The current issue and full text archive of this journal is available on Emerald Insight at:
www.emeraldinsight.com/0263-5577.htm

Managing big data in coal-fired power plants: a business intelligence framework

Jongsawas Chongwatpol
NIDA Business School, National Institute of Development Administration, Bangkok, Thailand

Received 23 November 2015
Revised 1 March 2016
Accepted 22 April 2016

Abstract

Purpose – Many power producers are looking for ways to develop smarter energy capabilities to tackle challenges in the sophisticated, non-linear dynamic processes due to the complicated operating conditions. One prominent strategy is to deploy advanced intelligence systems and analytics to monitor key performance indicators, capture insights about the behavior of the electricity generation processes, and identify factors affecting combustion efficiency. Thus, the purpose of this paper is to outline a way to incorporate a business intelligence framework into existing coal-fired power plant data to transform the data into insights and deliver analytical solutions to power producers.

Design/methodology/approach – The proposed ten-step business intelligence framework combines the architectures of database management, business analytics, business performance management, and data visualization to manage existing enterprise data in a coal-fired power plant.

Findings – The results of this study provide plant-wide signals of any unusual operational and coal-quality factors that impact the level of NO_x and consequently explain and predict the leading causes of variation in the emission of NO_x in the combustion process.

Research limitations/implications – Once the framework is integrated into the power generation process, it is important to ensure that the top management and the data analysts at the plants have the same perceptions of the benefits of big data and analytics in the long run and continue to provide support and awareness of the use of business intelligence technology and infrastructure in operational decision making.

Practical implications – The key finding of this study helps the power plant prioritize the important factors associated with the emission of NO_x; closer attention to those factors can be promptly initiated in order to improve the performance of the plant.

Originality/value – The use of big data is not just about implementing new technologies to store and manage bigger databases but rather about extracting value and creating insights from large volumes of data. The challenge is to strategically and operationally reconsider the entire process not only to prepare, integrate, and manage big data but also to make proper decisions as to which data to select for the analysis and how to apply analytical techniques to create value from the data that is in line with the strategic direction of the enterprise. This study seeks to fill this gap by outlining how to implement the proposed business intelligence framework to provide plant-wide signals of any unusual operational and coal-quality factors that impact the level of NO_x and to explain and predict the leading causes of variation in the emission of NO_x in the combustion process.

Keywords Big data, Business analytics, Business intelligence, Coal-fired, NO_x

Paper type Research paper

1. Introduction

One of the most important tasks for coal-fired power producers is not only to ensure an uninterrupted energy supply but also to address concerns about the environmental impact of power generation. Power producers are facing economic, regulatory, and competitive challenges to deliver reliable and affordable energy while reducing the cost of operations. Additionally, the use of coal for electric power generation has led to a significant increase in the emissions of toxic substances such as CO_2, NO_x, and SO_x (Chatzimouratidis and Pilavachi, 2012; Chongwatpol and Phurithititanapong, 2014;

Industrial Management & Data Systems
Vol. 116 No. 8, 2016
pp. 1779-1799
© Emerald Group Publishing Limited
0263-5577
DOI 10.1108/IMDS-11-2015-0473

Venkatesh *et al.*, 2012; Waldner *et al.*, 2013). Thus, significant concerns have been raised about the environmental impact on climate and air quality of toxic emissions associated with the power generation process.

Many technologies and advanced control systems have been developed and implemented to improve the efficiency of power plant assets while lowering the level of toxic substance emissions to the environment. For instance, a recent study on the implementation of a low temperature economizer to improve the adsorption and oxidation capacity of the fly ash in coal-fired power plants shows that efficiency in the removal of total and elemental mercury increased by 42.87 and 18.85 percent, respectively (Zhou *et al.*, 2015). Coal-fired power plants utilizing the ultra-super critical technology report energy efficiency up to 46 percent and reduction of CO_2 emissions (Liu *et al.*, 2013) and Campbell (2013) finds a 0.15-0.84 percent efficiency increase when combustion control optimization is implemented to adjust coal and air flow to optimize steam in the turbine and generator (Campbell, 2013). Additionally, emerging environmental protection technologies such as denitration, desulfurization, and dedusting equipment are installed at the plants to alleviate the negative impact on the environment (Xu *et al.*, 2011).

However, coal-fired power plants involve very sophisticated, non-linear dynamic processes because of the complicated operating conditions with the different levels of equipment and sub-processes, the huge boiler architecture, and different coal characteristics. Even with advanced technologies implemented at the energy site, it is still very difficult to understand the energy consumption behavior and to effectively control the operational parameters in order to minimize the energy needed to produce the electrical energy output and the environmental impact on climate and air quality of toxic emissions associated with the power generation process.

Thus, not surprisingly, many power producers are looking for ways to develop smarter energy capabilities to tackle those challenges and one prominent strategy is to deploy advanced intelligence systems and analytics to monitor key performance indicators, capture insights about the behavior of the electricity generation processes, and identify factors affecting combustion efficiency. Fortunately, the existing IT infrastructure has the capability to collect vast amounts of data not only from the boiler, turbine, and combustion processes, but from the fuel properties such as coal qualities and the chemical composition of coal. However, even with having access to vast troves of data and content, also called big data, many enterprises still do not know how to turn the data into actionable strategies for better business decisions. Additionally, the traditional descriptive statistics, conventional regression analysis, or even optimization methods currently employed at the plant site are not designed to handle and support big data and analytics.

In other words, big data has attracted the attention of both business leaders and academics. However, the sudden rise of big data has left them unprepared to use it effectively, especially in the energy industry. Business intelligence is introduced as one of the application areas of growing importance in supporting business decisions. The concept of business intelligence opens the door to numerous opportunities to deal with big data not only by integrating platforms to handle more complex, unstructured data with emerging data sources, but by emphasizing the analytical process enabled by big data as well (Phillips-Wren *et al.*, 2015). Although a handful of studies have employed business intelligence in exploring and analyzing large amounts of complex data, the implications of the business intelligence schema in the energy industry are yet to be well developed and established. This study seeks to fill this gap and outlines a way to incorporate a business intelligence framework into existing coal-fired power plant data to transform the data into insights and deliver analytical solutions to power producers. A case study from a coal-fired power plant in Thailand has been conducted to explore the implications of business intelligence and business analytics in the energy industry. The power plant currently seeks to use data generated by the installed stack emission monitoring platforms and the top priority is compliance with air pollutant emission standards. Since NO_x is one of the key contributors to Thailand's pollution, the main focus of this study is on corrective actions and preventive maintenance related to factors that have a great impact on NO_x emission.

Although this study originated within the context of a specific coal-fired power plant, the challenges of managing the distributed energy resources and enhancing environmental protections at the plant are quite typical of what is observed in most power plants. The business intelligence framework presented in this study can provide fruitful avenues for exploring the implications of big data and analytics in the energy industry, contributing to both knowledge and practice. In particular, this framework may encourage practitioners with a standard-reactive data management mindset to proactively incorporate massive data volumes in their analyses. The rest of this study is organized as follows. After a brief literature review in Section 2, Section 3 presents the proposed business intelligence framework of this study. This framework is applied to the electricity generation process in a coal-fired power plant, along with the current problem found in the combustion process in Section 4. Results and discussion on implementing the business intelligence framework to analyze, monitor, and manage NO_x emissions are presented, followed by lesson learned in Section 5, managerial and practical contributions in Section 6, and the conclusions in Section 7.

2. Brief literature review

Big data is defined as "a relative term describing a situation where the volume, velocity and variety of data exceed an organization's storage or compute capacity for accurate and timely decision making" (SAS, 2012; Garter, 2016). Similarly, big data has been defined as "high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation" (Gartner, 2016). Big data has recently become a hot issue and is expanding in many domains, from customer relationship management, retail management, and social media marketing to process improvement and supply chain management (Stubbs, 2014). On the other hand, business intelligence and business analytics have emerged as the techniques, technologies, methodologies, systems, and applications that help business enterprises make timely business decisions by analyzing the vast amount of critical business data collected through various enterprise systems, gaining critical insights from the structured data, and mining unstructured user-generated contents (Chen *et al.*, 2012). Turban *et al.* (2011) define business intelligence as an umbrella term that combines architectures, tools, databases, analytical tools, applications, and methodologies to aid in decision making. The architecture of business intelligence consists of four main components: a data warehouse, business analytics, business performance management (BPM), and a user interface (Turban *et al.*, 2011). Many studies have addressed various aspects of business intelligence and business analytics applications in the era of big data. Mayer-Schonberger and Cukier (2013) and Minelli *et al.* (2012) illustrate the applications of big data analytics by enabling competitive differentiation to discover and solve business problems (Mayer-Schönberger and Cukier, 2013; Minelli *et al.*, 2012).

