Monitoring Flames in an Industrial Boiler Using Multivariate Image Analysis

Honglu Yu and John F. MacGregor
McMaster Advanced Control Consortium, Dept. of Chemical Engineering, McMaster University, Hamilton, Ontario L8S 4L7, Canada

DOI 10.1002/aic.10164
Published online in Wiley InterScience (www.interscience.wiley.com).

An on-line digital imaging system is developed for monitoring flames in an industrial boiler system. The information extracted from the RGB flame images is used to predict the performance of the boiler system. A practical method based on multivariate image analysis techniques and partial least squares is developed to efficiently extract information from the rapidly time varying flame images, and to predict boiler performance, NO_x and SO_2 concentration in the off-gas, and the energy content of an incoming waste fuel stream. The approach is very general and can be applied to a wide range of combustion processes. © 2004 American Institute of Chemical Engineers *AIChE J,* 50: 1474–1483, 2004
Keywords: flames, boilers, image analysis, performance monitoring, principa-component analysis, partial least squares

Introduction

Combustion plays an important role in many industrial processes. The efficiency of the combustion usually has a great influence on the economics of the process and on its environmental impact.

In combustion processes fuel and oxidizer (typically air) are mixed and burned. Generally, two categories can be identified based on whether the fuel and oxidizer are mixed first and then burned (premixed) or whether combustion and mixing occurs simultaneously (nonpremixed). Each of these categories is further subdivided based on whether the fluid flow is laminar or turbulent. Among the four classes, turbulent nonpremixed combustion processes are of interest in many industrial applications. They appear in jet engines, diesel engines, steam boilers, and furnaces, for example. The use of the nonpremixed combustion is widespread because it is safer to handle than premixed combustions. However, nonpremixed combustions include more complex chemistry and are harder to control. Unless very sophisticated mixing techniques are used, nonpremixed flames show a yellow luminescence, caused by glowing soot particles. The colors of the flames indicate the combustion region and the temperature of the field. This latter feature allows for the possibility of using color flame images to monitor the combustion process.

Visualization methods have been used to study combustion flames in laboratories. Shimoda et al. (1990) reported a combustion diagnostic system where the radiation energy and temperature profiles of flames were quantitatively identified from the flame images and the concentrations of unburned carbon and NO_x in the exhaust gas were estimated in a coal-fired boiler. Yamaguchi et al. (1997) developed a fiber-optical imaging sensor to detect the air/fuel ratio in a premixed-type gas-fired model combustor by monitoring the radiant intensities of flames over three spectral bands. Huang et al. (1999) set up a flame-flicker monitoring system, where the flicker of a gaseous flame was quantified by computing the oscillation of the radiant intensity of individual pixels within the luminous region of flame images. Lu et al. (2000) designed and evaluated an instrument system for monitoring, characterization, and evaluation of fossil-fuel–fired flames in a utility boiler. Geometrical and luminous parameters of the flame were determined from the images. Wang et al. (2002) reported a method to predict NO_x emissive concentration for a coal boiler by using color flame images and neural networks. However, few of the above studies used turbulent nonpremixed flames as objects.

In many industrial furnace and boiler systems, television systems have been installed to monitor the flame. However, most of the time the only information the flame images are providing is whether the flame is burning. Because of the

Correspondence concerning this article should be addressed to J. F. MacGregor at macgreg@mcmaster.ca.

highly turbulent combustion, the flames are always bouncing around, and even the most experienced operator will have difficulty in judging the performance of the combustion.

In this situation, a monitoring system based on image analysis can become very helpful. This is the motivation of the research carried out in this study. An industrial steam boiler is used as the study object. To analyze the color images, a feature extraction based on multivariate image analysis (MIA) techniques using principal-component analysis (PCA) was developed. Nine features are extracted from the PCA score plot space for each image. Further analysis of the image features and the process measurements are performed using PCA and partial least squares (PLS) to help understand the relationship between the feature variables and the process variables, and to predict the performance of the boiler.

System Setup

The flame monitoring system is shown in Figure 1. The steam boiler studied in this article uses both the waste liquid streams from other processes and natural gas as fuels. Therefore, the overall composition of the fuel often changes dramatically. An analog color camera was installed in the boiler and is connected with a monitor for displaying the live images. In this research, the analog signals were recorded by a normal VCR and then a video card converted the signals on the video tapes into the digital images. The resulting images are RGB (red-green-blue) color images, with the size of 120 $\times$ 160 pixels. Considering the processing time, the imaging sample time is set as 1 frame/s.

Case Studies

Case descriptions

Two cases are studied herein. Case I covers a 114-min period (see Figure 2). In this period of time, only liquid fuel was used. In the first half of this period of time, the liquid fuel flow rate was decreased from 1.5 to 0.75 kg/s; then in the second half of the period, it was increased back to 1.5 kg/s. The steam generated followed the same trend as the fuel flow. Case II covers a 56-min period. In this case, both liquid fuel and natural gas were used. During this period of time, the flow rate of liquid fuel was gradually reduced from 0.889 kg/s to 0 and at the same time the flow rate of natural gas was increased from 1.60 to 1.99 m^3/s, to keep the steam generated at the desired level. The flow rate changes of the fuels are shown in Figure 3a and the resulting change of the steam flow rate is shown in Figure 3b. In both cases, the air/fuel ratio of the boiler was automatically modulated based on the preset control scheme.

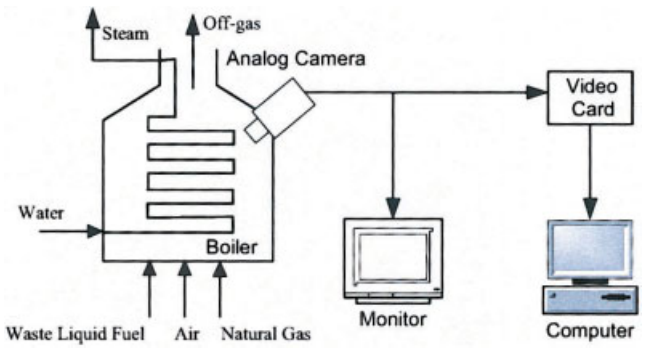


Figure 1. Flame monitoring system.

