TECH BRIEF

Data Science Summer Fellows Program

Summer 2021

Modernizing Water and Wastewater
Treatment through Data Science
Education & Research

Metro Wastewater Reclamation District – PAA Disinfection

Sophie Kearney, Baylor University
Rachel Walker, Colorado School of Mines
Joseph Yu, Baylor University

SUMMARY

Peracetic acid (PAA) disinfection is used to inactivate biological contaminants, including E.coli, in wastewater treatment. However, it is more expensive than traditional chlorination methods. Accurate dosing is important to minimize the costs faced by wastewater treatment plants. This Tech Brief discusses ways to predict the efficiency of PAA disinfection by investigating the contact time, CT, to maximize efficiency of the wastewater treatment facility. We found that certain environmental factors impact CT and that by accounting for these factors, we can improve the accuracy of CT calculation and its correlation to log removal of E. coli.

INTRODUCTION

Traditional disinfection in wastewater treatment is achieved by adding chlorine to water, however the process generates problematic disinfection byproducts (DBPs). The concentration of DBPs can be reduced if peracetic acid (PAA) disinfection methods were applied to the wastewater disinfection process instead of chlorination.[1,2] On the other hand, the cost of the PAA disinfection methods are more expensive compared to the traditional chlorination methods. The PAA disinfection is controlled by the dosing concentration in the plant, so to optimize the dosing protocols, the factors that affect the process must be identified. Two crucial variables are the Pre-disinfection and Post-disinfection count of *E. coli*. Another important variable is contact time (CT), defined as the integral of the PAA decay equation from time, t_o to t. CT shows how long the disinfectant is in contact with the wastewater, so a larger CT suggests greater contact of PAA with wastewater and is a crucial indicator of the disinfection process. The goal of this project is to isolate the variables that could affect the PAA disinfection performance and CT calculation.

FACILITY SYSTEM DESCRIPTION

The wastewater treatment facility is located in Denver, belonging to Metro Wastewater. The PAA disinfection system is employed to eliminate bacteria and viruses in the water, and is located at the end of the wastewater treatment process during disinfection before discharging.

Source: Google Maps 2020

Figure 1: Birds Eye view of the PAA Disinfection site. In Figure 1, PAA concentration analyzers were placed at the location indicated as the square, the pre-disinfection E. coli sample count was taken at where the circle is located, and the post-disinfection E. coli sample count was taken at where the triangle is shown.

The plant operates by dosing PAA into the wastewater and allowing the water to travel through the serpentine (S-shaped) basin to reach the final treatment processes before discharging as shown in Figure 1. The concentration of PAA dosing is determined using data that was gathered from the PAA analyzer and the following equations:

$$C_t = (C_0 - D) * e^{-kt} \quad (1)$$

$$CT = \frac{(C_0 - D)}{k} * (1 - e^{-kt}) \quad (2)$$

The target CT is set by the plant and the PAA dosing concentration C_o is then calculated using equations 1 and 2.

It is crucial for the dosing control system to have accurate PAA concentration readings from the analyzers, and the assumption is that at any given time point there is at least one PAA analyzer functioning and collecting readings from the wastewater.

DATA DESCRIPTION

The data provided by Metro Wastewater was collected from April 15th, 2020 to March 18th, 2021. Previously mentioned was the daily data from the PAA analyzers and the wastewater sampling for the *E.coli* count calculation. In addition, sensors record the PAA dosing flow rate and the total plant flow rate simultaneously. These measurements extend from every minute to roughly every thirty minutes. Downstream analyzers record PAA concentration data in time intervals ranging from fifteen seconds to every few minutes. Effluent nutrient concentration data such as ammonia, nitrogen, and phosphate was measured by sensors and recorded hourly. Total Suspended Solids (TSS) and water temperature was provided as hourly averages. Furthermore, basin measurements of nutrient concentrations are recorded every few days. Decay coefficient (k), used in calculating CT, was observed using jar testing.

Various points within the data set contain missing values. Most often this was a result of sensors being down for maintenance or calibrations.

EXPLORATORY DATA ANALYSIS

Log removal of *E.coli* was calculated using the following equation:

$$(LogRemoval) = log(1 - \frac{N_t}{N_0}) \quad (3)$$

N_t and N_o represent Post and Pre Disinfection E. coli, respectively. We analyzed correlations between log removal and other covariates, specifically CT, but we did not find any significant relationships. Figure 2 shows that CT, ammonia, and temperature do not follow the same trend as log removal.

