Metropolitan Water Reclamation District of Greater Chicago

Welcome to the February Edition of the 2021 M&R Seminar Series

NOTES FOR SEMINAR ATTENDEES

- All attendees' audio lines have been muted to minimize background noise.
- A question and answer session will follow the presentation.
- Please use the Chat feature to ask a question via text to All Panelists.
- The presentation slides will be posted on the MWRD website after the seminar.
- ISPE has approved this seminar for one PDH. Certificates will only be issued to participants who attend the entire presentation.

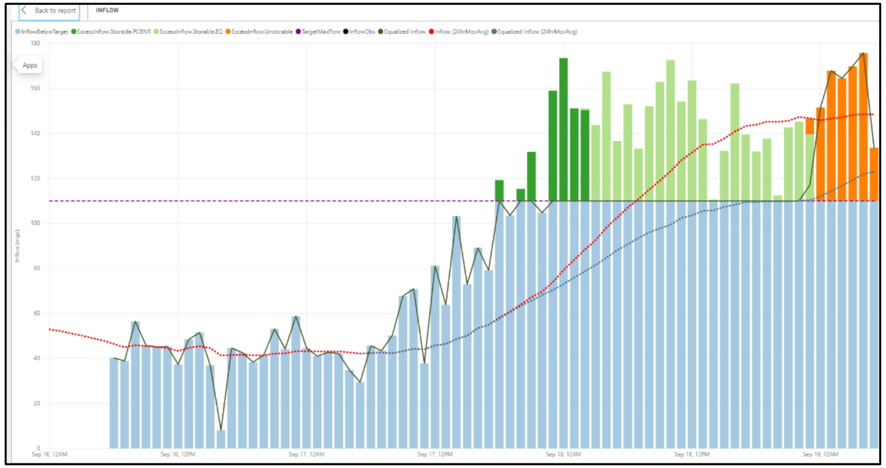
Katya Bilyk, P.E.

Senior Associate

Hazen and Sawyer, Raleigh, North Carolina

Katya Bilyk is a Senior Associate for Hazen and Sawyer in Raleigh, North Carolina. She received a bachelor of science degree in civil engineering from Virginia Tech and a master of science degree in environmental engineering from the University of North Carolina, Chapel Hill. She has 20 years of experience in the industry and focuses on wastewater process design, modeling, and optimization. Ms. Bilyk is actively involved in WEF activities related to these topics and has published and/or organized more than 40 papers and workshops on nutrient removal. In recent years she has used Python software to apply machine learning and advanced data analytics to the water industry. She is a professional engineer licensed in North Carolina, Virginia and New York.

Hazen



Hazen Predict

Predictive Analytics: The Next Step in Water Resource Recovery Facility Operation and Optimization

Presenter: Katya Bilyk, PE, Hazen and Sawyer
February 26, 2021

Agenda

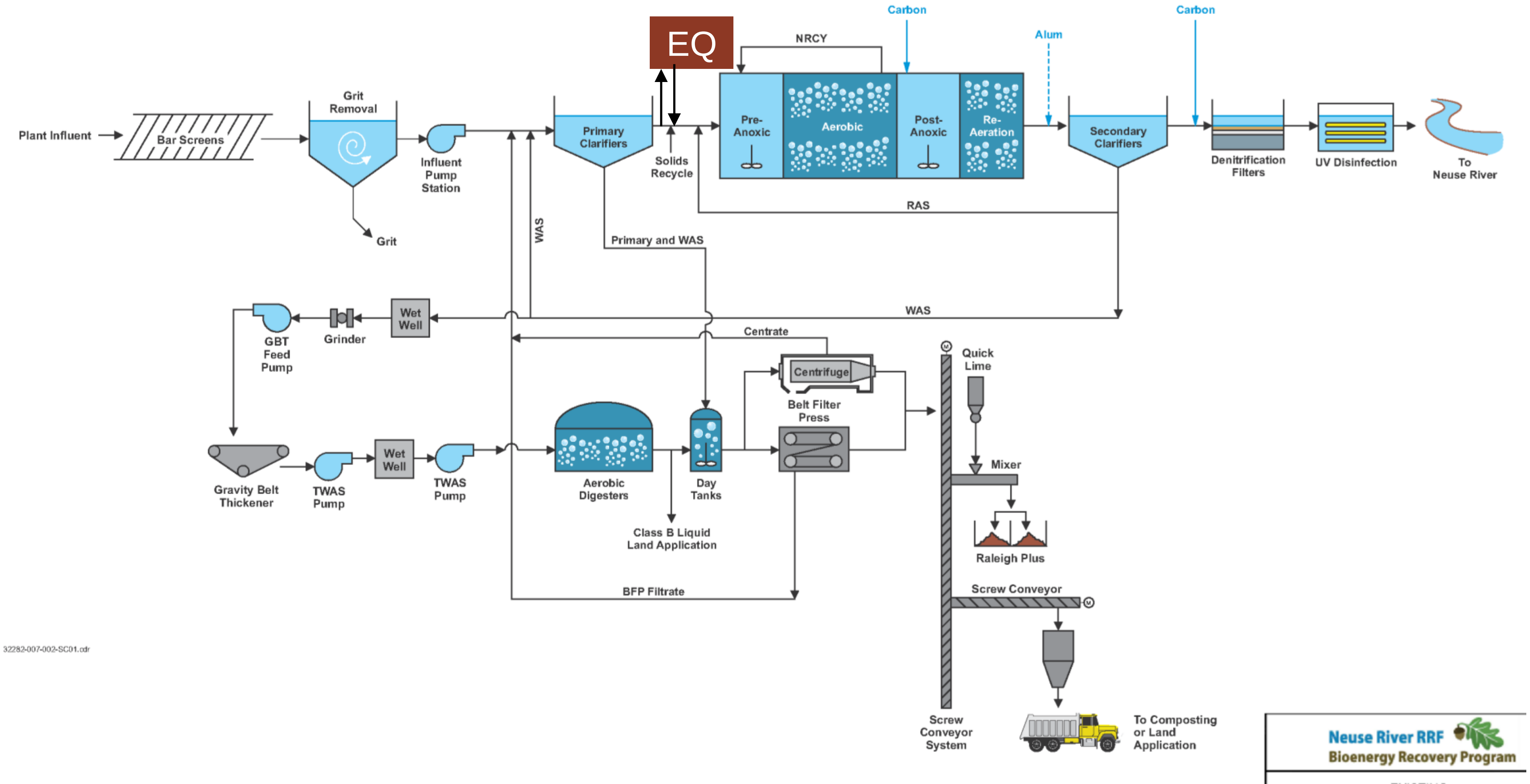
- Overview
- Model development
- Model deployment
- Interactive tool Raleigh Water staff use to view model results
- Model performance
- Summary
- Other opportunities

Overview

The NRRRF is Located in Raleigh, NC, Permitted To Treat 75 mgd, And Must Meet Strict Nutrient Limits

- Annual Average, Load-Based TN Allocation
  - **Current TN Allocation: 687,373 lbs/year**
  - **3 mg/L TN at 75 mgd**
- Quarterly average TP limit
  - **2.0 mg/L**
- Monthly average NH3-N limits
  - **1.0 mg/L summer / 2.0 mg/L winter**
- Stringent BOD5 limits

NRRRF Utilizes A 4-Stage Biological Nutrient Removal Process



Real Time Process Control Programs Were Implemented in 2018 To Optimize Operations

DO / Ammonia-Based Aeration Control	Load Based Equalization	Automated Chemical Feed (Carbon / Alum)	Clarifier Optimization Program
Better match air demands to biological demand	Operate EQ basin based on target ammonia loads (vs flow)	Use on-line nutrient sensors for real-time control of chemical dosages	Use on-line MLSS analyzer for real-time solids loading rate
Reduce blower demand and supplemental carbon demands	Improve efficiency of biological process	Optimize chemical dosage rates	Real-time prompts for # of Clarifiers needed
Feb 2018	Sept 2018	Nov 2018	Fall 2018

Real-time Process Controls Have Improved Effluent Quality, Reduced Operating Costs, and Automated Routine Decision Making

ABAC, Ammonia-load EQ, and Nutrient-Paced Carbon Feed

TN Reduction	2.2 mg/L → 1.8 mg/L
Air demand reduction	10%
Methanol reduction	>40%
Reduction in number of blowers	1 – 2250 hp unit
Electrical savings	$300,000/yr
Methanol savings	$200,000/yr
Total savings	$500,000/yr

2017 - 2019

ROI for Raleigh Water on Real-time Process Controls < 1 Year

- Real-time process controls were implemented in 2017
  - **Instruments - $124,000**
  - **Integration - $191,000**
  - **Engineering - $0**
  - **Total investment - $315,000**
- **ROI < 1 year**

- Electrical savings $300,000/yr
- Chemical savings $200,000/yr
- TN reduction from 2.2 to 1.8 mg/L
- Nitrogen credits not used valued at $1.3M

Predictive Analytics Using Machine Learning Was Identified to Improve Operational Efficiency During Wet Weather Events

- 75 mgd
- Average daily flow of 48 mgd
- Hydraulic capacity of 225 mgd
- Highest hourly flow recorded 184 mgd
- 32 MG EQ basin

- Aim of the project was to predict influent flows 72-hours in advance
  - **How can we best leverage our existing infrastructure to optimize treatment during a wet weather event?**
  - **How can we tie-in this program with other wet weather management programs like secondary clarifier guidance program?**

- Deliverable = ML-driven predictive tool Raleigh Water interacts with via Power BI

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

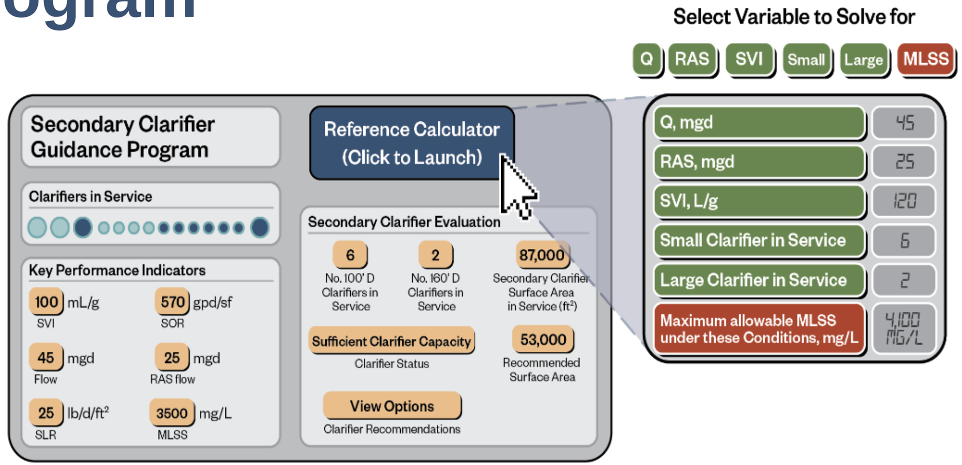
- Provides 6.4 MG of EQ per tank and 1-2 are usually available

Divert to EQ

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

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

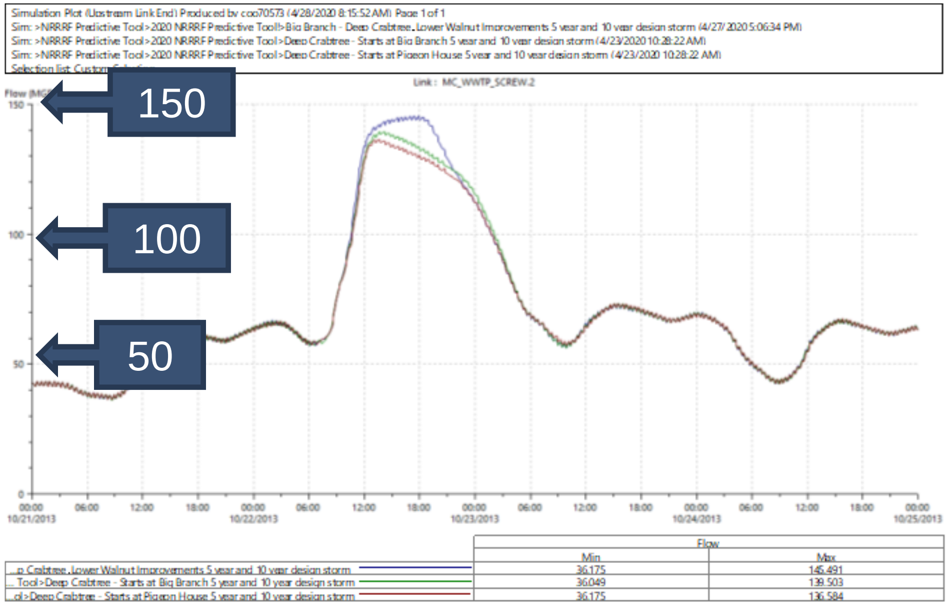
- Linear regression equation derived from many SPAs
- Estimates required # clarifiers
- Calculator can solve for 5th variable



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

Why did Other Strategies Fall Short?