The key concept of big data and analytics is to use analytical techniques to describe, explore, and analyze large and complex data set that requires advanced data storage, management, and visualization technologies (Chen *et al.*, 2012). Davenport (2006) presents the value of business analytics and how enterprises use analytics to build competitive strategies and extract maximum value from their business processes. Wu *et al.* (2014) propose a big data processing framework consisting of three tiers: big-data mining platform in Tier 1, big-data semantics and application knowledge in Tier 2, and big-data mining algorithms in Tier 3. The study shows that to unleash the full power of big data, high-performance computing platforms, demand-driven aggregation of information sources, data mining and analysis techniques, and user-interest modeling are required (Wu *et al.*, 2014). Gartner identifies four types of analytics capability which help enterprises move from traditional descriptive analytics (what happened?) to advanced diagnostic analytics (why did it happen?), predictive analytics (what will happen?), and prescriptive analytics (how can we make it happen?) (Rivera and Meulen, 2014).

Improving power plant performance and reducing the environmental impact as a result of combustion process through the applications of business intelligence and analytics are still ongoing areas of interest, particularly in the energy industry. A recent survey of the use and performance of big data in the energy industry by PennEnergy and IBM reports that 61.2 percent of power respondents use data for complex modeling, simulations, and analytics but 89.8 percent of those respondents are not satisfied with the modeling performance. Additionally, most respondents agree that the mission-critical activity is to provide data and insights to large numbers of operational decision makers (PennEnergy, 2015). Many studies also report the benefits of data mining techniques such as neural networks, fuzzy logic, decision trees, and regression analysis for monitoring boiler efficiency, modeling energy consumption, and monitoring the emission of CO_2, SO_x, or NO_x (Kusiak and Song, 2006; Liu *et al.*, 2013; Schlechtingen *et al.*, 2013; Song and Kusiak, 2007). Zheng *et al.* (2009) apply various data mining techniques to predict NO_x emissions at a coal-fired utility boiler. Combined with other optimization algorithms through particle swarm optimization (PSO), genetic algorithm (GA), and ant colony optimization (ACO), the results indicate the optimum operating parameters for effectively reducing NO_x emissions (Zheng *et al.*, 2009). Wang *et al.* (2012) apply GA and support vector regression (SVR) methods not only to model the energy consumption but also to optimize the operational parameters such as reheat steam pressure, gas temperature, oxygen content in gas, and steam temperature in large coal-fired power plants (Wang *et al.*, 2012). Xu *et al.* (2011) employ the grey relational analysis (GRA) and the analytic hierarchy process (AHP) method to assess the multi-objective performance of power plants (Xu *et al.*, 2011). The study shows that coal-fired power plants can achieve multioptimization of their economic, thermal, and environmental performance simultaneously.

These studies offer just a few of the implications of implementing the concepts of big data, business intelligence, and business analytics to improve power plant performance considering both the operational and the environmental aspects of power generation. The business intelligence framework and the proposed ten steps to accomplish this big data project are outlined in the next section.

3. Big data and the proposed business intelligence framework

The importance of big data has grown tremendously even in the domain of the coal-fired energy industry. Many energy professionals seek to examine how big data facilitate better and smarter decision making to keep them ahead of the competition by driving efficiencies, reducing the costs of operations, or enhancing environmental protection. The types and volume of data captured at the energy sites continue to grow but the key decision is how to use the massive amount of data to create new insights for energy producers.

In addition to the three Vs (volume, variety, and velocity) of big data mentioned above, other dimensions have been mentioned in the literature: veracity, referring to the imprecise and unreliability inherent in the data (Gokhale, 2011); variability, another important aspect of big data observed in the data flow rates, such as inconsistency of observed data over time with periodic peaks and troughs (SAS, 2012); and value, the intrinsic value obtained by analyzing large volumes of data (ORACLE, 2015). All of these dimensions of big data can be applied in this case study. The data collected from the operational control processes are reported in multiple terabytes or even petabytes, since the power plant runs 24 hours a day, seven days a week; the plant has been collecting the data from multiple sources, including the combustion, steaming, and flue-gas treatment processes and the characteristics of coal; the complexity of the power generation processes has led to an unprecedented rate of data creation; and the important task for the plant is to derive value from the diverse, irregular, and unstable data by identifying the factors that lead to the variation in the production process. Thus, to unlock the potential value of big data, organizations need to turn such data into meaningful information, which is the driving force for implementing analytical tools and techniques for managing high-volume, velocity, variety, veracity, or variability data.

In this study, big data refers to data captured from the power generation processes that are related to the boiler, burner, pulverizer, turbine, coal hopper, pump, condenser, generator, steaming process, and the quality of the coal. Capturing this data generates a total of 120 variables that can be used for developing analytical models and statistical analysis. The database management system is designed to record and store data every minute, meaning that enormous amounts of terabyte data are pouring into the system and the challenge is to understand and take advantage of its value. Figure 1 presents a business intelligence framework that combines the architectures of database management, business analytics, BPM, and data visualization to manage existing enterprise data in a coal-fired power plant. Particularly in the business analytics layer, high-performance analytics though data mining approaches are applied and integrated to guide the analyst to draw knowledge from the data and answer the following types of business analytics questions: what happened, why did it happen, what will happen, and how can we make it happen.

The first step is to define appropriate business goals and the objectives for the analysis in the domain of the energy industry (see Figure 1). For instance, process or data analysts at the coal-fired power plant might be interested in developing predictive models to support management decision making, and the focus may be on identifying the causes of variations in the fuel energy input needed to produce electrical energy output. The second step is to properly set up the key performance measurements so that the analysts know which metrics or indicators are used to measure, monitor, and manage the performance of the electrical generation process. Usually the BPM layer identifies the key performance indicators, such as heat rate (kJ/kWh), turbine temperature, cost/benefit analysis, the level of toxic substances such as NO_x, SO_x, and opacity level in part per million (ppm), which must be linked to the organizational strategies. The next task, in Step 3, is to collect the data relevant to the problem or business domain. For instance, heat rate (kJ/kWh), which is defined as the amount of fuel energy needed to produce

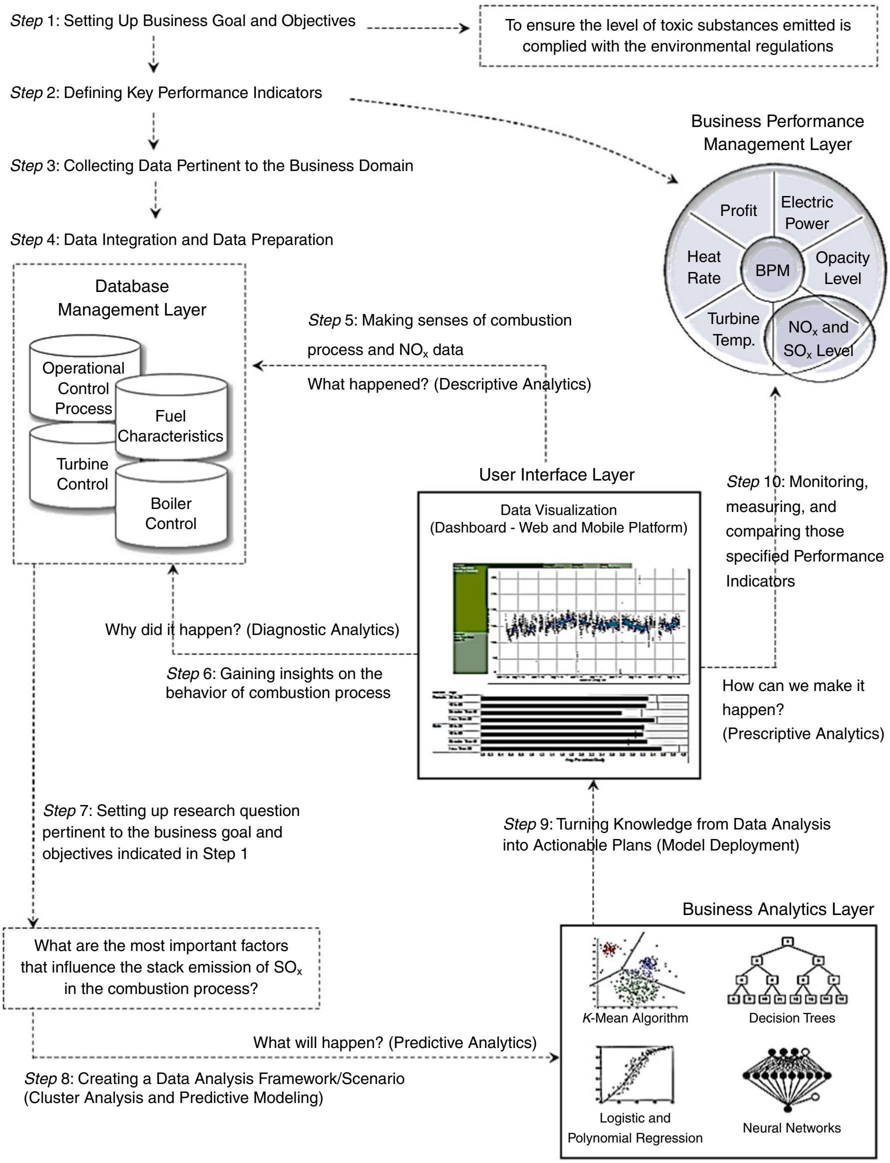

Figure 1.
Business intelligence framework

1 kWh of electrical energy output, is the key factor in measuring the overall efficiency of the power plant. The analysts realize that the lower the heat rate, the more efficiently the plant is operating. Thus, it is very important to identify and collect data on factors related to coal quality and to machines and processes that may affect variations in the heat rate. In Step 4, preparing the data, it may take some time to clean, format, and organize data, especially when the data come from different data sources (the database management layer). Problems of data duplication, data inconsistency, and data errors are usually found while preparing or integrating data before further analysis.

