Neurocomputing 102 (2013) 154–162

Contents lists available at SciVerse ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Extreme learning machine based genetic algorithm and its application in power system economic dispatch

Hongming Yang [a], Jun Yi [a], Junhua Zhao [b,*], ZhaoYang Dong [c]

[a] School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410114, China
[b] School of Electrical Engineering, Zhejiang University, Hangzhou, Zhejiang, China
[c] Centre for Intelligent Electricity Networks, The University of Newcastle, Australia

ARTICLE INFO

Available online 6 June 2012

Keywords:
Extreme learning machine
Genetic algorithm
Power system economic dispatch

ABSTRACT

In this paper a novel optimization algorithm, which utilizes the key ideas of both genetic algorithm (GA) and extreme learning machine (ELM), is proposed. Traditional genetic algorithm employs genetic operations, such as selection, mutation and crossover to generate the optimal solution. In practice, the child solutions generated by crossover and mutation are largely random and therefore cannot ensure the fast convergence of the algorithm. To tackle the weakness of traditional GA, the ELM is introduced to estimate the nonlinear functional relationships between the parent population and child population generated by genetic operations. The trained downward-climbing and upward-climbing ELMs are then employed to generate candidate solutions, which forms the new population together with the solutions given by genetic operations. The proposed algorithm is applied to the power system economic dispatch problem. As demonstrated in case studies, the modified genetic algorithm is able to locate local minima faster and escape from local minima with a greater probability. The proposed algorithm can therefore ensure the faster convergence and provide more economical dispatch plans.

1. Introduction

The genetic algorithm (GA) is a well-known random search and global optimization method based on the idea of natural selection and evolution. Rather than relying on the gradient information, it searches the optimal solution by simulating the natural evolution process. GA is proven to be a suitable method for solving large scale optimization problems which are nonlinear, non-convex and non-continuous. It has demonstrated several significant advantages, such as strong robustness, convergence to global optimum and parallel search capability. On the other hand, traditional GA may also suffer the problems of premature convergence and slow global convergence [1]. To overcome these weaknesses, many researchers have conducted in-depth studies on the genetic operators, control parameters and procedure of traditional GA, and consequently proposed a number of modified algorithms. In [2], a modified GA is introduced based on the theory of biology immunity; by utilizing the concepts of immunity and vaccination, the proposed algorithm is able to effectively avoid the degenerative phenomenon in the evolution process. In [3], a novel algorithm is proposed to adaptively adjust the probabilities of mutation and crossover based on the current fitness value. By using the means of adaptive mutation and crossover, the proposed algorithm has a faster convergence speed; it however suffers from the premature convergence problem. In [4], a novel algorithm is proposed by integrating genetic algorithm and simulated annealing. Due to the strong local search capability of simulated annealing, the modified GA has a faster convergence speed; the adoption of simulated annealing however incurs stricter restrictions on the optimization problem.

The extreme learning machine (ELM) is a new algorithm for training feed-forward neural networks proposed by Guangbin Huang [5,7–16]. By adjusting the weights of connections between its nodes, neural networks are able to estimate the nonlinear functional relationships between the input and output variables. Different from traditional training algorithms for neural networks, ELM randomly determines the connection weights between the input layer and hidden layer, and obtains the connection weights between the hidden layer and output layer analytically. Therefore, ELM can effectively avoid the slow training speed and over-fitting problems suffered by traditional neural network training algorithms [5–16]. An evolutionary approach is proposed in [7] to utilize the GA search capability to enhance ELM. In this paper, because of its superior computational efficiency and generalization ability, we employ ELM to develop a novel evolution strategy. Based on the new evolution strategy, we propose a modified genetic algorithm. The proposed algorithm employs ELM to estimate the complex nonlinear relationships between the parent and child populations generated by genetic operators such as crossover and mutation. The trained ELMs are

* Corresponding author.
E-mail address: fuxiharp@gmail.com (J. Zhao).

0925-2312/$ - see front matter
http://dx.doi.org/10.1016/j.neucom.2011.12.054

then used to adjust the evolution direction of GA, thus improve its convergence speed and global search capability.

Power system economic dispatch (ED) is a large-scale, high-dimensional and nonlinear optimization problem, which aims at minimizing the system generation cost subject to a variety of technical and economical constraints [17]. Traditionally, economic dispatch problems are solved by mathematical programming methods, which usually fail to locate the global optimum due to the nonconvexity of the economic dispatch problem. Also, mathematical programming places strict restrictions on the problem; for example, the problems must be continuous and differentiable. On the contrary, genetic algorithm is well-known for its global search capability, and is suitable for solving discontinuous and nondifferentiable problems. Therefore, in this paper we apply the proposed ELM based genetic algorithm in the power system ED problem. Case study results demonstrate that the modified GA can effectively solve the ED problem.

2. Extreme learning machine

Extreme learning machine is a novel algorithm for training single-layer feed-forward neural networks [5,7–15]. The topological structure of an ELM network can be shown in Fig. 1.

Denote the numbers of nodes in the input, hidden and output layers of the network as U,R and O. Given the input vector $\mathbf{x}=[x_1,x_2,\ldots,x_U]^T$, the output $\mathbf{y}=[y_1,y_2,\ldots,y_O]^T$ of a single-layer feed-forward neural network can be expressed as [5]

$$y_o=\sum_{i=1}^{R}\beta_{io}\,G_i(\boldsymbol{\omega}_i,b_i,\mathbf{x})\quad (o=1,2,\ldots,O) \tag{1}$$

where $\boldsymbol{\omega}_i=[\omega_{1i},\omega_{2i},\ldots,\omega_{Ui}]$ represents the connection weights between the input layer and ith node in the hidden layer; b_i is the bias of the ith hidden node; $\boldsymbol{\beta}_i=[\beta_{i1},\beta_{i2},\ldots,\beta_{iO}]$ represents the connection weights between the ith node in the hidden layer and the output layer; $G_i(\boldsymbol{\omega}_i,b_i,\mathbf{x})$ is the output of the ith hidden node. We have $G_i(\boldsymbol{\omega}_i,b_i,\mathbf{x})=g(\boldsymbol{\omega}_i\cdot\mathbf{x}+b_i)$, where $g(\cdot)$ is the activation function.

The ELM network uses the linear activation function in the input and output layers, and employs the nonlinear activation function in the hidden layer, such as Sigmoid function, Sine function, Hard Limit function, Triangular basis function and Radial basis function [16]. Among them Sigmoid function is widely used and can be expressed as

$$g(s)=\frac{1}{1+\exp(-s)} \tag{2}$$

where s is the input of the hidden node, and $g(s)$ is the output.