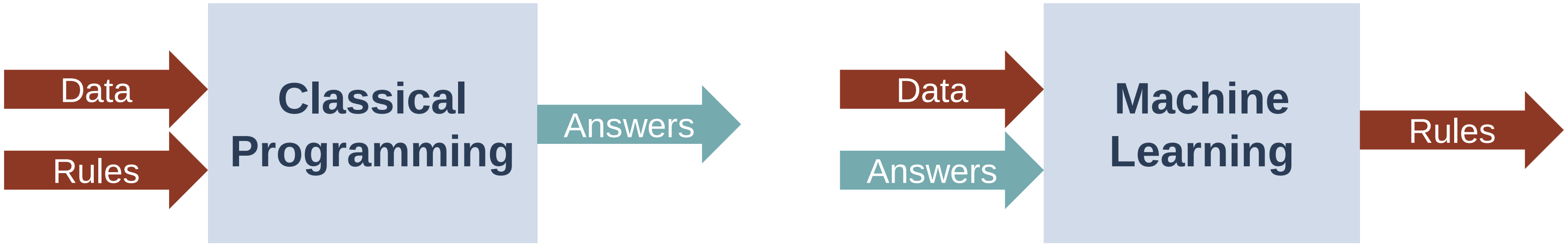
- Prior to this project, staff used pump station data to estimate peak flow and had 30-60 minutes of advance warning
- Flow monitors in collection system aren't predictive
  - **Doesn't tell you if flows will increase or decrease**
- City has a calibrated collection systems model but no way to currently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



Collection system model output, manually generated.

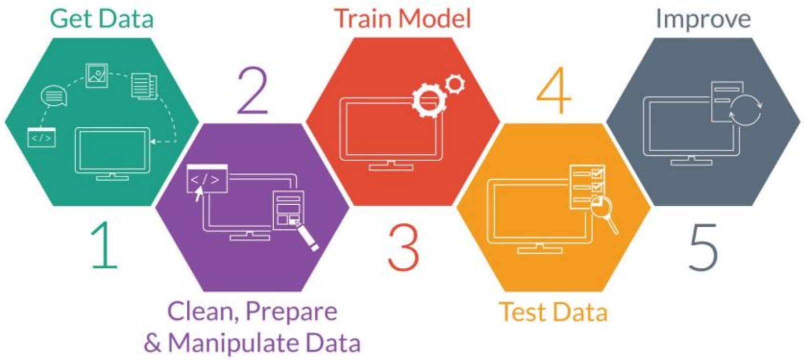
Machine Learning Overview

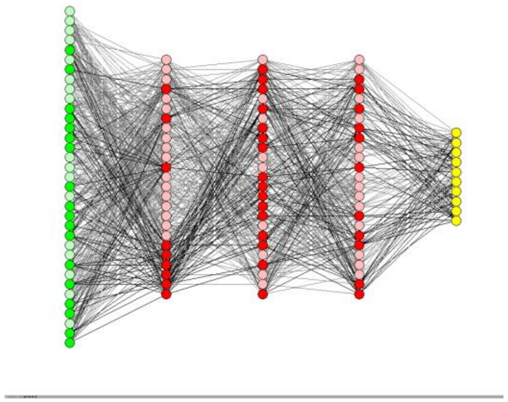
Machine Learning is Well Suited for Creating Predictive Tools because it can make Accurate Predictions without Explicitly Being Programmed to Do So



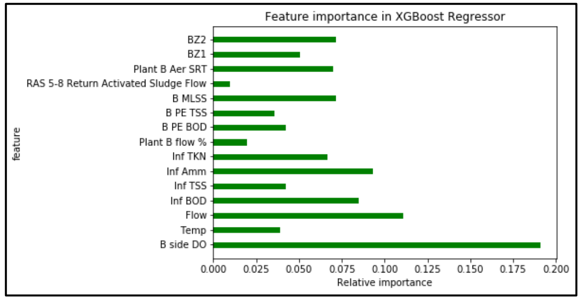
Machine Learning is an Alternative to Traditional Mechanistic Models

- 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.

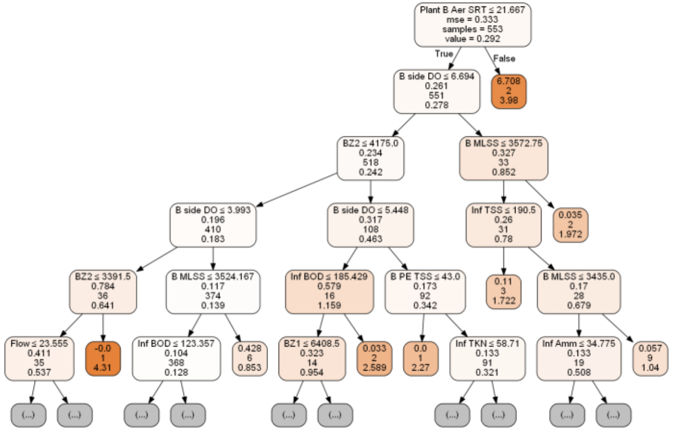




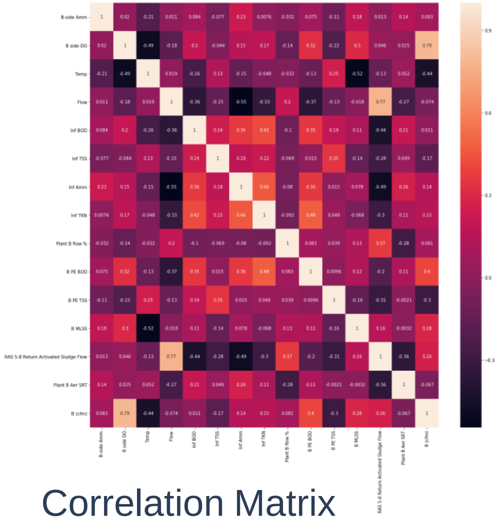
Examples of Machine Learning Tools



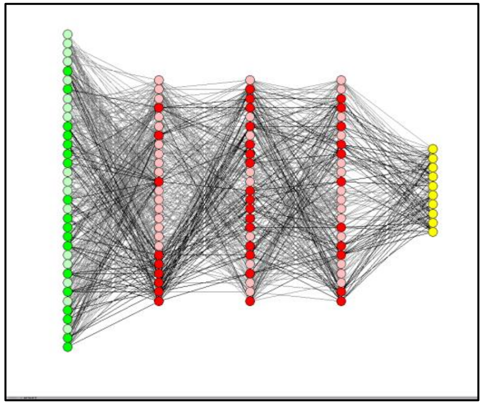
Feature Importance



Decision Trees

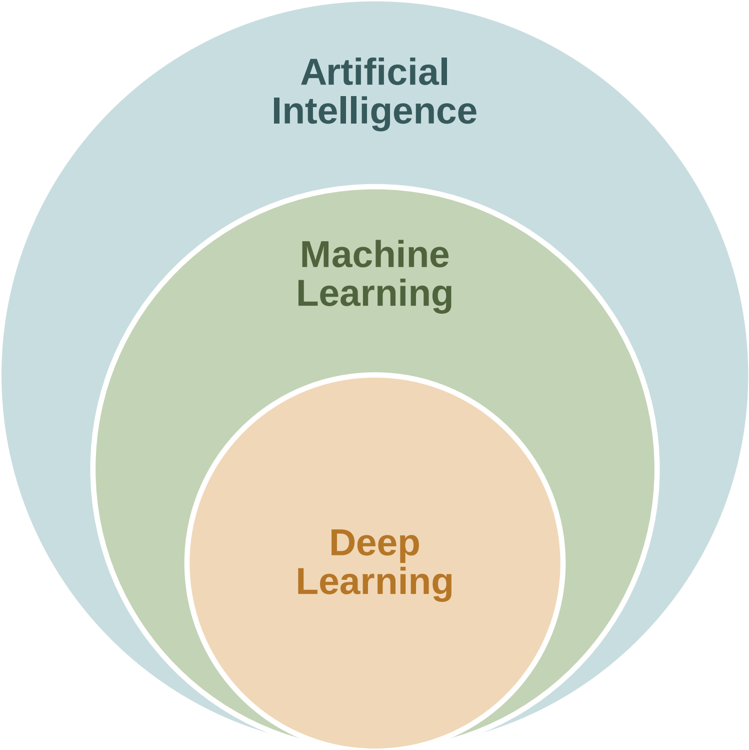


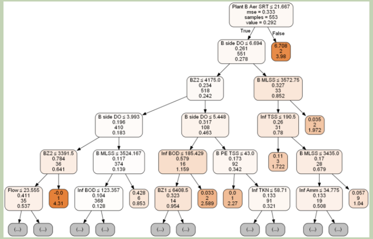
Correlation Matrix



Neural Networks
(deep learning)