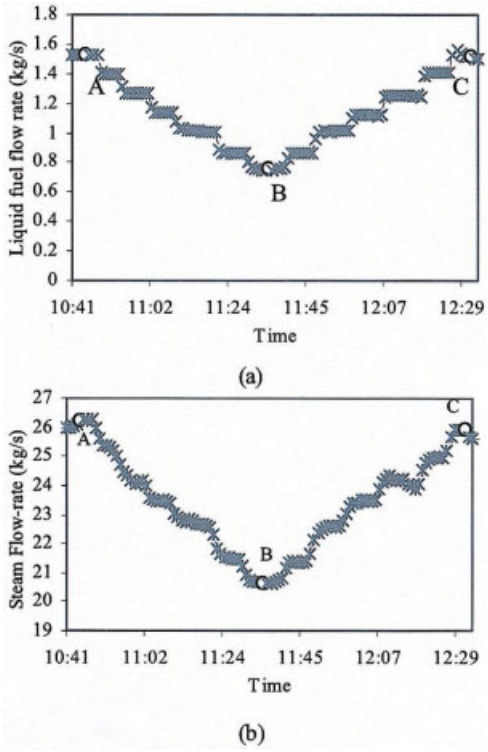


Figure 2. Flow rates of fuel and steam for case I: (a) flow rate of liquid fuel; (b) flow rate of steam generated.

Flame images

For case I a total of 6840 frames were obtained and for case II 3360 frames were obtained. Table 1 lists the process conditions corresponding to the points (A–F) marked in Figures 2 and 3. Some sample images are shown in Figure 4. For each point, two consecutive images with a 1-s time difference are shown. It is reasonable to assume that during this 1 s, the feed and composition conditions in the combustion process did not change. It can be observed that the flames in the boiler appeared highly turbulent, with the images changing significantly over every 1-s interval. This poses considerable difficulty in trying to extract stable information about the combustion process.

Extracting Features from Flame Images

Basic theory of multivariate image analysis

A flame image is an RGB color image, in which the color of each pixel is characterized by the numerical values (normally integers from 0 to 255) of its R, G, and B channels. A color image can then be expressed as a multivariate image composed of three variables (R, G, and B channels). In this work, a practical method based on MIA techniques (Geladi and Grahn, 1996) is developed to extract feature variables from rapidly time varying flame images.

MIA techniques are based on multiway PCA. Without con-

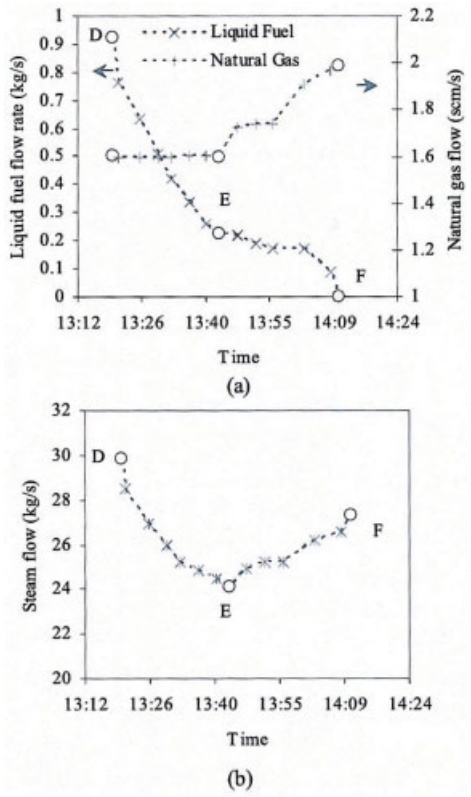


Figure 3. Flow rates of fuels and steam for case II: (a) flow rate of liquid fuel and natural gas; (b) flow rate of steam generated.

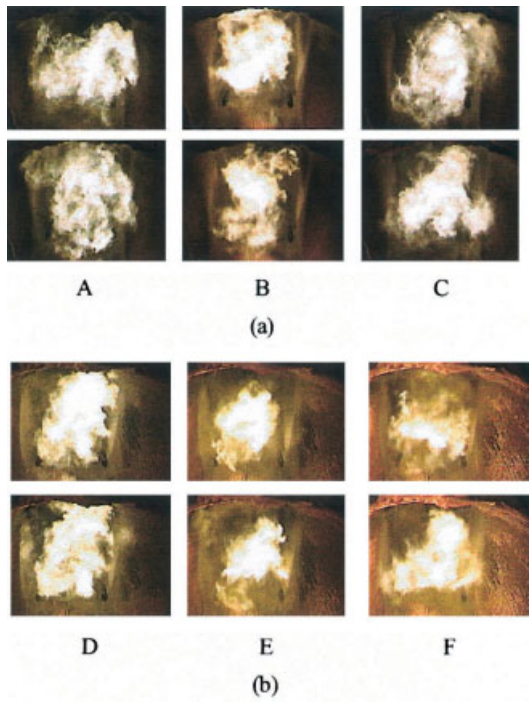


Figure 4. Sample images taken 1 s apart at different conditions: (a) Case I; (b) Case II.

sidering the spatial coordinates of pixels, we can unfold the image matrix and express it as a two-way matrix

$$\underline{\mathbf{I}}_{N_{row}\times N_{col}\times 3} \xrightarrow{unfold} \mathbf{I}_{N\times 3} = \begin{bmatrix} c_{1,r} & c_{1,g} & c_{1,b} \\ \vdots & \vdots & \vdots \\ c_{i,r} & c_{i,g} & c_{i,b} \\ \vdots & \vdots & \vdots \\ c_{N,r} & c_{N,g} & c_{N,b} \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_i \\ \vdots \\ \mathbf{c}_N \end{bmatrix}$$

where $\underline{\mathbf{I}}$ is the three-way image data array with image size $N_{row}\times N_{col}$. $\mathbf{I}$ is the $N\times 3$ unfolded two-way image matrix, where $N = N_{row}\times N_{col}$. N is the number of pixels in the image. $c_{i,r}$, $c_{i,g}$, $c_{i,b}$ ($i = 1, \ldots, N$) are the intensity values of the R, G, and B channels for pixel i. $\mathbf{c}_i$ ($i = 1, \ldots, N$) is the ith row vector of $\mathbf{I}$, which represents the color values of pixel i.

