HARNESSING MACHINE LEARNING AI MODELS FOR PREDICTING BIOLOGICAL OXYGEN DEMAND IN SECONDARY EFFLUENT OF WATER RESOURCE RECOVERY FACILITY

Muhammad Hassnain[1], Sarada M.W. Lee[2], Muhammad Rizwan Azhar[1]

1. School of Engineering, Edith Cowan University, 270 Joondalup Dr, Joondalup, Western Australia, Australia.
2. Perth Machine Learning Group, Western Australia, Australia.

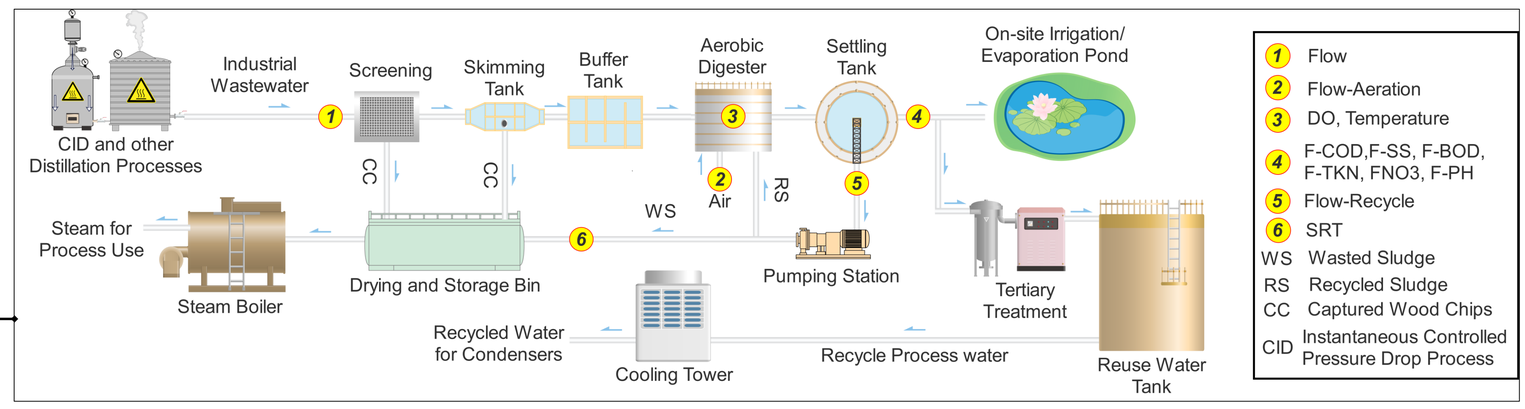
1 INTRODUCTION

- Effluent Biological Oxygen Demand (F-BOD) is a key indicator of wastewater treatment performance.
- Traditional 5-day BOD testing delays operational feedback, limiting timely interventions.
- Industrial effluents, particularly from essential oil manufacturing, pose unique challenges due to high organic loads, acidity, and variability.
- This study investigates machine learning (ML) models as “soft sensors” to predict F-BOD in real-time, reducing reliance on time-intensive laboratory testing.

Objectives

- Develop and validate ML/AI models for accurate F-BOD prediction. Identify the key operational and chemical parameters that influence effluent quality.
- Integrate predictive models into a graphical user interface (GUI) for Engineers & operator use.
- Simulate operational boundaries to support real-time process optimisation, compliance and efficient process control.

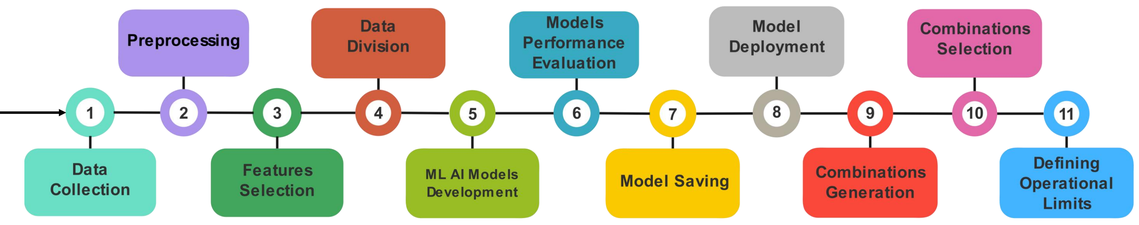
Process flow diagram for the Industrial Water Resource Recovery Facility (WRRF), treating essential oil wastewater.



2 METHODS

- **Dataset:** 19 years (2005-25) of lab and instrumentation records from the WRRF. (~54K samples with 12 features)
- **ML Models:** 16 ML models tested, including regression, boosting, tree and neural network based models.
- **Evaluation:** R^2, RMSE, MAE; feature importance and ablation studies.
- **Explainability and User Interfaces**: SHAP for interpretability. Simulation GUI for engineers & operators.

