International Journal of Soft Computing 10 (1): 87-93, 2015
ISSN: 1816-9503

ANN And Gradient Based Optimal Power Flow

[1]R. Prathiba and [2]M. Balasingh Moses
[1]Department of Electrical and Electronics Engineering,
Easwari Engineering College, Chennai-89, India
[2]Department of Electrical and Electronics Engineering, Anna University, Trichy, India

Abstract: Optimal Power Flow (OPF) is essential and is employed to attain desired performance in any field of engineering. It plays an important role in power system operation and planning. In deregulated environment of power sector, it is of increasing importance for determination of electricity prices. Traditional optimization techniques have limitations in obtaining the global solution. With a non-monotonic solution surface, classical methods are highly sensitive to starting points and frequently converge to local optimal solution or diverge altogether. This study describes an artificial neural network-GD based optimal power flow which is highly constrained non convex optimization problems with single objective, the fuel cost and compares its results with well known classical methods in order to prove its validity for present deregulated power system analysis.

Key words: Optimal power flow, gradient descent, artificial neural network, deregulation, error weightage

INTRODUCTION

Global optimization of non-continuous, non-linear functions arising from real world complex engineering problems which may have large number of local minima and maxima is quite challenging. A number of deterministic approaches based on branch and bound and real algebraic geometry are found to be successful in solving these problems to some extent. Of late, stochastic and heuristic optimization techniques have emerged as efficient tools for global optimization. It has been applied to a number of engineering problems in diverse fields and one such field is power system optimization. The power system is a complex network used for generating and transmitting electric power. It is expected to operate with consumption of minimal resources giving maximum security and reliability. The Optimal Power Flow (OPF) problem is an important tool to help the operator achieve these goals by providing the optimal settings of all controllable variables. The various objectives of OPF problem are:

- Minimization of cost of generation
- Minimization of transmission losses
- Maximization of social benefit

A number of mathematical optimization techniques have been proposed in literature to solve the OPF problem. For decades, conventional optimization techniques such as Linear Programming (LP), Quadratic Programming (QP), Newton Method and Interior Point Methods have been used for solving optimal power flow problem (Dommel and Tinney, 1968; Tinney and Hart, 1967; Momoh *et al*., 1994). LP Method requires that the objective function and constraints have linear relationship which may lead to loss of accuracy. The Newton Method (Tinney and Hart, 1967) suffers from the difficulty in handling inequality constraints. Conventional methods are not efficient in handling problems with discrete variables. The combinatorial search approaches, branch and bound and cutting plane algorithms which are usually used to solve the Mixed Integer Programming Model are non-polynomial and all suffer from the curse of dimensionality making them unsuitable for large scale OPF problems. In recent years global optimization techniques such as Genetic Algorithms (GA), Evolutionary Programming (EP), Particle Swarm Optimization (PSO) and differential evolution have been proposed to solve the OPF problem (Yang *et al*., 1996; Storn and Price, 1997; Momoh *et al*., 1999; Goldberg, 1989).

The cost of generation optimization problem has to be formulated as a mixed integer, non-linear problem. Artificial neural network in combination with Gradient Method presented here uses a less stochastic approach to problem solving. ANN is a machine learning based classifier capable of solving non-linear problems and incorporating constraints of varying nature. Back

Corresponding Author: R. Prathiba, Department of Electrical and Electronics Engineering, Easwari Engineering College, Chennai-89, India

Propagation algorithm is used in the decision making process as to the validity of the constraints and the optimization is achieved independently by gradient descent. But the GD is invoked based on the decision made by ANN and therefore two independent algorithms are cascaded to form a hybrid. The validity of the proposed method is tested on standard IEEE 30 bus systems. Results obtained using ANN-GD and a conventional optimization technique is also provided for comparing the performance of the proposed method.

OPTIMAL POWER FLOW PROBLEM

The Optimal Power Flow (OPF) is a static, non-linear and non-convex optimization problem which determines a set of optimal variables from the network state, load data and system parameters. Optimal values are computed in order to achieve a certain goal such as generation cost or transmission line power loss minimization subjected to equality and inequality constraints. In general the OPF problem can be presented as follows:

$$\text{Min } f(x, u) \tag{1}$$

$$\text{s.t } g(x, u) = 0 \tag{2}$$

$$h(x, u) \leq 0 \tag{3}$$

$$x_{min} \leq x \leq x_{max} \tag{4}$$

$$u_{min} \leq u \leq u_{max} \tag{5}$$

where, f(x, u) is the objective function that typically includes total generation cost (active power dispatch) or total losses in transmission system (reactive power dispatch). Generally (x, u) represents the load flow equations and h(x, u) represents transmission line limits. The vector of dependent and control variables are denoted by x and u, respectively.