At this point, the data analysts can start making some sense of the performance measures, such as heat rate, combustion process, or NO_x data (Step 5). Dashboard, web, or mobile platforms (the user interface layer) can facilitate such data exploration activities so that the analysts understand how these measures are captured and are able to figure out what is happening to those measures from the historical data. The next task, in Step 6, is for the analysts to look for insights into the data, whether they can discover any patterns based on the preliminary analysis of the targeted performance measures. For instance, they may detect an upward trend or cyclic patterns in the level of NO_x from the previous five months of data. They might form engineering teams to find the root causes for the variation or patterns in such performance measure.

Step 7 is to set up the specific research questions relevant to the business goal and objectives indicated in Step 1. This process helps guide the direction of what methodologies or statistical analysis techniques should be used to answer the research questions. For instance, the research question might be "What are the most important factors that influence the overall efficiency and heat rates of power plants?" Step 8 focusses on creating a data analysis scenario where a variety of data mining techniques (business analytics layer) such as neural networks, logistic and polynomial regression, and decision tree models can be employed. Cluster analysis can also be used to segment combustion data into sub-groups that share similar characteristics before building models to explain and predict the behavior of heat rate, level of NO_x, or other targeted variables. In Step 9, the key findings from data analysis, either from the data exploration in Steps 5 and 6 or from the predictive models in Step 8, are used to provide feedback to the analysts so that actionable plans can be developed to minimize variation in the energy generation processes. Any corrective actions or preventive maintenance on factors affecting the targeted performance measures can then be managed and controlled (Step 10).

Coal-fired power plants

Coal power is one of the established electricity sources providing a vast quantity of inexpensive and reliable power. In Thailand, the demand from both tourism and service sectors is increasing approximately at 10 per cent per year and the government is planning to build another coal-fired facility to meet the growing demand (Williams, 2014). In this case study of a coal-fired power plant in Thailand, we explore the implications of a business intelligence and analytics framework within the energy industry. The coal-fired power plant in this study, "BLCP Power Station," is located in Rayong Province, built with a total generating capacity of 1,434 Megawatts. The plant is under the Independent Power Producer (IPP) program providing electricity to the Electricity Generating Authority of Thailand (EGAT). The plant has implemented many emerging technologies such as electrostatic precipitators (ESP), flue gas desulfurization (FGD), stack emission monitoring platforms, and low NO_x (nitrogen oxide) burners, to improve operational efficiency, reduce the technical risks, and ensure compliance with environmental regulations.

Figure 2 presents the flow of the electricity generation and flue-gas treatment process in a coal-fired power plant. The electricity generation process starts when coal is crushed into a fine powder in the coal bunker to increase the surface area for the burning process. The powdered coal is then burned at a high temperature in the combustion chamber of a boiler. The hot gases and heat energy produced convert the water in tubes lining the boiler into steam, which is used to spin turbines to generate electricity. The byproduct of this combustion process is the flue gas, which is discharged into the air. This flue gas

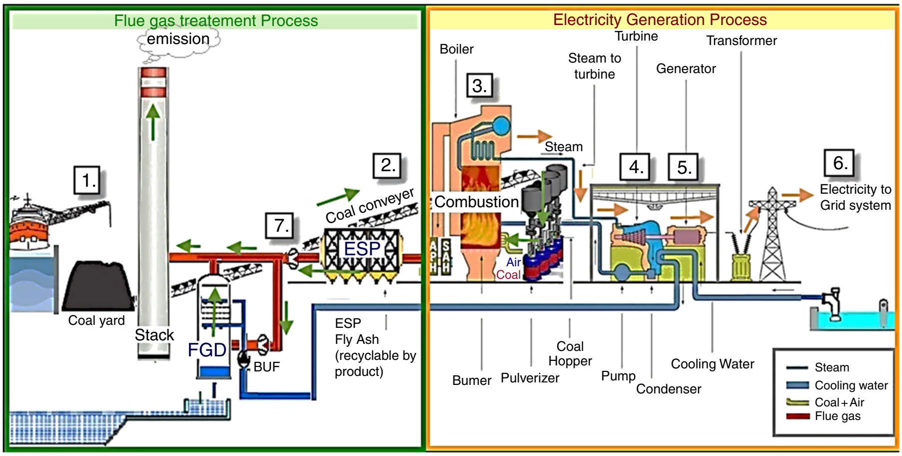

Figure 2.
Coal-fired power plant process

contains water vapor, fly ash, and many toxic substances such as carbon dioxide, NO_x, and SO_x. To comply with the environmental regulations, this flue gas must be treated appropriately by installed protective equipment.

4. Results and discussion

Step 1: setting up business goal and objectives

Although the environmental control systems in this coal-fired power plant have become increasingly standardized, the plant is always looking for ways to proactively monitor plant operations so that the level of toxic substance emissions complies with the mandatory standards for environmental protection. Thus, the goal of this big data project is to promote corrective actions and preventive maintenance on those factors that are the leading causes of variation in toxic-substance levels.

Step 2: defining key performance indicators

The coal-fired power plant is responsible for installing, calibrating, operating, and maintaining a continuous toxic-substance monitoring system in the stack outlet. Currently, the plant observes great variation in NO_x readings, which should not exceed 241 ppm, during the energy production process and the engineering teams conclude that such variation may be caused by differences in the amount of fuel energy used to meet the demand for electricity or by burning coal from different suppliers. Thus, the level of NO_x in the flue gas exhaust emissions is the key performance indicator in measuring the overall efficiency of the power plant in this study.

Step 3: collecting data pertinent to the business domain

The stack emission monitoring platform allows the database management subsystems to record and store data on the level of NO_x every minute. Thus, a total of 2,365,200 records from approximately four-and-a-half years of power plant operation can be explored. The data set that is used to determine the important factors influencing the NO_x level contains a total of 120 variables, which include the data necessary to calculate the total boiler output, fuel heat input, turbine output, and stack emissions as well as the operations control of each major component in the process. See Table I for examples of the variables used in this study.