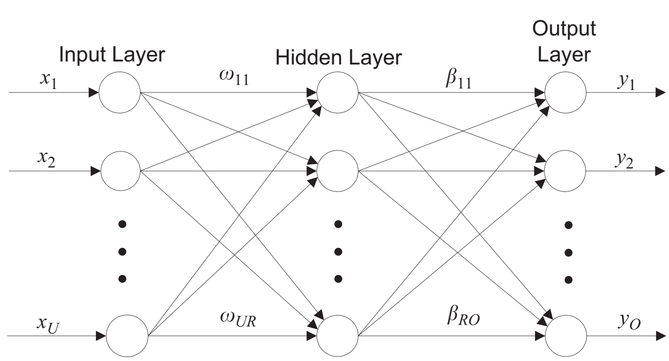

Fig. 1. The topological structure of the ELM network.

The objective of training ELM is to minimize the empirical error and structural error by assigning the best weights $\boldsymbol{\omega}_i,\boldsymbol{\beta}_i$. Therefore, given the training data set $(\mathbf{X}_j,\mathbf{Y}_j),\cdot j=1,2,\ldots,W$, the training of ELM is a nonlinear optimization problem, whose objective can be given as

$$\min E(\boldsymbol{\omega}_i,\boldsymbol{\beta}_i)=\sum_{j=1}^{W}\|\mathbf{y}_j-\mathbf{Y}_j\| \tag{3}$$

where $\mathbf{y}_j=[y_{j1},y_{j2},\ldots,y_{jO}]^T$ is the outputs of the ELM network given inputs $\mathbf{X}_j=[x_{j1},x_{j2},\ldots,x_{jU}]^T$; $\mathbf{Y}_j=[Y_{j1},Y_{j2},\ldots,Y_{jO}]^T$ represents the real values of the corresponding dependent variables.

When weights $\boldsymbol{\omega}_i$ and bias b_i are randomly assigned, solving optimization problem (3) is equivalent to solving Eq. (4) for its least square solution $\boldsymbol{\beta}$. Based on Moore Penrose's generalized inverse matrix theory, the weights can be solved analytically as $\boldsymbol{\beta}=\mathbf{G}^{+}\cdot\mathbf{Y}$, where $\mathbf{G}^{+}$ represents the generalized inverse matrix of the hidden layer output matrix $\mathbf{G}$ of the ELM network [5,9].

$$\mathbf{G}\boldsymbol{\beta}=\mathbf{Y} \tag{4}$$

$$\mathbf{G}=\begin{bmatrix} G_1(\boldsymbol{\omega}_1,b_1,\mathbf{X}_1),\ldots G_R(\boldsymbol{\omega}_R,b_R,\mathbf{X}_1) \\ \ldots \\ G_1(\boldsymbol{\omega}_1,b_1,\mathbf{X}_W)\ldots G_R(\boldsymbol{\omega}_R,b_R,\mathbf{X}_W)\end{bmatrix}_{W*R} \tag{5}$$

$$\boldsymbol{\beta}=\begin{bmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\\ \ldots\\ \boldsymbol{\beta}_R\end{bmatrix}_{R\times O},\quad \mathbf{Y}=\begin{bmatrix}\mathbf{Y}_1{}^T\\ \mathbf{Y}_2{}^T\\ \ldots\\ \mathbf{Y}_W^T\end{bmatrix}_{W\times O}, \tag{6}$$

3. Modified genetic algorithm based on ELM

3.1. Basic genetic algorithm

Genetic algorithm refers to a class of random search methods for global optimization. It uses a population of strings to encode the initial candidate solutions. GA then employs genetic operators (selection, mutation, crossover) to generate new populations based on the initial population, and gradually evolves towards the best solution [17–19]. The convergence speed of GA is closely related to the procedure and parameters of the genetic operators such as selection, mutation and crossover. The procedure of basic GA is illustrated in Fig. 2.

In GA, the aim of selection operator is to select the candidate solutions with higher fitness values to be inherited in the next iteration. The crossover operator exchanges some chromosomes of the parent solutions with crossover probability P_c to obtain child solutions. Generally speaking, the selection and crossover operators enhance GA's ability to search local optima. On the other hand, the mutation operator randomly changes the chromosomes of some candidate solutions with mutation probability P_m. The mutation operator can increase the diversity of the population, thus improving the global search capability of the algorithm [19].

In practice, the solutions generated by crossover and mutation are largely random, therefore the convergence speed cannot be guaranteed. Also, when GA reaches the neighborhood of the global optimum, it will easily suffer from the premature convergence problem. To tackle these disadvantages of basic GA, ELM is employed to find better evolution directions. The ELM is firstly trained to estimate the nonlinear functional relationship between the parent and child populations [10–12]. It is then used to adjust the search direction of GA and improve its convergence property.

3.2. Modified genetic algorithm based on ELM

Based on the nonlinear function estimation capability of ELM and the concepts of genetic algorithm, we propose two new evolution

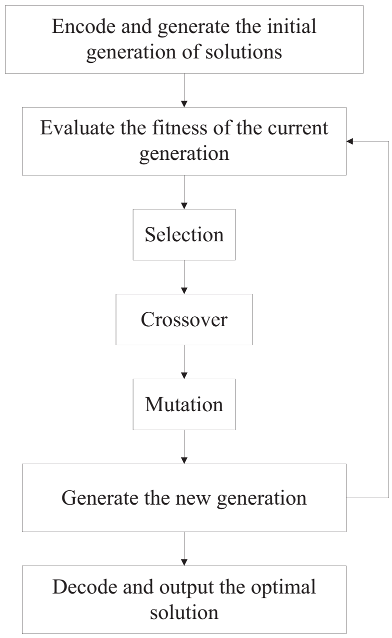

Fig. 2. The procedure of basic genetic algorithm.

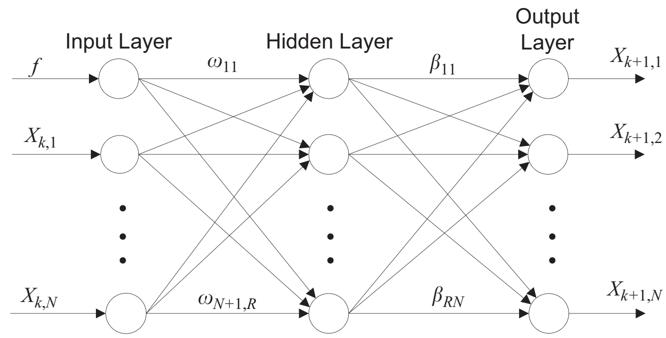

Fig. 3. Structure of the downward-climbing ELM network.

