selected MVA flows and n is the number of selected ORPD control variables. In this example, the number of input features are set to m=10 and n=19. Therefore, the number of input features is 29. Then, principal component analysis (PCA) is applied to eliminate redundant input information so that the number of input features is reduced and the learning performance is increased.

Table 6.3 Two-class partition of data

Database	Training set	Testing set
5000	4000	1000
Two-class partition		
Class A (secure)	1623	390
Class B (insecure)	2377	610

Based on the input vector shown in (6.10), LVQ as a voltage security classifer is trained by the data whose details are listed in Table 6.3. Performance of LVQ classifier of 1000 unforeseen testing samples is shown in Table 6.4. The performance of the developed LVQ is assessed by (4.3) to (4.5) of Chapter 4. In this study, we find that the effectiveness of the developed LVQ is quite moderate. One of potential reasons of such non-superior performance of LVQ could be that we have assumed too many parameter changes in the phase of generating the knowledge base data. This could introduce a very high degree of complexity to the learning ability of LVQ. Further investigations are required on this aspect.

Table 6.4 Performance evaluation of the classification

Classification of testing set by LVQ		
	Class A	Class B
Samples of class A (390)	349	41
Samples of class B (610)	156	454
Classification performance evaluation		
Success rate	80.3%	(803/1000)
False alarm rate	10.51%	(41/390)
Missed alarm rate	25.57%	(156/610)

Ten operating conditions which were correctly classified by LVQ as voltage insecure are now considered for the voltage security enhancement scheme. Fig. 6.9 shows the active power loss P_{loss} before and after the ACO optimization process of ten different operating

states whereby Fig. 6.10 depicts the corresponding voltage seucrity margin (VSM). The parameters of ACO used in this paper are given in Table 6.5.

Table 6.5 ACO parameter settings

Parameter	Value
Archive size (*k*)	40
Number of ants (*n_ant*)	30
Convergence rate factor (*q*)	0.2
Pheromone evaporation (ξ)	0.99

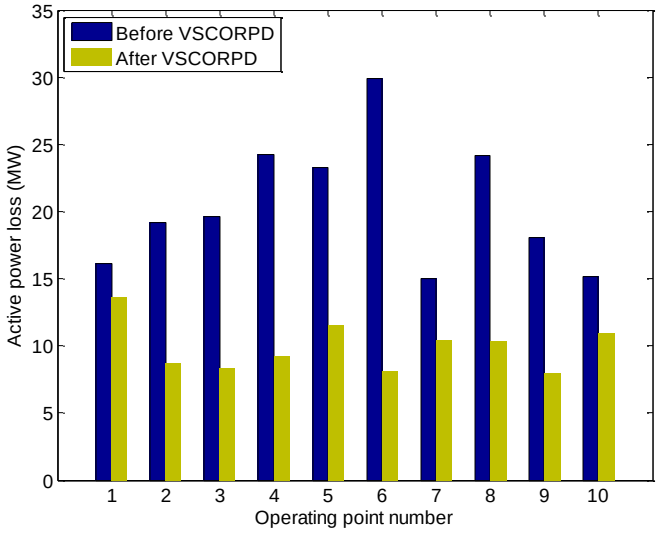


Fig. 6.9 Comparison of active power losses before and after VSCORPD

It can obviously be observed that with optimal setting of control parameters P_{loss} of all operating conditions are significantly reduced. At the same time, all VSMs which were previously below the threshold limit (VSM^{limit} = 0.25) and classified as insecure states are now enhanced because VSM of all states are now greater than 0.25.

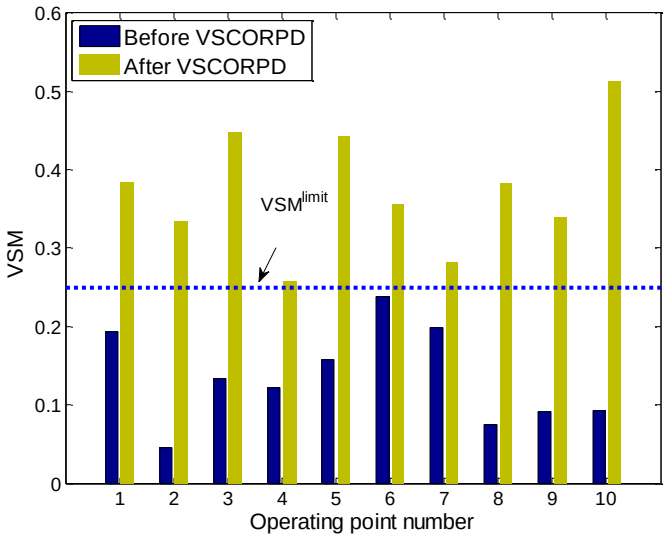


Fig. 6.10 Comparison of voltage stability margins before and after VSCORPD

The active power loss of the current-best (p_{best}) and best-so-far (g_{best}) solutions are plotted against the generation number in to demonstrate the convergence property of the ACO for solving VSCORPD. The ACO terminates if one of the following criterion has been reached:

- After 150 iteraions, there is no further improvement of g_{best};
- the maximum number of generation (G_{max} = 200) is reached.

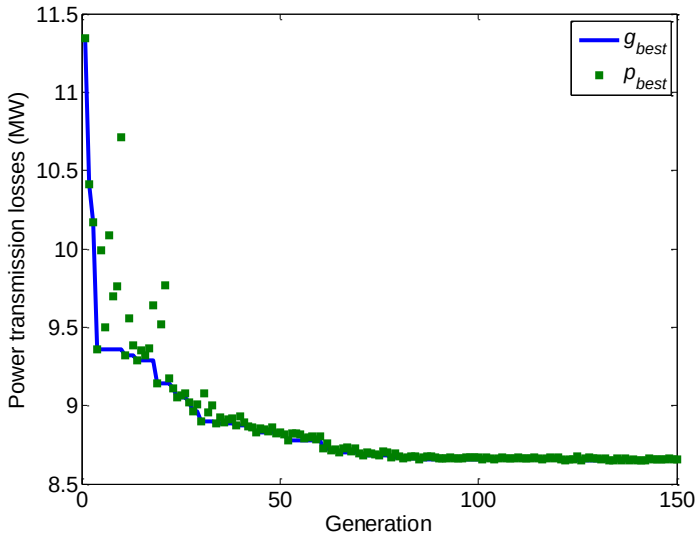


Fig. 6.11 Convergence characteristics of ACO in VSCORPD

The variations of discrete control variables such as transformer tap positions over the optimization process is shown in Fig. 6.12. The variation of selected generator bus voltages which are continuous variables is shown in Fig. 6.13.

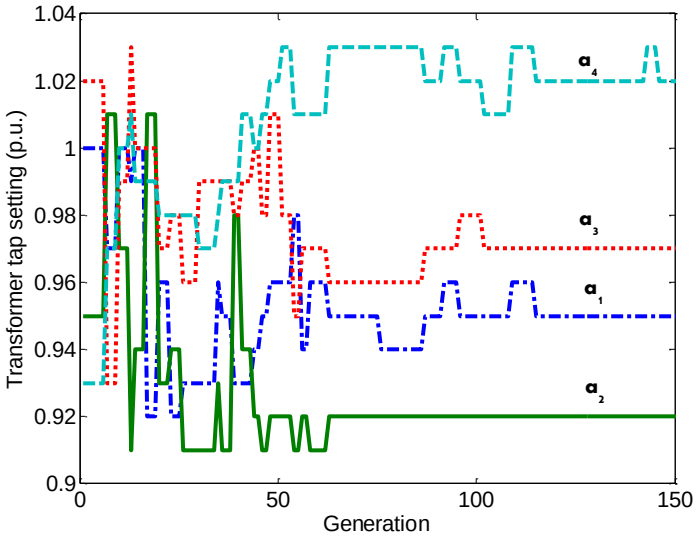


Fig. 6.12 Variation of discrete control variables

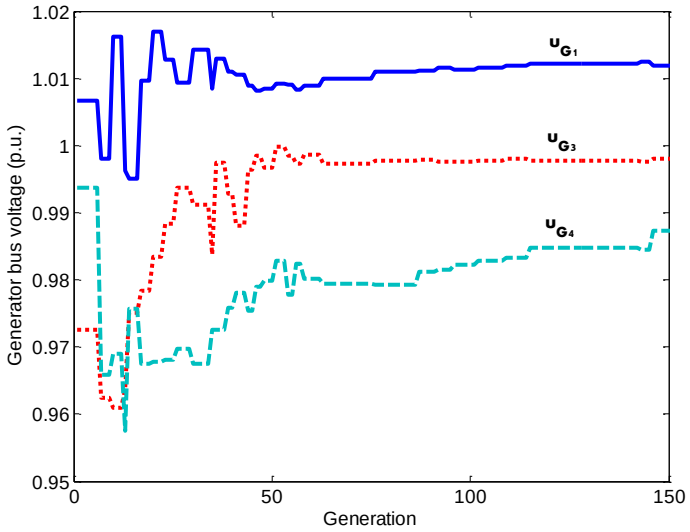


Fig. 6.13 Variation of continuous variables

These two figures show that the control variable barely unchanged when reaching the maximum number of generations. This is the common scenario indicating that the EA is reaching the optimal or a quasi-optimal solution.

