Monitoring of Drum-boiler Process Using Statistical Techniques

Seshu K. Damarla and Madhusree Kundu

Abstract—**Present study addresses the monitoring of drum boiler process. Methodologies; based on clustering time series data and moving window based pattern matching have been proposed for the detection of fault in the chosen process. Design databases were created for the process by simulating the developed process model. A modified k-means clustering algorithm using similarity measure as a convergence criterion has been adopted for discriminating among time series data pertaining to various operating conditions. The proposed distance and PCA based combined similarity along with the moving window approach were used to discriminate among the normal operating conditions as well as detection of faults for the processes taken up.**

Index Terms—**Drum boiler, k-means clustering, Moving window, PCA, Pattern matching.**

I. Introduction

Monitoring a chemical process is a challenging task because of their multivariate and highly correlated nature. The data based approaches; supervised learning; unsupervised learning and multivariate statistical techniques rather than the model based approaches are convenient for process monitoring. New methodologies; based on clustering time series data and moving window based pattern matching have been proposed for detection of normal as well as faulty conditions in the process. Data collection has become a mature technology over the years but the interpretation of process historical database has become an active area of research [1-3]. A modified k-means clustering algorithm using similarity measures as a convergence criterion has been used for clustering datasets pertaining to different operating conditions including faulty one. At a given time period of interest; for a multivariate time series data or template data, a similar pattern can be located in the historical database using the proposed pattern matching algorithm. Both the pattern matching and clustering time series data are useful for successful monitoring of the process including fault detection and its analysis.

Drum boiler is crucial benchmark process in view of modeling and control system design. This process was addressed by Gordon Pellegrinetti & Joseph Bentsman (1996), K. J. Astrom & R. D. Bell (2000), Wen Tan et al (2002), and K. Jawahar & N. Pappa (2005) [1-4] regarding modeling and control aspects. In the present work drum-boiler process has been monitored using similarity measures. The model has been derived using first principles and is characterized by few physical parameters. The parameters used in the model are those from a Swedish Power Plant. The values of all the parameters used were adapted from the work of K. J. Astrom & R. D. Bell. Sixteen (16) numbers of datasets belonging to four operating conditions including an abnormal operating condition were generated to evaluate the performance of the proposed techniques.

Manuscript received April 21, 2011; revised June 15, 2011.

Seshu K. Damarla, M.Tech student: E-mail: Seshu.chemical@gmail.com

Madhusree Kundu, Associate Professor, Department of Chemical Engineering, National Institute of Technology, Rourkela, India.

E-mail:mkundu@nitrkl.ac.in,madhushreek@yahoo.com,Fax:0661-2462 999

II. Similarity Factors

A. PCA Similarity

Principal component analysis is multivariable statistical technique to reduce the dimensionality of the large datasets by transforming set of original correlated variables into a new set of uncorrelated variables. These new uncorrelated variables capture the maximum variance in the dataset and are linear combinations of the original variables. PCA was successfully applied by Kourti & MacGregor (1996) and Martin & Morris (1996) to cluster multivariate time series data [5 & 6]. Krzanowski (1979) developed PCA similarity factor by choosing largest k principal components of each multivariate time series dataset that describe at least 95 % of variance in the each dataset [7]. These principal components are Eigen vectors of the covariance matrix. The PCA similarity factor between two datasets is defined by equation (1).

$$S_{PCA} = \frac{1}{k}\sum_{i=1}^{k}\sum_{j=1}^{k} Cos^2\theta_{ij} \qquad (1)$$

where k is the number of selected principal components in both datasets, θ_{ij} is the angle between the i^{th} principal component of X_1 and jth principal component of X_2. When first two principal components explain 95% of variance in the datasets, S_{PCA} may not capture the degree of similarity between two datasets because it weights all PCs equally. Obviously S_{PCA} has to modify to weight each PC by its explained variance. The modified S_{PCA}^{λ} is defined as

$$S_{PCA}^{\lambda} = \frac{\sum\sum(\lambda_i^{(1)}\lambda_j^{(2)})Cos^2\theta_{ij}}{\sum_{i=1}^{k}\sum_{j=1}^{k}\lambda_i^{(1)}\lambda_j^{(2)}} \qquad (2)$$

where $\lambda_i^{(1)}, \lambda_i^{(2)}$ are the eigen values of the first and second datasets respectively.

B. Distance Similarity

In addition to above similarity measure, distance similarity factor can be used to cluster multivariate time series data. Distance similarity factor compares two datasets that may have similar spatial orientation. The process variables pertaining to different operating conditions may have similar principal components.

The distance similarity factor is defined as

$$S_{dist} = 2\times\frac{1}{\sqrt{2\pi}}\int_{\phi}^{\infty} e^{\frac{-z^2}{2}} dz = 2\times[1-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\phi} e^{\frac{-z^2}{2}} dz \qquad (3)$$

where $\phi = \sqrt{(\bar{X}_2-\bar{X}_1)\sum_1^{*-1}(\bar{X}_2-\bar{X}_1)^T}$, $\bar{x}_2$ & $\bar{x}_1$ are sample means row vectors.

Σ_1 is the covariance matrix for dataset X_1 and $\sum_1^{*-1}$ is pseudo inverse of X_1 . Dataset X_1 is assumed to be a reference dataset. In equation (3), a one side Gaussian distribution is used because $\Phi \geq 0$. The error function can be calculated by using any software or standard error function tables. The integration in equation (3) normalizes S_{dist} between zero and one.