OBJECTIVE FUNCTION

Let the objective function to be minimized is given:

$$F = \sum_i Fi(Pg_i)$$

This is the sum of operating cost over all controllable power sources. $Fi(Pg_i)$ is the generation cost function for Pg_i generation at bus i. The cost is optimized with the following constraints. The inequality constraint on real power generation at bus i:

$$Pg_i^{min} \leq Pgi \leq Pg_i^{max} \tag{6}$$

where, Pg_i^{min} and Pg_i^{max} are respectively minimum and maximum values of real power generation allowed at generator bus i. The power flow equation of the power network:

$$g(v, \varphi) = 0 \tag{7}$$

Where:

$$g(v, \varphi) = \begin{cases} P_i(v, \varphi) - P_i^{net} \\ Q_i(v, \varphi) - Q_i^{net} \\ P_m(v, \varphi) - P_m^{net} \end{cases} \tag{8}$$

Where:

P_i and Q_i	=	Calculated real and reactive power for PQ bus i
P_i^{net} and Q_i^{net}	=	Specified real and reactive power for PQ bus i
P_m and P_m^{net}	=	Calculated and specified real powers for PV bus m
V and φ	=	Voltage magnitude and phase angles at different buses

The inequality constraint on reactive power generation Qg_i at each PV bus:

$$Qg_i^{min} \leq Qg_i \leq Qg_i^{max} \tag{9}$$

where, Qg_i^{min} and Qg_i^{max} are respectively minimum and maximum value of reactive power at PV bus i. The inequality constraint on voltage magnitude V of each PQ bus:

$$V_i^{min} \leq V_i \leq V_i^{max} \tag{10}$$

where, V_i^{min} and V_i^{max} are respectively minimum and maximum voltage at bus i. The inequality constraint on phase angle φ_i of voltage at all the buses i:

$$\varphi_i^{min} \leq \varphi_i \leq \varphi_i^{max} \tag{11}$$

where, φ_i^{min} and φ_i^{max} are, respectively minimum and maximum voltage angles allowed at bus i. MVA flow limit on transmission line:

$$MVAf_{ij} \leq MVAf_{ij}^{max} \tag{12}$$

where, $MVJf_{ij}^{max}$ is the maximum rating of transmission line connecting bus i and j.

HEURISTIC OPTIMIZATION

The methodology adopted for computing OPF using ANN and gradient descent method is explained with the help of a general block diagram as shown in Fig. 1. A large number of patterns were generated considering a wide range of load variation randomly at each bus. The input features were selected to reduce the dimensionality of the input as well as size of neural network using constraint violations alone. The selected inputs as well as outputs were normalized and supervised learning was applied for accurate estimation of the fuel cost using multilayer perception. The structure of the MLP network used in this research is shown in Fig. 2.

Training set generation: For generating the training set, a large number of differing inequality constraints such as active, reactive power flow, voltage magnitude and phase angle violations are considered for various values of power from each generator and the corresponding network parameters.

Input features: The inputs selected for training the neural network are the in-equality constraints violations from the power determined by the gradient descent from the previous step.

