Clustering-Based Performance Optimization of the Boiler–Turbine System

Andrew Kusiak, *Member, IEEE*, and Zhe Song

***Abstract*—In this paper, two optimization models for improvement of the boiler–turbine system performance are formulated. The models are constructed using a data-mining approach. Historical process data is clustered and the discovered patterns are selected for performance improvement of the boiler–turbine system. The first model optimizes a widely used performance index, the unit heat rate. The second model minimizes the total fuel consumption while meeting the electricity demand. The strengths and weaknesses of the two models are discussed. An industrial case study illustrates the concepts presented in the paper.**

***Index Terms*—Boiler–turbine system, clustering, data mining, performance optimization.**

I. Introduction

IMPROVING performance of a boiler–turbine unit is of interest to the energy industry due to increasing fuel costs. The system performance depends on the accuracy of models and the selected performance metrics.

Performance optimization of a boiler–turbine system is usually considered in two phases. The first is the design and implementation of a control system before the power plant becomes operational. The second is the use of the performance test code [e.g., American Society of Mechanical Engineers (ASME) performance test code] to periodically evaluate the system performance to update the operating parameters (set points) of the controllers. Kuprianov [13] discussed different objective functions to improve boiler thermal efficiency and reduce emissions based on certain test codes (or "a test code"). Farhad *et al*. [10] demonstrated the use of the ASME performance test code in reducing fuel and energy consumption.

Numerous modeling approaches of boiler–turbine systems have focused on using the first principle, e.g., thermodynamics. Researchers applied energy and material balance, material flow, and chemistry to derive models in the form of differential equations. Typical benchmark nonlinear models of boilers and turbines can be found in [2], [3], [8], and [22]. Ben-Abdennour and Lee [5] reported test results of a fuzzy fault accommodation controller. Moon and Lee [20] presented a fuzzy controller that can update the fuzzy rules adaptively by a simple set-point error-checking process. Espinosa *et al*. [9] applied fuzzy logic to identify the boiler–turbine system and implemented it to reduce overshooting and settling time. Yu and Xu [31] discussed the feasibility and efficacy of applying a feedback linearization technique to a nonlinear boiler–turbine model for control of steam pressure and electricity output. Tan *et al*. [28] attempted to determine control settings where distances between the nonlinear system and its corresponding linearization model were minimal; thus, the linear controller's performance was guaranteed.

Other applications of boiler–turbine control can be found in [17], [18], and [23]. The results published in the literature are not based on benchmark nonlinear boiler–turbine models. Fuzzy logic and autotuning techniques were used by [17]. A model predictive control approach [24] was illustrated in the papers by [18] and [23]. Such a technique generally uses an accurate model to predict the system behavior based on the changing inputs, and calls for the continuous solving of a quadratic programming optimization problem.

Although the literature reports progress in controlling boiler–turbine systems, the existing approaches usually are expensive to implement due to uncertainty involved in operating such systems. System errors accumulate due to the assumptions made in modeling. Also, control systems are usually designed to ensure system stability and fast response. System performance metrics, e.g., fuel consumption, are usually not well integrated in the control system. The performance test code is widely used to monitor performance; however, it involves a number of constants that are difficult to obtain, which may cause unreliable test results.

The research reported in this paper focuses on metacontrol of a single boiler–turbine unit to reduce fuel consumption while satisfying megawatt load constraints. The economic and emission aspects are not considered in this paper. They can be integrated into the load dispatch optimization models [1]. However, the approach presented in this paper is general, and it allows for solving models with a variety of objectives and constraints. It is obvious that controllers do not fully capture the boiler–turbine system dynamics due to process changes, e.g., boiler aging. Opportunities exist to adjust (bias) controllable parameters to improve performance. A data-driven approach is proposed to generate control settings to improve the performance of the boiler–turbine system.

II. Optimization Models and Performance Metrics

A. Performance Criteria

Before the optimization models are presented, three performance metrics are discussed. A performance criterion directly impacts the optimization result. A widely used metric for boiler–turbine unit (steam turbine) performance is the *unit heat rate*

Manuscript received July 20, 2006; revised October 29, 2007. This work was supported in part by the Iowa Energy Center under Grant IEC 04-06. Paper no. TEC-00362-2006.

The authors are with the Intelligent Systems Laboratory, Department of Mechanical and Industrial Engineering, The University of Iowa, Iowa City, IA 52242-1527 USA (e-mail: andrew-kusiak@uiowa.edu; zhe-song@uiowa.edu).

Digital Object Identifier 10.1109/TEC.2007.914183

0885-8969/$25.00 © 2008 IEEE

(UHR) in

$$\text{UHR} = \frac{\text{Heat Input Rate to Turbine}}{\text{Generator Electricity Output} \times \text{BE}} \quad (1)$$

where the boiler efficiency (BE) is calculated from the heat-loss metric [25], [27]

$$\text{BE}(\%) = 100\% - \text{Boiler Heat Loss } \%. \quad (2)$$

A lower value of the UHR implies higher boiler–turbine system performance. Equations (1) and (2) call for accurate values (calculated or measured) of the heat input (e.g., the heat contained in the fuel, the heat of the entering air) and accurate values of various heat losses [25], e.g., the heat loss due to dry gas or heat loss due to the moisture in fuel. It is obvious that the UHR and BE are susceptible to errors in real industrial environments.

Besides the UHR, another performance metric is used in the electric power industry, the fuel electricity rate *(*FER*)*, which is the ratio of fuel British Thermal Unit (Btu) rate (Btu/h) and the electricity produced by a generator (in megawatt), as shown in

$$\text{FER} = \frac{\text{Fuel Btu Rate}}{\text{Electricity Output}}. \quad (3)$$

A power plant may use different types of fuel (e.g., coal and biomass) with varied Btu content, which is suitably captured by the numerator of (3). It is also easy to see from (3) that a low FER value of the boiler–turbine system is desired. Thus, for a fixed load (electricity produced), a boiler–turbine unit with lower FER burns less fuel. Note that (3) involves two parameters, the electricity output that is accurately measured, and the fuel Btu rate, which can be accurately estimated.