C. Combined Similarity Factor

The combined similarity factor (SF) combines S_{PCA}^{λ} and S_{dist} using weighted average of the two quantities and used for clustering of multivariate time series data.. The combined similarity is defined as

$$SF = \alpha_1 S_{PCA}^{\lambda} + \alpha_2 S_{dist} \qquad (4)$$

The selection of α_1 and α_2 is up to the user but ensure that the sum of them is equal to one. In this work we selected the values of α_1 & α_2 are 0.67 and 0.33.

D. K-means Clustering Using Similarity Factors

The time series data pertaining to various operating conditions were discriminated and classified using the following similarity based K-means algorithm.

Given: Q datasets, $\{x_1, x_2,...x_q....x_Q\}$ to be clustered into k clusters :

1. Let j^{th} dataset in the i^{th} cluster be defined as x_i^j . Computation of the aggregate dataset $X_i (i = 1,2,....k)$, for each of the k clusters as,

$$X_i = [(x_1^{(i)})^T ....(x_j^{(i)})^T ....(x_{Qi}^{(i)})]^T \qquad (5)$$

where Q_i is the number of datasets in the database. Note that $\sum_{i=1}^{k} Q_i = Q$.

2. Calculation of the dissimilarity between dataset x_q and each of the k aggregate datasets $X_i, i = 1,2....k$ as,

$$d_{i,q} = 1 - SF_{i,q} \qquad (6)$$

where $SF_{i,q}$ is similarity between q^{th} dataset and i^{th} cluster described by equation 4. Let the aggregate dataset X_i in equation 6 be the reference dataset. Dataset x_q is assigned to the cluster to which it is least dissimilar. Repetition of the aforesaid steps for Q numbers of datasets.

E. Clustering Performance Evaluation

Some key definitions were introduced by Singhal &. Seborg (2005) to evaluate the performance of the clusters obtained using similarity factors [8]. Assuming the number of operating conditions is N_{op} and the number of datasets for operating condition j in the database is N_{DBj} . Cluster purity is defined to characterize each cluster in terms of how many numbers of datasets for a particular operating condition present in the i^{th} cluster.

Cluster purity is defined as,

$$P_j \overset{\Delta}{=} (\frac{max_j N_{ij}}{N_{pi}}) \times 100\% \qquad (7)$$

where N_{ij} is the number of datasets of operating condition j in the i^{th} cluster and N_{pi} is the number of datasets in the i^{th} cluster.

Cluster efficiency measures the extent to which an operating condition is distributed in different clusters. This method is to penalize the large values of K when operating condition j distributed into different clusters. Clustering efficiency is defined as,

$$\eta = (\frac{max_i N_{i,j}}{N_{DBj}}) \times 100\% \qquad (8)$$

where N_{DBj} is the number of datasets for operating condition j in the database. Large number of datasets of operating condition present in a cluster can be considered as dominant operating condition.

III. MOVING WINDOW BASED PATTERN MATCHING

In this approach, the snapshot or template data with unknown start and end time of operating condition moves through historical data in a sample wise or dataset wise manner and the similarity between them is characterized by distance and PCA based combined similarity factor [9-11]. The historical data windows with the largest values of similarity factors are collected in a candidate pool and are called records to be analyzed by the process Engineer. For the present work, the historical data window moved one observation at a time, with each old observation is getting replaced by new one (sample wise manner). Pool accuracy, Pattern matching efficiency and overall effectiveness of pattern matching are important metrics that quantify the performance of the proposed pattern matching algorithm. The proposed pattern matching technique is consisting three steps which are follow as:

1. Specification of the snapshot (variables and time period)
2. Comparison between snapshot and periods of historical windows using moving window
3. Collection of historical windows with the largest values of similarity factors

N_P: The size of the candidate pool. N_P is the number of historical data windows that have been labeled ''similar'' to the snapshot data by a pattern matching technique. The data windows collected in the candidate pool are called records.

N_1 = number of records in the candidate pool that are exactly similar to the current snapshot data, i.e. having a similarity of 1.0/or number of correctly identified record.

N_2 = number of records in the candidate pool that are not correctly identified.

$N_P = N_1 + N_2$

N_{DB}: The total number of historical data windows that are actually similar to the current snapshot. In general, $N_{DB} \neq N_P$

Pool accuracy=$(N_1 / N_P) \times 100\%$

Pattern matching efficiency=$[1-((N_P - N_1) / N_{DB})] \times 100\%$

Pattern matching algorithm efficiency =$(N_P / N_{DB}) \times 100\%$

A large value of Pool accuracy is important in case of detection of small number of specific previous situations from a small pool of records without evaluating incorrectly identified records. A large value of Pattern matching efficiency is required in case of detection of all of the specific previous situations from a large pool of records. The proposed method is completely data driven and unsupervised; no process models or training data are required. The user should specify only the relevant measured variables.

IV. Drum Boiler

A. The Brief Introduction of Boiler System

The utility boilers in the thermal/nuclear power plants are water tube drum boilers. This type of boiler usually comprises two separate systems. One of them is the steam–water system, which is also called the water side of the boiler. In this system preheated water from the economizer is fed into the steam drum, and then flows through the down-comers into the mud drum.

