TECH BRIEF

Modernizing Water and Wastewater Treatment through Data Science Education & Research

Data Science Summer Fellows Program

Summer 2021

Aqua Aerobic Systems, Inc. – Modeling Backwash Volume of a Cloth Filter During Stormwater Events

Henry Burch, Baylor University
Kate Burgess, Baylor University
Tyler Krause, Colorado School of Mines

SUMMARY

Combined sewer systems in the US present capacity issues for wastewater treatment facilities. Organizations such as *Aqua-Aerobic Systems, Inc.* are addressing these challenges by inventing creative solutions such as their cloth media filter technology. By collecting online and lab data of filter performance, *Aqua-Aerobic Systems, Inc.* has enabled us to predict patterns in backwash volume during wet weather events. This briefing describes the primary treatment process, the data wrangling techniques, the predictive models, and the results of our statistical analysis. We generated linear and polynomial models, and in our analysis, the polynomial model predicted backwash volume per minute more accurately, with a 3.1% reduction (improvement) in the error (RMSE), compared to the linear model.

INTRODUCTION

The primary goal of this study is to predict the backwash volume per minute of the cloth filtration system using hydraulic loading rates, solids loading rates, and the use of coagulants.

The secondary goal is to analyze the overall percent of water used for backwash cycles during storm weather events.

FACILITY SYSTEM DESCRIPTION

Aqua-Aerobic Cloth Media Filters have provided the solution for high-performance tertiary treatment for more than 20 years. The OptiFiber Cloth Media is exclusively engineered for both water and wastewater treatment. OptiFiber is suitable for a variety of applications and provides maximum solids removal over a wide range of particle sizes [1].

Operation of the Aqua-Aerobic Cloth Media Filter is completed in three phases: filtration, backwash, and solids wasting. During filtration, influent water enters the filter tank and completely submerges the cloth filters. As water passes through the filters, solids accumulate on the cloth media. Over time, solids form a mat on the surface of the filters, resulting in inhibited flow and causing the filter tank level to rise. When the water reaches a predetermined level, a backwash cycle is initiated. In this backwash process, clean effluent water is pulled back through the filter collecting the solids with it. Backwash water is sent to a solids collection area while normal filtration continues.

DATA DESCRIPTION

All data were collected during twelve wet weather events via either lab data collection or through an online supervisory control and data acquisition sensor (SCADA).

Lab analysis was conducted on 24 aggregated samples for Events 3 - 12. During Event 2, 20 aggregated samples were collected and analyzed. Each aggregated sample was the average of three composite samples. The first eight samples were collected over 15-minute periods, the next four samples were collected over 30-minute periods, and the last 12 samples were collected over one-hour periods. Variables of importance in the lab data were the solids loading rate (SLR) (lbs/ft^2/day), total suspended solids (TSS) (mg/L), and coagulation addition. Due to the inconsistent sampling frequency, our team converted variables of importance into per-minute units.

The SCADA data were collected in five-second intervals. Each event occurred over a period from one to three days. The variables of importance from the SCADA data were backwash volume (gpm) and the hydraulic loading rate (HLR) (gpm/ft^2.)

EXPLORATORY DATA ANALYSIS

The first step in the data exploration process was to visualize the relationship between the different variables. Through some preliminary plotting, we were able to identify all outliers and unusual values in the data. In doing so, our team observed that many values in Event 1 differed substantially from the values and relationships seen in the other 11 events. For example, in Figure 1 the cumulative backwash volume spikes downward unlike other events. In order to maintain consistency in our work, we discarded the data from Event 1.

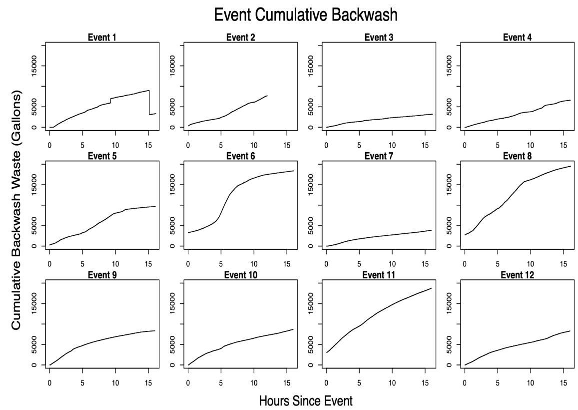

Figure 1: Influent Turbidity Of All 12 Events

Our next step was to explore the relationships between variables of importance in greater depth. Figure 2 displays how the percent of water used during backwashes changes over the hours since an event started.

