Available online at www.sciencedirect.com
SCIENCE DIRECT®
Fuel 82 (2003) 693–703
FUEL
www.fuelfirst.com

Constrained optimization of combustion in a simulated coal-fired boiler using artificial neural network model and information analysis☆

Ji-Zheng Chu[a], Shyan-Shu Shieh[b], Shi-Shang Jang[c,*], Chuan-I Chien[c], Hou-Peng Wan[d], Hsu-Hsun Ko[d]

[a]*Department of Automation, Beijing University of Chemical Technology, Beijing, People's Republic of China*
[b]*Department of Occupational Safety and Hygiene, Chang Jung University, Tainan, Taiwan, ROC*
[c]*Department of Chemical Engineering, National Tsing-Hua University, 101, Section 2 Kuang Fu Road, Hsin-Chu 300, Taiwan, ROC*
[d]*Energy and Resources Laboratories, ITRI, Hsin-Chu, Taiwan, ROC*

Received 19 March 2002; accepted 4 October 2002; available online 14 November 2002

Abstract

Combustion in a boiler is too complex to be analytically described with mathematical models. To meet the needs of operation optimization, on-site experiments guided by the statistical optimization methods are often necessary to achieve the optimum operating conditions. This study proposes a new constrained optimization procedure using artificial neural networks as models for target processes. Information analysis based on random search, fuzzy c-mean clustering, and minimization of information free energy is performed iteratively in the procedure to suggest the location of future experiments, which can greatly reduce the number of experiments needed. The effectiveness of the proposed procedure in searching optima is demonstrated by three case studies: (1) a bench-mark problem, namely minimization of the modified Himmelblau function under a circle constraint; (2) both minimization of NO_x and CO emissions and maximization of thermal efficiency for a simulated combustion process of a boiler; (3) maximization of thermal efficiency within NO_x and CO emission limits for the same combustion process. The simulated combustion process is based on a commercial software package CHEMKIN, where 78 chemical species and 467 chemical reactions related to the combustion mechanism are incorporated and a plug-flow model and a load-correlated temperature distribution for the combustion tunnel of a boiler are used.

Keywords: Combustion; Constrained optimization; Information analysis; Artificial neural networks

1. Introduction

Although there have been a lot of experimental and theoretical studies on the basic physical and chemical principles of a boiler's operation, and great advance has been made in understanding various aspects of the operation, it is still impracticable, as Faravelli and co-workers [1] have just pointed out, to couple detailed fluid dynamics and kinetics in the combustion system design even with the continuous increase of computer power, not to mention to simulate a boiler and its various auxiliary subsystems as a whole.

If fuel and environment conditions are specified, the thermal efficiency of a given boiler depends mainly on the air to fuel ratio, and on the distribution of air and fuel to burners at different locations if two or more burners are used. For different fuels and different furnace configurations, the best air to fuel ratio and the best air and fuel distributions are surely different, which may be roughly estimated by analysis but can only be determined accurately by testing runs. As for the control of trace pollutant emissions (in terms of ppm or ppb NO_x, SO_2, etc.) of a boiler, all of such uncertain factors as fuel compositions, complexity of flow and temperature fields in the burning chamber, and the complicated mechanism of chemical reactions make model prediction highly unreliable.

The inability of theoretical analysis makes empirical approaches necessary to explore the optimum operation conditions of a boiler. Nordin et al. [2] used fractional factorial design procedure and identified the important factors of NO_x reduction in a fluidized bed combustor (FBC) with primary measures and selective non-catalytic

* Corresponding author. Tel.: +886-3-5735294; fax: +886-3-5715408.
E-mail address: ssjang@che.nthu.edu.tw (S.S. Jang).
☆ Published first on the web via Fuelfirst.com—http://www.fuelfirst.com

0016-2361/03/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0016-2361(02)00338-1

reduction. Henttonen et al. [3] performed an optimizing control of NO_x and SO_2 emissions in the FBC process based on static models which were regressed from experimental data. George [4] suggested a boiler test system in order to monitor NO_x emissions through optimizing various operation conditions in a comprehensive and systematical manner, and claimed that achieving NO_x emission reduction targets does not necessarily mean equipment retrofits or poor boiler performance.

The capability of artificial neural networks (ANNs) as a universal modeling tool has been widely recognized in the last 20 years, and Cybenko [5] showed that ANNs could approximate any arbitrary non-linear functions. ANNs offer an alternative approach to model process behavior, as they do not require a priori knowledge of the process phenomena. They learn by extracting imbedded patterns from data that describe the relationship between the inputs and the outputs in any given process phenomenon. When appropriate inputs are applied to an ANN, the ANN acquires 'knowledge' from the environment in a process known as 'learning'. As a result, the ANN assimilates information that can be recalled later. ANNs are capable of handling complex and non-linear problems, processing information rapidly and reducing the engineering effort required in model development.

ANNs have been successfully applied to a variety of problems such as process fault diagnosis, system identification, pattern recognition, process modeling and control, and statistical time series modeling. Kalogirou [6] gave a review on the application of ANNs in energy systems. Reifman et al. [7] developed an intelligent emissions controller for fuel gas re-burning in coal-fired power plant, and in their study, a feed-forward neural network (FFN) was used to model the static non-linear relationships between the distribution of injected natural gas into the upper region of the furnace of a coal-fired boiler and the corresponding oxides of nitrogen emissions exiting the furnace. Zhou et al. [8] used an ANN model and genetic algorithms to optimize low NO_x pulverized coal combustion. Other instances of ANNs applied in solving combustion problems include: modeling the temporal evolution of a reduced combustion system [9], predicting coal ash fusion temperature [10], predicting coal/char combustion rate [11], conducting model predictive control in thermal power plant [12], etc.

It is obvious that optimization of complex systems such as the combustion process of a boiler is a trial-and-error process. In such an iterative process, experimental test is performed, and the test data are analyzed, and further test is suggested based on the analysis, and such iterations continue until satisfactory performance is achieved.

In this study, a novel optimization procedure is proposed by extending the experimental design method developed by the authors [13] to constrained cases, and is used to search for the best operation conditions of a simulated coal-fired combustion process. The proposed optimization procedure uses ANNs to model the relationship of the performance index with various operating variables, searches the built response surface under constraints to produce candidate points of next batch of test, determines the test points of next batch through information analysis. This procedure works iteratively and optimum conditions are expected after several batches of test. The major advantages of the proposed procedure are its abilities to deal with multi-variables, to precisely determine the number and location of future test experiments, to consider the non-analytical constraints, and to locate multiple optima.

