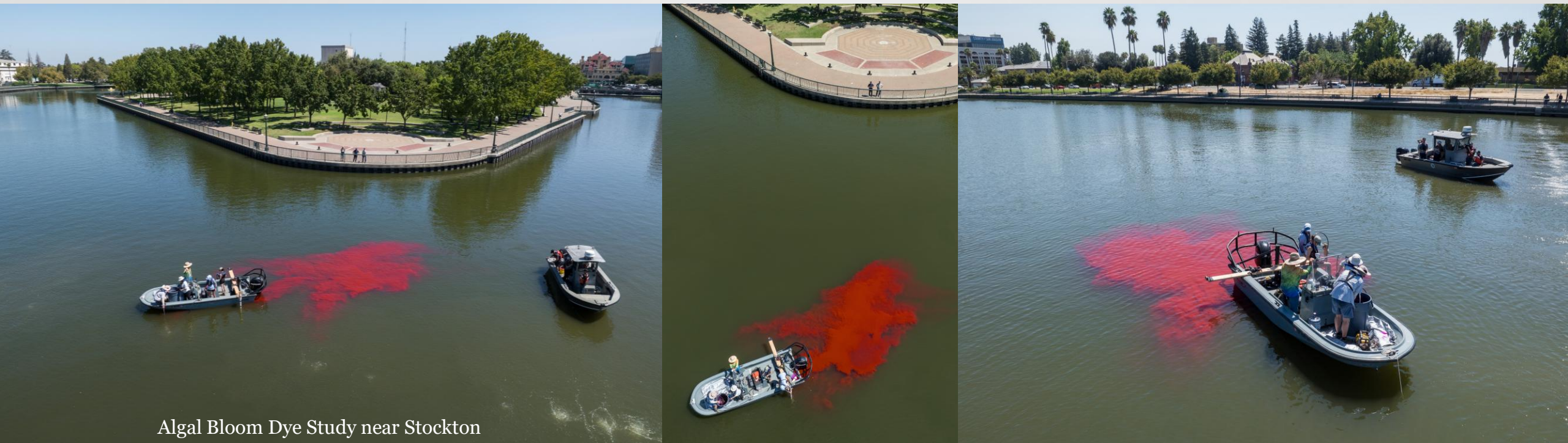


Predicting the Microcystis Visual Index (MVI) using a Data-driven Approach to Support Harmful Algal Bloom (HAB) Management in the Delta

32nd CWEMF Annual Meeting, April 22, 2026



Algal Bloom Dye Study near Stockton



Presentation Outline

01

Background & Motivation

HABs in the Delta, MVI reclassification

02

Study Data & Methods

*1006 samples, 8 features, 8 ML models,
Improvements in the HAB work 2026*

03

ML Model Performance

Combined score, threshold optimisation

04

Ensemble Predictions

Top 4-models ensemble, uncertainty analysis

05

Feature Importance Results

Permutation Importance, SHAP analysis

06

Key Findings

Consensus across methods, analysis outcomes

07

Future Work

Clustering, UMAP & PCA, misclassification analysis

- * MVI – Microcystis Visual Index
- * SHAP – Shapley Additive exPlanations
- * PCA – Principal Component Analysis
- * UMAP – Uniform Manifold Approximation and Projection



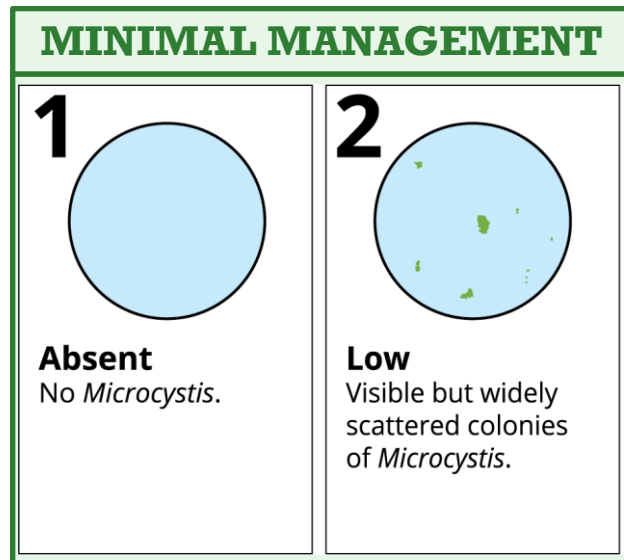
Background — HABs in the Sacramento–San Joaquin Delta

Why does it matter?

