An Application of Data Mining in Energy Industry

Jongsawas Chongwatpol
NIDA Business School, National Institute of Development Administration
jongsawas.c@ics.nida.ac.th

Extended Abstract/Poster Papers

1. INTRODUCTION

Many power producers are facing challenges from competitive pressure not only to ensure an uninterrupted and reliable energy supply but also to comply with the environmental regulation. The use of coal for electric power energy has addressed concerns about the environmental impact of the emissions of toxic substances [1-4], resulting in the environmental impact on climate and air quality. Although many technologies and advanced energy generation systems have been introduced and implemented to help controlling the level of toxic substance emissions to the environment, it is still very difficult to understand the energy consumption behavior and how the power generation process have a great impact on the quality of toxic emissions. This study seeks to fill this gap and outline a way to incorporate data mining techniques into existing coal-fired power plant data to monitor plant operations so that the level of toxic substance emissions complies with the mandatory standards for environmental protection.

2. COAL-FIRED POWER PLANT

A case study from a coal-fired power plant in Thailand has been conducted to explore the implication of data mining techniques to promote both corrective action and preventive maintenance at the power plant site. Currently, the plant is observed a great variation of the level of toxic substances and the top priority is to comply with the air pollutant emission standards. The plant is operated to provide electricity to the government with the maximum generating capacity of 1,434 megawatts. The process of the electricity generation starts when the crushed, powder-like, coal is burned at the very high temperature in the combustion chamber of a boiler. The hot gases produced then convert the water in the boiler into steam in order to spin the turbines to generate electricity. On the other hand, the byproduct of this combustion process is the flue gas, which contains many toxic substances such CO_2, NO_x, SO_x, or fly ash. This flue gas must be treated appropriately before being discharged into the air to comply with environmental standard. The power plant has installed, calibrated, and maintain a continuous toxic-substance monitoring system in the two stack outlets. However, the plant observed a great variation in NO_x readings from the flue gas exhaust emissions. Surprisingly, the NO_x readings on both outlets are dispersed differently even though the combustion process remains the same (see Fig 1 and Fig 2). Thus, the research question of this study is set up as follow: What are the important factors that have a great impact on the level of NO_x in the flue gas exhaust emissions on both stack outlets?

Keywords: Data Mining, Energy Industry, Coal-fired, NO_x

3. RESEARCH METHODOLOGY

To answer this research question, this study examine whether more complex analytical data mining techniques such as decision tree, neural network, stepwise polynomial regression, and support vector machine can better explain and predict the variation of NO_x level, which may be caused by differences in the operational control process or suppliers of coal. The dataset is partitioned into 50% for training and validation before applying those data mining techniques to predict and explain the leading causes of variation in NO_x level. The dataset contains variables from boilers, burner, turbine, pulverizer, pump, generator, transformer, and cooling water as well as chemical composition in coal such as Al_2O_3, CaO, Fe_2O_3, Ti_2O, and Mn_3O_4. We follow the CRISP-DM (Cross Industry Standard Process for Data Mining) model as a guideline for diagnosing the variation of NO_x level. CRISPDM breaks down this data mining project into six phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment [5].

4. DISCUSSION AND CONCLUSION

The results of this study not only provide a great signals of any unusual operational and coal-quality factors that influence the level of NO_x but also help explain and predict the leading causes of variation in the emission of NO_x in the combustion process. As presented in Fig.1 and Fig.2, a high NO_x level (above 241 ppm) requires immediate attention and corrective action is needed to comply with the environmental regulation. However, our interest is in the groups with an intermediate NO_x level, a level which is above an average of approximately 140 ppm but has not exceeded the threshold limit. This group can potentially shows signs of problems if the current electricity generation process is not changed or controlled. The stepwise regression and decision tree models show that main steam pressure, main steam temperature, Eco Outlet Gas O_2, Mn_3O_4, Ti_2O, and Na_2O are among factors that have a very high impact on the NO_x level. The plant manager can now minimize emissions of NO_x by promoting preventive maintenance on those key operational control process factors and coal properties, which consequently helps improve the overall performance of the power plant.