6.3.1.3. A model considering costs of adjusting devices

To achieve a practical VSCORPD solution, the number of control actions should not be too large. Excessive operations of discrete control devices, such as OLTC positions, generally shorten life expectancy of the device and thereby incur more maintenance costs and excessive operational burdens. Therefore, a realistic ORPD problem formulation should take this aspect into consideration [124, 125]. The number of switching operations of OLTC and capacitor can be constrained to ensure that the number of operations is less than the maximum daily allowance.

In this thesis, we take a different model. The objective of our VSCORPD is to minimize the cost of energy losses and the cost of adjusting discrete control devices. The latter cost is defined by the multiplication between the difference of device settings of two successive time steps *t*-1 to *t* and the unit operating cost of each device. System security constraints and sufficient VSM are taken into consideration. The optimization is carried out for every planning time step. Therefore, the VSCORPD problem at the time step *t* can be formulated as a combined objective problem as shown below:

$$\text{Minimize } C_W \Delta t P_L^t + \sum_{i=1}^{NT} C_{Ti} \left| a_i^t - a_i^{t-1} \right| + \sum_{i=1}^{NQ} C_{Qi} \left| q_i^t - q_i^{t-1} \right| \qquad (6.11)$$

$$\text{Subject to: (6.3)-(6.8) for all time steps}$$

where C_w is the electricity price (\$/kWh); P_L^t is the total power transmission loss at the time step t (kW); Δt is the time interval (usually 1 hour); C_{Ti} is the unit operating cost of transformer i (\$); a_i^t is the tap position of transformer i at the time step t; C_{Qi} is the unit operating cost of shunt compensator i (\$); q_i^t is the setting of capacitor i at the time step t.

The implementation procedure of this model can be summarized as shown in Fig. 6.14. The set of optimal control variables is determined by ACO. Power flow is carried out in every fitness evaluation based on two sets of inputs; the ACO trial vector representing system control variables and the demand and generation patterns. Power flow dependent variables, such as load bus voltages and line flows are computed and given back to ACO as constraint values. The NN developed in Section 4.4.1 of Chapter 4 is used here to estimate VSM in (6.8). Additional inputs can be given to the NN to enhance the accuracy.

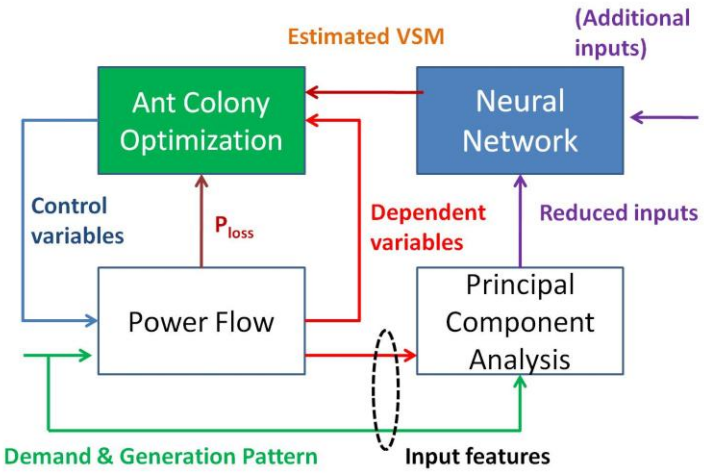


Fig. 6.14 Conceptual diagram of the proposed method

The effectiveness of this model is verified by solving VSCORPD of a daily operation of the modified IEEE 30-bus test system. The active and reactive load profile for a day with 15 minute sampling interval is shown in Fig. 6.15. The generators are assumed to supply the load with a fixed participation factor throughout the daily operation.

It is the requirement that 6% of voltage stability margin has to be guaranteed at every operating state ($VSM^{\text{limit}} = 1.06$).The unit operating cost of transformer of all transformers C_{Ti} is assumed to be equal to \$10/change where all capacitors have the identical unit operating cost C_{Qi} of \$3/change. The electricity price C_W is assumed to be 6 cent/kWh.

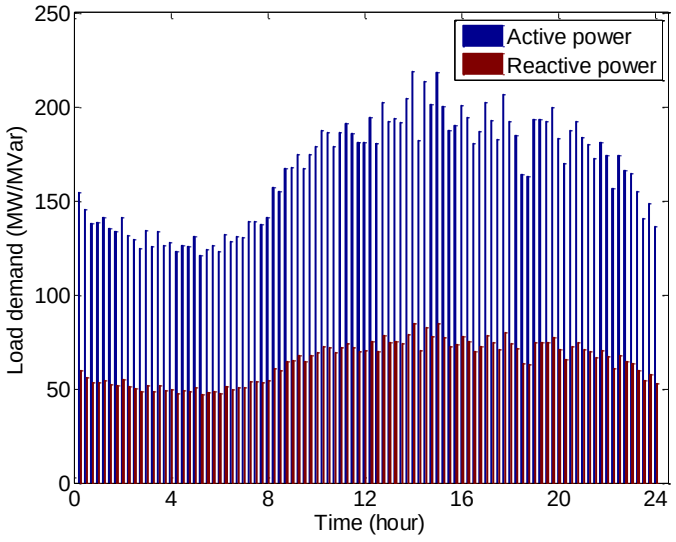


Fig. 6.15 Daily load profile (15-minute sampling interval)

Three control schemes namely ACO #1 to ACO #3 are set up where different objective functions are minimized whereby the same set of constraints is considered. The ACO #1 is the proposed model as shown in (6.11). In ACO #2, the traditional model of loss minimization is applied. The objective function of ACO #2 is the first term of (6.11). The number of discrete control actions is minimized in ACO #3 where the objective function can be represented by the last two terms of (6.11). The ACO parameters are set up as follows; n_{arch} = 20, n_{ant} =10; gen^{max} = 300. The termination criterions for ACO are similar to the ones discussed in section 6.3.1.2. The ACO stops if one of them has been reached.

Following the given demand and generation patterns, CPF analyses reveal that six operating conditions at 14:00, 14:30, 15:00,15:30, 17:15 and 17:30 have VSMs less than the VSM^{limit}. All three control schemes ACO #1 to ACO #3 were applied to solve for the optimal control scheme at every operating time interval. VSMs based on the optimal control settings determined from each control scheme are shown in Fig. 6.16. It can be observed that all three control schemes are successful in improving the VSM to be greater than VSM^{limit} as it is shown that there is no operation condition with VSM less than 1.06.

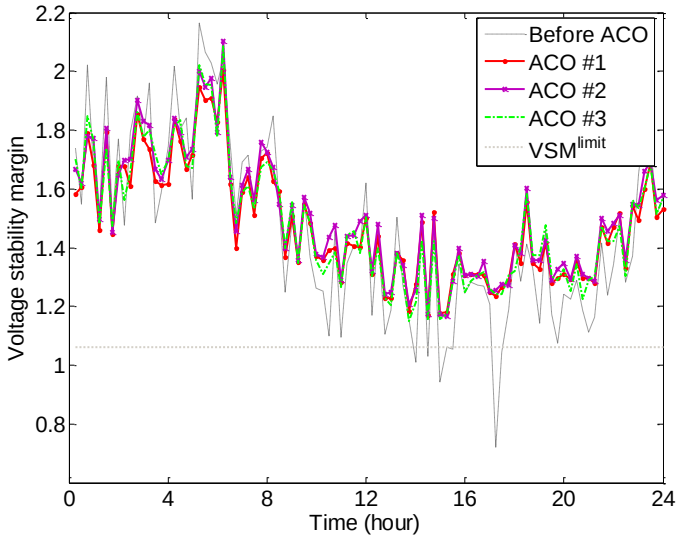


Fig. 6.16 Voltage stability margins of different cases

Some of optimal control settings of discrete control devices during the 24-hour operating horizon are graphically shown here. The daily variation of tap positions of transformers T_1 and T_4 are shown in Fig. 6.17. Obviously, ACO #2 results in a relative larger number of

device operations because the cost of adjustment is neglected. Such excessive operation significantly reduces the life time expectancy of the device. It is practically feasible to frequently change the tap position. On the other hand, ACO #1 and ACO #3 can help reduce the number of operations by considering costs in the objective function. The similar observation can be drawn from the operation of two selected capacitors C_1 and C_8 as shown in Fig. 6.18. It can be observed that capacitor settings change comparatively more frequently than transformer taps because C_{Qi} is approximately three times cheaper than C_{Ti}.