- Cyanobacteria HAB growth **increases** with **warming climate** in the **Delta**.
- **Produced toxins** affect fish, pets, wildlife, and humans — **liver cancer** and **neurological damage** risk.
- **Five classes** of **Microcystis Visual Index (MVI)** were reclassified to **binary** MVI Classes for actionable HAB management: **Minimal** (MVI=1,2) vs **Mandatory** Management (MVI=3–5).

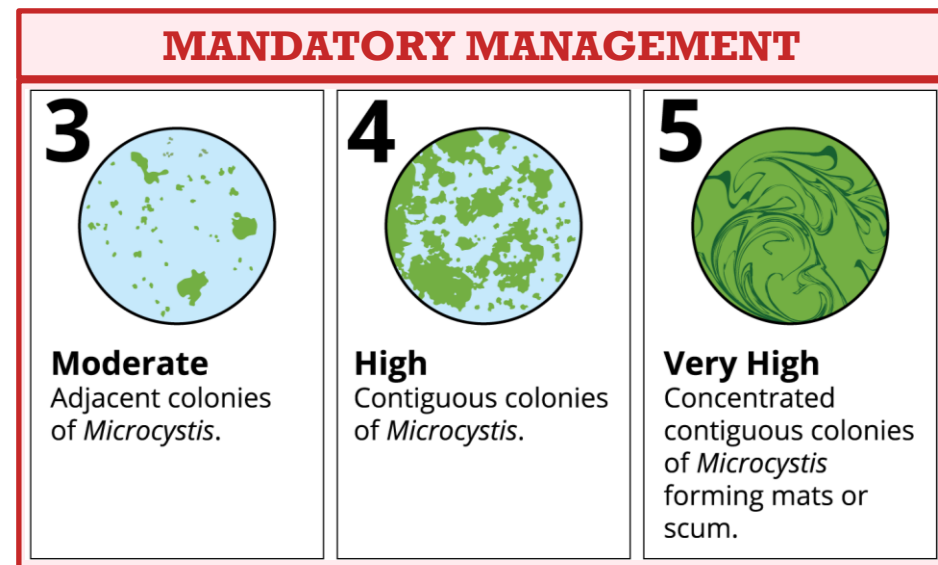
MVI Reclassification

Disclaimer: **Mandatory Management** here indicates that *management* is deemed *necessary*.



~85% of samples

Source: Flynn et al. 2022 |
Bouma-Gregson et al. 2024



~15% of samples ← **class imbalance**

Study Data

1,006 Samples

*Discrete Daily Samples
(Summer & Fall 2014–2022)*

**35 Sites
168 Locations**

Sacramento–San Joaquin Delta

8 Input Features

*Water quality, nutrients, turbidity
antecedent flow and velocity*

8 Input Environmental Features

Water Quality

- ✓ Water Temperature (°C)
- ✓ Specific Conductivity (µS/cm)

Nutrients

- ✓ Dissolved Ammonia
- ✓ Dissolved Nitrate & Nitrite
- ✓ Dissolved Orthophosphate

