

Task 5.17.1.3: Non-Parametric Statistics for RO LRV Modeling

- Intact reverse osmosis (RO) removes pathogens by over 99.999%. Online surrogates can show a minimum log reduction value (LRV) that is being achieved at any time. However, the sensors for these surrogates are prone to outliers, skewing the data distribution.
- We proposed an advanced statistical method called the Shewhart Sign Chart. This method reduced alarms by ignoring non-consecutive outliers (Figure 1).

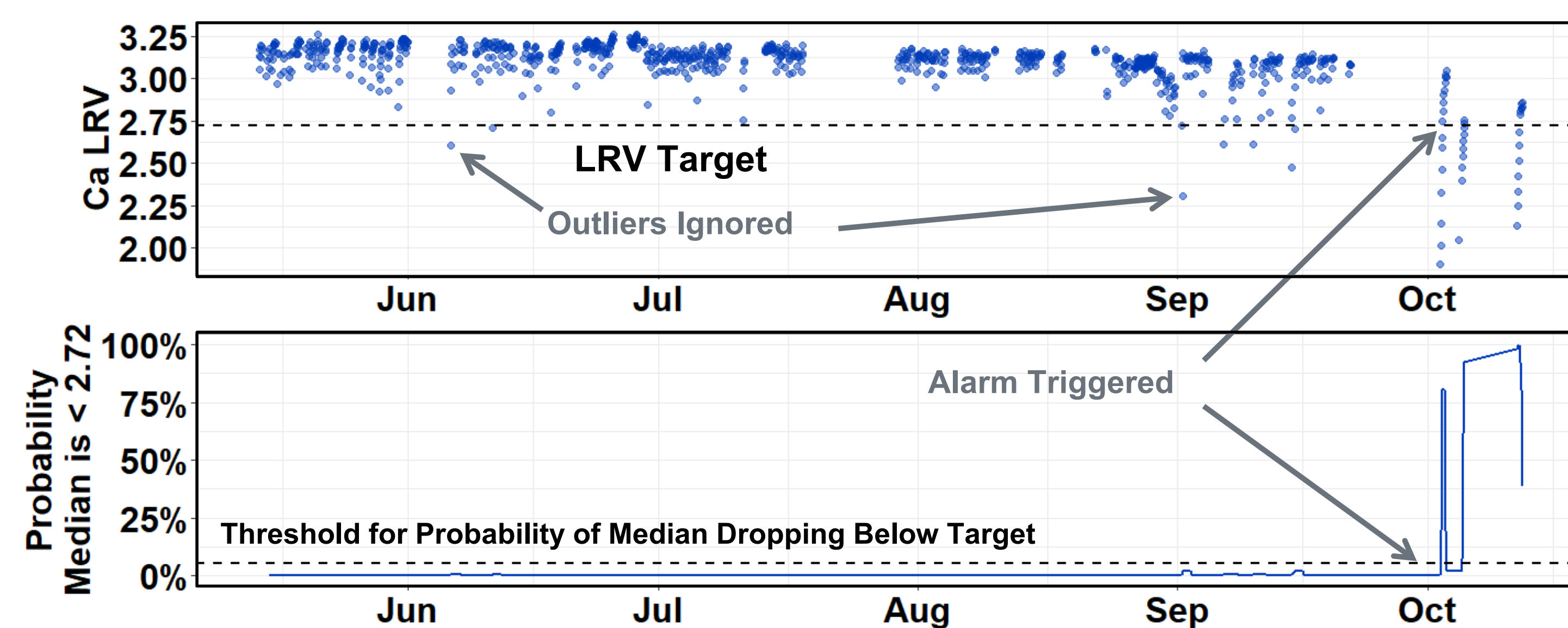


Figure 1

- In 2024-2025, this method was implemented on one full-scale RO train via a cloud dashboard at OCWD. It is now being expanded to all RO trains at OCWD with the following data architecture (Figure 2).