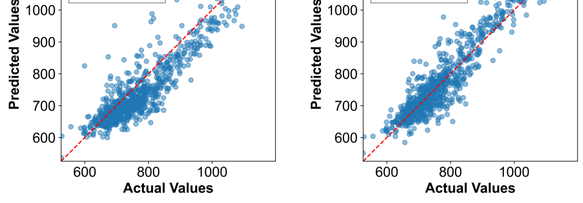
Models Development & Deployment Pipeline



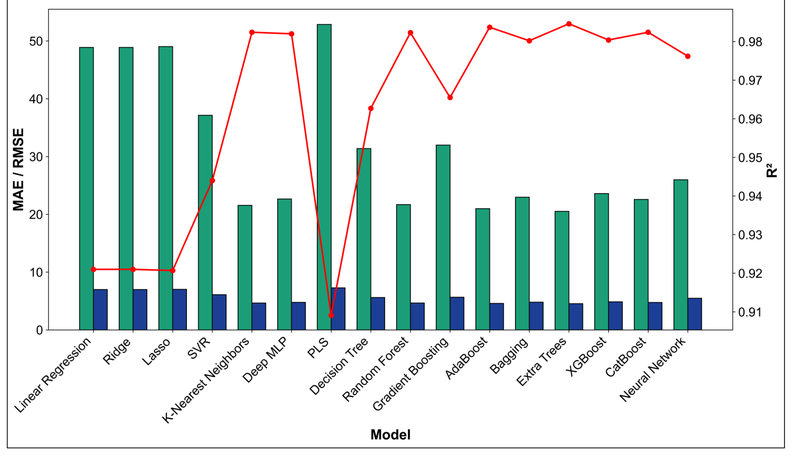
3 RESULTS

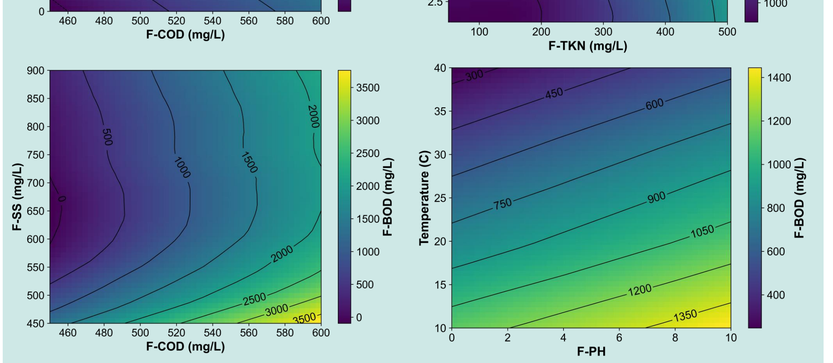
- **Accurate BOD predictions:** Extra Trees achieved the best accuracy (**$R^2 \approx 0.96$**). **Best for simulation:** Neural Networks applied to operational simulation. **Key predictors:** F-COD, influent flow, recycle sludge flow. **Contour plots revealed:** High F-COD + high flow → F-BOD spikes. Low pH / low temp → microbial inhibition/lower F-BOD.
- **Impact:** Soft-sensed predictions available instantly vs a 5-day lab test. Accelerated decision-making and process control.

Scatter Plots of Predicted vs Actual F-BOD for Two Top-Performing ML Models



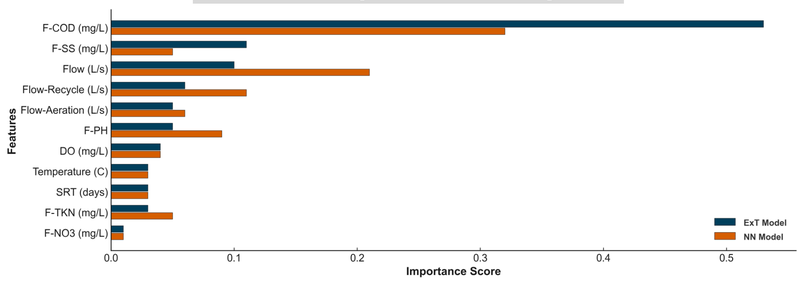
Model Performance Comparison across Various ML Models



