Hazen

Exciting Applications of Machine Learning in Water Industry

June 8, 2023

Speaker: Micah Blate, PE

Co-author: Katya Bilyk, PE

Agenda

Machine Learning Overview

Wet Weather Flow Management at Neuse River RRF

Predicting Cake Solids

Summary

Machine Learning Overview

We've Been Able to Automate a Lot of Important Decisions at WRFs and the Result has been Lower Operating Costs and Better Effluent Quality

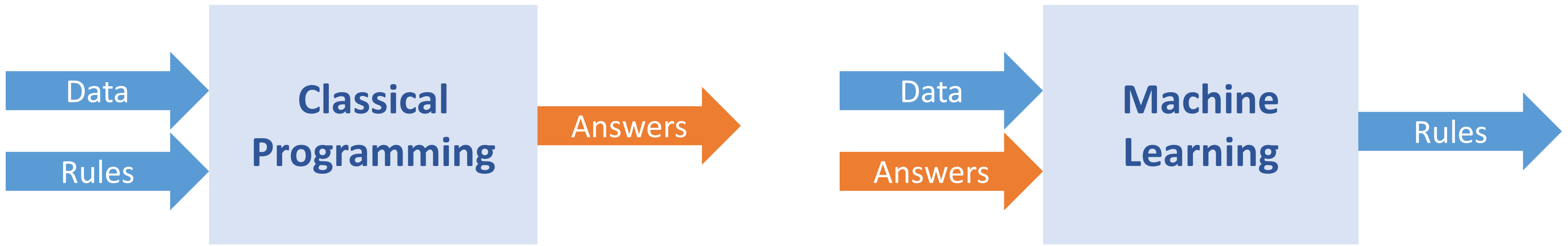
Program	Description	Goal
ABAC	Better match air demand to biological demand	Reduce blower demand and supplemental carbon demands
Aerobic SRT Control	Real-time SRT adjustment based on SRT, pH, DO	Improve efficiency of nitrification
Load Based Equalization	Operate EQ basin based on target ammonia loads (vs flow)	Improve efficiency of biological process
Automated Chemical Feed (Carbon)	Use online nutrient sensors for real-time control of chemicals	Optimize chemical dosage rates
Secondary Clarifier Guidance	Use online MLSS analyzer for real-time solids loading rate	Real-time prompts for No. of Clarifiers needed

Machine Learning has the Potential to Greatly Improve Operational Efficiency with Data-Driven Decision-Making Tools

- There are still a lot of decisions we make manually
- If we had models trained to real data, we could empower operators to make optimal decisions

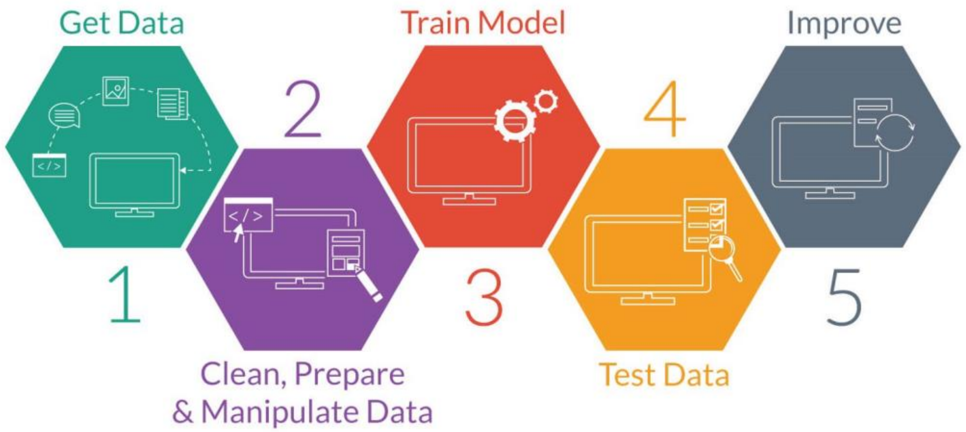
- Operator-made Decisions
  - Dewatering
    - How much polymer should I add?
    - What %TS can I expect?
    - Can I adjust my equipment to obtain better performance?
  - Wet weather
    - How high will the flow get during the next rainfall event?
    - How much additional volume will I have to treat?
    - When should I start flow to EQ. another basin, or store RAS?
  - Process
    - How many pounds of solids should I waste today?
    - How can I optimize nutrient removal, SVI, etc?

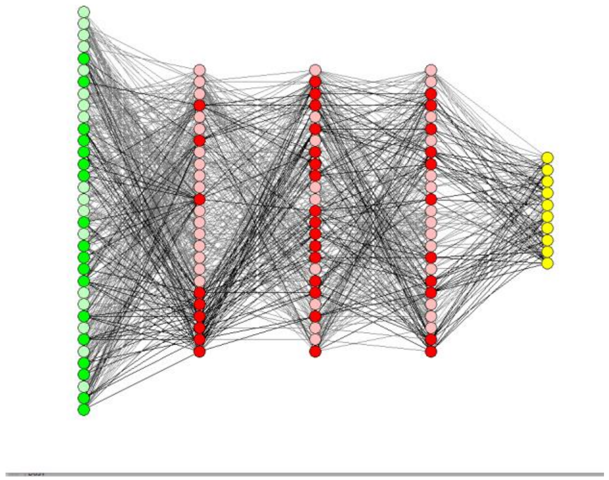
Machine Learning is Well Suited for Creating Predictive Tools because it can make Accurate Predictions without Explicitly Being Programmed to Do So



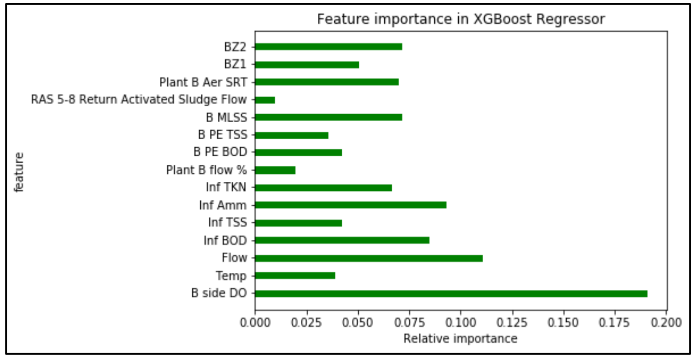
Machine Learning is an Alternative to Traditional Mechanistic Models

- ML uses algorithms, assign weights to independent variables, then seeks to minimize error in predicting a dependent variable
- Uses open source computer programming languages like Python
- Used in many fields including medicine, banking, finance, physics, etc.