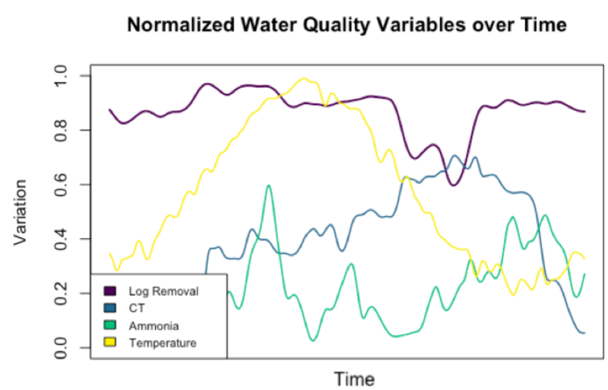


Figure 2: Normalized water quality variables to visualize a lack of correlation with the log removal trend.

Equation 2 indicates that k plays a large role in calculating CT. Accurately calculating CT is constrained by the accuracy of the k value. For example, when analyzing correlations between CT and phosphorus, k can increase or decrease the CT in a way that could represent a false correlation between two variables (Figure 3).

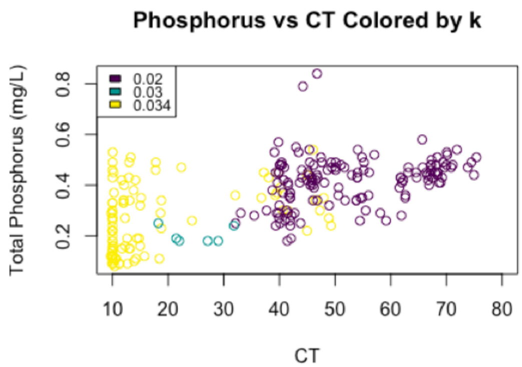


Figure 3: Phosphorus vs Metro CT Colored by k, the increasing CT could be caused by k, not phosphorus.

STATISTICAL ANALYSIS and RESULTS

Adjustment to Water Travel Time

Treated wastewater during disinfection is constantly moving through the serpentine basins at a variable rate. To adjust for the varying flow, we implemented a simple time model that accounts for the total plant flow rate:

$$(BasinVolume) = (FlowRate) * (\Delta t) \quad (4)$$

$$\Delta t = \frac{(BasinVolume)}{(FlowRate)} \quad (5)$$

This estimates the time required for the same body of water to travel from the PAA dosing point to the PAA analyzer that are both shown in Figure 1.

Using the relationship that was provided from Metro Wastewater, the PAA dosing flow rate readings and PAA concentration analyzer measurements were matched by the time interval that most closely match the Δt calculated.

Linear Regression Model

$$\frac{N_t}{N_o} = (1 - \beta)e^{\left(-k_p CT^m\right)} + (\beta)e^{(-k_d CT)} \quad (6)$$

The mechanistic model (Equation 6) of PAA efficacy suggests a correlation between *E.coli* removal and CT. We created a linear model to analyze this relationship, but the R^2 value is 0.044 (Table 1), which does not indicate a strong relationship with the current CT.

Model	R^2
Log removal vs provided CT	0.044
Provided CT vs relevant covariates	0.802
Re-calculated CT vs relevant covariates	0.846

Table 1: R^2 value comparisons for different linear regression models

To improve this correlation, we investigated the accuracy of CT. We created a linear model of the provided CT and environmental factors that could influence CT accuracy, such as water quality or nutrient concentrations. This model had an R2 of 0.0802 (Table 1), which suggests that these environmental factors influence CT and can also predict CT.

Using the adjusted time, CT was recalculated and another linear model was created that had a slightly stronger relationship to environmental factors, specifically ammonia, TSS, and temperature. Therefore, these variables should be taken into consideration when calculating CT, and may improve CT accuracy and allow for a better fit to the mechanistic model.

K Sensitivity Analysis

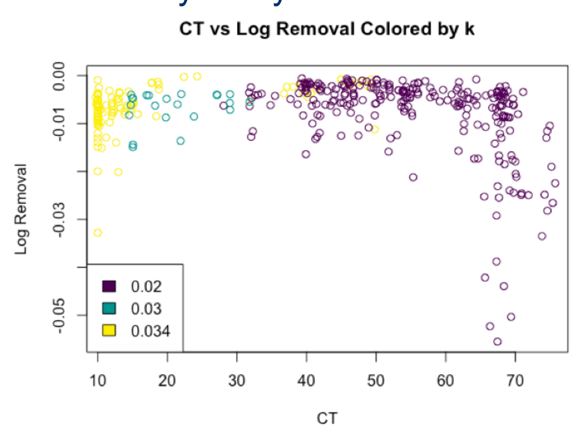


Figure 4: Metro Wastewater CT vs. Log Removal of E. coli colored by the decay coefficient (k) values.

As a result of its importance in the mechanistic model, we hypothesized changing k values may be impacting CT calculation.

There seems to be very little correlation between CT and log removal, but when visualizing this relationship while also considering k (Figure 4), it seems like the change in CT values could be caused by the k changing 3-4 times in our time interval.