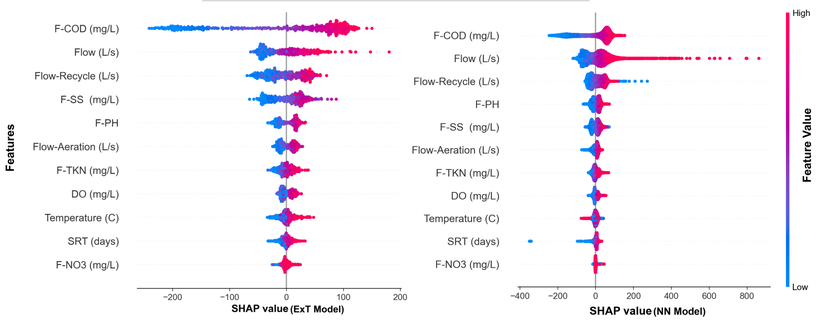


Contour plots showing the effects of key operational and influent parameters on F-BOD concentrations in the industrial WRRF.

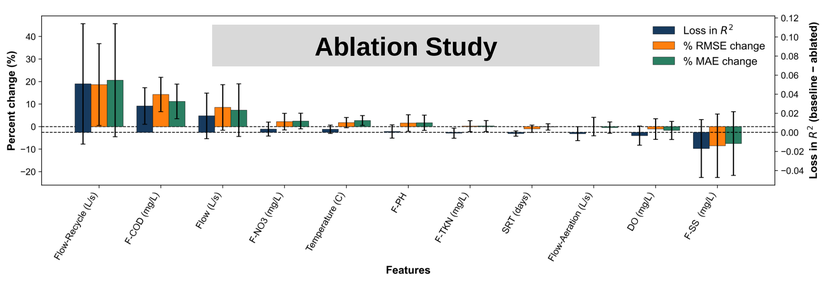
Feature Importance Analysis



SHAP Analysis



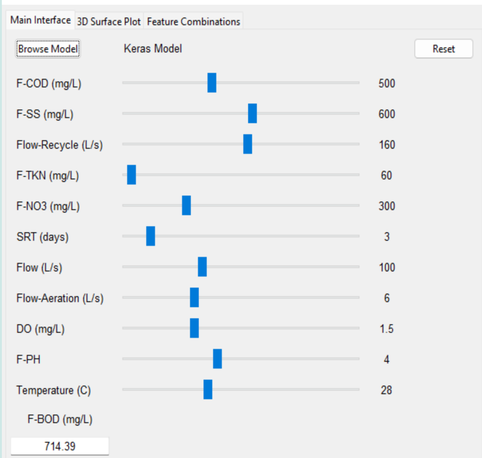
Ablation Study



Simulated operational boundary conditions based on key input features, corresponding target F-BOD limits, and their associated process implications and broader impacts when accepted thresholds are exceeded.

Operational Limit	Input Features Conditions Set	Process Implication	Impact
F-BOD (mg/L) > 200 mg/L	Flow-Aeration* (L/s) < 4 DO (mg/L) < 2.5 Other conditions are determined by the BOD upper limit.	Over aeration	Energy wastage
F-BOD (mg/L) < 2500	F-COD (mg/L) < 520 F-SS (mg/L) > 550 Flow-Recycle* (L/s) F-TKN (mg/L) < 400 F-NO3 (mg/L) > 600 SRT* (days) < 10 Flow* (L/s) < 250 Flow-Aeration* (L/s) > 1.5 DO (mg/L) > 1 F-PH* > 10 and < 4 Temperature (C) > 40 and <15	Organic load surge to the tertiary treatment plant and algal blooms in the onsite irrigation pond	Potential oxygen depletion, imbalance in aquatic ecosystems, increased sulphide generation in the conveyance system, eutrophication with higher nutrient load, proliferation of pathogens, and operational issues in tertiary treatment.

The simulation GUI with the NN model delivers instant F-BOD predictions, enabling operators and engineers to quantify the effects of input features and take prompt process-control actions when approaching or exceeding allowable limits.



4 CONCLUSION

- ML models can replace routine BOD_5 tests as a soft sensor, saving cost and time.
- Provides real-time insight for WRRF operators/Engineers.
- Improves safety by reducing lab sampling.
- Scalable to other industrial and municipal WRRFs.
- **Future work:** Multivariate predictions, SCADA integration.

REFERENCES & ACKNOWLEDGEMENT

Hassnain, M., Lee, S. M. W., & Azhar, M. R. (2025). *Soft sensing of biological oxygen demand in industrial wastewater using machine learning models.* **Journal of Water Process Engineering, 78**, 108699. https://doi.org/10.1016/j.jwpe.2025.108699

Scan the QR code to access the full paper online

I, Muhammad Hassnain, gratefully acknowledge the **School of Engineering, Edith Cowan University**, for sponsoring and supporting this research and my participation at **Next Water 2025**.

Scan the QR code to email the presenters.