B. Optimization Models

Optimizing a boiler–turbine unit studied in this research is essentially equivalent to solving a metacontrol problem. The solution provides values of controllable variables, such as feeder speed, fan speed, preheat coil temperature, and so on.

Assume a boiler–turbine system can be described by a triplet $(\boldsymbol{u}, \boldsymbol{x}, \boldsymbol{v})$, where $\boldsymbol{u} \in R^l$ is a vector of l controllable variables, $\boldsymbol{v} \in R^m$ is a vector of m noncontrollable variables (e.g., outside air temperature, river water temperature), and $\boldsymbol{x} \in R^k$ is a vector of k system state variables (e.g., temperature, megawatt load, UHR, and turbine heat rate) [6], [26]. Most state variables are measured, but some are calculated. The state variables are also called response variables, as they change according to the changes of controllable and noncontrollable variables. Highly correlated state variables can be removed from consideration due to the redundant information.

Assume that the boiler–turbine system is represented as $\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{u}, \boldsymbol{v})$, where $\boldsymbol{f}(.)$ is a function capturing the process in the steady state [26]. The $\boldsymbol{x} = \boldsymbol{f}(\boldsymbol{u}, \boldsymbol{v})$ can be also expressed as: $x(1) = f_1(\boldsymbol{u}, \boldsymbol{v}), x(2) = f_2(\boldsymbol{u}, \boldsymbol{v}), \ldots, x(k) = f_k(\boldsymbol{u}, \boldsymbol{v})$.

Let $= [0\,0\,0 \ldots 1 \ldots 0\,0\,0]_{1\times k}$ be a vector projecting $\boldsymbol{x}$ into the desired performance metric, such as the UHR or FER, thus $\boldsymbol{C}x = \boldsymbol{C}f(\boldsymbol{u}, \boldsymbol{v})$. Similarly, matrix $\boldsymbol{D}$ with a suitable dimension can be defined to extract or linearly combine all the other state variables needed to be constrained except for the desired performance metric, $\boldsymbol{D}\boldsymbol{x} = \boldsymbol{D}\boldsymbol{f}(\boldsymbol{u}, \boldsymbol{v})$. The state variables include megawatt load, steam pressure, and so on.

The boiler–turbine performance optimization model is formulated next.

Model 1

$$\begin{aligned} &\arg\min_{\boldsymbol{u}} \boldsymbol{C}\boldsymbol{f}(\boldsymbol{u}, \boldsymbol{v}) \\ &\text{s.t.} \boldsymbol{u} \in \boldsymbol{U} \\ &\boldsymbol{D}\boldsymbol{f}(\boldsymbol{u}, \boldsymbol{v}) \in \mathbf{X}_D \end{aligned} \quad (4)$$

where $\mathbf{U}$ and $\mathbf{X}_D$ are the constraint set of controllable variables and the constraint set state variables, respectively. For example, the feeder speed is limited by its designed capacity, and megawatt load is determined by a contract.

Model 1 can accommodate the previously discussed performance criteria and is expressed in the following two different forms:

*Model 2 (*UHR *optimization)*

$$\begin{aligned} &\arg\min_{\boldsymbol{u}} \text{UHR}(\boldsymbol{u}, \boldsymbol{v}) \\ &\text{s.t. } \boldsymbol{u} \in \mathbf{U} \\ &\boldsymbol{D}f(\boldsymbol{u}, \boldsymbol{v}) \in \mathbf{X}_D. \end{aligned} \quad (5)$$

*Model 3 (*FER *optimization)*

$$\begin{aligned} &\arg\min_{\boldsymbol{u}} \text{FER}(\boldsymbol{u}, \boldsymbol{v}) \\ &\text{s.t.} \boldsymbol{u} \in \mathbf{U} \\ &\boldsymbol{D}f(\boldsymbol{u}, \boldsymbol{v}) \in \mathbf{X}_D. \end{aligned} \quad (6)$$

In practice, performance optimization of the boiler–turbine has to be considered for a fixed electricity output in a steady state. Thus, Model 3 can be transformed to minimize the fuel Btu rate (fuel input) subject to load demand.

*Model 4 (*FER *optimization)*

$$\begin{aligned} &\arg\min_{\boldsymbol{u}} \text{Fuel Btu Rate}(\boldsymbol{u}, \boldsymbol{v}) \\ &\text{s.t.} \boldsymbol{u} \in \mathbf{U} \\ &\boldsymbol{D}\boldsymbol{f}(\boldsymbol{u}, \boldsymbol{v}) \in \mathbf{X}_D. \end{aligned} \quad (7)$$

The fuel Btu rate can be inferred from the feeder speed and other fuel-related parameters. A practical issue arises here as to whether the total feeder speed is a good indication of the fuel Btu rate. One easy way to resolve this issue is to determine a correlation between the total feeder speed and the heat input to the turbine. High total feeder speed should lead to more heat into the boiler, and thus, more heat to the turbine. Fig. 1 shows the correlation between the total feeder speed of the boiler and heat input into the turbine. Seven data sets were randomly selected, each with 10 080 data points (1 week of data). It is obvious that the total feeder speed and the heat input to the turbine are highly correlated. Thus, the total feeder speed is a good approximation of the total fuel input to the boiler.