The objective of this study is to demonstrate the effectiveness of the proposed optimization procedure in searching for the optimum operation conditions of a boiler. In the following context, a simplified model for the combustion process of a coal-fired boiler is built, and the optimization procedure is introduced and is applied to optimize the operation conditions of the simulated combustion process. It should be noted that the intent of this paper is to describe the proposed optimization procedure and to show its power in improving the overall performance of a boiler's combustion process to achieve as high thermal efficiency and as low pollutant emissions as possible and not to provide an in-depth analysis of a boiler's combustion behavior.

2. Model of the combustion process of a coal-fired boiler

The simulated boiler has a 17 m high vertical combustion tunnel as shown in Fig. 1. There are two burners installed at elevations 2 and 4 m, and each burner's full load is designed to be 2752 kg-coal/h. Coal and air are introduced into the tunnel through the two burners. The amount of coal and air to each burner can be regulated by controllers which are not shown in Fig. 1.

Fuel used is pulverized coal whose element analysis is listed in Table 1. Air to the burning tunnel has a volume

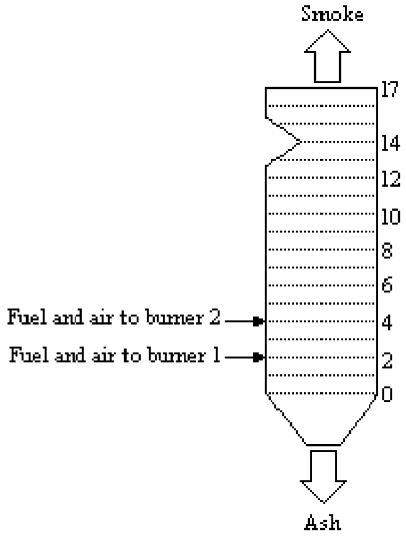

Fig. 1. Sketch of the burning tunnel of a boiler.

Table 1
Element analysis of coal used in calculation

Element	wt%	mol%
C	77.26	49.16
H	5.76	43.70
O	11.93	5.70
S	4.28	1.02
N	0.77	0.42
Total	100.00	100.00

composition of 78.4% N_2, 20.8% O_2, and 0.8% moisture. Theoretical consumption of air is 9.7529 N m^3/kg-coal to convert the coal into corresponding smoke components CO_2, H_2O, SO_2, and N_2. The low heat value of the coal is 32,387.7 kJ/kg-coal (or 7741 kcal/kg-coal).

The heat input to the tunnel from burner 1 and 2 are;

$$Q_1 = F_{c,1}\Delta H_c + F_{c,1}C_{p,c}(T_c - T_0) + V_{a,1}C_{p,a}(T_a - T_0) \quad (1)$$

$$Q_2 = F_{c,2}\Delta H_c + F_{c,2}C_{p,c}(T_c - T_0) + V_{a,2}C_{p,a}(T_a - T_0) \quad (2)$$

where, Q_1, Q_2 = input heat from burners 1 and 2, kJ/h; $F_{c,1}$, $F_{c,2}$ = pulverized coal to burners 1 and 2, kg/h; $V_{a,1}$, $V_{a,2}$ = air to burners 1 and 2, N m^3/h (std conditions are 0 °C and 1 atm); T_0 = temperature at which heating value of coal is measured, 25 °C; T_c = temperature of the pulverized coal, K; T_a = temperature of the feed air, K; $C_{p,c}$ = heat capacity of the pulverized coal, kJ/(kg K); $C_{p,a}$ = heat capacity of the feed air, kJ/(kg K); ΔH_c = low heating value of the pulverized coal, kJ/kg.

The heat loss caused by the smoke leaving the combustion tunnel can be determined as follows;

$$Q_s = V_s\{y_{s,CO}\Delta H_{CO} + C_{p,s}(T_s - T_0)\} \quad (3)$$

where, Q_s = heat loss with the leaving smoke, kJ/h; V_s = flowrate of smoke, N m^3/h; $y_{s,CO}$ = concentration of CO in the smoke, volume fraction; ΔH_{CO} = low heating value of CO, kJ/N m^3; $C_{p,s}$ = heat capacity of smoke, kJ/(N m^3 K), calculated from smoke composition; T_s = temperature of smoke at the outlet of the combustion tunnel.

As a primary approximation, the thermal efficiency, η, used in this study is defined as follows;

$$\eta = \left(1 - \frac{Q_s}{Q_1 + Q_2}\right) \times 100 \quad (4)$$

where the heat loss by unburnt carbon in the ash is assumed negligible.

A commercial software package CHEMKIN developed by Leeds University (Reaction Design, 2000. http://www.reactiondesign.com), was used to calculate the concentrations of NO_x and CO in the outlet smoke. One-dimension plug-flow through the tunnel was assumed in the calculation of CHEMKIN. For simplicity, homogeneous reaction is also assumed in the combustion process that is to say, the fuel coal is regarded as a gaseous mixture of elements C, H, O, S, and N.

CHEMKIN requires to specify chemical reactions taking place in the combustion process. 467 chemical reactions were collected and incorporated into the calculation of the combustion process, where 78 chemical species and radicals are correlated by the reaction formula [14].

To start CHEMKIN, a temperature distribution along the combustion tunnel is also needed, which is set by the following equation;

$$T = \alpha_T T_r \quad (5)$$

where T = temperature at current conditions, K; T_r = temperature at reference conditions described by the curve shown in Fig. 2, K; α_T = correlation factor as a function of the heat input to the burning tunnel, which is set quite arbitrarily in this study by the following correlation;

$$\alpha_T = \begin{cases} \left(\dfrac{Q_1}{Q_{r,1}}\right)^{0.25} & \text{(for section 1)} \\ \left(\dfrac{Q_2 + Q_{s,1}}{Q_{r,2} + Q_{sr,1}}\right)^{0.25} & \text{(for section 2)} \end{cases} \quad (6)$$

$Q_{r,1}$, $Q_{r,2}$ = heat input from burner 1 and 2 as defined by Eqs. (1) and (2), respectively, at reference conditions; $Q_{s,1}$, $Q_{sr,1}$ = heat of smoke from section 1 to section 2, as defined by Eq. (3), at current and reference conditions, respectively.