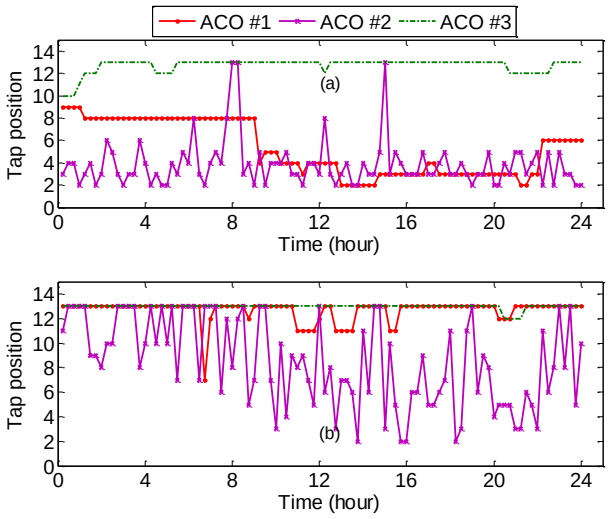


Fig. 6.17 Optimal tap position of (a) transformer T1 and (b) transformer T4

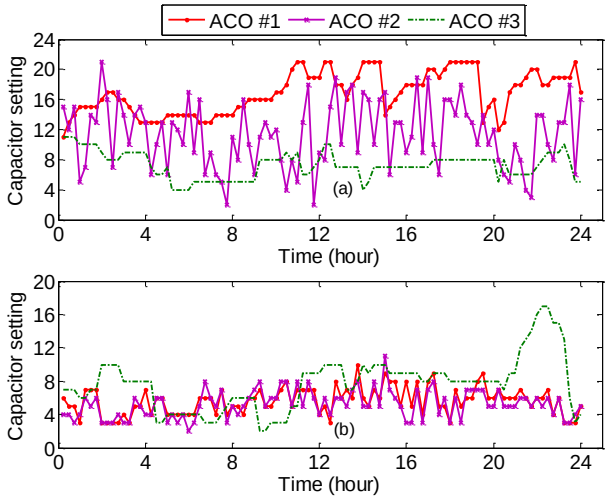


Fig. 6.18 Optimal setting of (a) capacitor C1 and (b) capacitor C8

The total numbers of operation action of transformer tap positions in 24 hours determined by each method are summarized in Table 6.6. Similarly, Table 6.7 summarizes the total operation times of capacitors in 24 hours. It is not surprising that ACO #3 result in the fewest operation times because the cost of adjustment is only the objective function. The number of tap operations found from ACO #1 is slightly greater than ACO #3 but still significantly less than ACO #2 in both cases of transformer and capacitors.

Table 6.6 Total operating times of tansformer tap positions

Method	T_1	T_2	T_3	T_4	$\sum T$
ACO #1	19	15	10	28	72
ACO #2	165	213	296	311	985
ACO #3	9	17	4	2	32

Table 6.7 Total operating time of capacitors

Method	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	$\sum C$
ACO #1	82	75	58	90	128	156	119	137	845
ACO #2	463	242	243	151	244	207	392	155	2097
ACO #3	52	67	67	48	64	75	72	67	512

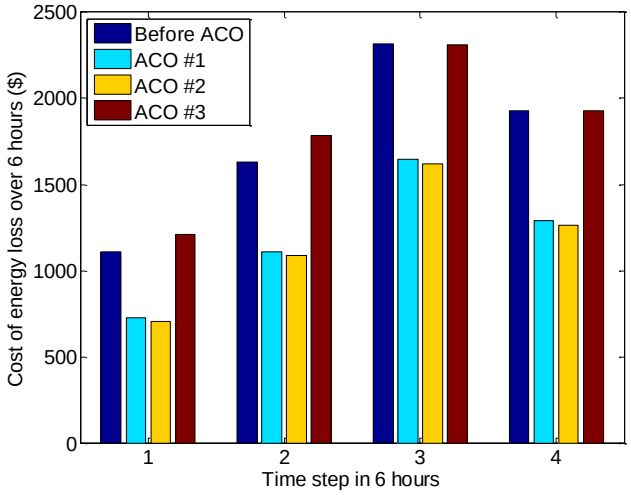


Fig. 6.19 Cost due to energy loss during six-hour periods

Based on the results presented so far, it seems that ACO #3 outperforms the others. Costs due to energy losses C_{loss} over the period of 6 hours before and after three control schemes are shown in Fig. 6.19. It can be observed that in every time step C_{loss} of ACO #3 is relatively equal or even higher than the C_{loss} of the initial operating condition (before ACO). Interestingly, C_{loss} of ACO #1 and ACO #2 are slightly different. The summation of C_{loss} at each time step results in the cost of energy losses for a daily operation $\sum C_{loss}$. Table 6.8 gives $\sum C_{loss}$ before the optimization and results of all control schemes. Percentage of cost reduction with respect to the case without optimization is shown in the second row of Table III. ACO #1 and ACO #2 have nearly identical performance in reducing C_{loss} while ACO #3 results in even greater C_{loss} represented by -5.98% of cost reduction. Therefore, it is very evident that ACO #1 outperforms the rests because of the ability to save C_{loss} and reduce the excessive operation of discrete control devices. The average CPU times used by each method at a time *t* are given in Table 6.9. ACO #3 is quite faster because the stopping criterion on no further improvement of solution quality is met. Also, there is no significant difference between CPU times of ACO#1 and ACO#2. This indicates that the proposed ACO#1 does not involve any further calculation complexity.

Table 6.8 Cost due to energy losses for the daily operation

	Before ACO	ACO #1	ACO #2	ACO #3
$\sum C_{loss}$($)	6991.3541	4882.1512	4678.2525	7409.7361
%	-	+30.17	+33.08	-5.98

Table 6.9 CPU time used by each method (s)

ACO #1	ACO #2	ACO #3
96.6507	96.2752	57.7555

6.3.2 Under-voltage load shedding

The solution of optimal load shedding involves the determination of the effective locations and optimal load reductions subject to various system constraints. This optimization task can be carried out in two stages: planning and operation. In the planning stage, system behaviors of different scenarios are analyzed and if necessary different control strategies may be determined. During the operation, an optimization algorithm is used to suggest the efficient operation scheme as per grid requirements. In the OPF framework, the main objective of the load shedding optimization problem is to minimize the cost of power interruption at buses while considering a number of system constraints as follows:

Minimize
$$f(\Delta p_d) = \sum_{i \in \mathbf{n}_s} C_i \cdot \left(\frac{\Delta p_{di}}{\partial \lambda / \partial p_{di}} \right) \tag{6.12}$$

subject to

a) Load bus voltage limits

- Base condition

$$u_{Li,b}^{\min} \le u_{Li,b} \le u_{Li,b}^{\max} \qquad \forall i \in \mathbf{s}_{PQ} \tag{6.13}$$

- Max. loading condition

$$u_{Li,m}^{\min} \le u_{Li,m} \le u_{Li,m}^{\max}$$

b) Line power flow limits

- Base condition

$$s_{Li,b}^{\min} \le s_{Li,b} \le s_{Li,b}^{\max} \qquad \forall i \in \mathbf{s}_{L} \tag{6.14}$$

- Max. loading condition

$$s_{Li,m}^{\min} \le s_{Li,m} \le s_{Li,m}^{\max}$$

c) Fixed power factor

$$\frac{\Delta p_{di}}{p_{di}^0} = \frac{\Delta q_{di}}{q_{di}^0} \qquad \forall i \in \mathbf{s}_{S} \tag{6.15}$$

d) Allowable load curtailment

$$\Delta p_{di}^{\min} \le \Delta p_{di} \le \Delta p_{di}^{\max} \qquad \forall i \in \mathbf{s}_{S} \tag{6.16}$$

e) Voltage stability margin limit

$$1 \le \lambda_0 + \sum_{i=1}^{N} \frac{\partial \lambda}{\partial p_{di}} \Delta p_{di} + \sum_{i=1}^{N} \frac{\partial \lambda}{\partial q_{di}} \Delta q_{di} \le 1.06 \tag{6.17}$$

where C_i is the power interruption cost at bus i (\$/kW) ; $\mathbf{s}_{PQ}$ is the set of load (PQ) buses; $\mathbf{v_L}$ is the set of transmission lines; $\mathbf{s}_S$ is the set of effective load buses selected for load shedding. The control variables in this load shedding problem are the active power load curtailment at effective buses represented by Δp_{di}. They are bounded by the minimum and maximum allowable loads can be curtailed at the bus i as shown in (6.16). The dependent variables are computed by the power flow calculation. Listed in (6.13)-(6.14) are the constraints of load bus voltages and line flows at base-case and at maximum loading conditions. The latter constraint ensures that the system operating condition will not violate the security limits. To simplify the problem, power factor at the load shedding buses are assumed to be unchanged. This can be achieved by proportionately curtailing the reactive power load Δq_{di} according to (6.15); where p_{di}^0 and q_{di}^0 are initial active and reactive power demand of the bus i, respectively. The value of λ is calculated based on the linear estimation technique discussed earlier in Section 4.3 of Chapter 4.

