Oregon State University

Machine Learning for Advanced Control and Prediction in Drinking Water Treatment: Coagulation and Flocculation

Perry Wilson

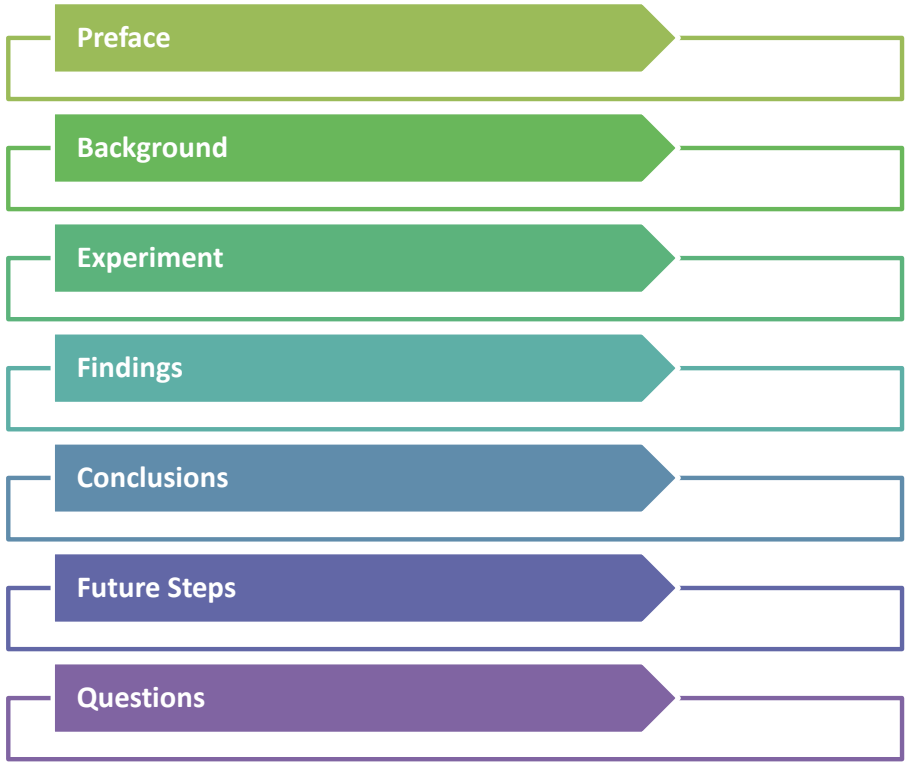
May 7, 2026

COLLEGE OF ENGINEERING

School of Chemical, Biological, and Environmental Engineering

Overview

Oregon State University
College of Engineering



Intended Audience

Oregon State University
College of Engineering

This presentation is meant for:

- Starting a ML project
- Beginners in ML
- Curious about the abilities of ML



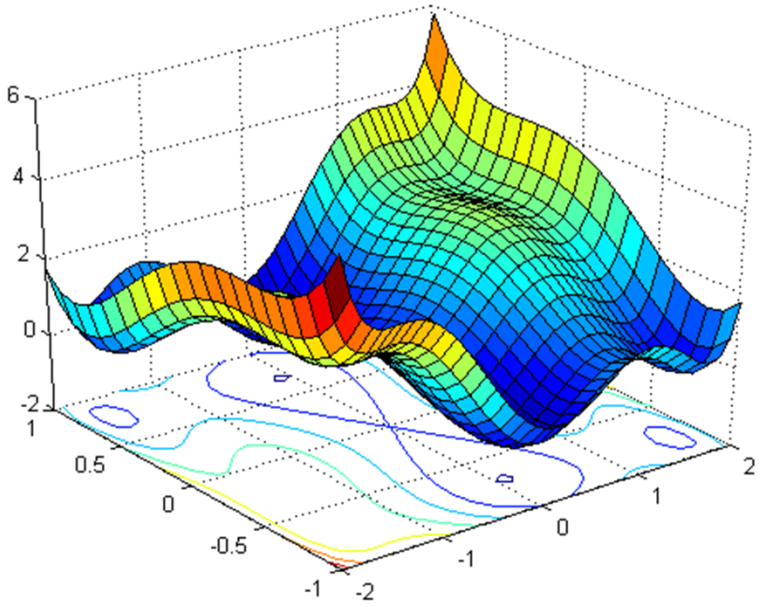
- Key Considerations
- General Workflow
- Helpful tools
- Common Challenges

Systems that Benefit from ML/AI

Oregon State University
College of Engineering

Difficult to Determine Optimization:

1. Highly Non-Linear Behavior
2. Hard to model mechanistically
3. Multi-Dimensional
4. Temporal Dependence (daily, seasonally, yearly)

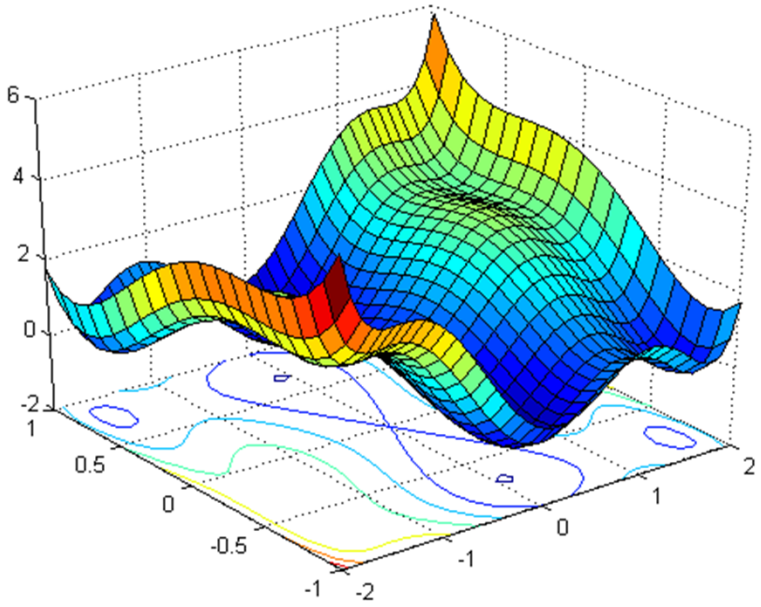


Systems that Benefit from ML/AI

Oregon State University
College of Engineering

Difficult to Determine Optimization:

1. Highly Non-Linear Behavior
2. Hard to model mechanistically
3. Multi-Dimensional
4. Temporal Dependence (daily, seasonally, yearly)



Trends captured with sensing

Example: Coagulation / Flocculation

Oregon State University
College of Engineering

Doses & settling determined with:

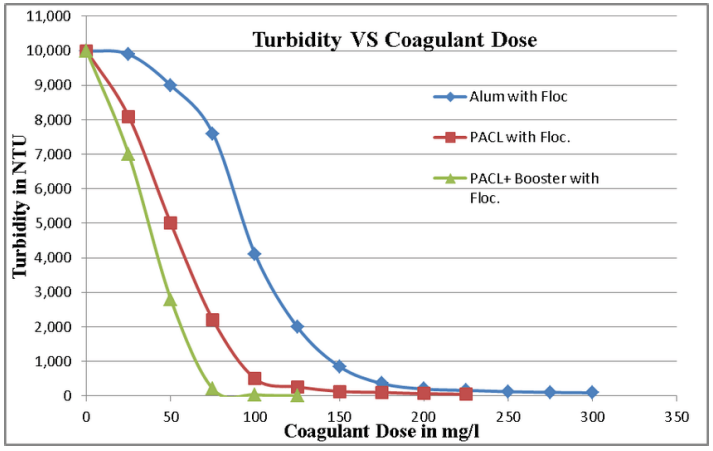
- Jar Tests
- Column Tests
- Doses Based on Operator Experience





Sensing in typical basin:

- Turbidity
- pH
- Temperature

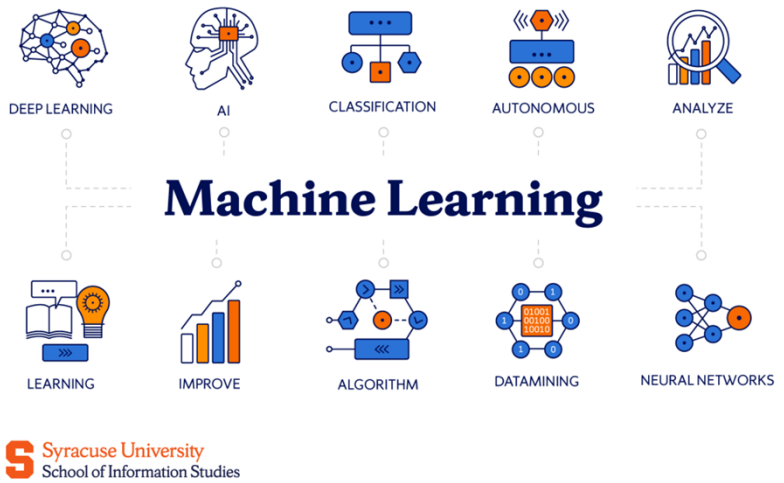


Project Discussed

Oregon State University
College of Engineering

Use data collected from a bench-scale physical coagulation flocculation basin to create a model that can forecast effluent turbidity eventually control coagulant dosing with machine learning (ML).



