Water Research 96 (2016) 155–165

Contents lists available at ScienceDirect

Water Research

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

Multivariate statistical analysis of water chemistry conditions in three wastewater stabilization ponds with algae blooms and pH fluctuations

Jack Wallace, Pascale Champagne*, Geof Hall

Department of Civil Engineering, Queen's University, Ellis Hall, 58 University Avenue, Kingston, Ontario, K7L 3N6, Canada

A R T I C L E I N F O

Article history:
Received 11 October 2015
Received in revised form
18 March 2016
Accepted 21 March 2016
Available online 24 March 2016

Keywords:
Algae
Lagoons
Principal components analysis
Wastewater stabilization ponds

A B S T R A C T

The wastewater stabilization ponds (WSPs) at a wastewater treatment facility in eastern Ontario, Canada, have experienced excessive algae growth and high pH levels in the summer months. A full range of parameters were sampled from the system and the chemical dynamics in the three WSPs were assessed through multivariate statistical analysis. The study presents a novel approach for exploratory analysis of a comprehensive water chemistry dataset, incorporating principal components analysis (PCA) and principal components (PC) and partial least squares (PLS) regressions. The analyses showed strong correlations between chl-a and sunlight, temperature, organic matter, and nutrients, and weak and negative correlations between chl-a and pH and chl-a and DO. PCA reduced the data from 19 to 8 variables, with a good fit to the original data matrix (similarity measure of 0.73). Multivariate regressions to model system pH in terms of these key parameters were performed on the reduced variable set and the PCs generated, for which strong fits ($R^2 > 0.79$ with all data) were observed. The methodologies presented in this study are applicable to a wide range of natural and engineered systems where a large number of water chemistry parameters are monitored resulting in the generation of large data sets.

1. Introduction

Wastewater stabilization ponds (WSPs), commonly implemented as a treatment approach in small and rural community facilities (Steinmann et al., 2003), are a potentially sustainable wastewater treatment technology with low capital costs, as well as low energy and maintenance requirements (Steinmann et al., 2003; Shammas et al., 2009). However, as semi-passive systems with high nutrient concentrations, WSPs can promote the excessive growth of algal biomass, particularly during the summer months, where temperatures and sunlight are highest (Steinmann et al., 2003). This growth, while contributing to significant nitrogen and phosphorus removal, can lead to increases in pH (Shammas et al., 2009), and may thus perpetuate water quality issues in the receiving environment(s). During peak photosynthetic activity, pH levels can rise to above 9.0, as CO_3^{2-} and HCO_3^- are consumed to produce CO_2 for cell growth and OH^- ions are left in excess (Gschlößl et al., 1998; Kayombo et al., 2005; Barsanti and Gualtieri, 2006; Veeresh et al., 2012). Gschlößl et al. (1998) and Steinmann et al. (2003) found constructed wetlands to be effective in buffering high pH levels generated in lagoons with excessive algae growth.

Prior to implementing system modifications such as constructed wetlands, preliminary characterization through multivariate statistical analysis can provide an understanding of the water chemistry processes occurring in WSPs and guidance for operational changes or further monitoring, as noted by other researchers (Miettinen et al., 2004; Ouali et al., 2009; Dong and Reddy, 2010). Although passive in their operation, WSPs are highly complex open systems, as they are subject to ambient environmental conditions and a number of interacting parameters, thus presenting unique challenges for modeling and performance prediction (Ouali et al., 2009). To understand the interactions between parameters from recorded field data, and to fully characterize WSPs, multivariate statistical modeling, such as Principal Components Analysis (PCA) and multivariate regressions, is appropriate (Ouali et al., 2009).

To date, much of the multivariate statistical work using wastewater data has largely focused on monitoring and process control, particularly for conventional treatment systems (Lee et al., 2004; Aguado and Rosen, 2008). However, the following highlights some of the research that has been conducted for passive or semi-passive treatment systems. Over a 7-year period, Miettinen et al. (2004) performed an intensive characterization of a lagoon

* Corresponding author.
E-mail addresses: jack.wallace@ce.queensu.ca (J. Wallace), pascale.champagne@queensu.ca (P. Champagne), gh26@queensu.ca (G. Hall).