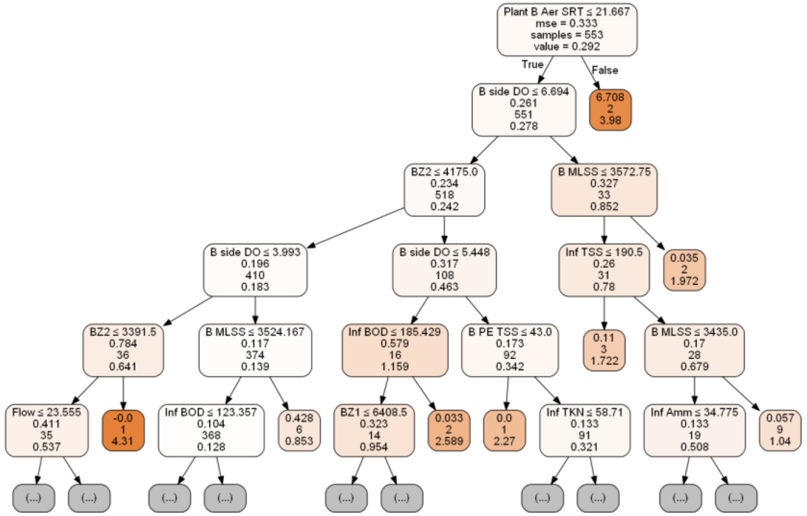




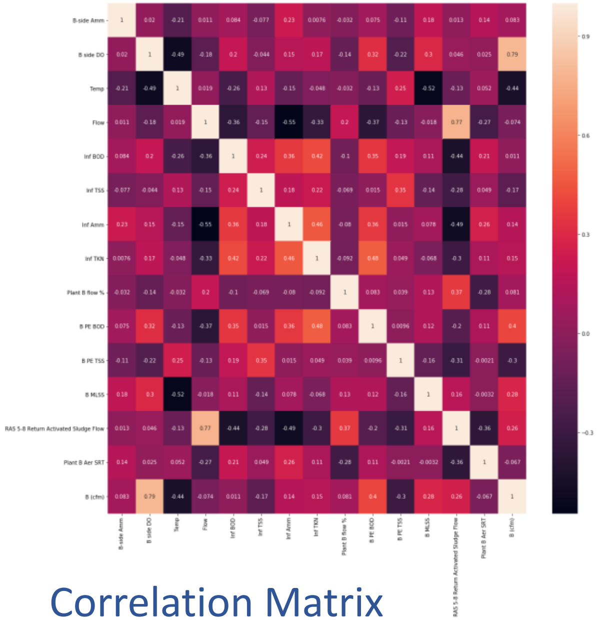
Examples of Machine Learning Tools



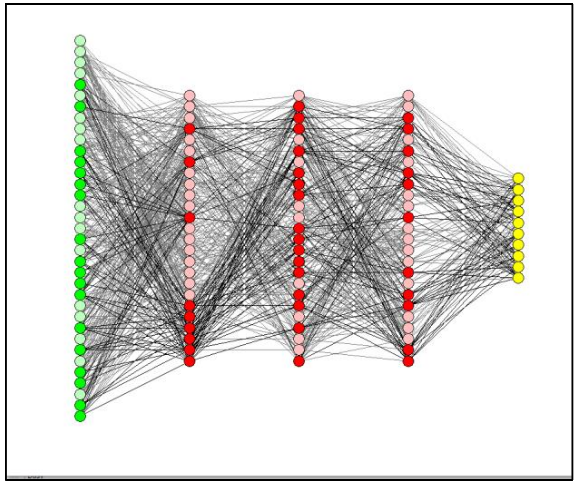
Feature Importance



Decision Trees



Correlation Matrix



Neural Networks (deep learning)

Steps to Deploying a Machine Learning Model

1. Obtain data to train the model
- Train the model
- Optimize model
- Develop schema for real-time data flow
- Develop real-time connectivity framework like SQL
- Develop user interface for model (Power BI)
- Connect all data pipelines
- Deploy model
- Review and revise model periodically

Wet Weather Flow Management at Neuse River RRF

Neuse River Resource Recovery Facility

- 75 mgd design capacity
- Nutrient limits
  - TN 3 mg/L
  - TP 2 mg/L
- Desire to harvest data to inform decision making process
  - Reduce operating costs
  - Improve quality and consistency of effluent
  - Automate decision making process

Current Wet Weather Standard Operating Protocol

Put 2 additional primaries online

- Provides 1.6 MG of EQ per clarifier

Put any additional BNR basins online

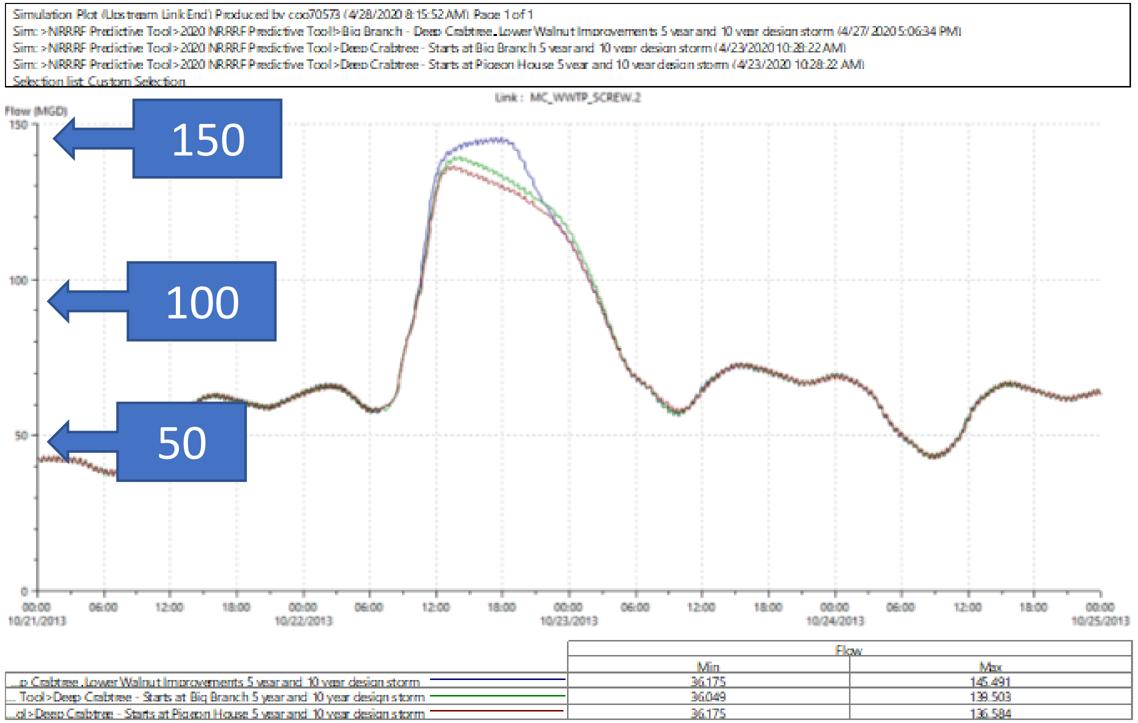
- Provides 6.4 MG of EQ per tank and 1-2 are usually available

Divert to EQ

- If 24- hour sustained flows > 150 mgd
- Subject to SVI and guidance program

Why did Current Strategies Fall Short?

- Currently staff use pump station data to estimate peak flow and have 30-60 minutes of advance warning
- Flow monitors in collection system aren't predictive
- Doesn't tell you if flows will increase or decrease
- City has a calibrated collection systems model but no way to currently utilize that tool in a real-time fashion

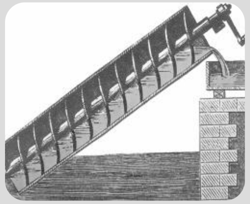


Collection system model output, manually generated.