Solving Model 4 optimally guarantees the minimum fuel input. However, the solution of Model *2* does not guarantee the minimum fuel input, rather the minimum UHR. Also, the speed of fuel input can be relatively easily estimated from the speed

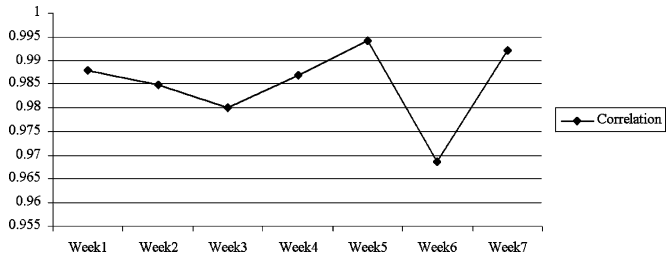

Fig. 1. Correlation between the total feeder speed and the heat input to the turbine.

of the fuel feeder. Note that the computed values of the UHR may involve large errors.

III. Data-Driven Methods

Analytical models of boiler performance are highly nonlinear, yet, modeling $\boldsymbol{f}(.)$ is fundamental to real-time optimization. Optimal solving of such models with classical optimization algorithms (e.g., nonlinear programming) is computationally expensive, especially since the models need be solved repeatedly for continuous performance improvement.

Data-driven models offer a viable alternative to analytical modeling. Among them, neural networks and fuzzy logic have found some applications [9], [11], [21]. Data-mining algorithms [30] are the latest addition to the data-driven methods of interest to the power industry. Some of the applications of data mining in the power industry are discussed in [7] and [14]–[16].

Both neural networks and fuzzy logic are good candidates for approximating and controlling nonlinear systems [9], [21]. However, the limiting factor of neural networks is a long training time when modeling large-scale and time-shifting processes. The main power of fuzzy logic modeling lies in its transformation of linguistic expressions into numeric values.

In this paper, Models 2 and 4 are solved by direct search of the centroid space. The analytical function $\boldsymbol{f}(.)$ is not required by the clustering-based approach. Historical process data is clustered, and patterns leading to high performance are stored. These patterns (cluster centroids) are further selected based on the current process's status and other operational constraints. The global optimum is not guaranteed; however, local optima leading to improved performance are determined.

Denote the high-dimensional data point recorded at time t from the boiler–turbine system as $P_t = [x(1)_t, \ldots, x(k)_t, u(1)_t, \ldots, u(l)_t, v(1)_t, \ldots, v(m)_t]^T$, where $x(1)_t$ is the first state variable's value at time t, $u(1)_t$ and $v(1)_t$ are controllable and noncontrollable variables, respectively, and the total dimension of the data point is $(k + l + m)$. Later $x(1)_t$ becomes the performance index UHR, and $x(2)_t$ denotes the *megawatt load*. To simplify the discussion, *megawatt load* is the only state variable to be constrained. However, the ideas introduced in this paper are easily generalizable to multiple constraints.

The historical training data set is $\{P_{t_0}, P_{t_1}, P_{t_2}, \ldots, P_{t_n}\}$ collected over a time horizon governed by the function $\boldsymbol{f}(.)$. The *K-means* clustering algorithm [19], [29] applied to the training data, produces a set of centroids capturing patterns corresponding to this function. Basic steps of the *K*-means algorithm based on notation presented in Table I are shown next.

TABLE I
Notation Used in the *K-means* Clustering Algorithm

Symbol	Description
C_i	The ith cluster
c_i	The centroid of cluster C_i
K	The number of clusters
q_i	The number of points in C_i

1) Select *K* points as initial centroids given a set of points;
2) **Repeat;**
3) Form *K* clusters by assigning each point to its closest centroid;
4) Recompute the centroid of each cluster;
5) **Until** the centroids do not change.

The centroid of a cluster C_i is computed as $c_i = 1/q_i \sum_{P_t \in C_i} P_t$. The radius of a cluster C_i is computed from $r_i = \sqrt{1/q_i \sum_{P_t \in C_i} (P_t - c_i)^2}$. Each cluster is composed of "similar" data points.

Let the *centroids* set be $\{c_1, c_2, c_3, \ldots, c_K\}$, where $c_1 = [x(1)_{c_1}, \ldots, x(k)_{c_1}, u(1)_{c_1}, \ldots, u(l)_{c_1}, v(1)_{c_1}, \ldots, v(m)_{c_1}]^T$. More details about *K-means* clustering and how to form centroids can be found at [29]. Basically it is assumed that $\boldsymbol{x}_c \approx \boldsymbol{f}(\boldsymbol{u}_c, \boldsymbol{v}_c)$ holds for centroids with bounded errors, where $\boldsymbol{x}_c$ represents the response variables of the centroid c, and $\boldsymbol{u}_c$ and $\boldsymbol{v}_c$ are the controllable and noncontrollable variables of the centroid c.

Observation: Let the points in $\{P_1, P_2, P_2, \ldots, P_n\}$ belong to centroid c, the centroid $c = (1/n)\sum_{i=1}^{n} P_i$, and the error between $x(1)_c$ and $f_1(\boldsymbol{u}_c, \boldsymbol{v}_c)$ is bounded, if f_1 is continuously differentiable at each point of an open set $S \subset R^{m+l}$.

Proof: From the definition of *K-means* algorithm, $x(1)_c = (1/n)\sum_{i=1}^{n} x(1)_i = (1/n)\sum_{i=1}^{n} f_1(\boldsymbol{u}_i, \boldsymbol{v}_i)$, where $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are the controllable and noncontrollable components of point P_i.

Therefore

$$\varepsilon = x(1)_c - f_1(\boldsymbol{u}_c, \boldsymbol{v}_c) = (1/n)\sum_{i=1}^{n} f_1(\boldsymbol{u}_i, \boldsymbol{v}_i) - f_1(\boldsymbol{u}_c, \boldsymbol{v}_c)$$

$$= (1/n)\sum_{i=1}^{n} [f_1(\boldsymbol{u}_i, \boldsymbol{v}_i) - f_1(\boldsymbol{u}_c, \boldsymbol{v}_c)].$$

For simplicity, $(\boldsymbol{u}, \boldsymbol{v})$ is denoted as variable ω, and then $\varepsilon = (1/n)\sum_{i=1}^{n} [f_1(\boldsymbol{\omega}_\iota) - f_1(\boldsymbol{\omega}_c)]$.