Although the temperature profile shown in Fig. 2 was taken from a pilot boiler and Eqs. (5) and (6) were formulated by roughly fitting the running data of that boiler, it must be emphasized that the temperature profile established in this way is rather crude and is just served to construct a qualitatively correct model of combustion.

For simplicity, we express the coal flow rate by an auxiliary quantity L_c, which is the percentage of actual coal flow rate in the burner's full load, and the air flow rate by air ratio α_a, which is defined as the ratio of actually fed air over that theoretically needed. Since there are two burners, CHEMKIN simulates burner 1 (section 1: 0–3 m of the burning tunnel), and burner 2 (section 2: 3–17 m of the tunnel) sequentially, and the calculation procedure is as the following:

(a) Give the flow rates of pulverized coal and air which are fed to burners 1 and 2, and determine the temperature

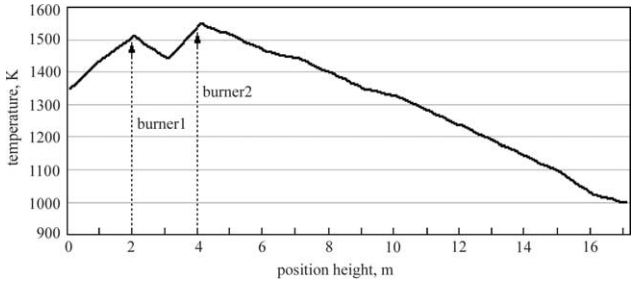

Fig. 2. Temperature distribution of the combustion tunnel at reference conditions ($L_{c,1} = L_{c,2} = 90\%$ and $\alpha_{a,1} = \alpha_{a,2} = 1.11$).

distribution in the combustion tunnel according to Eqs. (5) and (6).
(b) Sum up coal and air to burner 1 to form a single stream, and use CHEMKIN to determine the flow rate and composition of the flow at the outlet of section 1.
(c) Sum up coal and air to burner 2 and the flow leaving burner 1 to form a single stream, and use CHEMKIN to determine the flow rate and composition of the flow at the outlet of section 2.
(d) Calculate thermal efficiency and pollutant emissions.

It should be noted that the above model for combustion is roughly estimated in the following aspects. (a) Plug-flow is assumed through the combustion tunnel. (b) Homogenous reaction is used in the combustion process and the contribution of unburnt char in the ash to NO_x and CO formation is not included. (c) Temperature profile used in this study is just an estimate because it is correlated to complex heat transfer, flow field and other factors. (d) Thermal efficiency defined in Eq. (4) is incomplete, since heat loss caused by ash and water discharge, by incompletely burnt coal particles, and by surface convection with and radiation to the atmosphere, etc. is neglected. (e) The included chemical species and reactions are approximate to the true mechanism of chemical change in the combustion process. (f) Compositions of coal and air are simplified without considering ash and water.

The above considerations and assumptions in model building will surely produce some shift and distortion to the true relationship among all the variables of the combustion process. As a primary approximation, however, the model is based on the fundamental material and energy balance equations and includes all the main aspects of the combustion process, and it is our belief that the model is qualitatively correct and is complicated enough to serve as an example to demonstrate the effectiveness of an optimization procedure.

3. Constrained optimization guided by information analysis

For a general description of an optimization problem, we have a performance index to be minimized and it can be expressed by the following analytical function f

$$f = f(x, y) \tag{7}$$

with

$$y = g(x) \tag{8}$$

subjected to the constraints

$$H_a(x, y) \le 0 \tag{9}$$

$$H_n(x, y) \le 0 \tag{10}$$

where $x = [x_1, \ldots, x_{N_i}]$ is the vector of independent design/operation variables; $y = [y_1, \ldots, y_{N_o}]$ is the vector of the output variables which are functions of x, and the functions (g) can be a physical existence, e.g. a chemical plant, or a system of mathematical equations; H_a is constraints which can be analytically calculated from known x and y, and H_n is non-analytical constraints.

In the case that the original model, functions g, is expensive to be realized to get a response (it usually means to perform complicated experiment or to spend long computing time), it is worthwhile to spend more effort to explore the existing response data in order to make a wiser suggestion of the future generation of test points. There hence exists a compromise that the total effort spent both at making suggestion and getting response is minimal in the time limit of working schedule. Because of the rapid development of computing machines, it is almost always desirable to reduce the number of experiments in the optimization when experiment is needed to get a response.

Based on the above idea, we have the working logic of the proposed optimization procedure as sketched in Fig. 3. To initiate the optimization procedure, arrangement is given to the initial batch of experiments by using, for instance, an orthogonal array to all the identified independent design/operation variables over several arbitrarily defined parametric levels covering the estimated ranges. The whole procedure is iterative and is described stepwise in the following context.

3.1. Correlation of the experimental data with ANNs

After each batch of experiments, all the existing response data from the original model are correlated with an artificial neural network (ANN) which may be called a secondary model as contrasted to the original one. To have any usefulness, the secondary model must be built more easily and must be run more cheaply than the original model, and it is clear that ANNs are very suitable for this purpose. In addition constraints, which cannot be analytically expressed, are also modeled by ANNs.