Physical

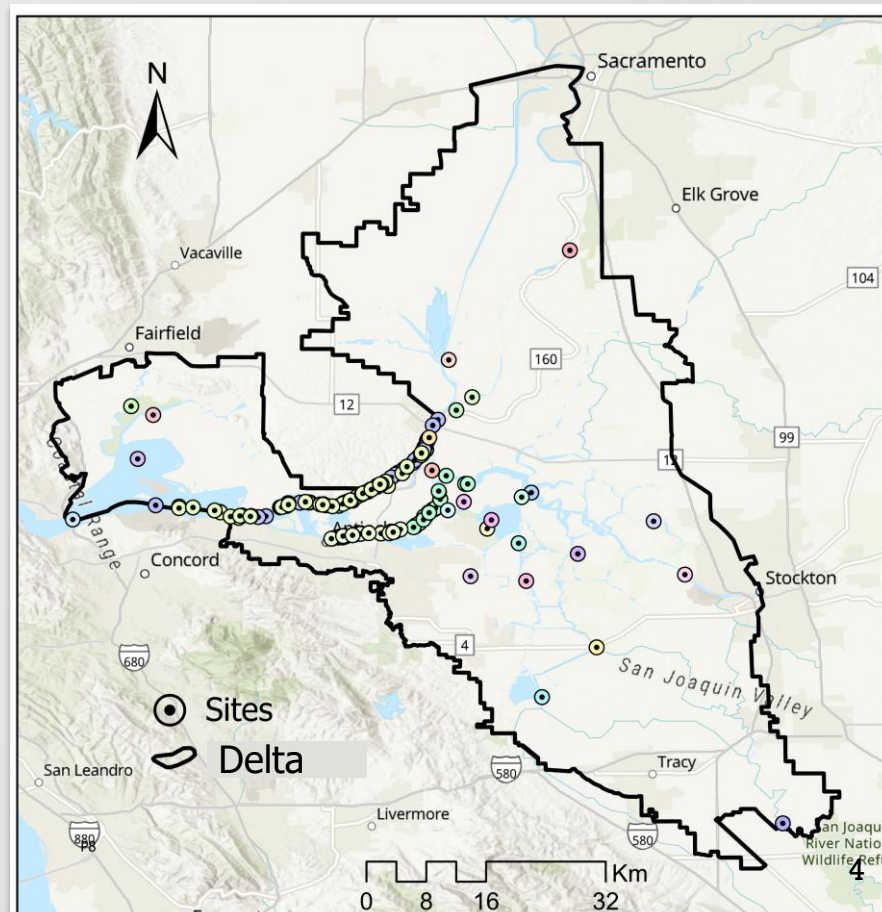
- ✓ Antecedent Flow (1-month Avg.) *
- ✓ Antecedent Velocity (1-month Avg.) *

** DSM2 Simulated*

Light

- ✓ Turbidity (FNU)

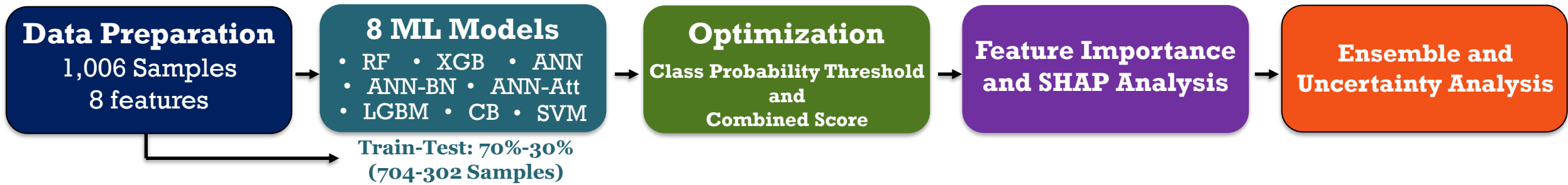
*Data Source: Colleagues from
Division of Integrated Science and Engineering*