Table 1. Sample Images Taken 1/s apart at Different Conditions

		Condition		
Case	Point	Liquid Fuel (kg/s)	Natural Gas (m^3/s)	Steam (kg/s)
I	A	1.5	0	26.1
	B	0.75	0	20.6
	C	1.5	0	26.1
II	D	0.917	1.61	30.6
	E	0.225	1.58	24.1
	F	0	2.00	27.2

Multiway PCA is equivalent to performing PCA on the unfolded image matrix $\mathbf{I}$

$$\mathbf{I} = \sum_{k=1}^{K} \mathbf{t}_k \mathbf{p}_k^T + \mathbf{E}$$

where K is the number of principal components, the $\mathbf{t}_k$ values are score vectors, and the corresponding $\mathbf{p}_k$ values are loading vectors. For an RGB color image, the maximum number of principal components is 3.

Because the row dimension of the $\mathbf{I}$ matrix is very large (equal to 19,200 for a 120×160 image space) and the column dimension is much smaller (equal to 3 for an RGB color image), a kernel algorithm is used to compute the loading and score vectors. In this algorithm the kernel matrix ($\mathbf{I}^T\mathbf{I}$) is first formed, and then singular value decomposition (SVD) is performed on this very low dimension matrix (3×3 for an RGB color image) to obtain loading vectors $\mathbf{p}_a$ ($a = 1, \ldots, K$). After obtaining loading vectors, the corresponding score vectors $\mathbf{t}_k$ are then computed using $\mathbf{t}_k = \mathbf{I}\mathbf{p}_k$. $\mathbf{t}_k$ is a long vector with length N. After proper scaling and rounding off, it can be refolded into the original image size and displayed as an image

$$s_{k,i} = \text{Round}\left(\frac{t_{k,i} - t_{k,\min}}{t_{k,\max} - t_{k,\min}} \times 255\right) \qquad i = 1, \ldots, N$$

$$(\mathbf{s}_k)_{N\times 1} \xrightarrow{refold} (\mathbf{T}_k)_{N_{row}\times N_{col}}$$

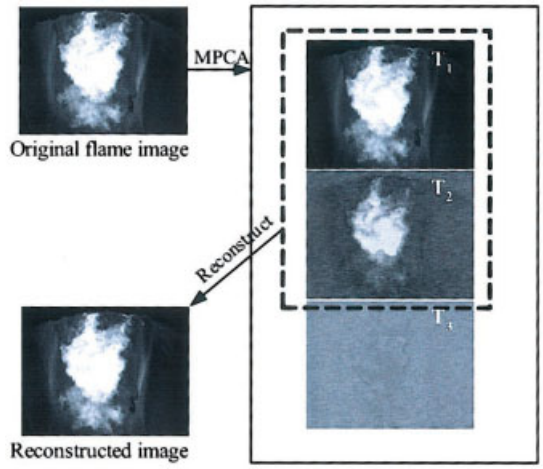


Figure 5. Score images for a sample flame image.

where $\mathbf{s}_k$ is the scaled and rounded score vector and $\mathbf{T}_k$ is the score image of component k. The values of $\mathbf{s}_k$ and $\mathbf{T}_k$ are integers from 0 to 255.

The original MIA is mainly for the analysis on a single multivariate image. To deal with a set of images, the kernel matrix is calculated as $\sum_j \mathbf{I}_j^T \mathbf{I}_j$ and then SVD is performed on this summation of covariance matrix to obtain loadings. It should be also pointed out that in this situation a common score scaling range ($t_{k,\min}$ and $t_{k,\max}$) should be used for all the images.

A total of 100 images were used in this study to compute the loading matrix and score scaling range. The first two components explained 99% of the total variance. In Figure 5, three score images are shown for a sample image. A reconstructed image from the first two components is also shown. It can be seen that the T_3 score image represents the residuals and contains little information. The reconstructed image from the first two components is almost the same as the original image.

Score plot space

Given that the first two components often explain most of the variance, instead of working in original three-dimensional (3-D) RGB space, working in the two-dimensional (2-D) orthogonal t_1– t_2 score space allows us to interpret the images much more easily.

Inspection of the t_1– t_2 score plot (a plot of t_1 values vs. t_2 values) is a common tool in general PCA analysis to give an overview of the whole system and/or detect clusters or outliers. However, when the studied objects are images, because of the large number of pixels, many of the pixels may have nearly identical t_1– t_2 values and plot on top of each other. A 256 × 256 histogram is used to describe the score plot space in this situation. This 2-D histogram, denoted as **TT**, can be obtained from scaled and rounded score vectors, $\mathbf{s}_1$ and $\mathbf{s}_2$. **TT** is a 256 × 256 matrix with elements computed as

$$TT_{i,j} = \sum_l 1 \qquad \forall\ l,\ s_{1l} = i,\ s_{2l} = j$$

$$i, j = 0, \ldots, 255$$

Figure 6 shows the corresponding s_1– s_2 score plots for the sample images shown in Figure 4. A darker shade indicates histogram bins with a low pixel intensity [black indicates histogram bins with no pixels having that (s_1, s_2) combination], and brighter colors indicate histogram bins with higher pixel intensities.

We can see from Figure 6 that for images captured at the same combustion condition, the score plot histograms of pixel intensities are very similar. However, as the combustion conditions change, the locations of pixels in the score plots change noticeably. This is the key result that enables one to use flame images to analyze and monitor the process for changing conditions. For any given process condition, even though the flame images are bouncing around, the PCA score space of those images is very stable, and it changes shape only with changing process conditions.