Apply the mean-value theorem [12],

$(1/n)\sum_{i=1}^{n} [f_1(\boldsymbol{\omega}_\iota) - f_1(\boldsymbol{\omega}_c)] = (1/n)\sum_{i=1}^{n} [\partial f_1/\partial \boldsymbol{\omega}_{\boldsymbol{\omega}=\alpha_\iota} (\boldsymbol{\omega}_\iota - \boldsymbol{\omega}_c)]$, where $\boldsymbol{a}_i$ is a point on the line segment joining $\boldsymbol{\omega}_\iota$ and $\boldsymbol{\omega}_c$. Then

$$|\varepsilon| = \left| (1/n)\sum_{i=1}^{n} \left[\frac{\partial f_1}{\partial \boldsymbol{\omega}}_{\boldsymbol{\omega}=\alpha_\iota} (\boldsymbol{\omega}_\iota - \boldsymbol{\omega}_c) \right] \right|$$

$$\leq \sqrt{\left(\sum_{i=1}^{n} \left[\frac{\partial f_1}{\partial \boldsymbol{\omega}}_{\boldsymbol{\omega}=\alpha_\iota} (\boldsymbol{\omega}_\iota - \boldsymbol{\omega}_c) \right]^2 \right) \Big/ n}.$$

Let $\varphi^2 = \max_{i=1\ldots n}\{(\partial f_1/\partial \boldsymbol{\omega}_{\boldsymbol{\omega}=\alpha_\iota})^2\}, \varphi \geq 0,$

$$\sqrt{\left(\sum_{i=1}^{n}\left[\frac{\partial f_1}{\partial \boldsymbol{\omega}}_{\boldsymbol{\omega}=\alpha_\iota}(\boldsymbol{\omega}_\iota - \boldsymbol{\omega}_c)\right]^2\right)\Big/ n} \leq \sqrt{\left(\varphi^2 \sum_{i=1}^{n}(\boldsymbol{\omega}_\iota - \boldsymbol{\omega}_c)^2\right)\Big/ n} = r\varphi$$

where

$$r = \sqrt{\left(\sum_{i=1}^{n}(\boldsymbol{\omega}_\iota - \boldsymbol{\omega}_c)^2\right)\Big/ n}$$

It is obvious that the error ε is bounded by φ, the maximum absolute gradient of the function f_1, and r, the cluster's radius projected on the $(\boldsymbol{u}, \boldsymbol{v})$ dimensions. By varying the radius of the cluster *C*, reasonable accuracy can be produced. An increase in the number of clusters *K* generally decreases the cluster radius. Based on the observation of Section III, the errors between $\boldsymbol{x}_c$ and $f(\boldsymbol{u}_c, \boldsymbol{v}_c)$ are also bounded.

Suppose current boiler–turbine system steady status is P_t. To minimize $x(1)_t$ and satisfy the *megawatt load* constraint (assume the load demand is *M*), search the centroids set and find the centroid c with the minimum $x(1)_c$ and $x(2)_c = M$ with some acceptable tolerance. After changing the controllable variables from $\boldsymbol{u}_t$ to $\boldsymbol{u}_c$, $\boldsymbol{x}_t$ should change toward $\boldsymbol{x}_c$ if the distance between $\boldsymbol{v}_t$ and $\boldsymbol{v}_c$ is small. However, in a dynamic system, the system's current states impact the desired states. Thus, the distance between $\boldsymbol{x}_t$ and $\boldsymbol{x}_c$ should be considered in the search process. Making large changes in the system input is not desired. Thus, the distance between $\boldsymbol{u}_t$ and $\boldsymbol{u}_c$ should also be considered. The Euclidean distance, i.e.,

$$\|\boldsymbol{u}_t - \boldsymbol{u}_c\| = \sqrt{\sum_{i=1}^{l}(u(i)_t - u(i)_c)^2}, \|\boldsymbol{v}_t - \boldsymbol{v}_c\| = \sqrt{\sum_{i=1}^{m}(v(i)_t - v(i)_c)^2},$$

$\|\boldsymbol{x}_t - \boldsymbol{x}_c\| = \sqrt{\sum_{i=1}^{k}(x(i)_t - x(i)_c)^2}$ has been selected as the preferred metric. The weighted Euclidean distance can be considered if the weights are needed to differentiate the importance of individual variables.

The process of minimizing the boiler–turbine performance criterion involves searching the nearest centroid with small $x(1)_c$ and satisfying all the constraints. Note that the clustering method learns the patterns from historical data, and therefore, each centroid's controllable settings are feasible. For FER optimization (Model 4), the search can be simplified to finding the nearest centroid satisfying all the constraints and with a smaller fuel input speed. Searching time in a centroid space is short, and this optimization process can be performed repeatedly, thus continuously improving the boiler–turbine performance.

TABLE II
TEN EXPERIMENTAL DATA SETS

Exp Data Set	Start Time Stamp	End Time Stamp
1	8/29/04 12:00 AM	9/4/04 11:59 PM
2	4/16/04 12:00 AM	4/22/04 11:59 PM
3	7/23/04 12:00 AM	7/29/04 11:59 PM
4	7/16/04 12:00 AM	7/22/04 11:59 PM
5	6/11/04 12:00 AM	6/17/04 11:59 PM
6	4/23/04 12:00 AM	4/29/04 11:59 PM
7	4/30/04 12:00 AM	5/6/04 11:59 PM
8	5/7/04 12:00 AM	5/13/04 11:59 PM
9	5/14/04 12:00 AM	5/20/04 11:59 PM
10	6/4/04 12:00 AM	6/10/04 11:59 PM