CALIFORNIA DEPARTMENT OF
WATER RESOURCES

ML Models and Modeling Workflow

MVI Binary Classification Prediction Modeling Workflow



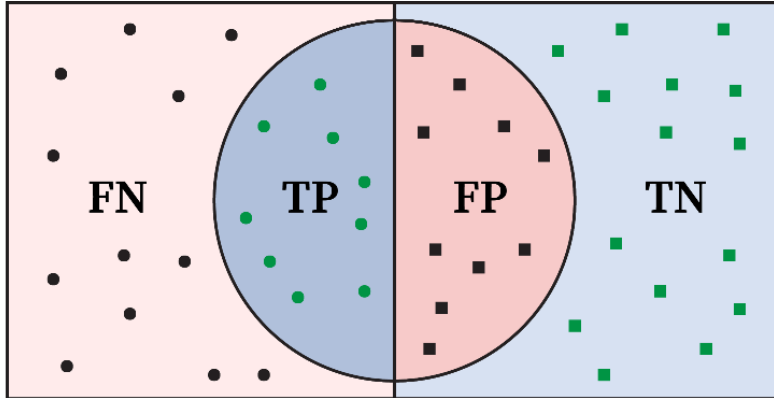
8 Models Trained & Evaluated

RF	XGB	ANN	ANN-BN	ANN-Att	LGBM	CB	SVM
Random Forest	XGBoost	ANN	ANN-BatchNorm	ANN-Attention	LightGBM	CatBoost	SVM

$$\text{Combined Score} = 0.35 \times \text{Accuracy} + 0.25 \times \text{Recall(Mandatory)} + 0.20 \times \text{Precision(Mandatory)} + 0.15 \times \text{Recall(Minimal)} + 0.05 \times \text{Precision(Minimal)}$$

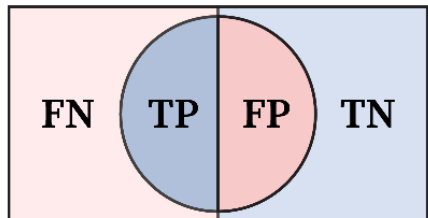
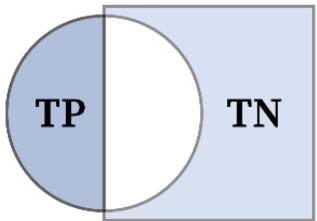
Combined Score