The diagram depicted in Fig. 1 shows the drum that holds water at saturation or near saturation condition and denser water flows through the down-comer into lower header by force of gravity. After being heated up further, it returns to the drum through the riser. Between lower and upper header, are stretch of tubes that constitute water walls and receive radiant heat from furnace. Water walls permit use of high temperature of furnaces and combustion rates. A part of water in these tubes and risers evaporates, with the result that the fluid in the riser is composed of a mixer of water and steam. The difference in density of water in the riser and down-comer provides the necessary motive force to set up circulation of water in the boiler system.

Boiler is a reservoir of energy. Amount of energy stored in each part is a complicated function of temperature and pressure. The model can now be developed in detail, expressing stored energy, input power and output power as function of control variables and sate variables. Global mass and energy balances capture much of the behavior of the system.

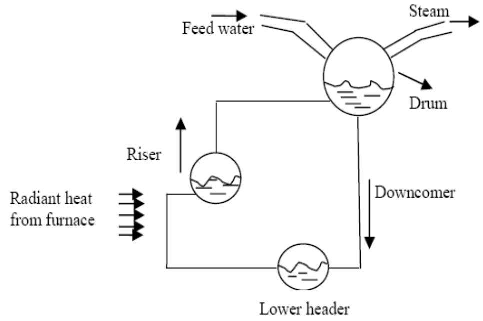

Fig. 1 Drum-boiler process

B. Modeling

Assumptions:

- The fundamental modeling simplification is that the two phases of the water inside the system are everywhere in saturated thermodynamic state.
- There is an instantaneous and uniform thermal equilibrium between water and metal everywhere.
- The energy stored in the steam and water is released or absorbed very rapidly when pressure changes. This is the key for understanding boiler dynamics. The rapid release of energy ensures that different parts of the boiler change their temperature in the same way.
- Steady state metal temperature is close to saturation temperature and the temperature differences are small dynamically.

Global mass and energy balance equations:

The inputs to the system are chosen to be

- Heat flow rate to the risers, Q
- Feedwater mass flow rate, q_f
- Steam flow rate, q_s

The outputs from the system are chosen to be:

- Drum level, L
- Drum pressure, P

A key feature of drum boiler is that there is an efficient energy and mass transfer between all parts that are in contact with steam. The mechanism responsible for heat transfer is boiling and condensation.

Global mass balance:

$$\frac{d[\rho_s V_{st} + \rho_w V_{wt}]}{dt} = q_f - q_s \tag{9}$$

Global Energy balance:

$$\frac{d[\rho_s V_{st} h_s + \rho_w V_{wt} h_w - P V_t + m_t c_p t_m]}{dt} = Q + q_f h_f - q_s h_s \tag{10}$$

Total Volume of Drum, Down-comer and Risers:

$$V_t = V_{st} + V_{wt} \tag{11}$$

where

ρ_s and ρ_w represent the densities of steam and water respectively,

h_s and h_w represent the enthalpies of steam and water per unit mass,

V_{st} and V_{wt} represent the total steam and water volume in the system,

V_t, is the total volume of the drum,

m_t, is the total metal mass,

t_m, is the metal temperature

A simple drum boiler model is obtained by combining equations (9), (10) and (11) with saturated steam tables. Mathematically the model is a differential algebraic system.

C. Proposed Second Order Boiler Model

To have a better insight into the key physical mechanism that affect the dynamic behavior of the system the state variable approach is considered. Drum pressure, P is chosen as a key state variable, since it is the most globally uniform variable in the system and is also easily measurable. The variables ρ_s, ρ_w, h_s, h_w are expressed as function of steam pressure using the steam table. The second state variable is chosen to be the total volume of water in the system i.e. V_{wt}.

Using equation (11), V_{st} is eliminated from equations (9) and (10).

The state equations then take the following form:

$$\left.\begin{aligned} e_{11}\frac{dV_{wt}}{dt} + e_{12}\frac{dP}{dt} &= q_f - q_s \\ e_{21}\frac{dV_{wt}}{dt} + e_{22}\frac{dP}{dt} &= Q + q_f h_f - q_s h_s \end{aligned}\right\} \quad (12)$$

where

$$e_{11} = \rho_w - \rho_s$$

$$e_{12} = V_{st}\frac{\partial \rho_s}{\partial P} + V_{wt}\frac{\partial \rho_w}{\partial P}$$

$$e_{21} = \rho_w h_w - \rho_s h_s$$

$$e_{22} = V_{st}\left(h_s\frac{\partial \rho_s}{\partial P} + \rho_s\frac{\partial h_s}{\partial P}\right) + V_{wt}\left(h_w\frac{\partial \rho_w}{\partial P} + \rho_w\frac{\partial h_w}{\partial P}\right) - V_t + m_t c_p\frac{\partial t_s}{\partial P} \quad (13)$$

Equations (12) and (13) constitute a state model of the second order boiler system. This simplistic model captures the gross behavior of the boiler quite well. In particular, it describes the response of drum pressure to changes in input power, feed-water and steam flow rates reasonably well. But the model has a serious deficiency. It doesn't capture the behavior of the drum water level, as the distribution of steam and water are not taken into account. The drum level control is difficult due to shrink and swells effects. The drum level may be defined as the level at which water stands in the boiler. The steam level is the space above the water level.

D. Distribution of Steam in Risers and Drum

The behavior of drum-level can best described by taking into account the distribution of water and steam in the system. The distribution of water and steam is considered separately for the riser section and the drum.

