Science of the Total Environment 598 (2017) 249–257

Contents lists available at ScienceDirect

Science of the Total Environment

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

Assessment of wastewater treatment facility compliance with decreasing ammonia discharge limits using a regression tree model

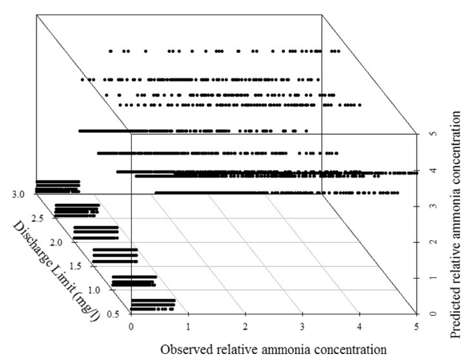
Bihu Suchetana [a,*], Balaji Rajagopalan [a,b], JoAnn Silverstein [a]

[a] Department of Civil, Environmental and Architectural Engineering, University of Colorado, 428 UCB, Boulder, CO 80309, USA
[b] Cooperative Institute for Research in Environmental Sciences, University of Colorado, CIRES Building, Rm 318, Boulder, CO 80309, USA

HIGHLIGHTS

- Wastewater treatment plant compliance modeled using regression trees.
- Compliance with current and future ammonia discharge limits studied.
- Previous month's compliance history chiefly affects compliance.
- Validation of regression trees shows median predictive accuracy of 70% or higher.
- Utility of proposed methodology in nutrient trading demonstrated.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 20 January 2017
Received in revised form 8 March 2017
Accepted 25 March 2017
Available online xxxx

Editor: D. Barcelo

Keywords:
Wastewater
Ammonia
Regression trees
Ammonia permit limit
Discharge permit compliance

ABSTRACT

A regression tree-based diagnostic approach is developed to evaluate factors affecting US wastewater treatment plant compliance with ammonia discharge permit limits using Discharge Monthly Report (DMR) data from a sample of 106 municipal treatment plants for the period of 2004–2008. Predictor variables used to fit the regression tree are selected using random forests, and consist of the previous month's effluent ammonia, influent flow rates and plant capacity utilization. The tree models are first used to evaluate compliance with existing ammonia discharge standards at each facility and then applied assuming more stringent discharge limits, under consideration in many states. The model predicts that the ability to meet both current and future limits depends primarily on the previous month's treatment performance. With more stringent discharge limits predicted ammonia concentration relative to the discharge limit, increases. In-sample validation shows that the regression trees can provide a median classification accuracy of >70%. The regression tree model is validated using ammonia discharge data from an operating wastewater treatment plant and is able to accurately predict the observed ammonia discharge category approximately 80% of the time, indicating that the regression tree model can be applied to predict compliance for individual treatment plants providing practical guidance for utilities and regulators with an interest in controlling ammonia discharges. The proposed methodology is also used to demonstrate how to delineate reliable sources of demand and supply in a point source-to-point source nutrient credit trading scheme, as well as how planners and decision makers can set reasonable discharge limits in future.

* Corresponding author.
E-mail address: bihu.suchetana@colorado.edu (B. Suchetana).

http://dx.doi.org/10.1016/j.scitotenv.2017.03.236
0048-9697/

1. Introduction

In the United States, the discharge of pollutants from wastewater treatment facilities to surface waters is regulated by the National Permit Discharge Elimination System (NPDES) created in 1972, under the Clean Water Act. NPDES permits are typically issued and enforced by individual States to control pollutant levels in wastewater discharges by setting effluent concentration limits for regulated constituents. These limitations may be technology based or water quality based (USEPA, 2010). Technology based effluent limits (TBELs) are based on conventional secondary treatment technologies, and set national maximum limits for Biochemical Oxygen Demand (BOD), Total Suspended Solids (TSS) and pH in discharging waters. Water Quality Based Effluent Limits (WQBELs) are set at the state level to achieve desired water quality standards in a particular receiving water, and typically include limits for effluent nitrogen and phosphorus, varying geographically and often seasonally. Spatial variability in WQBELs typically depends on local aquatic life conditions, designated uses of receiving waters and dilution factors while seasonal variations are due to seasonal changes in streamflow, water temperature and sensitivity of the developmental stages of aquatic organisms (Rossman, 1989).

After the USEPA's decision to decline issuing TBELs for nutrients (Shapiro, 2012), many states have considered new uniform statewide permit limits for nutrients, e.g. Colorado (CDPHE, 2012) and Utah (Daigger et al., 2014). Concern about the ability of plants to comply with lowered permit limits using current infrastructure and the cost of implementation of nutrient removal technology has led some states to consider exemptions and allow a transition period before new limits take effect. For example, the State of Colorado exempts plants with design capacity less than one million gallons per day (MGD) (~4000 m^3/d) from new nitrogen and phosphorus limits (CDPHE, 2012).