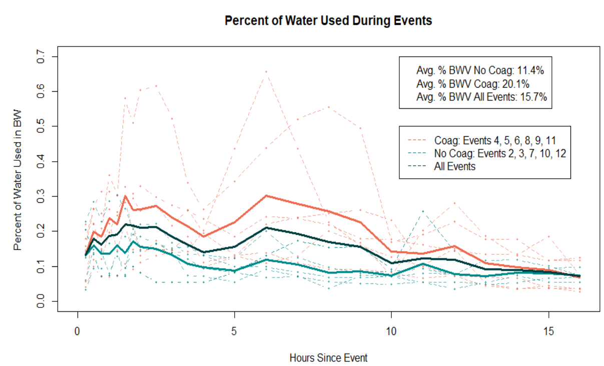

Figure 2: Percent of Backwash Used During Backwashes Over Time

This plot illustrates the general observation that the percentage of water used during backwashes decreases as time goes on. It also revealed that events not containing a coagulant have lower water percentage usage than events that do. These findings led us to investigate how specific variables affect the behavior of the backwash volume over time.

RESULTS and ANALYSIS

In order to predict the backwash volume per minute used during wet weather events, we utilized linear and polynomial regression. For both the linear and polynomial models, we used SLR, HLR, and coagulant as predictor variables. Predictor variables are independent variables used in regression modeling to help give information on a specified dependent variable. SLR and HLR were standardized by their standard deviation.

Linear Regression Model

To determine the accuracy of our linear model, we plotted the predicted versus true values, as seen in Figure 3, to show how close our model predicted the values for backwash volume per minute. The diagonal line on the plot represents the direct relationship between the predicted values and the true values, meaning that if a point falls on that line, then the predicted value equals the true value. The data points are spread out across the plot, with more accuracy for the lower values, but less accuracy for the higher values.

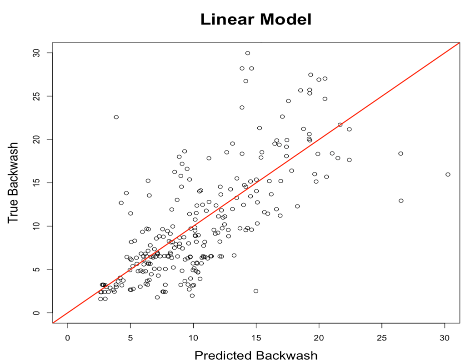

Figure 3: True vs Predicted Values for Linear Model

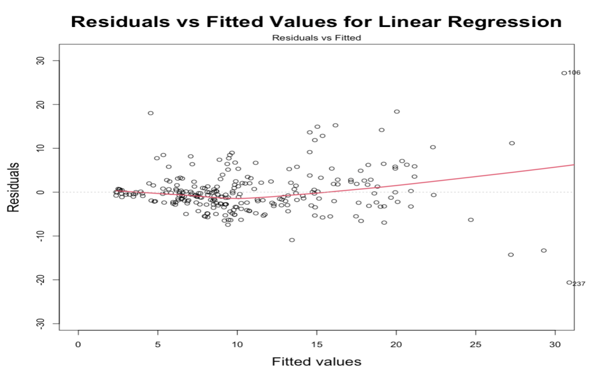

Figure 4: Residuals vs Fitted Values for Linear Model

The residuals versus fitted values plot in Figure 4 shows us the curvature of the data. If the model fit the data perfectly, the data would form a straight line, thus telling us there is room for improvement in the model.

The linear model's coefficients and corresponding significance values are shown in Table 1 below.

Predictor	Coefficient	Significance
SLR ($lbs/ft^2/day$)	0.67209	> 2e-16
HLR (gpm/ft^2)	0.26820	0.164
Coagulant	5.03978	6.54e-13

Table 1: Linear Model Coefficients and Significance Values

A coefficient can be described as a one unit increase in the variable that results in a unit increase or decrease in backwash volume per minute that is specified by the coefficient value. A coefficient with a smaller significance value is more significant in the model.

Polynomial Regression Model

The curved shape of the residuals versus fitted plot from the linear model led us to consider including higher order terms in our model. A polynomial model is linear in its coefficients, but it includes variable terms of higher degrees. After running the new polynomial model, we created the same plots as with the linear model as shown in Figures 5 and 6.