Distribution of steam and water in risers

The steam-water distribution varies along the risers. In the riser section water exists in two phases namely the liquid-phase i.e. water and the vapor-phase i.e. steam. The mass fraction or dryness fraction of a liquid-vapor mixture must be defined prior to further discussion. In a liquid-vapor mixture, x is known as the quality.

$$x = \frac{m_v}{m_v + m_l}$$

where , m_v and m_l are the masses of vapor and liquid respectively in the mixture. The value of x varies between 0 and 1. In order to determine the average density of steam-water mixture in the riser, it is necessary to define the void fraction. The void fraction α of a two phase mixture is a volumetric quantity and is given as: α =(volume of vapor)/(volume of vapor + volume of liquid).

α and x are related by:

$$\alpha = \frac{1}{1+\left(\frac{1-x}{x}\right)\varphi} \text{ or } x = \frac{1}{1+\left(\frac{1-x}{x}\right)\frac{1}{\varphi}} \quad (14)$$

where, $\varphi = \frac{\vartheta_f}{\vartheta_g}s$. v_f and v_g are the specific volumes of saturated liquid and vapor respectively. s is the slip ratio of two-phase mixture. The two phases of the mixture do not travel at the same speed. Instead there is a slip between them, which causes the vapor to move faster than liquid. s is a dimensionless number, greater than 1. It is defined as the ratio of average vapor velocity to the average liquid velocity, at any cross-section of the riser. The slip ratio is neglected in the present work, as it doesn't have a major influence on the fit to experiment data.

The behavior of two-phase flow is complicated and is typically modeled by partial differential equations. Keeping a finite dimensional model it is assumed that that shape of the distribution is known. The assumed shape is based on solving the partial differential equations in the steady state. There exists a linear distribution of steam-water mass ratio along the risers. The ratio varies in the following form:

$$\alpha_m(\xi) = \alpha_r\xi, \; 0 \le \xi \le 1 \quad (15)$$

where ξ is a normalized length coordinate along the risers and α_r is the mass ratio at the riser outlet. The volume and mass fractions of steam are related through

$$\alpha_v = f(\alpha_m) \quad (16)$$

where, $$f(\alpha_m) = \frac{\rho_w\alpha_m}{\rho_s + (\rho_w - \rho_s)\alpha_m} \quad (17)$$

For modeling the drum-level the total amount of steam in the drum is to be obtained. The governing equation is the average steam volume fraction in the risers, which is obtained by integrating the equation (17) over the limits 0 to 1 can be given as:

$$\bar{\alpha}_v = \int_0^1 \alpha_v(\xi)\,d\xi = \frac{1}{\alpha_r}\int_0^1 f(\alpha_r\xi)\,d(\alpha_r\xi) = \frac{1}{\alpha_r}\int_0^{\alpha_r} f(\xi)\,d\xi = \frac{\rho_w}{(\rho_w-\rho_s)}\left(1 - \frac{\rho_s}{(\rho_w-\rho_s)\alpha_r}\ln\left(1 + \frac{\rho_w-\rho_s}{\rho_s}\alpha_r\right)\right) \quad (18)$$

The partial derivatives of $\bar{\alpha}_v$ with respect to drum pressure and steam mass fraction are obtained as:

$$\left.\begin{aligned} \frac{\partial\bar{\alpha}_v}{\partial P} &= \frac{1}{(\rho_w-\rho_s)^2}\left(\rho_w\frac{\partial\rho_s}{\partial P} - \rho_s\frac{\partial\rho_w}{\partial P}\right)\left(1 + \frac{\rho_w}{\rho_s}\frac{1}{1+\eta} - \frac{\rho_s+\rho_w}{\eta\rho_s}\ln(1+\eta)\right) \\ \frac{\partial\bar{\alpha}_v}{\partial\alpha_r} &= \frac{\rho_w}{\rho_s\eta}\left(\frac{1}{\eta}\ln(1+\eta) - \frac{1}{1+\eta}\right) \\ \text{where } \eta &= \frac{\alpha_r(\rho_w-\rho_s)}{\rho_s} \end{aligned}\right\} \quad (19)$$

The transfer of mass and energy between steam and water by condensation and evaporation is a key element in the modeling. When modeling the phases separately the transfer must be accounted for explicitly, hence the joint balance equations for water and steam are written for the riser section.

Lumped parameter model:

The global mass balance for the riser section is:

$$\frac{d[\rho_s\bar{\alpha}_v V_r + \rho_w(1-\bar{\alpha}_v)V_r]}{dt} = q_{dc} - q_r \quad (20)$$

where, q_r, is the total mass flow rate out of the risers., q_{dc}, is the total mass flow rate into the risers.

The global energy balance for the riser section is:

$$\frac{d[\rho_s h_s \bar{\alpha}_v V_r + \rho_w h_w (1-\bar{\alpha}_v) V_r - P V_r + m_r c_p t_s]}{dt} = Q + q_{dc} h_w - (\alpha_r h_c + h_w) q_r \quad (21)$$

Circulation flow:

In the natural circulation boiler the flow rate is driven by density gradients in the risers and down-comers. The flow through the down-comer (q_{dc}) can be obtained from a momentum balance.

The equation consists of three terms namely the internal term, driving force that in this case is the density difference and frictional force in the flow through pipes.