http://dx.doi.org/10.1016/j.watres.2016.03.046
0043-1354/

Nomenclature

chl-a	chlorophyll-a
cBOD	carbonaceous biochemical oxygen demand
COD	chemical oxygen demand
CRC	Canada Research Chair program
CRD	Collaborative Research and Development grant
CREATE	Collaborative Research and Training Experience Program
d	days
DO	dissolved oxygen
HRT	hydraulic retention time
L	liter
m	meter
m^2	square meter
m^3	cubic meter
mg	milligram
μg	microgram
N	nitrogen
NH_3	unionized ammonia
NH_4^+	ammonia
NO_2^-	nitrite
NO_3^-	nitrate
ORP	oxidation reduction potential
ρ	Spearman ρ correlation coefficient
P	phosphorus
Q	similarity measure
PC	Principal Components
PCA	Principal Components Analysis
PLS	Partial Least Squares
R^2	regression coefficient
NSERC	National Sciences and Engineering Research Council of Canada
STEWARD	Systems Training and Education in Water Assets Research Development
TIP	total inorganic phosphorus
TKN	total Kjeldahl nitrogen
TSS	total suspended solids
WPCP	Amherstview Water Pollution Control Plant
WSP	wastewater stabilization pond
WSP#1	Amherstview WPCP WSP#1 effluent
WSP#2	Amherstview WPCP WSP#2 effluent
WSP#4	Amherstview WPCP WSP#4 effluent

system comprising 11 ponds in series for 19 parameters. The study included chlorophyll-a (chl-a), but excluded phosphorus species, which is often the limiting nutrient for algae growth in freshwater systems (Wang et al., 2010). The study allowed for the identification of lagoons where the most significant chemistry changes and operational challenges were occurring, and the parameters involved. Additionally, the multivariate statistical methods used (PCA and parallel factor analysis) were verified to be effective in modeling the environmental data set and both methods produced consistent results. Ouali et al. (2009) conducted a study focusing on a multivariate analysis of three parameters (oxygen demand, suspended solids and temperature) over a 12-year period for an aerated lagoon, in order to evaluate treatment performance and highlight the importance of multivariate statistical techniques for large datasets. The authors used PCA to reduce the dataset from six variables to three principal components (PCs), which explained 74% of the variability in the data. The study reaffirmed the benefit and practicality of multivariate statistical analysis in exploratory evaluation of complex water chemistry systems. Dong and Reddy (2010) examined the relationships between nitrogen, phosphorus, chemical oxygen demand, and microbial diversity and growth with PCA for two anaerobic lagoons treating swine wastewater. Through PCA analysis, the authors identified a gradual shift in microbial populations and compositions along different steps of the wastewater treatment process, and confirmed the strong relationship between microbial diversity and these key water chemistry parameters.

This study aims to add to the body of literature that pertains to water chemistry analyses, in the context of pH and algae growth interactions. A novel application of multi-variate regression analysis to supplement PCA interpretation is proposed in this study to further advance the understanding of water chemistry in WSP and similar systems. As noted above, WSPs are highly complex systems with many dynamic water chemistry interactions, and the use of multivariate statistical techniques can provide valuable guidance and insight into understanding and diagnosing operational issues associated with water quality. To this end, the study consisted of a preliminary evaluation and statistical analysis of a system of three WSPs operating in series that have experienced consistently high pH levels and excessive algae growth during the summer months since 2003. The characterization of a wide spectrum of water quality parameters (nutrients, oxygen demand, solids, chl-a) in a WSP system using multivariate statistical methods to explain the relationships between algae growth and changes in pH has not been reported. Zang et al. (2011) noted the need for further research for these simply engineered but complex water quality systems with relatively low (under 10 μg/L) chl-a concentrations. Over a 4-month period from mid-summer (May 22, 2012; peak algae growth) into mid-fall (October 30, 2012; algae die-off), a comprehensive set of water chemistry parameters were collected and analyzed using correlations analysis, PCA, and PC and partial least squares (PLS) regression analysis. These analyses were performed to provide a deeper understanding of the underlying water chemistry conditions and of the principal parameters affecting pH levels during the algae growth period, such that long-term system modifications and monitoring could be designed and integrated. Specifically, PCA followed by regression analysis were performed on a dataset to determine the parameters that influence pH in the WSPs, such that treatment could be tailored according to these relationships in the future. Another key element of this work was the statistical confirmation of the relationship between excessive algae growth and the high pH levels observed. Additionally, it is expected that identification of the principal parameters contributing to system dynamics and performance, through the application of the multivariate statistical techniques presented, will allow for more targeted water quality parameter monitoring in the future.

2. Methods

2.1. System overview

In this study, three WSPs at the Amherstview municipal water pollution control plant (WPCP) in eastern Ontario, Canada, were grab sampled over a 4-month period (July–October 2012). The monitoring period was selected based on the period of algae