FN = False Negative
TP = True Positive

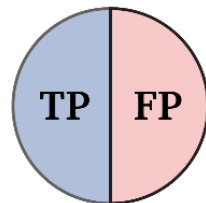


FP = False Positive
TN = True Negative

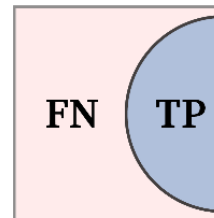
(i) Accuracy



(ii) Precision



(iii) Recall

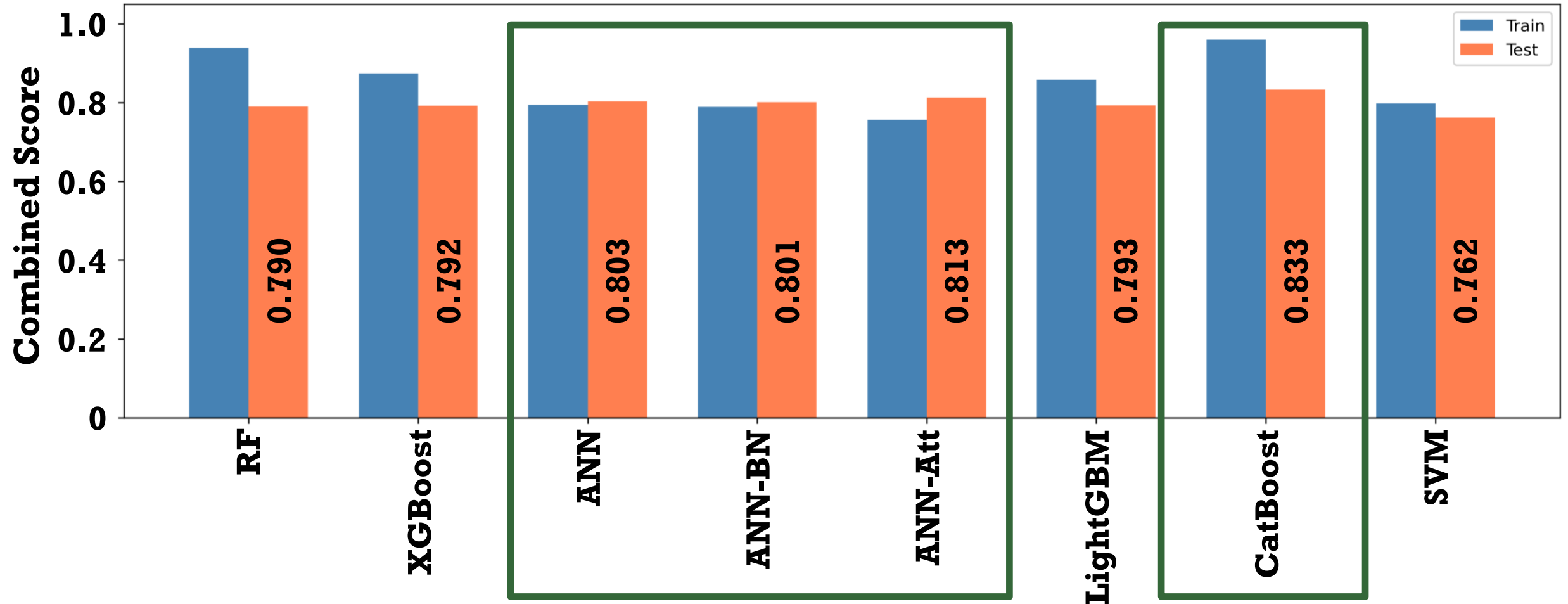


Combined Score

- **Accuracy – 35%**
- **Recall (Mandatory) – 25%**
- **Recall (Minimal) – 15%**
- **Precision (Mandatory) – 20%**
- **Precision (Minimal) – 5%**

ML Model Performance – MVI Binary Classification

Train vs. Test Combined Score for All 8 Models



ML Model Performance – MVI Binary Classification (Cont.)

Model	Train Combined	Test Combined	Recall Mandatory	Recall Minimal	Test Accuracy	Class Separation Probability Threshold
Random Forest	0.939	0.790	0.67	0.93	0.89	0.30
XGBoost	0.874	0.792	0.76	0.90	0.88	0.25
ANN	0.794	0.803	0.96	0.82	0.84	0.55
ANN-BatchNorm	0.789	0.801	0.91	0.84	0.85	0.50
ANN-Attention	0.756	0.813	0.91	0.86	0.86	0.85
LightGBM	0.858	0.793	0.82	0.87	0.86	0.60
CatBoost	0.959	0.833	0.73	0.95	0.91	0.50
SVM	0.798	0.762	0.82	0.83	0.83	0.25

★ Recall(Mandatory) is the most critical metric — missing a HAB event has severe safety consequences

Ensemble Predictions – Top 4 ML Models

Strategy 1: Simple Average

Equal weight mean probability
from top 4-models

Strategy 2: Weighted Average

Models weighted by their
Combined Scores on Test data

Strategy 3: Majority Vote

Mandatory Management; if ≥ 3 of
top 4-models predict Mandatory

Top 4 model selection has been done based on Test Combined Score ranking.

Uncertainty Classification

LOW

Std. < 0.05

Low uncertainty — reliable
prediction

MEDIUM

$0.05 \leq \text{Std.} < 0.15$

Moderate uncertainty — review
recommended

HIGH

Std. ≥ 0.15

High uncertainty — field
verification needed

Uncertainty estimated as Standard Deviation (std) of prediction probabilities for all test samples by each of the top 4 models.



Ensemble Predictions – Confusion Matrices

- ✓ Simple Average and Weighted Average ensemble strategies demonstrated **more accurate Minimal Management** category.
- ✓ Majority Vote ensemble strategy showed **more accurate Mandatory Management** category.