The common feed-forward ANNs of three layers are used in this study, and their general architecture is depicted elsewhere [13]. The input variables of the network are the design/operation variables $x^p = [x_1^p, \ldots, x_{N_i}^p]$. The hidden layer consists of N_h neurons, whose output is given by

$$h_j(x^p) = \sum_{i=1}^{N_i} w_{ji}^h x_i{}^p + b_j^h, \qquad j = 1, 2, \ldots, N_h \tag{11}$$

The outputs of the whole ANN represent the predicted response (or controlled) variables, $y^p = [y_1^p, \ldots, y_{N_o}^p]$, and are determined by

$$\hat{y}_k(x^p) = \sum_{j=1}^{N_h} w_{kj}^o z[h_j(x^p)] + b_k^o, \qquad k = 1, 2, \ldots, N_o \tag{12}$$

where N_o is the dimensions of the output vector, the input and output component of the pth data pair are defined by

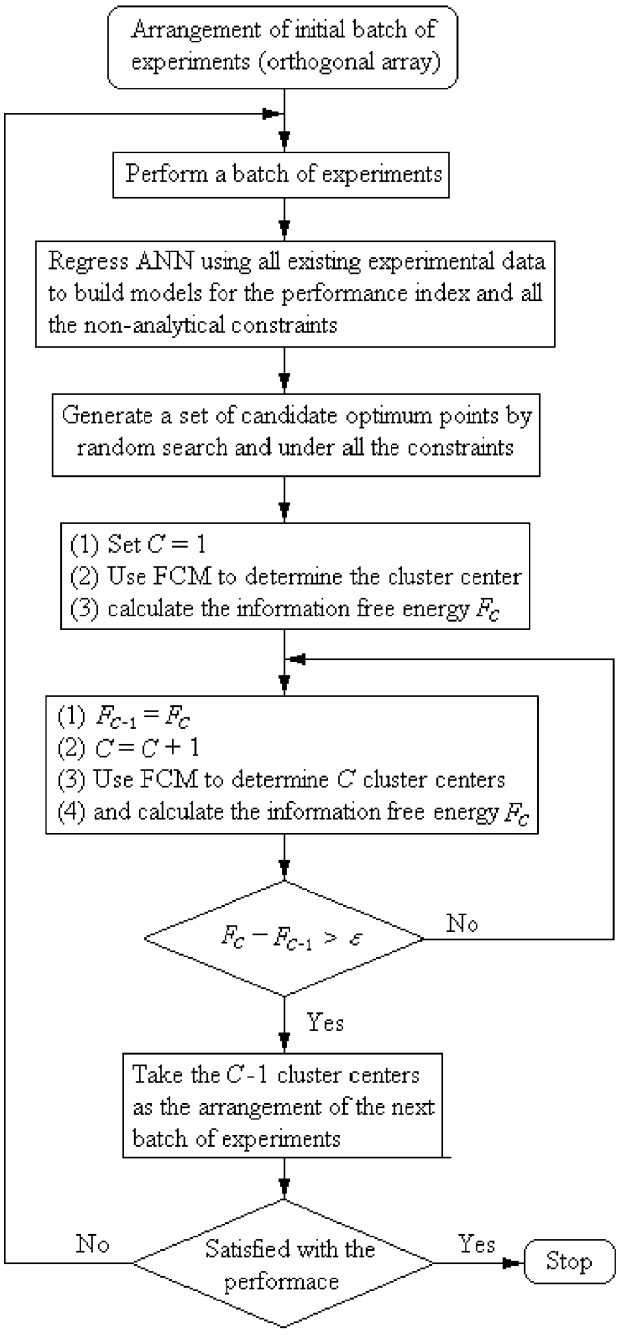

Fig. 3. Optimization algorithm guided by information analysis.

$\{x^p, y^p\}$; w^{o}_{kj} and w^{h}_{ji} are weights between the output and the hidden layers and weights between the hidden and the input layers, respectively; b^{o}_{k} and b^{h}_{j} are biases in the output and hidden layers, respectively; and z is the output in the hidden layer. The hyperbolic tangent activation function [15] is used. The sum square error, E, which represents the deviation of the predicted values from targeted ones is used to evaluate the ability of the network

$$E = \frac{1}{P}\sum_{p=1}^{P}\sum_{k=1}^{N_o}[y_k(x^p) - \hat{y}_k(x^p)]^2 \tag{13}$$

where P is the number of experimental data pairs. The pseudo-Gauss-Newton method [16,17] is used for training. Due to the small number of training data, a statistical technique called the leave-one-out (LOO) cross-validation scheme [18] is used.

3.2. Production of candidate optimum points

This step is to get an overall understanding of the response surface represented by the secondary model in the whole range of design/operation variables with a random search method. The result of this step is the production of a near-optimum population of points, or candidate optimum points, fulfilling all the constraints.

In product and process development the feature of interest is the optimal operating condition. Multiple local optima are frequently encountered. It is often necessary to rate alternative local optima based on secondary objectives such as safety, robustness, and so on. Therefore a non-gradient search technique is adopted here. Implementation of random search in this step is stated in detail elsewhere [13]. Random search converges theoretically to the global optimum, and several non-gradient optimization methods are detailed by Jang et al. [19].

3.3. Suggesting the locations of further experiments according to the result of information analysis on the candidate optimum points

In this step, iterations are performed to suggest the most promising points where optima may locate. In the iterations, a quantity called information free energy (F) is minimized by selecting the number of clusters about the candidate optimum points produced in the previous step. The cluster centers serve as the suggested points of next batch of experiments.

For the system of N points generated in the previous step, if the centers of C clusters are determined by the fuzzy c-means algorithm [13,20] and if a fuzzy membership of the kth point with respect to the center of the ith cluster is calculated as μ_{ik} $(k = 1, 2, \ldots, N;\ i = 1, 2, \ldots, C)$, the information entropy is defined and evaluated as

$$S = \frac{1}{N}\left(\sum_{i=1}^{C}\sum_{k=1}^{N}\mu_{ik}\ln\mu_{ik} - \sum_{i=1}^{C}N_i\ln N_i\right) \tag{14}$$

where $N_i = \sum_{k=1}^{N}\mu_{ik}$ is called the fuzzy number of data of the ith cluster.

Another quantity called information energy is defined as

$$U = \sum_{j=1}^{C}\frac{N_j}{N}f[c^i, \hat{y}(c^i)] - f_{\min} \tag{15}$$

where $f_{\min}$ is the value of the minimum f recorded in the random search and $f[c^i, \hat{y}(c^i)]$ is the performance index evaluated at the center of the ith cluster whose location is c^i.

It is clear that U is just the expected value of the performance index.

The indices of information entropy and information energy are measures of how well a set of cluster centers represents the data points and how well a set of cluster centers performs if it is chosen as the next set of experiments, respectively. In order to avoid inconsistency of results from considering them separately, a balance is needed and is achieved by defining a composite information index, the information free energy (F) as follows;

$$F = U - TS \tag{16}$$

where temperature T is a normalization factor and is defined as

$$T = \frac{f_{\max} - f_{\min}}{\ln P} \tag{17}$$

where $f_{\max}$ is the value of the maximum f recorded in the random search and P is the total number of existing experiments.