Each location in the s_1– s_2 score plot represents a certain kind of color (ignoring the residual variation in t_3). We can compute the color for each location by

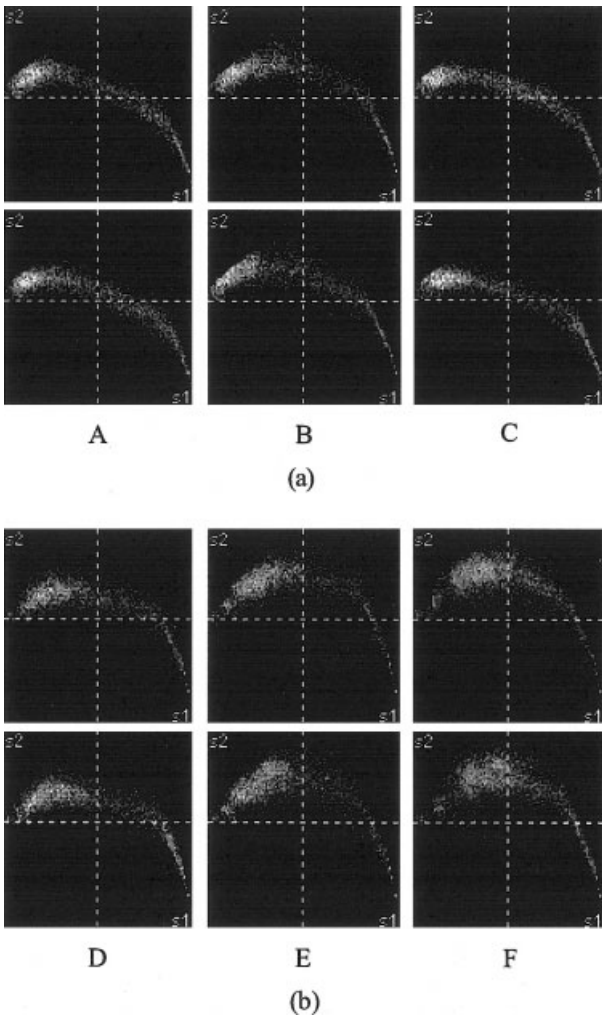


Figure 6. Score plots for the sample images shown in Figure 4 (order is the same): (a) Case I; (b) Case II.

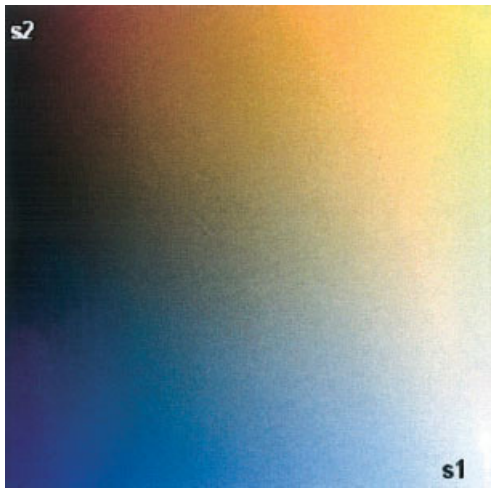


Figure 7. Color plane for s_1– s_2 score plot.

$$[R \quad G \quad B]_{ij} = t_1(i)\mathbf{p}_1^T + t_2(j)\mathbf{p}_2^T \qquad i, j = 0, \ldots, 255$$
$$t_1(i) = i(t_{1,\max} - t_{1,\min}) + t_{1,\min} \qquad t_2(j) = j(t_{2,\max} - t_{2,\min}) + t_{2,\min} \quad (1)$$

A color plane then can be computed for the score plot space, as shown in Figure 7. We can see the relation between the position in score plot and the color. Given that the colors of the flames are different, under different combustion conditions, we observe movement in the score plot space.

Extraction of features from score plot

As seen in the last section, the score plot contains important information about the flame and the corresponding process conditions; however, directly monitoring the process based on the appearance of the score plots is not practical. This is because it is difficult to monitor a time series process by watching changes in a 2-D matrix, and even if observers were able to detect some changes happening in the score plot, it is hard to connect such changes with changes in the process variables. Therefore, we need to extract features from the score plot that have more physical meaning and to further relate those features with the process measurements to help us understand more about the flames and the combustion process.

Basically, the features we are discussing here can be divided into two categories. The first category is called *luminous features*, including the flame luminous region area, flame brightness, uniformity of flame brightness, and the average brightness of nonluminous area. The second category is called *color features*, including average color of the whole image, average color of the flame luminous region, and the number of colors appearing in the flame region.

Flame Luminous Region. The flame luminous region is extracted from the image by choosing a mask in the score plot space. The boundary of the mask is easily obtained by a trial and error process, whereby one selects a mask area in the score plot, selects the pixels lying under it and highlights them in the image space, and iterates until one obtains a mask that extracts the feature of interest. More details on this common masking procedure can be found in references on MIA (Geladi and Grahn 1996). The final mask selected for extracting the flame luminous region is shown as the gray area in Figure 8a. To illustrate the extraction ability of the mask, a sample flame image is shown in Figure 8b. If we set all pixels falling outside this mask to have a gray color, the image shown in Figure 8c is obtained. We can see that the luminous flame region is separated from the other part of the image.

Let us define a 256 × 256 binary matrix **M** that fits over the 256 × 256 score plot as a matrix which has an element equal to 1 if the element location lies under the luminous flame mask and 0 if it does not. This binary matrix **M** can then be used to easily define several useful features of the flame image.

Luminous Features

(1) Luminous region area (A). The area of the flame image that corresponds to the luminous region can be easily computed by counting the number of the pixels falling into the flame mask. This can be defined mathematically using the binary matrix **M** as

$$A = \sum_{i,j} TT_{i,j} \qquad \forall\ (i, j),\ M_{i,j} = 1$$

(2) Flame brightness (B). Flame brightness can be obtained by integrating the luminous intensity level contributed from all pixels falling inside the luminous region. The luminous intensity level for any location in the score plot is computed by converting the color plane obtained by Eq. 1 to a gray-scale plane with elements corresponding to the luminous intensity at each location. The relationship is

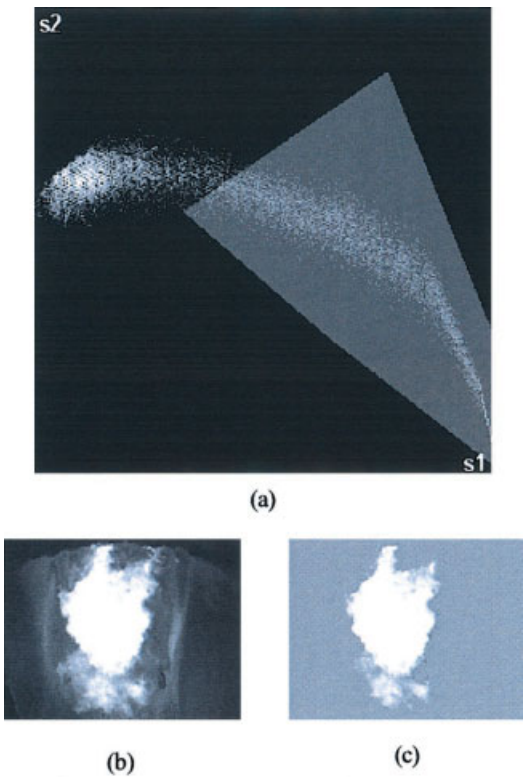


Figure 8. Illustration of flame region mask: (a) score plot and mask, (b) one sample image, (c) the flame region decided by the mask.