Because the power system may become unstable (λ<1) after a severe disturbance, therefore the load shedding algorithm must be able to bring the system back to the boundary of stable operation (λ=1). However, it may not be necessary in a practical viewpoint to guarantee a great distance to the collapse. Therefore, the maximum stability margin of 6% is set (λ=1.06) up as expressed in (6.17).

The proposed load shedding scheme is applied again to the IEEE 30-bus test system to verify its effectiveness. As mentioned earlier that load shedding should be only used in a very critical situation, therefore a critical operating point with λ<1 is created. This is accomplished by doubling the power demand from the base case to 566.8 MW. Also, the N-1 contingency analysis reveals that the outage of the line connected between buses 28 and 27 is the most severe with λ=0.7533 which is an unstable operating condition. The load shedding program is a fast and powerful measure to counteract this problem. Therefore, the optimal control actions will be determined to bring this operating point from the unstable operating zone to the stable one.

Costs due to power interruption incurred by power consumers in different sectors according to [126] are given in Table 6.10. According to the sensitivity analysis presented in Chapter 4, only five load locations are selected to participate in the load shedding program. The permissible amounts of load that can be curtailed at each bus are listed in Table 6.11. The composition of each load is also presented in that table. For example, the load bus 30 is of the 0.6t+0.2i+0.2r composition. This means 60% of total demand of this bus comes from the transportation (t) sector, 20% from the industrial (i) sector and 20% from the residential (r) sector. Costs per 100 kW power interruptions at every bus showing different cost characteristics are also listed in Table 6.11.

Table 6.10 Cost due to power interruption in different sectors

Interruption cost ($/kW)			
Transportation (t)	Industrial (i)	Commercial (c)	Residential (r)
16.42	13.93	12.87	0.15

Table 6.11 Load shedding limits and load configuration

Bus	$\Delta p_{di,min}$(pu)	$\Delta p_{di,max}$(pu)	Configuration	Cost ($/100 kW)
23	0	0.032	0.5c+0.5r	651
24	0	0.087	0.3t+0.7i	1467.7
26	0	0.035	1r	15
29	0	0.024	0.4i+0.2c+0.4r	820.6
30	0	0.070	0.6t+0.2i+0.2r	1266.8

The control parameters of ACO are set to the values as shown in Table 6.12. Following the optimization process, the PV profile of the most critical bus (bus 30) obtained by the CPF is plotted in Fig. 6.20 against pre- and post-contigency (with no control actions) conditions. It is demonstrated that the proposed ACO-based load shedding method is able to restore voltage stability of the system while maintaining a number of constraints within their limits.

Table 6.12 Parameter settings of the ACO

Parameter	Value
Archive size (k)	40
Number of ants (n_{ant})	20
Number of replancement (n_{rp})	8
Convergence rate factor (q)	0.1
Pheromone evaporation (ξ)	0.99

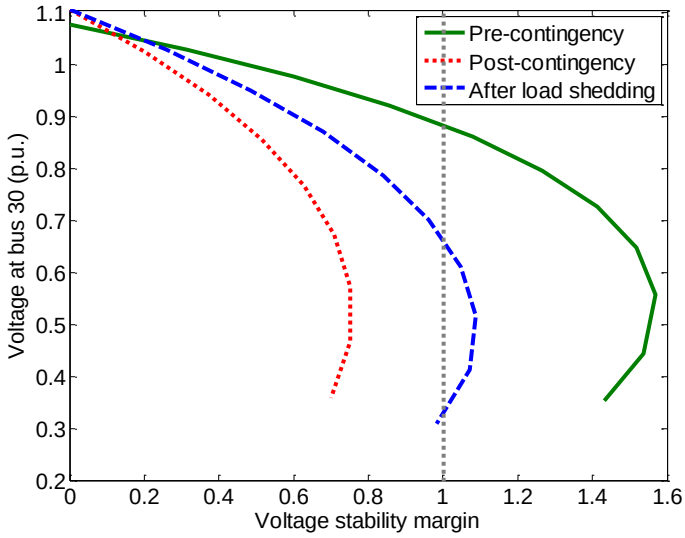


Fig. 6.20 PV curves of different operating conditions

The average convergence property obtained from 50 independent runs is shown in Fig. 6.21 (a). It is obvious that the proposed algorithm is capable of discovering the optimal solution at a very fast speed. Statistical evaluation has been performed and the histogram of the optimal objective value is depicted in Fig. 6.21 (b). It is obvious that the ACO nearly converges to the same solution. The optimal results of each independent run, the final λ after load shedding and CPU time are averaged and given in Table 6.13. Statistical values of the final objective values are shown in Table 6.14.

Table 6.13 Optimal solution and simulation time

Optimal control variables (kW)					λ	Time (s)
Δp_{d23}	Δp_{d24}	Δp_{d26}	Δp_{d29}	Δp_{d30}		
1.142	1.043	3493.378	2333.508	6081.638	1.046	144.425

Table 6.14 Statistics of 50 independent runs

Min	Mean	Max	Std.
52377.4765	52536.2404	53003.3872	142.9681

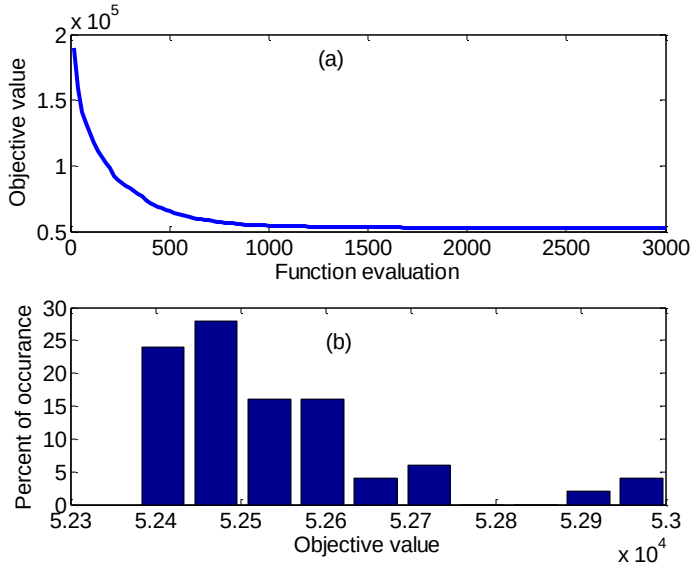


Fig. 6.21 Performance of ACO (a) average convergence property (b) histogram of the best objective values

6.4 Summary

This chapter proposes different optimal power flow (OPF) models for enhancing power system voltage stability in the framework of evolutionary computation. In some computationally expensive problems, the fitness function can be approximated by mapping models. This concept is called the meta-model. This chapter presents three model of this kind for solving voltage stability constrained optimal reactive power dispatch as the preventive

control measure. Voltage stability margin (VSM) is incorporated in addition to the constraints of the conventional OPF. Moreover, an optimal load shedding problem as the corrective measure to voltage collapse is proposed. VSM in this work is approximated by the sensitivity approach.

Chapter 7
Conclusion

7.1 Concluding observations

In the scope of this thesis, several applications of some computational intelligence (CI) methods have been developed to solve voltage stability assessment and control problems of electric power systems. With the limited information exchanges and restriction in computing time in real-time analysis, CI has become a very attractive technique for real-time operation of the future power grid. Following the observations and discussions made in the individual chapters, the following conclusions are made.

7.1.1 Online voltage stability assessment

There are a number of indices that can be used to determine voltage stability of the power system. Among them, the maximum loadability margin (MLM) is very popular due to its straightforward physical meaning. To determine this quantity, multiple power flows are inevitably required. Therefore, this process would be quite time-consuming for a large power system and especially when MLM has to be repeatedly calculated. This thesis implements two methods namely neural network (NN) and sensitivity approaches for estimate MLM rapidly while sacrificing acceptable errors. A new method for generating a large number of operating conditions was introduced. In this method, correlation of active and reactive power demands among load buses is considered and modeled by multivariate normal distribution. The variation of control variables, such as transformer tap positions, generator reactive power generation, etc is also taken into account.