Variable	Model Role	Measurement Level	Description	
NO_x Level	Target	Numerical	The level of Nox measured in PPM	
Date	ID	Nominal	Date and Time	
Operation Control Process	Input	Numerical	Operation Control Conditions:	
			1 3RY_SH_OUTLET MAIN STEAM PRESS (A)	47 UNIT_GROSS_MW
			2 3RY_SH_OUTLET MAIN STEAM PRESS (B)	48 MAIN_STEAM_FLOW
			3 3RY_RH_OUTLET STEAM PRESS	49 MAIN STEAM PRESS
			4 3RY RH OUTLET STEAM PRESS	50 MAIN STEAM TEMP
			5 3RY SH OUTLET MAIN STEAM TEMP-A	51 RH STEAM TEMP
			6 3RY SH OUTLET MAIN STEAM TEMP-B	52 FDF-A INLET VANE CD FB
			7 HOT REHEAT STEAM TEMP (ST INLET) (RH-2)	53 FDF-B INLET VANE CD FB
			8 HOT REHEAT STEAM TEMP (ST INLET) (LH)	54 IDF-A INLET VANE CD FB
			9 HOT REHEAT STEAM TEMP (ST INLET) (RH-1)	55 IDF-B INLET VANE CD FB
			10 ECO OUTLET GAS O2-A	56 BOOSTER FAN-A BLADE PICH TR SIGNAL
			11 ECO OUTLET GAS O2-B	57 BOOSTER FAN-B BLADE PICH TR SIGNAL
			12 ECO OUTLET GAS O2 (LOW SELECT)	58 BUF-A MOTOR CURRENT
			13 COAL FLOW-A	59 BUF-B MOTOR CURRENT
			14 COAL FLOW-B	60 FDF-A MOTOR CURRENT
			15 COAL FLOW-C	61 FDF-B MOTOR CURRENT
			16 COAL FLOW-D	62 IDF-A MOTOR CURRENT
			17 COAL FLOW-E	63 IDF-B MOTOR CURRENT
			18 COAL FLOW-F	64 AH_A INLET GAS TEMP (2)
			19 BURNER TILT	65 AH_B INLET GAS TEMP (2)
			20 ADJUSTABLE CONTROL DRIVE _AA_A DEM-1 RE	66 AH_A INLET GAS TEMP (1)
			21 ADJUSTABLE CONTROL DRIVE _AA_A DEM-2 RE	67 AH_B INLET GAS TEMP (1)
			22 ADJUSTABLE CONTROL DRIVE _AA_B DEM-2 RI	68 AH_A OUTLET FLUE GAS TEMP (1)
			23 ADJUSTABLE CONTROL DRIVE _AA_C DEM-1 RI	69 AH_A OUTLET FLUE GAS TEMP (2)
			24 ADJUSTABLE CONTROL DRIVE _AA_D DEM-1 RI	70 AH_A OUTLET FLUE GAS TEMP (3)
			25 ADJUSTABLE CONTROL DRIVE _AA_C DEM-2 RI	71 AH_B OUTLET FLUE GAS TEMP (1)
			26 ADJUSTABLE CONTROL DRIVE _AA_D DEM-2 RI	72 AH_B OUTLET FLUE GAS TEMP (2)
			27 AUX DAMPER _U_AA_A DEM-1 RB	73 AH_B OUTLET FLUE GAS TEMP (3)
			28 AUX DAMPER _U_AA_B DEM-1 RB	74 IDF-A INLET FLUE GAS TEMP
			29 AUX DAMPER _L_AA_A DEM-1 RB	75 IDF-B INLET FLUE GAS TEMP
			30 AUX DAMPER _L_AA_B DEM-1 RB	76 FGD GAS DUCT INLET GAS TEMP
			31 AUX DAMPER _U_AA_A DEM-2 RB	77 STACK INLET FLUE GAS TEMPERATURE
			32 AUX DAMPER _U_AA_B DEM-2 RB	78 PULVERIZER-A PRIMARY AIR FLOW
			33 AUX DAMPER _L_AA_A DEM-2 RB	79 PULVERIZER-B PRIMARY AIR FLOW
			34 AUX DAMPER _L_AA_B DEM-2 RB	80 PULVERIZER-C PRIMARY AIR FLOW
			35 AUX DAMPER _U_AA_C DEM-1 RB	81 PULVERIZER-D PRIMARY AIR FLOW
			36 AUX DAMPER _U_AA_D DEM-1 RB	82 PULVERIZER-E PRIMARY AIR FLOW
			37 AUX DAMPER _L_AA_C DEM-1 RB	83 PULVERIZER-F PRIMARY AIR FLOW
			38 AUX DAMPER _L_AA_D DEM-1 RB	84 TOTAL AIR FLOW TON PER HOUR
			39 AUX DAMPER _U_AA_C DEM-2 RB	85 AH_A OUTLET SECONDARY AIR FLOW
			40 AUX DAMPER _U_AA_D DEM-2 RB	86 AH_B OUTLET SECONDARY AIR FLOW
			41 AUX DAMPER _L_AA_C DEM-2 RB	87 FGD BOOSTER FAN-A FLUE GAS FLOW
			42 AUX DAMPER _L_AA_D DEM-2 RB	88 FGD BOOSTER FAN-B FLUE GAS FLOW
			43 AUX DAMPER _U_AA_C DEM-2 RB	89 FGD BOOSTER FAN FLUE GAS FLOW
			44 AUX DAMPER _U_AA_D DEM-2 RB	90 STACK INLET FLUE GAS FLOW
			45 AUX DAMPER _L_AA_C DEM-2 RB	91 FGD BYPASS DUCT FLUE GAS DIFF PRESS
			46 AUX DAMPER _L_AA_D DEM-2 RB	
Fuel Qualities	Input	Numerical	Coal qualities and chemical composition in coal	
			1 Weight average Al2O3 Input to Bolier	14 Weight average K2O Input to Bolier
			2 Weight average Ash Input to Bolier	15 Weight average MgO Input to Bolier
			3 Weight average CaO Input to Bolier	16 Weight average Mn3O4 Input to Bolier
			4 Weight average Carbon Input to Boiler	17 Weight average Na2O Input to Bolier
			5 Weight average Chloride Input to Bolier	18 Weight average Nitrogen Input to Bolier
			6 Weight average ESP K-Factor Input to Bolier	19 Weight average P2O5 Input to Bolier
			7 Weight average F.T. Input to Bolier	20 Weight average SiO2 Input to Bolier
			8 Weight average Fe2O3 Input to Bolier	21 Weight average SO3 Input to Bolier
			9 Weight average Fixed Carbon Input to Bolier	22 Weight average Ti2O Input to Bolier
			10 Weight average fuel ratio Input to Bolier	23 Weight average Total Mositure Input to Boiler
			11 Weight average Total Heat Input to Boiler	24 Weight average Total Sulphur Input to Bolier
			12 Weight average HGI Input to Boiler	25 Weight average Volatile Matter Input to Boiler
			13 Weight average Hydrogen Input to Bolier	26 Ash analysis percent Unburnt Carbon

Table I. An example of variables used in this study

Step 4: data integration and data preparation

Since the data are obtained from the different data acquisition systems in the plant, the first task in this step is to merge data, clean up any errors and inconsistencies, and exclude any outliers (extreme values) in the data. However, one of the limitations in data integration is that data are stored in different formats. For instance, a unit of measurement for many parameters is captured in the hour format and other parameters are recorded in the minute format. Thus, in order to prepare the data for analysis, some parameters are hourly averaged so that the problem of missing values is treated appropriately. The complete data record of 39,397 data readings is used for developing predictive models to accomplish the goal of this study: identifying factors influencing the variation in NO_x levels. Additionally, because of the multiple independent variables, multicollinearity is analyzed and treated appropriately on the basis of correlation analyses, variance inflation factors, and a review from experts in the energy industry.

Step 5: making senses of NO_x data

Exploring the historical data of the coal-fired power plant to understand how NO_x emissions behave can enable us to predict the efficiency of the flue gas treatment process. Figure 3 presents the scatter plots of NO_x emissions from the business intelligence dashboard developed on a mobile/web application using Tableau Software from January 2010 to June 2014. Clearly, a corrective action is needed for NO_x emission that is greater than the regulation limit of 241 ppm (Point 1) and the specification limit of 200 ppm (Points 2 and 3). However, any emission level that deviates greatly from the mean (Points 4, 6, and 7) is worth mentioning. Even though approximately 95 percent of the data samples are below the specification limit, 30 percent of them are in the range of 180-200 ppm, and an upward (Point 5) or downward trend can also be detected occasionally. This incident signals the data analysts to determine what factors drive such variations or patterns.

Step 6: gaining insights into the behavior of NO_x data

The power plant needs to supply electricity on a regular basis to satisfy the demand as a part of the power purchasing agreements with EGAT. As presented in Figure 4, approximately 88 percent of the data samples are in the range of 600-720 MW of electricity generation (see Point A, Figure 4). The distribution of NO_x (Point B) seems greatly dispersed as opposed to the generated electricity, which is quite stable at an average of 640 MW. Exploring deeper through other factors can give the analysts an idea as to whether other factors in the power generation process may be potential causes of increases in the NO_x level. The coal consumption rate (Point C), to meet the demand of electricity generation, is distributed similarly to that found in the electricity generation (Point A). However, the distributions of flue gas temperature and flow rate detected at the stack emission control (Points D and E) interestingly show in different patterns.

Step 7: setting up research question

The plant is responsible for operating a continuous flue-gas monitoring system to measure the level of NO_x but only traditional Excel-based regression analysis and descriptive statistics are used to analyze and monitor the plant performance. The following equations represent the overall reaction equation, currently employed at the plant to estimate the NO_x emission from the combustion process:

$$\begin{aligned} C_aH_bO_cN_dS_e + wH_2O + g(O_2 + 3.76N_2) \rightarrow\ & x_1CO + x_2CO_2 + x_3H_2O \\ & + x_4NO + x_5NO_2 + x_6SO_2 + x_7SO_3 + x_8O_2 + x_9N_2 + x_{10}C, \end{aligned}$$

where $S_{(s)} + O_{2(g)} = SO_{2(g)}$ is the primary reaction form of sulphur oxide; $SO_{2(g)} + 1/2O_{2(g)} = SO_3$ the secondary reaction; $1/2N_{(s)} + 1/2O_{2(g)} = NO_{(g)}$ the primary reaction form of nitrogen oxide; and $NO_{(g)}+1/2O_{2(g)} = NO_{2(g)}$ the secondary reaction at temperature above 1600°C or 2900°F.

The important task in this step is to develop predictive modeling to answer the following research question: "What are the most important factors that influence the stack emission of NO_x in the combustion process?"