Machine Learning Approach was Developed to Predict Flow up to 72-hours in Advance

Rainfall

Streamflow

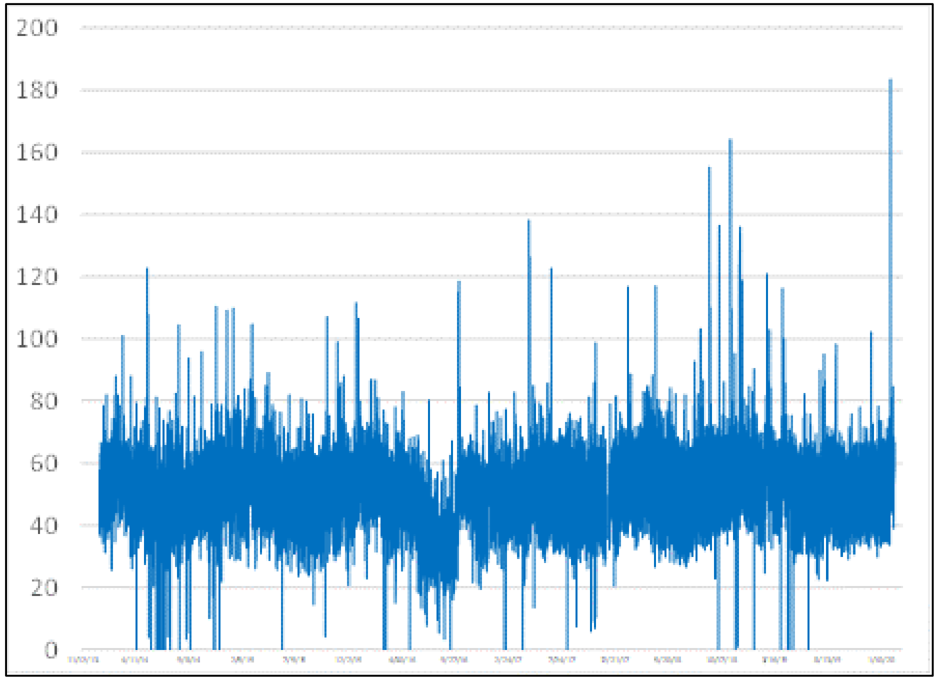


Past Influent Flow to NRRRF

Hour of Day

Collection System Improvements

Used python machine learning algorithms to train a model to 6+ years of influent flow data as a function of explanatory variables.

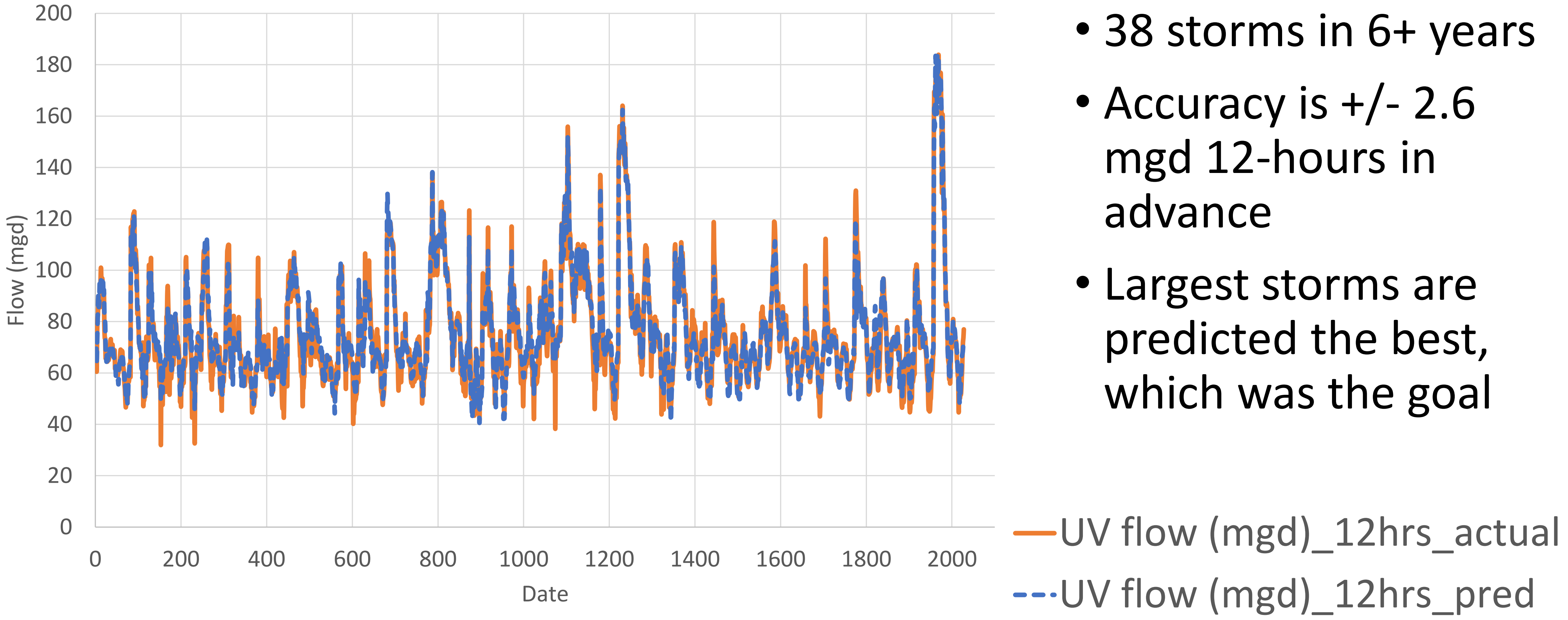


Sustained flows of 184 mgd experienced

Challenge meeting effluent TN and TP during wet weather events

Only 30-60 minutes of advance warning prior to this project

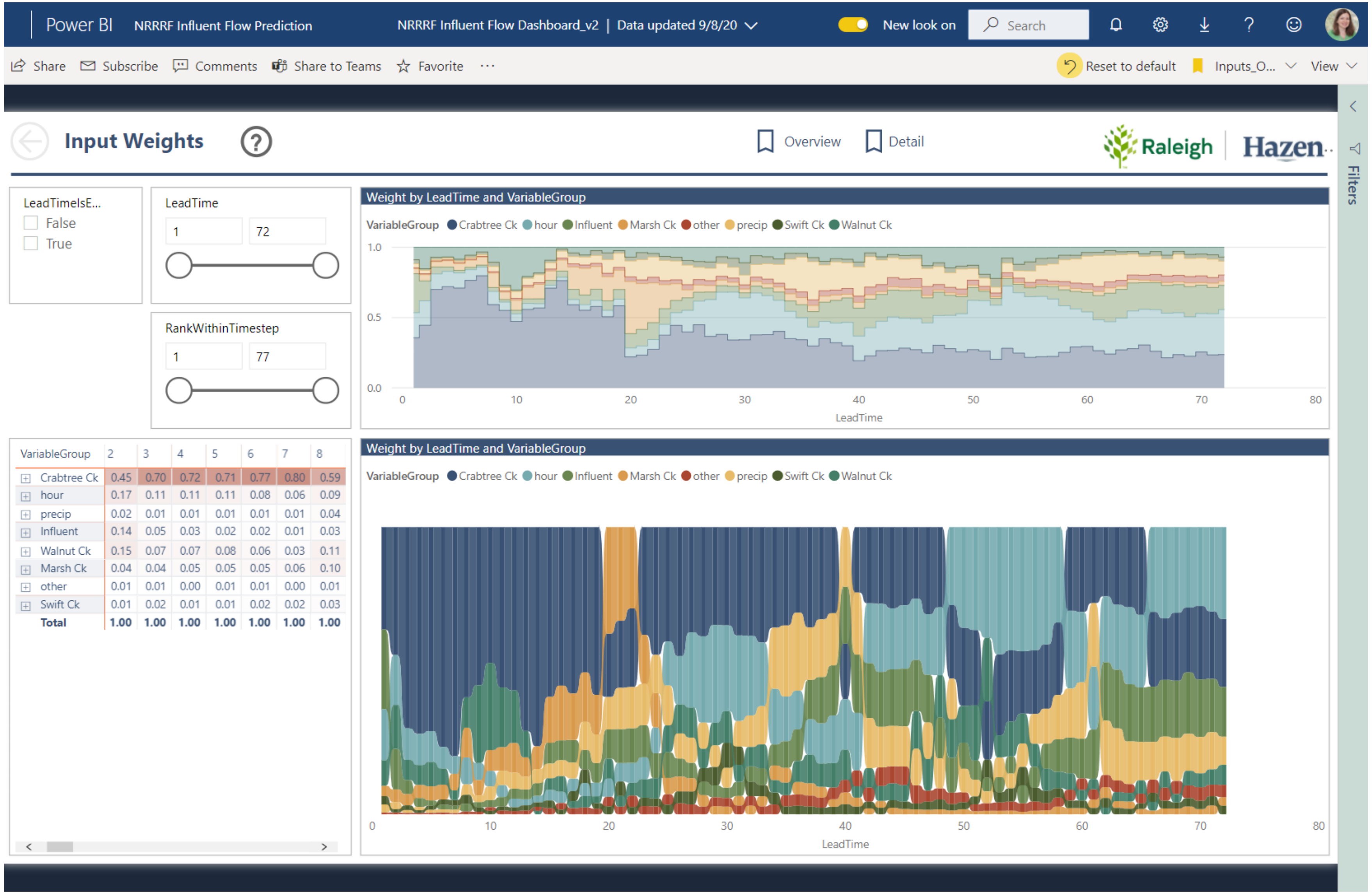
All Storms Predicted with Good Precision by the Model During Training