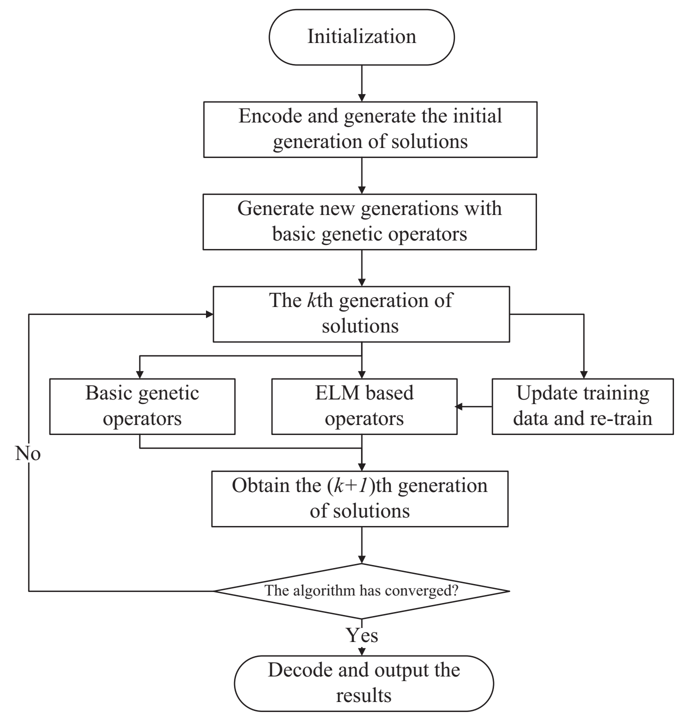

Fig. 4. The main procedure of the modified genetic algorithm.

operators based on ELM. As discussed above, the selection and crossover operators are related with the local search capability of GA, while the mutation operator aims at improving the global search capability. Similarly, the two ELM based operators are proposed to further improve the local and global search capabilities [12].

The first ELM based operator is named as the *downward-climbing* operator. The downward-climbing ELM estimates the functional relationship between the parent solutions and their child solutions that have highest fitness values and are generated by selection and crossover. Because selection and crossover are local search operators, the search direction given by downward-climbing ELM generally will be towards the local optimum. Also, since it is trained by the candidate solutions with highest fitness values, it will speed up the convergence to local optima compared with basic selection and crossover operators. Assume that in each iteration, N candidates, which have highest fitness values and are generated by selection and crossover, will be selected to train the downward-climbing ELM. Also, assume that the number of candidates generated by the downward-climbing ELM operator is M_d. Denote the candidate solutions, which are generated by the selection and crossover operators, in the kth and $(k+1)$th iterations as $X_{k,1},\ldots,X_{k,N}$ and $X_{k+1,1},\ldots,X_{k+1,N}$. f_k and f_{k+1} represent the highest fitness values of $X_{k,1},\ldots,X_{k,N}$ and $X_{k+1,1},\ldots,X_{k+1,N}$, respectively. The structure of the downward-climbing ELM network is illustrated in Fig. 3. The downward-climbing ELM will then be trained to estimate the functional relationship between $X_{k+1,1},\ldots,X_{k+1,M_d}$ and $f_k, X_{k,1},\ldots,X_{k,N}$. Once the weights $\boldsymbol{\omega}_i, \boldsymbol{\beta}_i$ of ELM network are known, the outputs of ELM operator $X_{k+1,1},\ldots,X_{k+1,M_d}$ can be expressed as

$$X_{k+1,o} = \sum_{i=1}^{R} \beta_{io} g\left(\omega_{i1} f + \sum_{j=2}^{N+1} \omega_{ij} X_{k,j} + b_i\right) \quad (o = 1,2,\ldots,M_d) \tag{7}$$

The second ELM based operator is called the *upward-climbing* operator. Similarly, the upward-climbing ELM will be trained based on the parent solutions and their child solutions that have the lowest fitness values and are generated by the mutation operator. The mutation operator randomly generates candidate solutions to enhance the population diversity and help GA escape from local optima. Therefore, by selecting the solutions, which are generated by mutation and have lowest fitness values, as the training instances, the upward-climbing operator can escape from local optima with a faster speed. The structure of the upward-climbing ELM network is similar to the downward-climbing ELM network as illustrated in Fig. 3.

The proposed modified genetic algorithm (as illustrated in Fig. 4) will integrate upward-climbing and downward-climbing operators with basic genetic operators to generate candidate solutions. In the first k iterations of the modified GA, only basic genetic operators (selection, crossover, mutation) will be employed. Meanwhile, the upward-climbing and downward-climbing ELMs will be trained with the first k iterations of candidate solutions. When training the two ELMs, the solutions from iteration i, $1 \le i \le k$ will be selected as the inputs (independent variables), while the solutions from iteration $i+q$ will be the outputs (dependent variables). In other words, the two ELMs are trained to estimate the nonlinear relationships between the parent solutions and their inter-iteration child solutions. Here q is the inter-iteration parameter of the modified GA which can control the convergence speed of two ELM operators.

After the kth iteration, the ELM operators will also be employed to generate candidate solutions, so as to change the search direction and the search range. In each iteration, the numbers of candidate solutions generated by selection, crossover and downward-climbing operators will be determined by their average fitness values in the last iteration. Also, the two ELMs will be re-trained continuously to ensure their effectiveness [13]. Denote M_s,M_c,M_m,M_d,M_u as the numbers of solutions generated by selection, crossover, mutation, downward-climbing and upward-climbing operators respectively in each iteration; let $M = M_s + M_c + M_d$ be the total number of solutions generated by selection, crossover and downward-climbing operators; K represents the maximum iteration number; ε represents the convergence criterion. The procedure of the modified GA can then be given as follows:

(1) Initialize algorithm parameters M,M_m,M_u,K,ε; these parameters will be fixed during the entire optimization process.
(2) Set iteration index $i = 0$, encode and generate the initial iteration X_0.
(3) When $i \le k$, set $M_d(i) = 0$ and $M_s(i) = M_c(i) = M/2$; employ selection, crossover and mutation operators to generate candidate solutions; meanwhile train downward-climbing and upward-climbing ELMs.
(4) When $i > k$, set $M_s(i) = M_c(i) = M_d(i) = M/3$.
(5) Employ both basic genetic operators and ELM based operators to generate solutions.
(6) Assume that in iteration i, the average fitness values of the solutions generated by selection, crossover and downward-climbing operators are respectively $f_s(i)$,$f_c(i)$,$f_d(i)$; then in iteration $i+1$, the numbers of solutions generated by the three operators can be recalculated as

$$M_s(i+1) = M f_s(i)/(f_s(i) + f_c(i) + f_d(i))$$

$$M_c(i+1) = M f_c(i)/(f_s(i) + f_c(i) + f_d(i))$$

$$M_d(i+1) = M f_d(i)/(f_s(i) + f_c(i) + f_d(i))$$

(7) Calculate the fitness difference $\varepsilon(i)$ between the best solutions in iteration i and $i+1$; repeat steps (5) and (6) if $\varepsilon(i) \ge \varepsilon$ and $i \le K$; otherwise decode and output the results.

