MWRD PAA - Preliminary Data Analysis

Kate Newhart

4/9/2019

# Introduction

The Robert W. Hite Treatment Facility, operated by the Metro Wastewater Reclamation District (MWRD) of Denver, CO, treats ~130 million gallons per day (MGD) of wastewater produced by ~2 million people from the Denver-metro area and is the largest wastewater treatment facility in the Rocky Mountain west. In an effort to reduce the cost of disinfection, a peracetic acid (PAA) system was installed to replace the existing chloramine system. However, due to variable influent *E. coli* concentrations to the disinfection system, it has been difficult to optimize the dosing of PAA to keep below *E. coli* limits of 126 (most probable number [MPN])/100 mL based on a 30-day geometric mean and 252 MPN/100 mL based on a 7-day geometric mean. In practice, PAA is overdosed to ensure that MWRD is meeting it’s discharge limit. The goal of this work is to identify correlations between upstream operating conditions in the secondary activated sludge system, *E. coli* concentrations, and PAA dosing.

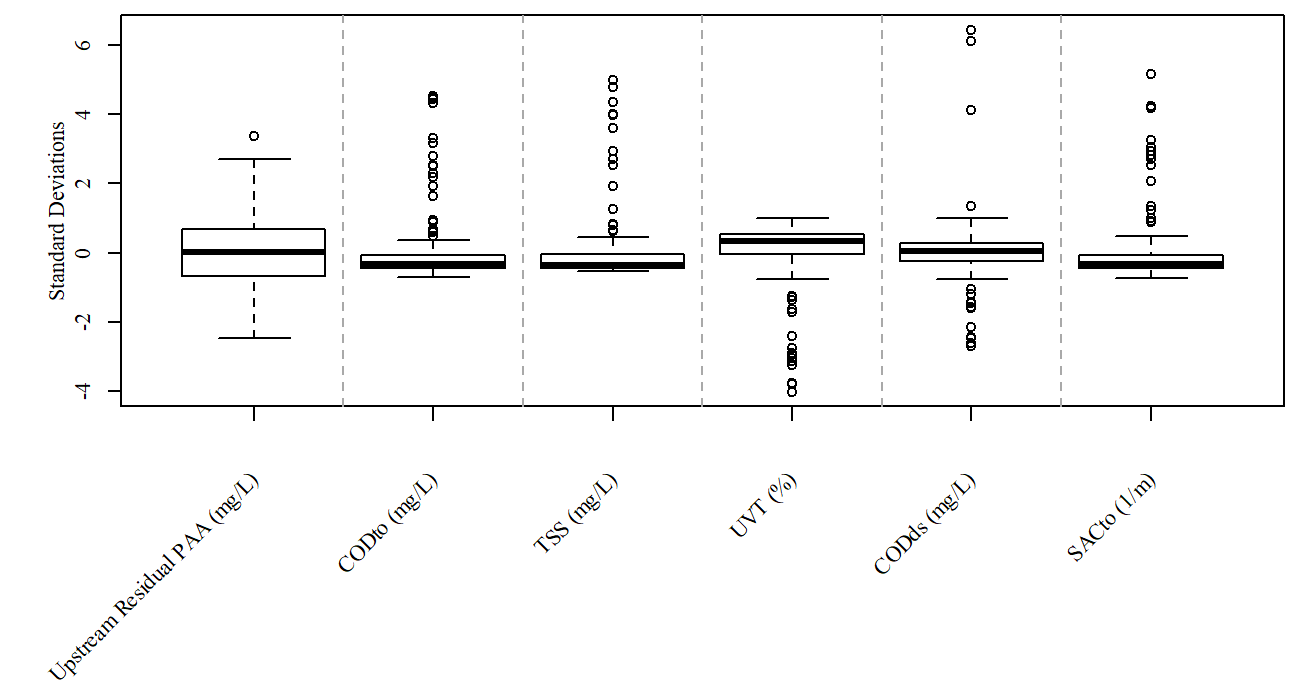
# Goals

Design a PAA disinfection dosing system that accounst for:

* Upstream secondary treatment performance
* Flowrate through the disinfection basin (i.e., hydraulic retention time or HRT)
* Flow conditions in receiving water body (e.g., low, mid-range, high, dry, moist)

# Questions

1. What effects pre-disinifection *E. coli*?
2. What effects PAA disinfection efficiency?

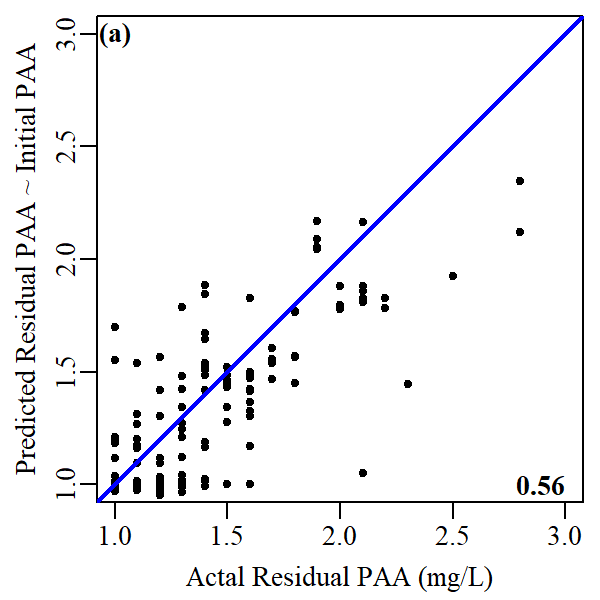
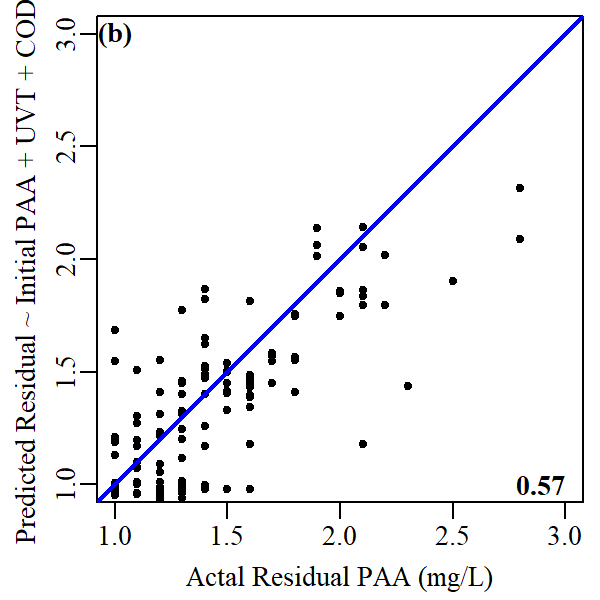