Step 8: creating a data analysis framework/scenario

To answer the specified research question, we examine whether more complex analytical models using several data mining methodologies and algorithms can be used

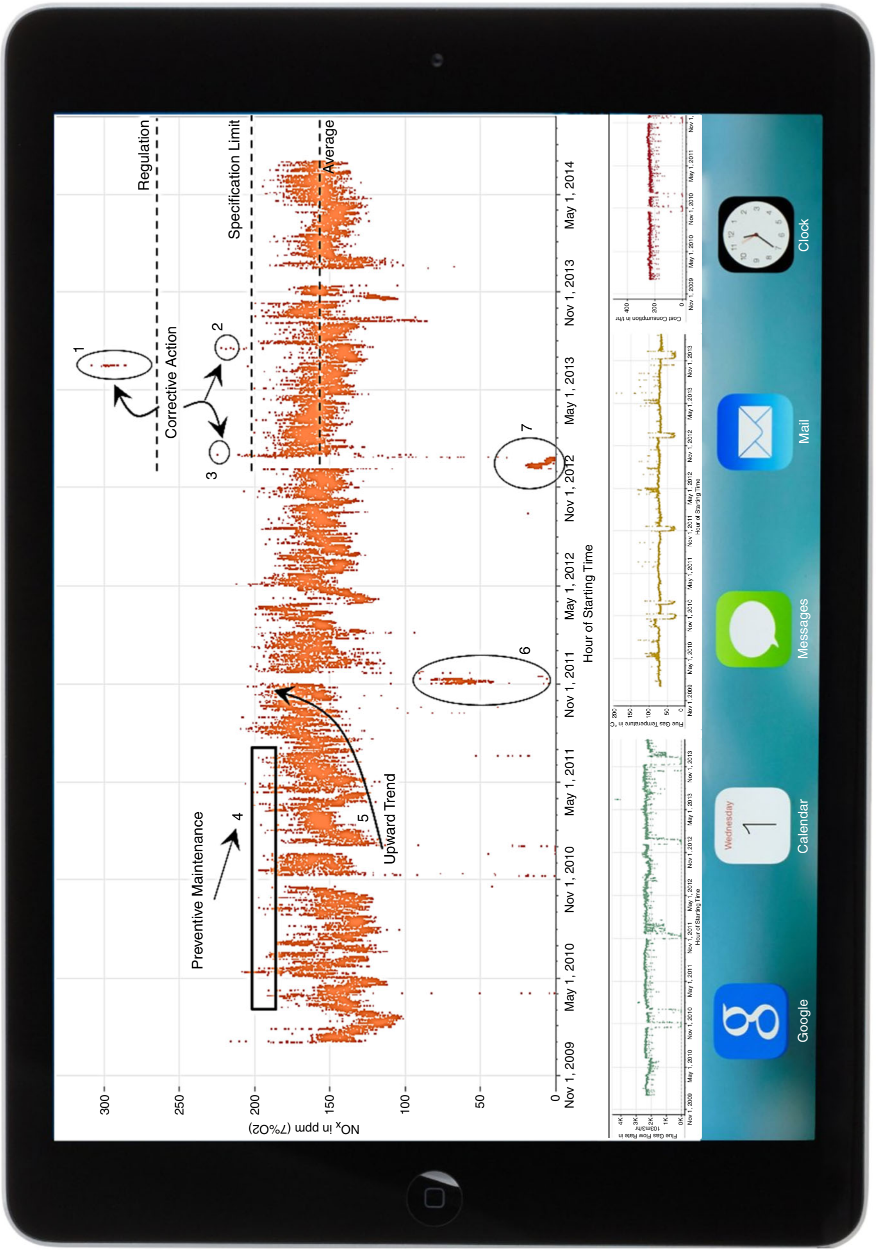

Figure 3.
The level of NO_x
from January 30,
2010 to June 30, 2014

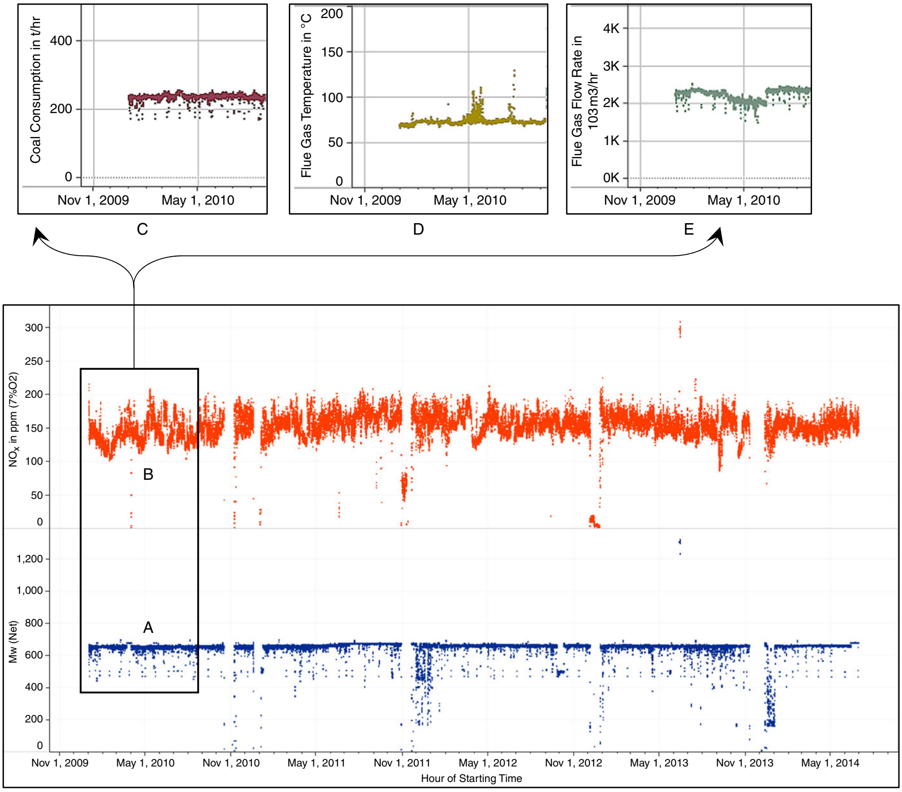

Figure 4.
Energy output (MW) and the level of NO_x

to explain and predict the relationship between the predictor variables (operational control process and the characteristic of coal quality) and the targeted variable (NO_x level). Data mining is a generic term encompassing a wide variety of statistical and analytical techniques to identify intrinsic patterns in data and interpret them into useful information within a particular context (Jackson, 2002; Turban *et al.*, 2011). A variety of data mining techniques is available in predictive analytics. However, since the purpose of this study is to focus on the integration of a business intelligence framework to improve power plant performance rather than proposing new prediction methods, only three popular predictive platforms (stepwise regression, decision tree, and artificial neural network models) are used to analyze data sets with multiple predictor variables (for a technical summary, including both algorithms and their applications for each data mining technique, see Jackson, 2002; Turban *et al.*, 2011). The predictive models are developed using SAS® Enterprise Miner™.

The data set is partitioned into 70 percent for training and 30 percent for validation. The training data set is used to build predictive models while the accuracy of the model fit is tested on the validation data set. Based on the Pearson's correlation results in Figure 5, increasing Motor Current, Unit_Gross_MW, and AH_A_Outlet_Secondary_Air_Flow produces a higher NO_x level. Similarly, we observe a decrease in the NO_x level, when AUX_DAMPER_U_AA_CDEM and Weight_average_Nitrogen_Input increase.