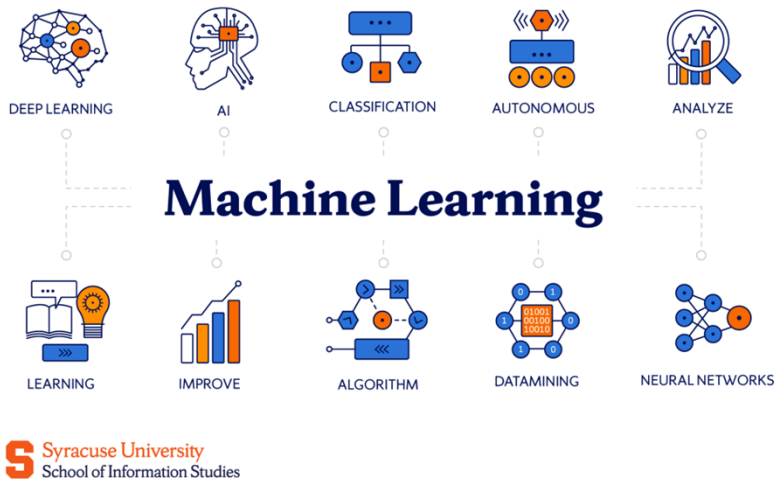


Project Discussed

Oregon State University
College of Engineering

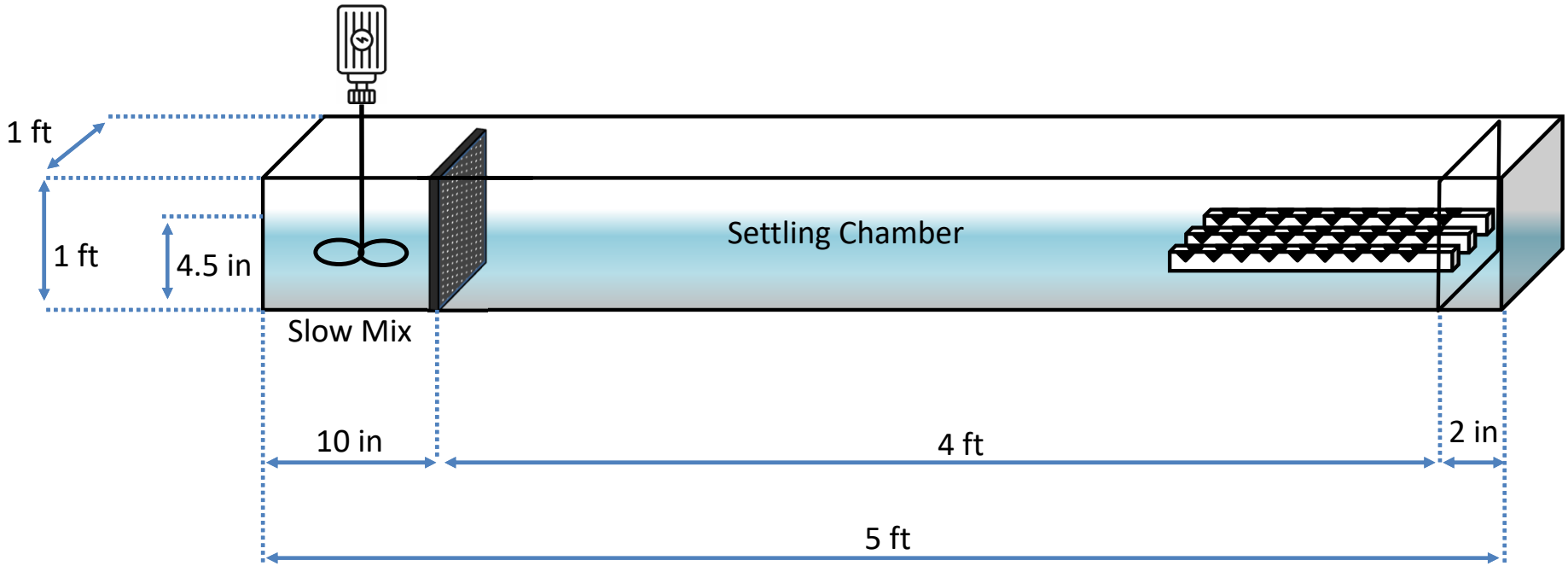
Use data collected from a bench-scale physical coagulation flocculation basin to create a model that can **forecast effluent turbidity** eventually control coagulant dosing with machine learning (ML).





Project Schematic

Oregon State University
College of Engineering



Value in Bench Scale Flow Reactor

Oregon State University
College of Engineering

Advantages of doing bench scale flow reactor:

1. Ability to change influent
   - Conditions Model Experiences
   - Water quality parameters
2. Flexibility in Design
3. Allowed to perform poorly
4. Next level from jar test

Setbacks of Bench Scale

Oregon State University
College of Engineering

Disadvantages of doing bench scale flow reactor:

1. Model creation is highly individualized
2. Difficult to replicate in full scale
3. Lack of historical data

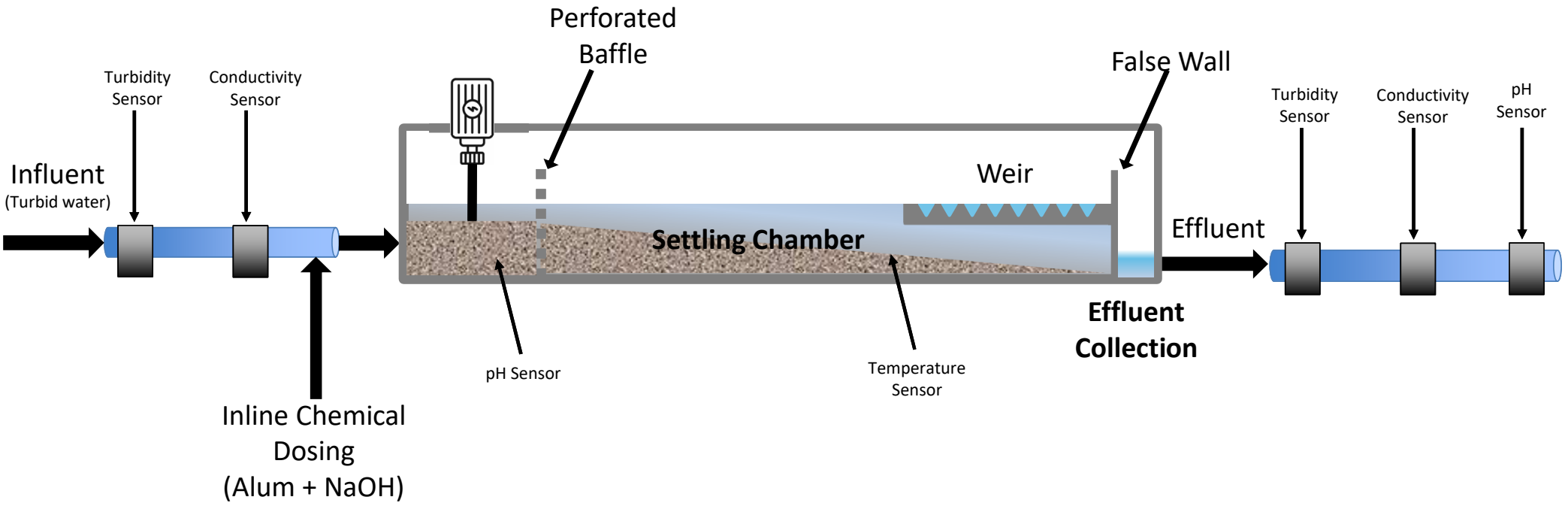
General Experimental Approach

Oregon State University
College of Engineering

1. Sensing Upgrades for Process of Interest
2. Real-time connection between sensing and model
3. Data collection, organization, and formatting
4. Build model, train and test

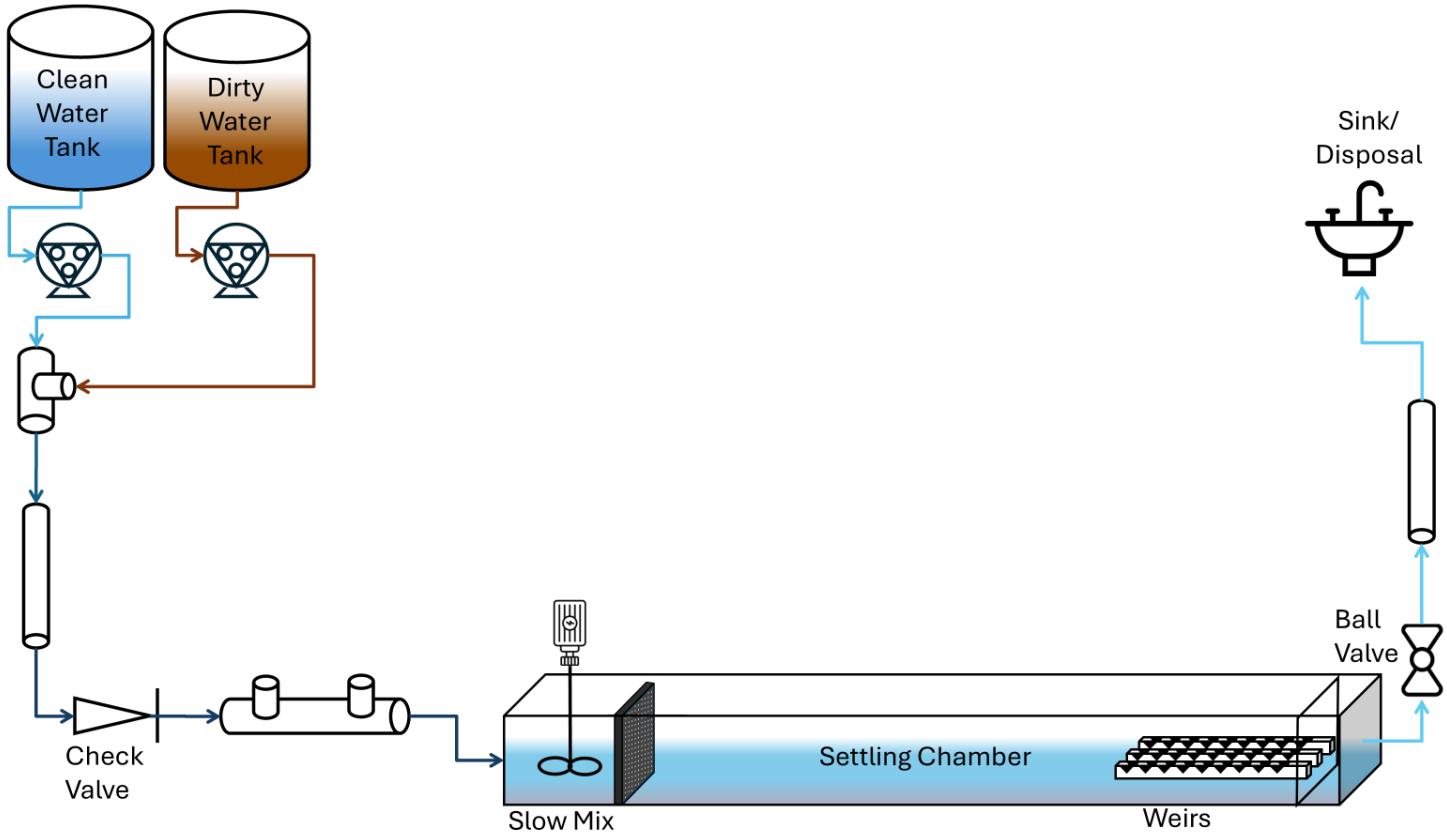
Experimental Setup – Sensing

Oregon State University
College of Engineering



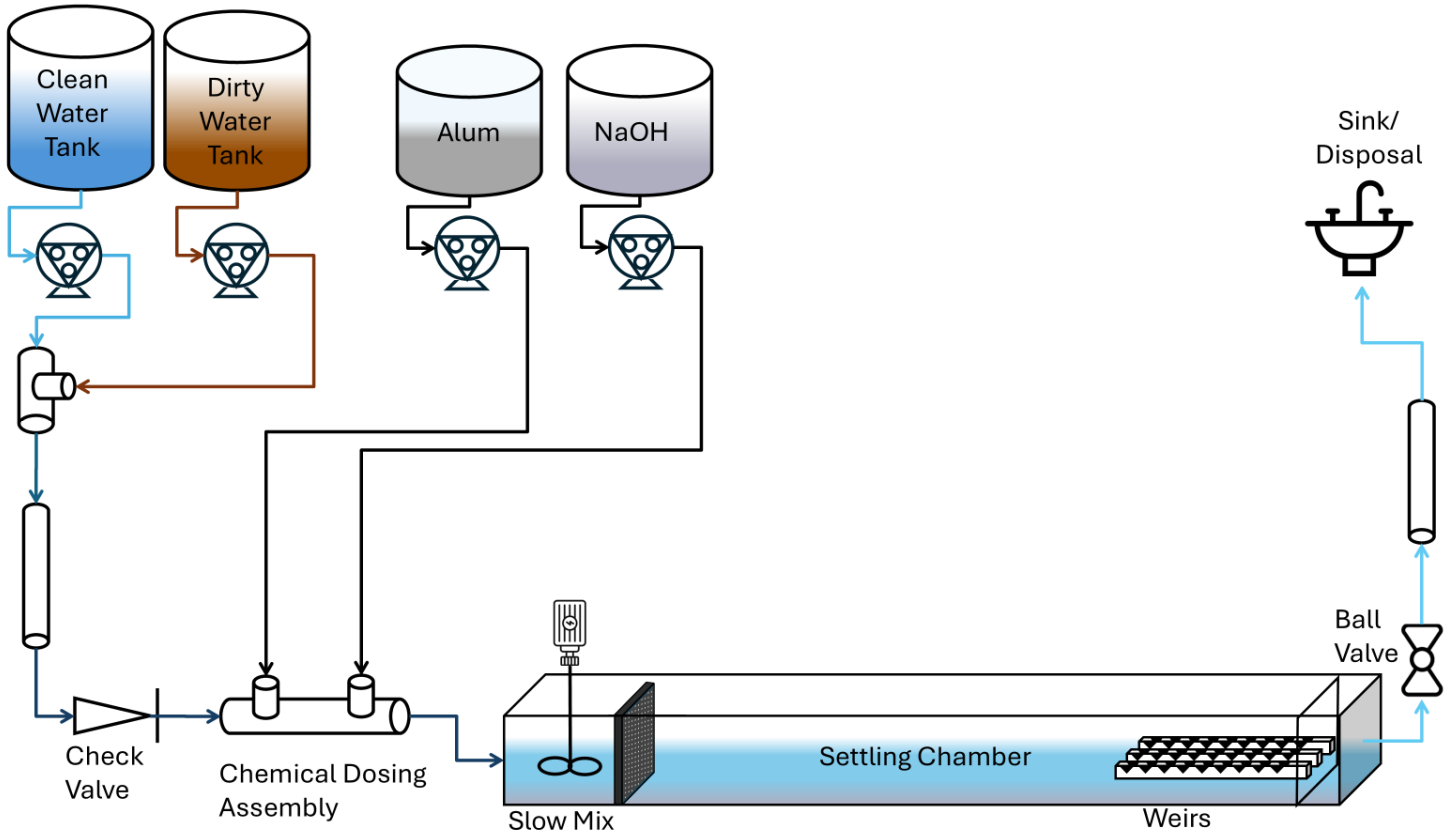
Experimental Setup - Basic

Oregon State University
College of Engineering



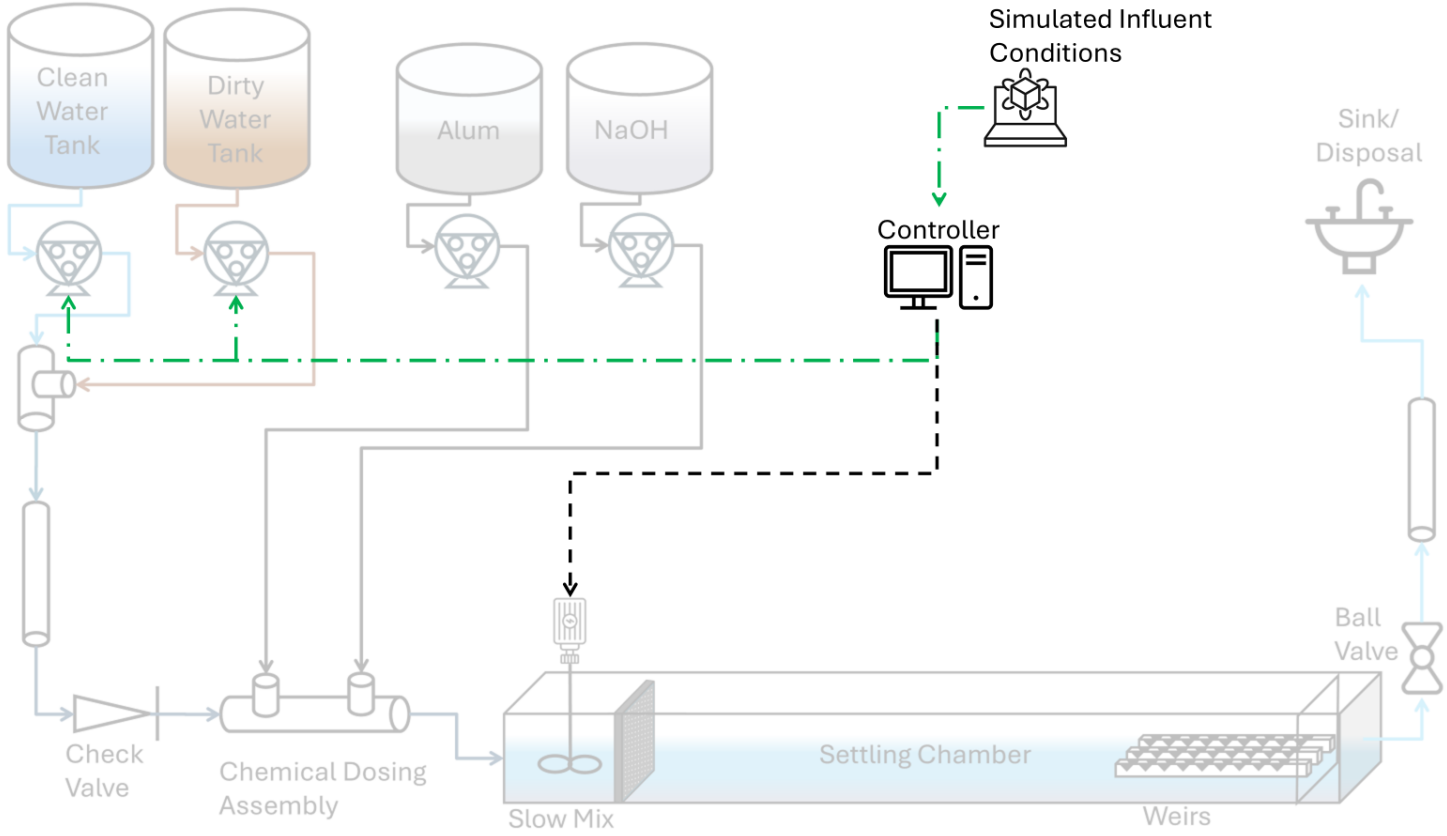
Experimental Setup - Inputs

Oregon State University
College of Engineering



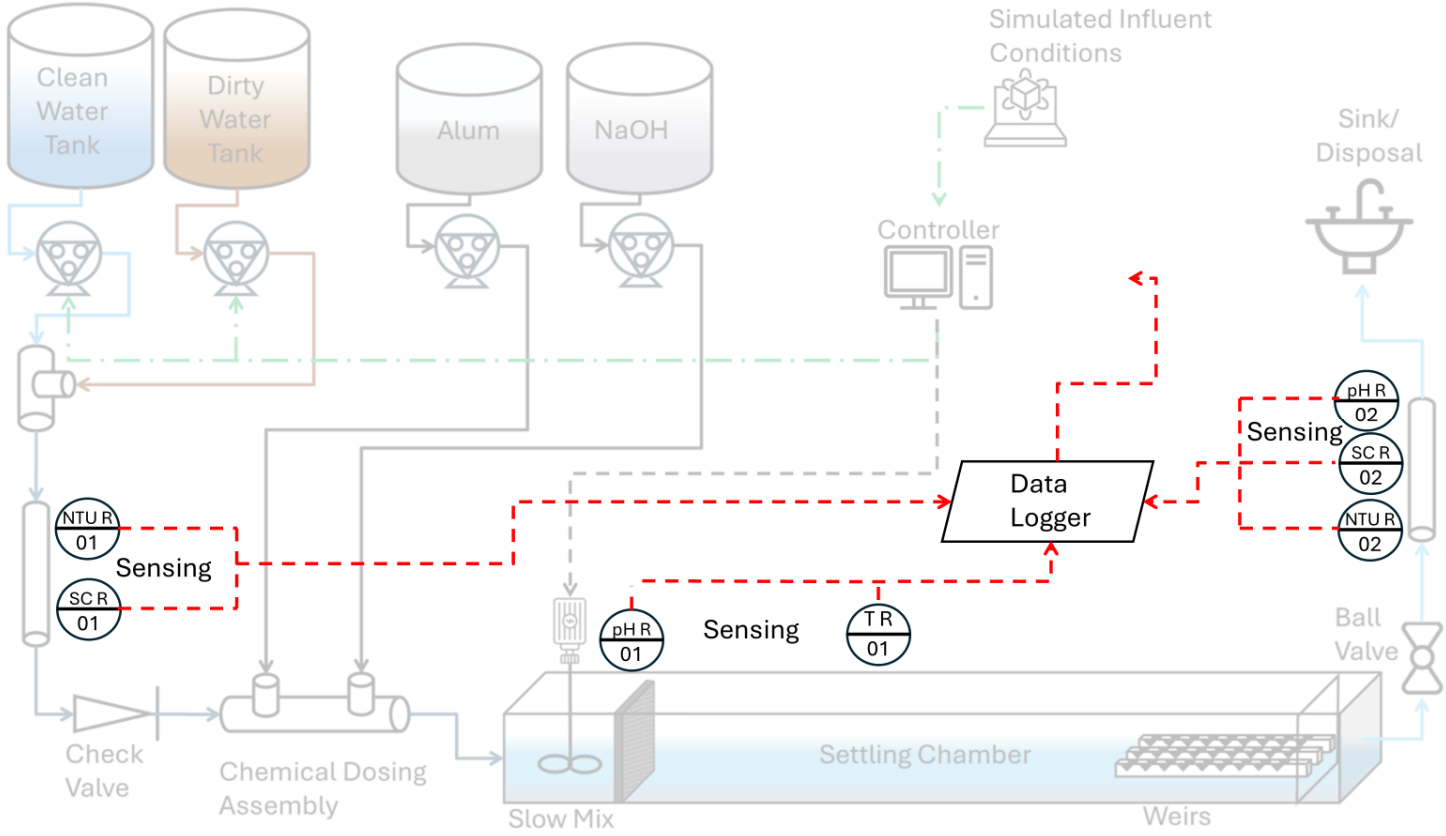
Experimental Setup – Conditions

Oregon State University
College of Engineering



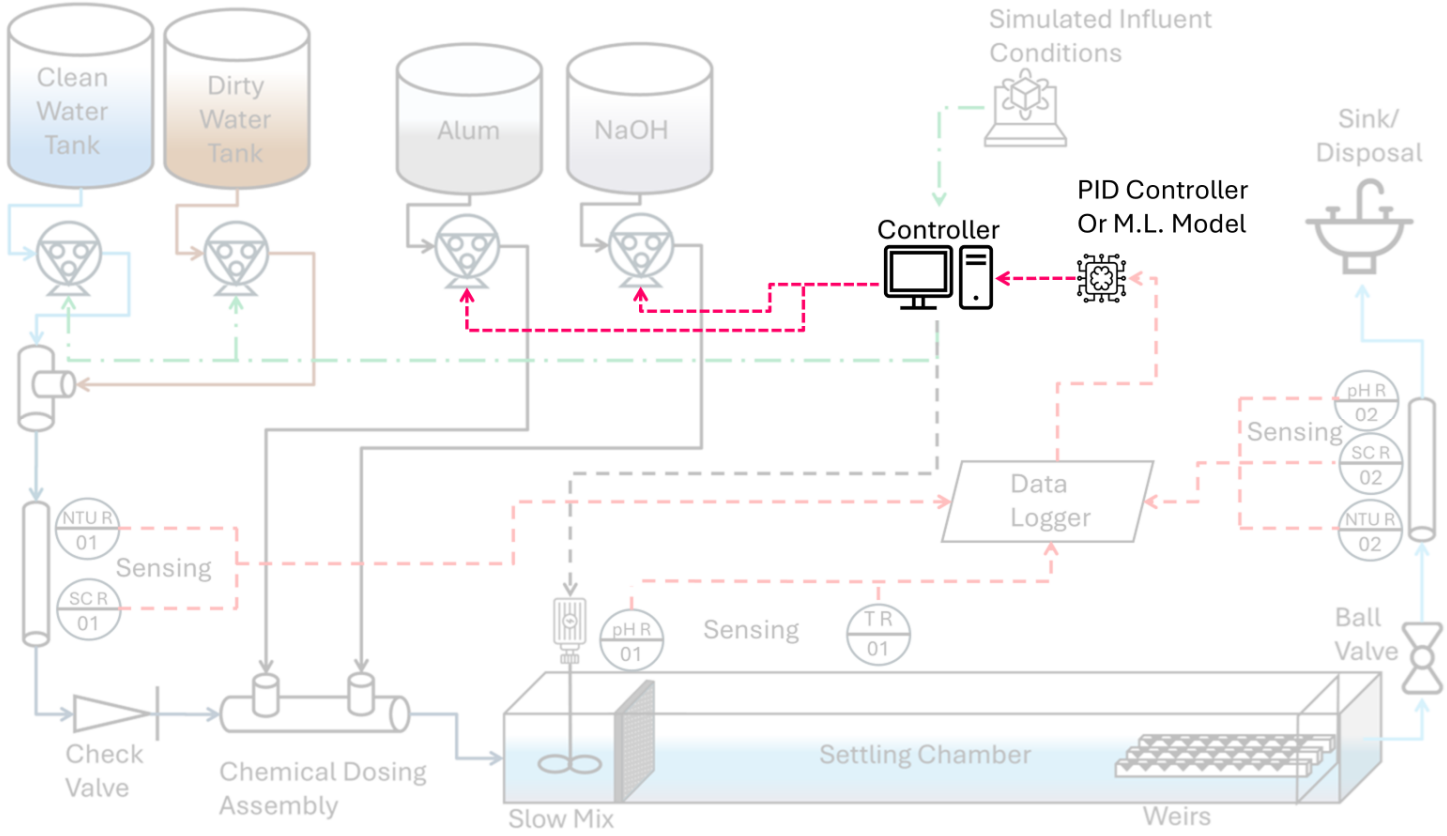
Experimental Setup - Sensing

Oregon State University
College of Engineering



Experimental Setup - Control

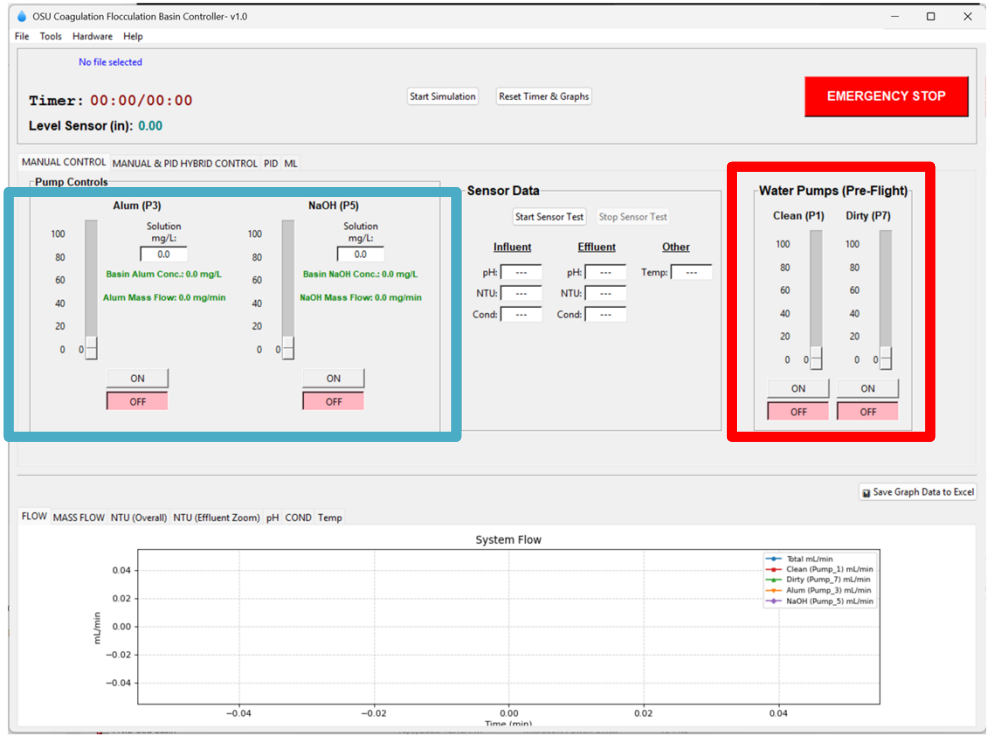
Oregon State University
College of Engineering



Experiment: Live Connection + Data

Oregon State University
College of Engineering

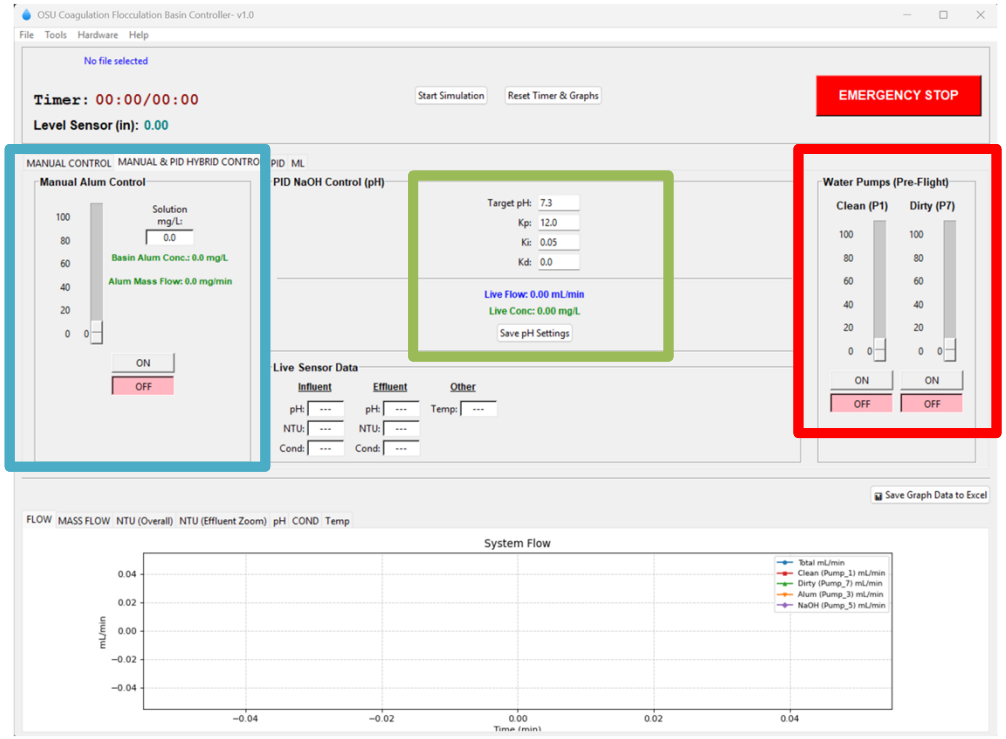
Manual Interface (Built Using Google Gemini)



Experiment: Live Connection + Data

Oregon State University
College of Engineering

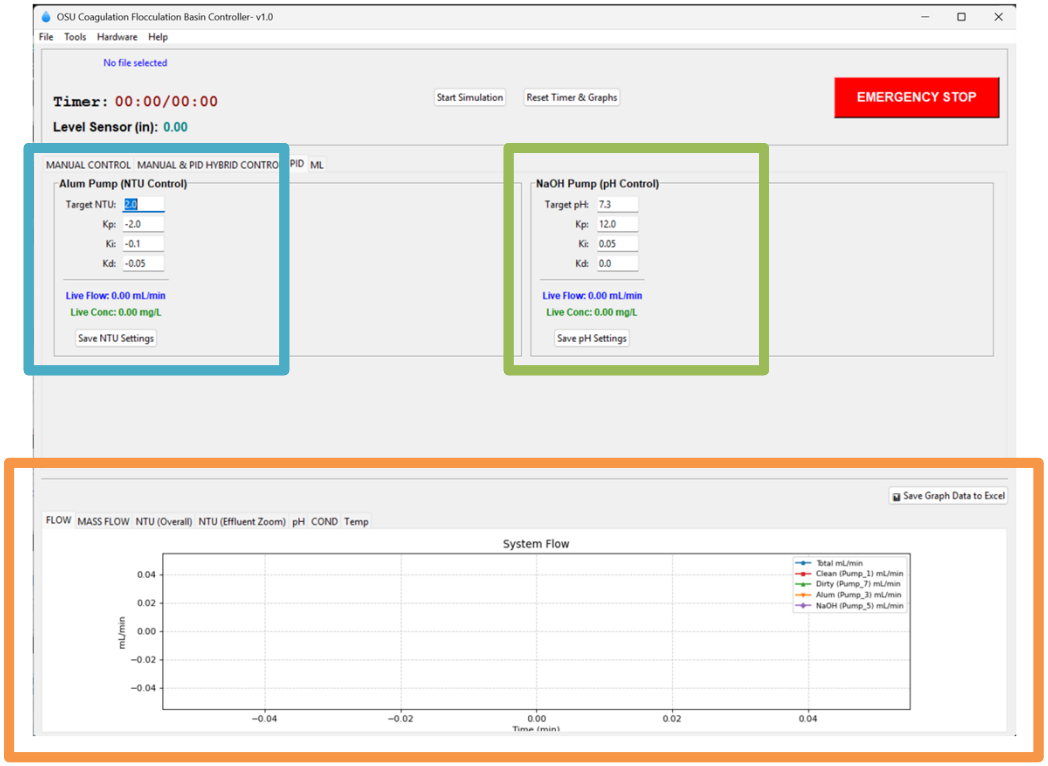
PID + Manual Interface



Experiment: Live Connection + Data

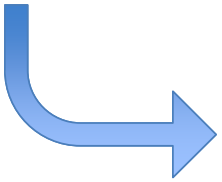
Oregon State University
College of Engineering

PID Interface



Experiment: Planning

1. Set influent turbidity
2. Set pH



Alternate Alum Dose

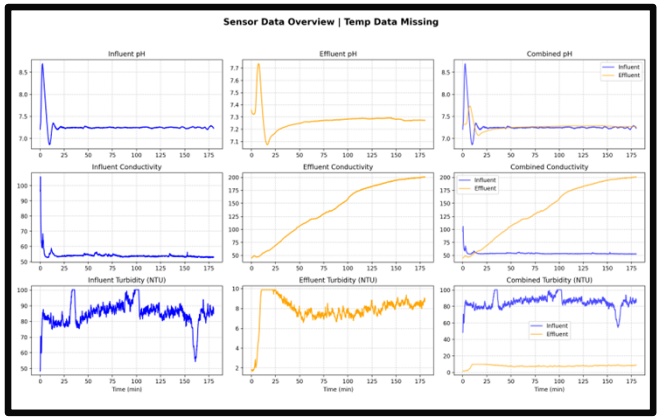
Oregon State University
College of Engineering

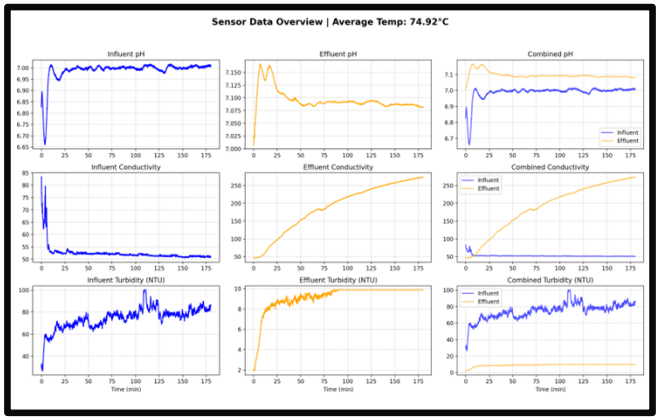
Experimental Planning:

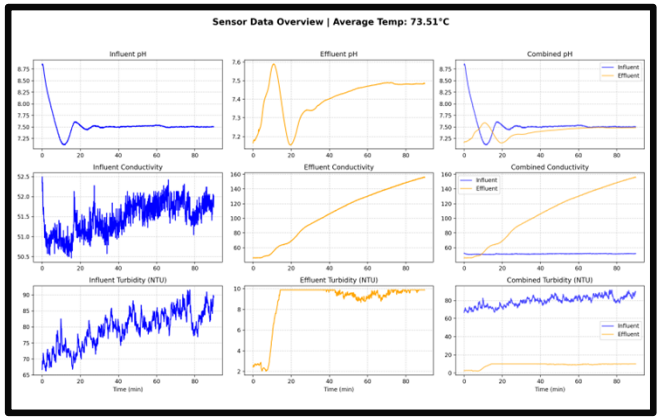
Date	pH	NTU in	Alum (mg/min)	Run Time (hours)
04/25	7.25	75	30	3
04/25	7.25	75	50	3
04/26	7.25	75	70	2.5
04/26	7	75	30	3
04/27	7	75	50	3
04/27	7	75	70	3
04/28	7.5	75	30	2
04/28	7.5	75	50	3
04/28	7.5	75	70	3
04/29	7.3	50	30	2
04/29	7.3	50	50	3
04/29	7.3	50	70	3

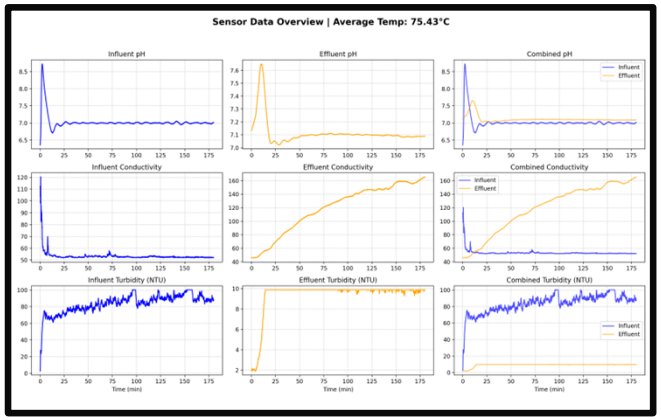
Experiment: Data collection

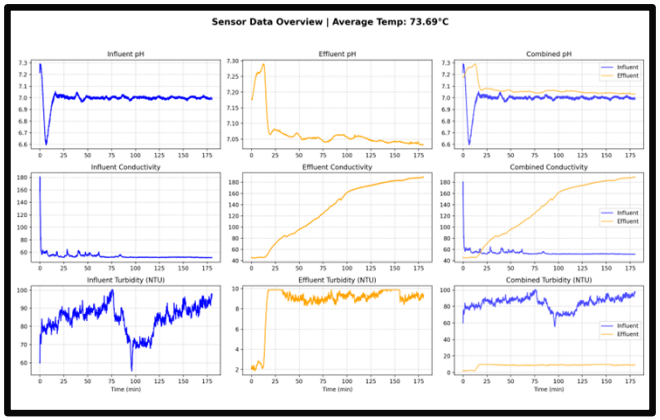
Oregon State University
College of Engineering

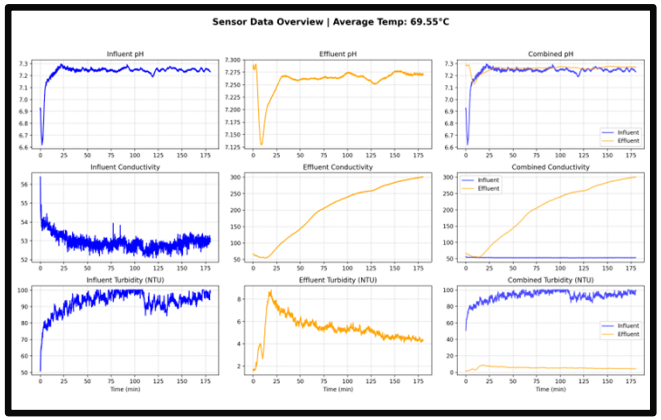












Experiment: Data Organization

Oregon State University
College of Engineering

Originally Collected Data (real time)

Time (HRT)	Turbidity	Time (Forecasted)	Turbidity (Forecasted)
0	0		
0.5	0		
1	**1.0**		
1.5	**2.5**		
2	**3.7**		

Data Shifting and Organization

Oregon State University
College of Engineering

Originally Collected Data (real time)		Same data but in the future	
Time (HRT)	Turbidity	Time (Forecasted)	Turbidity (Forecasted)
0	0	**1**	**1.0**
0.5	0	**1.5**	**2.5**
1	**1.0**	**2**	**3.7**
1.5	**2.5**	2.5	4.2
2	**3.7**	3	5.8

This example is a Forecast 1 HRT in the future

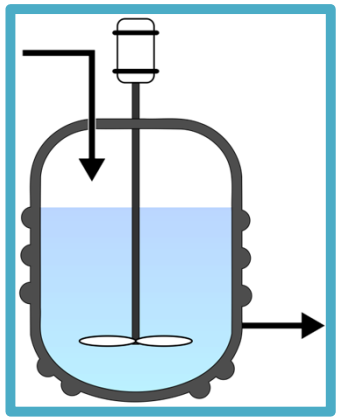
Data Shifting and Organization

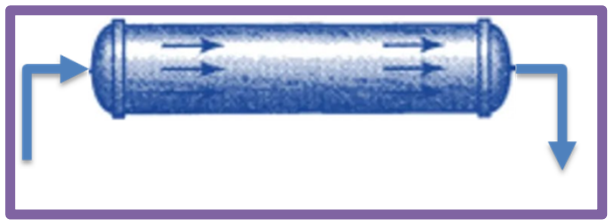
Oregon State University
College of Engineering

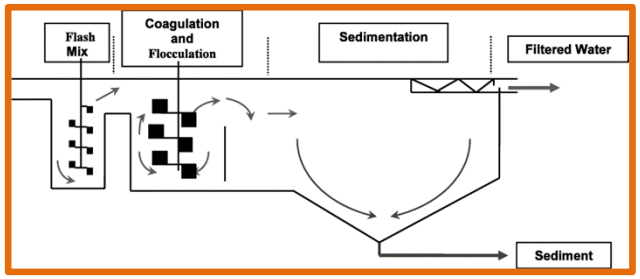
Real time data		Original data in future		Prediction Based on Historical Data	
Time (HRT)	Turbidity	Time (forecasted)	Turbidity (forecasted)	Time (predicted)	Turbidity (predicted)
0	0	**1**	**1.0**	**1**	.9
0.5	0	**1.5**	**2.5**	**1.5**	2.3
1	**1.0**	**2**	**3.7**	**2**	3.9
1.5	**2.5**	2.5	4.2	2.5	4.0
2	**3.7**	3	5.8	3	5.7

Experiment: Reactor Consideration

Oregon State University
College of Engineering



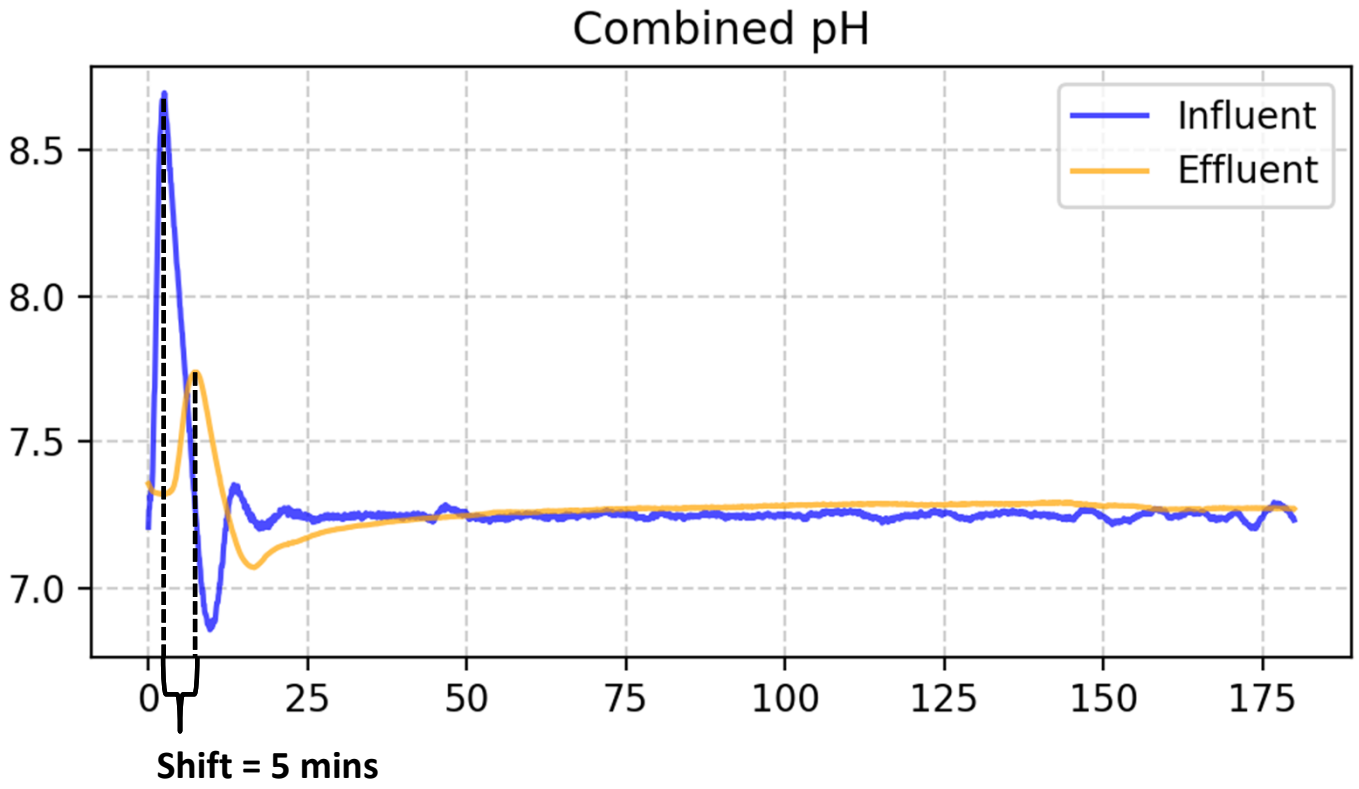




Reactor Type	Continuous Flow Stir Tank Reactor (CFSTR)	Plug Flow Reactor	**Hybrid/Non-ideal Reactor**
Changes Detected at Effluent	Almost immediately	1 HRT	**Observed**

Experiment: Data Shifting

Oregon State University
College of Engineering



Experiment: ANN Structure

Oregon State University
College of Engineering

Inputs (real time data):

Influent & Effluent Turbidity

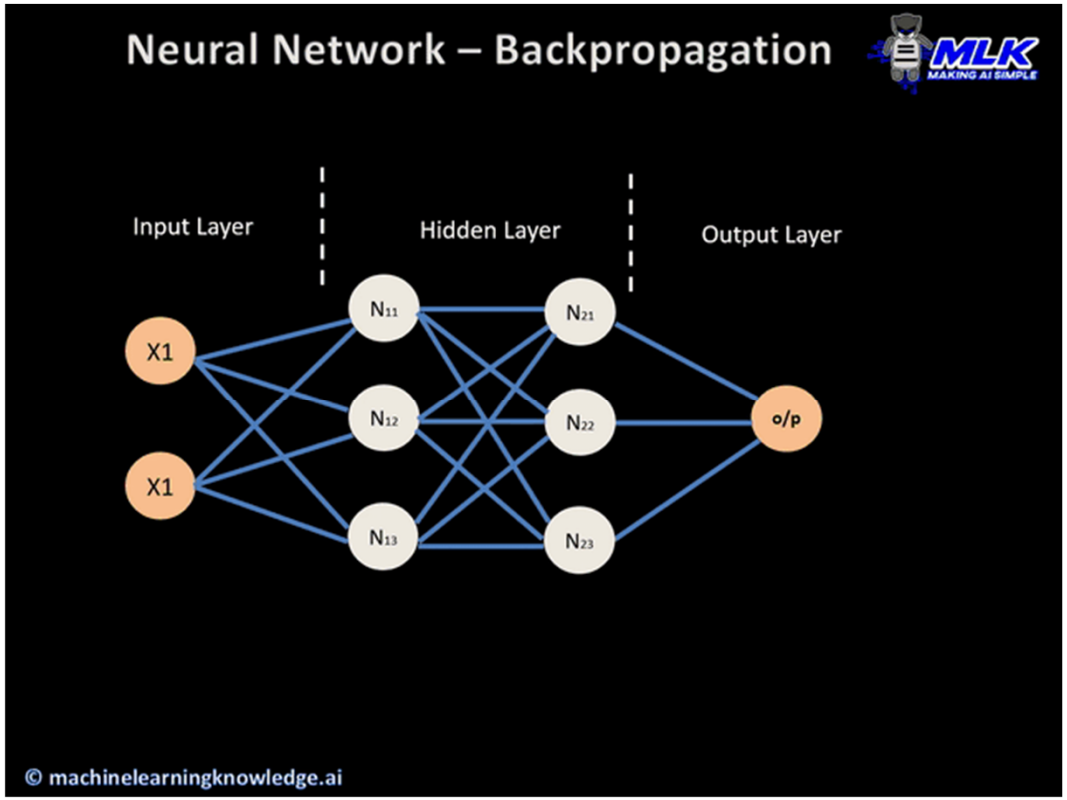
Influent & Effluent Conductivity

Influent & Effluent pH

Temperature

Total Flow Rate

Mass Flow chemicals



Output:

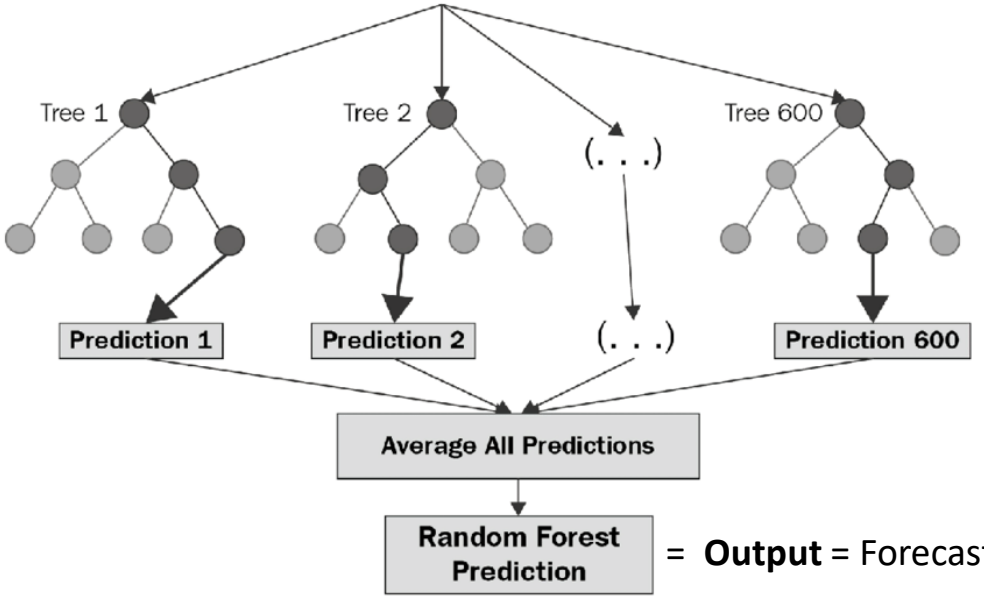
Forecasted Effluent Turbidity

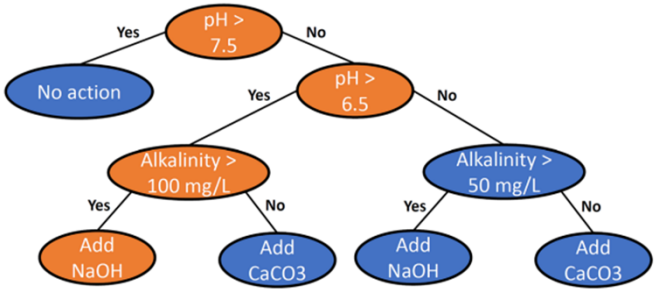
Experiment: RF Structure

Oregon State University
College of Engineering

Input:

influent + effluent real time values; flow rates; mass flow



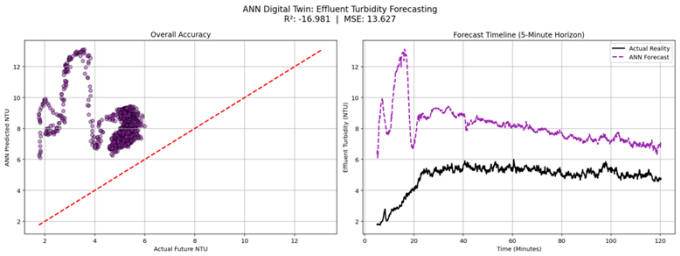


= **Output** = Forecasted Effluent Turbidity

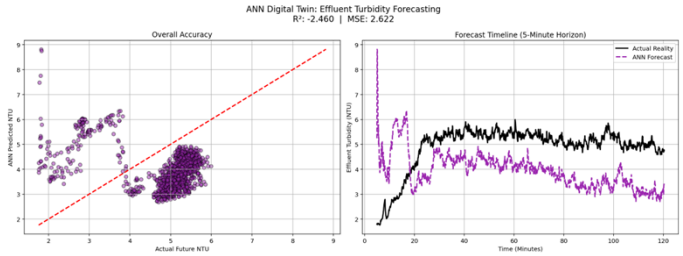
Model Training and Testing

Oregon State University
College of Engineering

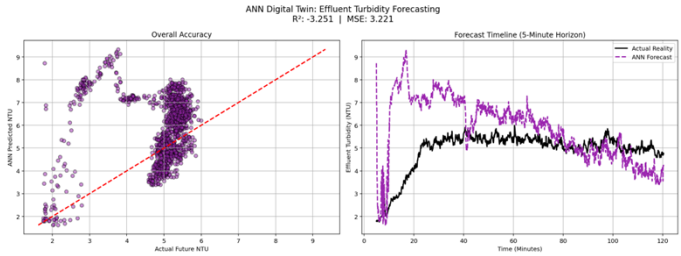
Model 1



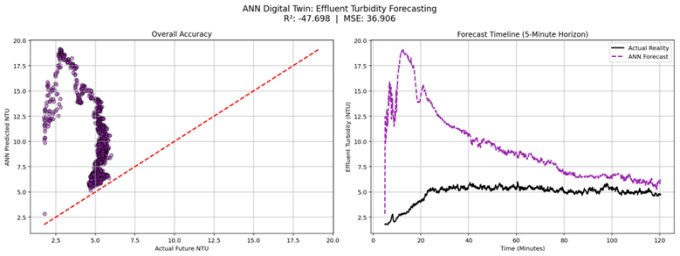
Model 2



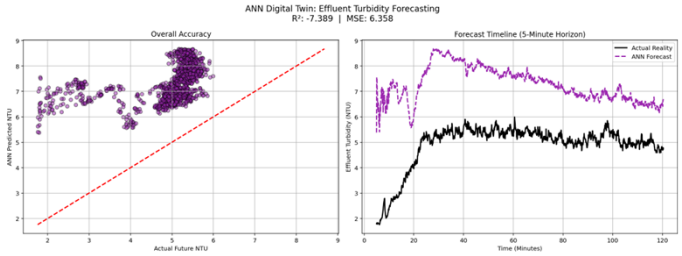
Model 3



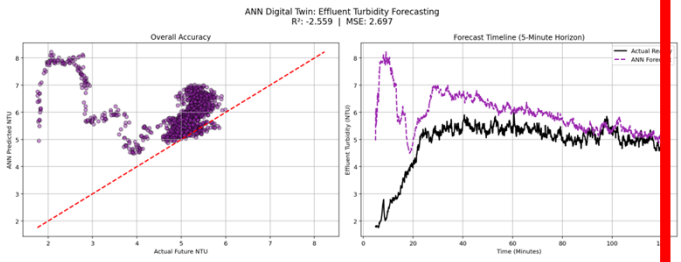
Model 4



Model 5



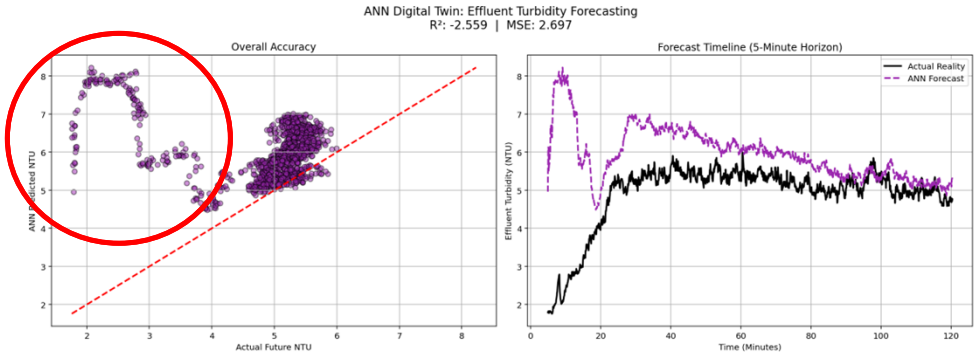
Model 6



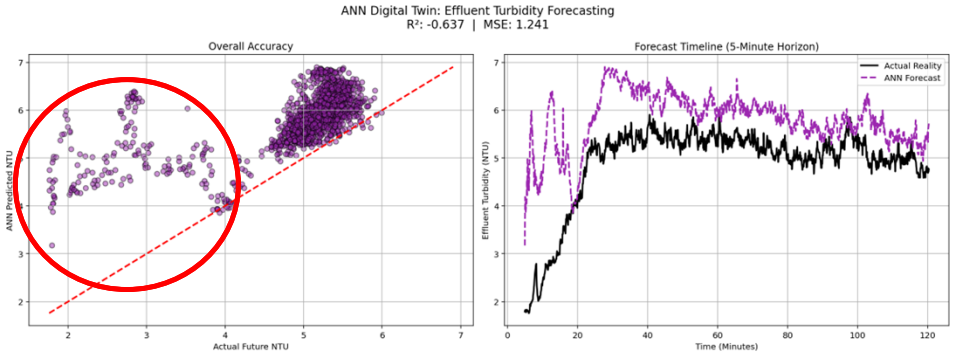
Model Training and Testing

Oregon State University
College of Engineering

Model 6



Model 7



Model missing key information here

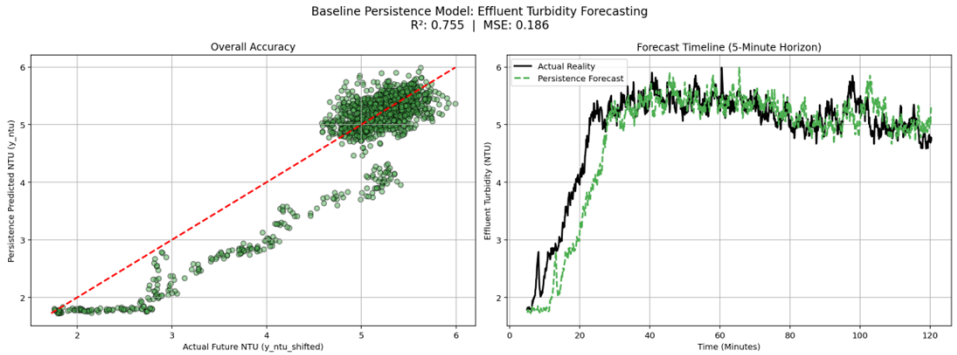
Added additional training data similar to this testing data.

Improved model as a result

Model Training and Testing

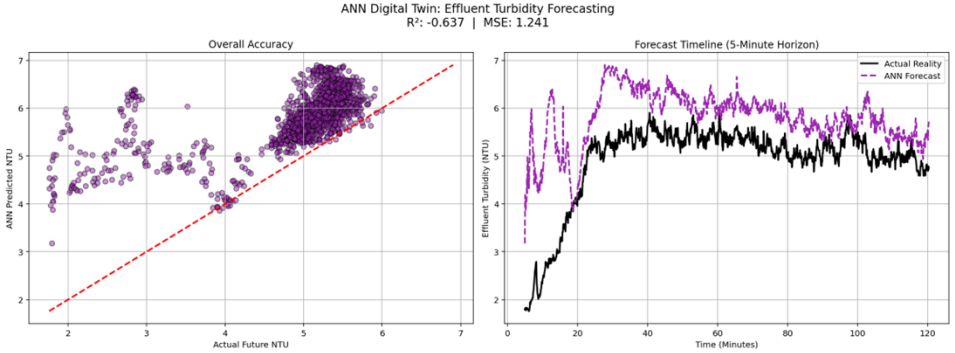
Oregon State University
College of Engineering

Persistence Model



PERSISTANCE = ASSUME CURRENT TURBIDITY VALUE IS SAME 5 MINUTES IN FUTURE

Model 7



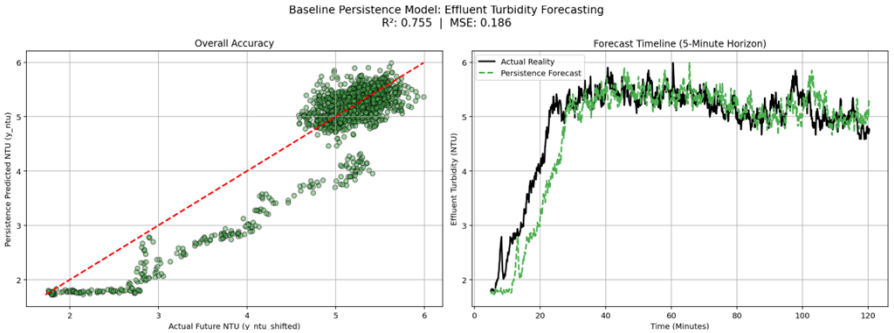
The model was built with too short of a forecast.

Could not outperform persistence

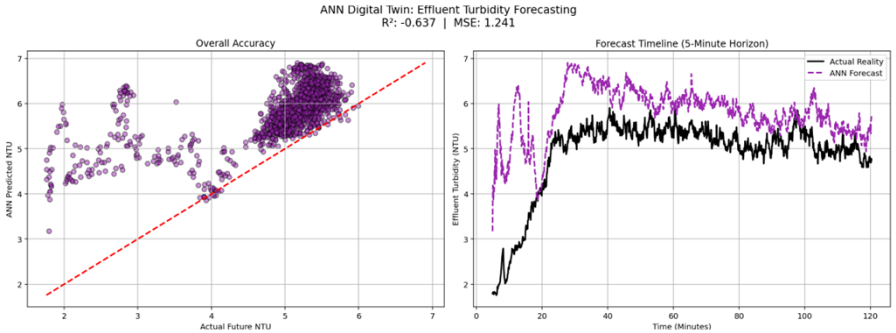
Models Evaluated

Oregon State University
College of Engineering

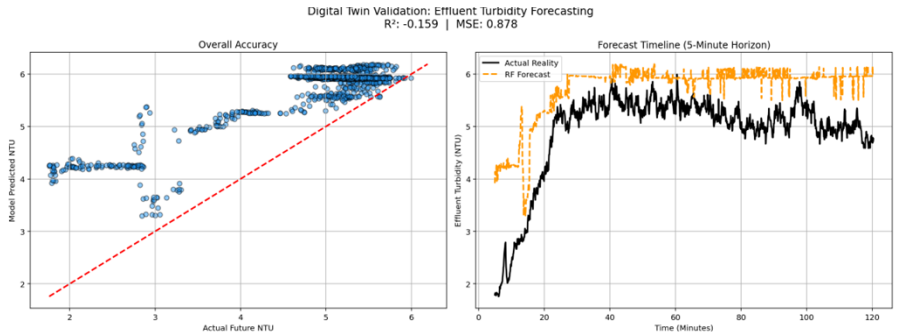
Persistence Model



ANN Model



Random Forest Model



Big takeaways

Oregon State University
College of Engineering

Sensing

SCADA Flexibility

Data Preparation

Data Space

Model Type

What's Next?

Oregon State University
College of Engineering

Collect More Data

- Train model that can forecast further into the future
- Create rich data space to work with
- Train different types of models (RNN, Linear Regression, etc...)

Forecast ML model integration into PID controller

- Shift into feedforward control
- Improvement in stability
- Decrease tendency to over-dose



Oregon State University
College of Engineering

Thank you

Questions

Perry Wilson – wilsoper@oregonstate.edu