Confusion Matrices - Ensemble

Simple Average

Recall (Minimal) 89.9% | Recall (Mandatory) 88.9%

True Class	Minimal Management	Mandatory Management
	Minimal Management	Mandatory Management
Minimal Management	231 (76.5%)	26 (8.6%)
Mandatory Management	5 (1.7%)	40 (13.2%)

Weighted Average

Recall (Minimal) 89.9% | Recall (Mandatory) 88.9%

True Class	Minimal Management	Mandatory Management
	Minimal Management	Mandatory Management
Minimal Management	231 (76.5%)	26 (8.6%)
Mandatory Management	5 (1.7%)	40 (13.2%)

Majority Vote

Recall (Minimal) 83.7% | Recall (Mandatory) 93.3%

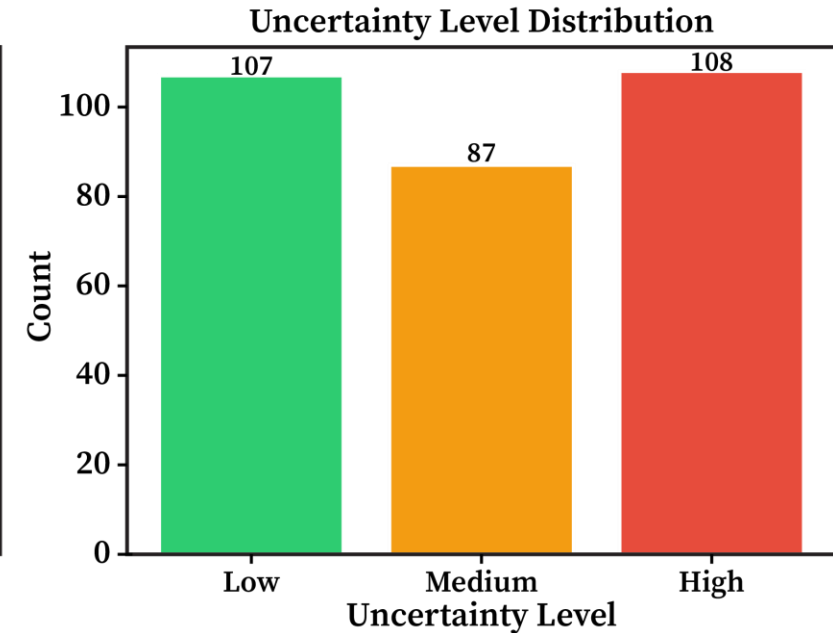
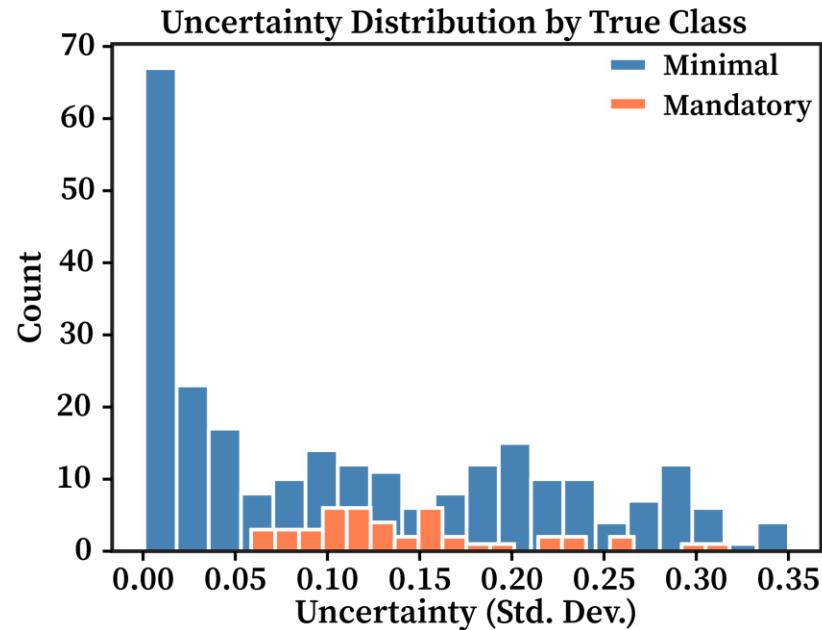
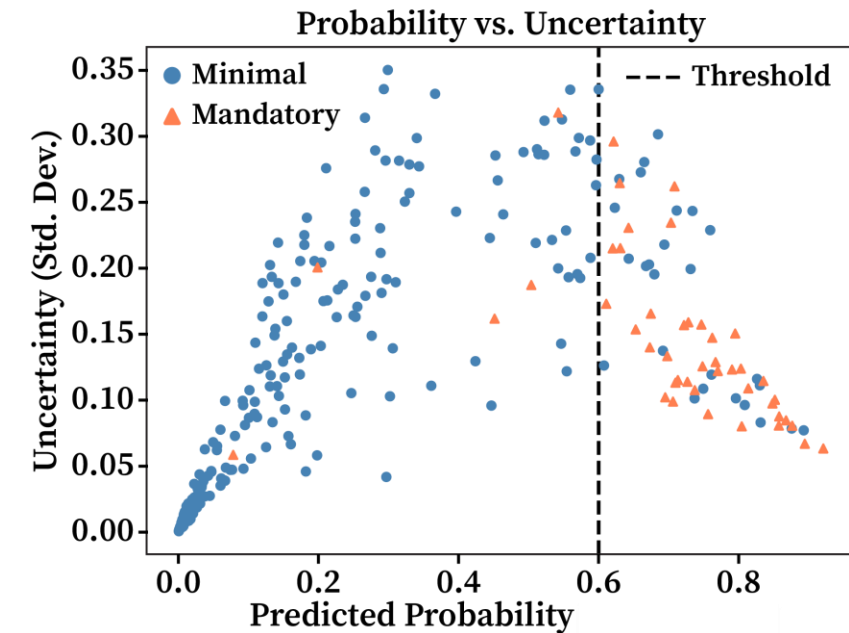
True Class	Minimal Management	Mandatory Management
	Minimal Management	Mandatory Management
Minimal Management	215 (71.2%)	42 (13.9%)
Mandatory Management	3 (1.0%)	42 (13.9%)

