Journal of Environmental Management xxx (2016) 1–10

Contents lists available at ScienceDirect

Journal of Environmental Management

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

Research article

Wastewater treatment aeration process optimization: A data mining approach

Ali Asadi, Anoop Verma*, Kai Yang

Department of Industrial and Systems Engineering, Wayne State University, Detroit, MI 48202, United States

A R T I C L E I N F O

Article history:
Received 26 April 2016
Received in revised form
13 July 2016
Accepted 14 July 2016
Available online xxx

Keywords:
Data-driven modeling
Energy optimization
Effluents
Aeration process
Data-mining

A B S T R A C T

Being water quality oriented, large-scale industries such as wastewater treatment plants tend to overlook potential savings in energy consumption. Wastewater treatment process includes energy intensive equipment such as pumps and blowers to move and treat wastewater. Presently, a data-driven approach has been applied for aeration process modeling and optimization of one large scale wastewater in Midwest. More specifically, aeration process optimization is carried out with an aim to minimize energy usage without sacrificing water quality. Models developed by data mining algorithms are useful in developing a clear and concise relationship among input and output variables. Results indicate that a great deal of saving in energy can be made while keeping the water quality within limit. Limitation of the work is also discussed.

1. Introduction

In order to clean wastewater from certain contaminants, wastewater treatment includes different methods and processes that energy intensive. Across USA, wastewater treatment facilities collect, treat, and release about 4 billion gallons of treated effluent per day from about 26 million homes, businesses, and recreational facilities nationwide (Electric Power Research Institute and Inc. (EPRI, 2002). Such moving and treating processes accounts for more than 4% of the US electricity consumption. Minimizing the energy use of WWTPs by just 10% could lead to an annual savings of $400 million or more (http://water.epa.gov/infr). Due to the environmental regulations, wastewater industries are primarily concerned with water quality. The energy consumption in WWTPs is mainly attributed to their heavy mechanical systems, such as the pump and air support systems which are responsible for moving and treating wastewater (Singh et al., 2012; Zhang et al., 2016). The air support system consists of a group of air blowers that provides oxygen to the aeration tanks for removing organic compounds and converting ammonia. Pump system and the air support system are typically 0.5-MW class mechanical equipment and accounts for more than 70% of the electricity consumption of WWTPs.

Traditionally, WWTP operations and designs are based on kinetic models or simulated data (Flores-Alsina et al., 2008; Sin et al., 2009). While such models have provided promising results, it requires some expert knowledge about different systems and subsystems within the process. Moreover, modeling of such systems heavily depends on the design of WWTPs and hence cannot be efficiently generalized.

In wastewater treatment plants, much effort and money is invested in operating and maintaining dense plant-wide measuring networks which is often untouched. With the proliferation of information technologies (IT), it is now possible to perform long term data archiving for analysis. The steadily growing amount of plant data fosters the avenues for plant wide analysis. Over the past few years, data-mining algorithms have gained tremendous popularity in industrial engineering sector consisting of numerous process and sub-processes. Successful applications of data-mining are visible in many domains such as semiconductor manufacturing (Kusiak, 2000; Tan et al., 2006), fault prognosis and diagnosis (Bae et al., 2003), information retrieval (Seo et al., 2001), transportation systems (Long and Li, 2015; Mashayekhy et al., 2015) and renewable energy (Krioukov, 2011; Lu et al., 2005). Few applications of data-mining algorithms in wastewater treatment industry have also been reported. In this regard, Maurice, et al. (Dixon et al., 2007) implemented a set of data-mining algorithms namely regression,

* Corresponding author.
E-mail address: anoop.verma@wayne.edu (A. Verma).

http://dx.doi.org/10.1016/j.jenvman.2016.07.047
0301-4797/© 2016 Published by Elsevier Ltd.

neural network (NNs), rule induction and visual analysis on TELEMAC project datasets to analyses and predict anaerobic digestion in the wastewater treatment plants. Garcia and Gonzalez (Garcıa and González, 2004) applied self-organized maps (SOM) and *k*-means clustering to develop wastewater supervision techniques in acidic chromic wastewater treatment plant. Researchers applied data mining in specific industrial wastewater process namely alcoholic beverage production (Dixon et al., 2007) and metallic industry wastewater (Garcıa and González, 2004). However, being a rather consistent combination of pollutants, and steady and predictable wastewater production, there approach cannot be generalized to other wastewater treatment plants. Gernaey et al (Gernaey et al., 2004). provided a comprehensive list of white box models for municipal wastewater systems. Authors also suggested using more advance data analytics techniques particularly artificial intelligence (AI) to understand and improve performance of wastewater treatment facilities. For improving the prediction accuracy, Chen and Chang (Chen et al., 2003) developed a hybrid control algorithm combining neural network (NN), Genetic Algorithm (GA) and Fuzzy Logic (FL). The control algorithm developed in their work can be used for optimizing and controlling systems when coping with online upset conditions. While their models provide good results, it demands a field expert to set up the fuzzy rules and also demands the full control of the system to be able to be implemented. Hernandez-del-Olmo (Hernández-del-Olmo et al., 2012; Hernandez-del-Olmo et al., 2012) applied AI techniques in order to improve the performance of wastewater plant. Authors utilized model-free reinforcement learning to minimize operational cost while keeping the quality of water within acceptable level. Despite, their methodology seems promising, the model is tested on simulated data with assumption of 70% sunny, 20% raining and 10% stormy days in a year. Tay and Zhang (Tay and Zhang, 1999) attempted to simulate a lab-scale anaerobic wastewater treatment system utilizing lab scaled wastewater treatment and simulated data. Villez et al. (Villez et al., 2008). used a two stage process to aiming to remove nitrogen and phosphorus from a pilot-scale SBR. Authors applied a multi-way principal component analysis (MPCA) process first to clean the data, and then they utilized LAMBDA based clustering method. Their method claims to converge fast but relies on visual inspection to detect outliers and erroneous data. Later, Verma et al. (Verma et al., 2013; Kusiak et al., 2013), utilized data-mining algorithms to predict total suspended solids and carbonaceous biochemical oxygen demand (CBOD) of an industrial wastewater treatment facility. Kusiak and Wei (Kusiak and Wei, 2012; Wei et al., 2012; Wei and Kusiak, 2015) developed a multi-objective model to optimize the activated sludge process in a WWTP and a significant energy saving was observed.