(a) Inertia force: $(L_r + L_{dc})\frac{dq_{dc}}{dt}$

(b) Driving force: $(\rho_w - \rho_s)\bar{\alpha}_v V_r g$

(c) Friction force: $\frac{k q_{dc}^2}{2\rho_w A_{dc}}$

Combining above three terms momentum balance equation is written as following:

$$(L_r + L_{dc})\frac{dq_{dc}}{dt} = (\rho_w - \rho_s)\bar{\alpha}_v V_r g - \frac{k q_{dc}^2}{2\rho_w A_{dc}} \quad (22)$$

Equation (29) is a first order system that has a time constant as given below:

$$\tau = \frac{(L_r + L_{dc}) A_{dc} \rho_w}{k q_{dc}} \quad (23)$$

The steady state relation for the system is given as:

$$\frac{1}{2} k q_{dc}^2 = \rho_w A_{dc} (\rho_w - \rho_s) \bar{\alpha}_v V_r g \quad (24)$$

Distribution of steam in the drum

The physical phenomenon in the drum is complicated. Steam enters from many riser tubes: feed water enters through a complex arrangement, water leaves through the down-comer tubes and steam through the steam valve. The geometry and flow patterns are complex. Basic mechanisms that occur in the drum are separation of water and steam and condensation.

The mass balance for the steam under the liquid level is given as:

$$\frac{d(\rho_s V_{sd})}{dt} = \alpha_r q_r - q_{sd} - q_{cd} \quad (25)$$

where, q_{cd} is the condensation flow, which is given by

$$q_{cd} = \frac{h_w - h_f}{h_c} q_f + \frac{1}{h_c}\left(\rho_s V_{sd} \frac{dh_s}{dt} + \rho_w V_{wd} \frac{dh_w}{dt} - (V_{sd} + V_{wd}) \frac{dP}{dt} + m_d c_p \frac{dt_s}{dt}\right) \quad (26)$$

The flow q_{sd} is driven by density difference of water and steam, and the momentum of the flow entering the drum. The expression for q_{sd} is an empirical model and is a good fit to the experimental data and is given as:

$$q_{sd} = \frac{\rho_s}{T_d}(V_{sd} - V_{sd}^0) + \alpha_r q_{dc} + \alpha_r \beta (q_{dc} - q_r) \quad (27)$$

where, V_{sd}^0, is the volume of steam in the drum in hypothetical situation when there is no condensation of steam in the drum and T_d is the residence time of steam in the drum.

Drum level:

Having accounted for distribution of steam below drum-level, now the drum-level is modeled. The drum level is composed of two terms,

- The total amount of water in the drum,
- The displacement due to changes of the steam-water ratio in the risers.

Derivation of the drum-level l measured from its normal operating level is given by:

$$l = \frac{V_{sd} + V_{wd}}{A_d} = l_w + l_s \quad (28)$$

where $l_w = \frac{V_{wd}}{A_d}$ and $l_s = \frac{V_{sd}}{A_d}$

$$V_{wd} = V_{wt} - V_{dc} - (1 - \bar{\alpha}_v) V_r \quad (29)$$

where, V_{wd}, is the volume of water in the drum, l_w, is the level variation caused by changes of amount of water in the drum, l_s, is the level variation caused by the steam in the drum, A_d, is the wet surface of the drum at the operating level.

E. The Model

Model is given by the following equations:

$$\left.\begin{aligned}
&\frac{d[\rho_s V_{st} + \rho_w V_{wt}]}{dt} = q_f - q_s \\
&\frac{d[\rho_s V_{st} h_s + \rho_w V_{wt} h_w - P V_t + m_t c_p t_m]}{dt} = Q + q_f h_f - q_s h_s \\
&\frac{d[\rho_s \bar{\alpha}_v V_r + \rho_w (1 - \bar{\alpha}_v) V_r]}{dt} = q_{dc} - q_r \\
&\frac{d[\rho_s h_s \bar{\alpha}_v V_r + \rho_w h_w (1 - \bar{\alpha}_v) V_r - P V_r + m_r c_p t_s]}{dt} = Q + q_{dc} h_w - (\alpha_r h_c + h_w) q_r \\
&\frac{d(\rho_s V_{sd})}{dt} = \alpha_r q_r - q_{sd} - q_{cd} \\
&q_{cd} = \frac{h_w - h_f}{h_c} q_f + \frac{1}{h_c}\left(\rho_s V_{sd}\frac{dh_s}{dt} + \rho_w V_{wd}\frac{dh_w}{dt} - (V_{sd} + V_{wd})\frac{dP}{dt} + m_d c_p \frac{dt_s}{dt}\right) \\
&l = \frac{V_{sd} + V_{wd}}{A_d} \\
&V_t = V_{st} + V_{wt}
\end{aligned}\right\} \quad (30)$$

The model as can be seen is a differential algebraic system. Since most available simulation software requires state equations, the state model is also derived.

F. State Variable Model

Prior to generation of database, linear model is required to design controllers. This was accomplished by the use of state space methodology. The selection of state variables is done in many different ways. A convenient way is to choose those variables as states, which have a good physical interpretation that describe storage of mass, energy and momentum. The variables used in this procedure are as follows:

1) State variables:

a. Drum pressure P

b. Total water volume of the system V_{wt}

c. Steam mass fraction α_r

d. Volume of steam in the drum V_{sd}

2) Manipulated inputs:

a. Heat flow rate to the risers Q

b. Feed water flow rate to the drum q_f

c. Steam flow rate from the drum q_s

3) Measured outputs:

a. Total water volume of the system V_{wt}

b. Drum pressure P

Direct synthesis controller design procedure has used for closed loop control system. Both the open loop and closed loop simulations were performed in order to generate the database. The drum-boiler process has been simulated for 1000 seconds with a sampling time of 2 seconds. Four distinct operating conditions were created by giving impulse, step and sinusoidal changes in the manipulated inputs in open loop as well as closed loop.