- 38 storms in 6+ years
- Accuracy is +/- 2.6 mgd 12-hours in advance
- Largest storms are predicted the best, which was the goal

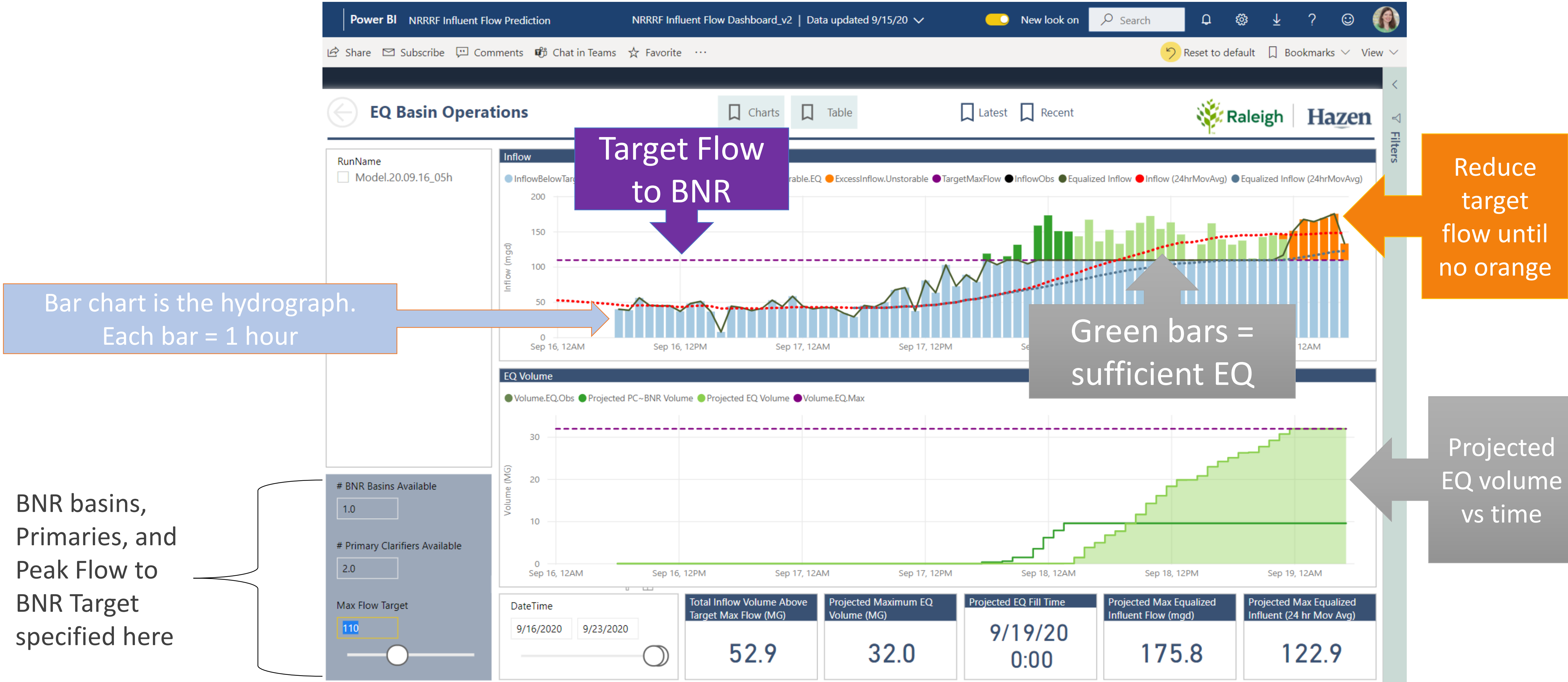
The Selected Model Is Fully Transparent

- Input weights screen shows how much influence each variable has
  - Periwinkle = Crabtree Creek streamflow
  - Light blue = hour of the day
  - Yellow = precipitation
- X-axis is time step from 1 to 72 hours
- Y-axis is percent influence



Model Prediction Screen – Updated Hourly

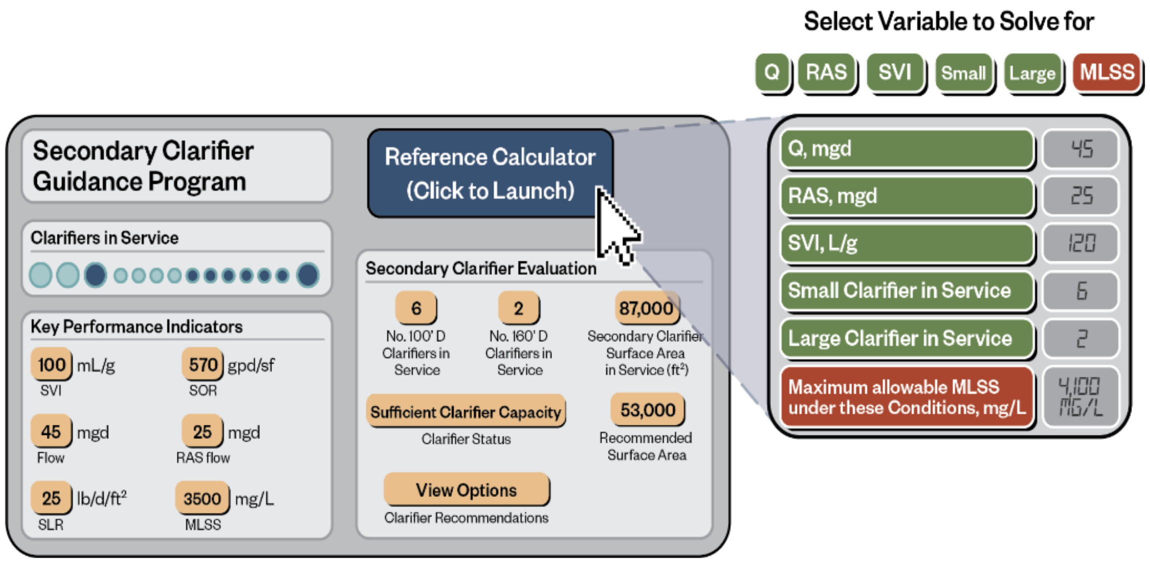
Hydrograph Incorporated into Dashboard for Plant Staff to Refine Operational Decisions Related to Wet Weather Management



BNR basins, Primaries, and Peak Flow to BNR Target specified here

Neuse River Resource Recovery Facility (NRRRF) Secondary Clarifier Guidance Program