TABLE III
CHARACTERISTICS OF THE INDUSTRIAL BOILER DATA SET

Process variable	Description	Process variable	Description
$u(1)$	Feeder A speed	$u(12)$	Lower damper position
$u(2)$	Feeder B speed	$u(13)$	Mill A fuel air damper position
$u(3)$	Feeder C speed	$u(14)$	Mill B fuel air damper position
$u(4)$	Feeder D speed	$u(15)$	Mill C fuel air damper position
$u(5)$	Air temp from preheat coil 1	$u(16)$	Mill D fuel air damper position
$u(6)$	Air temp from preheat coil 2	$u(17)$	ID fan A speed
$u(7)$	RHTR water flow	$u(18)$	ID fan B speed
$u(8)$	SHTR water flow	$v(1)$	Outside air temp
$u(9)$	Burner tilt degrees	$v(2)$	Circulating water inlet temp
$u(10)$	Upper damper position	$x(1)$	Unit heat rate
$u(11)$	Middle damper position	$x(2)$	Megawatt load

IV. INDUSTRIAL CASE STUDY

The data used in this project was generated from a 140-MW tangentially-fired boiler, 860 000 lbs/h, 2050 psi superheat, 759 000 lbs/h reheat, 1005 °F/1005 °F superheat/reheat temperatures.

Ten data sets were considered for the different experiments, each including 10 080 data points (7 days, see Table II). The raw data recorded from the boiler–turbine system was denoised and scaled. Other data preprocessing techniques can be applied to improve the quality of the data. Table III shows the process variables of the industrial data set used in this paper. The target-constrained response variable is the megawatt load.

To validate the proposed *K-means* clustering-based methods for optimizing a boiler–turbine performance, a virtual testing technique [16] was used for industrial data. An industrial data set collected for a period of 7 days at 1-min intervals was used. The data collected over the first 6 days (a training data set) was used to construct a *centroids* set. For each $P_t = [\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{v}_t]^T$ of the 1440 data points of the 7th day, centroid c is retrieved from the *Centroids* set based on some criteria (see Models 5 or 6). Then, values $\boldsymbol{u}_t$ are substituted with $\boldsymbol{u}_c$ (i.e., $P_t = [\boldsymbol{x}_t, \boldsymbol{u}_c, \boldsymbol{v}_t]^T$). Suppose the process model $\boldsymbol{f}(.)$ is known. Then, $\boldsymbol{x}_t$ is compared with $f(\boldsymbol{u}_c, \boldsymbol{v}_t)$ to see whether the performance is improved while the constraints are satisfied.

A neural network (NN) approach was used to capture the process model $\boldsymbol{f}(.)$ from the 7-day data set. Two NNs were trained, one to capture the function $x(1) = f_1(u(1), \ldots, u(18), v(1), v(2))$, the other to capture the function $x(2) = f_2(u(1), \ldots, u(18), v(1), v(2))$. The function f_1

TABLE IV
TEN CATEGORIES OF THE UHR

LT_9000	9000_9375	9375_9750	9750_10125	10125_10500
10500_10875	10875_11250	11250_11625	11625_12000	HT_12000

was used to predict whether the UHR would be reduced after applying the derived control settings. The function f_2 was used to predict whether the megawatt load would exceed the demand constraint. In this industrial case study, the megawatt load had to be constrained within *demand* $\pm$ 1 MW. In the test, the *demand* was assumed to be equal to a testing data point's megawatt load on the $7\mathrm{th}$ day.

A. Optimization of the UHR

To optimize the UHR in a centroids space, Model 2 is instantiated as Model 5.

Model 5

$$\begin{aligned}&\underset{c\in\mathrm{Centroids}}{\arg\min}\ |c-P_t|\\&\mathrm{s.t.}x(1)_c<x(1)_t\\&|x(2)_c-x(2)_t|\le 1\\&|v(1)_c-v(1)_t|=0,|v(2)_c-v(2)_t|=0.\end{aligned}\tag{8}$$

At time t, the current process status is P_t. Optimization of Model 5 involves finding in the *Centroids* set $\{c_1, c_2, c_3, \ldots, c_K\}$ centroid c minimizing the distance between c and P_t subject to various constraints. The searched centroid results in lower UHR than the one of P_t. The centroids $\{c_1, c_2, c_3, \ldots, c_K\}$ are extracted with the K-means algorithm applied to the training data set (the 6-day data set).

In computational experiments, the training data set was categorized into a number of subtraining data sets based on the UHR. Based on the domain expertise, the UHR was divided into ten categories (see Table IV). "LT_9000" means the UHR is lower than 9000, "9000_9375" means the UHR is greater than or equal to 9000, less than 9375.

The heuristic procedure of solving Model 5 for each point of the 7th day is as follows.

Step 1) Divide the 7-day data set (data set 1) into two data sets (data sets 2 and 3). Data set 2 consists of the day 1–day 6 data points. Data set 3 includes the 7th day data points.

Step 2) Categorize data set 2 into ten subsets based on the UHR levels, i.e., for each UHR level, there is a corresponding data set.

Step 3) For each subset, with $K = \lambda$ multiplied by the number of data points in the subset, apply the *K-means* algorithm to extract centroids; store them into the *Centroids* set.

Step 4) For each data point in data set 3, select centroids which satisfy the first two constraints of Model 5, then among those centroids, select a nearest centroid for the point and use the controllable variables' setting of the centroid to update the point's controllable variables. Save this "controlled" data point in the *controlled data set*.

TABLE V
MEGAWATT LOAD CATEGORIZED INTO 16 INTERVALS

LT_40	40_45	45_50	50_55	55_60	60_65	65_70
70_75	75_80	80_85	85_90	90_95	95_100	100_105
105_110	HT_110					

Note that the equality constraints in Model 5 are not necessarily satisfied in the experiments, otherwise, there would not be enough data points to be controlled. Here, λ is the clustering ratio heuristically determining the number of clusters.