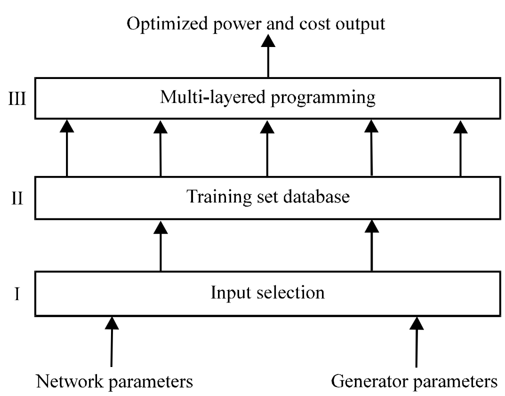

Fig. 1: The conceptual diagram of learnin model for MLP

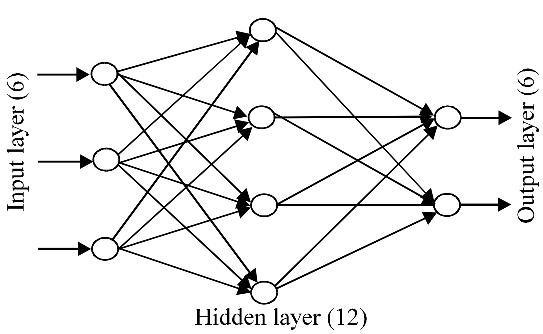

Fig. 2: Topology of a three layered MLP

Normalization of the data: Normalization of data is an important aspect for training of the neural network. Without normalization the higher valued input variables may tend to suppress the influence of smaller ones. To overcome this problem, neural networks are trained with normalized input data. The value of input variables is scaled between some suitable values [0-2]. In case of output variables, if the value varies over a wide range, the neural network is not able to map it correctly. The remedy is to scale the output between some suitable ranges [0-2].

Learning algorithm: Learning is a very important procedure in supervised Learning algorithms as they form the platform for the prediction process to occur. The Back Propagation algorithm is carried out by gradient descent for the learning process. Gradient-Descent Methods are among the most widely used of all function approximation methods and are particularly well suited to reinforcement learning. Gradient descent is a first-order optimization algorithm. To find a local minimum of a function using gradient descent, one takes steps proportional to the negative of the gradient (or of the approximate gradient) of the function at the current point. In this approach, the smooth problem is solved in the hope that the answer is close to the answer for the non-smooth problem. This process can sometimes be made rigorous with multiple random initializations.

Back propagation generally deals with using the intermediate output as the guideline for training the network. It uses the error between a predefined standard output and the current prediction of the network to fine tune the parameters of the network for future predictions.

The errors between standard and current predictions for the different possible parameters of the network in the range are determined and they form a database. The minimum sum-squared error in the database is determined using Gradient Descent Method with multiple initializations to ensure convergence to global minimum. The gradient descent moves with every step towards a possibly less error along the negative gradient of the function of dataset. Thus, it converges to a minimum which may be local or global. With this minimum as initialization at various points along the curve, the procedure is repeated and convergence to global minimum is ensured. The parameters corresponding to the minimum error are applied to the network and the same procedure is carried out until the sum-squared error goal is achieved.

The training data is used to generate the pre-defined standard output for a set of inputs. This set of inputs is programmed to increase with the number of test cases as the training data increases with the number of input values and their outputs. Therefore, the performance of the Learning algorithm and hence the prediction on the whole is bound to improve with time.

Stopping criteria: The algorithm stops its execution only under one of the two conditions:

- Sum squared error goal is achieved
- Maximum number of iterations is reached

The sum squared error goal is fixed at 0.05 to ensure proper accuracy of results and also to make sure that the results are obtained as a result of convergence and not merely based on reaching the limit of iterations. This is very essential because the algorithm may be on the verge of divergence due to various reasons and cutting it off abruptly may have adverse results on the efficiency of the algorithm.

So, it is very essential to obtain a suitable trade-off between the accuracy and iteration factor even though it sacrifices a small amount of accuracy. It ensures reliability and better efficiency than when the algorithm is abruptly cut-off. This value of sum-squared error goal is fixed manually by iterations and deciding which value performs the best in both the respects.

Gradient descent:

- Convergence to global minimum
- Iteration limit

The convergence to global minimum is considered the most suitable parameter for stopping criteria. But in cases where the convergence is restricted by the generator constraints the algorithm also stops with the iteration limit.

In such cases it is essential to maintain a proper learning rate. It determines how fast the algorithm converges to the minimum and sometimes looking for fast convergence may result in divergence and when the algorithm stops by maximum iteration, the value is away from the global minimum by a large distance. So, if there is a possibility of the algorithm stopping by iteration limit it is essential to make sure that the learning rate is as small as possible. It would not affect the results because convergence is not an essential parameter and the time also remains unchanged because it is constant when a certain number of iterations are invariably executed.

In this method, the learning rate is fixed at 1.1 to the normal value of 1. This is because the value of cost is a less steep function and therefore the enhancement of the gradient value by 10% will only aid the speeding of the convergence and not affect their accuracy in anyway.