$$L_{i,j} = [R \quad G \quad B]_{i,j} \begin{bmatrix} 0.299 \\ 0.587 \\ 0.114 \end{bmatrix} \qquad i, j = 0, \ldots, 255$$

in which L_{ij} is the luminous intensity level for location (i, j) in the score plot, and the conversion coefficient vector is the one that converts the element RGB signal into a corresponding luminous intensity (Matlab Image Processing Toolbox, MathWorks, Natick, MA).

The overall flame brightness is then calculated as

$$B = \sum_{i,j} TT_{i,j} L_{i,j} \qquad \forall\ (i, j),\ M_{i,j} = 1$$

(3) Uniformity of flame brightness (U). The uniformity of the flame brightness is defined as the standard deviation of the flame brightness throughout the luminous region

$$U = \sqrt{\frac{\Sigma_{i,j}\, TT_{i,j} L_{i,j}^2 - \Sigma_{i,j}\, TT_{i,j} L_{i,j}}{\Sigma_{i,j}\, TT_{i,j}}} = \sqrt{\frac{\Sigma_{i,j}\, TT_{i,j} L_{i,j}^2 - B}{A}} \qquad \forall\ (i, j),\ M_{i,j} = 1$$

(4) Average brightness of the nonluminous area (W). As one may observe from the sample images shown in Figure 4, the camera installed is facing the flame, which means we cannot see the length of the flame. Moreover, the flame image is a 2-D projection of a 3-D field and only part of the flame is captured in the image. Therefore, the brightness of the nonluminous area somehow gives information on the length and/or the volume of a flame. We compute the average brightness of the nonluminous area by

$$W = \frac{\Sigma_{i,j}\, TT_{i,j} L_{i,j}}{\Sigma_{i,j}\, TT_{i,j}} \qquad \forall\ (i, j),\ M_{i,j} = 0$$

Color Features

(1) Average color of the whole flame image (s_{1m}, s_{2m}). The average color of the whole flame image is expressed as the average location of pixels in the score plot

$$s_{1m} = \frac{\Sigma_{i,j}\, TT_{i,j} i}{N} \qquad s_{2m} = \frac{\Sigma_{i,j}\, TT_{i,j} j}{N} \qquad i, j = 0, \ldots, 255$$

in which N is the total number of pixels in the image.

(2) Average color of the flame luminous region (s_{1f}, s_{2f}). The average color of the flame luminous region is defined as the average location of the pixels belonging to the flame luminous region

$$s_{1f} = \frac{\Sigma_{i,j}\, TT_{i,j} i}{A} \qquad s_{2f} = \frac{\Sigma_{i,j}\, TT_{i,j} j}{A} \qquad \forall\ (i, j),\ M_{i,j} = 1$$

(3) Total number of colors appearing in the flame luminous region (N_c). The number of the different colors appearing in the flame region is the area of the flame region in the score plot space

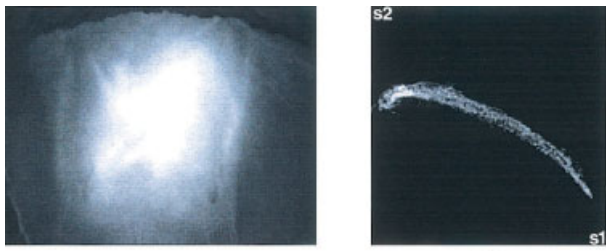


Figure 9. Effect of averaging flames in the image space: (a) an example averaged image (over 60 consecutive images), (b) the corresponding score plot.

$$N_c = \sum_{i,j} 1 \qquad \forall\ (i, j),\ (TT_{i,j} M_{i,j}) \neq 0$$

Filtering the Feature Variables. Because of the highly turbulent nature of the flame the calculated feature variables have a large variation from image to image. Furthermore, because the process characteristics are not changing rapidly, one is usually not interested in monitoring on a 1-s basis. Therefore, some form of image filtering should be performed to reduce the variability of the extracted features.

There are several possible filtering techniques that can be used at the different stages of the calculation. The first one is to perform the filtering operation directly in the image space. This is the most commonly used preprocessing by flame studies in the literature (Lu et al., 2000; Shimoda et al., 1990). However, in the highly turbulent circumstances, as we can see from an example averaged image (Figure 9a), averaging in the image space leads to a loss of the characteristics of the flame. In the score space, the shape of the score plot for the averaged image (see Figure 9b) has also been "distorted" compared to the single-frame images (the score plots of point A in Figure 6).

A second filtering technique is to perform filtering in the score plot space. This represents a much better solution because, as shown earlier, the score plot for each rapidly changing flame image remains very stable as long as the process conditions remain unchanged. Figure 10 shows an averaged score plot over 60 consecutive images. Compared to Figure 9b, this averaged score plot keeps the basic shape of the score plot of the individual image (A in Figure 6). The feature variables extracted from the averaged score plot are expected to summarize the characteristics of the flame during the short period of time over which the average is computed. However, the values of certain features extracted from these filtered score plots may have a different level from the features calculated directly from the raw images because the feature extraction calculations are not linear.

A third filtering approach is to apply a time domain filter to the extracted feature variables of each individual frame image. This approach has some advantages compared to averaging the score plots. First, it is much easier to handle the time series data such that at each time point, the data constitute a vector rather than a matrix. Second, the filtered feature values are directly related to the raw feature variables. The filter used here is an averaging filter with a window length 60.

