Engineering Applications of Artificial Intelligence 26 (2013) 1643–1651

Contents lists available at SciVerse ScienceDirect

Engineering Applications of Artificial Intelligence

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

Modeling and analysis of pumps in a wastewater treatment plant: A data-mining approach

Andrew Kusiak [a,*], Yaohui Zeng [a], Zijun Zhang [b]

[a] Department of Mechanical and Industrial Engineering, 3131 Seamans Center, The University of Iowa, Iowa City, IA 52242, USA
[b] Department of Systems Engineering and Engineering Management, City University of Hong Kong, P6600, 6/F, Academic 1, Hong Kong

ARTICLE INFO

Article history:
Received 23 April 2012
Received in revised form
27 January 2013
Accepted 3 April 2013
Available online 28 April 2013

Keywords:
Data mining
Pump modeling
Multi-layer perceptron neural network
Time series
Pump scheduling and controlling
Energy consumption

ABSTRACT

A data-mining approach is proposed to model a pumping system in a wastewater treatment plant. Two parameters, energy consumption and wastewater flow rate after the pumping system, are used to evaluate the performance of 27 scenarios while the pump was operating. Five data-mining algorithms are applied to identify the relationships between the outputs (energy consumption and wastewater flow rate) and the inputs (elevation level of the wet well and the speed of the pumps). The accuracy of the flow rate and energy consumption models exceeded 90%. The derived models are deployed to optimize the pump system. The computational results obtained with the proposed models are discussed.

1. Introduction

Wastewater treatment plants (WWTPs) are energy-consuming facilities, and it has been reported that they use 4% of the nation's electricity to move and treat water/wastewater (Clean energy opportunities in water and wastewater treatment facilities background and resources, 2009 Wastewater management fact sheet, 2006; Goldstein and Smith, 2002). Energy costs constitute 25–30% of the operational and maintenance (O&M) costs of water and wastewater facilities. Furthermore, the demand for electricity by WWTPs is expected to grow by approximately 20% over the next 15 years as population grows and environmental requirements become more stringent. Thus, conserving energy at WWTPs is a significant issue. Considerable energy savings can be achieved by optimally operating and monitoring wastewater treatment processes.

Since pump and blower motors account for more than 80% of WWTPs' energy costs (Wastewater management fact sheet, 2006), designing effective operational strategies for pumping systems can greatly reduce these costs. Research on pumping systems has been reported in the literature. Ormsbee and Lansey (1994) reviewed different models and approaches proposed for water-supply pumping systems. Bechwith and Wong (1995) used a genetic algorithm to solve the pump scheduling problem in a multi-source water supply system with multiple tanks. Barán et al. (2005) utilized a mass balance model and evolutionary computational algorithms to solve a multi-objective, pump-scheduling problem by minimizing four types of costs while satisfying the water demand and other constraints. More recently, Yang and Børsing (2010) developed a mixed-integer, non-linear, programming model for a simple, multi-pump, boosting system that included three variables: speed pumps, a simple water circular loop, and a storage tank. Wang et al. (2009) modeled pump scheduling in a water distribution system as a bi-objective optimization problem by taking into account pump operational costs and the land subsidence issue to reduce costs and address environmental concerns. The results obtained by the proposed genetic algorithm-based method have resulted in a wide range of schedules.

Modeling approaches investigated for pumping systems to date have been based predominantly on physics and mathematical programming. The assumptions that have been made limit the applicability of these models in industry. For example, Barán et al. (2005) assumed that the capacity of each pump, the pump discharge, electric energy consumption, and the power consumption of each pump combination were fixed for a period of 1 h. These assumptions neglected the dynamic characteristics of the pumping process and impeded the applicability of such models in industrial practice. Assuming static models of the pumps, Wang

* Corresponding author. Tel.: +1 3193355934.
E-mail addresses: andrew-kusiak@uiowa.edu (A. Kusiak), yaohui-zeng@uiowa.edu (Y. Zeng), zijzhang@cityu.edu.hk (Z. Zhang).

0952-1976/$ - see front matter
http://dx.doi.org/10.1016/j.engappai.2013.04.001

et al. (2009) demonstrated a potential for improvement of the efficiency of multi-pump systems.

In this paper, a data-mining approach for modeling pumping systems is introduced. As an emerging science with an abundance of successful applications in wind energy (Kusiak and Zhang, 2011; Zhang and Kusiak, 2012; Verma and Kusiak, 2012), HVAC systems (Kusiak et al., 2011; Kusiak and Li, 2010; Kusiak et al., 2010), and other areas (Berry and Linoff, 2004; Harding et al., 2006; Shah et al., 2006; Siegelmann and Sontag, 1994), data mining has proven to be a promising approach for modeling complex, dynamic, and non-linear systems. However, the research literature on utilizing data mining in modeling and optimizing pumping systems is sparse. In this research, data-mining approaches are applied to model the energy consumption and hydraulic workload (water flow rate after pumping) of a pumping system based on 15-min data collected by a municipal WWTP. Effectiveness of data-mining approaches in modeling pumping system is demonstrated. The developed models are applicable to scheduling operations of a pumping system.