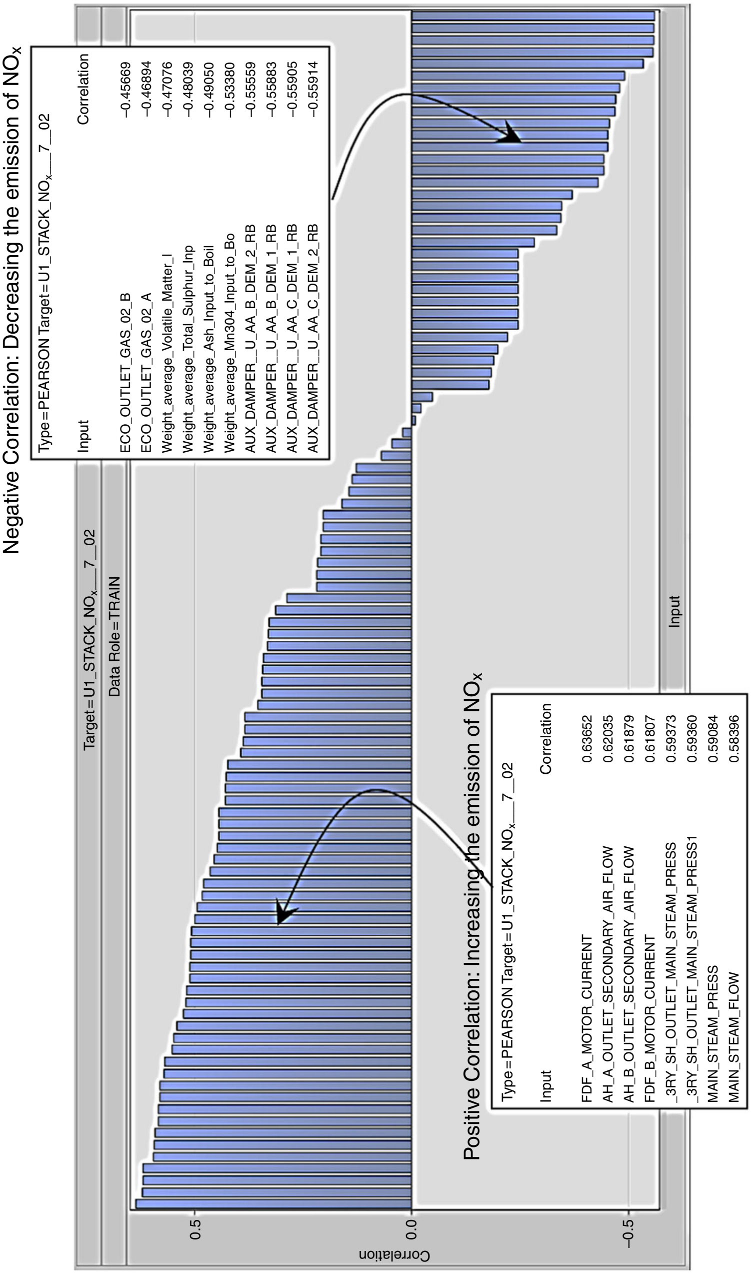

Figure 5.
NO_x and Pearson's
correlation

Step 9: turning knowledge from data analysis into actionable plans (model deployment)
In this step, the important task is to turn the knowledge from the analytical process into actionable plans. Since the coal combustion process is a complex non-linear phenomenon, the data analysts can start from the key finding from the decision tree and stepwise regression models to illustrate rule-based predictions and a regression equation and to understand and predict NO_x emissions (see Figure 6). The results indicate that the most important factors influencing the emission of NO_x are those related to coal properties, which include Mn_3O_4, Na_2O, and Ti_2O, followed by some of the operation control process parameters in the combustion process such as main steam pressure, air heater flue gas inlet and outlet temperatures, economizer outlet oxygen control, and pulverizer. Practically, the plant is now able to focus first on electricity generation with coal having the following properties: $Mn_3O_4 < 0.0254$, $Na_2O > 1.1335$, and $Ti_2O < 1.4452$. Thus, the "quality at the source" strategy can be promoted to ensure that the quality of coal is up to

Rule-based prediction from decision tree model

IF "Eco Outlet Gas O_2-A" is between 2.5746 and 3.68 ***AND*** "Weight Average HGI Input to Boiler" is less than 55.8941 ***AND*** "Aux Damper_L_AA_A Dem-1RM" is less than 62.965 ***AND*** "Weight Average Mn_3O_4 Input to Boiler" is less than 0.0254 ***THEN*** the expected NOx emission is 161.4832

IF "Eco Outlet Gas O2-A" is less than 2.5746 ***AND*** "Weight Average Na_2O Input to Boiler" is greater than 1.1335 ***AND*** "Weight Average Ti_2O Input to Boiler" is less than 1.4452 ***THEN*** the expected NOx emission is 148.2358

Polynomial Regression

Stack NOx Emission = - 251.8 - 0.7626(Adjust_Contorl_Drive_AA_C_D)
+ 0.8761(AH_A_inlet_Gas_Temp) +0.6903(AH_A_outet_Gas_Temp)
- 1.3539(AH_B_inlet_Gas_Temp) – 0.5702(Aux_Damper_U_AA_A_RB)
– 0.8045(Coal_Flow_A) – 0.2535(Coal_Flow_B) – 0.3303(Coal_Flow_C)
+ 15.4966(ECO_Outlet_Gas_O_2_Low_Select)
+ 2.2013 (FDF_A_Motor_Current) + 1.0684(Main_Steam_Press)
+ 0.4897(Pulverizer_A_Primary_Air_flow)
+ 0.2240(Pulverizer_C_Primary_Air_flow)
– 175.6(Weight_average_Mn_3O_4_Input_to_Boi)
– 13.0318(Weight_average_Na_2O_Input_to_Boi)
+ 30.5430(Weight_average_fuel_ratio_Input_to_Boi)

Where "Eco Outlet Gas O_2-A" is the economizer outlet oxygen control
"Aux Damper_L_AA_A Dem-1RM" is the auxiliary power at the Dem-1RM Stage
"AH_A_inlet_Gas_Temp" is the air heater flue gas inlet temperature for unit A
"AH_A_outet_Gas_Temp" is the air heater flue gas outlet temperature for unit A
"AH_B_inlet_Gas_Temp" is the air heater flue gas inlet temperature for unit B
"Main_Steam_Press" is main stream pressure
"Coal Flow" is the coal consumption flow rate for unit A and C
"Pulverizer_A_Primary_Air_flow" is the primary air flow for pulverizer #A
"Pulverizer_C_Primary_Air_flow" is the primary air flow for pulverizer #C
"Mn_3O_4" is Manganese Trioxide
"Na_2O" is Sodium Dioxide
"Ti_2O" is Titanium Dioxide

Figure 6.
Stepwise regression and decision tree models

the standard and those critical coal properties can be routinely monitored. The key findings also signal the power plant to pay more attention to factors in the operation control process, starting from monitoring the air heater flue gas inlet and outlet temperatures. When both inlet and outlet temperatures in the combustion chamber of the boiler increase, the probability is that the emission of NO_x significantly increases. The power plant may also try to improve operations by reducing the main steam pressure and the primary air flow for pulverizers A and C, which have a great impact on increasing the emission of NO_x as well.

Cluster analysis is used to segment combustion data into sub-groups based on the key operation control process parameters: main stream temperature, heat consumption, and flue gas flow pressure. *K*-means cluster analysis (Jackson, 2002; Turban *et al.*, 2011), which is a widely used method for cluster and segmentation analysis, is developed using SAS Enterprise Miner Workstation 13.1. The three-dimensional scatter plot in Figure 7 shows that the largest segment corresponds to the heat consumption rate, averaging between 1,400 and 1,700 kJ/kWhr with very high flue gas flow pressure and main steam temperature > 530°C. The results of cluster analysis show 4 segments. As presented in Figure 8, the distribution of red dots represents the overall distribution of all observed combustion data and the light-blue dots represent the distribution of the combustion data in each cluster. Cluster 1 lacks both flue gas flow pressure and heat consumption and only has a few observations slightly below-average on main steam temperature. Cluster 2 is slightly below-average on all parameters. Cluster 3 is largely categorized by the main steam pressure but has slightly below-average heat consumption and flue gas flow pressure. Lastly, Cluster 4 is close to the distribution centers for all parameters. These clusters signal the instability of the plant's power generation processes, especially when the combustion chamber of the