It is evident that with the increase of clusters, information free energy will decrease when C is small, then pass a minimum at a certain value of C, and finally increase. Therefore, the terminating criterion of information analysis iteration as suggested above is

$$F_C - F_{C-1} > \varepsilon \tag{18}$$

where ε is a small positive number.

From the above definitions for information entropy, energy and free energy, it is clear that entropy increases and energy decreases with the increase of clusters and that the information free energy is a compromise of them two. At early stages of optimization, temperature is high, information entropy dominates, and more experiments (cluster centers) are suggested to cover the whole feasible area uniformly. Thus, helps to build a complete secondary model. At final stages, temperature is low, and information energy is the predominant part of information free energy, which guides experiments conducted at possible optimum points. Early batches of experiments build an approximate outline of the variation of the objective function in the feasible area, and the final batches locate precisely the position of the optimum point. As a result, the final secondary model is very accurate in the neighborhood of the optimum point, but is only qualitatively correct in other places. The evolution of the secondary model will be demonstrated in a bench-mark test in Section 3.5. More detailed discussion on the meaning of information entropy and information free energy is found elsewhere [13].

3.4. Checking the convergence criterion

The final step of the iteration in the optimization procedure is to check the convergence criterion which usually means to see the difference between the newly produced minima and those produced last time. The procedure is terminated if the difference is within a preset tolerance, otherwise the procedure is carried out by performing new experiments at the suggested points from the above step.

It should also be noted that the prediction of the secondary model at the suggested test points will be very close to the experimental results at the end of converging process, which can also be a convergence criterion as it is clear in the comparison of the mesh surfaces of the Himmelblau function and its ANN counterpart (the secondary model) in the following bench-mark test.

3.5. A Bench-mark test to the proposed optimization procedure

As a primary test, the above optimization procedure is used to detect the minima of a bench-mark problem which is the modified Himmelblau function defined for $-5 \le x_1 \le 5$ and $-5 \le x_2 \le 5$

$$Z(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2 + x_1 + 3x_2 + 57 \tag{19}$$

under the constraint

$$x_1^2 + x_2^2 \le 3.5^2 \tag{20}$$

The mesh surface of Eq. (19) is shown in Fig. 4(a), and the constraint circle is depicted on the contour plot in Fig. 4(b). Note that we here assume that the constraint equation (20) is unknown and should be treated as a non-analytical constraint in minimization.

In the test calculation, the performance index is $f = Z(x_1, x_2)$; the inequality (20) is regarded as a non-analytical constraint $H_n = x_1^2 + x_2^2 - 3.5^2 \le 0$ and is also modeled by a feedfoward ANN.

As contrasted to the inability of traditional experimental design methods to build an accurate model for the modified Himmelblau function and to find the global optimum, the proposed procedure can detect all the four minima with ease in nine batches of experiments. The evolution history is shown in Figs. 5 and 6.

The converging process as shown in Fig. 6, vividly explains the meaning of information free energy as a criterion for suggesting future experimental conditions. For instance, the neural network model has only one minimum after the initial batch of experiments (see Fig. 6(b) batch 1), and the constraint is distorted from its original shape—a circle. If decrease in information energy is used as the only criterion, one experiment will be suggested in the next batch. However, information entropy calls for more experiments that are necessary to mold a more complete performance surface as shown in Fig. 6(c) for batch 1. It is shown in Fig. 6(a)–(c) that this algorithm captures all the four Kuhn–Tucker points of this particular problem. From the comparison of Fig. 4(a) and Fig. 5, it is clear that the feed-forward ANN's have a great power to capture

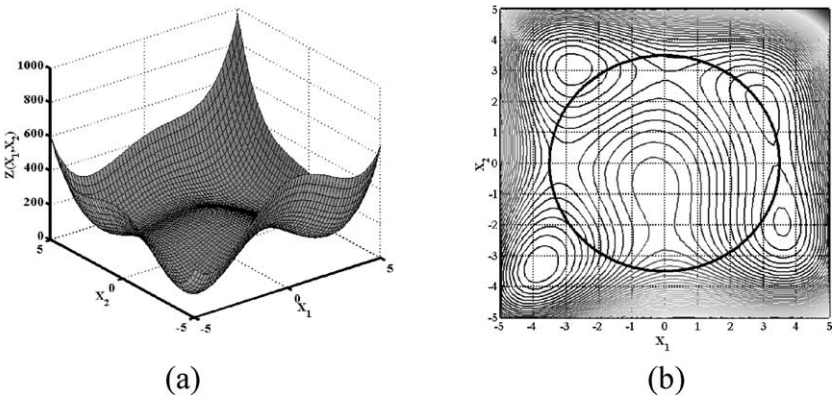

Fig. 4. (a) Mesh surface of the three-dimensional modified Himmelblau function; (b) the corresponding contour with a constraint circle.

the functional relationship of the modified Himmelblau function. The ANN model at the end of convergence is a close approximation to the Himmelblau function in the neighborhoods of the four minima and captures all the main features of the Himmelblau function.

4. Optimization of the coal-fired boiler

In the simulated combustion process as stated previously, there are four adjustable variables, namely, flow rates of coal and air in burner 1 and 2 ($F_{c,1}$, $F_{c,2}$, $V_{a,1}$, and $V_{a,2}$), or equivalently the burner's percentage load and air ratio ($L_{c,1}$, $L_{c,2}$, $\alpha_{a,1}$, and $\alpha_{a,2}$). Optimization here means that these four variables are tuned to achieve the best performance of the boiler under certain criterion and constraints. Two case studies have been carried out and the results are stated in Sections 4.1 and 4.2.

