Optimal Operation of Thermal Electric Power Production System without Transmission Losses: An Alternative Solution using Artificial Neural Networks based on External Penalty Functions

G.J. TSEKOURAS1, F.D. KANELLOS2, C.D.TSIREKIS3, N.E. MASTORAKIS1
1Department of Electrical & Computer Science, Hellenic Naval Academy Terma Hatzikyriakou, 18539, Piraeus
2 Production Engineering & Management Department, Technical University of Crete Technical University Campus, 89 Dyrrachiou, Athens, 10443 GREECE
3Independent Power Transmission Operator Chania, 73100 GREECE
Email: tsekouras_george_j@yahoo.gr, kanellos@mail.ntua.gr, ktsirekis@desmic.gr, mastor@wseas.org

Abstract: - The optimal economic operation of a power system is a crucial problem usually solved with Lagrange method. Alternatively, artificial neural networks (ANN) can be used for continuous, nonlinear, constrained optimization problems like the one examined in this paper. In this paper, a special ANN is developed for the determination of the optimal operation of an electric power production system without transmission losses. A specific case study is studied via the proposed method and the advantages and disadvantages of the method are discussed.

Key-Words: - ANN, economic dispatch, thermal power system, system marginal cost

1 Introduction
Nowadays, world increasingly concerns about climate changes, air quality and greenhouse gas emissions. This has led to stricter emissions regulations for electric power systems mainly concerning the thermal electric power generating units (coal, oil and gas). The use of the thermal units is inevitable because of their controllability and availability in contrast to alternative energy sources (i.e. wind-turbines, photovoltaic panels, hydro-turbines, etc) characterized by intermittent power production. However, for environment protection purposes financial penalties are imposed for Green House Gas (GHG) emissions offering a substantial economic advantage to clean alternative sources. Penalties vary for each thermal unit according to the GHG emissions and it is added to the operation economic cost of the unit. Taking into consideration this additional cost the optimal operation of the thermal units, known as economic dispatch, can be solved using the technique of Lagrange multipliers [1-2].

The objective of this paper is to present an alternative solution based on artificial neural networks specialized in the solution of continuous, nonlinear constrained optimization problems, based on external penalty functions [3]. In section 2 the mathematical base of the optimal operation of a thermal electric power system is formed. In section 3 the mathematical base of the developed ANN is presented and in section 4 the indicative case studies are presented. Finally, in section 5 the advantages and the disadvantages of the classic Lagrange technique and the presented method are commented.

2 Mathematical Bases for Optimal Operation of All-Thermal Power System without Losses
The optimum operation of the thermal and of the hydro-thermal continental power systems (with or without transmission losses) has been already thoroughly analyzed in the literature [1-2]. In this section a short review of the optimal operation of the thermal units without considering transmission system losses is presented, as in Fig. 1.

![Fig. 1. Power system with N thermal units in the same bus (without transmission loss).](image-url)
power, $P_{\text{min} \text{TH}_j}$ and the technically maximum active power $P_{\text{max} \text{TH}_j}$ of the unit. The respective constraints are formulated as:

$$P_{\text{min} \text{TH}_j} \leq P_{\text{TH}_j} \leq P_{\text{max} \text{TH}_j} \quad (1)$$

Where $j=1, 2, ..., N$.

The respective fuel cost of each unit is given by the function $F_{\text{TH}_j}(P_{\text{TH}_j})$, which is usually a polynomial of second or third order of $P_{\text{TH}_j}$. This means that:

$$F_{\text{TH}_j}(P_{\text{TH}_j}) = a_j + b_j \cdot P_{\text{TH}_j} + c_j \cdot P_{\text{TH}_j}^2 + d_j \cdot P_{\text{TH}_j}^3 \quad (2)$$

Where $a_j$, $b_j$, $c_j$, $d_j$ are the proper economic coefficients and they are considered known next.