2. Problem definition and data description

2.1. Problem definition

This research focused on investigating the performance of a pumping system in the wastewater treatment process of the Wastewater Reclamation Authority (WRA), Des Moines, Iowa, USA. The preliminary (initial) treatment facilities in the WRA include six, 55 MGD-class, variable-speed, raw wastewater pumps, five mechanical bar screens, and six aerated grit chambers. These facilities remove debris and large particles from the wastewater to prepare it for the biological treatment process that occurs later (Fig. 1). The bar screens remove paper, sticks, and other solids in the influent from the raw wastewater junction chamber to prevent these materials from damaging the equipment in the plant. Then, the pumping system conveys the influent into aerated grit chambers for grit removal.

Improved strategies for controlling pumping systems may lead to significant energy savings. Such strategies usually involve two issues, i.e., pump scheduling (determining which pumps should be operated and when they should be operated) and optimal control of the speed of the pumps. To address these two issues, pump models simulating performance of the pumping system are needed. Performance of a pumping system is evaluated by two parameters: energy consumption and water flow rate.

Establishing accurate models to evaluate the performance of a pumping system is challenging due to the system complexity. In this research, a data-mining approach is employed to model the pumping system by identifying its underlying performance, i.e., relationship between the system energy consumption, the water flow rates after pumping, and the pump speed. Data-mining algorithms extract such models from large datasets. Once the accuracy of these models is acceptable, they can be applied to optimize pump schedules and the speeds of pumps with the goal of saving energy. This paper addresses the process of modeling pumping systems, and the optimization component is considered for future research.

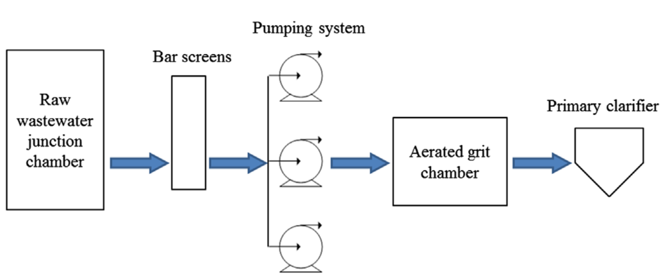

Fig. 1. Schematic diagram of the preliminary wastewater treatment process.

2.2. Data description and processing

The data used in this research were collected from July 20, 2011, to January 31, 2012. The pump speed, the energy consumption, the water flow rate after pumping, and the elevation level of the wet well were selected for developing data-mining models. The parameters that we used were sampled every 5 min and averaged over 15-min intervals. The maximum pump output was considered to be approximately 50 MGD. A higher wet-well elevation could push the maximum to about 55 MGD. Pump speed is measured in percentage of the maximum speed (500 RPM), and it can be transformed into RPM by multiplying the percentage value by a factor of 5.14. Since pumps are controlled by variable frequency drives (VFDs), pumping begins only when the pump achieves at least 80% of its speed to overcome the head pressure of the system. Thus, the pump is not considered to be working when its speed is less than 80% of the maximum speed. Data samples for which the pump speed was less than 80% were discarded.

The pumping system can be operated with various configurations of pumps. Pump configurations with more than 300 data points were considered in this research as sufficient information is desired by data-mining modeling. The available data contained errors caused by system failures, sensor malfunctions, and transmission errors. These errors led to missing data or data points that exceeded the extreme values of the parameters. Thus, the datasets were processed before use in modeling in order to account for abnormal, noisy, erroneous, missing, and irrelevant data. The processed datasets for each combination of pumps were divided into training sets (75%) and test sets (25%). The training sets were used to develop the energy consumption and water flow rate models with data-mining algorithms. The test sets were utilized to assess the accuracy and robustness of the developed models. Table 1 describes the datasets of 27 scenarios of pump combinations.

3. Model development and analysis

3.1. Energy consumption and flow rate models

In this section, energy consumption and flow rate models of each of the pump-combination scenarios are built by data-mining algorithms. The inputs of the energy consumption and the flow rate models are the same, namely, the wet well elevation level and the speed of each pump. The speed of each pump is represented by the percentage of that pump's maximum speed. The energy consumption is represented in kW since the time interval used for sampling was constant. Five scenarios (highlighted in Table 1), i.e., scenario 1 (pump 1 only), scenario 7 (pumps 1 and 3), scenario 19 (pumps 1, 4, and 5), scenario 26 (pumps 1, 3, 4, and 5), and scenario 27 (pumps 2, 3, 4, 5, and 6), were used to illustrate the modeling process of energy consumption and flow rate. Models for all other scenarios also were developed, and the results are presented in this paper.

The input and output parameters for the energy consumption model and the flow rate model of scenarios 1, 7, 19, 26 and 27 are listed in Tables 2–6. Based on the selected parameters, the energy consumption and flow rate models of five scenarios of the pumping system at time t are represented as Eqs. 1–10.

$$E_1(t) = f(v_1(t), l(t)) \quad (1)$$