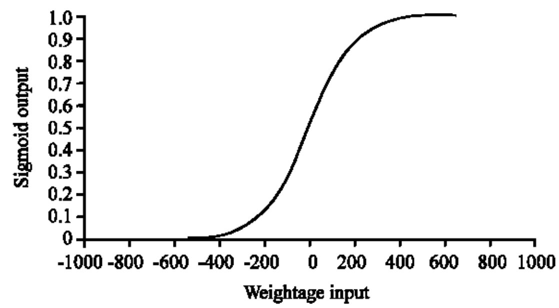

Fig. 3: Sigmoid function

Sigmoid function: Sigmoid function is a prediction function whose output lies in the range of 0 and 1. When the input is very negative, the output is zero. When the input is zero, output is 0.5 and when the input is highly positive, the output is one. Thus, based on the weight age allocation to various errors, the sigmoid function decides the importance of the errors and predicts the generator on/off status accordingly. It is given by the equation:

$$f(z) = \frac{1}{1+e^{-z}} \tag{13}$$

It can be represented graphically as in Fig. 3 which depicts the output range of the sigmoid function for the range of values from -1000 to 1000.

Solution algorithm: The Proposed Solution algorithm for ANN-Gradient Descent Based Optimal Power Flow (ANN-GD OPF) is given. Complete flow chart has been shown in Fig. 4.

Step 1: Prepare the data set of generator data, bus data, capacitor data, reactor data, transformer data and transmission line data

Step 2: Initialize number of generators (i = 1, 2, 3..., k) initialization of costcurve based on random guesses of power generate for each generator.
gen_pow = k (any random number)

Step 3: Obtain negative gradient of the function at each point of initialization and traverse the path of the curve to obtain global minimum for eac generator.
gen_pow = k*dF (Pi)/dt
F(Pi) is the cost function of the generator.

Step 4: Determine Y bus matrix, reactive and real power limits and other network and generator constraints

Step 5: Check if gen_pow > minimum generator power and gen_pow<maximum generator power. If condition is satisfied, proceed to next step. Else repeat steps 2-5 until conditions are satisfied.

Step 6: Determine the maximum and minimum limits of slack bus generation, bus voltage magnitudes and phase angles at all the buses and their corresponding power by running power flow analysis (Pg. max and Pg. min)

Step 7: Check the following constraints with gen_pow:

- Check the bus voltage violation
- Check the bus voltage phase angle
- Check the MVA flows violation for all the lines connected between bus i and j

- Check reactive power limits at all generator buses

if any of the above limits is violated proceed to next step. Else just output the power corresponding to the minimum cost (gen_pow)

Step 8: If the constraints are violated, calculate the error between constraint limits and obtained values. It is fed as input to the ANN input layer.

Step 9: Train the ANN through a dataset generated by all possible power values within the network constraints as determined in Step 6. The training is imparted through the Back Propagation algorithm that minimizes the error between predicted value and the closest value in the dataset. The weightage is allocated based on magnitude of error

a = min. F (e)

gen_wt = a*err_pg

Where: F (e) is the function of error between expected and predicted values in the dataset

a is the weightage factor

gen_wt is the error weightage

err_pg is the error from the input layer

Step 10: This gen_wt is the direct representation of error in various buses and is fed to the sigmoid function in the output layer which produces the on/off status of the generator as output.

Step 11: This on/off status gives the power allocation of generators that could be altered to overcome the constraint lapse. For example, [1 1 0 0 1 0] says that first, second and fifth generator power needs to be altered to obtain better results. The range of the output is 0 to 1. If the value is >0.6, we decide that the generator needs re-iteration.

Step 12: This output of ANN is set as guiding value to the GD. Repeat Steps 2-9 until all constraints are satisfied and no generator needs re-iteration.

Step 13: Repeat iteration for all generators and find the consolidated optimal power and the final generation cost.