For each time interval $DT$ with constant load demand $P_D$ the total fuel cost $F_{\text{tot}}$ of the power system is the sum of the individual fuel costs of the units, which is calculated as:

$$F_{\text{tot}} = \sum_{j=1}^{N} F_{\text{TH}_j}(P_{\text{TH}_j}) \quad (3)$$

Active power balance without considering transmission losses can be written for time interval $DT$, as:

$$\sum_{j=1}^{N} P_{\text{TH}_j} = P_D \quad (4)$$

The target is to determine the generating levels of the units such that the total fuel cost $F_{\text{tot}}$ is minimized and active power balance constraint is fulfilled. This means that the total fuel cost $F_{\text{tot}}$ should be minimized taking into consideration the respective constraint of eq. (4). This can be solved by including the eq. (4) in eq. (3) by using an unknown Lagrange multiplier, i.e. $\lambda$. Then eq. (3) is modified as:

$$L_{\text{tot}} = \sum_{j=1}^{N} F_{\text{TH}_j}(P_{\text{TH}_j}) - \lambda \cdot \left( \sum_{j=1}^{N} P_{\text{TH}_j} - P_D \right) \quad (5)$$

The optimality conditions are obtained by setting the partial derivatives of $L_{\text{tot}}$ with respect to $P_{\text{TH}_j}$ equal to 0:

$$\frac{\partial L_{\text{tot}}}{\partial P_{\text{TH}_j}} = 0 \Rightarrow \frac{dF_{\text{TH}_j}(P_{\text{TH}_j})}{dP_{\text{TH}_j}} - \lambda = 0 \Rightarrow \lambda = \frac{dF_{\text{TH}_j}(P_{\text{TH}_j})}{dP_{\text{TH}_j}} \quad \text{for } i = 1, 2, ..., N$$

$$\Rightarrow \lambda = \frac{dF_{\text{TH}_1}(P_{\text{TH}_1})}{dP_{\text{TH}_1}} = \ldots = \frac{dF_{\text{TH}_N}(P_{\text{TH}_N})}{dP_{\text{TH}_N}} \quad (6)$$

This means that during time interval $DT$ thermal units should share the load in a way that their respective incremental costs are equal to Lagrange coefficient, $\lambda$. Lagrange coefficient $\lambda$ is known as system marginal cost (SMC) and differs between time intervals with different loads. After the determination of $\lambda$, the generating levels of the units $P_{\text{TH}_j}$ can be calculated.

If the function cost is a polynomial of second order ($d_j = 0$ for all $i$), eq. (6) is simplified as:

$$\lambda = \frac{dF_{\text{TH}_j}(P_{\text{TH}_j})}{dP_{\text{TH}_j}} = b_j + 2 \cdot c_j \cdot P_{\text{TH}_j} \quad \text{for } j = 1, 2, ..., N \quad (7)$$

Resolving eq. (7) the generating level of the $j^{th}$ unit can be calculated as a function of the respective system marginal cost $\lambda$:

$$P_{\text{TH}_j} = \frac{\lambda - b_j}{2 \cdot c_j} \quad \text{for } j = 1, 2, ..., N \quad (8)$$

System marginal cost $\lambda$ is obtained if the active power balance (eq. (4)) constraint is used as following:

$$(4) \& (8) \Rightarrow \sum_{j=1}^{N} \lambda - \frac{b_j}{2 \cdot c_j} = P_D \Rightarrow \lambda \cdot \left( \sum_{j=1}^{N} \frac{1}{2 \cdot c_j} \right) - \frac{b_j}{2 \cdot c_j} = P_D \Rightarrow \lambda = \left( P_D + \sum_{j=1}^{N} \frac{b_j}{2 \cdot c_j} \right) \left( \sum_{j=1}^{N} \frac{1}{2 \cdot c_j} \right) \quad (9)$$

If the function cost is a polynomial of third order, eq. (6) is simplified as:

$$\lambda = b_j + 2 \cdot c_j \cdot P_{\text{TH}_j} + 3 \cdot d_j \cdot P_{\text{TH}_j}^2 \quad \text{for } j = 1, 2, ..., N \quad (10)$$

From eq. (10) the generating level of the $j^{th}$ unit can be calculated as a function of the respective system marginal cost $\lambda$:

$$P_{\text{TH}_j} = \frac{-c_j + \sqrt{c_j^2 + 3 \cdot d_j \cdot (\lambda - b_j)}}{3 \cdot d_j} \quad \text{for } j = 1, 2, ..., N \quad (11)$$

Using Gauss-Seidel classic technique the determination of $\lambda$ is achieved as following:

1) Initialization of $\lambda^{(0)}$ and $\lambda^{(1)}$.