**Figure 1.** Scaled boxplot of online PAA data and Carbovis data

# Methods

## Linear regression

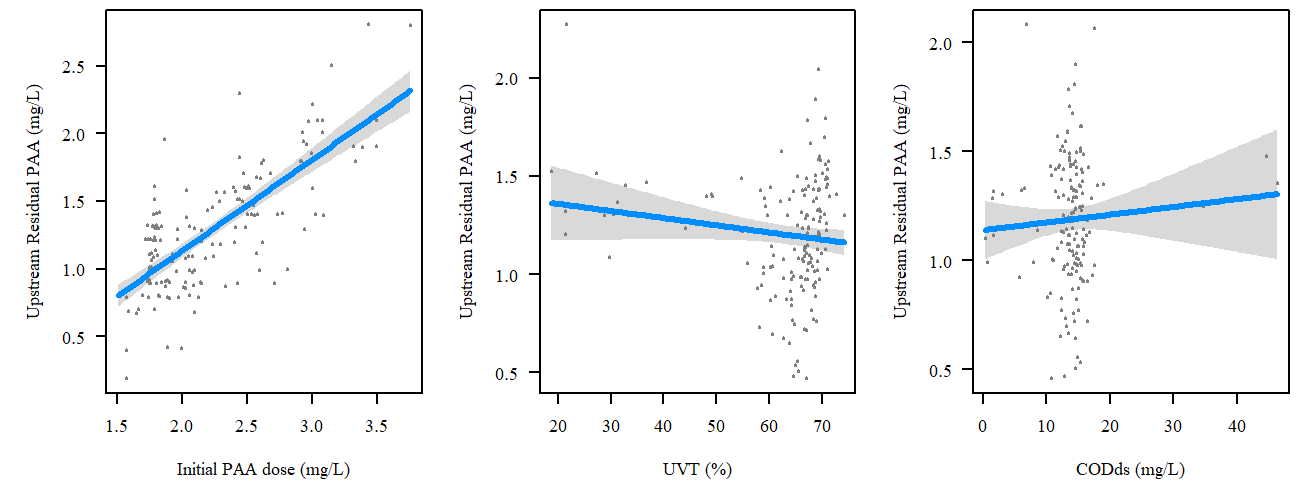
In simple linear regression, a model is constructed of a response variable (**Y**) that is a linear function of other variables (*xi*). The linear regression model assumes that **Y** is normally distributed, errors are normally distributed and independent, and **X** has constant variance.

**Figure 2.** Linear model fit for upstream residual PAA given (a) the initial dosing concentration of PAA and (b) the initial dosing concentration of PAA, UV transmitance, and CODds. Black circles represent actual observations. Blue line represents a perfect model fit. R-squared value in lower right. Figures show a strong relationship between initial PAA dose and the upstream residual, but linear impacts of other online water quality variables are minimal.

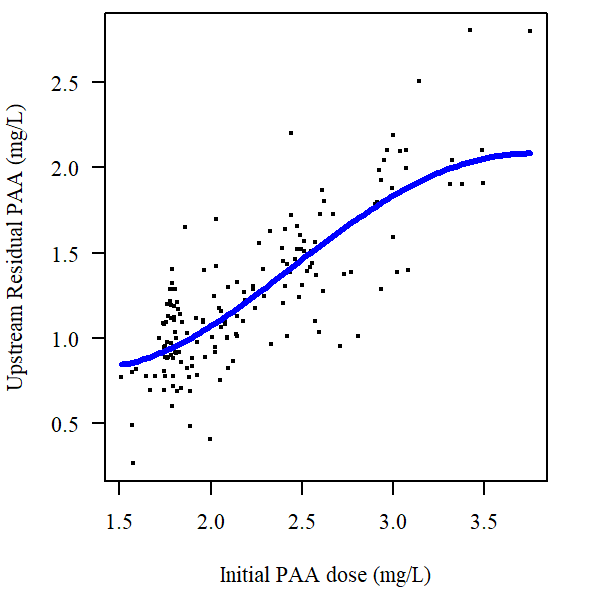
## Generalized Linear Model

The general linear model (LM) requires that the response variable follows the normal distribution whilst the generalized linear model (GLM) is an extension of the LM that allows the specification of models whose response variable follows different distributions. No difference was found between a LM fit with the initial calculated PAA dose, UVT, and CODds and a GLM fit with the same predictor variables.



**Figure 3.** GLM model constructed for upstream residual PAA concentration (mg/L) calculated from initial calculated PAA dose, UVT, and CODds.

## Non-Regression Model



**Figure 4.** Nonregression model (support vector machine) for predicting upstream residual PAA. # Supplementary Information ## Figures ## Tables