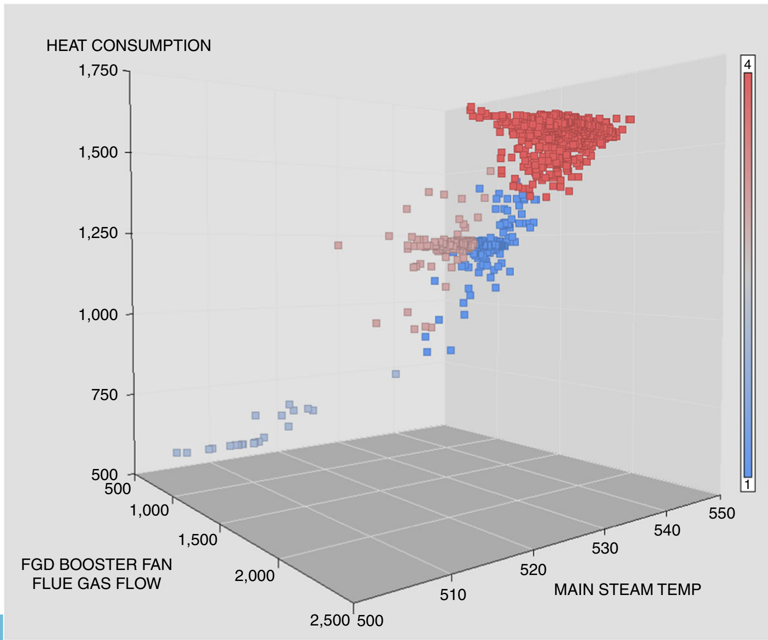

Figure 7.
Three-dimension scatter plot and cluster groups

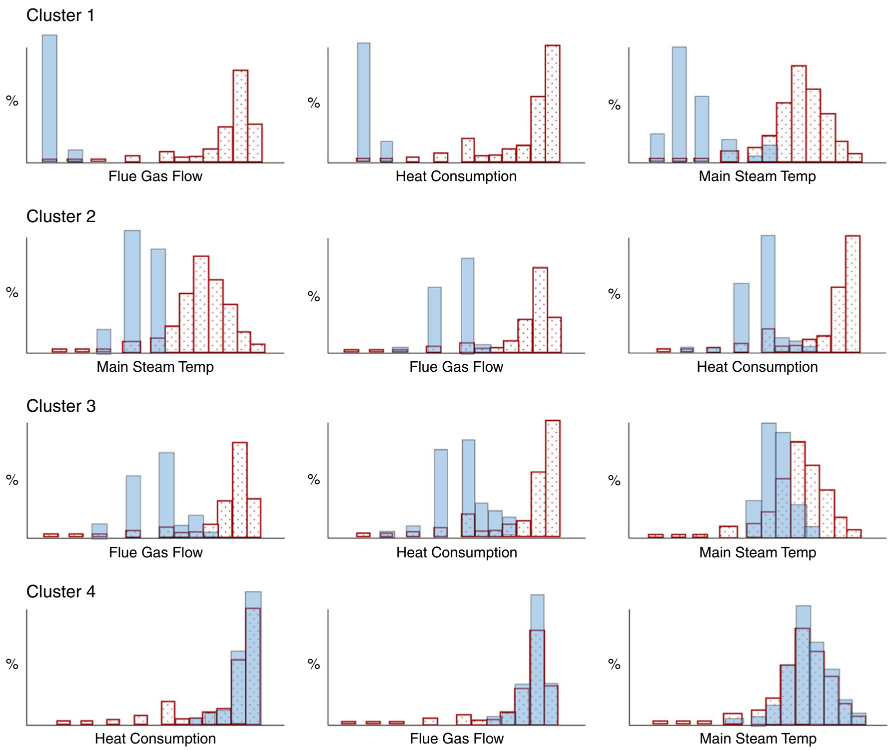

Figure 8.
Cluster profiles

boiler requires significantly higher or lower than average heat rate consumption and the pressure and temperature, which are measured after the water in tubes lining the boiler is converted into steam, vary greatly from the designed threshold. Such irregularity may impact the level of NO_x in the flue gas discharged into the air.

The pictorial representation of the neural network architecture is presented in Figure 9. A neural network, which is a mathematical and computational model, is considered a well-known biologically inspired predictive modeling technique developed for pattern recognition and data classification (Jackson, 2002; Turban *et al.*, 2011). The architecture consists of various input variables related to both operation control process parameters (main steam pressure, main steam temperature, O_2ECOout, and heat rate, etc.) and coal properties (Ti_2O, Na_2O, Mn_3O_4, and CaO, etc.). An example of the equation in Figure 9 for explaining and predicting the level of NO_x (Y) is presented with one hidden layer and three hidden units where each input variable and each hidden unit has its own weight. The transfer function, also called the Tanh Sigmoid function, combines all the inputs into a single value for each hidden unit and then calculates the output value of the NO_x level. The variables with high coefficient estimates in the neural network model, including AH_A_inlet_Gas_Tem, Eco Outlet Gas O_2-A, Pulverizer_A_Primary_Air_flow, Mn_3O_4, and Na_2O, are among the important variables influencing the NO_x level. When the observed values of those parameters deviate significantly from the allowance threshold, promoting preventive maintenance on the machines related to those parameters or encouraging routine inspection of the critical coal properties helps reduce the chance of increasing the NO_x level.

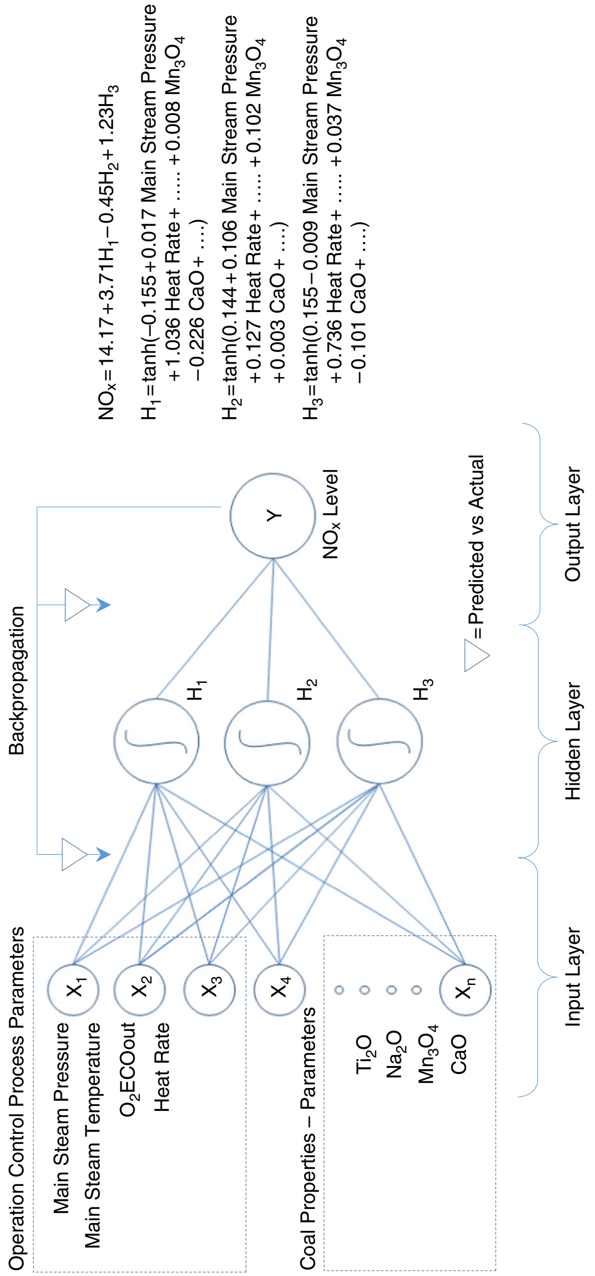

Figure 9.
A neural
network model

Step 10: monitoring, measuring, and comparing the level of NO_x

This finding helps the power plant prioritize the important factors associated with the emission of NO_x; closer attention to those factors can be promptly initiated in order to improve the performance of the plant. Our results, presented in Step 9, show that advanced business analytics through data mining approaches are capable of predicting the emission of NO_x, given sufficient data with the proper input variables. When coal properties and preventive maintenance on key operation control process parameters from the decision tree and stepwise regression models are properly managed, the chances of exceeding NO_x environmental standards can be reduced. The power plant can then use the existing databases along with those analytical models to accurately predict any other toxic substances to monitor plant performance and especially to comply with environmental regulations.

5. Lesson learned

Currently, a single-enterprise data warehouse is used to process large amounts of data while big data analytics software is installed on an ad-hoc basis, a situation which requires substantial effort from both the power plant's IT personnel and data analysts. Engineering teams at the shop-floor operations are the key players responsible for identifying the data that should be acquired and stored and what data are needed in the designated problem domain (Step 3 in the proposed framework). For instance, one problem domain is prediction of the heat rate, which is the amount of fuel energy input needed to produce electrical energy outputs vs the prediction of the toxic level of the flue gas exhaust emissions, which must comply with the mandatory standards for environmental protection.

The NO_x prediction project is considered the best practice and is communicated among departments. Consequently, top management has a roadmap for exploring a variety of available big data technologies such as cloud, Hadoop, and noSQL DB and has pledged more dedicated resources to monitor the defined key performance indicators (Step 2) through the developed dashboards (see both the user interface and BPM layers in the business intelligence framework in Figure 1).

However, the first difficulty for big data initiatives is data management (Step 4) because not all data is stored in the central repository and some data are available in different formats. Thus significant time is required for database administrators to prepare, integrate, and clean the data to ensure its quality for data analytics implementation. The second difficulty is business analytics, where transforming the focus on descriptive analysis of small-structured historical data to a focus on advanced predictive modeling for large amounts of complex data requires time and resources from experts or staff with strong technological skills. Both shop-floor staff and executives are more comfortable with the advanced data visualization (Steps 5 and 6) and often try to skip Steps 7 to 9 in the framework. Thus, it is important to empower the staff to experiment with analytical tools in order to develop analytical and technological skill sets so that they can exploit the full spectrum of big data analytics.

Lastly, although the proposed business intelligence framework and the results of this study help create awareness of big data technology for both engineering staff and executives and the attitude toward big data is positive, still, not all of these people are fully engaged in experimenting with big data initiatives. Only certain assigned groups are proactive and try to collect feedback from peers for improvement, which is why it is important that big data projects be sponsored by the top management and that the value of big data be well recognized.