2) Determination of the generating levels of the units $P_{\text{TH}_j}$ based on eq. (11).

3) Calculation of the active power imbalance $e^{(k)}$:

$$e^{(k)} = \sum_{j=1}^{N} P_{\text{TH}_j}^{(k)} - P_D \quad (12)$$

4) Step to next iteration:

$$k = k + 1 \quad (13)$$

5) Check of the first iteration: if $k=1$, step 2 is executed else continue with step 6.

6) Check for convergence: if $|e^{(k)}| \leq$ convergence limit then step 8 is executed else we continue with step 7.

7) Determination of $\lambda^{(k)}$: Based on Newton-Raphson method the $\lambda^{(k)}$ is determined as:

$$\lambda^{(k)} = \lambda^{(k-1)} - \frac{\lambda^{(k-1)} - \lambda^{(k-2)} \cdot e^{(k-2)}}{e^{(k-1)} - e^{(k-2)}} \cdot e^{(k-1)} \quad (14)$$

Afterwards, step 2 is executed.

8) After the determination of $\lambda$ and the generating
levels of the units, \( P_{Thh} \), the \( N \) inequalities from the technical constraints should be checked. If a violation of an inequality exists, i.e. \( P_{Thh} < P_{minThh} \), the output active power is set to the respective value of the violated constraint, i.e. \( P_{Thh} = P_{minThh} \). Afterwards the rest load demand \( P_D \) is calculated as \( P_D = P_0 - P_{Thh} \) and the optimization process is repeated from step 1 without considering the respective unit, i.e. the unit 2.

The Gauss-Seidel method can be also applied in case of second order polynomial function cost and in the third iteration the active power imbalance error becomes zero.

3 Mathematical Base for Continuous, Nonlinear, Constrained Optimization Problems with Exterior Penalty Functions solved with ANNs

3.1 General

A constrained minimization problem can be stated as the following one: “Find \( \bar{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \) which minimizes the continuous, non-linear scalar function \( f(\bar{x}) = f(x_1, x_2, \ldots, x_n) \) subject to the inequality constraints \( g_i(\bar{x}) \geq 0 \) for \( i = 1, 2, \ldots, m \).” It is noted that inequality constraint \( w(\bar{x}) \leq 0 \) can be replaced by the constraint inequality \( -w(\bar{x}) \geq 0 \), while each equality constraint can be represented equivalently by two inequality constraints, i.e. \( h(\bar{x}) = 0 \) can be replaced by the two inequalities \( h(\bar{x}) \geq 0 \) and \( h(\bar{x}) \leq 0 \). The aforementioned problem can be transformed to an unconstrained problem having the following energy pseudo-cost function:

\[
E(\bar{x}) = f(\bar{x}) + \sum_{i=1}^{m} k_i \cdot P_{ig}(g_i(\bar{x}))
\]

(15)

Where \( k_i \) are the multiplication penalty factors \( (k_i > 0) \) and \( P_{ig}(g_i(\bar{x})) \) are the quadric penalty function terms defined by:

\[
P_{ig}(g_i(\bar{x})) = \left( \min \left\{ 0, g_i(\bar{x}) \right\} \right)^2
\]

(16)

\[
P_{ig}(g_i(\bar{x})) = \begin{cases} 0, & \text{if } g_i(\bar{x}) \geq 0 \\ >0, & \text{if } g_i(\bar{x}) < 0 \end{cases}
\]