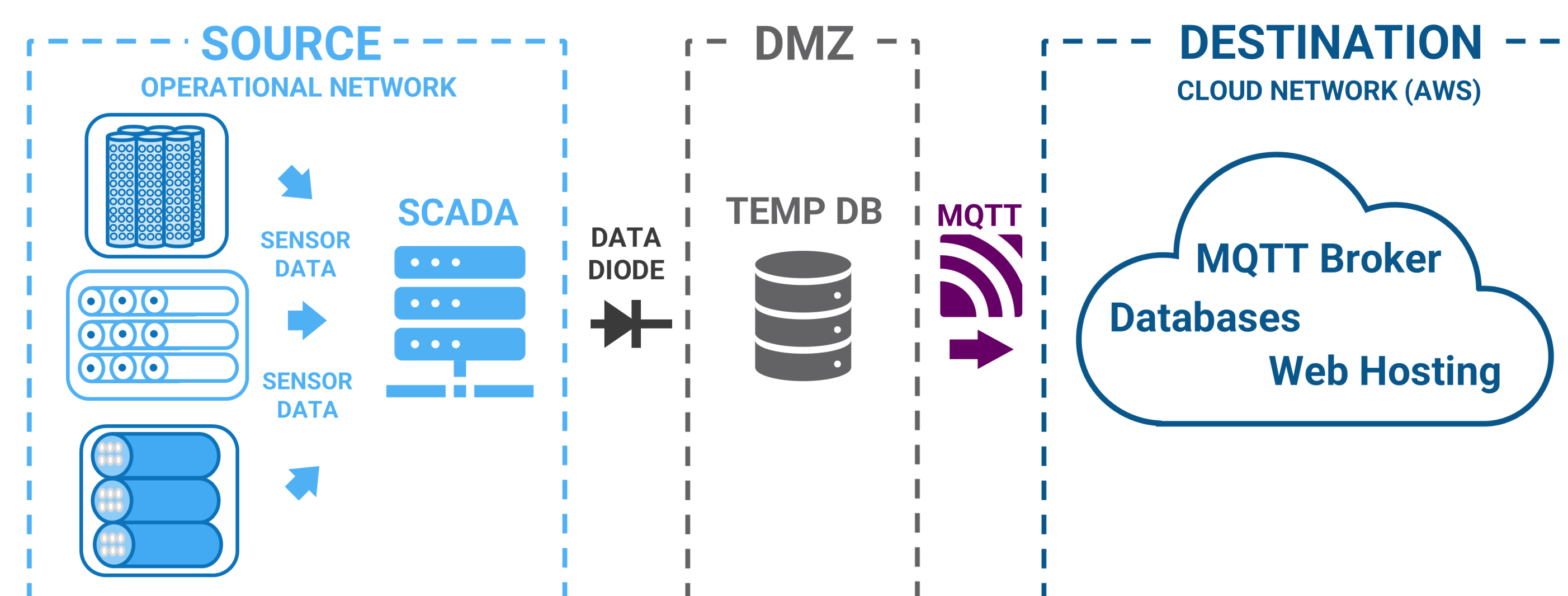


Figure 2

Task 5.17.1.4: Supervised Machine Learning Regression for N-Nitrosodimethylamine (NDMA)

- NDMA is a key disinfection byproduct in reuse. Monitoring NDMA in real-time is not possible with commercial sensors. UV/AOP is operated conservatively, assuming the maximum historical NDMA concentration. NDMA machine learning (ML) soft sensors could make potable reuse more resilient and efficient by dynamically updating UV/AOP control to reflect actual influent NDMA concentrations.
- An ML model for NDMA was accurate within ± 3 ng/L based on prior data from OCWD (Figure 3).
- This accuracy would correspond to a 21% UV energy saving, even with a 95% safety factor, with a 0.69 ng/L NDMA target.