G. *Derivation of state equations*

The pressure and water dynamics are obtained from the global mass and energy balances equations (9) and (10). Combining these equations the state variable form is obtained as given by the set of equations (11). The riser dynamics is given by the mass and energy balance equations (20) and (21) are further simplified by eliminating 'q_r' and multiplying equation (20) by '$(h_w + \alpha_r h_c)$' and adding it to equation (21).

The resulting expression is given as:

$$h_c(1-\alpha_r)\frac{d(\rho_s\bar{\alpha}_v V_r)}{dt} + \rho_w(1-\bar{\alpha}_v)V_r\frac{dh_w}{dt} - \alpha_r h_c\frac{d[\rho_w(1-\bar{\alpha}_v)V_r]}{dt} + \rho_s\bar{\alpha}_v V_r\frac{dh_s}{dt} - V_r\frac{dP}{dt} + m_r c_p\frac{dt_s}{dt} = Q - \alpha_r q_{dc} h_c \quad (31)$$

If the state variables 'P' and 'α_r' are known, the riser flow rate 'q_r' can be computed from equation (27). This can be given as:

$$q_r = q_{dc} - V_r\frac{\partial[(1-\bar{\alpha}_v)\rho_w + \rho_s\bar{\alpha}_v]}{\partial P}\frac{dP}{dt} + V_r(\rho_w-\rho_s)\frac{\partial\bar{\alpha}_v}{\partial\alpha_r}\frac{d\alpha_r}{dt} \quad (32)$$

The drum dynamics can be captured by the mass balance equation (25). The expression for 'q_r', 'q_{sd}' and 'q_{cd}' are substituted in equation (25). The resulting simplified expression is given as:

$$\rho_s\frac{dV_{sd}}{dt} + V_{sd}\frac{d\rho_s}{dt} + \frac{1}{h_c}\left[\rho_s V_{sd}\frac{dh_s}{dt} + \rho_w V_{wd}\frac{dh_w}{dt} - (V_{sd}+V_{wd})\frac{dP}{dt} + m_d c_p\frac{dt_s}{dt}\right] + \alpha_r(1+\beta)V_r\frac{d[(1-\bar{\alpha}_v)\rho_w + \rho_s\bar{\alpha}_v]}{dt} = \frac{\rho_s}{T_d}\left(V_{sd}{}^0 - V_{sd}\right) + \frac{h_f - h_w}{h_c}q_f \quad (33)$$

The state equations are written as:

$$\left.\begin{aligned}
&e_{11}\frac{dV_{wt}}{dt} + e_{12}\frac{dP}{dt} = q_f - q_s\\
&e_{21}\frac{dV_{wt}}{dt} + e_{22}\frac{dP}{dt} = Q + q_f h_f - q_s h_s\\
&e_{32}\frac{dP}{dt} + e_{33}\frac{d\alpha_r}{dt} = Q - \alpha_r h_c q_{dc}\\
&e_{42}\frac{dP}{dt} + e_{43}\frac{d\alpha_r}{dt} + e_{44}\frac{dV_{sd}}{dt} =\\
&\qquad\frac{\rho_s}{T_d}\left(V_{sd}{}^0 - V_{sd}\right) + \frac{(h_f - h_w)}{h_c}q_f
\end{aligned}\right\} \quad (34)$$

where

$$e_{11} = (\rho_w - \rho_s)$$

$$e_{12} = V_{st}\frac{\partial\rho_s}{\partial P} + V_{wt}\frac{\partial\rho_w}{\partial P}$$

$$e_{21} = \rho_w h_w - \rho_s h_s$$

$$e_{22} = V_{st}\left(h_s\frac{\partial\rho_s}{\partial P} + \rho_s\frac{\partial h_s}{\partial P}\right) + V_{wt}\left(h_w\frac{\partial\rho_w}{\partial P} + \rho_w\frac{\partial h_w}{\partial P}\right) - V_t + m_t c_p\frac{\partial t_s}{\partial P}$$

$$e_{32} = \left(\rho_w\frac{\partial h_w}{\partial P} - \alpha_r h_c\frac{\partial\rho_w}{\partial P}\right)(1-\alpha_r)V_r + \left((1-\alpha_r)h_c\frac{\partial\rho_s}{\partial P} + \rho_s\frac{\partial h_s}{\partial P}\right)\bar{\alpha}_v V_r + (\rho_s + (\rho_w-\rho_s)\alpha_r)h_c V_r\frac{\partial\bar{\alpha}_v}{\partial P} - V_r + m_r c_p\frac{\partial t_s}{\partial P}$$

$$e_{33} = ((1-\alpha_r)\rho_s + \alpha_r\rho_w)h_c V_r\frac{\partial\bar{\alpha}_v}{\partial\alpha_r}$$

$$e_{42} = V_{sd}\frac{\partial\rho_s}{\partial P} + \frac{1}{h_c}\left(\rho_s V_{sd}\frac{\partial h_s}{\partial P} + \rho_w V_{wd}\frac{\partial h_w}{\partial P} - V_{sd} - V_{wd} + m_d c_p\frac{\partial t_s}{\partial P}\right) + \alpha_r(1+\beta)V_r\left(\bar{\alpha}_v\frac{\partial\rho_s}{\partial P} + (1-\bar{\alpha}_v)\frac{\partial\rho_w}{\partial P} + (\rho_w-\rho_s)\frac{\partial\bar{\alpha}_v}{\partial P}\right)$$

$$e_{43} = \alpha_r(1+\beta)(\rho_s-\rho_w)V_r\frac{\partial\bar{\alpha}_v}{\partial\alpha_r}$$

$$e_{44} = \rho_s$$

It is noted that the state space model obtained has an interesting lower triangular structure where state variables can be grouped as: $\left(\left((V_{wt}, P), \alpha_r\right), V_{sd}\right)$. The variables inside each parenthesis can be computed independently. Model is thus a nest of second, third and fourth order model. The second order model describes drum pressure and total water volume in the system. The third order model captures the steam dynamics in the risers and the fourth order model also describes the accumulation of steam below the water surface in the drum dynamics.