If the first partial derivatives of the constraints \( g_i(\bar{x}) \) are continuous in \( \mathbb{R}^n \), then this will also happen for the respective penalty functions of eq. (16), as:

\[
\frac{\partial P_{ig}(g_i(\bar{x}))}{\partial x_j} = 2 \cdot \left\{ \min \left\{ 0, g_i(\bar{x}) \right\} \right\} \cdot \frac{\partial g_i(\bar{x})}{\partial x_j}
\]

(17)

The minimization of the energy function is succeeded by using a gradient technique [3]. The respective system of differential equations \( \forall j = 1, 2, \ldots, n \) is:

\[
dx_j = -\mu_j \left( \frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} k_i \cdot S_i \cdot \frac{\partial g_i(\bar{x})}{\partial x_j} \right)
\]

(18)

Where,

- \( t \) is the auxiliary parameter of time,
- \( \mu_j \) is the ANN learning rate lower limited by a constant positive number \( \mu_0 \) defined by the following equation:

\[
\mu_j(t) = \max \{ \mu_{init,j} \cdot \exp \left( \frac{t}{T_{\mu,j}} \right), \mu_{0,j} \}
\]

(19)

with \( \mu_{init,j} > 0, \mu_{0,j} > 0, T_{\mu,j} > 0 \)

- \( k_i \) is the multiplication penalty factor upper limited by a constant positive number \( k_0 \) defined by the following equation:

\[
k_i(t) = \min \{ k_{init,i} \cdot \exp \left( \frac{t}{T_{k,i}} \right), k_{0,i} \}
\]

(20)

with \( k_{init,i} > 0, k_{0,i} > 0, T_{k,i} > 0 \)

- \( S_i \) is a proper step function, which implements the minimization problem of (16):

\[
S_i(g_i(\bar{x})) = \begin{cases} 0, & \text{if } g_i(\bar{x}) \geq 0 \\ 1, & \text{if } g_i(\bar{x}) < 0 \end{cases}
\]

(21)

The block diagram of a constrained nonlinear optimization ANN based on the penalty method is shown in Fig. 2 and is proposed by Chua et al [6].

3.2 Optimal operation of all-thermal power system without losses

The target is to minimize the total fuel cost \( F_{tot} \) of the power system given by eq. (3). In this minimization problem the unknown variable vector \( \bar{x} \) consists of the generating levels of the units \( P_{Th} \) while the function \( f(\bar{x}) \) represents total fuel cost function, \( F_{tot} \):

\[
\bar{x} = (P_{Th1}, P_{Th2}, \ldots, P_{Th,N})^T \in \mathbb{R}^N
\]

(22)

\[
f(\bar{x}) = \sum_{i=1}^{N} \left( a_i + b_i \cdot P_{Th} + c_i \cdot P_{Th}^2 + d_i \cdot P_{Th}^3 \right)
\]

(23)

The technical constraints of the thermal units of eq. (1) are transformed to the following inequalities:

\[
g_i(\bar{x}) = P_{maxTh} - P_{Th} \geq 0
\]

(24)

\[
g_{i+n}(\bar{x}) = P_{Th} - P_{minTh} \geq 0
\]

(25)

Where \( i = 1, 2, \ldots, N \).

The equality of the active power balance without considering the losses of the transmission system given by eq. (4) is transformed to the following to inequalities:
Fig. 2. Block diagram of a constrained nonlinear optimization ANN based on the penalty method.

\[
g_{2_{N+1}}(\bar{x}) = \sum_{j=1}^{N} P_{ij} - P_D \geq 0 \tag{26}
\]

\[
g_{2_{N+2}}(\bar{x}) = P_D - \sum_{j=1}^{N} P_{ij} \geq 0 \tag{27}
\]

For \( N \) generating units 2\( \cdot N + 2 \) inequality constraints are obtained.