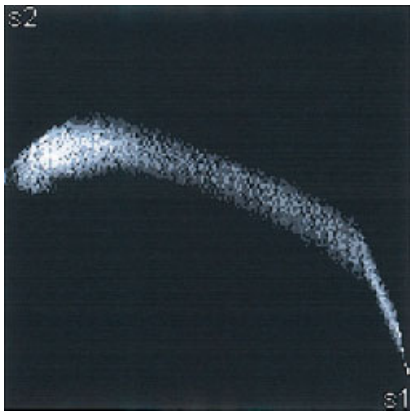


Figure 10. Averaged score plot (over 60 consecutive images).

Results

Feature variables for two case studies

Figure 11 plots three feature variables for the two case studies. Raw features computed from each individual image are shown as dark gray lines. Two types of filtered features are also shown: filtered values of the individual features from each image (indicated by the lighter gray lines) and features computed from the averaged score plots (indicated by the black lines). Filtered data by both filters are almost identical for all the feature variables (such as Figure 11a and c), except for the number of colors (N_c) in the flame luminous region (Figure 11b). The value of N_c computed from the averaged score plot is much higher than those obtained by filtering the N_c values from the time series of the individual image values. However, the two filtered values have similar trends. This is expected, given that the total number of colors in all 60 images averaged in the score plot must be much larger than that in any individual image. In the following computations, only filtered data from the individual image feature time series are used. Using filtered features obtained from the average score plot leads to similar results.

In case I, the liquid fuel flow rate was first decreased and then increased back to the original value. All the feature variables (such as in Figure 11a) follow the same trend or the inverse trend of the liquid fuel flow rate change in Figure 2.

In case II, both liquid fuel and natural gas flow rate were changing. In this case the trends of feature variables are different. Some of the trends are similar to the liquid fuel flow rate change, such as s_{2m} (the average s_2 value of the whole image; Figure 11c), whereas other feature trends are similar to the natural gas flow rate change, such as s_{2f} (the average s_2 value of the flame luminous region). To get a better understanding of the relation between these image feature variables and the process measurements (steam flow and liquid fuel flow rates), principal-component analysis (PCA) and partial least squares (PLS) studies are performed in the next section.

Monitoring the process using multivariate statistical methods

Analysis of the Data Sets Using PCA. In this section, two PCA models are obtained, one for the case I data and another for the case II data. The purpose is to reveal the relationship between the changes of the fuel flow rates and the flame properties. The variables included in the PCA models are the nine filtered flame feature variables and the fuel flow rates.

In the PCA model for case I, the first principal component explained 88.4% of the total variance. This is reasonable because only one factor (flow rate of liquid fuel) was changing in this case. The loading plot of PCA (Figure 12) shows the relative importance of the variables to this principal compo-

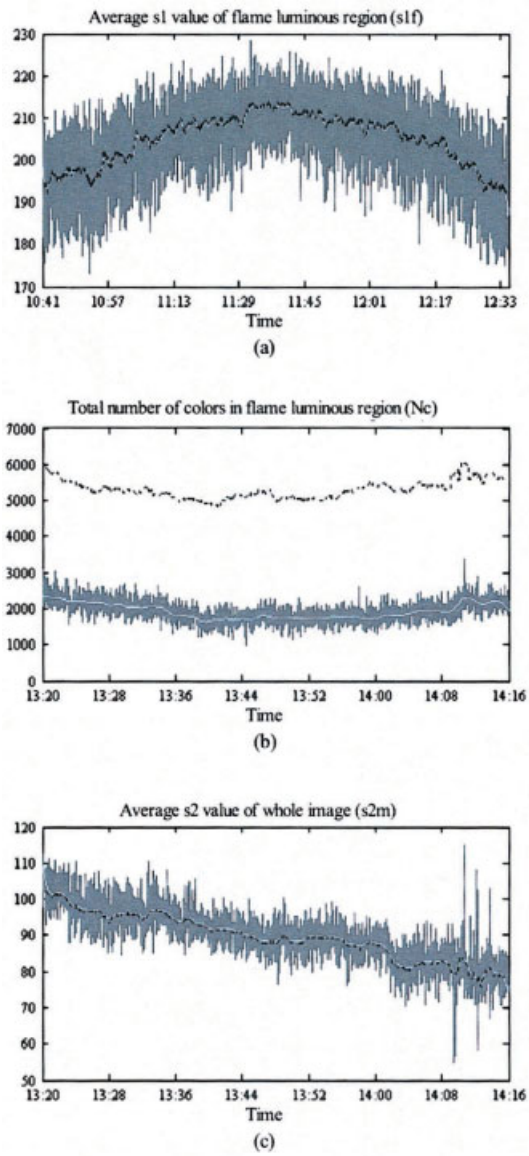


Figure 11. (a) One of the feature variables for case I: average s_1 value of flame luminous region ($s_{1[infi]f}$); (b) One of the feature variables for case II: total number of colors in flame luminous region (N_c); (c) One of the feature variables for case II: average s_2 value of whole image ($s_{2[infi]m}$).

The darker gray lines represent raw values from each 1-s frame image, the lighter gray lines represent the filtered values obtained from directly filtering the raw feature variables, and the black lines represent the filtered values obtained from averaged score plot.

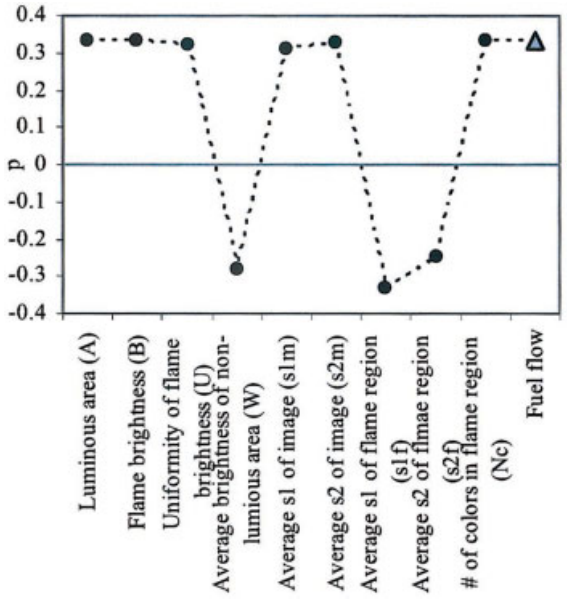


Figure 12. Loading plot of first component of PCA model for case I.

nent. All the features contribute and are highly correlated with the fuel flow. We can see that when the liquid fuel flow increases so does the flame brightness (B), the luminous area (A), and so forth, whereas at the same time the average scores (s_{1f}, s_{2f}) decrease.