4.1. Case 1: thermal efficiency maximization and NO_x and CO minimization

In this case study, we try to detect the optimum operation conditions of the simulated combustion process

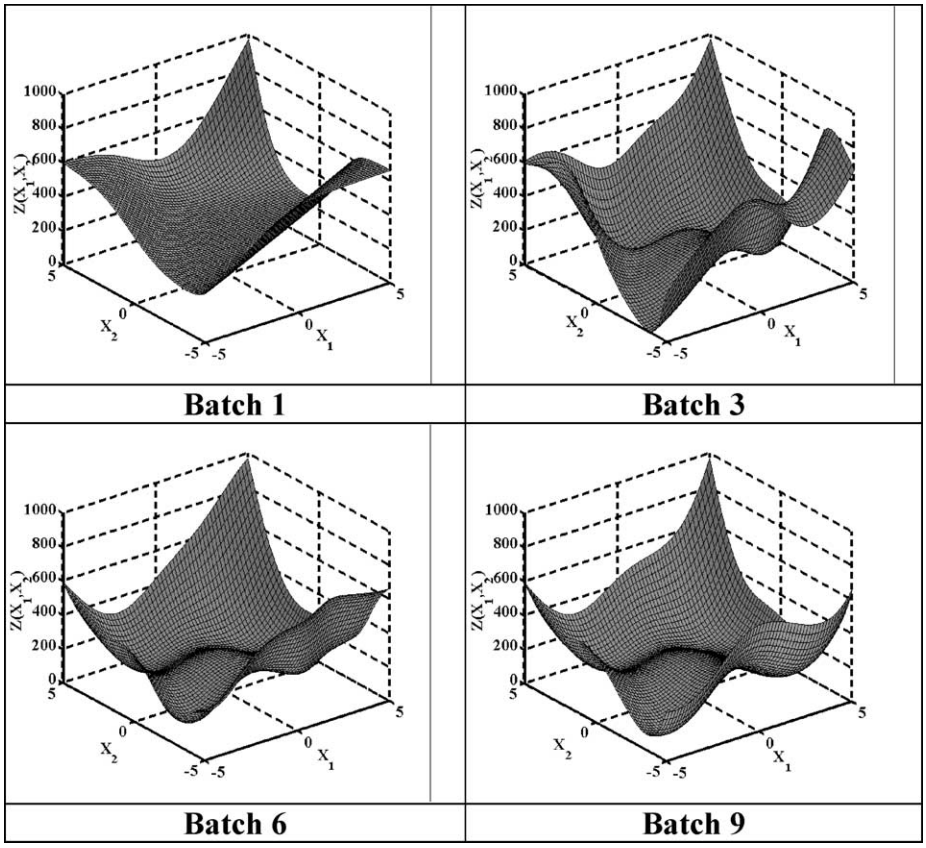

Fig. 5. Mesh surfaces of the output to the two inputs of the ANN model trained based on the accumulated data for different batch number.

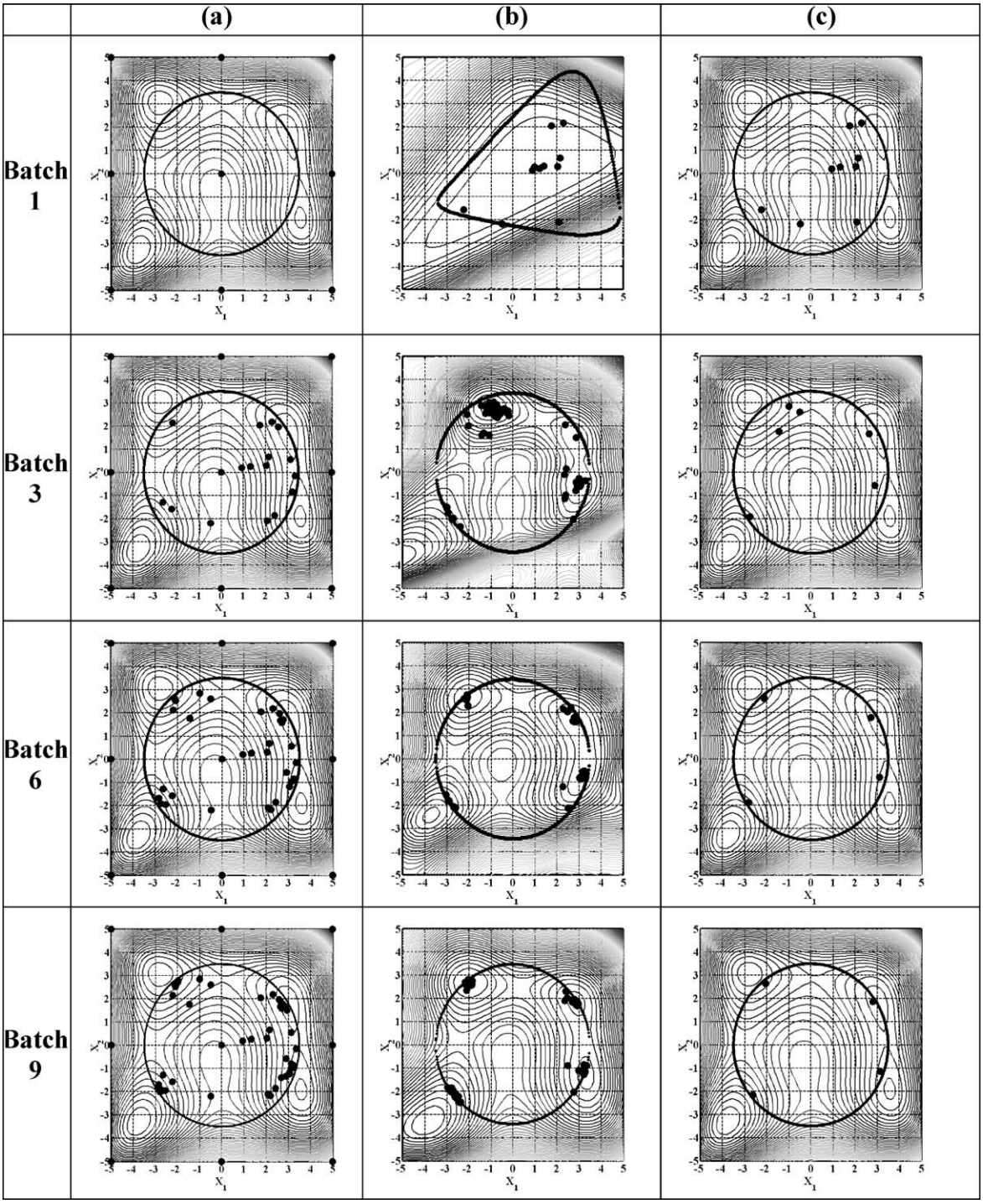

Fig. 6. Minimization of the modified Himmelblau function under a circle constraint. (a) Existing experimental points against the contour of Eq. (19) and the constraint inequality (20). (b) Corresponding model contour whose solid points represent the local minimum points found. (c) The next batch of the experimental points suggested by the proposed procedure, against the contour of Eq. (19).