H. *Equilibrium Values*

The steady state solution of the state model of equation (34) is given by:

$$\left.\begin{aligned}
&q_f = q_s\\
&Q = q_s h_s - q_f h_f\\
&Q = q_{dc}\alpha_r h_c\\
&V_{sd} = V_{sd}^0 - \frac{T_d(h_w - h_f)}{\rho_s h_s}q_f
\end{aligned}\right\} \quad (35)$$

where, q_{dc} is given by equation (31) i.e.

$$q_{dc} = \sqrt{\frac{2\rho_w A_{dc}(\rho_w-\rho_s)g\bar{\alpha}_v V_r}{K}}$$

A convenient way to find the initial values is to first specify steam flow rate q_s and steam pressure P. the feed water flow rate q_f and input power Q are given by first two equations of equation (35) and the steam volume in the drum is given by the last equation of (35). The steam quality α_r is obtained by solving the nonlinear equations:

$$\left.\begin{aligned}
&Q = \alpha_r h_c\sqrt{\frac{2\rho_w A_{dc}(\rho_w-\rho_s)g\bar{\alpha}_v V_r}{K}}\\
&\bar{\alpha}_v = \frac{\rho_w}{\rho_w-\rho_s}\left(1 - \frac{\rho_s}{(\rho_w-\rho_s)\alpha_r}\ln\left(1 + \frac{\rho_w-\rho_s}{\rho_s}\alpha_r\right)\right)
\end{aligned}\right\} \quad (36)$$

The equilibrium values of the state variables:

1.Drum pressure P=8.5 MPa

2.Total water volume of the system V_{wt}=57.5 m^3

3.Steam mass fraction α_r=0.051

4.Volume of steam in the drum V_{sd}=4.8 m^3

The equilibrium values of the input variables are assumed as:

1. Heat flow rate to the risers Q=80.40437506e6 MW
2. Feed water flow rate to the drum q_f=32.00147798 kg/sec
3. Steam flow rate from the drum q_s=32.00147798 kg/sec

I. Parameter

An interesting feature of the model is that it requires only a few parameters to characterize the system. The parameter values that are considered are those from a Swedish plant. The plant is an 80 MW unit. The model is characterized by the parameters given in Table 1.

V. Results & Discussions

Four numbers of different operating conditions for Drum-boiler process are listed in Table 2. Sixteen numbers of datasets were generated by perturbing the process with pseudo random binary signal (PRBS) generator. Signal to noise ratio were maintained as 10.0. Negative step change in feed water flow rate induced continuous decrease in total water volume and continuous increase in drum pressure that represents abnormal scenario. Fig.2 shows the response of abnormal process. The combined similarity measures are capable of detecting faulty operating condition as cluster 3 and datasets pertaining to various other operating conditions were also identified correctly. The performance of clustering is presented in Table 3. The proposed moving window based sample wise pattern matching technique perfectly identified all the snapshot data including normal and faulty data which find their similarity with the same process data present in the historical database as shown in Table 4.

VI. Conclusions

Sample wise moving window based pattern matching algorithm emancipated encouraging fault detection for drum-boiler process. Clustering time series data was the approach used in the present work for discrimination among various operating conditions including the detection of faulty situations with 100 % accuracy. Efficient modeling and simulation of the drum-boiler process were the key factors behind the generation of design databases required for the successful implementation of the proposed monitoring techniques. In this regard the detail Drum-boiler process modeling including state space model for the process was developed and presented.

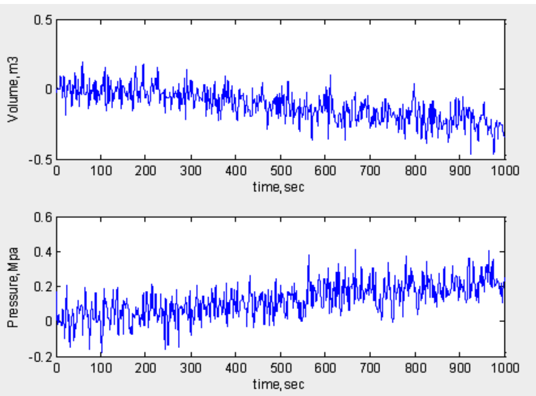

Fig. 2 Response of faulty process

TABLE 1 Values of The Drum-Boiler Model Parameters

S. NO	Parameter name	Notation	Value with units
1	Riser volume	V_r	37 m^3
2	Downcomer volume	V_{dc}	11 m^3
3	Total volume	V_t	88 m^3
4	Drum area at normal operating level	A_d	20 m^2
5	Downcomer flow area	A_{dc}	0.355 m^2
6	Riser metal mass	m_r	160e3 kg
7	Total metal mass	m_t	300e3 kg
8	Drum metal mass	m_d	100e3 kg
9	Friction coefficient in downcomer-riser loop	k	25
10	Empirical q_{sd} coefficient	β	0.3
11	Residence time of steam in drum	T_d	12 sec
12	Bubble volume coefficient	V_{sd}^0	10 m^3
13	Acceleration due to gravity	g	9.81 m/sec^2
14	Drum volume	V_d	40 m^3