Common Types of Machine Learning Algorithms

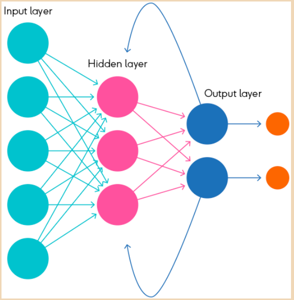




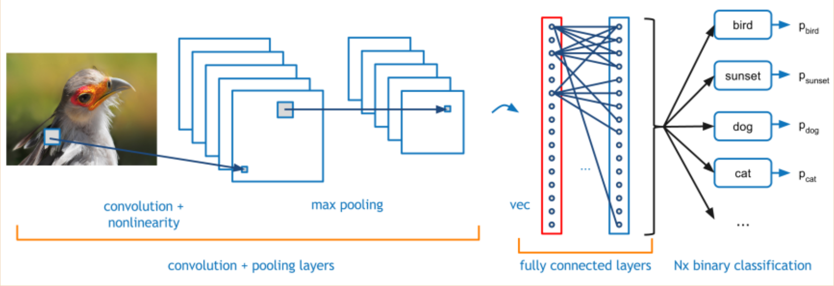
Decision Trees

Random Forest Models

XGBoost

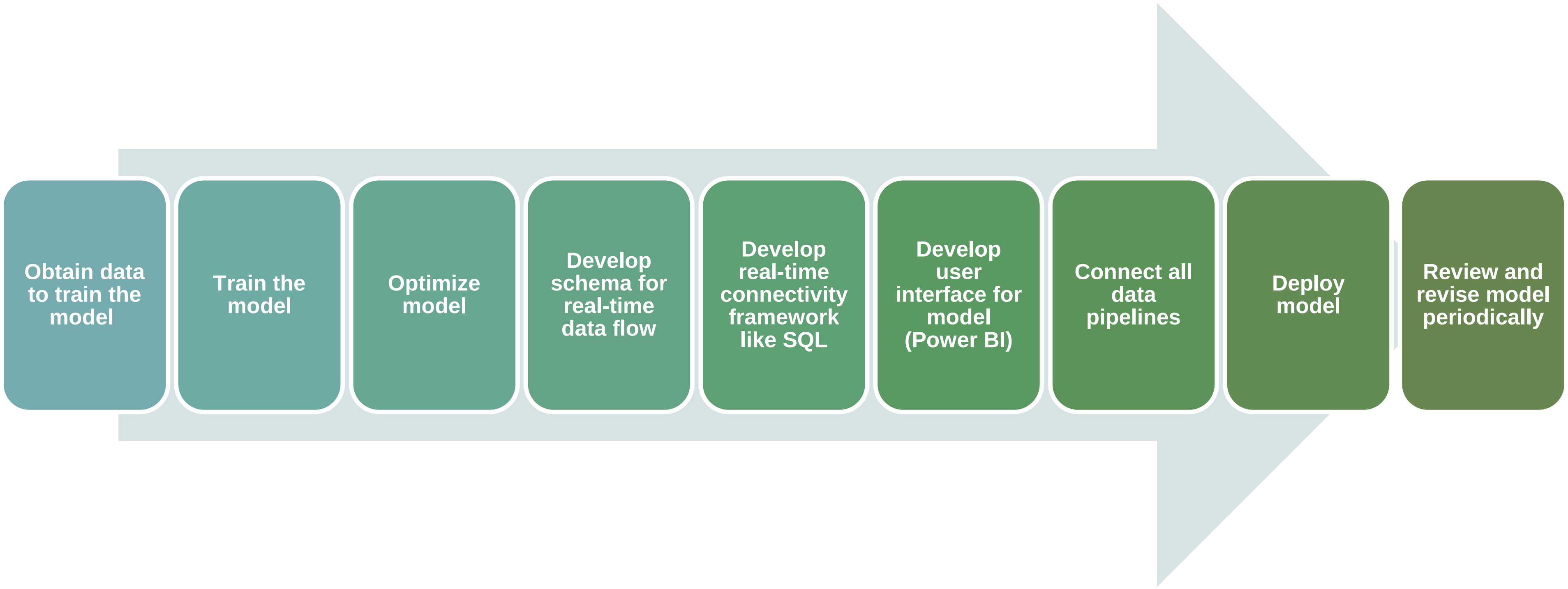


Recurrent Neural Network



Convolutional Neural Network

Steps to Deploying a Machine Learning Model



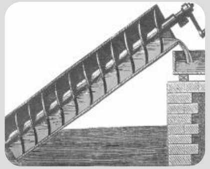
Early Explorations with Machine Learning Led to The Raleigh Water Project

Raleigh Water Model Development

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

Rainfall

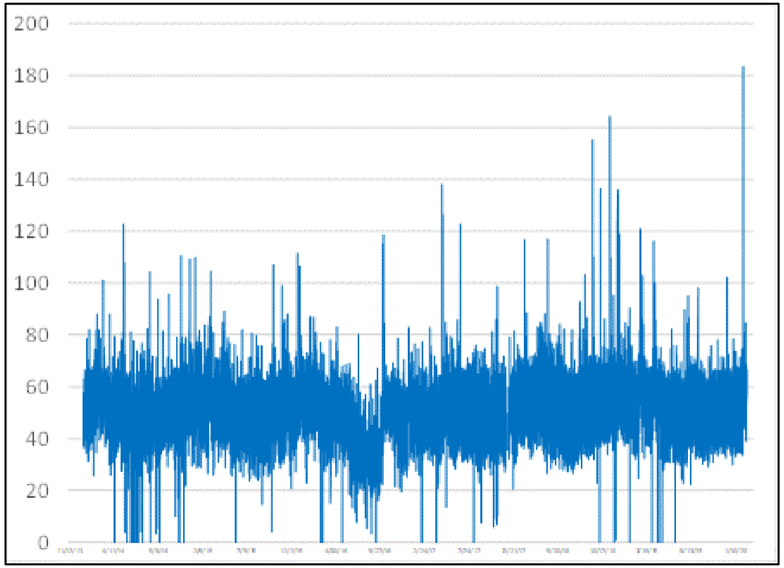
Streamflow



Past Influent Flow to NRRRF

Hour of Day

Collection System Improvements



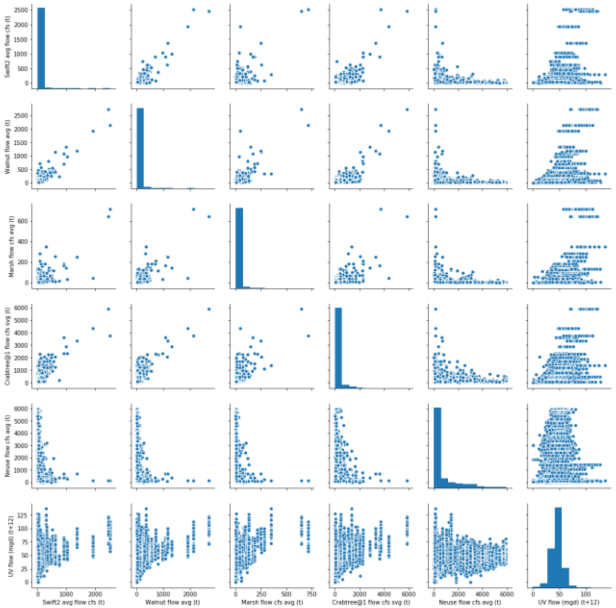
Used python machine learning algorithms to train a model to 6+ years of influent flow data as a function of explanatory variables.

Sustained flows of 184 mgd experienced

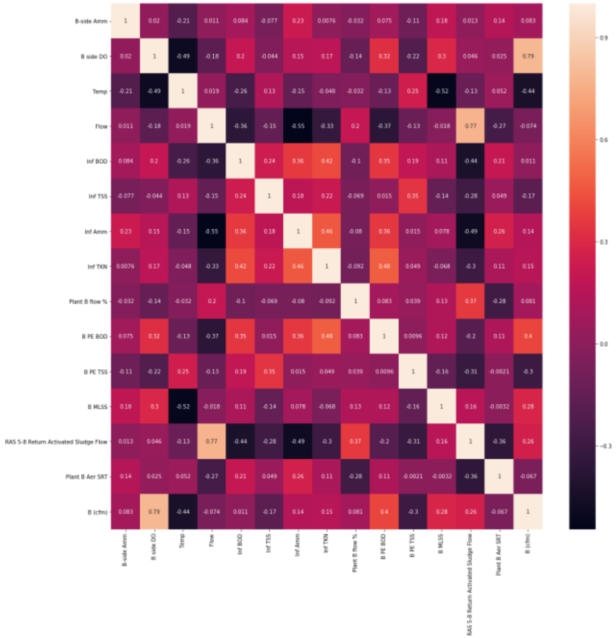
Challenge meeting effluent TN and TP during wet weather events

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

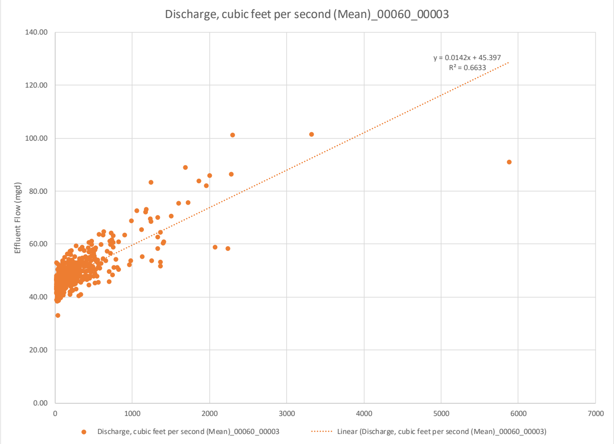
Used Exploratory Data Analysis Tools In Python To Select The Right Variables For Use In This Model