4. Power system economic dispatch model considering coal-fired and hydro plants

4.1. The probabilistic analysis of the water inflow of single hydro plant

In power system analysis, the Pearson type III (P-III) distribution is widely employed to model the water inflow of a single hydro plant. Its probability density function can be given as [20]

$$\phi(I) = \begin{cases} \frac{1}{\beta^{\alpha}\Gamma(\alpha)}(I-\delta)^{\alpha-1}e^{-(I-\delta)/\beta} & \delta \le I < +\infty \\ 0 & -\infty < I < \delta \end{cases} \tag{8}$$

where I denotes the water inflow; $\Gamma(\alpha) = \int_0^{\infty} e^{-t}t^{\alpha-1}\,dt$ is the Gamma function; α,β,δ are the parameters of the P-III distribution and can be estimated with the Maximum Likelihood Estimation (MLE).

4.2. The joint distribution of the water inflows of multiple hydro plants

Multiple hydro plants in an area can be influenced by similar weather conditions; their water inflows thus are not independent. Therefore, it is necessary to consider the correlations between the inflows of multiple plants and accurately estimate their joint distribution. The Copula function can represent the joint distribution of multiple water inflows with their marginal distributions [21]. The joint distribution of multiple hydro plant inflows can be illustrated in Fig. 5.

As clearly observed in Fig. 5, the joint distribution of water inflows exhibits the upper fat-tail property. Therefore, it is appropriate to be modeled with the Gumbel-Copula function as follows [22]

$$\begin{aligned}\Phi(I_1,\ldots,I_\Omega) &= C(F_1(I_1),\ldots,F_\Omega(I_\Omega)) \\ &= \exp\left\{-[(-\ln F_1(I_1))^{\theta} + \cdots + (-\ln F_\Omega(I_\Omega))^{\theta}]^{1/\theta}\right\}\end{aligned} \tag{9}$$

where $\Phi(\cdot)$ represents the joint distribution; I_1,I_2,…,I_Ω are the water inflows of each plant, and all follow the P-III distribution;

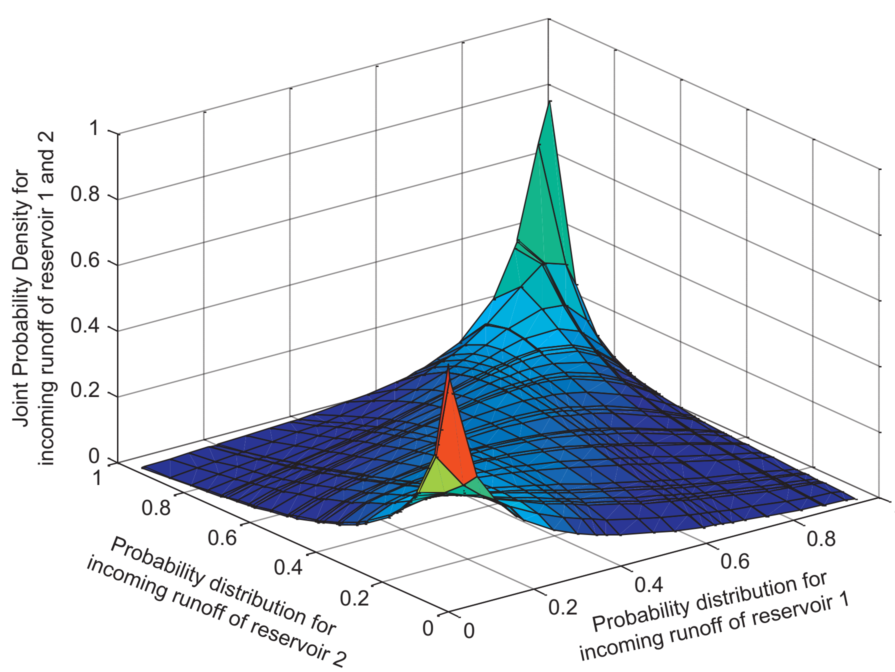

Fig. 5. The joint distribution of multiple water inflows.

$F_1(I_1),F_2(I_2),\ldots,F_\Omega(I_\Omega)$ denote their respective marginal distributions; $C(\cdot)$ is the Gumbel-Copula function; θ is the parameter of the Copula function and can be estimated with the MLE.

4.3. The power system economic dispatch model based on chance constraints

The water inflows of hydro plants are largely uncertain. In this paper, the uncertainty of water inflows is handled by employing chance constraints; in other words, the dispatch plan should satisfy the technical constraints with a certain probability [23]. The optimization objective of the proposed ED model is set as minimizing the overall generation cost:

$$\min f(P_{gi},P_{hi},Q_{fi},Q_{hi}) = \sum_{i=1}^{\Omega}(a_i P_{gi}^2 + b_i P_{gi} + c_i) \tag{10}$$

where Ω denotes the number of nodes in the power network; a_i, b_i, c_i are the cost parameters of the coal-fired plant at node i; P_{gi}, P_{hi} are respectively the power output of coal-fired plant and power output of hydro plant at node i; Q_{fi}, Q_{hi} are respectively the water outflows of the hydro plant for electricity generation and flood discharge purposes. P_{gi}, P_{hi}, Q_{fi}, Q_{hi} are all decision variables.

The following constraints will be considered in the proposed model:

(1) The power balance constraint

$$\sum_{i=1}^{\Omega}(P_{hi} + P_{gi} - P_{di}) = 0 \tag{11}$$

where P_{di} is the power demand at node i.

(2) System operating constraints

$$P_{gi}^{\min} \le P_{gi} \le P_{gi}^{\max} \tag{12}$$

$$P_{hi}^{\min} \le P_{hi} \le P_{hi}^{\max} \tag{13}$$

$$\Pr\{P_{hi} \le P_{Hi}\} \ge \gamma_1 \tag{14}$$

$$P_{Hi} = 9.81\eta_i Q_{fi} H_i \tag{15}$$

$$H_i = \psi_i(v_i) = \psi_i(I_i + V_i - Q_{fi} - Q_{hi}) - H_i^{\text{down}} \tag{16}$$

$$Q_{fi}^{\min} \le Q_{fi} \le Q_{fi}^{\max} \tag{17}$$

$$0 \le Q_{hi} \le Q_{hi}^{\max} \tag{18}$$

where $P_{gi}^{\max}$, $P_{gi}^{\min}$ are the upper and lower limits of power output of coal-fired plant i; $P_{hi}^{\max}$, $P_{hi}^{\min}$ are the upper and lower limits of power output of hydro plant node i; $\Pr\cdot$ represents the probability of an inequality; P_{Hi} is the maximum power output of hydro plant i; γ_1 is the pre-assigned confidence level; η_i is the power output parameter of hydro plant i; H_i is the water head of hydro plant i; $I_1,I_2,\cdots,I_i,\cdots,I_\Omega$ are the water inflows of hydro plants and follow the Gumbel-Copula function (9); V_i represents the initial water volume of hydro plant i before dispatch; $\psi_i(\cdot)$ represents the functional relationship between the water head and water volume v_i of hydro plant i; H_i^{down} is a parameter for calculating water head; $Q_{fi}^{\min}$,$Q_{fi}^{\max}$,$Q_{hi}^{\max}$ are the limits of electricity generation and flood discharge outflows.