- Linear regression equation derived from many SPAs
- Estimates required # clarifiers
- Calculator can solve for 5th variable



$$C\ [ft^2] = 981*Q + 909*SVI - 530*Q_{RAS} + 34.2*MLSS - 193{,}090$$

Secondary Clarifier Guidance Program Screen Allows Real-Time Determination of Secondary Clarifiers And RAS Flow Needed

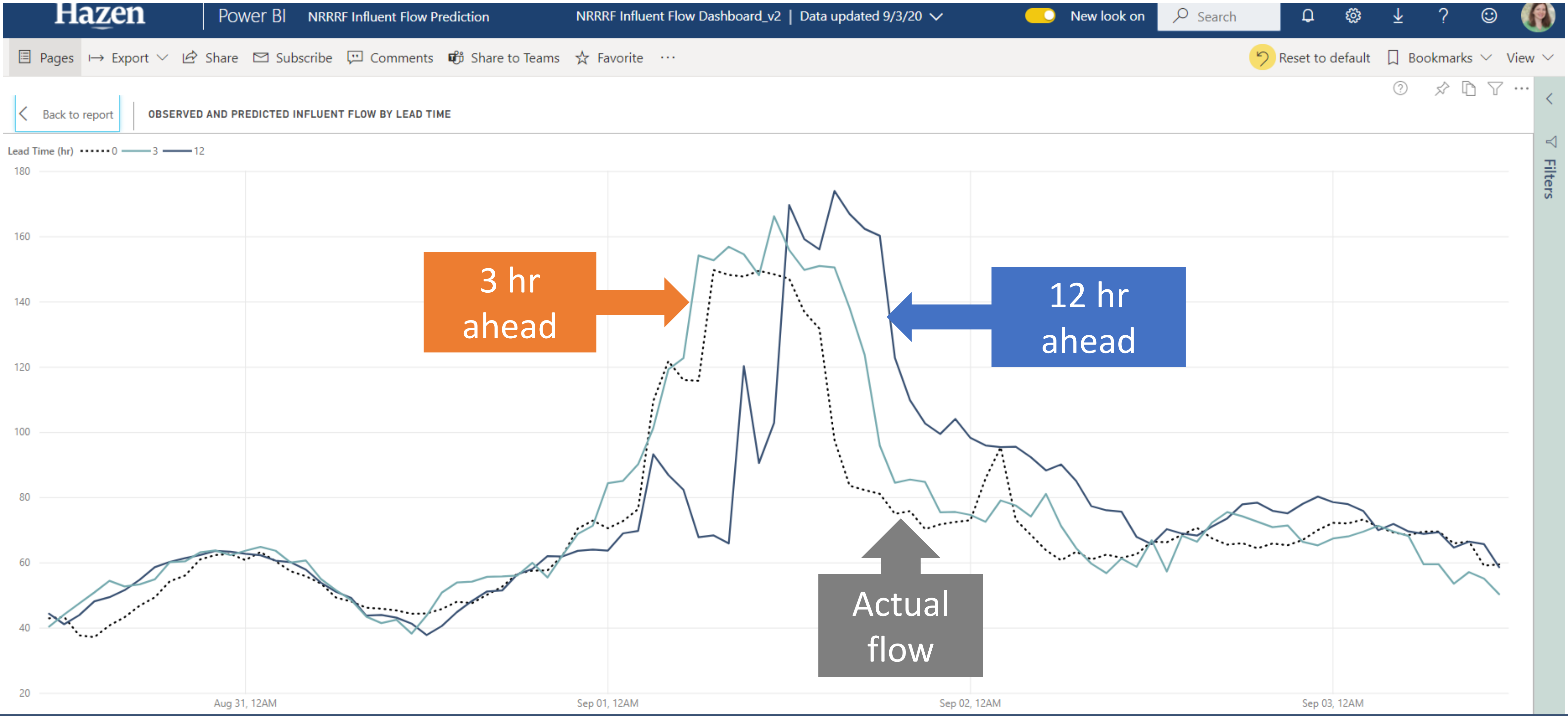
Left – displays key performance indicators for past 72 hours.

Top center – displays past flow (blue colors), projected flow (green), and maximum allowable flow (red) with all secondary clarifiers in service.

Right – calculator tool that allows operators to solve for any variable

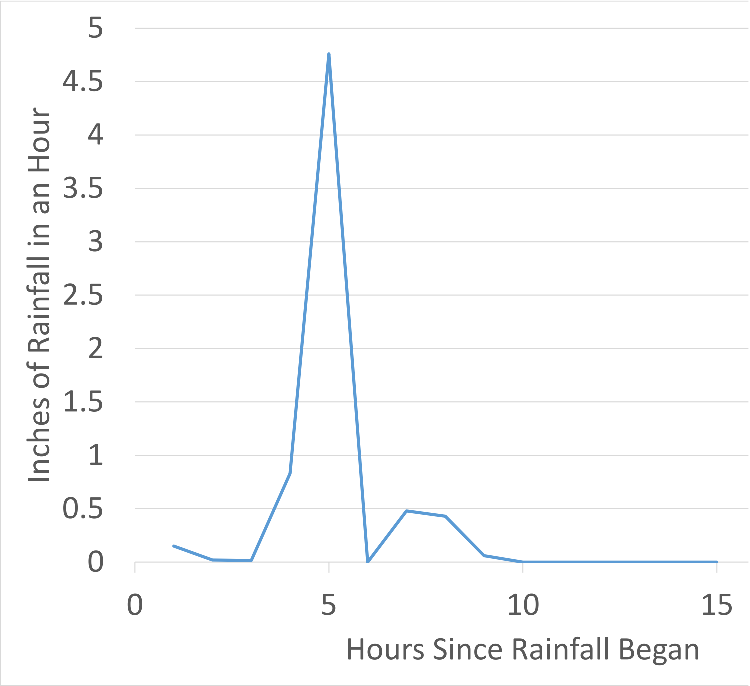
Bottom center –KPIs and combinations of small and large clarifiers that meet the criteria in the calculator tool.

Another Good Recent Prediction for a Recent 150 mgd Wet Weather Event (6.7" Rain in 9 hours) Was Well Predicted