Pairwise plots



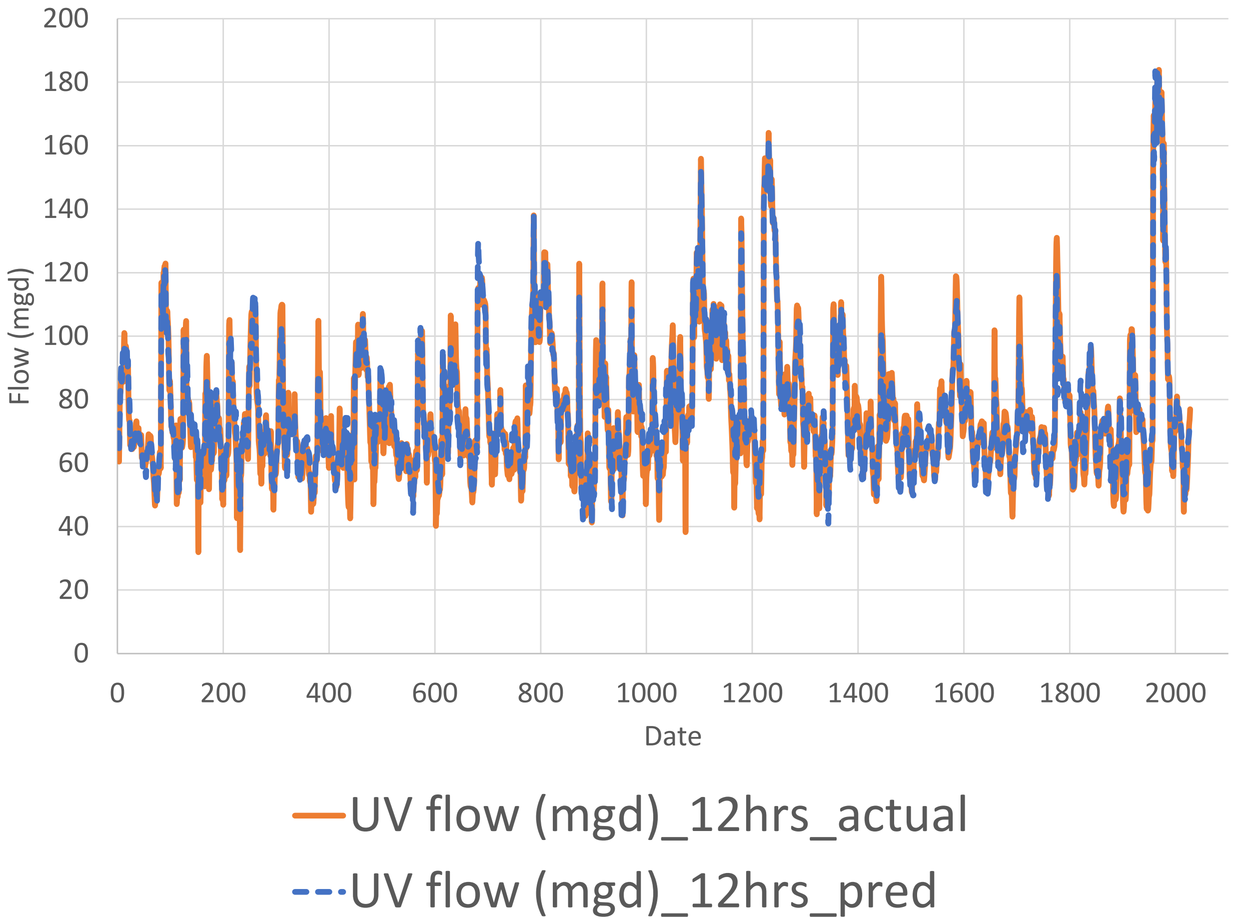
Correlation tables



Linear regression

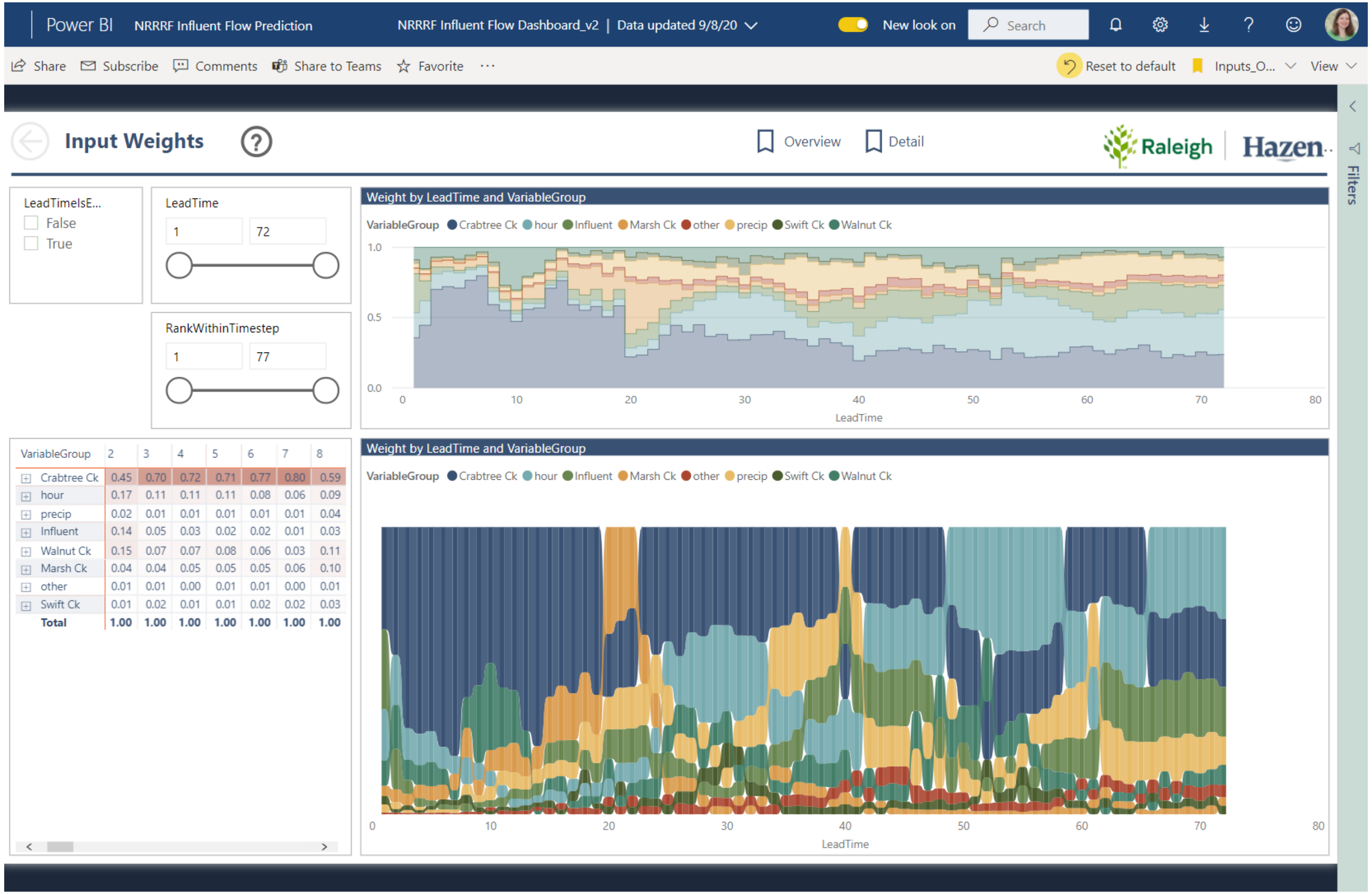
All Storms Predicted with Good Precision by the Model During Training

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

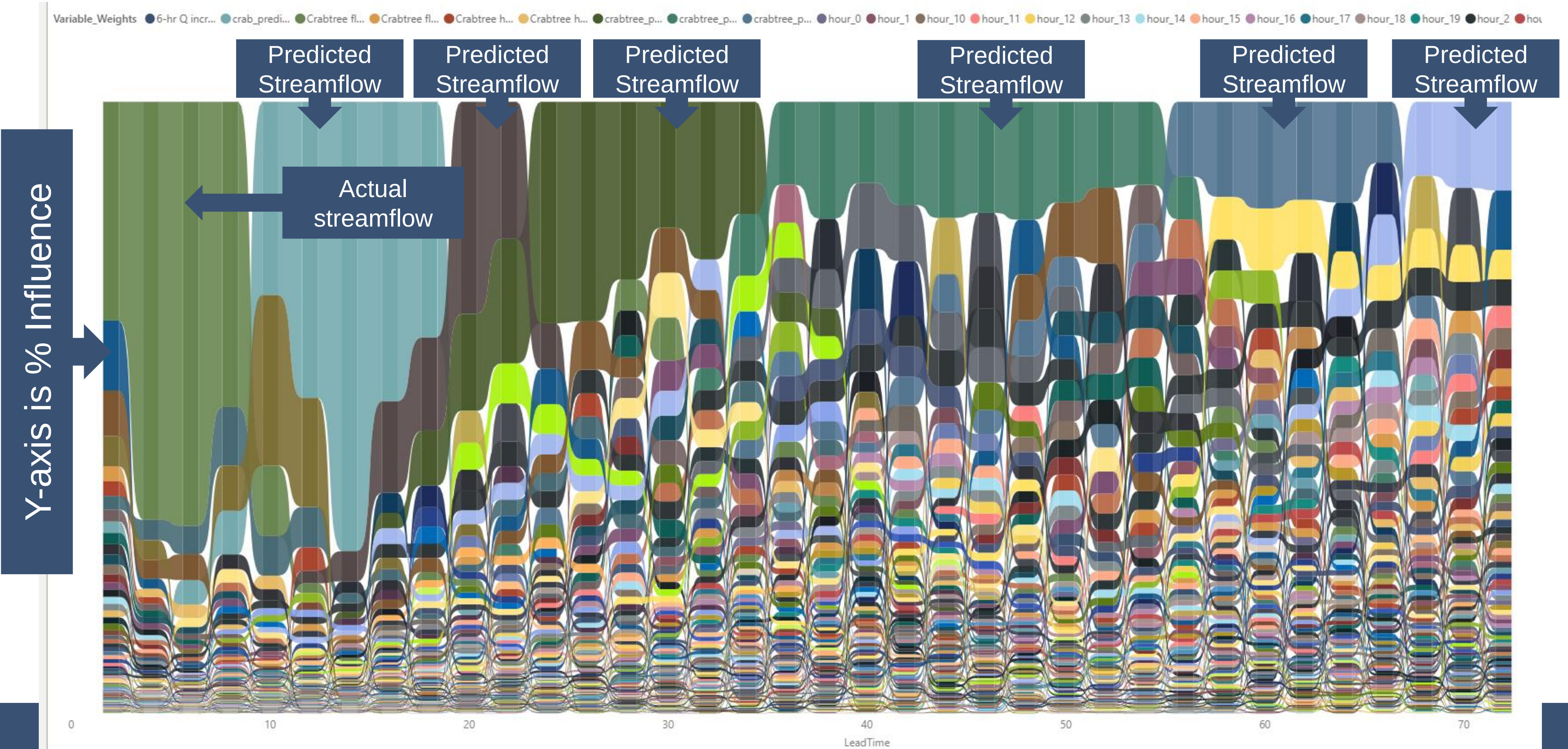


The Selected Model Is Fully Transparent

- 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

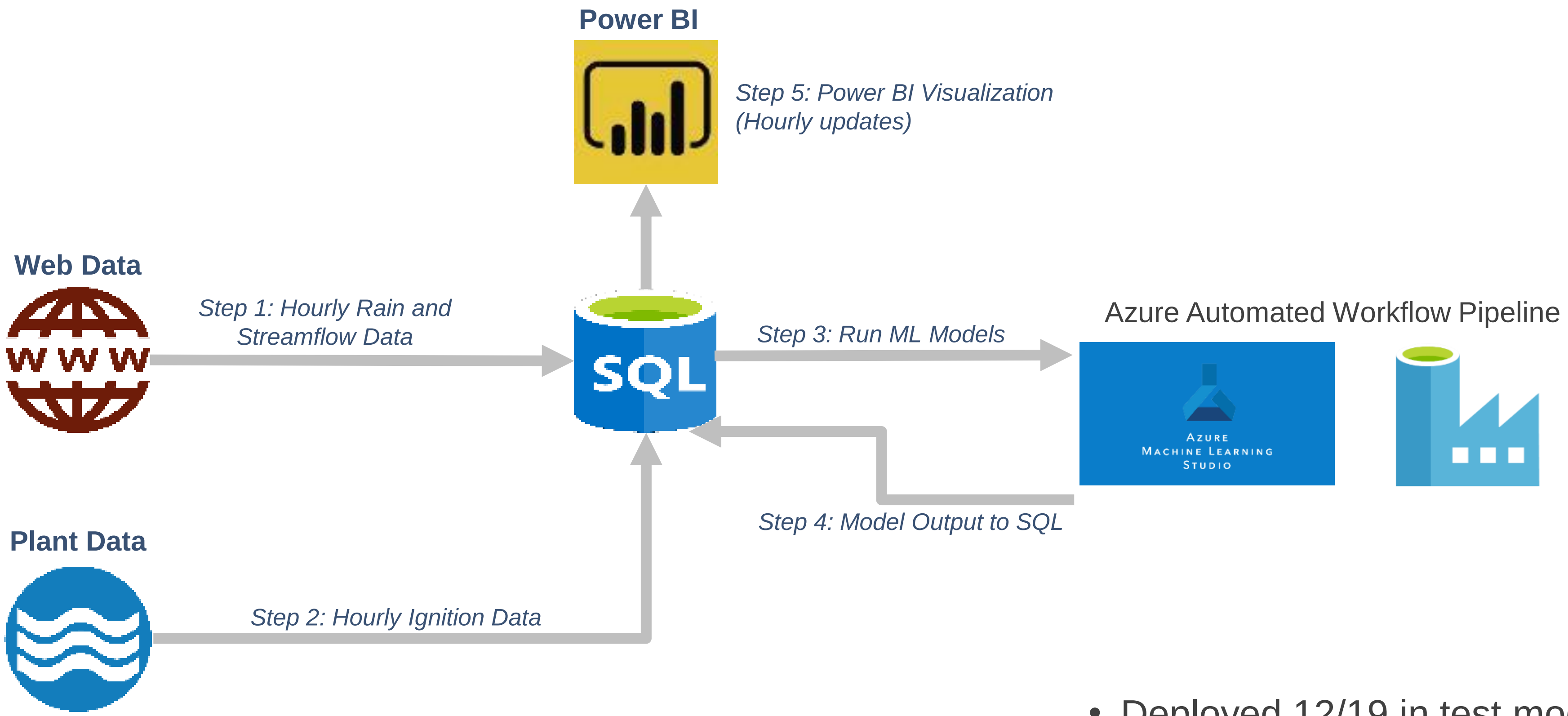


Closest Predictions Rely More on Actual Streamflow Data and Farther Away Predictions Rely on Predicted Streamflow and Rainfall Totals



Model Deployment

Data Architecture



- Deployed 12/19 in test mode
- Finalized 7/20

Model Reliability and Maintenance of Automated Pipeline

Component	Weather data	USGS data	Ignition Data	SQL	Power BI	Automated Machine Learning Pipeline
Source	Dark Sky and NWS	USGS	Raleigh Water	Hazen	Hazen	Hazen
Reliability	98.2%	98.7%	96.7%	Same as Azure	Same as Azure	99.95%

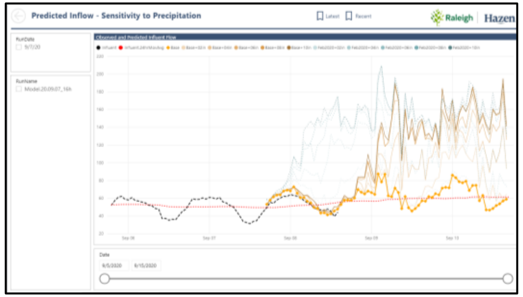
Cloud/Azure

Interactive Tool Raleigh Water Staff Use To View Model Results

Final Deliverable Has 16 Screens

1. Cover
2. Inventory
3. Model prediction
4. Model sensitivity
5. EqOps
6. Secondary clarifier guidance
7. Model performance
8. Model QC
9. Plant Ops
10. USGS
11. Precip
12. XY
13. Timeseries
14. Map
15. InputWeights
16. Inputs