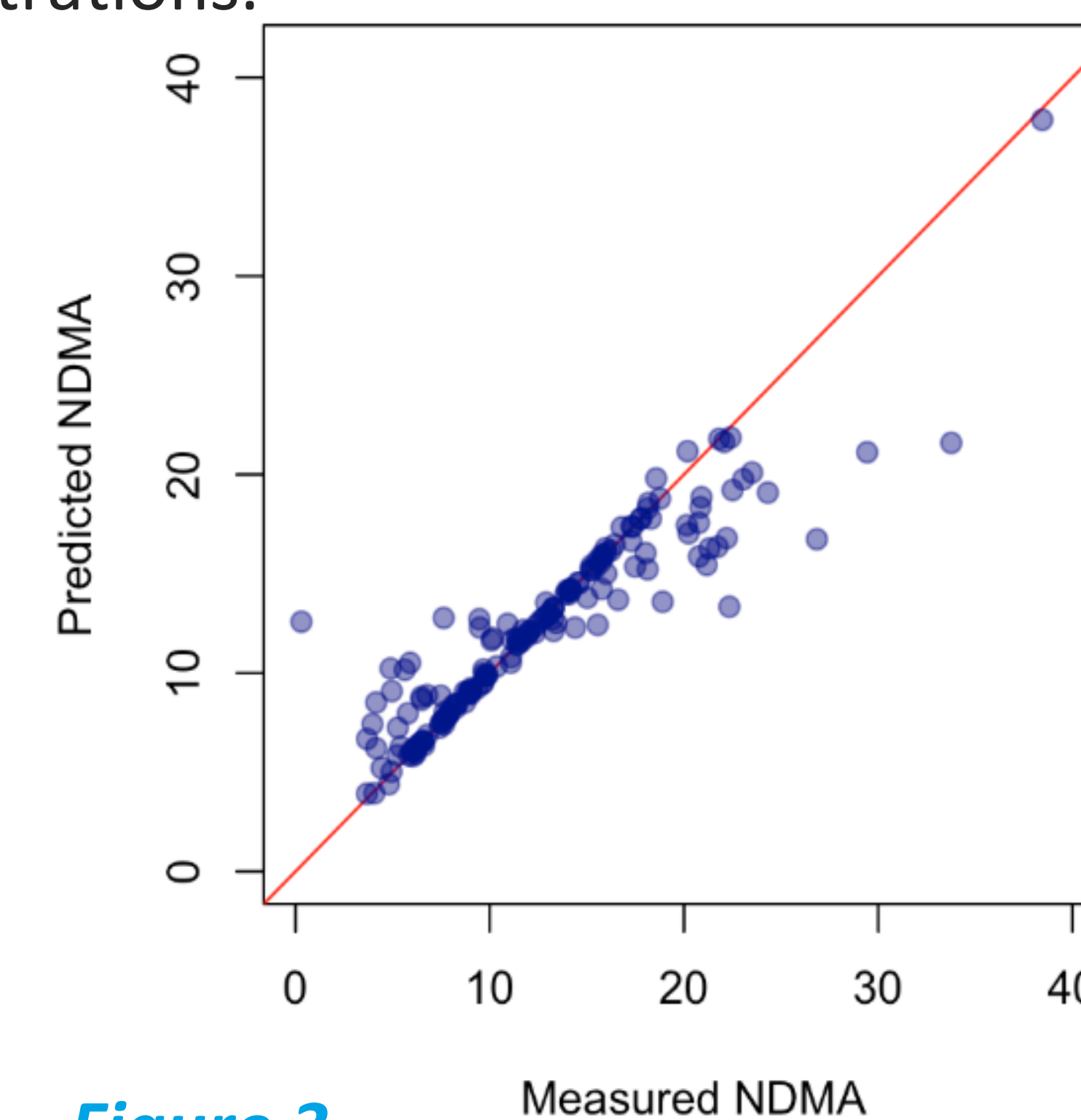


Figure 3

LVMWD Results

- We collected 182 NDMA samples over 10 months to replicate this model at LVMWD, which has an NDMA target for surface water augmentation reuse in California.
- The LVMWD model was less accurate, due to a spike in NDMA concentrations in the training set that the features could not explain.
- The dataset is being expanded with historic data to determine if more time points improve the model.