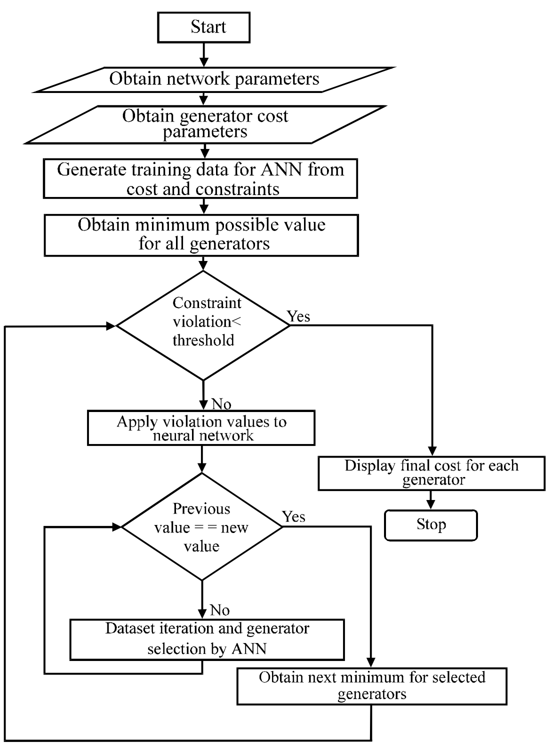

Fig. 4: Flow chart of ANN-GD-OPF

SIMULATION RESULTS

The proposed ANN-GD-OPF algorithm is applied to IEEE-30 bus test system whose data has been given in (http://www.ee.washington.edu/research/pstca.). For this analysis the NN parameters as shown in Table 1 has been taken. This algorithm is applied to quadratic cost characteristic equation as per the following equation and the value of cost coefficients given in Table 2 has been taken for analysis:

$$Fi(Pg_i) = \alpha_i \times Pg_i^2 + \beta_i \times Pg_i - \gamma_i \quad (14)$$

The cost function of the generator and its corresponding cost curve is initialized randomly. The method of negative gradients is used to move towards the optimum of each curve until a minimum point is reached. The power and the cost corresponding to the minimum are determined. Determination of Y bus matrix and other network parameters are carried out as initialization. The various basic generator constraints and network constraints are set up. The optimal power is checked for the generator constraints if violated next local optima is obtained using Gradient Descent Method and then proceed to network constraints. The network parameters for the power obtained in the previous steps are obtained. The error between parameters and the limits are fed to the NN input layer. For IEEE 30 bus system the trained NN has one input layer (6 neurons-one per constraint), one hidden layer (12 neurons) and one output layer (6 neurons-one per generator).

The weight age is allocated to each error based input based on the back propagating on the dataset obtained by using all network data and their corresponding network parameters. The back propagation minimizes the error between the predicted network parameters from the obtained power and the actual power within constraints.

Table 1: ANN parameters

ANN parameters	Values
Layer size	3
Maximum number of epochs to train	5000
Learning rate	1.1
Sum-squared error goal	0.05

Table 2: Cost coefficients of generating units

		Cost coefficients	
Generators	Bus number	$\alpha_{i(\$/MW^2h)}$	$\beta i_{(\$/MW}hi_{(\$/}h)$
G1	1	0.00375	2.000 (0)
G2	2	0.01750	1.750 (0)
G3	5	0.06250	1.000 (0)
G4	8	0.00834	3.250 (0)
G5	11	0.02500	3.000 (0)
G6	13	0.02500	3.000 (0)

Table 3: Simulation results of NN and GD based OPF

Generator	Bus number	Power (MW)	Cost ($/h)
1	1 (reference)	181.1100	485.22
2	2	41.9844	104.32
3	5	16.6004	33.82
4	8	23.0070	79.19
5	11	9.9990	32.49
6	13	12.0000	39.60
Total	-	284.7000	774.65

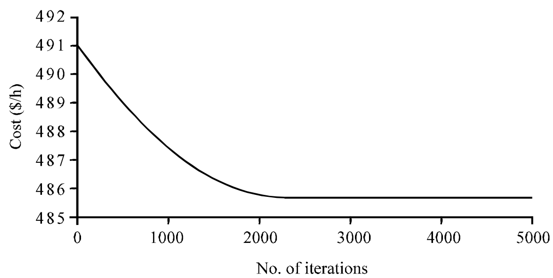

Fig. 5: Gradient cost converagence descent (Generator G1)

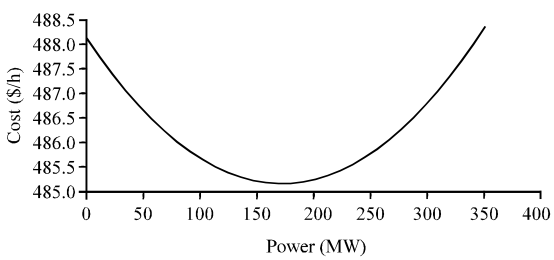

Fig. 6: Cost post curve of generator G1

Thus, the weight age is more to high violation of constraints and relatively low to marginal lapse. Thus, the exact bus and therefore the exact generator where the constraints are relaxed is obtained.