7.1.2 Evolutionary algorithm

Evolutionary algorithms have been proven in the past few decades as an effective means for solving many complex and real-world optimization problems. It is quite well-known that control parameters play a very significant role in the quality of solutions. Toward the line of developing a powerful and parameter-less optimization algorithm, a recent variant of differential evolution namely JADE was modified to eliminate the only remaining control

parameter; the population size. The so-called JADE-vPS algorithm adaptively changes the number of individuals that the algorithm evolves in each generation denoted as *NP*. More individuals are added up to the population if it is necessary. The maximum and minimum boundaries of *NP* are introduced to balance diversification and intensification of the search. Statistical results based on non-parametric tests reveal that JADE-vPS outperforms the conventional DE and an adaptive DE in most problems. Compared to the original JADE, JADE-vPS outperforms in all unimodal test problems and some multimodal ones. These results encourage us to explore new methods to further improve the search capability.

7.1.3 Voltage stability constrained optimal power flow

VSCOPF is an optimization problem with non-convex and discontinuous cost functions capable of determining a voltage secure operating point. VSCOPF is formulated as a nonlinear programming problem and solved by various EAs. The proposed method considers both active and reactive power flow equations to determine state variables. Control variables consist of active and reactive power outputs and transformer tap positions. In this formulation, the proposed method is able to analyze active and reactive power optimization at the same time. According to simulation results of the modified IEEE-30 bus test system, the developed JADE-vPS algorithm produces very robust results in terms of robustness and solution quality. Moreover, the average convergence properties also show that JADE-vPS performs very well and reliably in locating an optimal solution without specifying any control parameters.

7.1.4 Countermeasures

For a computationally expensive problem such as determining countermeasures, an EA may become sluggish due to the large number of fitness evaluations including standard power flow calculation and stability margin computation. To solve this problem, the actual but expensive fitness function is approximated by a faster but less accurate model. Three models were proposed to determine the optimal control variables as preventive measures to voltage collapse. This problem is termed voltage stability constrained optimal reactive power dispatch (VSCORPD) in this thesis.

In the first model, there is no actual fitness evaluation involved and all constraints including the voltage stability margin (L-index) were approximated by designated NNs. The

results showed significant computational reduction in terms of CPU times while sacrificing acceptable errors. A two stage approach was proposed as the second model. An operating condition is classified by an offline trained learning vector quantization (LVQ) whether or not it is voltage stable. For an unstable one, the appropriate control settings are determined by VSCORPD. A feed-forward NN assists VSCORPD in the second stage. Few unstable operating conditions that were correctly classified by the LVQ were selected for the stability improvement. It is shown that the stability margin of all these operating conditions was successfully enhanced beyond the minimum requirement. The third model for VSCORPD minimized the actions of discrete control devices in addition to the conventional energy losses for daily operation of a power system. Control and energy costs are incorporated and minimized subject to various system security and stability constraints. Simulation results demonstrated that the proposed method can reduce the energy cost and the total number of control actions for a daily operation while at the same time guaranteeing sufficient stability margin.

A load shedding model was developed as a corrective control measure that helps restore the power system from an unstable condition to a stable one. The objective of this model is to minimize the cost of power interruption. Sensitivity of MLM with respect to active power demand at the bus is incorporated into the objective function so that the influence of effective load buses can be captured in the optimization process. Robustness of the method was clearly shown from the results of 50 independent runs. The proposed method shows the ability to restore voltage stability of the power system with minimum effects to the load.

7.2 Future research directions

There are a number of issues that can be addressed in the future works on computational intelligence for voltage stability assessment and control as listed below.

- Neural networks (NNs) or any other learning tools can be developed to approximate the MLMs that are restricted by different limits such as generator reactive power, power transmitted over the lines, Hopf bifurcation, etc.
- Recurrent NNs can be applied for predicting dynamic voltage profiles of power systems.

- FACTS devices can be incorporated in the optimization technique to enhance voltage stability. Appropriate size and location of the device should be identified.
- The search capability of the proposed JADE-vPS algorithm can be further improved.
- A robust machine learning method that is less insensitive to changes in system conditions, such as changes in network topology, load demand and missing data should be developed.
- Countermeasures could be considered as a stochastic optimization problem to capture various uncertainties.

References

[1] V. Ajjarapu, *Computational Techniques for Voltage Stability Assessment and Control*: Springer Verlag, 2006.

[2] S. P. Teeuwsen, "Oscillatory Stability Assessment of Power Systems using Computational Intelligence," Ph.D. Thesis, Faculty of Engineering, University of Duisburg-Essen, Duisburg, Germany, 2005

[3] "Voltage Stability Assessment: Concepts, Practices and Tools," in *Special Publication*, August 2003.

[4] T. Van Cutsem and C. Vournas, *Voltage Stability of Electric Power Systems*: Kluwer Academic, 1998.

[5] P. Kundur, *Power System Stability and Control*: McGraw-Hill, 1994.

[6] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. A. Canizares, N. Hatziargyriou, D. J. Hill, A. Stankoviv, C. W. Taylor, T. Van Cutsem, and V. Vittal, "Definition and Classification of Power System Stability," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 1387-1401, May 2004.

[7] C. W. Taylor, *Power System Voltage Stability*: McGraw-Hill Inc., 1994.

[8] B. H. Lee and K. Y. Lee, "A Study on Voltage Collapse Mechanism in Electric Power Systems," *IEEE Trans. Power Syst.*, vol. 6, no. 3, pp. 966-974, Aug. 1991.

[9] B. H. Lee and K. Y. Lee, "Dynamic and Static Voltage Stability Enhancement of Power Systems," *IEEE Trans. Power Syst.*, vol. 8, no. 1, pp. 231-238, Feb. 1993.

[10] X. Lei and D. Retzmann, "Static and dynamic approaches for analyzing voltage stability," *Euro. Trans. Electr. Power* vol. 16, pp. 277-296, 2006.

[11] G. K. Morrison, B. Gao, and P. Kundur, "Voltage Stability Analysis using Static and Dynamic Approaches," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 1159-1171, Aug. 1993.

[12] B. Gao, G. K. Morison, and P. Kundur, "Voltage Stability Evaluation using Modal Analysis," *IEEE Trans. Power Syst.*, vol. 7, no. 4, pp. 1529-1542, Nov. 1992.

[13] Y. Mansour, W. Xu, F. L. Alvarado, F. L. Alvarado, and C. Rinzin, "SVC Placement using Critical Modes of Voltage Stability," *IEEE Trans. Power Syst.,* vol. 9, no. 2, pp. 757-763, May 1994.

[14] C. A. Canizares and F. L. Alvarado, "Point of collapse and continuation methods for large ac/dc systems," *IEEE Trans. Power Syst.,* vol. 8, no. 1, pp. 1-8, Feb. 1993.

[15] V. Ajjarapu and C. Christy, "The Continuation Power Flow: A Tool for Steady State Voltage Stability Analysis," *IEEE Trans. Power Syst.,* vol. 7, no. 1, pp. 416-423, Feb. 1992.

[16] H. D. Chiang, A. J. Flueck, K. S. Shah, and N. Balu, "CPFLOW: A practical tool for tracing power system steady-state stationary behavior due to load and generation variations," *IEEE Trans. Power Syst.,* vol. 10, no. 2, pp. 623-634, May 1995.

[17] P. A. Löf, T. Smed, G. Andersson, and D. J. Hill, "Fast Calculation of a Voltage Stability Index," *IEEE Trans. Power Syst.,* vol. 7, no. 1, pp. 54-60, Feb. 1992.

[18] C. A. Canizares, A. C. Z. d. Souza, and V. H. Quintana, "Comparison of Performance Indices for Detection of Proximity to Voltage Collapse," *IEEE Trans. Power Syst.,* vol. 11, no. 3, pp. 1441-1450, Aug. 1996.

[19] R. J. Avalos, C. A. Canizares, F. Milano, and A. J. Conejo, "Equivalency of Continuation and Optimization Methods to Determine Saddle-Node and Limit-Induced Bifurcation in Power Systems," *IEEE Trans. Circuits and Syst.-I: Regular Papers,* vol. 56, no. 1, pp. 210-223, Jan. 2009.

[20] W. Rosehart, C. Roman, and A. Schellenberg, "Optimal Power Flow with Complementarity Constraints," *IEEE Trans. Power Syst.,* vol. 20, no. 2, pp. 813-822, May 2005.

[21] T. Van Cutsem, "A method to compute reactive power margins with respect to voltage collapse " *IEEE Trans. Power Syst.,* vol. 6, no. 1, pp. 145-156, Feb. 1991.

[22] F. Milano, C. A. Canizares, and A. J. Conejo, "Sensitivity-Based Security-Constrained OPF Market Clearing Model," *IEEE Trans. Power Syst.,* vol. 20, no. 4, pp. 2051-2060, Nov. 2005.

[23] F. Milano, C. A. Canizares, and M. Invernizzi, "Multi-objective optimization for pricing security in electricity markets," *IEEE Trans. Power Syst.,* vol. 18, no. 2, pp. 596-604, May 2003.

[24] T. VanCutsen, "A method to compute reactive power margins with respect to voltage collapse " *IEEE Trans. Power Syst.*, vol. 6, no. 1, pp. 145-156, Feb. 1991.

[25] E. Vaahedi, Y. Mansour, C. Fuchs, S. Granville, M. d. L. Latore, and H. Hamadanizadeh, "Dynamic Security Constrained Optimal Power Flow/ Var Planning " *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 38-43, Feb. 2001.

[26] A. A. El-Dib, H. K. M. Youssef, M. M. El-Metwally, and Z. Osman, "Maximum loability of power systems using hybrid particle swarm optimization," *Elec. Power Syst. Research*, vol. 76, pp. 485-492, 2006.

[27] S. H. Goh, Z. Y. Dong, and T. K. Saha, "Locating Voltage Collapse Points using Evolutionary Computation Techniques," in *2007 IEEE Congr. Evol. Compt. (CEC 2007)*, 2007, pp. 2923-2930.

[28] V. Balamourougan, T. S. Sindhu, and M. S. Sachdev, "Technique for online prediction of voltage collapse," *IEE Gener. Transm. Distrib.*, vol. 151, no. 4, pp. 453-460, July 2004.

[29] P. Kessel and H. Glavitsch, "Estimating the Voltage Stability of a Power System," *IEEE Trans. Power Delivery*, vol. PWRD-1, no. 3, pp. 346-352, July 1986.

[30] T. Q. Tuan, J. Fandino, N. Hadjsaid, J. C. Sabonnadiere, and H. Vu, "Emergency load shedding to avoid risks of voltage instability using indicators," *IEEE Trans. Power Syst.*, vol. 9, no. 1, pp. 341-351, Feb. 1994.

[31] M. Glavic and T. Van Cutsem, "Wide-Area Detection of Voltage Instability From Synchronized Phasor Measurements. Part I: Principle," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1408-1416, Aug. 2009.

[32] K. Vu, M. Begovic, D. Novosel, and M. M. Saha, "Use of Local Measurements to Estimate Voltage Stability Margin," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1029-1035, Aug. 1999.

[33] I. Simon, G. Verbic, and F. Gubina, "Local voltage-stability index using Tellegen's theorem," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1267-1275, Aug. 2006.

[34] B. Milosevic and M. Begovic, "Voltage Stability Protection and Control using a Wide-Area Network of Phasor Measurements," *IEEE Trans. Power Syst.*, vol. 18, no. 1, pp. 121-127, Feb. 2003.

[35] M. H. Haque, "Online monitoring of maximum permissible loading of a power system within voltage stability limits " *IEE Gener. Transm. Distrib.,* vol. 150, no. 1 pp. 107-112, Jan. 2003.

[36] M. H. Haque, "Use of V-I characteristic as a tool to assess the voltage stability limit of a power system," *IEE Gener. Transm. Distrib.,* vol. 151, no. 1, pp. 1-7, Jan. 2004.

[37] M. Nizam, A. Mohammed, and A. Hussain, "Performance Evaluation of Voltage Stability Indices for Dynamic Voltage Collapse Prediction," *Journal of Applied Sciences,* vol. 6, no. 5, pp. 1104-1113, 2006.

[38] M. Tripathy and S. Mishra, "Bacteria Foraging-Based Solution to Optimize Both Real Power Loss and Voltage Stability Limit," *IEEE Trans. Power Syst.,* vol. 22, no. 1, pp. 240-248, Feb. 2007.

[39] T. Kohonen, *Self-Ogranization and Associative Memory*, 2 nd. ed. Berlin: Springer, 1988.

[40] J. Zhang and A. C. Sanderson, "JADE: Adaptive Differential Evolution with Optional External Archive," *IEEE Trans. Evol. Compt.,* vol. 13, no. 5, pp. 945-958, Oct. 2009.

[41] S. P. Teeuwsen, I. Erlich, and M. A. El-Sharkawi, "Feature Reduction for Neural Network based Small-Signal Stability Assessment," in *Power System Computation Conference (PSCC) 2002*, Sevilla, Spain, 2002.

[42] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, UK: Oxford University Press, 2007.

[43] R. Fisher, "The use of mulitple measurements in taxonomic problems," *Annals of Eugencies,* vol. 7, pp. 179-188, 1936.

[44] I. T. Jollife, *Principal Component Analysis*. New York: Springer-Verlag, 1986.

[45] R. Eberhart and Y. Shi, *Computational Intelligence*. MA, USA: Morgan Kaufmann, 2007.

[46] G. N. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret, "Wind power forecasting using advanced neural networks models," *IEEE Trans. Ener. Conv.,* vol. 11, no. 4, pp. 762-767, Dec. 1996.

[47] D.Chen and R. R. Mohler, "Neural-Network-Based Load Modeling and Its Use in Voltage Stability Analysis," *IEEE Trans. Control Syst. Tech.,* vol. 11, no. 11, pp. 460-470, July 2003.

[48] A. Konar, *Computational Intelligence: Principles, Techniques and Applications*. Heidelberg: Springer Verlag, 2005.

[49] T. Bäck, D. B. Fogel, and Z. Michalewicz, "Handbook of Evolutionary Computation ": Oxford Univ. Press., 1997.

[50] D. Ashlock, *Evolutionary Computation for Modeling and Optimization*: Springer Verlag, 2006.

[51] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press, 1975.

[52] R. L. Haupt and S. E. Haupt, "Practical Genetic Algorithms," 2, Ed.: John Wiley, 2004.

[53] Z. Michalewicz, *Genetic Algorithm + Data Structure = Evolution Programs*, 2 ed. New York: Springer, 1995.

[54] L. J. Eshelman and D. J. Shaffer, "Real-coded genetic algorithms and interval schemata " in *Foundation of Genetic Algorithms* D. L. Whitley, Ed. San Mateo, CA: Morgan Kaufmann, 1993.

[55] M. Dorigo, V. Maniezzo, and A. Colorni, "The Ant System: Optimization by a Colony of Cooperating Agents," *IEEE Trans. Syst. Man and Cybern. Part B-Cybernatics*, vol. 26, no. 1, pp. 1-13, 1996.

[56] F. O. d. Franca, G. P. Coelho, F. J. v. Zuben, and R. R. d. F. Attux, "Multivariate Ant Colony Optimization in Continuous Search Spaces," in *GECCO'08*, Atlanta, Georgia, USA, 2008, pp. 9-16.

[57] K. Socha and M. Dorigo, "Ant colony optimization for continuous domains," *Euro. J. of Oper. Res.*, vol. 185, pp. 1155-1173, 2008.

[58] R. Storn and K. Price, "Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization* vol. 11, pp. 341–359, 1997.

[59] J. Liu and J. Lampinen, "On setting the control parameter of the differential evolution method," in *the 8th International Conference on Soft Computing (MENDEL)*, Brno, Czech Republic, 2002.

[60] K. Price, R. Storn, and J. Lampinen, *Differential Evolution - A Practical Approach to Global Optimization*: Springer Verlag, 2005.

[61] D. Zaharie, "Critical values for the control parameters of differential evolution algorithms," in *8th International Mendel Conference on Soft Computing*, Brno, Czech Republic, 2002.

[62] J. Brest, S. Greiner, B. Boskovic, M. Mernik, and V. Zumer, "Self-Adpting Control Parameters in Differential Evolution: A Comprehensive Study on Numerical Benchmark Problems," *IEEE Trans. Evol. Compt.*, vol. 10, no. 6, pp. 646-657, Dec. 2006.

[63] A. K. Qin, V. L. Huang, and P. N. Suganthan, "Differential Evolution Algorithm With Strategy Adaptation for Global Numerical Optimization," *IEEE Trans. Evol. Compt.*, vol. 13, no. 2, pp. 398-417, April 2009.

[64] J. Brest and M. S. Maucec, "Population size reduction for the differential evolution algorithm," *Applied Intelligence*, vol. 29, pp. 228-247, 2008.

[65] N. S. Teng, J. Teo, and M. H. A. Hijazi, "Self-adaptive population sizing for a tune-free differential evolution," *Soft Computing*, vol. 13, pp. 709-724, 2009.

[66] J. Teo, "Exploring dynamic self-adaptive populations in differential evolution," *Soft Computing*, vol. 10, pp. 673-686, 2006.

[67] P. N. Suganthan, N. Hansen, J. J. Liang, K. Deb, Y.-P. Chen, A. Auger, and S. Tiwari, "Problem Definitions and Evaluation Criteria for the CEC 2005: Special Session on Real-Parameter Optimization," IEEE CEC 2005 May 2005.

[68] J. J. Liang, P. N. Suganthan, and K. Deb, "Novel composition test functions for numerical optimization," in *IEEE Swarm Intelligence Symposium*, Pasadena, CA, USA, June 2005.

[69] W. C. Zhong, J. Liu, M. Z. Xue, and L. C. Jiao, "A multiagent genetic algorithm for global numerical optimization," *IEEE Trans. Syst. Man and Cybern. Part B-Cybernatics*, vol. 34, pp. 1128-1141, April 2004.

[70] S. Siegel, *Nonparameteric Statistics for The Behaviorial Sciences*: McGraw-Hill Kogakusha, Ltd, 1956.

[71] Z. Michalewicz, "A Survey of Constraint Handling Techniques in Evolutionary Computation Methods," in *the 4th Annual Conference on Evolutionary Programming*, pp. 135-155.

[72] E. Tsang, *Foundation of Constraint Satisfaction*. London and San Diego: Academic Press, 1993.

[73] O. Yeniay, "Penalty Function Methods for Constrained Optimization with Genetic Algorithms," *Math. and Compt. Appli.*, vol. 10, no. 1, pp. 45-56, 2005.

[74] A. Homaifar, S. H. Y. Lai, and X. Qi, "Constrained optimization via genetic algorithms," *Simulation*, vol. 62, pp. 242-254, 1994.

[75] Z. Michalewicz, "Genetic algorithms, numerical optimization and constraints," in *The sixth international conference on genetic algorithms*, 1996, pp. 151-158.

[76] B.Tessema and G. G. Yen, "A self adaptive penalty function based algorithm for constrained optimization," in *Proc. IEEE Congr. Evolut. Comput.*, July 2006, pp. 246-253.

[77] D. J. Sobajic and Y. H. Pao, "Artificial Neural-Net Based Dynamic Security Assessment for Electric Power Systems," *IEEE Trans. Power Syst.*, vol. 4, no. 1, pp. 220-228, Feb. 1989.

[78] B. Jayasekara and U. D. Annakkage, "Transient Security Assessment using Multivariate Polynomial Approximation," *Elect. Power Syst. Res.*, vol. 77, pp. 704-711, 2007.

[79] H. Sawhney and B. Jeyasurya, "A feed-forward artificial neural network with enhanced feature selection for power system transient stability assessment," *Elect. Power Syst. Res.*, vol. 76, pp. 1047-1054, 2006.

[80] B. A. Archer, U. D. Annakkage, B. Jayasekara, and P. Wijetunge, "Accurate Prediction of Damping in Large Interconnected Power System with the Aid of Regression Analysis," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1170-1178, Aug. 2008.

[81] S. Chauhan and M. P. Dave, "Kohonen Neural Network Classifier for Voltage Collapse Margin Estimation," *Elect. Power Comp. and Syst.*, vol. 25, pp. 607-619, 1997.

[82] M. M. Salama, E. M. Saied, M. M. Abou-Elsaad, and E. F. Ghariany, "Estimating the voltage collapse proximity indicator using artificial neural network," *Energy Conv. & Mang.*, vol. 42, pp. 69-79, 2001.

[83] C. Rehtanz, "Visualization of voltage stability in large electric power systems," *IEE Proc. Gener. Transm. Distr.,* vol. 146, no. 6, pp. 573-576, Nov. 1999.

[84] S. Kamalasadan, A. K. Srivastava, and D. Thukaram, "Novel Algorithm for Online Voltage Stability Assessment Based on Feed Forward Neural Network," in *IEEE PES Gen. Meeting*, 2006.

[85] T. S. Chung and Y. Fu, "A Sequential Feature Selection Based Neural Network Approach to Dynamic Voltage Stability Estimation," *IEEE Power Engr. Review Letters*, pp. 51-55, Jan. 2002.

[86] V. R. Dinavahi and S. C. Srivastava, "ANN Based Voltage Stability Margin Prediction," in *IEEE Power Engr. Soc. Summer Meetings*, July 2001, pp. 1275-1280.

[87] A. A. El-Keib and X. Ma, "Application of Arificial Neural Networks in Voltage Stability Assessment," *IEEE Trans. Power Syst.,* vol. 10, no. 4, pp. 1890-1896, Nov. 1995.

[88] C. W. Liu, C.-S. Chang, and M.-C. Su, "Neuro-Fuzzy Networks for Voltage Security Monitoring Based on Synchronized Phasor Measurements," *IEEE Trans. Power Syst.,* vol. 13, no. 2, pp. 326-332, May 1998.

[89] M. L. Scala, M. Trovato, and F. Torelli, "A Neural Network-Based Method for Voltage Security Monitoring," *IEEE Trans. Power Syst.,* vol. 11, no. 3, pp. 1332-1341, Aug. 1996.

[90] D. Q. Zhou, U. D. Annakkage, and A. D. Rajapakse, "Online Monitoring of Voltage Stability Margin using an Artificial Neural Network," *IEEE Trans. Power Syst.,* vol. 25, no. 3, pp. 1566-1574, Aug. 2010.

[91] S. Chakrabarti and B. Jeyasurya, "Multicontingency Voltage Stability Monitoring of Power Systems Using Radial Basis Function Networks," in *ISAP05*, 2005, pp. 314-319.

[92] A. Maiorano and M. Trovato, "A neural network-based tool for preventing control of voltage stability in multi-area power system," *Neurocomputing*, vol. 23, pp. 161-176, 1998.

[93] S. Chakrabarti and B. Jeyasurya, "On-line Voltage Stability Monitoring Using Artificial Neural Network," in *Power Engineering, Large Engineering Syst. Conf. (LESCOPE-04)*, 2004, pp. 71-75.

[94] D. Niebur and A. J. Germond, "Power System Static Security Assessment using the Kohonen Neural Network Classifier," *IEEE Trans. Power Syst.*, pp. 270-277, 1991.

[95] N. J. Higham, "Computing the Nearest Correlation Matrix- A Problem from Finance," School of Mathematics, University of Manchester, Manchester, UK 2002.

[96] K. Yang and J. Trewn, "Chp. 3 Introduction to Multivariate Random Variables, Normal Distribution, and Sampling Properties," in *Multivariate Statistical Methods in Quality Management*, 1st ed: McGraw Hill Professional, 2004.

[97] H. Saadat, *Power System Analysis*, 1st International ed. Singapore: WCB/ McGraw Hill, 1999.

[98] F. Milano, L. Vanfrett, and J. C. Morataya, "An Open Source Power System Virtual Laboratory: The PSAT Case and Experience," *IEEE Trans on Education*, vol. 51, no. 1, pp. 17-23, Feb. 2008.

[99] H. W. Dommel and W. F. Tinney, "Optimal Power Flow Solutions," *IEEE Trans. Power App. Syst.*, vol. PAS-87, no. 10, pp. 1866-1876, Oct. 1968.

[100] M. R. AlRashidi and M. E. El-Hawary, "Applications of computational intelligence techniques for solving the revived optimal power flow problem," *Elect. Power Syst. Res.*, vol. 79, pp. 694-702, 2009.

[101] X. Xia and A. M. Elaiw, "Optimal dynamic economic dispatch of generation: A review," *Elect. Power Syst. Res.*, vol. 80, pp. 975-986, 20101.

[102] B. H. Chowdhury and S. Rahman, "A review of recent advances in economic dispatch," *IEEE Trans. Power Syst.*, vol. 5, no. 4, pp. 1248-1259, 1990.

[103] H. R. Cai, C. Y. Chung, and K. P. Wong, "Application of Differential Evolution Algorithm for Transient Stability Constrained Optimal Power Flow," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 719-728, May 2008.

[104] D. Gan, R. J. Thomas, and R. D. Zimmermann, "Stability-Constrained Optimal Power Flow," *IEEE Trans. Power Syst.*, vol. 15, no. 2, pp. 535-540, May 2000.

[105] R. Wang and R. H. Lasseter, "Re-dispatching Generation to Increase Power System Security Margin and Support Low Voltage Bus," *IEEE Trans. Power Syst.*, vol. 15, no. 2, pp. 496-501, May 2000.

[106] W. J. Tang, M. S. Li, Q. H. Wu, and J. R. Saunders, "Bacterial Foraging Algorithm for Optimal Power Flow in Dynamic Environment," *IEEE Trans. Cir. and Syst.-I: Regular Papers*, vol. 55, no. 8, pp. 2433-2442, Sep. 2008.

[107] D. C. Walters and G. B. Sheble, "Genetic algorithm solution of economic dispatch with valve point loading," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 1325-1331, Aug. 1993.

[108] N. Sinha, R. Chakrabarti, and P. K. Chattopadhyay, "Evolutionary programming techniques for economic load dispatch," *IEEE Trans. Evolut. Comput.*, vol. 7, no. 1, pp. 83-94, Feb. 2003.

[109] C. L. Chiang, "Improved genetic algorithm for power economic dispatch of units with valve-point effects and multiple fuels," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1690-1699, 2005.

[110] A. Pereira-Neto, C. Unsihuay, and O. R. Saavedra, "Efficient evolutionary strategy optimisation procedure to solve the nonconvex economic dispatch problem with generator constraints," *IEE Gener. Transm. Distrib.*, vol. 152, no. 5, pp. 653-660, Sep. 2005.

[111] S. O. Orero and M. R. Irving, "Economic dispatch of generators with prohibited operating zones: a genetic algorithm approach " *IEE Gener. Transm. Distrib.*, vol. 143, no. 6, pp. 529 - 534 Nov. 1996.

[112] Z. L. Gaing, "Particle swarm optimization to solving the economic dispatch considering the generator constraints," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1187-1195, Aug. 2003.

[113] T. Amraee, A. M. Ranjbar, B. Mazafari, and N. Sadati, "An enhanced under-voltage load-shedding scheme to provide voltage stability," *Elec. Power Syst. Research*, vol. 77, pp. 1038-1046, 2007.

[114] C. M. Affonso, L. C. P. d. Silva, F. G. M. Lima, and S. Soares, "MW and MVar Management on Supply and Demand Side for Meeting Voltage Stability Margin Criteria," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1538-1545, Aug. 2004.

[115] L. D. Arya, V. S. Pande, and D. P. Kothari, "A technique for load-shedding based on voltage stability consideration," *Elec. Power and Ener. Syst.*, vol. 27, pp. 506-517, 2005.

[116] L. Shi and K. Rasheed, "A survey of fitness approximation methods applied in evolutionary algorithms," in *Computational Intelligence in Expensive Optimization Problems*, Y. Tenne and C.-K. Goh, Eds. Heidelberg: Springer Verlag 2010, pp. 3-28.

[117] Y. Jin and J. Branke, "Evolutionary Optimization in Uncertain Environments - A Survey," *IEEE Trans. Evol. Compt.*, vol. 9, no. 3, pp. 303-317, June 2005.

[118] Y. Jin, "A Comprehensive Survey of Fitness Approximation in Evolutionary Computation," *Soft Computing*, vol. 9, no. 1, pp. 3-12, 2005.

[119] C. A. Canizares, "Applications of Optimization to Voltage Collapse Analysis," in *IEEE/PES Summer Meeting*, San Diego, USA, July 1998.

[120] G. Y. Wu, C. Y. Chung, K. P. Wong, and C. W. Yu, "Voltage stability constrained optimal dispatch in deregulated power systems," *IEE Genr. Transm. Distrb.*, vol. 1, no. 5, pp. 761-768, 2007.

[121] H. Song, B. Lee, and Y.-H. Moon, "Reactive optimal power flow incorporating margin enhancement constraints with nonlinear interior point method," *IEE Genr. Transm. Distrb.*, vol. 152, no. 6, pp. 961-967, Nov. 2005.

[122] K. Meng, H. G. Wang, Z. Y. Dong, and K. P. Wong, "Quantum-inspired particle swarm optimization for valve-point economic load dispatch," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 215-222, Feb. 2010.

[123] X. Lin, A. K. David, and C. W. Yu, "Reactive power optimization with voltage stability consideration in power market systems," *IEE Genr. Transm. Distrb.*, vol. 150, no. 3, pp. 305-310, May 2003.

[124] Y.-J. Zhang and Z. Ren, "Optimal Rective Power Dispatch Considering Costs of Adjsting the Control Devices," *IEEE Trans. Power Syst.*, vol. 20, no. 3, pp. 1349-1356, Aug. 2005.

[125] S. A. Soman, K. Parthasarathy, and D. Thukaram, "Curtailed Number and Reduced Controller Movement Optimization Algorithms for Real-Time Voltage/ Reactive Power Control," *IEEE Trans. Power Syst.*, vol. 9, no. 4, pp. 2035-2041, Nov. 1994.

[126] P. J. Balducci, J. M. Roop, L. A. Schienbein, J. G. Desteese, and M. R. Weimar, "Electrical power interruption cost estimates for individual industries, sectors and U.S. economy," Pacific Northwest National Lab Feb. 2002.

Publication List

Papers published in conferences

1. **W.Nakawiro** and I. Erlich,“ Online Voltage Stability Monitoring using Artificial Neural Network”, In proceeding International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, April 2008 Nanjing, China
2. **W.Nakawiro** and I. Erlich,“ A Combined GA-ANN Strategy for Solving Optimal Power Flow with Voltage Security Constraint”, In proceeding IEEE Asia-Pacific Power and Energy Engineering Conference APPEEC, March 2009 Wuhan, China
3. **W.Nakawiro** and I. Erlich,“ Voltage Security Assessment and Control System using a Hybrid Intelligent Method”, In proceeding PowerTech 2009 IEEE Power & Energy Society, June 2009 Bucharest, Romania
4. **W.Nakawiro** and I. Erlich,“ Optimal Load Shedding for Voltage Stability Enhancement by Ant Colony Optimization,” In proceeding Intelligence System Application to Power Systems (ISAP) 2009, November 2009 Curitiba, Brazil
5. V.S. Pappala, **W.Nakawiro** and I. Erlich, “Preventive Optimal Control of Wind Farm Reactive Sources,” In Proceeding 2010 Transmission and Distribution Conference and Exposition, New Orleans, USA, April 2010
6. **W.Nakawiro** and I. Erlich, “A Hybrid Method for Voltage Stability Constrained Optimal Reactive Power Dispatch”, In Proceeding IEEE PES General Meeting, July 26 - July 29, 2010, Minneapolis, Minnesota, USA.
7. I. Erlich, G.K. Venayagamoorthy, and **W.Nakawiro**, “A Mean-Variance Optimization Method”, In Proceeding 2010 IEEE Congress on Evolutionary Computation,” July 26 - July 29, 2010,Barcelona, Spain

Papers under preparation

8. **W.Nakawiro** and I. Erlich, “Effective Constraint Handling in Evolutionary Algorithm for Non-convex Economic Dispatch”, in the final revision before submitting to IET Proceeding in Generation, Transmission and Distribution

9. **W.Nakawiro** and I. Erlich, "A parameter-less differential evolution method for voltage stability constrained optimal power flow", submitted to 2011 Power System Computation Conference (PSCC 2011)
10. **W. Nakawiro,** I. Erlich and Y. Nithipattrarat, "Effects of Equality Constraint Handling on Economic Dispatch using Differential Evolution Algorithms", submitted to ISAP 2011, Greece
11. **W.Nakawiro**, J. Rueda, I. Erlich, "A Mean-Variance Mapping Optimization Method for Optimal Power Flow: A Comparative Study", submitted to 2011 International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT 2011)
12. I. Erlich, **W.Nakawiro** and M. Martínez, "Optimal Dispatch of Reactive Sources in Wind Farms", submitted to 2011 IEEE PES General Meeting (PESGM 2011)

Curriculum Vitae

Personal Data

Name:	Worawat Nakawiro
Gender:	Male
Date of birth:	16th March 1981
Place of birth:	Bangkok, Thailand
Nationality:	Thai
Marital status:	Single
Permanent address:	4/309 Lane 14, Sahakorn village, Serithai Road, Klongkum, Bungkum, Bangkok, 10240, Thailand

School education

1988-1993	Primary school Anuban Piboonwej School, Bangkok, Thailand
1993-1998	Secondary school Bodin Decha (Sing Singhasaenee) School, Bangkok, Thailand

Study

1998-2002	Undergraduate student Department of Electrical Engineering, Faculty of Engineering, Thammasat University, Bangkok, Thailand
2002-2004	Graduate student Electric Power System Management Program, Energy Field of Study Asian Institute of Technology, Pathumthani, Thailand
2006-Present	Research assistant Institute of Electric Power Systems, University of Duisburg-Essen, Duisburg, Germany

Professional experience

04.2002-10.2002	Full-time Lecturer School of Engineering, Bangkok University, Bangkok, Thailand (under the Faculty Development Program)
10.2002-03.2003	Part-time Instructor Department of Electrical Engineering, Thammasat University, Bangkok, Thailand
08.2004-03.2006	Research Associate Energy Field of Study, Asian Institute of Technology, Pathumthani, Thailand