The first partial derivatives of the minimization function and of the constraints are the following:

\[
\frac{\partial f}{\partial P_{ij}} = b_j + 2 \cdot c_j \cdot P_{ij} + 3 \cdot d_j \cdot P_{ij}^2 \tag{28}
\]

\[
\frac{\partial g_i}{\partial P_{ij}} = \begin{cases} 
 1, & j = i - N \\
 0, & j \neq i - N 
\end{cases} \quad i = N + 1, \ldots, 2 \cdot N \tag{29}
\]

\[
\frac{\partial g_{2_{N+1}}}{\partial P_{ij}} = 1 \tag{30}
\]

\[
\frac{\partial g_{2_{N+2}}}{\partial P_{ij}} = -1 \tag{31}
\]

Where \( j = 1, 2, \ldots, N \).

The ANN technique used in this study is an iterative process, where eq. (18) can easily be transformed into the following discrete-time iterative equation:
\[
P_{THj}^{(s+1)} = \left\{ \begin{array}{l}
P_{THj}^{(s)} - \mu_j \left( \frac{\partial f}{\partial P_{THj}} \right) x_j^{(s)} - \mu_j \left( \frac{\partial g_j}{\partial P_{THj}} \right) x_j^{(s)} \\
\sum_{i=1}^{m} k_i x_i \cdot S_i \cdot t_i \cdot g_j \left( x_j^{(s)} \right) \end{array} \right. \quad (33)
\]

Where, \( \ell \) is the discrete variable for time or alternatively the “training epochs of ANN technique”. The proposed ANN training process will be terminated if one of the following two stopping criterions is reached:

- the generating levels of the units \( P_{THj} \) are stabilized (the variation between two epochs should be smaller than \( \text{limit convergence} \))
- the maximum number of epochs is exceeded (larger than \( \text{max\_epochs} \)).

4 Application of the Proposed Method

4.1 Case study

A simple thermal electric power system without transmission losses is assumed. It consists of three thermal units with the following fuel cost functions and technical constraints:

\[
F_{TH1}(P_{TH1}) = \left\{ \begin{array}{l}
10^3 P_{TH1} + 0.03 P_{TH1}^2 + 10^{-4} P_{TH1}^3 \\
\text{m.u.} \\
\text{h}
\end{array} \right.
\]

\[
F_{TH2}(P_{TH2}) = \left\{ \begin{array}{l}
8760 + 75 P_{TH2} + 0.1 P_{TH2}^2 + 5 \times 10^{-5} P_{TH2}^3 \\
\text{m.u.} \\
\text{h}
\end{array} \right.
\]

\[
F_{TH3}(P_{TH3}) = \left\{ \begin{array}{l}
4800 + 90 P_{TH3} + 0.03 P_{TH3}^2 + 8 \times 10^{-5} P_{TH3}^3 \\
\text{m.u.} \\
\text{h}
\end{array} \right.
\]

\[
50 \leq P_{TH1} \leq 350 \text{ [MW]}
\]

\[
50 \leq P_{TH2} \leq 350 \text{ [MW]}
\]

\[
50 \leq P_{TH3} \leq 450 \text{ [MW]}
\]

Where “m.u.” is the monetary unit.

4.2 Application of the ANN method

Two load levels are examined: 460 MW and 260 MW. The last one activates the technical constraints of the thermal units.

In all cases of ANN training the following parameters values have been used after a few tests so that the algorithm remains stable:

\[
\begin{align*}
\mu_{init} &= 10^{-3} \quad \mu_{j,0} = 0.06 \quad T_{\mu_{0,j}} = 200 \text{ for } j=1, 2, 3 \\
k_{init} &= 10 \quad k_{0,j} = 100 \quad T_{k_{0,j}} = 215 \text{ for } i=1, \ldots, 8 \\
\text{limit convergence} &= 10^{-4} \quad \text{max\_epochs} = 100000
\end{align*}
\]

The initialization of the generating levels of the units is done as following:

\[
P_{THj} = \frac{P_{min_{THj}} + P_{max_{THj}}}{2} \quad \text{for } j = 1, 2, 3 \quad (34)
\]