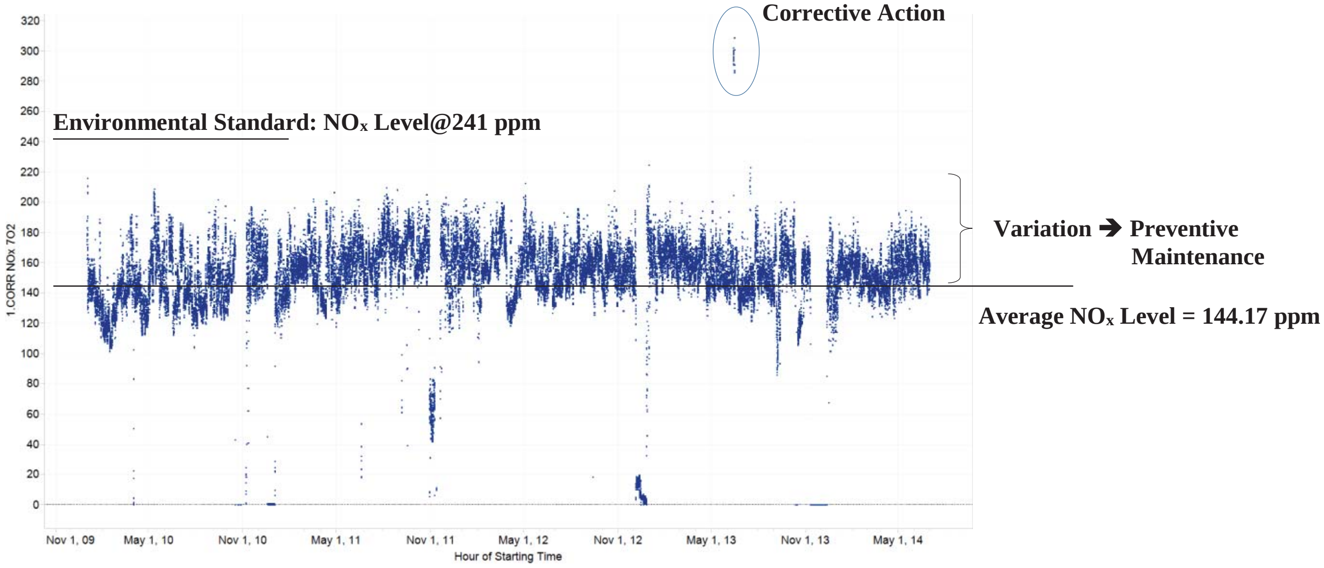

Fig. 1. NO_x level at the stack outlet #1

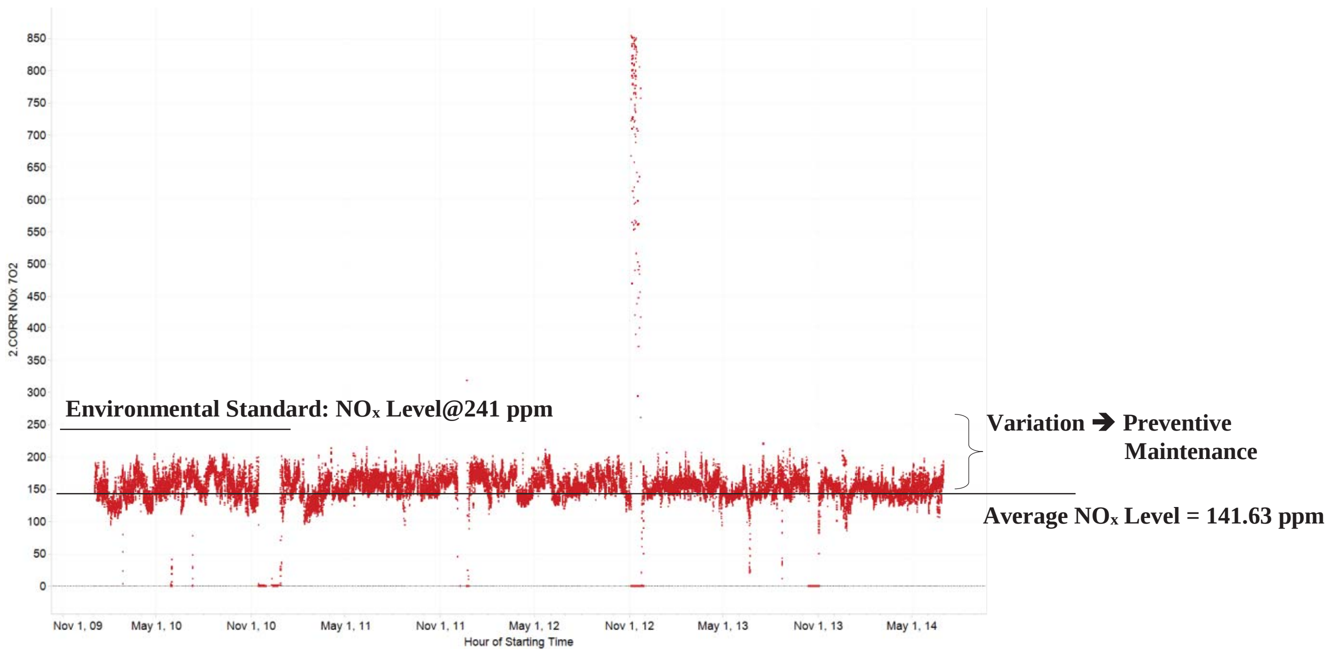

Fig. 2. NO_x level at the stack outlet #2

REFERENCES

[1] A. I. Chatzimouratidis and P. A. Pilavachi, "Decision support systems for power plants' impact on the living standard," *Energy Conversion and Management,* vol. 64, pp. 182-198, 2012.

[2] J. Chongwatpol and T. Phurithititanapong, "Applying analytics in the energy industry: A case study of heat rate and opacity prediction in a coal-fired power plant," *Energy,* vol. 75, pp. 463-473, 2014.

[3] A. Venkatesh, P. Jaramillo, W. M. Griffin, and H. S. Matthews, "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, pp. 1-9, 2012.

[4] M. Waldner, R. Halter, A. Sigg, B. Brosch, H. Gehrmann, and M. Keunecke, "Energy from Waste - Clean, efficient, renewable: Transitions in combustion efficiency and NO_x control," *Waste management,* vol. 33, pp. 317-326, 2013.

[5] C. Shearer, "The CRISP-DM model: The new blueprint for data mining," *Journal of data warehousing,* vol. 5, pp. 13-22, 2000.