Estimating k for CT calculation

Decay is a crucial part of the calculation of CT, so we then improved the precision of decay by using a variable k influenced by environmental factors. We identified the variables that were the most likely to impact k: water temperature, TSS, and the nutrient concentrations in the water. Metro acknowledged that TSS and water temperature were known to impact k. This supported our predictive linear model of CT using dependent variables which were not directly used in calculating CT. As a result, we created a k that adjusted to changes in TSS, temperature, and ammonia by scaling this data to various ranges of acceptable k values. Ammonia was used to represent nutrient concentrations because it had the most complete set of data. We also created a vector of variable ks that combined multiple environmental factors. We found that calculating CT with a variable k influenced by environmental factors improved the relationship between CT and log removal, which is supported by the mechanistic model of disinfection (Equation 6).

CONCLUSIONS

The empirical model suggests that CT is influenced by environmental factors, however we do not recommend solely using this empirical model to predict CT without first basing it on a larger dataset.

We found that adjusting the time using the flow rate improves the accuracy of the calculated CT compared to the provided CT. When adjusting calculations to CT, we also found that we can create a model that more precisely fits the data and also is slightly more correlated to log removal.

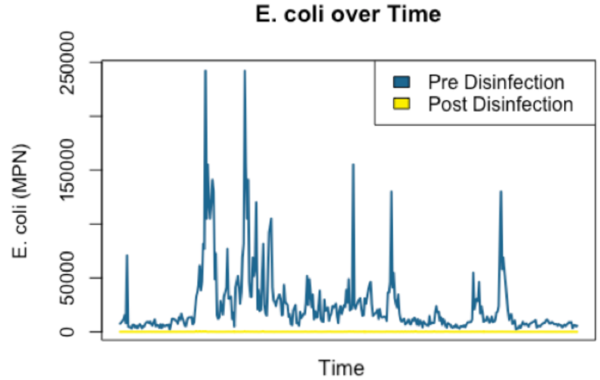


Figure 5: Pre and Post Disinfection of E. coli over time.

We did not find a strong relationship through empirical models between log removal and CT or any other relevant covariates. Low variation in log removal values could cause the weak correlations seen in empirical models. Pre-disinfection *E. coli* varies over time, but Post-disinfection *E. coli* is almost always very close to zero with little variation (Figure 5), and this is reflected in log removal. Through a k sensitivity analysis, we found that variations in k greatly impact CT calculation over a static k. This suggests that a variable k would accurately reflect the wastewater conditions and improve PAA dosing.

Accounting for environmental factors like ammonia, TSS, and water temperature improves our calculation of CT by adjusting the variables used to derive CT, such as k. Although we only found a slightly better correlation between the newly calculated CT and log removal, by recalculating CT, we improved the overall accuracy, thus further research can be done to investigate whether increasing the frequency of water quality observations such as secondary effluent TSS and water nutrient loading improves the correlation between CT and log removal. Additionally, there may be a more complex relationship between CT and log removal that could be investigated.

REFERENCES

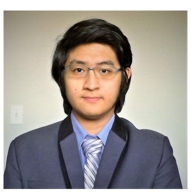
(1) Crebelli, R.; Conti, L.; Monarca, S.; Feretti, D.; Zerbini, I.; Zani, C.; Veschetti, E.; Cutilli, D.; Ottaviani, M. Genotoxicity of the Disinfection By-Products Resulting from Peracetic Acid or Hypochlorite-Disinfected Sewage Wastewater. Water Res. 2005, 39 (6), 1105–1113.

(2) Monarca, S.; Richardso, S. D.; Feretti, D.; Grottolo, M.; Thruston, A. D.; Zani, C.; Navazio, G.; Ragazzo, P.; Zerbini, I.; Alberti, A. Mutagenicity and Disinfection ByProducts in Surface Drinking Water Disinfected with Peracetic Acid. Environ. Toxicol. Chem. 2002, 21 (2), 309–318.

AUTHORS

Sophie Kearney, Science Research Fellow concentrating in Bioinformatics in the Class of 2024.

Rachel Walker, Studying Applied Mathematics at the Colorado School of Mines, she is continuing her education getting a master's degree from the Colorado School of Mines in the fall. She enjoys sitting around doing nothing.



Joseph Yu, class of 22' of Baylor University majoring in Bioinformatics and on the pre-med track.

ACKNOWLEDGEMENTS

Metro Wastewater:

Josh Goldman-Torres,
JGoldman-Torres@mwrd.dst.co.us

Kate Newhart,
KNewhart@mwrd.dst.co.us

Amanda Hering PhD., Baylor University

Doug Nychka PhD.,
Colorado School of Mines

Aurora Waclawski

Sweta Rai