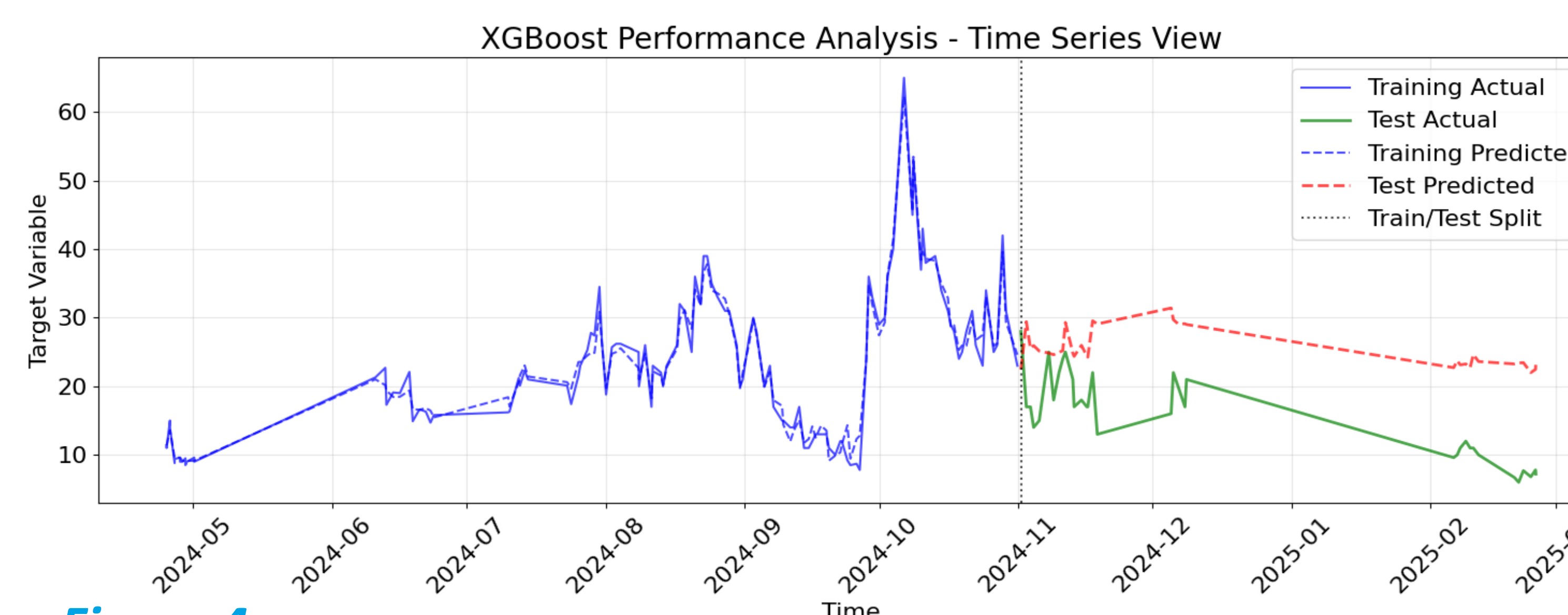


Figure 4

NAWI CONNECTIONS

Period of Performance: 2022-2026

Challenge Area/Topic Area: Municipal Wastewater; Autonomous Water; Resilient Treatment

NAWI Leverage: This project is collaborating with NAWI Project 5.08 Advanced Process Controls – Autonomous Control and Optimization via a Project Enhancement Grant (PEG).

KEY FINDINGS AND CONCLUSIONS

Task 1.3

Key Findings:

- Most RO LRV surrogates evaluated (Ca, S, Sr, TOC, and Peak C Fluorescence) were not normally distributed; most were left-skewed and S was bimodal.
- Left-skewed data would cause excessive alarms if assumed normal.
- A Shewhart Sign Chart could detect true events in 3 hr while reducing alarms by >50% compared to alarms on single points.

Conclusions:

- Shewhart Sign Chart for RO LRV monitoring would meet the **Reliability & Availability** pipe parity metric by reducing false alarms by 50%.
- Scale-up is feasible through real-time cloud dashboards.

Task 1.4

Key Findings:

- NDMA concentrations are highly correlated to pressure drops across RO membranes and influent ammonia concentrations.
- The optimal model inputs depend on model type. PCA is an effective method of dimension reduction for process units like RO-banks with many correlated features. In this case, all PCA-based models performed as well or better than mRMR, a feature selection approach.
- While different model types may compare similarly on average, many predict extreme outliers in simulated real-time forecasts (e.g., support vector machines and k-nearest neighbors).
- Model performance depends on location, due in part to different levels of data variance.

Conclusions:

- ML-based adaptive UV control could save energy at some locations.
- More data is needed to explain seasonal NDMA patterns at other locations.

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Alliance for Water Innovation (NAWI), funded by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), Industrial Efficiency and Decarbonization Office under Funding Opportunity Announcement DE-FOA-0001905. This work was also completed in part by OCWD for the U.S. Bureau of Reclamation under Contract No. 0704-0188.