B. Optimization of the Fuel Electricity Rate

The objective function of Model 4 is to minimize the *fuel Btu rate* input to the boiler for the four feeders. For FER optimization, the research question can be stated as follows "Given a megawatt, can one determine a smaller total feeder speed?" In most boiler–turbine control systems, overshooting is common due to the changing combustion and energy transformation processes. For example, previously calculated boiler air and fuel controller parameters may no longer be valid. In this case study, each of the four feeders had a maximum speed of 10.25 RPM (rotations per minute). Next, Model 4 is instantiated as Model 6.

Model 6

$$\begin{aligned}&\underset{c\in\mathrm{Centroids}}{\arg\min}\ |c-P_t|\\&\mathrm{s.t.}\sum_{i=1}^{4}u(i)_c<\sum_{i=1}^{4}u(i)_t\\&|x(2)_c-x(2)_t|\le 1\\&|v(1)_c-v(1)_t|=0,|v(2)_c-v(2)_t|=0.\end{aligned}\tag{9}$$

To simplify computation, the megawatt load was categorized into different levels shown in Table V. The boiler–turbine unit in this case study has a maximum load of 120 MW. It usually runs between 40 MW and 110 MW. The category "LT_40" means lower than 40 MW, "HT_110" means higher than 110 MW, and "40_45" means higher than 40 MW, smaller than or equal to 45 MW.

The heuristic procedure for clustering-based FER optimization is as follows.

Step 1) Divide the 7-day data set (data set 1) into two data sets (data sets 2 and 3). Data set 2 consists of the day 1–day 6 data points. Data set 3 consists of the 7th day data points.

Step 2) Categorize data set 2 into 16 subsets based on the megawatt levels, i.e., each megawatt load level is associated with a corresponding data set.

Step 3) For each subset, with $K = \lambda$ multiplied by the number of data points in the subset, apply the *K-means* algorithm to extract centroids, store them into the *Centroids* set.

Step 4) For each data point in data set 3, select out centroids that satisfy the first two constraints in Model 6. Then, among those centroids, select a nearest centroid for the point and use the controllable variables' setting of the centroid to update the point's controllable

TABLE VI
RESULTS FROM UHR OPTIMIZATION WITH $\lambda = 0.05$

Exp Data Set	Average UHR Change	%Change	# of Instances with Decreased UHR	# of Instances Controlled	Hit %
1	-161.38	-1.56%	1394	1437	97.00%
2	-336.71	-3.52%	325	354	91.81%
3	-45.4	-0.45%	799	1426	56.00%
4	4.98	0.06%	512	1439	35.58%
5	-53.34	-0.53%	1045	1328	78.69%
6	-130.87	-1.27%	480	543	88.39%
7	-157.29	-1.56%	1148	1212	94.70%
8	-97.4	-0.96%	1202	1403	85.67%
9	-112.07	-1.13%	822	823	99.87%
10	-33.25	-0.34%	944	1224	77.12%

TABLE VII
RESULTS FROM UHR OPTIMIZATION WITH $\lambda = 0.07$

Exp Data Set	Average *UHR* Change	%Change	# of Instances with Decreased *UHR*	# of Instances Controlled	Hit %
1	-145.02	-1.40%	1379	1437	95.96%
2	-113.72	-1.12%	225	354	63.56%
3	-70.41	-0.70%	1227	1426	86.04%
4	2.46	0.03%	543	1439	37.73%
5	-52.86	-0.52%	1041	1328	78.39%
6	-110.34	-1.07%	428	543	78.82%
7	-145.26	-1.44%	1098	1212	90.59%
8	-100.59	-1.00%	1209	1403	86.17%
9	-106.82	-1.07%	821	823	99.76%
10	-42.93	-0.44%	960	1224	78.43%

variables. Save this "controlled" data point in the *controlled data set*.

Note that the satisfaction of the equality constraints of Model 6 is not guaranteed, as there might not be enough data points to be controlled.

C. Comparison of the Results and Discussion

The two *controlled data sets* for UHR and FER optimization were evaluated for opportunities to reduce the UHR or total feeder speed for the fixed megawatt load. One concern was whether the UHR would decrease for the fixed megawatt load. Such a concern did not apply to the FER optimization, as the total feeder speed was controllable. The only concern was to satisfy the megawatt load constraint.

Tables VI and VII show the results of the UHR optimization for the *controlled data set*. For example, in the data set of 1440 instances (data points), the procedure of Section IV-A identified 1437 data points on the 7th day for Experiment 1. These points could be controlled to lower the UHR for a fixed megawatt load. An NN was used to predict the UHR of the controlled 1437 data points. The average UHR change for 1437 data points was 161.38. From Table VI, 1394 of the 1437 controlled data points in Experiment 1 resulted in a lower UHR. The remaining ones predicted a somewhat higher UHR. This implies that after the controllable variables were modified according to the computed centroid, the UHR could increase. The "Hit%" in Table VI refers to the percentage of the controlled data points with a decreased UHR. The "% Change" expresses the relative average change of the UHR based on the original UHR.

Tables VIII and IX illustrate the results of the FER optimization for a *controlled* data set. In Experiment 1 of Table VIII, 800 data points were identified by the procedure of Section IV-B. Each data point was controlled with a decreased total feeder speed. The average total feeder speed was decreased by 0.35. The "% Change" reflects the average relative change of the total feeder speed based on the original total feeder speed.

TABLE VIII
RESULTS FROM FER OPTIMIZATION WITH $\lambda = 0.05$

Exp Data Set	Average Feeder Speed Change	%Change	# of Controlled Instances	# of Decrease More Than 0.5 RPM
1	-0.35	-1.83%	800	261
2	-1.21	-3.44%	1384	1104
3	-2.68	-7.34%	1461	1115
4	-0.86	-2.65%	1412	843
5	-0.93	-3.96%	1315	903
6	-0.39	-1.49%	1094	283
7	-1.13	-4.16%	1273	1044
8	-0.79	-2.28%	1368	868
9	-0.64	-1.81%	1401	678
10	-0.39	-0.99%	125	42

TABLE IX
RESULTS FROM FER OPTIMIZATION WITH $\lambda = 0.07$

Exp Data Set	Average Feeder Speed Change	%Change	# of Controlled Instances	# of Decrease More Than 0.5 RPM
1	-0.45	-2.28%	919	381
2	-1.26	-3.47%	1406	1052
3	-2.86	-7.80%	1275	1140
4	-0.99	-3.00%	1412	887
5	-0.72	-3.05%	1062	586
6	-0.4	-1.43%	1235	351
7	-1.06	3.90%	1283	1037
8	-0.81	-2.34%	1366	888
9	-0.59	-1.68%	1408	611
10	-0.41	-1.04%	162	53

TABLE X
CHANGE IN MEGAWATT LOAD WITH $\lambda = 0.05$

Exp Data Set	*UHR* Optimization		*FER* Optimization	
	Average MW Change	%Change	Average MW Change	%Change
1	6.63	11.78%	0.61	1.15%
2	11.47	18.23%	1.84	2.30%
3	5.32	5.80%	2.17	2.11%
4	3.94	4.77%	1.02	1.08%
5	3.99	5.98%	1.17	1.50%
6	2.33	3.64%	1.16	1.47%
7	4.09	5.34%	1.63	2.03%
8	1.61	1.78%	1.04	1.15%
9	0.95	1.04%	0.72	0.76%
10	6.19	6.80%	0.6	0.51%

Tables VI–IX show that there are significant opportunities to decrease the UHR or the total feeder speed from the industrial data sets.

Tables X and XI compare the absolute megawatt load changes based on the two optimization models. The adjusted controllable variables have led to the megawatt load changes according to the underlying combustion and energy conversion principles. It is obvious that the UHR optimization model leads to larger load changes. The FER model could satisfy the $\pm$ 1 MW constraint. Increasing the clustering ratio λ from 0.05 to 0.07 decreases the load changes in UHR optimization. The reason is that increasing the number of clusters decreases the clusters' radius, thus improving the prediction accuracy of a centroid based on the observation of Section III. Increasing λ does not significantly affect the load changes in the FER optimization. The reason is that in the procedure of Section IV-B, the data sets were already

TABLE XI
CHANGE IN MEGAWATT LOAD WITH $\lambda = 0.07$

Exp Data Set	*UHR* Optimization		*FER* Optimization	
	Average MW Change	%Change	Average MW Change	%Change
1	4.77	7.79%	1.07	1.85%
2	8.38	13.39%	1.58	1.83%
3	2.08	2.45%	2.33	2.29%
4	3.9	4.72%	1.04	1.11%
5	3.99	5.99%	1.14	1.55%
6	2.11	3.25%	1.24	1.42%
7	3.33	4.52%	1.69	2.10%
8	1.41	1.51%	1.02	1.12%
9	0.995	1.09%	0.79	0.83%
10	5.5	5.96%	0.65	0.55%

classified based on the load intervals. Thus, each data set to be clustered had more homogeneous data points in terms of the load distribution. However, for the UHR optimization procedure, data sets were classified based on the UHR intervals. Each data set had much variability of load distribution.

V. CONCLUSION

In this paper, a clustering approach to modeling a boiler–turbine system was presented. The clustering-based approach did not use a process model to improve the boiler–turbine performance. It extracted centroids from historical data and improved the boiler–turbine performance by searching the centroid space. The clustering-based approach can be easily implemented, is computationally efficient, and easily adapts to the process changes by online learning from the current process data. It is a good candidate for real-time performance optimization of boiler–turbine systems. All centroids were extracted from the historical data. If suitable data was not available, the clustering-based approach could not generate better solutions than the existing ones. Experiments showed that sufficient data usually exist to significantly improve process performance.

The UHR and FER optimization models are of particular interest as both have the potential to improve the efficiency of the boiler–turbine system. The UHR model did not directly optimize the fuel input to the boiler. Compared with the FER model, the UHR model resulted in larger load changes. The UHR is a widely used metric in the power industry. As the UHR is vulnerable to different types of errors, the UHR model did not fully guarantee reduced fuel consumption. The industrial case study showed that optimizing the UHR could even increase fuel consumption. The FER model provides a viable alternative for reducing fuel consumption.

REFERENCES

[1] M. A. Abido, "Multiobjective evolutionary algorithms for electric power dispatch," *IEEE Trans. Evol. Comput.*, vol. 10, no. 3, pp. 315–329, Jun. 2006.

[2] K. J. Åström and R. D. Bell, "Drum-boiler Dynamics," *Automatica*, vol. 36, no. 3, pp. 363–378, 2000.

[3] R. D. Bell and K. J. Åström, "Dynamic models for boiler–turbine–alternator units: Data logs and parameter estimation for a 160 MW unit," Lund Inst. Technol., Lund, Sweden, Tech. Rep. TFRT-3192, 1987.

[4] W. L. Brogan, *Modern Control Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[5] A. Ben-Abdennour and K. Y. Lee, "An autonomous control system for boiler–turbine units," *IEEE Trans. Energy Convers.*, vol. 11, no. 2, pp. 401–406, Jun. 1996.

[6] B. W. Bequette, *Process Control: Modeling, Design and Simulation*. Upper Saddle River, NJ: Pearson, 2003.

[7] A. Burns, A. Kusiak, and T. Letsche, "Mining transformed data sets," in *Knowledge-Based Intelligent Information and Engineering Systems*, Vol. I, R. Khosla, R. J. Howlett, and L. C. Jain, Eds. Heidelberg, Germany: Springer, 2004, pp. 148–154.

[8] F. P. de Mello, "Boiler models for system dynamic performance studies," *IEEE Trans. Power Syst.*, vol. 6, no. 1, pp. 66–74, Feb. 1991.

[9] J. Espinosa, J. Vandewalle, and V. Wertz, *Fuzzy Logic, Identification and Predictive Control*. London, U.K.: Springer-Verlag, 2005.

[10] S. Farhad, M. Younessi-Sinaki, and M. Saffar-Avval, "Energy saving in operating steam power plants based on ASME performance test code," in *Proc. 2005 ASME Power Conf.*, Chicago, IL, pp. 1381–1385.

[11] K. M. Hangos, R. Lakner, and M. Gerzson, *Intelligent Control Systems: An Introduction with Examples*. Amsterdam, The Netherlands: Kluwer, 2001.

[12] H. K. Khalil, *Nonlinear Systems*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[13] V. I. Kuprianov, "Applications of a cost-based method of excess air optimization for the improvement of thermal efficiency and environmental performance of steam boilers," *Renew. Sustainable Energy Rev.*, vol. 9, no. 5, pp. 474–498, 2005.

[14] A. Kusiak, A. Burns, and F. Milster, "Optimizing combustion efficiency of a circulating fluidized boiler: A data mining approach," *Int. J. Knowl.-Based Intell. Eng. Syst.*, vol. 9, no. 4, pp. 263–274, 2005.

[15] A. Kusiak and S. Shah, "A data-mining-based system for prediction of water chemistry faults," *IEEE Trans. Ind. Electron.*, vol. 53, no. 2, pp. 593–603, Apr. 2006.

[16] A. Kusiak and Z. Song, "Combustion efficiency optimization and virtual testing: A data-mining approach," *IEEE Trans. Ind. Inf.*, vol. 2, no. 3, pp. 176–184, Aug. 2006.

[17] S. Y. Li, H. B. Liu, W. J. Cai, Y. C. Soh, and L. H. Xie, "A new coordinated control strategy for boiler–turbine system of coal-fired power plant," *IEEE Trans. Control Syst. Technol.*, vol. 13, no. 6, pp. 943–954, Nov. 2005.

[18] D. Li, T. Chen, H. J. Marquez, and R. K. Gooden, "Life extending control of boiler–turbine systems via model predictive methods," *Control Eng. Pract.*, vol. 14, no. 4, pp. 319–326, 2006.

[19] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Stat. Probab.*, Berkeley, CA: Univ. of California Press, vol. 1, 1967, pp. 281–297.

[20] U. C. Moon and K. Y. Lee, "A boiler–turbine system control using a fuzzy auto-regressive moving average (FARMA) model," *IEEE Trans. Energy Convers.*, vol. 18, no. 1, pp. 142–148, Mar. 2003.

[21] M. Norggaard, O. Ravn, N. K. Poulsen, and L. K. Hansen, *Neural Networks for Modeling and Control of Dynamic Systems*. London, U.K.: Springer-Verlag, 2000.

[22] G. Pellegrinetti and J. Bentsman, "Nonlinear control oriented boiler modeling–A benchmark problem for controller design," *IEEE Trans. Control Syst. Technol.*, vol. 4, no. 1, pp. 57–64, Jan. 1996.

[23] G. Poncia and S. Bittanti, "Multivariable model predictive control of a thermal power plant with built-in classical regulation," *Int. J. Control*, vol. 74, no. 11, pp. 1118–1130, 2001.

[24] J. A. Rossiter, *Model-Based Predictive Control: A Practical Approach*. New York: CRC, 2003.

[25] G. G. Rajan, *Optimizing Energy Efficiencies in Industry*. New York: McGraw-Hill, 2003.

[26] D. E. Seborg, T. F. Edgar, and D. A. Mellichamp, *Process Dynamics and Control*, 2nd ed. New York: Wiley, 2003.

[27] H. Taplin, *Combustion Efficiency Tables*. Lilburn, GA: Fairmont, 1991.

[28] W. Tan, H. J. Marquez, T. W. Chen, and J. H. Liu, "Analysis and control of a nonlinear boiler–turbine unit," *J. Process Control*, vol. 15, no. 8, pp. 883–891, 2005.

[29] P. N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Reading, MA: Pearson Education/Addison Wesley, 2006.

[30] I. H. Witten and F. Eibe, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed. San Francisco, CA: Morgan Kaufmann, 2005.

[31] D. Yu and Z. Xu, "Nonlinear coordinated control of drum boiler power unit based on feedback linearization," *IEEE Trans. Energy Convers.*, vol. 20, no. 1, pp. 204–210, Mar. 2005.

Andrew Kusiak (M'90) received the B.S. degree in precision engineering and the M.S. degree in mechanical engineering from Warsaw University of Technology, Warsaw, Poland, in 1972 and 1974, respectively, and the Ph.D. degree in operations research from the Polish Academy of Sciences, Warsaw, in 1979.

He is currently a Professor at the Intelligent Systems Laboratory, Department of Mechanical and Industrial Engineering, The University of Iowa, Iowa City. He is the author or coauthor of several technical papers published in journals sponsored by professional societies, such as the Association for the Advancement of Artificial Intelligence, the American Society of Mechanical Engineers, and the Institute of Industrial Engineers. He is the author of numerous books. He speaks frequently at international meetings, conducts professional seminars, and consults for industrial corporations. His current research interests include applications of computational intelligence in automation, energy, manufacturing, product development, and healthcare.

Prof. Kusiak is an IIE Fellow. He has served on the editorial boards of over 40 journals and is the Editor-in-Chief of the *Journal of Intelligent Manufacturing*.

Zhe Song received the B.S. and M.S. degrees from the China University of Petroleum, DongYing City, China, in 2001 and 2004, respectively. He is currently working toward the Ph.D. degree in industrial engineering at the Intelligent Systems Laboratory, The University of Iowa, Iowa City.

His current research interests include data mining, computational intelligence, optimization, and their applications in process and discrete manufacturing industry.