TABLE 2: Various Operating Conditions for Drum Boiler Process

Operating condition	Parameter range	No. of datasets
1. Impulse change in q_f and q_s (open loop)	$4 \le q_f \ge 6$ $2 \le q_s \ge 3$	2
2. Simultaneous step changes in three manipulated inputs	$4 \le q_f \ge 10$ $2 \le q_s \ge 5$ $15 \le Q \ge 25$	7
3. Negative change in feed water flow (fault)	$4 \le q_f \ge 6$	4
4. Simultaneous sinusoidal change in three inputs	High, medium and low frequencies	3

TABLE 3: Combined Simil Arity Factor Based Clustering of Different Operating Conditions

Cluster No.	NP	P	Op. Cond. 1	Op. Cond. 2	Op. Cond. 3	Op. Cond. 4
1	2	100	1	0	0	0
2	5	100	0	2	0	0
3	4	100	0	0	3	0
4	3	100	0	0	0	4

TABLE 4: MOVING WINDOW BASED PATTERN MATCHING PERFORMANCE

Snapshot Op. Cond.	Np	N1	N2	P, Pattern matching accuracy	ζ, Pattern matching efficiency	Efficiency Of Algorithm
Op. Cond. 1	1	1	0	100	100	100
Op. Cond. 2	1	1	0	100	100	100
Op. Cond. 3	1	1	0	100	100	100
Op. Cond. 4	1	1	0	100	100	100

ACKNOWLEDGMENT

Authors would like to thank to Prof. Palash Kundu, EE Dept., Jadavpur University, Kolkata, India for providing motivation for this work.

REFERENCES

[1] G. Pellegrinetti, J. Bentsman, "Nonlinear control oriented boiler modeling-A Benchmark Problem for Controller Design," *ieee transactions on control systems technol*ogy, vol. 4, no. **1,** January 1996
[2] K. J. Astrom, R. D. Bell, "Drum-boiler Dynamics," Automatica 36 (2000) pp. 363-378.
[3] W. tan, H. J. Marquez, T. Chen, "Multivariable Robust Controller Design for A Boiler System," *IEEE Transactions on control systems technology*, vol. 10, no. 5, September 2002.
[4] K. Jawahar, N. Pappa, "Selftuning Nonlinear Controller,**"** proceedings of the 6th WSEAS Int. Conf. on Fuzzy Systems, Lisbon, Portugal, June 16-18, 2005 pp. 118-126.
[5] T. Kourti, J. F. MacGregor, "Multivariate SPC methods for process and product monitoring," *J. Quality Tech.* 1996, 28: pp. 409–428.
[6] E. B. Martin, A. J. Morris, "An Overview of Multivariate Statistical Process Control in Continuous and Batch Performance Monitoring," *Trans. Inst. Meas. and Control*. 1996, 18: pp. 51–60.
[7] WJ. Krzanowski, "Between-groups Comparison of Principal Components,*" J. Amer. Stat. Association*. 1979, 74: pp. 703–707.
[8] A. Singhal, D. E. Seborg, "Clustering Multivariate Time Series Data," *J. Chemometrics*. 2005, 19: pp. 427-438.
[9] M. C. Johannesmeyer, A. Singhal, D. E. Seborg, "Pattern Matching in Historical Data,*" AIChE J. September* 2002, 48: pp. 2022-2038.
[10] A. Singhal, D. E. Seborg, "Pattern matching in multivariate time series databases using a moving window approach," *Ind. Eng. Chem. Res.* 2002, 41: pp. 3822-3838.
[11] A. Singhal, D. E. Seborg, "Matching Patterns from Historical Data Using PCA and Distance Similarity Factors," Proceedings of the 2001 American Control Conference; *IEEE: Piscataway,* NJ. 2001, pp. 1759-1764.

Madhusree Kundu was born in Kolkata, 21 January 1965. She did her B.S.C. in Chemistry from university of Calcutta, India. Dr. Kundu got her B.Tech and M.Tech. in Chemical Engineering from the same university in the years 1990 & 1992, respectively. She completed her Doctoral study from Indian Institute of Technology, Kaharagpur, India and was awarded Ph.D. in the year 2005.

She served as a Process Engineer in the Simon carves India Ltd. during 1993-1998. She served as a lecturer and then Assistant Professor in The Birla Institute of Technology & Science, Pilani, India during 2004-2006. Presently, she is the Associate Professor, Department of Chemical Engineering, NIT, Rourkela, Orissa, India. She has published a couple of research articles and book chapters to her credit. Fluid Phase Equilibria, Advanced Process Control, Fault Detection and Diagnosis, & Process Monitoring are the current research interests of her.

Dr. Kundu is the life member of Indian Institute of Chemical Engineers (IIChE).

Seshu K.Damarla, M.Tech Student,Department of Chemical Engineering, NIT, Rourkela, Orissa, India. e-mail: seshu.chemical@gmail.com. He got his B.Tech Degree from Bapatla Engineering College, Andhra Pradesh, India.