In case of 460 MW demand load the generating levels of the units are obtained after 12650 epochs, approximately. The analytical results of the generating levels, the power balance error (difference between power produced by the units and load demand), the epochs and the total cost of the system are registered in Table 1. Unfortunately, the active power balance is approximately satisfied, as the respective error is almost 0.8 MW. This happens as the limit of convergence depends on the variation of the generating levels and not on balance of equality, as happening in the classical Lagrange method of section 2. In Figures 3 and 4 the generating levels of the thermal units and the total fuel cost with respect to the training epochs are presented. It is proven by these Figs that the method is stable for the chosen parameters values.

In case of 260 MW load demand the generating levels of the units are obtained after 820 epochs,
approximately. The respective results are registered in Table 1, while in Figs. 5 and 6 the graphics of the thermal units generating levels and the total fuel cost with respect to the epochs are shown, respectively. The active power balance error is approximately 0.2 MW which is smaller than in the previous case.

Fig. 5. Case of demand load 260 MW: the generating levels of the thermal units with respect to the number of epochs.

Fig. 6. Case of demand load 260 MW: the total fuel cost with respect to the number of epochs.

4.3 Comparison with classical technique

In Table 1 the results of the generating levels of the thermal units, the power balance error (difference between produced power by the units and the load demand), the epochs and the total cost of the system are registered for both of previously examined cases (460 MW and 260 MW of load demand) and both minimization techniques (ANN and the classical Lagrange method of section 2).

It is concluded that:
- Both methods give practically the same results.
- The ANN based method needs significant more epochs for convergence than the classical technique.
- The ANN based method does not satisfy exactly the active power balance constraint. This happens as the limit of convergence depends on the variation of the generating levels and not on balance of equality, as happening in the classical Lagrange method.
- The ANN based method seems to be slightly superior to the classical technique in terms of fuel cost minimization. This happens because in both cases the total production of the generating of the units is slightly smaller than the total load demand.

5 Conclusions

This paper presents an ANN based optimization method applied to a thermal electric power system without considering transmission losses. The proposed method is specialized in continuous, nonlinear, constrained optimization problems with external penalty functions. The respective results are satisfactory if the proper parameters of learning rate and penalty multiplicative factors are properly selected. However, the proposed method is characterized by the three following drawbacks: (1) the number of iterations (epochs) needed is significant larger than those required in classical Lagrange technique, (2) active power balance constraint is not satisfied completely, (3) the system marginal cost (SMC), which is a crucial parameter for the operation of the power system is not directly obtained. It should be calculated afterwards in contrast to the classical Lagrange optimization technique where it is determined within the optimization process (see eq. (14)). In future work, the proposed method can be improved by using augmented Lagrange multiplier ANN method, which can easily remove the last two disadvantages.

<table>
<thead>
<tr>
<th>Method</th>
<th>ANN</th>
<th>Classic</th>
<th>ANN</th>
<th>Classic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target Load</td>
<td>460 MW</td>
<td>260 MW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P_{TH1}$ [MW]</td>
<td>110.97</td>
<td>110.58</td>
<td>49.99</td>
<td>50.00</td>
</tr>
<tr>
<td>$P_{TH2}$ [MW]</td>
<td>157.50</td>
<td>157.83</td>
<td>109.27</td>
<td>108.78</td>
</tr>
<tr>
<td>$P_{TH3}$ [MW]</td>
<td>190.80</td>
<td>191.58</td>
<td>100.52</td>
<td>101.22</td>
</tr>
<tr>
<td>Error [MW]</td>
<td>-0.73</td>
<td>0.01</td>
<td>-0.22</td>
<td>0.00</td>
</tr>
<tr>
<td>Epochs[-]</td>
<td>12647</td>
<td>8</td>
<td>824</td>
<td>13</td>
</tr>
<tr>
<td>Cost[m.u./h]</td>
<td>61171</td>
<td>61252</td>
<td>40232</td>
<td>40254</td>
</tr>
</tbody>
</table>
References:


