ARTICLE IN PRESS

Energy xxx (2014) 1–11

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Applying analytics in the energy industry: A case study of heat rate and opacity prediction in a coal-fired power plant

Jongsawas Chongwatpol*, Thanrawee Phurithititanapong

NIDA Business School, National Institute of Development Administration, 118 Seri Thai Road, Bangkapi, Bangkok 10240, Thailand

A R T I C L E I N F O

Article history:
Received 23 April 2014
Received in revised form
31 July 2014
Accepted 1 August 2014
Available online xxx

Keywords:
Analytics
Coal-fired
Power plant
Data mining
Heat rate
Opacity

A B S T R A C T

Power producers are looking for ways not only to improve the efficiency of power plant assets but also to alleviate concerns about the environmental impact of power generation without compromising their market competitiveness. To meet this challenge, advanced analytics through data mining approaches has been applied to predict and explain the leading causes of variation in heat rates and the opacity of the flue gas exhaust emissions, which are key factors in measuring the overall efficiency of the power plant. The results from a case study from a coal-fired power plant in Thailand demonstrate the practical validity of our approaches. The key finding contributes to identifying parameters that are the key determinants of excessive heat rates and opacity levels. Thus, corrective and preventive actions related to those parameters can be regularly evaluated and monitored.

1. Introduction

Coal-fired power producers are looking for ways not only to improve the efficiency of power plant assets but also to address concerns about the environmental impact of power generation without compromising their market competitiveness. Many technologies and advanced control strategies have been introduced to meet this challenge to further enhance plant efficiency. For instance, power plants utilizing USC (ultra-super-critical) coal-fired plant technology indicate energy efficiency up to 46% and reduction of CO_2 emissions [1]. Campbell (2013) reports 0.15%–0.84% efficiency increase when combustion control optimization is implemented to adjust coal and airflow to optimize steam in the turbine and generator [2]. Emerging environmental protection technologies such as dedusting, desulfurization, or denitration equipments are being installed at plants to alleviate the negative impact on the environment [3].

The main goal of power plants is to ensure an uninterrupted energy supply with lower levels of toxic substance emissions to the environment. However, coal-fired power plants are usually associated with different types of coal characteristics and with sophisticated dynamic-nonlinear behavior involving the levels of equipment, sub-processes or sub-systems. Thus, it is difficult to understand the energy consumption behavior or identify factors that affect combustion efficiency through traditional descriptive statistics, conventional regression analysis, or even optimization methods that are limited to a small number of parameters.

Advanced analytics through data mining approaches has been applied successfully in many fields across all industries, from customer relationship management, behavioral profiling, healthcare, and genome analysis to supply chains [4–6]. Data mining is a generic term encompassing a wide variety of techniques to identify intrinsic patterns in data and interpret them into useful information within a particular context. Data mining incorporates both statistical and analytical techniques to effectively and efficiently understand and use data [7,8]. Particularly for power plant communities, many studies address the benefits of data mining techniques to model energy consumption, to monitor boiler efficiency, and to monitor the emission of CO_2, SO_x, or NO_x [1,9–13].

The idea of applying analytics for improving power plant performance is not new; however, this study offers alternatives for improving the efficiency of power plants and reducing the environmental impact of power generation by using analytics techniques such as stepwise regression, decision trees, and neural networks. Our study differs from the others in its focus on (1) heat rate - the amount of energy input needed to produce the electrical energy output and (2) the opacity of the emissions - the opacity of coal-fired power plants' stack emissions. We believe that reducing

* Corresponding author.
E-mail addresses: jong_tn@hotmail.com, jongsawas.c@ics.nida.ac.th (J. Chongwatpol).

http://dx.doi.org/10.1016/j.energy.2014.08.002
0360-5442/

Please cite this article in press as: Chongwatpol J, Phurithititanapong T, Applying analytics in the energy industry: A case study of heat rate and opacity prediction in a coal-fired power plant, Energy (2014), http://dx.doi.org/10.1016/j.energy.2014.08.002

the heat rate in the combustion process will enable power plants to generate the same amount of electricity with less fuel, resulting in decreasing stack emission opacity and emissions of toxic pollutants from burning fossil fuels.

Thus, in exploring the potential of analytics in the power plant industry, the following research questions arise: “What are the most important factors that influence the overall efficiency and heat rate of power plants?” and “What factors have a great impact on stack emission opacity when the quality of the coal in the combustion process changes?” A case study from a coal-fired power plant in Thailand has been conducted to answer these research questions. Unlike many studies that concern only the operational parameters related to the boiler, turbine, or combustion processes, this study analyzes factors associated with the fuel properties—coal qualities and the chemical composition of coal—as well. The main purpose is to examine whether more complex analytical models using several data mining methodologies and algorithms can better predict and explain the variation of heat rate and stack opacity emission.

Although this study originated within the context of a specific coal-fired power plant, the situation depicted at the plant is quite typical of what is observed in most power plants. The case presented in this study can provide fruitful avenues for exploring the application of analytics in the energy industry, contributing to both knowledge and practice. The rest of this study is organized as follows. After a brief literature review in Section 2, Section 3 presents the research methodology of this study. Heat rate and stack emission opacity prediction are discussed in Sections 4 and 5, respectively. The last section contains the conclusion and directions for future research.

2. Brief literature review

Improving power plant performance is an ongoing area of interest to the energy industry. Much of the focus has been on increasing boiler efficiency, improving combustion processes, and reducing the environmental impact of stack emissions. Liu et al. (2013) use fuzzy neural network methods to model the thermal power plant unit in a 1000 MW power plant ultra-super-critical boiler system. Compared to the traditional recursive least squares (RLSs) method, the proposed methods indicate superior capability to predict electric power output, steam pressure, and separator outlet steam temperature [1]. Xu et al. (2011) propose a method to assess the multi-objective performances of power plants. By employing GRA (grey relational analysis) and the AHP (analytic hierarchy process), coal-fired power plants can achieve multi-optimization of their thermal, environmental, and economic performance simultaneously [3]. Sueyoshi and Goto (2013) propose using DEA (data environment analysis) to analyze the environmental performance of coal-fired power plants by examining the legal validity of the U.S. Clean Air Act [14]. Many other analytical techniques, including a Gaussian generalized function criterion, ANP (analytical network process), and life cycle assessment are also used in decision-making mechanisms to determine electrical energy needs considering the economical, technological, social, and life quality aspects [15–18]. Moradi and Bakhtiari-Nejad (2011) propose a LTI (linear time invariant) model to improve boiler unit performance using an optimum robust minimum-order observer. By controlling the critical factors such as drum pressure, steam temperature, or water level of drum, power plants can benefit from lower design costs and higher accuracy in demand estimation resulting in more efficient operational performance [19].

In analytics specifically, a few studies address the application of data mining techniques in the power plant industry. For instance, Song and Kusuak (2007) describe how to integrate data mining approaches with constraint-based control of boiler efficiency to reduce variance in the megawatt load in the combustion process [11]. Schlechtingen et al. (2013) use data mining approaches to monitor the wind turbine power curve. The main goal is to predict the performance of the turbines to ensure that electrical power output meets the demand and to detect any deviations that cause financial loss in the operations [10]. Wang et al. (2012) apply SVR (support vector regression) and GA (genetic algorithm) methods to model the energy consumption and to optimize the operational parameters such as steam temperature, reheat steam pressure, gas temperature, and oxygen content in gas in large coal-fired power plants [12]. Zheng et al. (2009) build predictive models using various data mining techniques to predict NO_x emissions at a coal-fired utility boiler. Combined with other optimization algorithms through ACO (ant colony optimization), GA (genetic algorithm), and PSO (particle swarm optimization), the results indicate the optimum operating parameters for effectively reducing NO_x emissions [13]. Kusuak and Song (2006) apply cluster analysis, decision tree, and neural network models to improve the combustion efficiency of a coal-fired boiler. A K-Means clustering algorithm is used to classify combustion efficiency into subgroups based on boiler characteristics such as the age of the boiler, heat ball location, and boiler installation. Predictive modeling is then developed to predict boiler efficiency and the results are compared to those from a traditional linear regression model [9]. Other data mining related techniques such as fuzzy association rule mining, sequence analysis, principal component analysis, and partial least squares techniques are also applied to monitor power plant performance [20,21].

These studies are just a few examples of efforts to improve power plant performance considering both operational and environmental aspects of power generation. This study focuses on identifying factors associated with operational parameters and the quality of coal that influence the variation of heat rate and opacity observed in the power generation. The research methodology and problem scenarios are presented in the next section.

3. Research methodology

We follow the CRISP-DM (Cross Industry Standard Process for Data Mining) methodology as a guideline in structuring the data mining project for diagnosing defects in the energy industry. CRISP-DM breaks down this data mining project into six phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment [22,23].

3.1. Business understanding

A case study of a coal-fired power plant in Thailand has been conducted to explore the benefit of applying analytical methods in the energy industry. This coal-fired power plant provides the electricity to the EGAT (electricity generating authority of Thailand) to meet the country's rapidly increasing demand for electricity under the IPP (independent power producers agreement), a long-term PPA (power purchase agreement). Many emerging technologies for improving the quality of bituminous coal, FGD (flue gas desulfurization), ESP (electrostatic precipitators), and NO_x (low nitrogen oxide) burners, are implemented at the power plants to improve operational efficiency, to reduce the technical risks, and to ensure compliance with environmental regulations.

Fig. 1 presents the flow of the electricity generation process in a coal-fired power plant. First, coal is conveyed to the coal bunker and is crushed by the mill into a fine powder, increasing the surface area to speed up the burning process. The powdered coal is then fed into the combustion chamber of a boiler, where it is burned at a high temperature. The output of this process is steam, which is then