The weight age multiplied with the input gives the suitable values for each generator status. Sigmoid function is applied on the data and then the final generator status is determined. This generator on/off status gives the load allocation of generators that could be altered to overcome the constraint lapse. For example, [1 1 0 0 1 0] says that first, second and fifth generator power needs to be altered to obtain better results. These generator curves are re-iterated and the new minimum values are obtained. The procedure is repeated until the constraint errors are nullified and the power values settle within a satisfactory margin.

The simulation results for the proposed algorithm are given in Table 3. The convergence of this proposed algorithm is shown in Fig. 5. The value of cost of the Generator (G1) with respect to the number of iterations is plotted and it is observed that the convergence occurs at $485.2 which validates the results obtained for Generator G1 in Table 3. The number of iterations required to obtain global minimum is 2320. The figure shows that the cost remains almost constant even if the numbers of iterations are increased further. In order to validate that the cost obtained for Generator G1 corresponds to its global minimum, the cost power function curve of the generator is plotted in Fig. 6. It shows that the power corresponding to the cost $485.22 is 181.11MW and is the least possible value for the given generator, subject to the constraints.

It is observed that the GD reaches smoothly about the minimization of cost without any round-about courses, thus making sure that the convergence is very fast while also ensuring that the global minimum is reached. The constant corresponding to the proportionality of the negative gradients in GD is called learning rate and it determines the time taken for convergence to occur. But making it too large may lead to divergence in a few cases. So, a proper trade-off needs to be obtained between speed and reliability and this learning rate is the parameter to be tampered with. Iteratively, researchers fix the learning rate parameter at unity for this algorithm.

Table 4: Comparison of OPF Methods

Approach	Minimum generation cost ($/h)
Steepest descent-OPF	803.549568
EP-OPF	803.571862
GA-OPF	803.915839
Matlab optimization tool box	803.550000
ANN-GD-OPF	774.650000

COMPARISON OF RESULTS

Table 4 gives the comparison of the optimal point (i.e. minimum cost) in all the above four methods applied to same data IEEE 30-bus test system (http://www.ee.washington.edu/research/pstca).

CONCLUSION

The ANN-GD based OPF algorithm with single objective solution has been proposed in this study. Single objective optimization problems like active power problem with cost minimization objective have been solved. Simulation results shows that the proposed scheme is capable of handling mixed integer as well as non-linear optimization problems. This method is capable of obtaining better solutions when compared to EP and GA (Yuryevich and Wong, 1999; Devaraj and Yegnanarayana, 2005; Sood, 2007). Thus, the proposed ANN-OPF Method is most suitable for incorporating new objective functions and constraints arising from deregulated environment of modern power systems.

REFERENCES

Devaraj, D. and B. Yegnanarayana, 2005. Genetic-algorithm-based optimal power flow for security enhancement. Proc. IEE. Generat. Transmission Distribut., 152: 899-905.

Dommel, H.W. and W.F. Tinney, 1968. Optimal power flow solutions. IEEE Trans. Power Apparatus Syst., 87: 1866-1876.

Goldberg, D.E., 1989. Genetic Algorithms in Search Optimization and Machine Learning. Addison-Wesley, New York, USA.

Momoh, J.A., R. Adapa and M.E. El-Hawary, 1999. A review of selected optimal power flow literature to 1993. I. Nonlinear and quadratic programming approaches. IEEE Trans. Power Syt., 14: 96-104.

Momoh, J.A., S.X. Guo, E.C. Ogbuobiri and R. Adapa, 1994. The quadratic interior point method solving power system optimization problems. IEEE Trans. Power Syst., 9: 1327-1336.

Sood, Y.R., 2007. Evolutionary programming based optimal power flow and its validation for deregulated power system analysis. Int. J. Elect. Power Energy Syst., 29: 65-75.

Storn, R. and K. Price, 1997. Differential evolution: A simple and efficient heuristic for global optimization over continuous spaces. J. Global Optim., 11: 341-359.

Tinney, W.F. and C.E. Hart, 1967. Power flow solution by newton's method. IEEE Trans. Power Apparatus Sys., 86: 1449-1460.

Yang, H.T., P.C. Yang and C.L. Huang, 1996. Evolutionary programming based economic dispatch for units with non-smooth fuel cost function. IEEE Trans. Power Syst., 11: 112-118.

Yuryevich, J. and K.P. Wong, 1999. Evolutionary programming based optimal power flow algorithm. IEEE Trans. Power Syst., 14: 1245-1250.