NRRRF Maintained Its Excellent Effluent Quality During This 6.7” Rainfall Event

- 148 mgd peak hour flow
- 3.1 flow peaking factor
- 6.7” rain in 9 hours

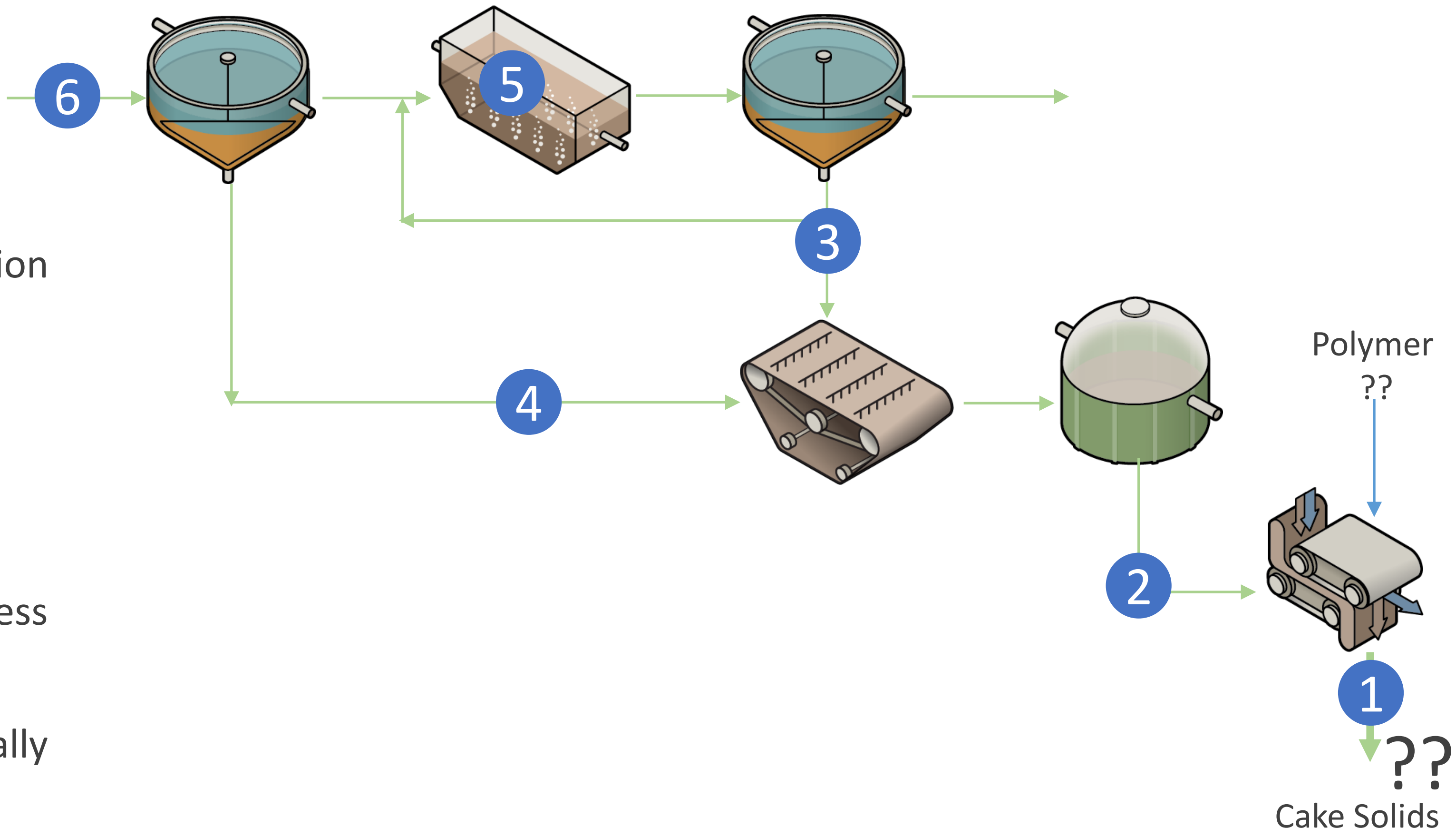


Date	Influent Flow (mgd)	Effluent Flow (mgd)	Effluent TSS (mg/L)	Effluent TP (mg/L)	Effluent Ammonia (mg/L)	Effluent TN (mg/L)
8/31	54	-	BDL	0.14	BDL	-
9/1	106	72	-	-	-	-
9/2	68	61	BDL	BDL	BDL	1.9

Dewatering Case Study

The journey sludge took to reach dewatering is very important

- Types of water associated with floc
- Floc bound water capacity (g Water/g VSS)
  - **Associated with particulates**
  - **Associated with colloidal material**
- Free ion (divalent cations) composition (charge and bonding capacity)
- VS/TS ratio
- Digestion chemistry
- Mechanistic modeling is still being developed (Sumo by Dynamita Process Modeling)
- Some needed data plants don't usually collect



Exploratory Questions: Is it possible to predict the cake TS% as a function of past data trends? What variables contribute to this prediction?

Identify Potential Parameters

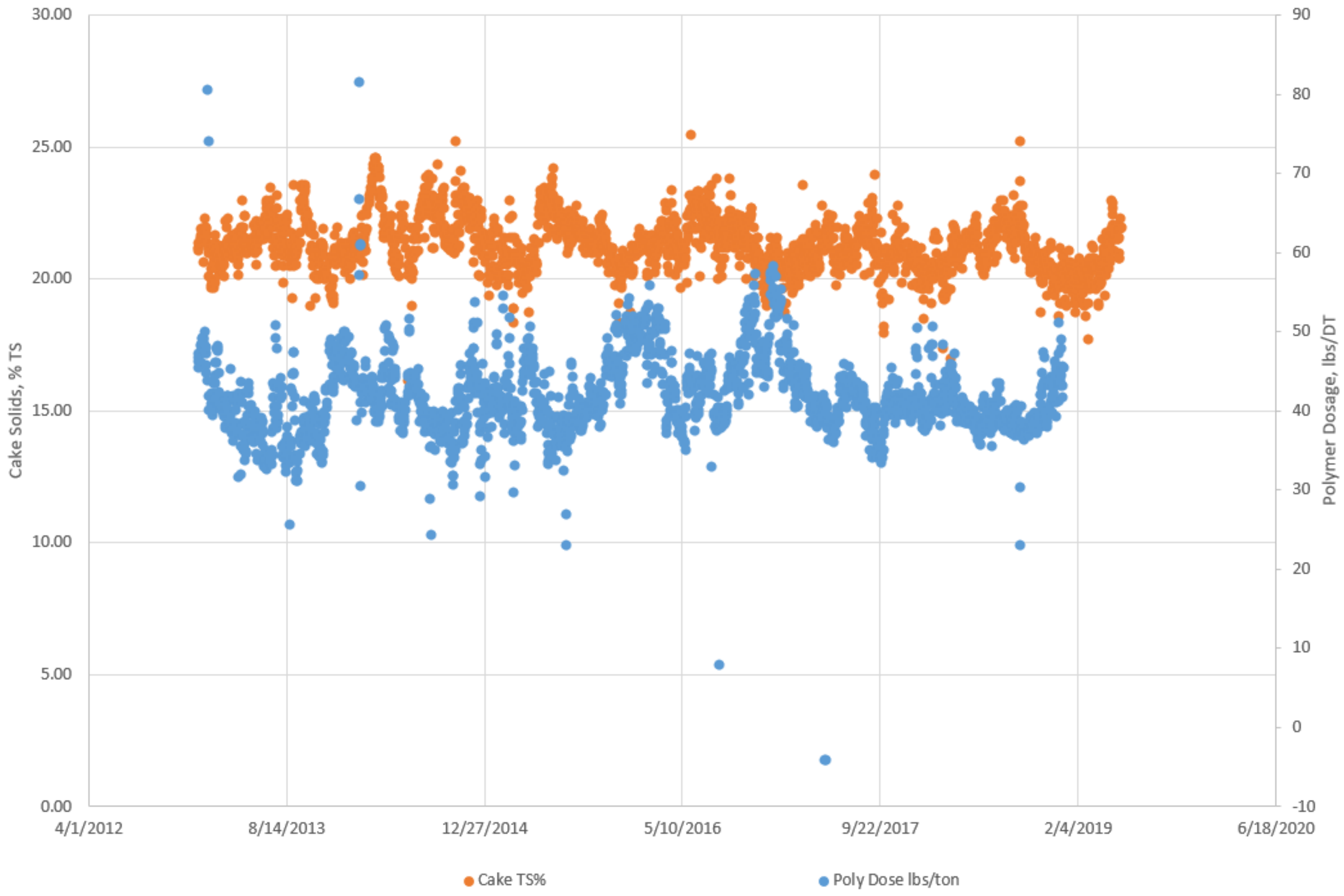
Evaluate/ Analyze Parameters

Develop Predictive Tools

Iterate and Refine Tools

Machine learning can use the history of the sludge to predict dewatering

- We sought a dataset with reliable historical data, spanning many years, with significant variation in % TS
- Explored whether different machine learning models could be used to find an empirical relationship between explanatory variables and dewaterability