as stated in Section 2, in the range of 80–100% the full load of the boiler. We pursue a set of operation conditions at which the emissions of NO_x and CO is low and the thermal efficiency is high. That is to say, our target is clean air and high thermal efficiency. As a demonstration, the objective function to be minimized is formulated as

$$f = \frac{y_{s,NO_x} - a_{NO_x}}{b_{NO_x} - a_{NO_x}} + \frac{y_{s,CO} - a_{CO}}{b_{CO} - a_{CO}} + \frac{b_\eta - \eta}{b_\eta - a_\eta} \tag{21}$$

where y_{s,NO_x} = concentration of NO_x in the smoke, volume fraction; $y_{s,CO}$ = concentration of CO in

the smoke, volume fraction; η = thermal efficiency as defined in Eq. (4); a = lower limit of a variable specified by its subscript; b = upper limit of a variable specified by its subscript.

In Eq. (21), a's and b's are scale factors to make all the three terms in the equation equally important. The values for them can be roughly estimated from theoretical analysis and previous experience.

To make the simulated combustion process more realistic, noise is introduced into the simulated combustion process by adding a random change in the coal's composition

$$z_H = z_{H,0} \times (1 + 0.035\alpha) \quad (22a)$$

$$z_C = z_{C,0} - 0.035\alpha z_{H,0} \quad (22b)$$

where, $z_{C,0}$ and $z_{H,0}$ are the average compositions of carbon and hydrogen in the coal, z_C and z_H are those used in experiments, and α is a random number in the range $[-1,1]$.

We can start the optimization procedure by forming an orthogonal array of operation conditions as suggested in Taguchi method [21], to cover the 80–100% load range in three operation levels of the four operating variables ($L_{c,1}$, $L_{c,2}$, $\alpha_{a,1}$, and $\alpha_{a,2}$). Table 2 lists the conditions and results of the eighteen 'experiments' in the initial batch. The best operating condition among the initial experimental data is No. 17 whose objective function $f = 53$, while the worst one is No. 3 whose f is 200.

Table 3 documents the evolving history of the optimization process. In Table 3, the conditions and results of the experiment which has the lowest objective function (f) are listed for each batch of experiments, and detailed intermediate results of the convergence process are listed elsewhere [14]. Three out of the last four experiments in Table 3 have close values of objective function f, which indicates the convergence of the optimization procedure if we remember the random fluctuation in the coal's compositions. The optimum operating conditions are thus located at;

Burner 1:	
Load percentage, $L_{c,1}$	89%
Air ratio, $\alpha_{a,1}$	1.011
Burner 2:	
Load percentage, $L_{c,2}$	89%
Air ratio, $\alpha_{a,2}$	1.11

At these conditions, the thermal efficiency is 72.44%, and concentrations of NO_x and CO are $y_{s,NO_x} = 173$ ppm, and $y_{s,CO} = 0.132$ ppm.

Table 2
Orthogonal experiments in the initial (1st) batch for Case 1

No.	Operation variables				Emissions[a]		η (%)	f
	$L_{c,1}$ (%)	$\alpha_{a,1}$	$L_{c,2}$ (%)	$\alpha_{a,2}$	NO_x (ppm)	CO (ppm)		
1	80	1.01	80	1.01	182	16.4	75.05	112
2	90	1.11	90	1.11	207	0.104	71.09	89
3	100	1.21	100	1.21	258	0.0775	66.74	200
4	80	1.11	80	1.21	204	0.0716	71.37	82
5	90	1.21	90	1.01	195	0.0951	71.18	74
6	100	1.01	100	1.11	189	0.158	70.67	73
7	80	1.21	90	1.11	244	0.0693	70.56	138
8	90	1.01	100	1.21	246	0.0998	70.20	144
9	100	1.11	80	1.01	186	0.101	72.14	52
10	80	1.11	100	1.11	252	0.0957	70.65	146
11	90	1.21	80	1.21	192	0.0738	69.21	94
12	100	1.01	90	1.01	186	15.3	72.87	136
13	80	1.01	90	1.21	197	0.0946	71.36	74
14	90	1.11	100	1.01	216	0.121	71.44	95
15	100	1.21	80	1.11	198	0.0646	69.73	94
16	80	1.21	100	1.01	189	0.106	71.17	67
17	90	1.01	80	1.11	199	0.156	73.33	53
18	100	1.11	90	1.21	206	0.0786	68.96	114

[a] Concentration at STP and 6 vol% O_2.

Table 3
Summary of the experiments in the nine batches for Case 1

No.[a]	Operation variables				Emissions[b]		η (%)	f
	$L_{c,1}$ (%)	$\alpha_{a,1}$	$L_{c,2}$ (%)	$\alpha_{a,2}$	NO_x (ppm)	CO (ppm)		
23/2	90	1.110	80	1.110	187	0.0828	71.67	58
25/3	92	1.110	84	1.088	196	0.0887	71.49	71
28/4	81	1.095	90	1.033	205	0.141	72.72	67
29/5	99	1.022	92	1.026	183	0.197	72.29	47
34/6	88	1.095	83	1.019	194	0.151	73.01	50
36/7	90	1.024	89	1.100	172	0.121	72.21	35
40/8	99	1.033	88	1.041	190	0.195	72.49	53
45/9	89	1.011	89	1.110	173	0.132	72.44	33
50/10	91	1.022	89	1.097	173	0.134	72.38	35

[a] Experiment series number/batch number.
[b] Concentration at STP and 6 vol% O_2.

4.2. Case 2: thermal efficiency maximization under NO_x and CO constraints

In real production and management, emissions of NO_x, CO and other pollutants are not necessarily minimized as long as regulation is not violated. Therefore, the boiler should be operated to maximize thermal efficiency under the constraints of $y_{s,NO_x} < 210$ ppm and $y_{s,CO} < 1$ ppm.