(3) Hydro plant security constraints

$$\Pr\{I_i + V_i - Q_{fi} - Q_{hi} \le V_i^{\max}\} \ge \gamma_2 \tag{19}$$

$$\Pr I_i + V_i - Q_{fi} - Q_{hi} \ge V_i^{\min} \ge \gamma_3 \tag{20}$$

where $V_i^{\max}$, $V_i^{\min}$ represent the maximum and minimum water volume of hydro plant i; γ_2,γ_3 are the pre-assigned confidence levels.

(4) Transmission line power constraints

$$-P_l^{\max} \le \sum_{i=1}^{\Omega}\rho_{li}(P_{gi} + P_{hi} - P_{di}) \le P_l^{\max} \quad (l = 1,2,\ldots,L) \tag{21}$$

where ρ_{li} represents the sensitivity factor of the power injected at node i to the transmitted power of line l; $P_l^{\max}$ is the transmitted power limit of line l; L is the number of lines in the power network.

4.4. Handling chance constraints

To strictly calculate the chance constraints, it is necessary to calculate multiple integral, which is difficult when there are many hydro plants involved. Therefore, in this paper we employ the sample average approximation method to handle chance constraints, and transform the original stochastic optimization problem into a deterministic problem [23–25].

By employing function $\max(\cdot)$, a series of constraints can be transformed into an equivalent constraint as follows [25]

$$\{P_{hi} - 9.81\eta_i Q_{fi} H_i \le 0,\ i = 1,\cdots,\Omega\} := \max_{1 \le i \le \Omega}\{P_{hi} - 9.81\eta_i Q_{fi} H_i\} \le 0 \tag{22}$$

where

$$\Psi_1(P_{hi}) = \max_{1 \le i \le \Omega} P_{hi} - 9.81\eta_i Q_{fi} H_i$$

By randomly sampling J groups of water inflows $I_j = [I_{1j},\ldots,I_{\Omega j}]^T$ $(j = 1,\cdots J)$ for the Ω hydro plants from the Gumbel-Copula joint distribution (8), and using the average value to replace the probability:

$$\Pr\{\Psi_1(P_{hi}) \le 0\} = E[Flag(\Psi_1(P_{hi},I))] = \frac{1}{J}\sum_{j=1}^{J} Flag(\Psi_1(P_{hi},I_j)) \tag{23}$$

where $Flag(t)$ is the indicator function, which equals 1 when $t \le 0$, and equals 0 otherwise. The chance constraint (14) can then be

Table 1
Parameters of coal fired plants.

Node	a_i(\$/MW^{2}h)	b_i(\$/MWh)	c_i(\$/h)	$P_{gi}^{\min}$(MW)	$P_{gi}^{\max}$(MW)
1	0.043	20	0	0	330
2	0.25	20	0	0	140
3	0.01	40	0	0	100

Table 2
Parameters of hydro plants.

Node	η_i	$Q_{hi}^{\max}$(m^3/s)	$P_{Hi}^{\min}$(MW)	$P_{Hi}^{\max}$(MW)	$V_i^{\min}$(m^3)	$V_i^{\max}$(m^3)	V_i(m^3)	H_i^{down}(m)
6	0.84	0.8×10^4	56.6	270	5.88×10^9	6.34×10^9	6.0×10^9	41.6
8	0.8	0.6×10^4	33.4	225	2.75×10^9	2.9×10^9	2.8×10^9	162.0

transformed into deterministic constraint

$$\frac{1}{J}\sum_{j=1}^{J} Flag(\Psi_1(P_{hi}, I_j)) \geq \gamma_1 \quad (24)$$

Similarly, chance constraints (19) and (20) can also be transformed into deterministic constraints

$$\frac{1}{J}\sum_{j=1}^{J} Flag(\Psi_2(Q_{fi}, Q_{hi}, I_{ij})) \geq \gamma_2 \quad (25)$$

$$\frac{1}{J}\sum_{j=1}^{J} Flag(\Psi_3(Q_{fi}, Q_{hi}, I_{ij})) \geq \gamma_3 \quad (26)$$

where

$$\Psi_2(Q_{fi}, Q_{hi}, I_{ij}) = \max_{1 \leq i \leq \Omega} I_{ij} + V_i - Q_{fi} - Q_{hi} - V_i^{\max}$$

$$\Psi_3(Q_{fi}, Q_{hi}, I_{ij}) = \max_{1 \leq i \leq \Omega} Q_{fi} + Q_{hi} + V_i^{\min} - I_{ij} - V_i$$

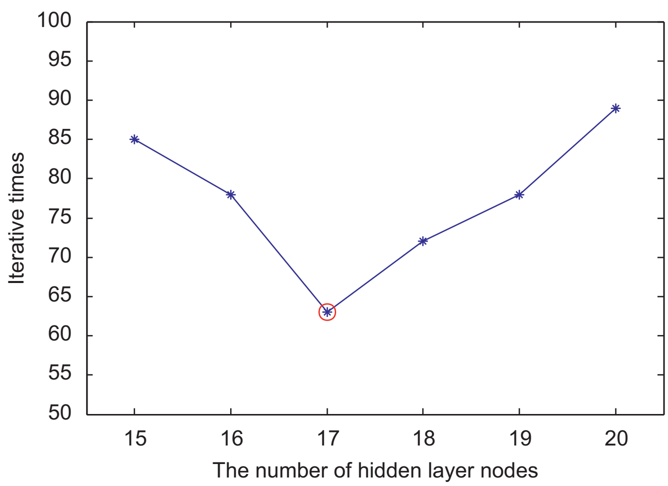

Fig. 6. The number of hidden layer nodes vs. convergence speed.

By replacing chance constraints (14) and (19)–(20) with constraints (24)–(26), the power system ED problems (10)–(21) are transformed into a deterministic optimization problem. The proposed modified GA can then be employed to solve this problem.

5. Case studies

The proposed ED model and modified GA will be tested with the IEEE 14-node system [26]. The IEEE 14-node system includes 5 generator nodes. Three coal-fired plants are located at nodes 1, 2 and 3. The parameters of the coal-fired plants are illustrated in Table 1.

At nodes 6 and 8, there are two hydro plants whose parameters are shown in Table 2. Also, for the two hydro plants, the functional relationships between water head and water volume

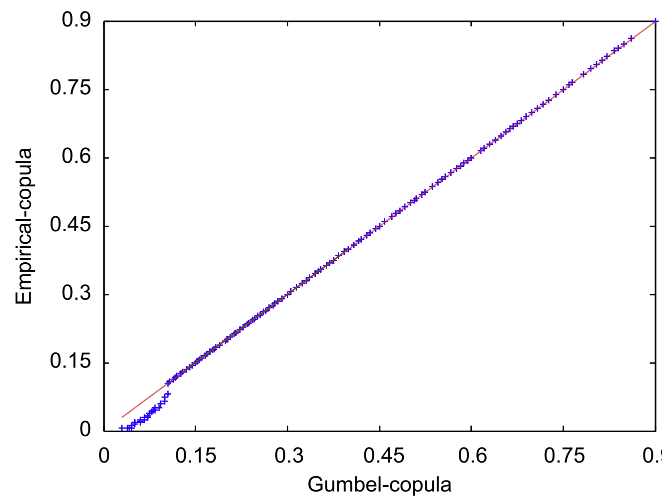

Fig. 8. The QQ plot of the joint distribution of water inflows.

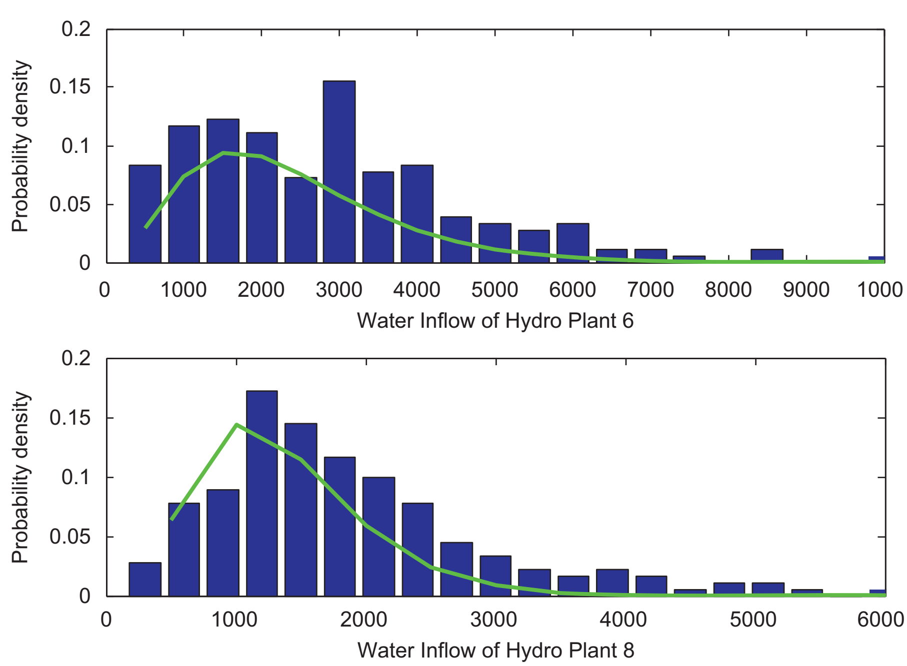

Fig. 7. The fitted distributions of hydro plants water inflows.

are

$\psi_6(v_6) = -64.7v_6^2 + 717.3v_6 - 1935,\ \psi_8(v_8) = -4.61v_8^2 + 32.69v_8 + 132.65$

The confidence levels of the chance constraints are set as $\gamma_1 = \gamma_2 = \gamma_3 = 0.98$.

The parameters of the proposed GA are set as follows. For both the downward-climbing and upward-climbing ELMs, the input layer and output layer have 10 and 9 nodes respectively. The number of nodes in the hidden layer will be determined based on the performance. The population sizes of the modified GA are set as $M = 50$, $M_m = 10$, $M_u = 10$ respectively. The maximum iteration number is $K = 200$. The convergence criterion is set as $\varepsilon = 10^{-5}$. The probabilities of crossover and mutation are set as $P_c = 0.8$ and $P_m = 0.1$. The inter-iteration parameter is set as $q = 3$. The relationship between the number of hidden nodes and the iteration number of the modified GA needed for converging is shown in Fig. 6. As illustrated, when the number of hidden nodes is set as 17, the modified GA converges fastest [14–15].

5.1. The joint distribution of the water inflows of two hydro plants

The water inflow data of the two hydro plants are the real data obtained from two hydro plants in Hunan Province, China. The parameters of the P-III distribution can be estimated from the water inflow data. The fitted distributions of two hydro plants are illustrated in Fig. 7.

Based on the P-III distributions of water inflows of two hydro plants, the Gumbel-Copula function can be employed to estimate the joint distribution of water inflows. The joint parameter θ can be estimated as 1.36. To verify the accuracy of the Gumbel-Copula function, the quantile–quantile (QQ) plots of empirical joint distribution and Gumbel-Copula joint distribution are shown in Fig. 8. As observed, the Gumbel-Copula function can well approximate the joint distribution of two water inflows.

5.2. The results of the power system economic dispatch

We investigate the performance of the proposed ED and modified GA in the following four aspects:

(1) *The impact of the joint distribution on economic dispatch*
Based on the Gumbel-Copula joint distribution of the water inflows of hydro plants 6 and 8, the ED model is solved with the modified GA. The obtained system generation cost is 2967.2$. On the other hand, if the water inflows of two plants are assumed to be independent, the corresponding system generation cost is 3361.3$, as illustrated in Table 3.

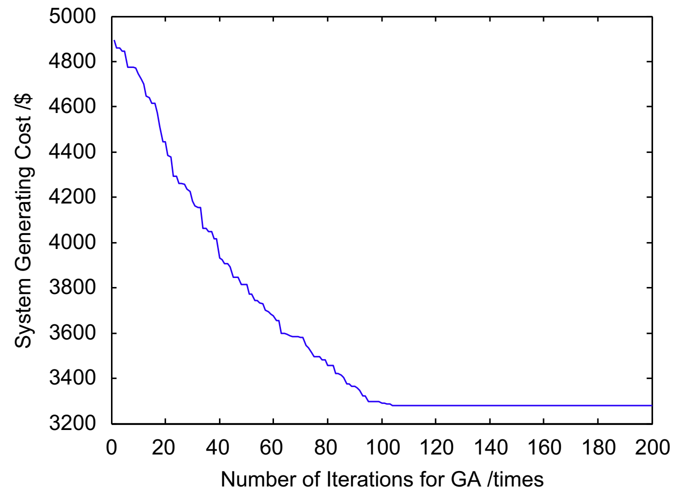

Fig. 9. The relationship between system generation cost and iteration number for basic GA.