Exploration of Explanatory Variables to Predict %TS

Parameters believed to potentially impact dewaterability

Current Values

Polymer Dose

Influent:

- Flow
- TKN
- BOD
- TSS/VSS
- Temperature

Secondary System

- SVI
- MLSS
- RAS Flow/Conc

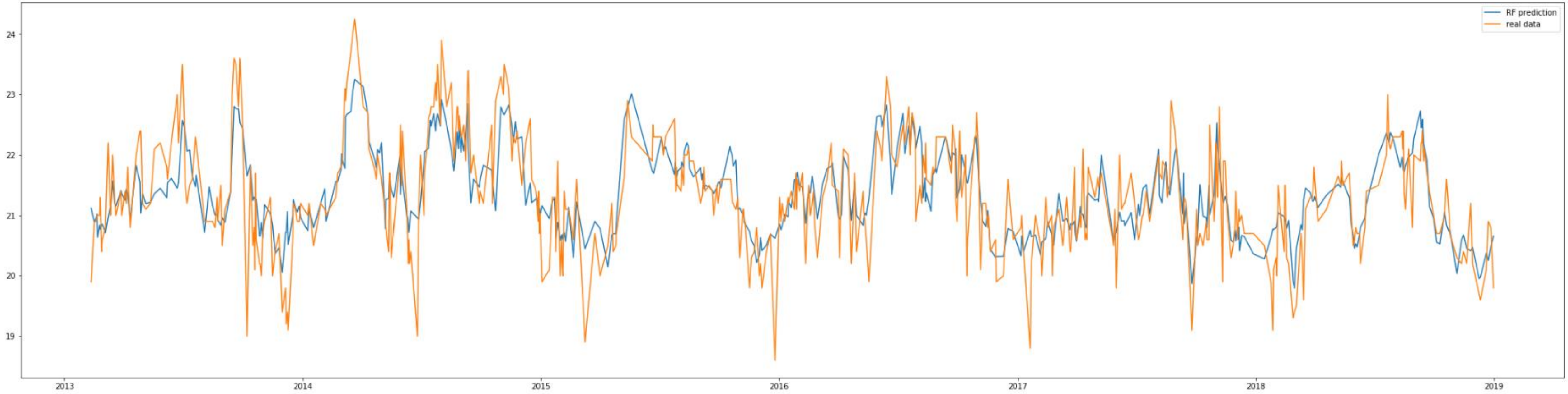
Solids Processing

- Digester VSR
- HRT/SRT
- Influent Load
- TPS/WAS Ratio
- CNash

Trends also Important as Past Actions Influence Solids Treatment

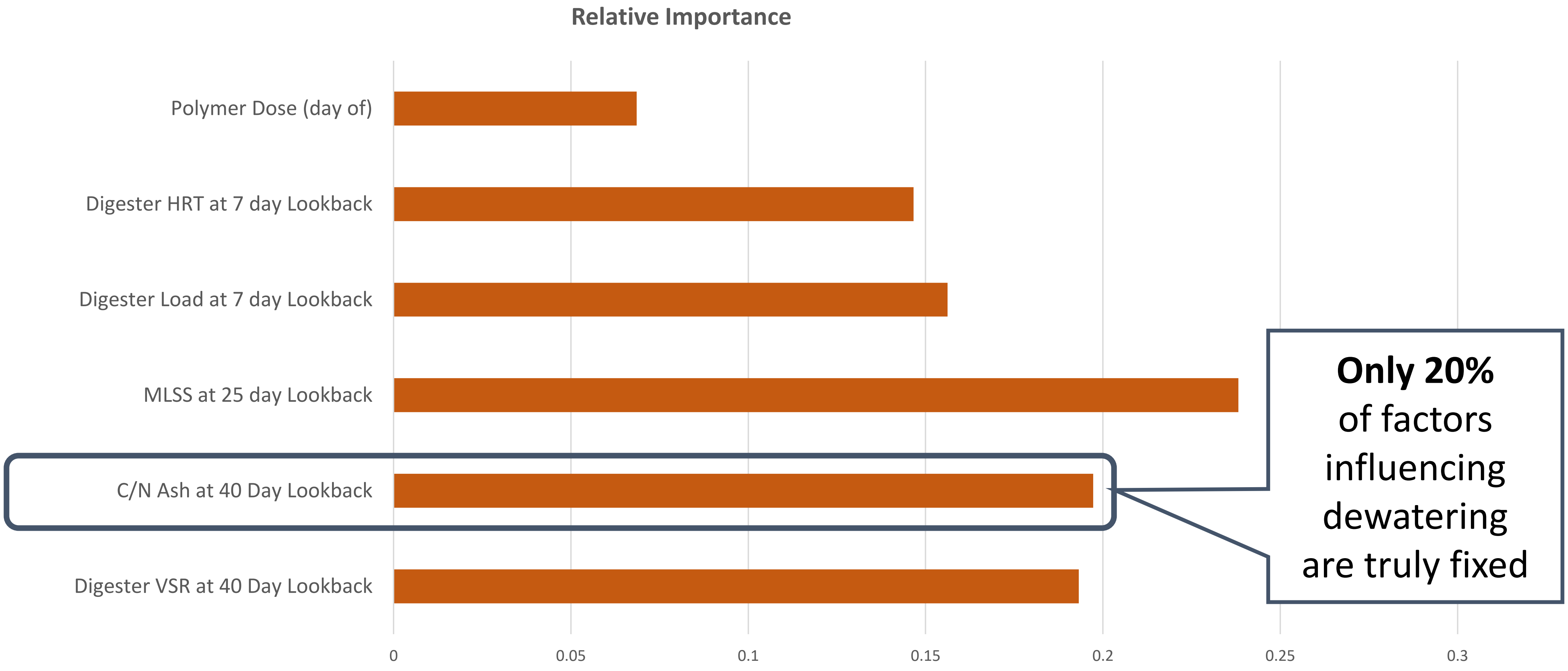
- 7-day
- 14-day
- 25-day
- 30-day
- 40-day
- Applied to all variables except polymer dose