6. Managerial and practical contributions

While business intelligence has been introduced as a new research paradigm in many disciplines to make use of large quantities of complex data, we have seen very few applications in the energy industry that fully explore its capabilities. Specifically, much attention has been paid to the value that power plants could create through the use of big data, business intelligence, and analytics, which have helped top-performing organizations make successful decisions in promoting both corrective actions and preventive maintenance in the operations.

Additionally, while the need for advanced predictive analytics to analyze ample data for timely decision making in improving operations and complying with the environmental standards is apparent, such analytics are not available in traditional statistical analysis methods such as conventional, Excel-based, regression modeling and descriptive statistics. The uniqueness of this study is the business intelligence framework proposed to gain insight into the behavior of operational control processes and the characteristics of coal that have a great impact on the performance of power plants. The case study presented here is just an example of how to feasibly leverage big data analytics from enormous amounts of combustion data through data visualization and advanced predictive modeling to identify the factors that are the leading causes of variation in toxic-substance (NO_x) levels; thus, closer attention to those factors can be promptly initiated in order to improve the performance of the power plant.

It is hoped that this study provides guidance for the development of business intelligence and analytics in the energy industry. Combined with the proposed framework, such development presents a tremendous potential for immediate large-scale applications for solving similar problems, e.g., the emission of other toxic substances such as CO_2 and SO_x, identification of electricity consumption and energy efficiency, electricity usage in the power distribution system, or even defects in the production processes, and other key performance indicators that show significant variations in the operations.

7. Conclusions

The use of big data is not about implementing new technologies to store and manage bigger databases but rather about extracting value and creating insights from large volumes of data. The challenge is to strategically and operationally reconsider the entire process not only to prepare, integrate, and manage big data but also to make proper decisions as to which data to select for the analysis and how to apply analytical techniques to create value from the data that is in line with the strategic direction of the enterprise. This study seeks to fill this gap by outlining how to implement the proposed business intelligence framework presented in Section 3 to provide plant-wide signals of any unusual operational and coal-quality factors that impact the level of NO_x and to explain and predict the leading causes of variation in the emission of NO_x in the combustion process. The concept of the business intelligence framework presented in this study is not just only for large enterprises but can be applied to smaller enterprises that aim to extract more value from their existing data assets and IT resources. Once the framework is integrated into the power generation process, the next step is to ensure that the top management and the data analysts at the plants have the same perceptions of the benefits of big data and analytics in the long run and continue to provide support and awareness of the use of the business intelligence technology and infrastructure in operational decision making.

References

Campbell, R.J. (2013), "Increasing the efficiency of existing coal-fired power plants", available at: www.fas.org/sgp/crs/misc/R43343.pdf (accessed April 16, 2014).

Chatzimouratidis, A.I. and Pilavachi, P.A. (2012), "Decision support systems for power plants' impact on the living standard", *Energy Conversion and Management*, Vol. 64 No. 1, pp. 182-198.

Chen, H., Chiang, R.H. and Storey, V.C. (2012), "Business intelligence and analytics: from big data to big impact", *MIS Quarterly*, Vol. 36 No. 4, pp. 1165-1188.

Chongwatpol, J. and Phurithititanapong, T. (2014), "Applying analytics in the energy industry: a case study of heat rate and opacity prediction in a coal-fired power plant", *Energy*, Vol. 75 No. 1, pp. 463-473.

Davenport, T.H. (2006), "Competing on analytics", *Harvard Business Review*, Vol. 84 No. 1, pp. 98-107.

Gartner (2016), "Gartner IT glossary: big data", available at: www.gartner.com/it-glossary/big-data (accessed March 1, 2016).

Gokhale, V. (2011), "2011 IBM tech trends report: the clouds are rolling in […] is your business ready?", available at: www.ibm.com/developerworks/community/blogs/ff67b471-79df-4bef-9593-4802def4013d/entry/2011_ibm_tech_trends_report_the_clouds_are_rolling_in_is_your_business_ready5?lang=en (accessed March 1, 2016).

Jackson, J. (2002), "Data mining: a conceptual overview", *Communications of the Association for Information Systems*, Vol. 8 No. 1, pp. 267-296.

Kusiak, A. and Song, Z. (2006), "Combustion efficiency optimization and virtual testing: a data-mining approach", *IEEE Transactions on Industrial Informatics*, Vol. 2 No. 3, pp. 176-184.

Liu, X., Kong, X., Hou, G. and Wang, J. (2013), "Modeling of a 1000MW power plant ultra super-critical boiler system using fuzzy-neural network methods", *Energy Conversion and Management*, Vol. 65 No. 1, pp. 518-527.

Mayer-Schönberger, V. and Cukier, K. (2013), *Big Data: A Revolution that Will Transform How We Live, Work, and Think*, Houghton Mifflin Harcourt, New York, NY.

Minelli, M., Chambers, M. and Dhiraj, A. (2012), *Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses*, John Wiley & Sons, Hoboken, NJ.

ORACLE (2015), "An enterprise architect's guide to big data", available at: www.oracle.com/technetwork/topics/entarch/articles/oea-big-data-guide-1522052.pdf (accessed March 1, 2016).

PennEnergy (2015), "The Use and Performance of Big Data in the Energy Sector", available at: www-01.ibm.com/common/ssi/cgi-bin/ssialias?infotype=SA&subtype=WH&htmlfid=XBW03025USEN#loaded (accessed February 22, 2014).

Phillips-Wren, G., Iyer, L.S., Kulkarni, U. and Ariyachandra, T. (2015), "Business analytics in the context of big data: a roadmap for research", *Communications of the Association for Information Systems*, Vol. 37 No. 1, pp. 448-472.

Rivera and Meulen (2014), "Gartner says advanced analytics is a top business priority", available at: www.gartner.com/newsroom/id/2881218 (accessed February 22, 2014).

SAS (2012), "Big data meets big data analytics", available at: www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/big-data-meets-big-data-analytics-105777.pdf (accessed February 22, 2015).

Schlechtingen, M., Santos, I.F. and Achiche, S. (2013), "Using data-mining approaches for wind turbine power curve monitoring: a comparative study", *IEEE Transactions on Sustainable Energy*, Vol. 4 No. 3, pp. 671-679.

Song, Z. and Kusiak, A. (2007), "Constraint-based control of boiler efficiency: a data-mining approach", *IEEE Transactions on Industrial Informatics*, Vol. 3 No. 1, pp. 73-83.

Stubbs, E. (2014), *Big Data, Big Innovation: Enabling Competitive Differentiation Through Business Analytics*, John Wiley & Sons, Hoboken, NJ.

Turban, E., Sharda, R. and Delen, D. (2011), *Decision Support and Business Intelligence Systems*, Pearson Education, Inc, Upper Saddle River, NJ.

Venkatesh, A., Jaramillo, P., Griffin, W.M. and Matthews, H.S. (2012), "Implications of changing natural gas prices in the United States electricity sector for SO_2, NO_X and life cycle GHG emissions", *Environmental Research Letters*, Vol. 7 No. 3, pp. 1-9.

Waldner, M., Halter, R., Sigg, A., Brosch, B., Gehrmann, H. and Keunecke, M. (2013), "Energy from waste – clean, efficient, renewable: transitions in combustion efficiency and NO_x control", *Waste Management*, Vol. 33 No. 2, pp. 317-326.

Wang, N., Zhang, Y., Zhang, T. and Yang, Y. (2012), "Data mining-based operation optimization of large coal-fired power plants", *AASRI Procedia*, Vol. 3 No. 1, pp. 607-612.

Williams, D. (2014), "Thai Minister says coal-fired plant critical to meet demand", available at: www.powerengineeringint.com/articles/2014/10/thai-minister-says-coal-fired-plant-critical-to-meet-demand.html (accessed August 9, 2016).

Wu, X., Zhu, X., Wu, G.-Q. and Ding, W. (2014), "Data mining with big data", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 26 No. 1, pp. 97-107.

Xu, G., Yang, Y.-p., Lu, S.-y., Li, L. and Song, X. (2011), "Comprehensive evaluation of coal-fired power plants based on grey relational analysis and analytic hierarchy process", *Energy Policy*, Vol. 39 No. 5, pp. 2343-2351.

Zheng, L.-G., Zhou, H., Cen, K.-F. and Wang, C.-L. (2009), "A comparative study of optimization algorithms for low NOx combustion modification at a coal-fired utility boiler", *Expert Systems with Applications*, Vol. 36 No. 2, pp. 2780-2793.

Zhou, Z.-J., Liu, X.-W., Zhao, B., Chen, Z.-G., Shao, H.-Z., Wang, L.-L. and Xu, M.-H. (2015), "Effects of existing energy saving and air pollution control devices on mercury removal in coal-fired power plants", *Fuel Processing Technology*, Vol. 131 No. 1, pp. 99-108.

Corresponding author
Jongsawas Chongwatpol can be contacted at: jongsawas.c@ics.nida.ac.th