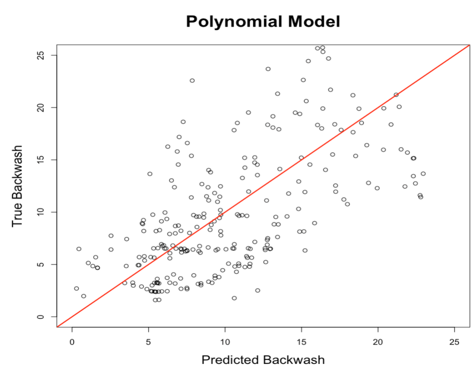

Figure 5: True vs Predicted Values for Polynomial Model

Figure 5 shows the predicted versus true values for the polynomial model. There is a more consistent distribution of the data points in Figure 5 compared to Figure 3, which means the model fits better throughout the entirety of the data set.

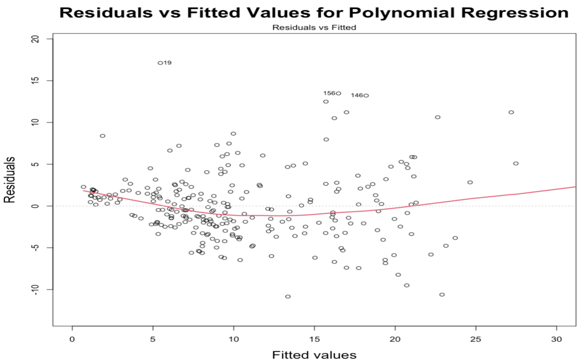

Figure 6: Residuals vs Fitted Values for Polynomial Model

Figure 6 shows us the residuals versus fitted plot, which further emphasizes the point that the polynomial model better fits the data. The important difference between this plot and the linear model counterpart is the more consistent distribution of the points. The polynomial model has a RMSE value of 5.29, which shows improvement over the linear model and reveals that the polynomial model more accurately predicts the true values for backwash volume per minute.

As with the linear model, we found the coefficients and their corresponding significance values as seen in Table 2 below.

Predictor	Coefficient	Significance
SLR^2 (lbs/ft^2/day)	3.35650	> 2e-16
HLR (gpm/ft^2)	30.41222	7.82e-05
SLR × HLR	3.08856	> 2e-16
SLR^2 × HLR	-1.36073	> 2e-16
HLR^2	-13.64166	5.71e-05
HLR^3	1.79099	8.58e-05
Coagulant	3.9929	1.07e-12

Table 2: Polynomial Model Coefficients and Significance values

After our models were finalized, we used their respective coefficients to create a spreadsheet that allows for *Aqua-Aerobic Systems, Inc.* to predict backwash volume per minute. The spreadsheet takes in inputs of SLR, HLR, and whether a coagulant is used or not, and it outputs a predicted backwash volume per minute value for each model. They can also adjust the confidence level that produces a measure of the uncertainty in the prediction. This spreadsheet is important because it enables Aqua-Aerobic Systems to give new clients a prediction of the backwash volume per minute given the client's SLR and HLR and helps to determine the appropriate size for the system.

CONCLUSIONS

Our first goal was to predict backwash volume per minute utilizing wastewater variables such as HLR, SLR, and the addition of coagulant. As seen above, the polynomial model predicts the backwash volume per minute more accurately than the linear model. To show our findings, our team designed an excel sheet that will allow the user to input HLR and SLR values and will use the models to calculate the predicted backwash volume per minute intervals.

One challenge we faced was analyzing data measured over inconsistent time intervals, and if feasible, future research would benefit from taking measurements that are equally spaced in time.

REFERENCES

[1] "AquaPrime® - Aqua-Aerobic Systems: Cloth Media Filtration." *Aqua*, 31 May 2019, www.aqua-aerobic.com/filtration/cloth-media/aquaprime/.

AUTHORS

Henry Burch

Studying Data Science and Statistics at Baylor University.

Kate Burgess

Studying Biochemistry and Computer Science at Baylor University.

Tyler Krause

Studying Statistics and Data Science at Colorado School of Mines.

ACKNOWLEDGEMENTS

We would like to give a special thanks to:
Amanda Hering PhD., Baylor University
Doug Nychka PhD., Colorado School of Mines
Tzahi Cath PhD., Colorado School of Mines
Luke Durell, Baylor University
Tayler Elwell, Colorado School of Mines
Hattie Means, Baylor University
Aqua Aerobic Systems, Inc. Professionals,
John Dyson, Product Channel Manager
Terry Reid, Director of Research and Development