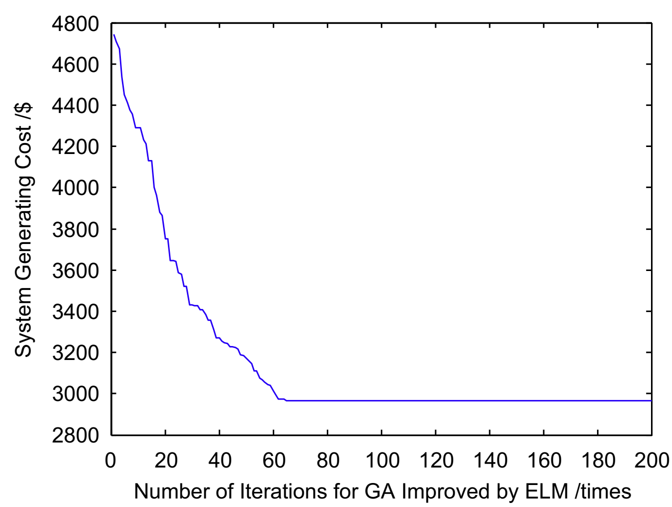

Fig. 10. The relationship between system generation cost and iteration number for modified GA.

Table 3
The dispatch results considering different assumptions of water inflow dependency.

Dependency between two water inflows	I_1 (m^3/s)	I_2 (m^3/s)	Q_{f6} (m^3/s)	Q_{h6} (m^3/s)	Q_{f8} (m^3/s)	Q_{h8} (m^3/s)	Power output (MW)		System generation cost ($)
							6	8	
Independent	6980.2	3183.2	3206.35	2067.61	1185.67	548.74	51.479	91.586	3361.3
Copula	7377.6	3341.0	3652.47	2304.81	1350.71	650.37	54.571	94.256	2967.2

Table 4
Performance comparison between basic and modified GA.

Algorithm	Power output (MW)					K	Computational time (s)	System generation cost ($)
	1	2	3	6	8			
Basic GA	85.338	19.893	12.744	51.025	90.000	104	7.93	3289.8
Modified GA	80.577	18.175	11.421	54.571	94.256	63	5.85	2967.2

Table 5
The performance of modified GA by selecting different q.

q	Power output (MW)					K	Computational time (s)	System generation cost ($)
	1	2	3	6	8			
2	81.332	19.658	12.527	53.324	92.159	78	6.51	2988.9
3	80.577	18.175	11.421	54.571	94.256	63	5.85	2967.2
4	82.832	20.542	12.001	52.367	91.258	86	7.27	3023.1
5	80.843	19.852	12.247	54.001	92.057	81	6.74	2979.6

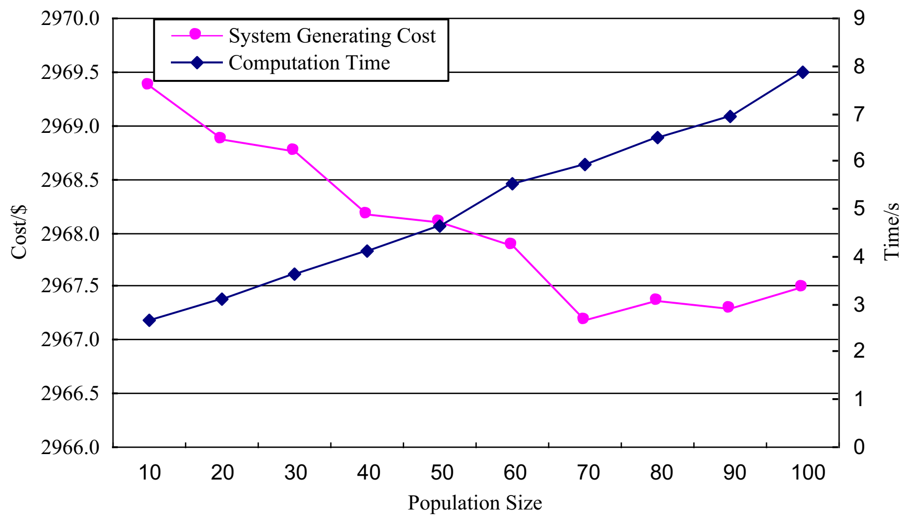

Fig. 11. The impact of overall population size on the performance of modified GA.

Comparing the results, it is clear that by employing the Gumbel-Copula function, the system generation cost can be reduced.

(2) *Performance comparison of optimization algorithms*
The proposed ED model is solved with basic and modified GA respectively. The corresponding dispatch plans, system generation costs, iteration numbers and computational time are given in Table 4. Also, the relationship of system generation cost vs. iteration number is plotted in Figs. 9 and 10. As observed, compared with corresponding values of the basic GA, the modified GA can decrease the system generation cost by 322.6$. Also, its iteration number and computational time are 41 and 2.08 s smaller than the basic GA. The results demonstrate that the modified GA outperforms basic GA in terms of both accuracy and convergence speed.

(3) *The impact of inter-iteration parameter on the performance of modified GA*
The dispatch results obtained by selecting different inter-iteration parameter q are given in Table 5. As observed, the optimal dispatch plan is obtained when q is set as 3. Therefore, it is important to select q appropriately so as to improve the performance of the modified GA.

(4) *The impact of population size on the performance of modified GA*
The overall population size $M+M_m+M_u$ has significant impact on the performance of the modified GA. To investigate its impact, the relationships of population size vs. computational time, and population size vs. system generation cost are illustrated in Fig. 11. As clearly observed, as the population size increases, the system generation cost will be reduced. However, the computational time will increase linearly with the population size. Therefore, to achieve a tradeoff between accuracy and computational efficiency, the best choice of $M+M_m+M_u$ is around 70.

6. Conclusion

This paper presents a modified genetic algorithm based on extreme learning machine. To utilize the exceptional function estimation capability of ELM, the proposed algorithm employs two ELM based evolution operators to enhance the local and global search capabilities of basic GA. By continuously retraining the ELM based operators, and using them to guide the search direction, the convergence property of GA can be improved.