The PCA model for the case II data has two significant components because there were two sources of variation. The first component explained 52.3% and the second component explained 42.6% of the total variance. From the scatter plot of the loadings ($p_1 - p_2$) in Figure 13, we can see that s_{2m} (average s_2 value of the image) has positive correlation with liquid fuel flow rate, W (average brightness of nonluminous area) has a positive correlation, and s_{2f} (average s_2 value of the flame luminous area) has a negative correlation with natural gas flow rate. This conclusion is confirmed by the feature variable time

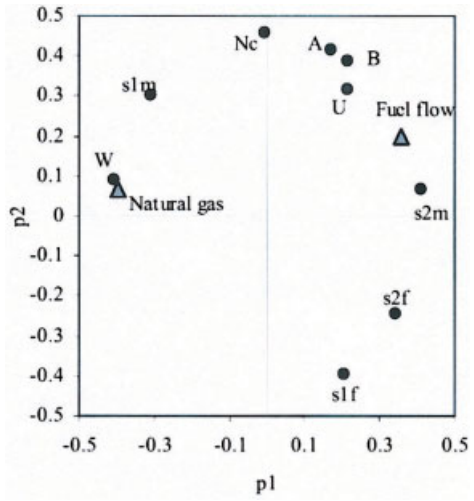


Figure 13. $p_1 - p_2$ loading plot of PCA model for the data of case II.

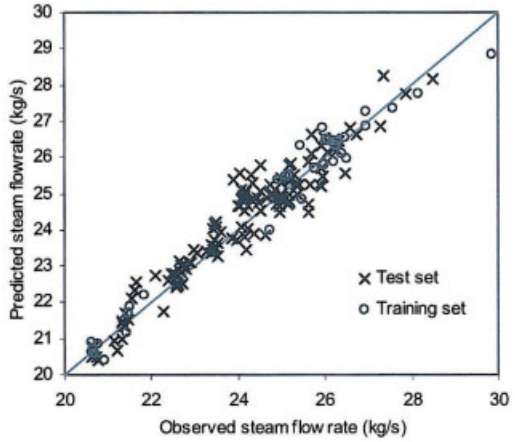


Figure 14. Predicted vs. Observed steam flow rates.

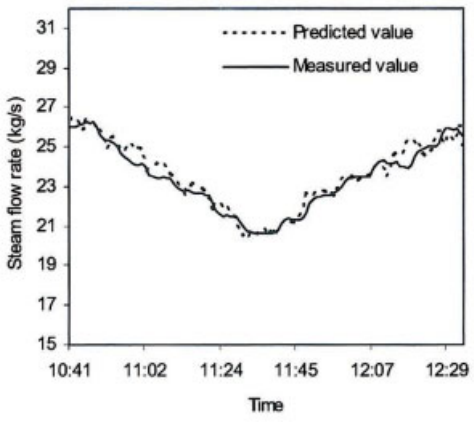
series plots (for example, s_{2m} shown in Figure 11c has a similar trend to the change of liquid fuel flow rate shown in Figure 3a).

Prediction of Flow Rate of Steam Generated. In this section, a single linear PLS model is built using the combined data from both case I and case II to demonstrate that information extracted from the flame images also is sufficient to obtain an accurate prediction of the steam flow rate from the boiler. All nine feature variables from flame images are used as predictors. A total of 33 observations are used as training set and 133 observations are used as a test set.

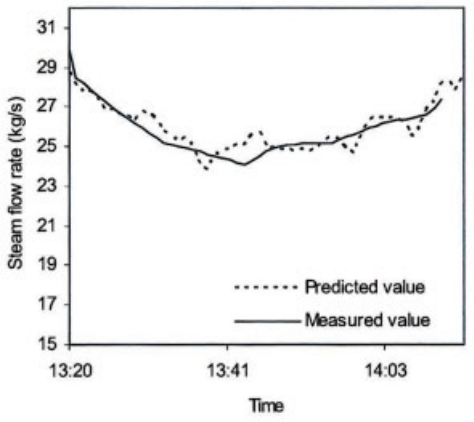
Eight latent variables were determined to be significant by cross-validation. The root mean square of prediction error (RMSE) for the training set is 0.4167 kg/s and for the test set is 0.5 kg/s, demonstrating that the model can predict results from new images almost as well as it fits the training data. The prediction vs. observation plot is shown in Figure 14. Figure 15a and 15b are the time series plots of the predicted and observed steam flow rates for the two case studies, respectively. We can see that the steam flow rates are well predicted.

Prediction of NO_x and SO_2 Concentration in the Off-Gas. It was shown in the above section that by using the features extracted from images as regressors, a PLS regression model can be developed to accurately predict the steam flow rate from the boiler. Although, by itself, prediction of the steam flow rate is not important (this flow rate is easily measured), the fact that it could be predicted so well just from 2-D RGB images of the boiler flame, using no other information, is extremely important. It demonstrates that there is a tremendous amount of information in these images that, if combined with additional process data, could greatly improve the inferential models typically used to monitor the operating performance and emissions from these combustion processes.

In this section and the next section, a preliminary study on the prediction of the NO_x and SO_x concentrations in the off-gas and on the prediction of the energy content in the waste fuel stream for the same boiler system is presented. In this study, 12 videotapes were collected during a 2-month period of time. Each videotape contained one-half hour of recorded flame images. The average process variables during the time when the videotapes were recorded were also collected. The measured NO_x and SO_2 concentrations in the off-gas were obtained from an on-line analyzer. The detailed description of data can be found in Yu (2003).