The operating variables are the same as in Case 1. It should be noted that in this case, two feed-forward ANNs are used to build the above two inequality constraints from existing experimental data. Table 4 lists the converging process of the proposed procedure. The last three experiments in Table 4 have the similar operating conditions, emission concentrations, and thermal

Table 4
Summary of the experiments in the five batches for Case 2

No.[a]	Operation variables				Emissions[b]		η (%)
	$L_{c,1}$ (%)	$\alpha_{a,1}$	$L_{c,2}$ (%)	$\alpha_{a,2}$	NO_x (ppm)	CO (ppm)	
20/2	100	1.012	100	1.048	200	0.341	71.81
24/3	85	1.130	81	1.019	177	0.0965	72.75
28/4	81	1.070	83	1.016	183	0.180	73.76
31/5	80	1.083	81	1.023	178	0.124	73.75
33/6	81	1.110	83	1.030	196	0.140	73.24

[a] Experiment series number/batch number.
[b] Concentration at STP and 6 vol% O_2.

efficiency. The optimum operation conditions are summarized as;

Burner 1:	
Load percentage, $L_{c,1}$	81%
Air ratio, $\alpha_{a,1}$	1.070
Burner 2:	
Load percentage, $L_{c,2}$	83%
Air ratio, $\alpha_{a,2}$	1.016

At the optimum operating conditions, the thermal efficiency is 73.76%, and concentrations of NO_x and CO are $y_{s,NO_x} = 183$ ppm, and $y_{s,CO} = 0.180$ ppm. Because the variation in the coal's compositions, the NO_x and CO emissions are not right in the edge of regulation limit.

The above optimal operating conditions mean that the loads of both burners are at the low end with medium excess O_2 for the first burner and very low excess O_2 for the second burner, which is quite reasonable to the physical and chemical fundamentals.

5. Discussion and conclusion

It has been known that delaying the mixing of the combustion air with the fuel (i.e. air staging) is an effective means to reduce production of NO_x in the combustion process [22]. Also we know that there exists a best air ratio at which thermal efficiency achieves the maximum for a given fuel and a given combustion equipment. From these basic principles of combustion, the optimum conditions for the above two cases are justified. The improvement of thermal efficiency from 72.4% in Case 1 to 73.8% in Case 2 shows the great significance in optimizing the operation of boilers.

The effectiveness of the proposed optimization procedure is clearly testified if we consider the complexity of the simulated system. In the first case, the minimum is located for a four-variable system within 10 batches by calling on the original model (or performing test) 50 times, and in the second case, the maximum is located for a four-variable system under two constraints within six batches by calling on the original model 33 times.

It should be noted that the results of optimization depend on various boundary conditions such as coal types, furnace configurations, seasonal atmospheric conditions, etc. The above case studies only serve as examples to demonstrate our optimization strategy.

Acknowledgements

The authors thank the financial support to this work by Industrial Technology Research Institute, Energy and Resource Laboratories, Taiwan, ROC, through the grant 90A0098J2.

References

[1] Faravelli T, Bua L, Frassoldati A, Antifora A, Tognotti L, Ranzi E. A new procedure for predicting NO_x emissions from furnaces. Comput Chem Engng 2001;25:613.
[2] Nordin A, Eriksson L, Ohman M. NO reduction in a fluidized bed combustor with primary measures and selective non-catalytic reduction. Fuel 1993;74(1):128.
[3] Henttonen J, Kojo IV, Kortela U. Optimizing control of NO_x and SO_2 emissions in the FBC process. J Inst Energy 1992;65:118.
[4] George EM. NO_x control through boiler optimization. Power Engng 1994;Feburary:29.
[5] Cybenko G. Approximation by superpositions of a sigmoidal function. Math Control Sign Syst 1989;2:303.
[6] Kalogirou SA. Applications of artificial neural networks in energy systems: a review. Energy Conversion Mgmt 1999;40:1073.
[7] Reifman J, Feldman EE, Wei TYC, Glickert RW. An intelligent emissions controller for fuel lean gas re-burn in coal-fired power plants. J Air Waste Mgmt Assoc 2000;50:204–51.
[8] Zhou H, Cen K, Mao J. Combining neural network and genetic algorithms to optimize low NO_x pulverized coal combustion. Fuel 2001;80:2163.
[9] Blasco JA, Fueyo N, Dopazo C, Ballester J. Modeling the temporal evolution of a reduced combustion chemical system with an artificial neural network. Combust Flame 1998;113:38.
[10] Yin C, Luo Z, Ni M, Cen K. Predicting coal ash fusion temperature with a back-propagation neural network model. Fuel 1998;77:1777.
[11] Zhu Q, Jones JM, Williams A, Thomas KM. The predictions of coal/char combustion rate using an artificial neural network approach. Fuel 1999;78:1755.
[12] Prasad G, Swidenbank E, Hogg BW. A local model networks based multivariable long-range predictive control strategy for thermal power plant. Automatica 1998;34(10):1185.
[13] Chen J, Wong DSH, Jang S-S, Yang S-L. Product and process development using artificial neural-network model and information analysis. AIChE J 1998;44(4):876.
[14] Chien CI. Operation optimization of the boiler of a cogeneration plant: higher thermal efficiency and lower NO_x emissions. Thesis of Master Degree, National Tsinghua University, Taiwan; 2001.
[15] Kalman BL, Kwasny SC. Why Tanh? Choosing a sigmoidal function. Int Joint Conf Neural Networks, Baltimore, MD 1992;.
[16] Hertz J, Krogh A, Palmer RG. Introduction to the theory of neural computation. New York: Addison-Wesley; 1991.
[17] Gorodkin J, Hansen LK, Krogh A, Savrer C, Winther O. A quantitative study pruning by optimal brain damage. Int J Neural Syst 1993;4:159.

[18] Allen DM. The relationship between variable selection and data augmentation and a method for prediction. Technometrics 1974; 16(1):125.
[19] Jang J-SR, Sun C-T, Mizutani E. Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. Englewood Cliffs, NJ: Prentice-Hall; 1997. p. 173, chapter 7.
[20] Bezdek JC, Hathaway RJ, Sabin MJ, Tucker WT. Convergence theory for fuzzy c-means: counter examples and repairs. IEEE Trans Syst, Man, Cybernet 1987;SCM-17(5):873.
[21] Taguchi G. Introduction to quality engineering. Asian Productivity Organization; 1986.
[22] Hesselmann GJ. Optimization of combustion by fuel testing in a NO_x reduction test facility. Fuel 1997;76(13):1269.