The proposed algorithm is applied in the power system economic dispatch problem. The results of case studies demonstrate that:

(1) The join distribution of the water inflows of multiple hydro plants can be accurately approximated with the Gumbel-Copula function. Also, the uncertainties of the power system ED problem can be well handled with chance constraints.
(2) The proposed algorithm outperforms the basic GA in terms of both accuracy and computational efficiency in the ED problem.
(3) To further improve the convergence property of the modified GA, it is important to appropriately tune its parameters, such as the inter-iteration parameter and the population size.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (71071025, 70601003, 51107114), the Outstanding Youth Science Foundation of Hunan Province (10JJ1010), the Program for New Century Excellent Talents in University of China (NCET-08–0676), the Open Innovation Platform Foundation of Hunan College (10K003), and National Outstanding Youth Science Foundation of China (70925006).

References

[1] B.P. Buckles, F.E. Perry, Genetic Algorithms, IEEE Computer Society Press, Los Alamitors, California, 1992.
[2] L.C. Jiao, L. Wang, A novel genetic algorithm based on immunity, IEEE Trans. Syst. Man Cybern.—Part A: Syst. Hum. 30 (5) (2000) 552–561.
[3] M. Srinivas, L.M. Patnaik, Adaptive probabilities of crossover and mutation in genetic algorithms, IEEE Trans. Syst. Man Cybern. 24 (1994) 656–667.
[4] G. Magyar, M. Johnsson, O. Nevalainen, An adaptive hybrid genetic algorithm for the three-matching problem, IEEE Trans. Evol. Comput. 4 (2) (2000) 135–146.
[5] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (1–3) (2006) 489–501.
[6] I. Mrazova, D.H. Wang, Improved generalization of neural classifiers with enforced internal representation, Neurocomputing 70 (16–18) (2007) 2940–2952.
[7] Q.Y. Zhu, A.K. Qin, P.N. Suganthan, G.B. Huang, Evolutionary extreme learning machine, Pattern Recognition 38 (2005) 1759–1763.
[8] G.B. Huang, D.H. Wang, Y. Lan, Extreme learning machines: a survey, Int. J. Mach. Learn. Cybern. 2 (2011) 107–122.
[9] G.B. Huang, Q.Y. Zhu, K.Z. Mao, C.K. Siew, P. Saratchandran, N. Sundararajan, Can threshold networks be trained directly? IEEE Trans. Circuits Syst. II: Express Briefs 53 (3) (2006) 187–191.
[10] G.B. Huang, L. Chen, Enhanced random search based incremental extreme learning machine, Neurocomputing 71 (16–18) (2008) 3460–3468.
[11] G.B. Huang, D.H. Wang, Advances in extreme learning machines, Neurocomputing 74 (16) (2011) 2411–2412.
[12] G.B. Huang, Q.Y. Zhu, C.K. Siew, Real-time learning capability of neural networks, IEEE Trans. Neural Networks 17 (4) (2006) 863–878.
[13] L. Chen, G.B. Huang, H.K. Pung, Systemical convergence rate analysis of convex incremental feedforward neural networks, Neurocomputing 72 (10–12) (2009) 2627–2635.
[14] G.B. Huang, M.B. Li, L. Chen, C.K. Siew, Incremental extreme learning machine with fully complex hidden nodes, Neurocomputing 71 (4–6) (2008) 576–583.
[15] G.B. Huang, L. Chen, C.K. Siew, Universal approximation using incremental constructive feedforward networks with random hidden nodes, IEEE Trans. Neural Networks 17 (4) (2006) 879–892.
[16] G.B. Huang, H. Zhou, X. Ding, R. Zhang, Extreme learning machine for regression and multi-class classification, IEEE Trans. Syst. Man Cybern.—Part B (September 2011).
[17] Xiaohong Ernan Ni, Renhou Li Guan, Scheduling hydrothermal power systems with cascaded and head-dependent reservoirs, IEEE Trans. Power Syst. 14 (3) (1999) 1127–1132.
[18] Jin Yang Yunfeng, Miao Yang Xiaohong, A modified genetic algorithm and its application in optimizing in the function, Artif. Intell. Recognition Technol. (2007) 1118–1119.
[19] D.E. Goldberg, Genetic Algorithm in Search, Optimization and Machine Learning, Addison-Wesley, Massachusetts, 1989.
[20] G.Y. Jin, A review of hydrologic frequency analysis, Adv. Water Sci. 10 (1999) 319–327.
[21] B. Renard, M. Lang, Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology, Adv. Water Sci. 30 (4) (2007) 897–912.
[22] H. Valizadeh Haghi, M. Tavakoli Bina, M.A. Golkar, S.M. Moghaddas-Tafreshi, Using Copulas for analysis of large datasets in renewable distributed generation: PV and wind power integration in Iran, Renewable Energy 35 (9) (2010) 1991–2000.
[23] J.K. Wu, J.Q. Zhu, G.T. Chen, H.L. Zhang, A hybrid method for optimal scheduling of short-term electric power generation of cascaded hydroelectric plants based on particle swarm optimization and chance-constrained programming, IEEE Trans. Power Syst. 23 (4) (2008) 1570–1579.
[24] A. Shapiro, Monte Carlo sampling methods, Handb. Oper. Res. Manage. Sci. 10 (2003) 353–425.
[25] Anton J. Kleyyweght, Alexander Shapiro, Tito Homen-de-Mello, The sample average approximation methods for stochastic discrete optimization, Soc. Ind. Appl. Math. 12 (2) (2001) 479–502.
[26] Divshali, P. Hasanpor, Hosseinian, S.H. Azadani, E. Nasr, M. Abedi, Application of bifurcation theory in dynamic security constrained optimal dispatch in deregulated power system, Electr. Eng. (2011) 1–10.

Hongming Yang received her M.Sc. degree in Electrical Engineering from Wuhan University, China in 1997. She received her Ph.D. degree in electrical engineering from Huazhong University of Science and Technology, China in 2003. Currently, she is a Research Associate in the Hong Kong Polytechnic University and also a full Professor in Changsha University of Science and Technology, China. Her research interests include power system analysis and power markets.

Jun Yi was born in Changsha, China in 1984. He received his B.Eng.degree in automation from Changsha University of Science and Technology, China in 2008. He is currently pursuing his M.S. degree in Electrical engineering from Changsha University of Science and technology. His main research interests include power system analysis and power market.

JunHua Zhao obtained his Bachelor and Ph.D. degrees from Xi'an Jiaotong University, China and The University of Queensland, Australia, in 2003 and 2007, respectively. He is now a postdoctoral research fellow at School of Electrical Engineering, Zhejiang University, China. Dr. Zhao's research interests include power system analysis and computing, smart grid, stochastic methods and their applications in power engineering.

Zhaoyang Dong obtained his Ph.D. from The University of Sydney, Australia in 1999. He is now the chair professor at the Centre for Intelligent Electricity Networks, The University of Newcastle. His research interests include power system security assessment, power system planning and management, electricity market simulation and risk management, data mining, computational intelligence and their application in power engineering. Dr. Dong is a Senior Member of IEEE.