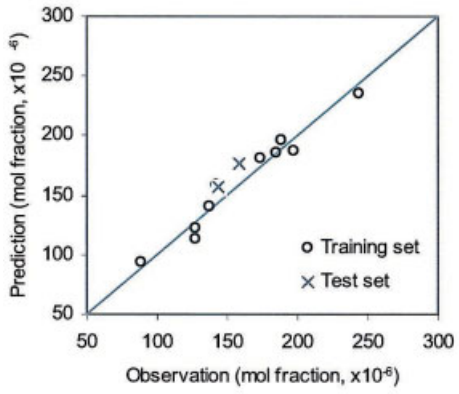


(a)

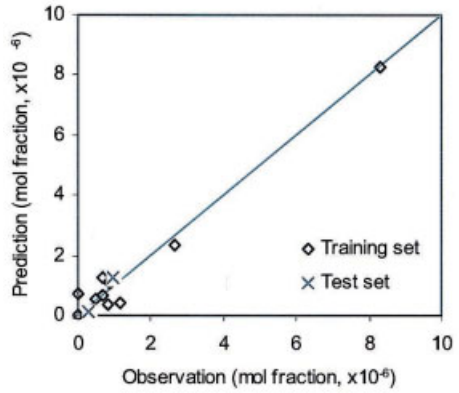


(b)

Figure 15. Time series of predicted and measured steam flow rates: (a) Case I, (b) Case II.



(a)



(b)

Figure 16. Prediction vs. observation plots for the NO_x (a) and SO_2 (b) concentrations.

Two PLS models were developed to predict the emissive concentrations of NO_x and SO_2, respectively. For each model, 10 samples were used as a training set, whereas two samples were used for evaluation (test set). The prediction vs. observation plots are shown in Figure 16. The results show good agreement between the predicted and the measured data. Moreover, by incorporating other process measurements, such as temperatures, pressures, and flow rates, the prediction results were not improved. This implies that the flame images contain the greatest amount of useful information that is correlated with the NO_x and SO_2 concentrations.

Prediction of Energy Content of Waste Fuel Stream. In this section, the same data set as in the last section was used to predict the heat of combustion of the incoming waste fuel stream. The heat of combustion values were measured off-line in a laboratory. A PLS model was built to predict the product of heat of combustion and the liquid fuel flow rate. The heat of combustion can then be obtained by dividing the PLS model prediction by the measured liquid fuel flow rate. In addition to the feature variables extracted from the flame images, the measured flow rates of liquid fuel and natural gas were also included as predictors. This is shown in the following equation:

$$\hat{Q} = [F_{lf} \quad F_{ng} \quad \mathbf{v}]\,\dot{\gamma} \text{ where } Q = H_{lf}F_{lf}$$

$$\hat{H}_{lf} = \hat{Q}/F_{lf}$$

and where $\mathbf{v}$ is the feature vector extracted from the image data, F_{ng} is the natural gas flow rate, F_{lf} is the flow rate of liquid fuel,

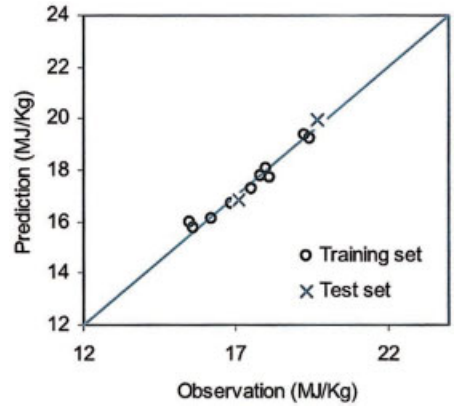


Figure 17. Prediction vs. observation plot for prediction of the heat of combustion of the liquid fuel.

$\dot{\gamma}$ is the model regression coefficient vector, and H_{lf} is the heat of combustion.

Ten samples are used as a training set and two samples are used as a test set. Figure 17 shows the prediction vs. observation plots for both models. We can see that the model has good prediction performance.

Conclusion

A combustion monitoring system based on color (RGB) images of the flame was developed for an industrial steam boiler. Experimental results demonstrate that the system is capable of characterizing the luminous and color parameters of a flame qualitatively and quantitatively, and of predicting the performance of the boiler system over a wide range of conditions.

The key methodology in this flame monitoring system is developed based on multivariate image analysis using multiway principal-component analysis methods. The rapidly time varying flame images were shown to give very stable score plot histograms in the principal-component space. These score plots were very stable for all flames imaged under constant process conditions, but changed in a consistent way whenever the process conditions changed. Masking methods were used to extract a number of features of the flames from the principal-component score plots. These features were shown to be highly related to the process fuel feed rates. Using these features as regressors, several partial least squares regression models were developed to accurately predict the steam flow rate from the boiler, the NO_x and SO_2 concentrations in the off-gas, and the energy content of the incoming waste liquid stream. This study demonstrates that there is a tremendous amount of information in these flame images, which can be successfully used to monitor the operating performance and the emissions of the combustion processes.

Literature Cited

Geladi, P., and Grahn, H., *Multivariate Image Analysis*, Wiley, Chichester, UK (1996).

Huang, Y., Y. Yan, G. Lu, and A. Reed, "On-line Flicker Measurement of Gaseous Flames by Image Processing and Spectral Analysis," *Meas. Sci Technol.*, **10**, 726 (1999).

Lu, G., Y. Yan, and D. D. Ward, "Advanced Monitoring, Characterisation and Evaluation of Gas Fired Flames in a Utility Boiler," *J. Inst. Energy*, **73**, 43 (2000).

Shimoda, M., A. Sugano, T. Kimura, Y. Watanabe, and K. Ishiyama, "Prediction Methods of Unburnt Carbon for Coal Fired Utility Boiler Using Image Processing Technique of Combustion Flame," *IEEE Trans. Energy Convers.*, **5**(4), 640 (1990).

Wang, F., X. J. Wang, Z. Y. Ma, J. H. Yan, Y. Chi, C. Y. Wei, M. J. Ni, and K. F. Cen, "The Research on the Estimation for the NO_x Emissive Concentration of the Pulverized Coal Boiler by the Flame Image Processing Technique," *Fuel*, **81**, 2113 (2002).

Yamaguchi, T., K. T. V. Grattan, H. Uchiyama, and T. Yamada, "A Practical Fiber Optic Air-Ratio Sensor Operating by Flame Color Detection," *Rev. Sci. Instrum.*, **68**(1), 197 (1997).

Yu, H., "Development of Vision-based Inferential Sensors for Process Monitoring and Control," Ph.D. Thesis, McMaster University, Hamilton, Ontario, Canada (2003).

Manuscript received Dec. 13, 2002, and revision received Sep. 3, 2003.