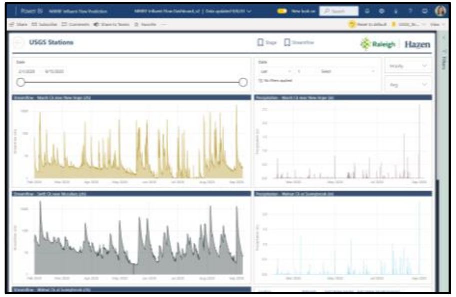
Model prediction



Sensitivity to Rainfall Amount

Select flow above which to utilize EQ

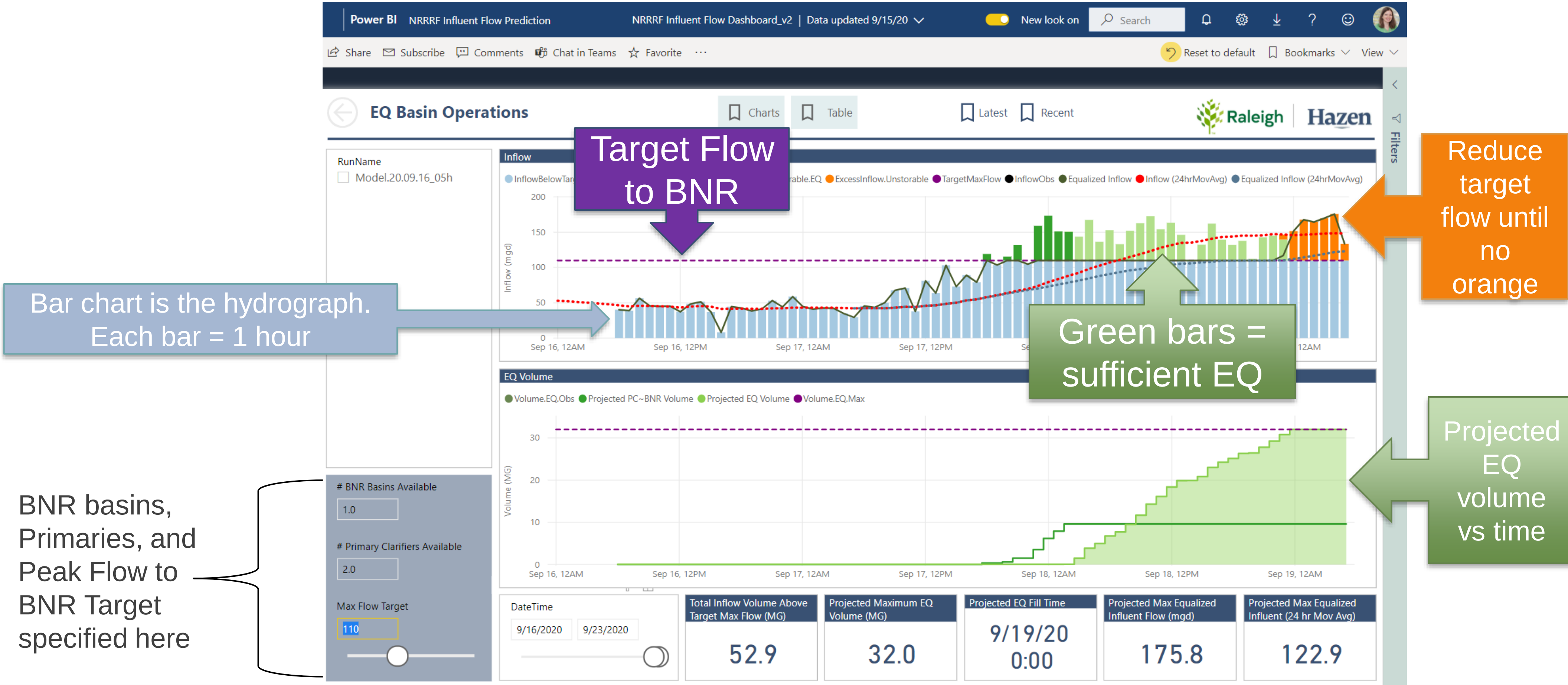
Secondary Clarifier Guidance Program to estimate # SCs needed



Monitor USGS Streamflow

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

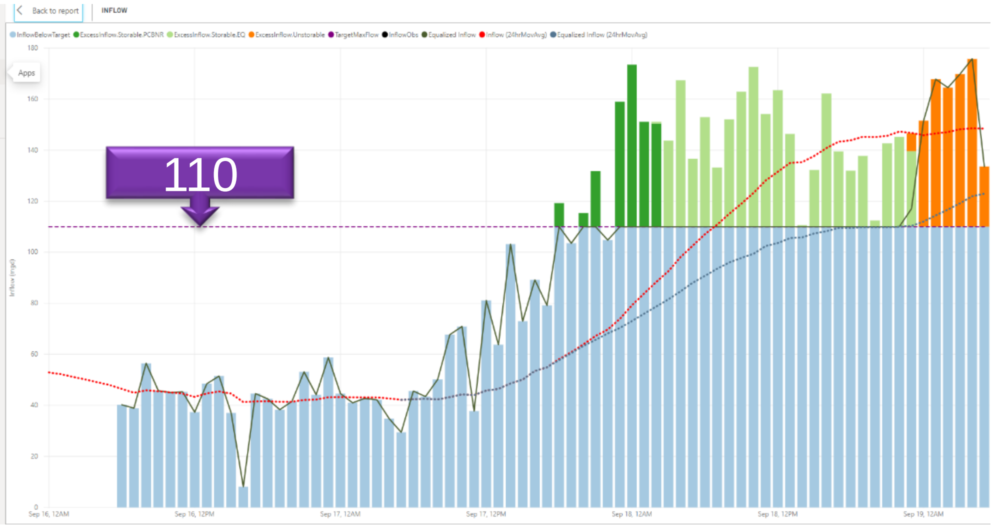
Example of How the EQ Management Tool Works

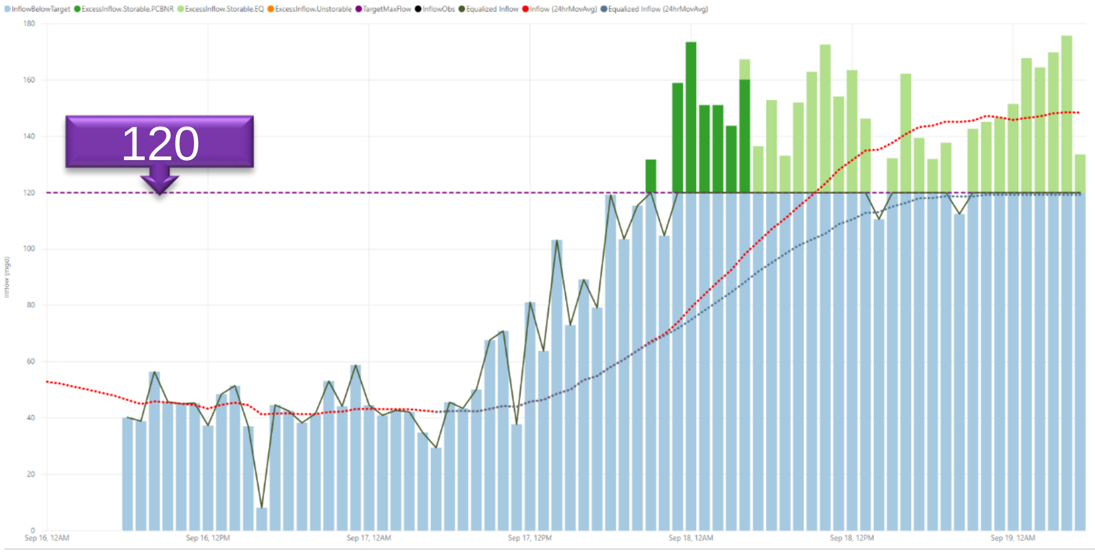
Flow threshold set to 110 mgd.

There is insufficient EQ capacity.

Flow threshold set to 120 mgd.

There is adequate capacity. Strategy is to divert flows when Q > 120 mgd.





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

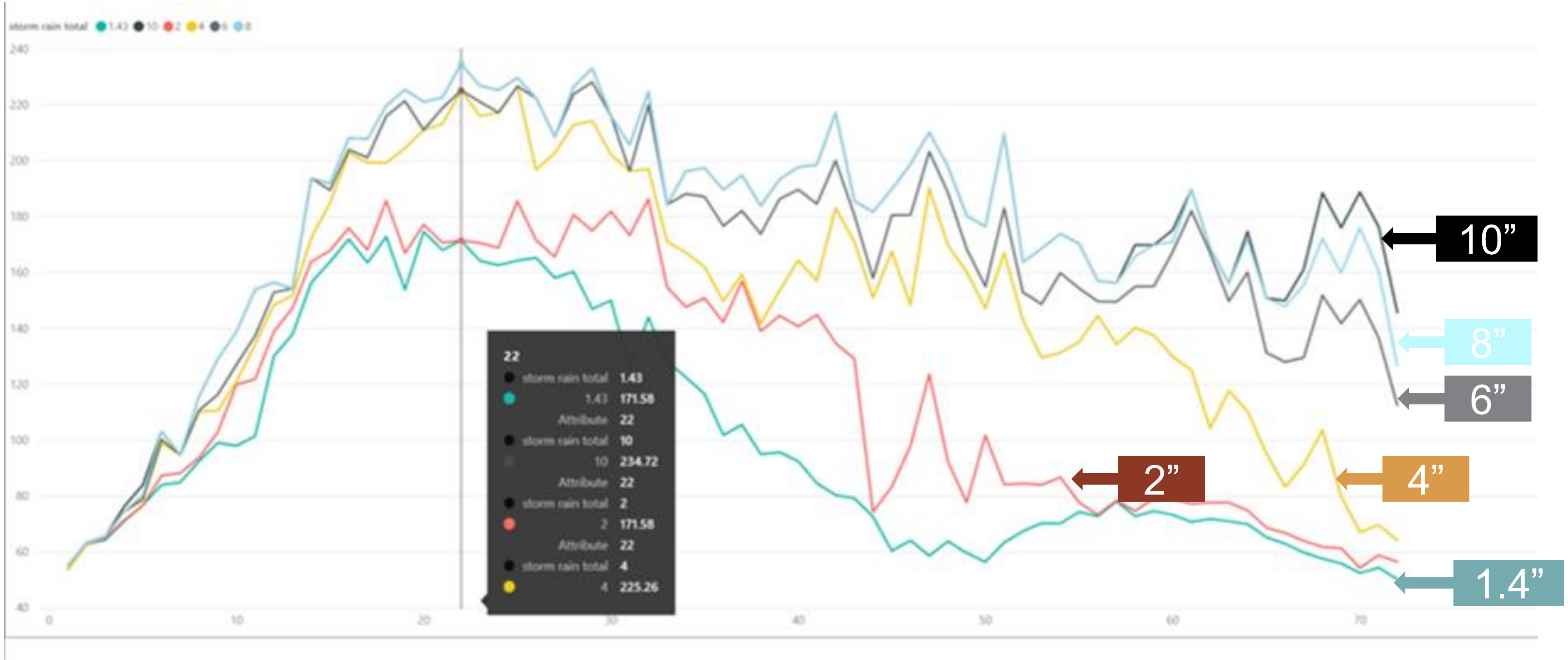
Left – displays key performance indicators for past 72 hours.

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

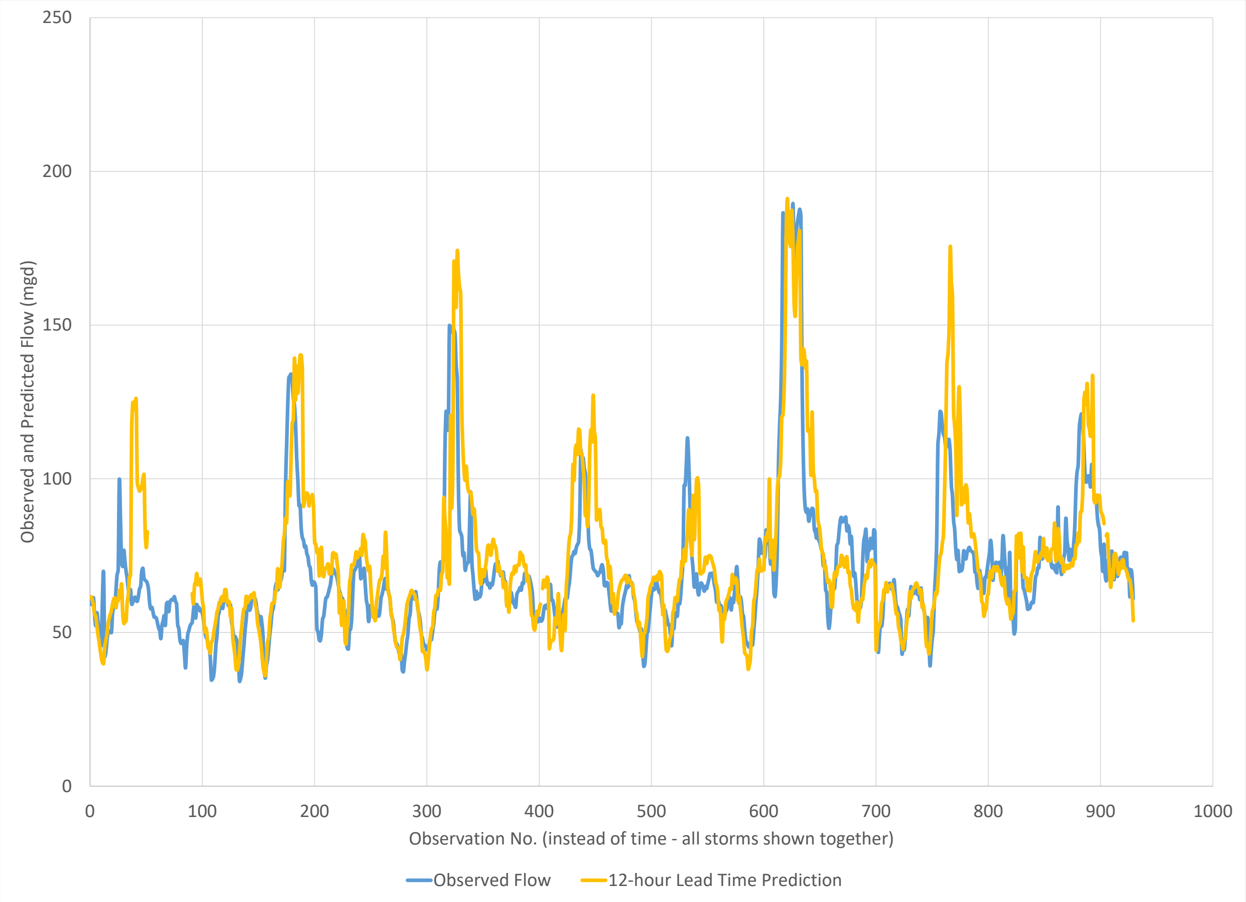
Bottom center –KPIs and combinations of small and large clarifiers that meet the criteria in the calculator tool.

Model Includes Sensitivity Analysis to Account for Uncertainty in Rainfall Quantity and Timing



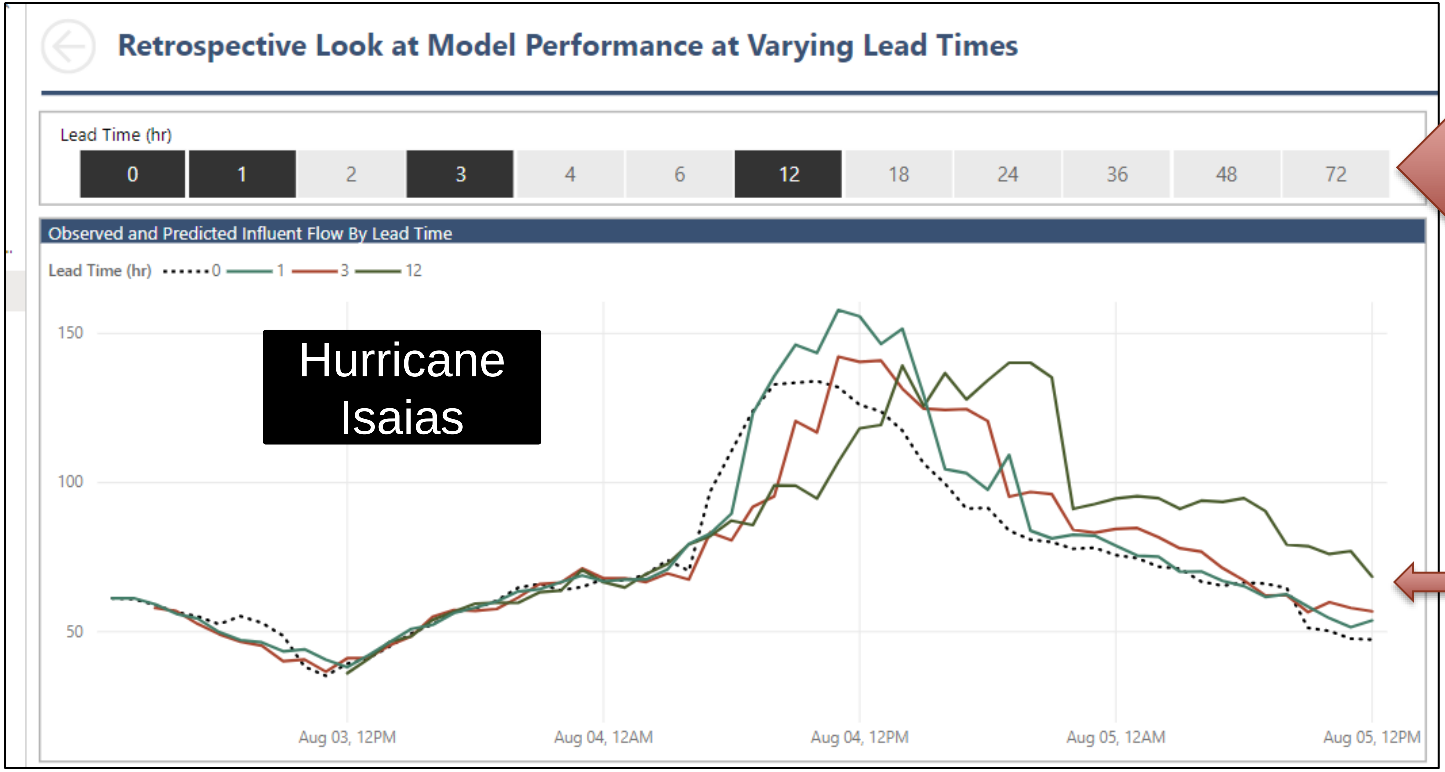
Model Performance

There Have Been 8 Major Storm Events Since the Model Was Deployed in July 2020



- All well predicted
  - **Blue – observed**
  - **Yellow – predicted 12-hours in advance**
- Wet weather EQ used 5 times
  - **Volume ranged 12.6 – 26.8 MG**
  - **Never exceeded 32 MG**
- Models errs on the side of being conservative
  - **This is because 10+ hours away model depends more on predicted streamflow and rainfall**
  - **Model w/i 10 hours of event depends more on actual streamflow and rainfall**

Model Accuracy Post-Deployment Has Been Very Good

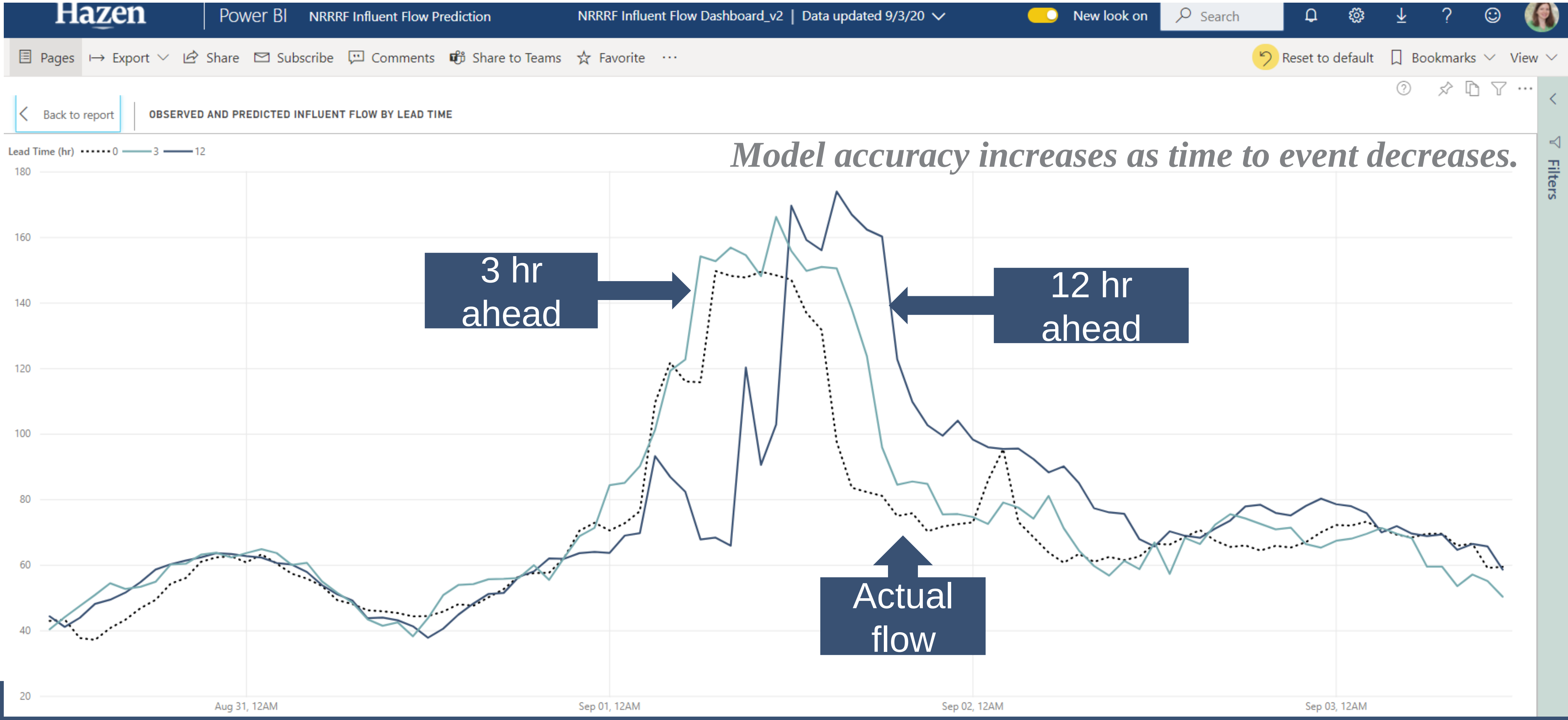


Operator selects forecast lead time to compare to actual flow

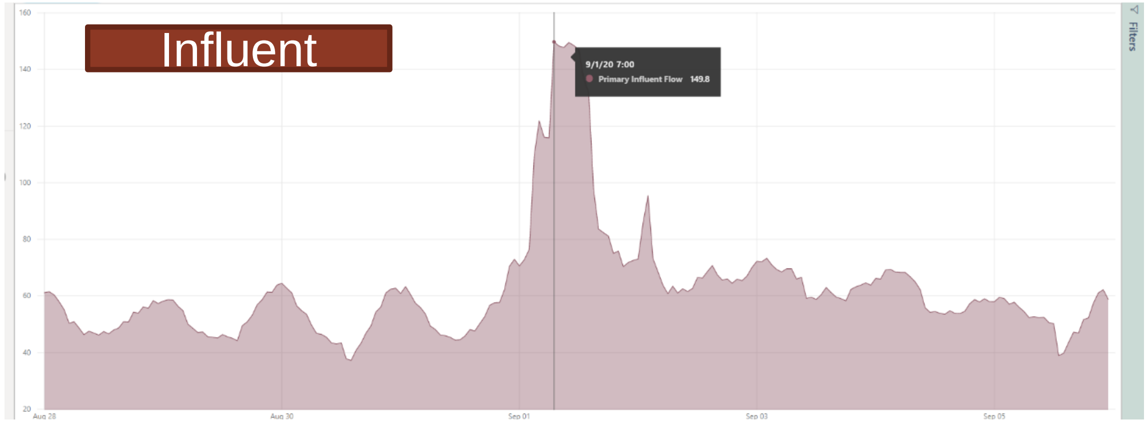
Results plotted here

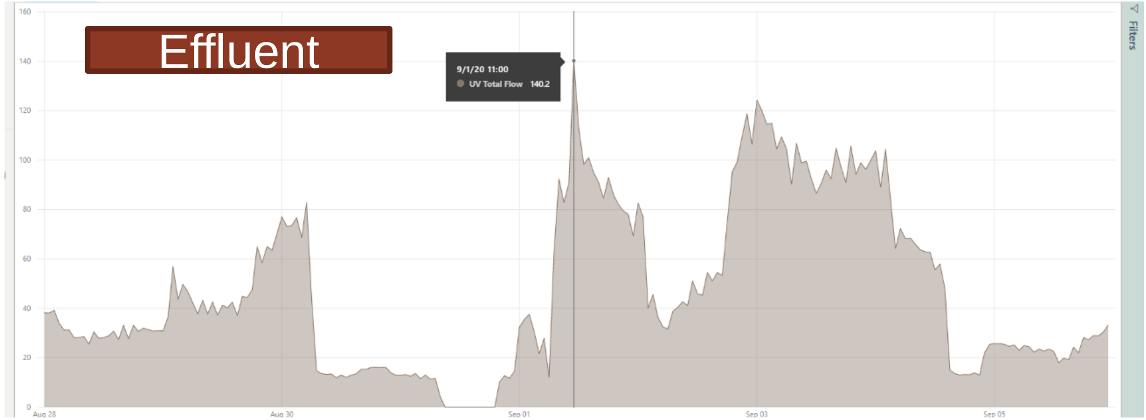
Model accuracy increases as time to event decreases.

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



Operators Used the Prediction to Implement their Wet Weather SOP Beginning with PCs, BNRs, then EQ

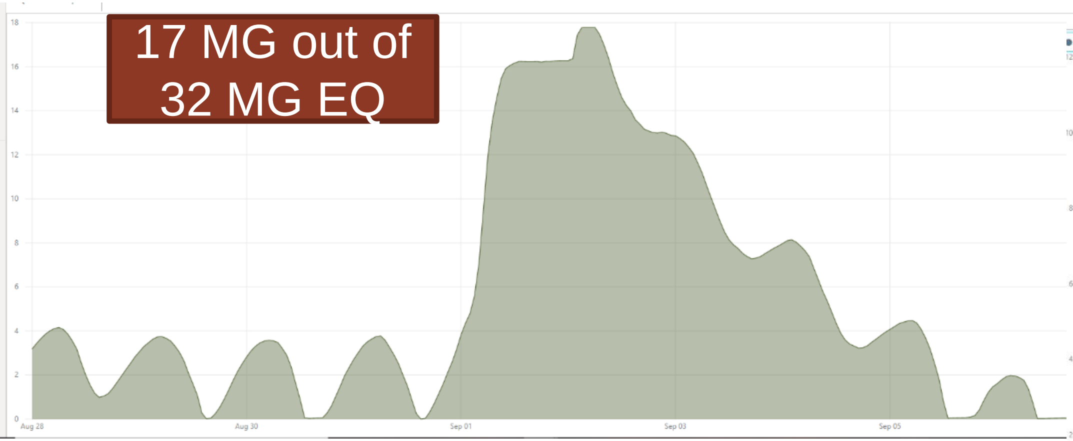


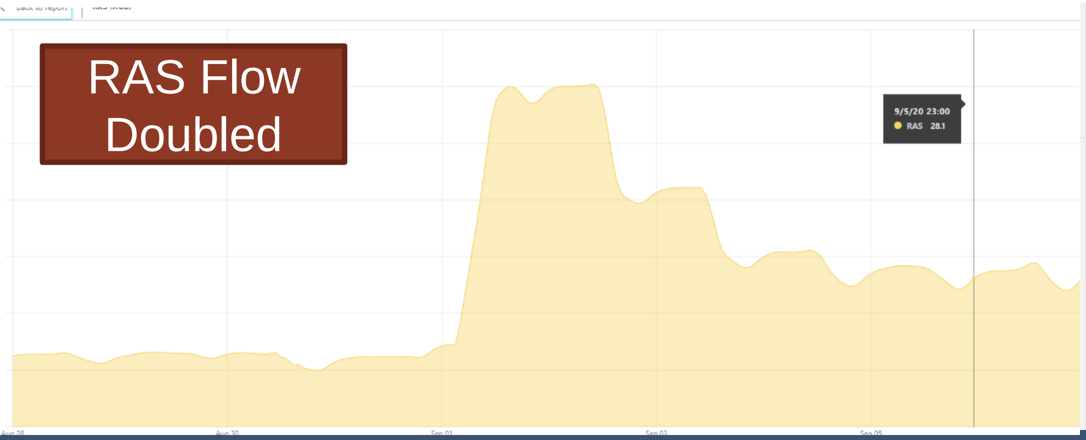


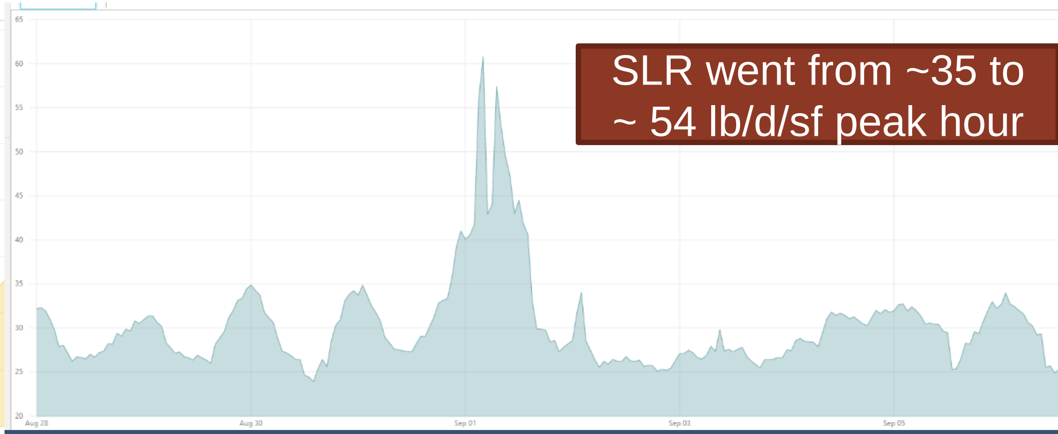
Date	Influent Flow (mgd)	Effluent Flow (mgd)
9/1	106	69
9/2	68	61
9/3	65	101
9/4	53	51

- Put 2 primaries online
- Put 1 BNR basin online
- Utilized 17 MG of EQ
- Peak hour flow 148 mgd
- Peak hour SLR 54 lb/d/sf
- Peak hour SOR 1500 gpd/sf

They Also Doubled RAS Flows, Put 6 SCs into Service, and Their Strategy Left Some EQ Volume Still Available



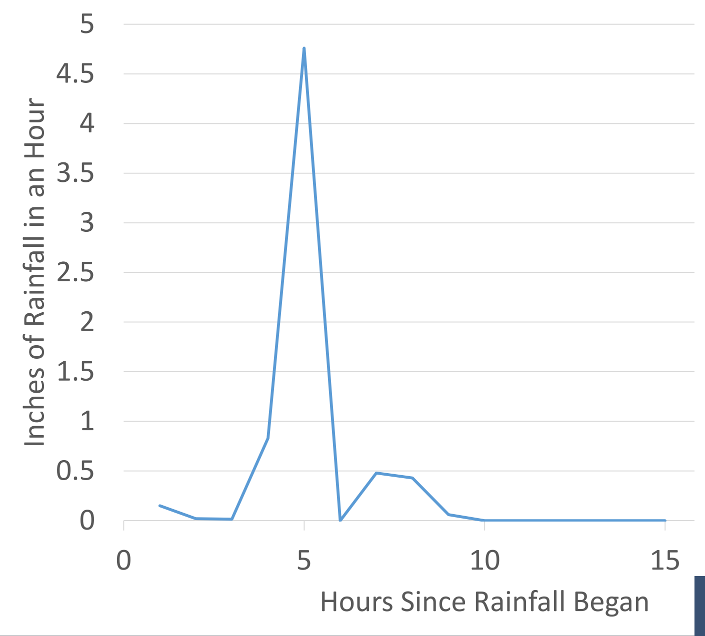




NRRRF Maintained Good Effluent Quality During This 6.7” Rainfall Event

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	48	BDL	-	0.14	-
9/1	106	69	-	-	-	-
9/2	68	61	BDL	0.54	BDL	1.9

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

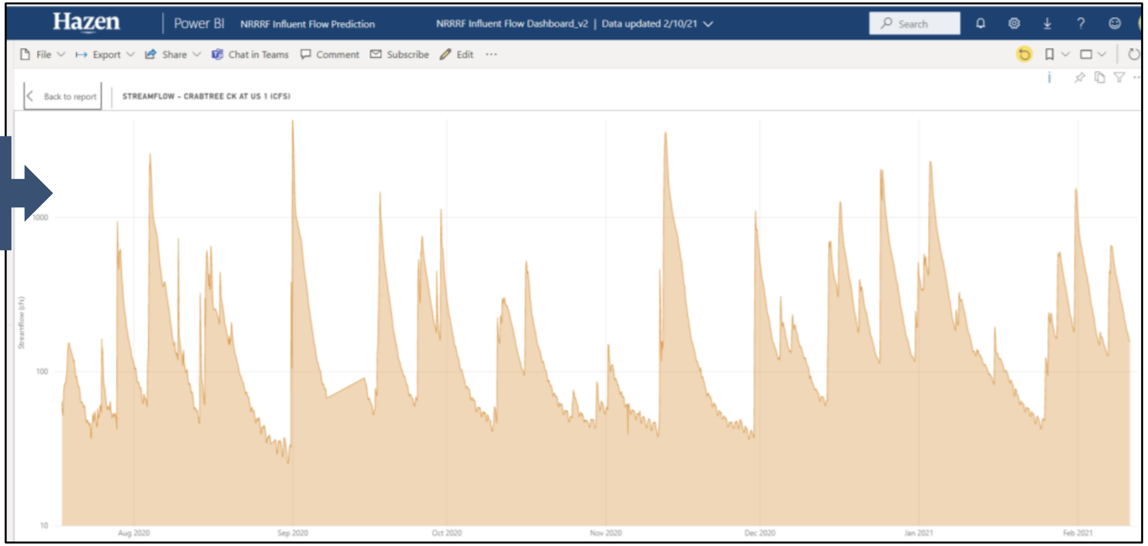


A Few More Thoughts On The Visualization

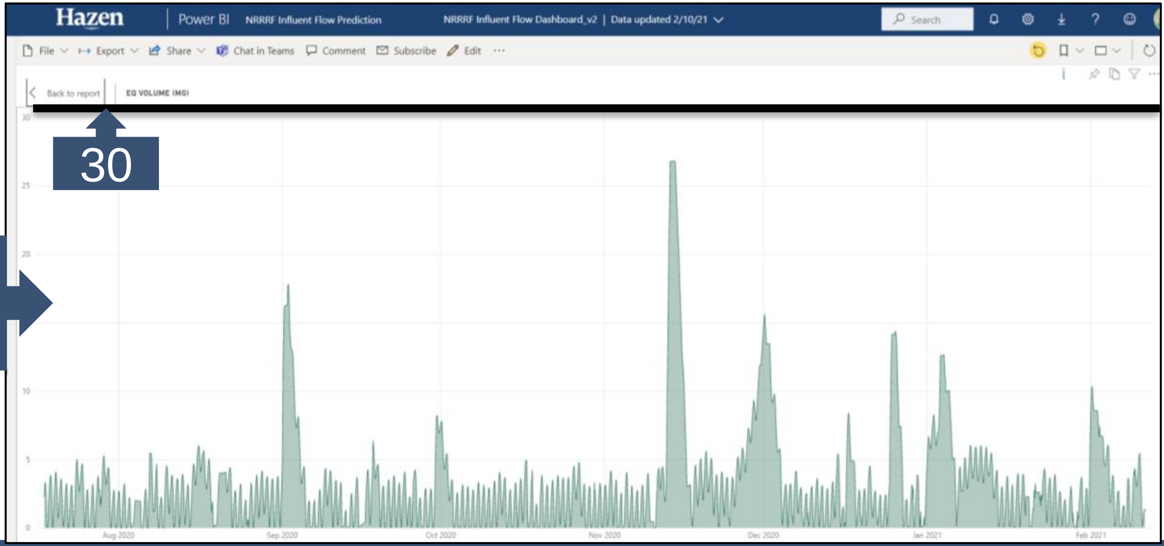
"The resultant model and Power BI dashboard are an extremely valuable tool that provides utility staff near-real-time visualizations of key data, such as current operating parameters and stream flood stages as well as future flow predictions.