Random Forest Prediction was Most Accurate

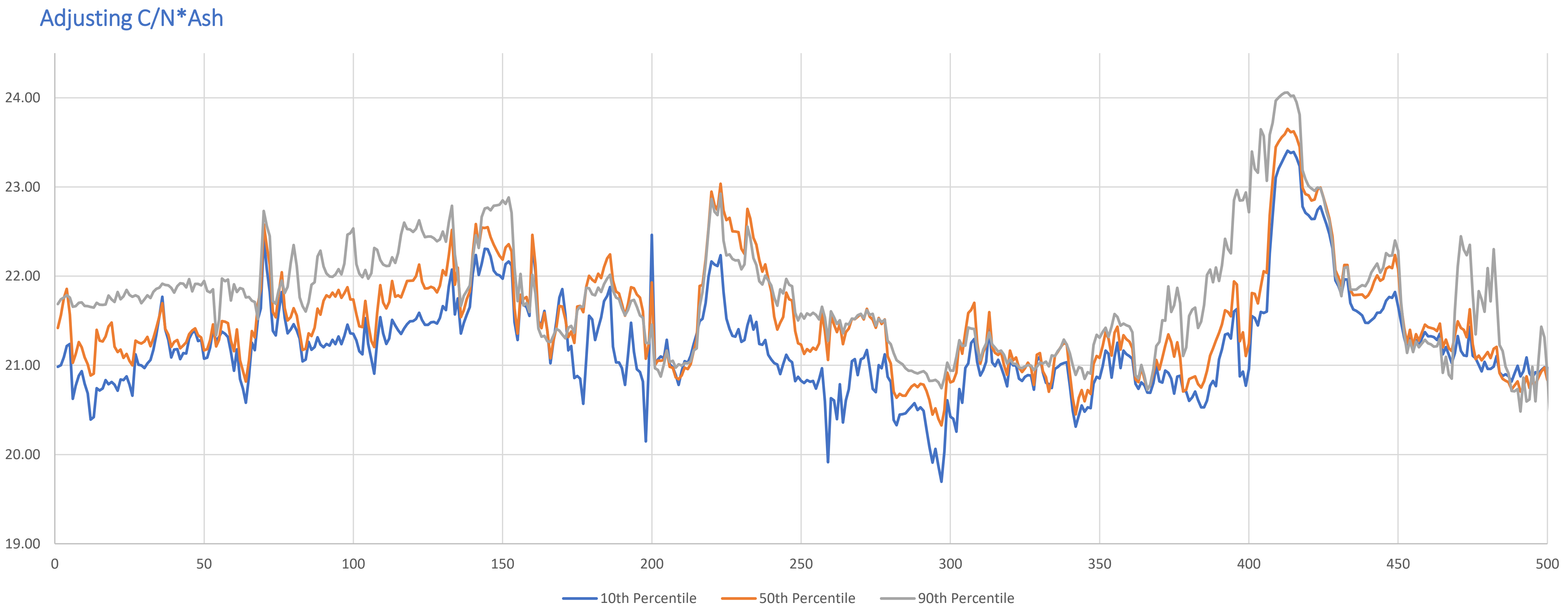


Parameter	Unit
Mean Absolute Error	% TS: +/- 0.4%

Key Variables Predicting Dewaterability and Their Relative Importance



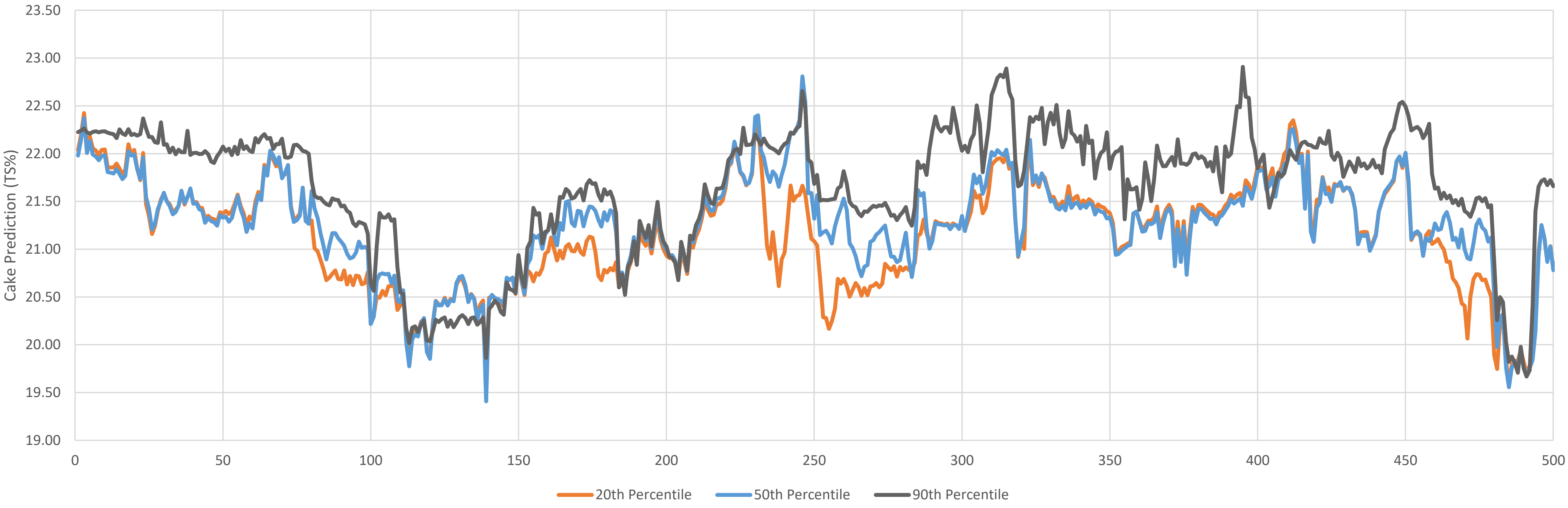
Sensitivity Analysis on C/N*Ash Shows Expected Relationship that a Higher Ratio → Higher %TS



*The model predicts improved dewaterability with a higher C/N*ash ratio, which is consistent with research*

Model Also Shows that the Longest HRT → Higher %TS

Adjusting Digester HRT



Model developed at this plant suggest that longer HRT (potentially more volatile destruction) leads to better dewaterability

How Would This Tool be Used In Real Life?

Big Picture Insights

- Learn how your plant behaves
- Verify those conclusions are sound
- Iterate and revise model until the conclusions make sense

Planning

- Estimate annual operating costs
- Identify potential efficiency losses*
- Identify seasonal trends

Day-to-Day

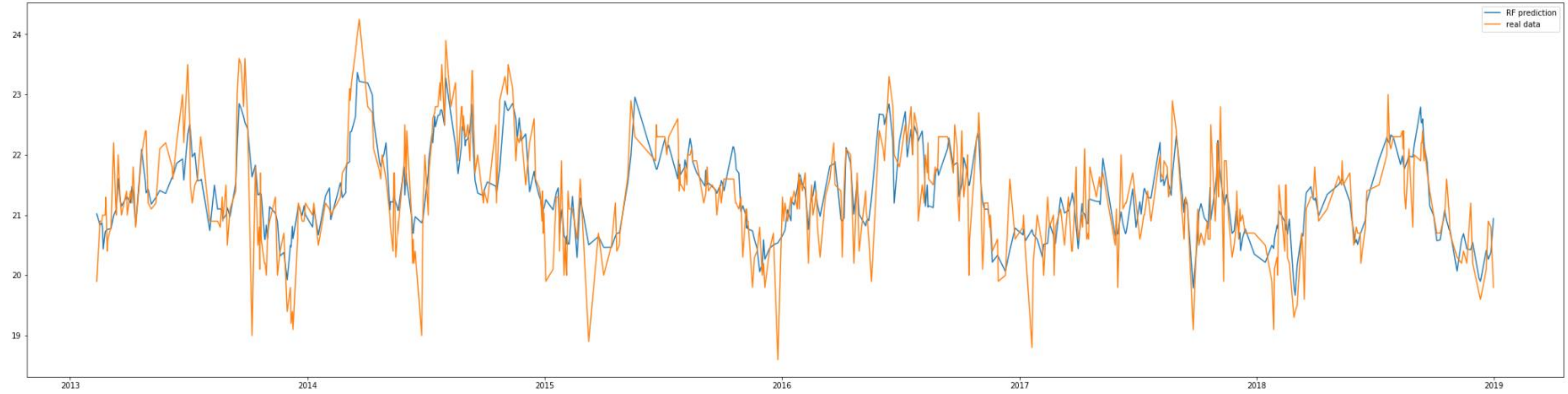
- Predict %TS
- Optimize dewatering machine settings
- Increase lb polymer/DT if low %TS expected

*For example, %TS is lower than model predicts, HRT in digester is the same, but perhaps mixing became less efficient, resulting in a change in state (the role/importance of HRT).

Summary

Machine Learning Solutions Offer Great Promise in the Water Industry

- Can quickly and easily train models with plant data to understand complex relationships
- Tools can be deployed in Power BI to compliment existing operating system
- Tools can save money and optimize treatment
- Machine learning can be paired with digital twins for enhanced utility



Questions

Micah Blate, PE

mblate@hazenandsawyer.com