The effectiveness of lower nutrient discharge limits in achieving improved water quality is dependent on compliance with these limits as well as adequate enforcement measures. Unfortunately, a number of studies indicate appreciable levels of non-compliance with existing limits and spotty enforcement across the country (Cassady, 2004; USEPA Office of Inspector General, 2011). Moreover, the USEPA expects available funds for NPDES permit enforcement to decrease in the coming years (USEPA, 2015). Beyond fiscal limits, there are often political and legal obstacles to both public agency and citizen-initiated enforcement actions (Andreen, 2007; Steinzor, 2003). Given the uncertainty around compliance with permit limits and challenges to enforcement measures across the country, achieving compliance with new, more stringent nutrient permits and associated water quality improvement may be a challenging task.

Uncertainty in compliance can also affect nutrient trading, a market-based strategy for watershed-scale nutrient management consistent with the Total Maximum Daily Load (TMDL) framework used by the EPA and State regulators (Branosky et al., 2011; Hoag and Hughes-Popp, 1997; Jones et al., 2010; Lal, 2010; Ribaudo et al., 2005; Virginia Department of Environmental Equality, 2016). Studies of the feasibility of nutrient credit trading and attempts to implement a market approach have focused on trades between point (wastewater discharges) and non-point (primarily agriculture) sectors and results have been generally disappointing. Factors that limit inter-sector trading are differing economic incentives, regulatory structures, and even differences in the ability to verify credits (Hoag and Hughes-Popp, 1997; King and Kuch, 2003). The most acknowledged success of emissions trading is the cap-and-trade system to reduce acid rain by limiting sulfur dioxide (SO_2) emissions from coal-fired power plants enabled by the Clean Air Act Amendments of 1990. Factors associated with the success of the acid rain program include similar incentives to control emissions across the market and willingness of stakeholders, especially policymakers and regulators to allow flexibility in reaching environmental goals (Chan et al., 2012). A similar market-based approach for nutrient credit trading among wastewater treatment plants will rely on a sufficient pool of potential sellers – plants that reliably discharge well-below permit limits, and potential buyers – plants that are at risk for violating permits. However, uncertainty in nutrient discharges from both buyers and sellers is a major challenge to establishing reliable markets for nutrients (Dennison et al., 2012; USEPA, 2007). We propose a diagnostic framework, which can evaluate compliance to nutrient limits as a function of readily measurable variables, can provide timely guidelines for estimating the reliability of individual treatment plants with benefits for utilities and policy makers.

Recent research using statistical models based on long-term effluent data has identified the sources of variability in compliance with effluent limits (Suchetana et al., 2016; Weirich et al., 2015a,b; Weirich et al., 2011). While these models were well suited for predictive purposes, they could not offer any operational or diagnostic insights to improve compliance. For this study, the trend of increasing regulation of nutrient discharges was the basis for focus on compliance with ammonia limits. This necessitated building a data-driven diagnostic framework that could distinguish which variables affect compliance to both existing and more stringent ammonia discharge standards, as well as promote strategies for individual wastewater treatment facility operators to improve compliance and achieve water quality standards for nutrients at a reasonable cost. To this end, we develop a regression tree-based diagnostic framework to evaluate compliance of wastewater treatment plants to current and future ammonia discharge standards. The new approach in this study is the first methodological attempt to evaluate compliance with ammonia limits using statistical modeling based on past performance data with the advantage of enabling consideration of more variables than previous approaches and output that can guide both treatment facility operators and environmental policy makers.

2. Methodology

The primary objective of the regression tree analysis is to propose a simple diagnostic framework to serve two related goals to aid water quality management decisions. First is to assess how more stringent discharge limits affect the fraction of dischargers able to comply with lower limits, in this case, for ammonia. Second is to test the relationship between the level of discharge limits and the creation of pools of "winners" (over-compliant dischargers) and "losers" (likely violators) who might form a market for nutrient trading to achieve water quality goals more cost effectively. Conventional regression methods are not able to achieve this objective in a direct manner. Moreover, application of random forest analysis enables determination of the relative importance of independent variables in achieving compliance with both current and future ammonia limits. The pruned regression tree provides a graphical identification of the relative importance of covariates and associated threshold values. A brief description of regression trees and random forests is provided below.

2.1. Regression trees

Classification and Regression trees (CART) (Breiman et al., 1984) are local regression models, which work in a recursive manner to partition the predictor space by delineating regions where the predictions of the dependent or response variable to a set of predictors is homogeneous. These trees first group the predictors and then a particular model is assigned to each of these groupings - the simple average is most commonly used (Elith et al., 2008). Local models have the advantage of representing the true relation between the covariates (i.e. predictors) and the response variable in each of the groups in the tree, as compared to global models where a single equation is used for the entire data set. If the response variable is continuous then it is called a regression tree; if categorical then it is called a classification tree. The CART algorithm is summarized in the Supplementary section for interested readers.

Trees grown to their full size tend to over-fit the data i.e. they tend to fit the training data so well that they cannot extrapolate efficiently (Elith et al., 2008). To overcome this problem, the size of the tree