The tool provides an interactive interface for quickly assessing current conditions and planning ahead for projected future conditions, which assists with making informed decisions, resulting in in greater efficiency and reliability in utilizing existing infrastructure, to effectively manage wet weather flows and continue to meet stringent effluent limits."   – Raleigh Water

Streamflow, log scale



EQ basin volume utilized



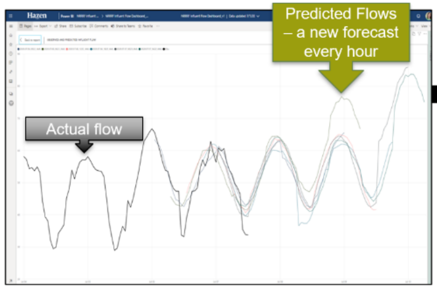
30

Summary

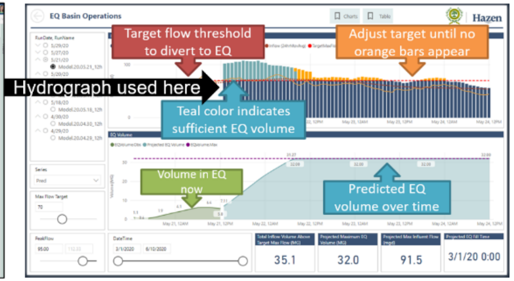
Lessons Learned from Deployment of Machine Learning at NRRRF

- NRRRF has realized significant operating cost savings with real-time process controls
- Machine learning has been utilized successfully to create a wet weather management tool

TN Reduction	Electrical savings	Methanol savings	Total savings
2.2 mg/L → 1.8 mg/L	$300,000/yr	$200,000/yr	$500,000/yr

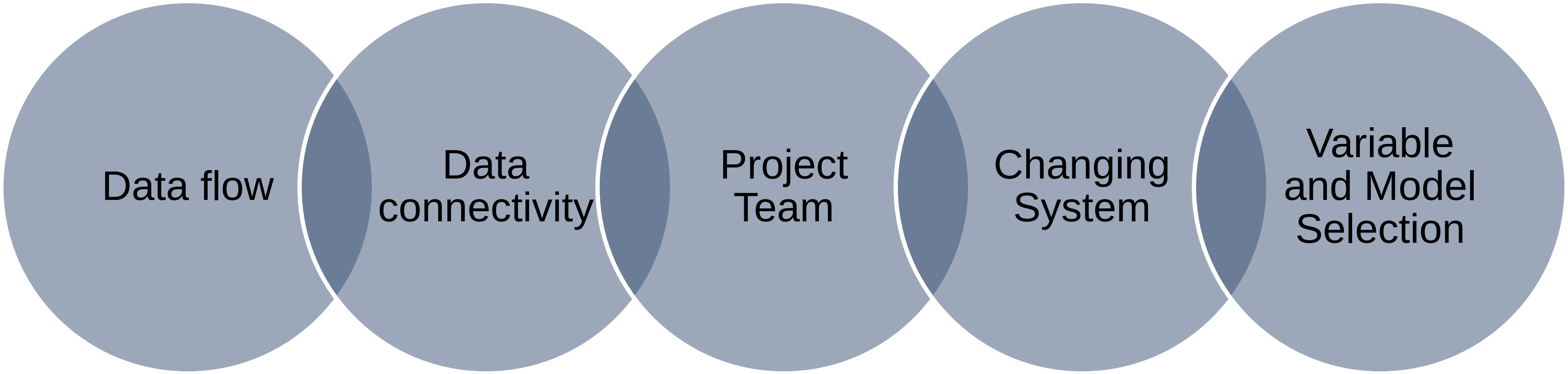


Used machine learning to train a model to 6+ years of influent flow data as a function of explanatory variables.



During wet weather, plant staff use to tool to determine the flow threshold above which to use equalization to minimize flow to the BNR process.

Important Considerations for Predictive Analytics Project

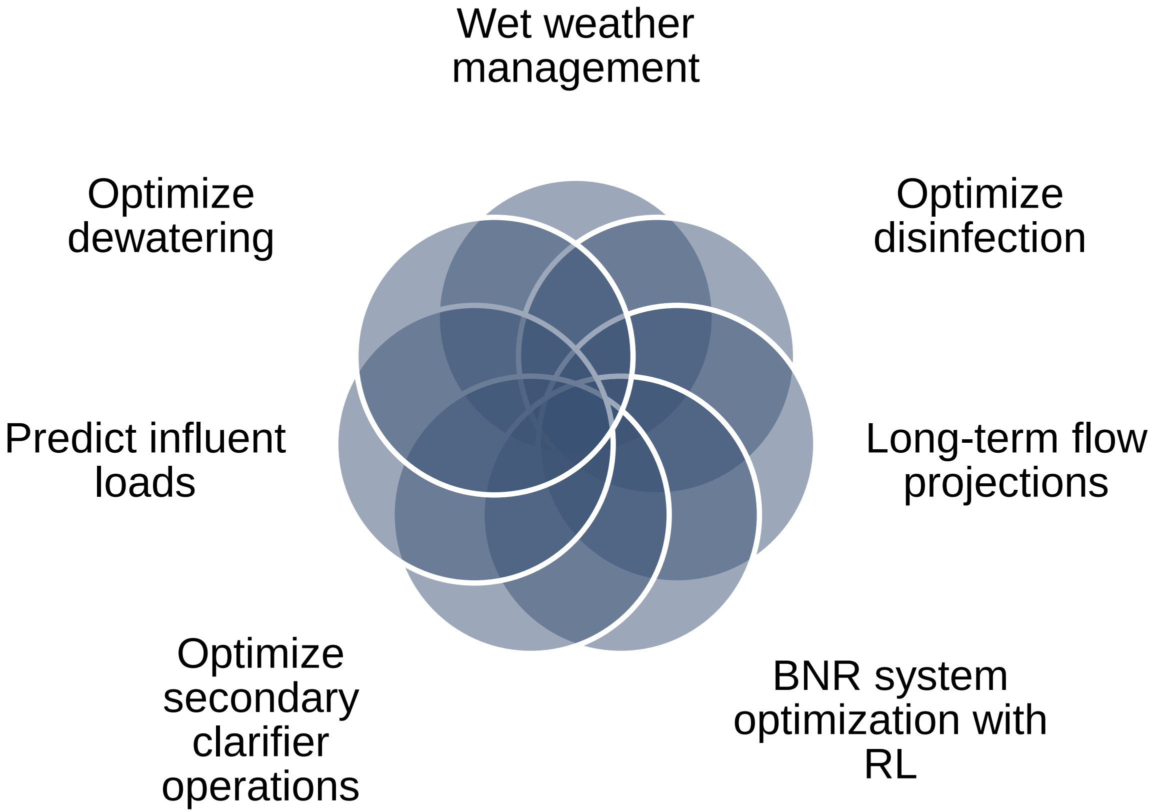


Acknowledgments

Many people from Raleigh Water and the Hazen team contributed to the success of this project including but not limited to: John Kiviniemi Nathan Howell, Mark Wessel, Steve Worley, Sharon Burch-Martin, Aaron Brower, Lisa Joseph, Michael Stevens, Adam Haggerty, Howard Hoyle, TJ Lynch, and Robert Massengill with Raleigh Water; and Grantley Pyke, Rod Moeller, Jamie MacDonald, Justin Irving, Luke Wang, and Alan Stone with Hazen and Sawyer. The project required coordination with the City of Raleigh's IT Department both initially and throughout the project for data flow management and plans for future deployment of the model. Sharon Burch-Martin led the coordination with the IT Department. Raleigh Water's systems integrator, CITI (Leo Jurado), also assisted with extracting hourly plant operating data from Ignition software to the Hazen Cloud. The authors would also like to thank Steve Cook with Black and Veatch for conducting those before and after simulations with the calibrated collection systems model and his insights into the performance of that model.

Other Opportunities

Future Intelligent Water System Goals for WRFs



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

Identify Potential Parameters

Evaluate/ Analyze Parameters

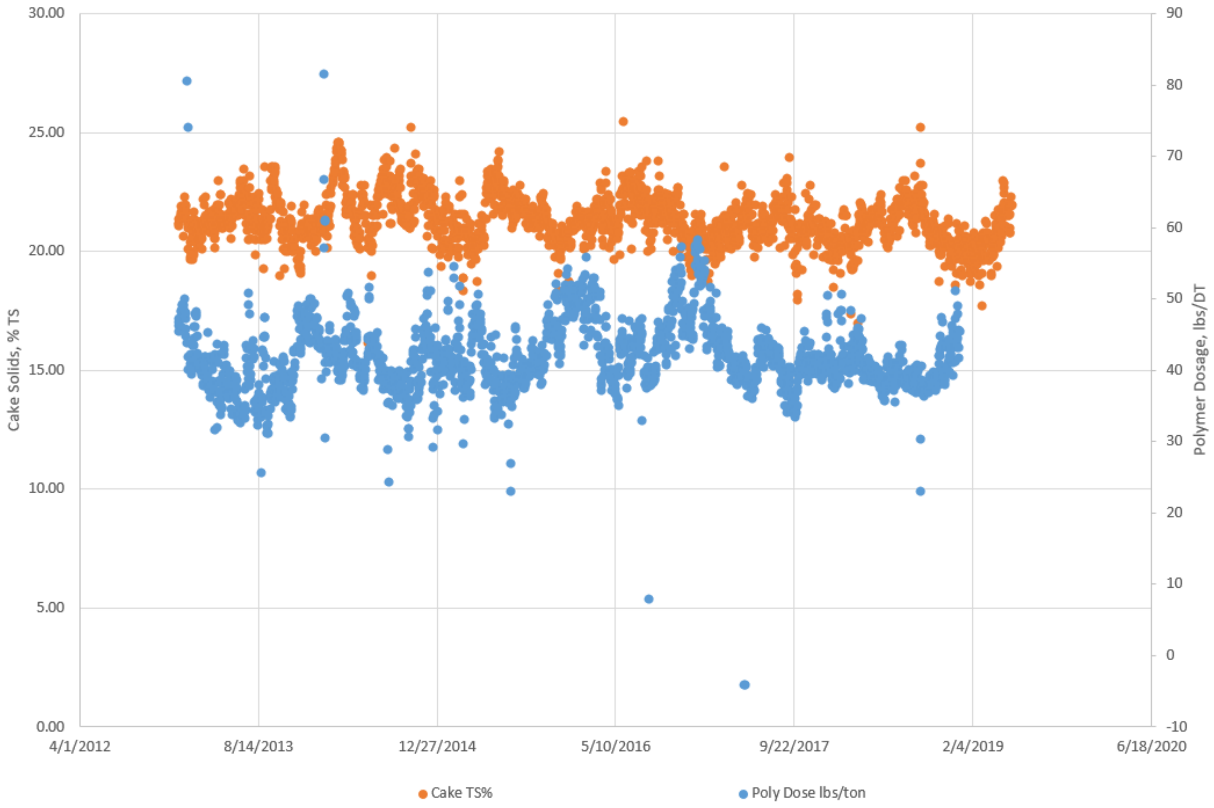
Develop Predictive Tools

Iterate and Refine Tools

How Would This Tool be Used In Real Life?

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

Random Forest Prediction was Most Accurate

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

Key Variables Predicting Dewaterability and Their Relative Importance

Sensitivity Analysis

Comparing Two %TS Prediction Models and Their Conclusions

Additional Machine Learning Applications Outside of WRFs

Predicting sewer pipe deterioration

Water treatment optimization

Water supply predictions

Predicting flood potential

Questions

Katya Bilyk

4011 Westchase Boulevard, Suite 300

Raleigh, NC 27607

kbilyk@hazenandsawyer.com

919-538-1276