The literature review above indicates lack of large scale, real studies on plant wide aeration process which do not need the full control of the system in order to be implementable as well as being accurate while keeping analysis understandable and explainable to decision makers. Even the published work in the literature that utilizes real world data has simplified models, i.e. effect of suspended phosphorous and dissolved phosphorous etc. is not analyzed (Kusiak and Wei, 2012; Wei et al., 2012; Wei and Kusiak, 2015). The research developed here aims to bridge the gap in the literature by performing analysis on aeration process of a treatment facility and developing easy to use and implementable data-driven models without scarifying the process accuracy.

The paper is organized as follows. In section 2, the description of the aeration process and related dataset is presented. Section 3 describes the proposed solution methodology along with the formulation of the optimization models. In section 4, results obtained from different optimization models are provided. Finally section 5 concludes the present analysis.

2. Data description

The industrial data used to perform the analysis was obtained from Detroit Water and Sewerage Department (DWSD), located in Detroit, MI. DWSD is the largest single-site wastewater treatment facility in the United States. It serves approximately 35% of the population of the State of Michigan, providing treatment of wastewater. DWSD distribute, treat and collects approximately 1.5 billion Gallons of water and wastewater per day (BGD) to be finally discharged into Detroit River. A generic flow diagram of the wastewater treatment process is shown in Fig. 1.

The collected wastewater enters the plant and passes through bar screens. Large items, such as rags and sticks, are screened out for later disposal. After screening, the influent wastewater enters a wet well and then is pumped to primary clarifiers. After a retention time of 1–2 h, scum floats to the surface where it is removed by a skimmer. Then, the wastewater is delivered by intermediate pumps to adjacent aeration tanks. In each aeration tank pure oxygen is provided by centrifugal blowers through bottom of thank. During normal operations, a required quantity of the sludge from the secondary clarifiers, called Returned Activated Sludge (RSL), enters the aeration tanks through sludge pumps. When the RSL and the wastewater are mixed, microorganisms in the activated sludge use oxygen provided by the fine bubble diffusers located on the bottom of the aeration basins to break down the organic matter. The remaining sludge from the secondary clarifiers and the sludge from the primary clarifiers are either pumped to the anaerobic digesters to produce biogas or fed to the incineration process and the final remaining is transported to the land field. The wastewater then enters cylindrical clarifiers for the secondary treatment. The settled sludge is returned back to the aeration basins for continuous supply of microorganisms. The water after being treated from secondary clarifiers is disinfected through chlorination and then discharged into the River.

The analysis presented here aims to improve the aeration process of the DWSD and hence the corresponding three years' worth of data is collected from the plant. The available data for the analysis was collected for the period of September 2012 to October 2014 (see Table 1). Data includes influent flow rate, influent pollutants, effluent pollutants, and aeration process parameters. The data is recorded at 1 h frequency, out of which two years of data is used for building the models and the last year data is used for model testing and validation. Despite the availability of advanced supervisory control and data acquisitions systems, the archiving of numerous parameters is done manually on a shift by shift basis. This poses issues in data quality, including, missing, and invalid values. In this study, the missing values are imputed based on the values recorded in previous time-periods.

3. Solution methodology

In this section, DWSD data (described earlier) is used to model the aeration process with an aim to optimize water quality and energy consumption. In the analysis, the dissolved oxygen (DO) is used as a controlled variable, whereas, influent flow rate, carbonaceous biochemical oxygen demand (CBOD), total suspended solids (TSS), total dissolved phosphorous (TDP), total suspended phosphorous (TSP) and air flow rate were uncontrollable. Due to strong correlation between Chemical Oxygen Demand (COD) and Biochemical Oxygen Demand (BOD) of municipal wastewater under normal operating condition (excluding big storms and flood), COD is not considered as an independent variable. DO is used as an indicator of energy consumption as most of the energy consumed and associated costs in the aeration process is derived from processes which results in increase DO. These processes may include

Please cite this article in press as: Asadi, A., et al., Wastewater treatment aeration process optimization: A data mining approach, Journal of Environmental Management (2016), http://dx.doi.org/10.1016/j.jenvman.2016.07.047