Predicted Class

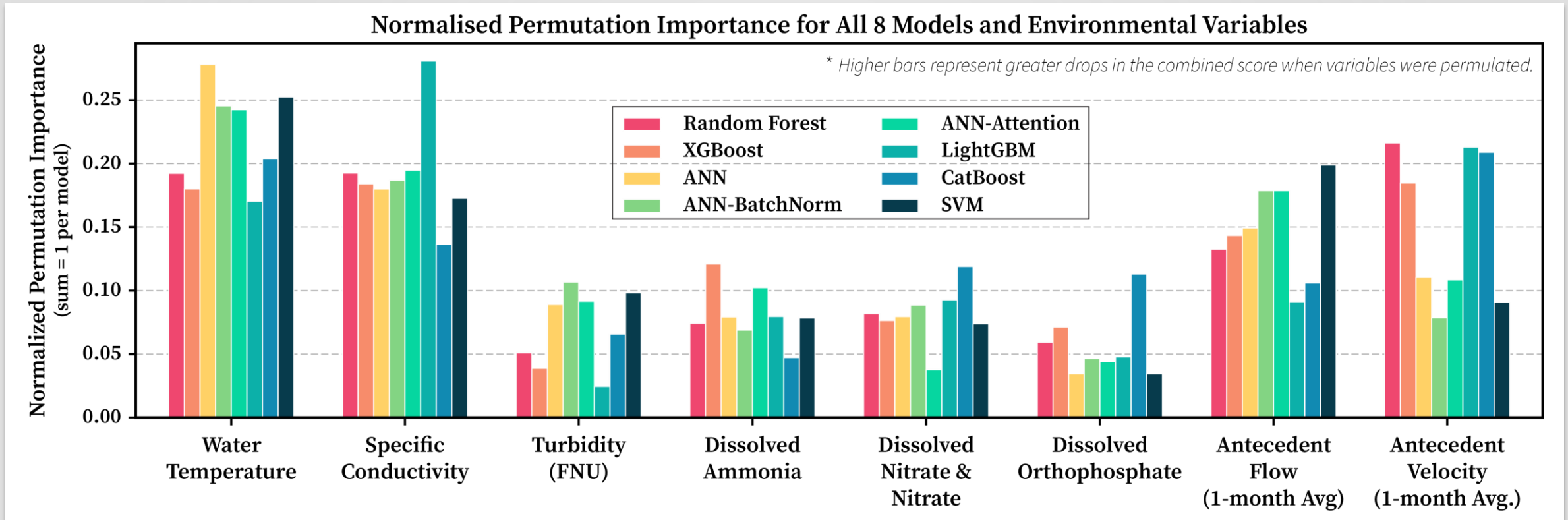
Predictions Uncertainties

- ✓ More uncertain predictions have predicted probability range between 0.35 - 0.65. Class probability threshold for ensemble predictions were optimized to 0.60.
- ✓ About one third (108) of the test sample predictions had High Prediction Uncertainty.

Top 4-Models' Ensemble Uncertainty Analysis

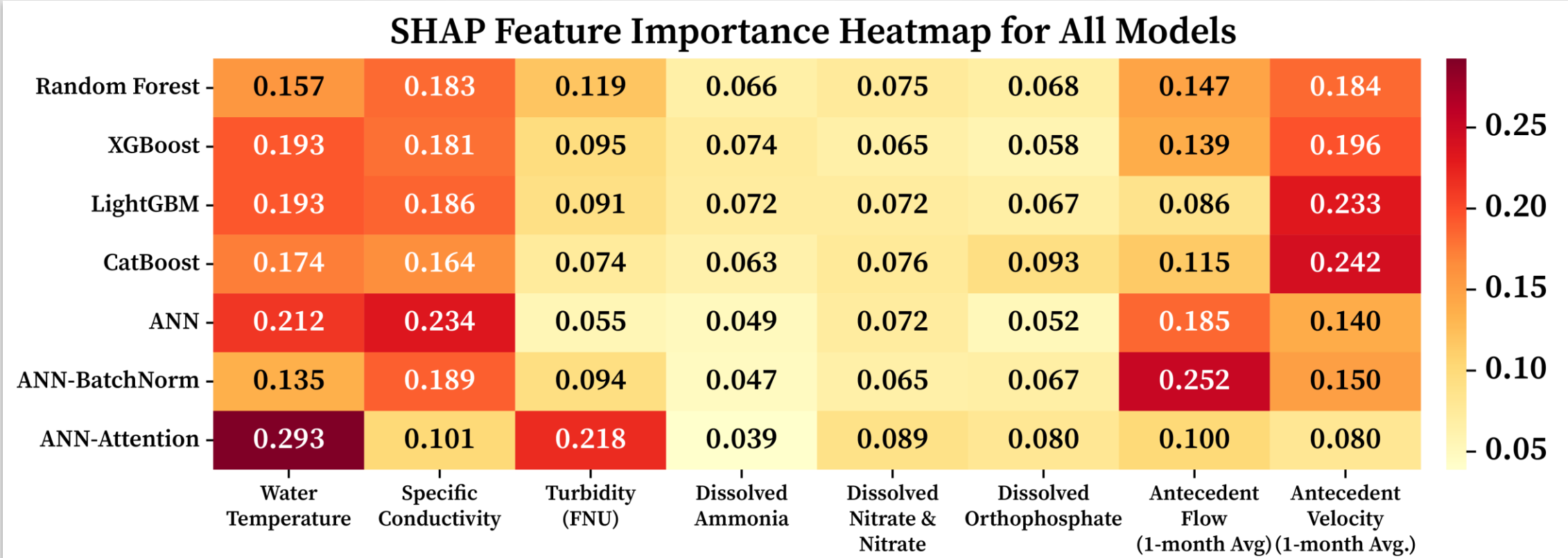


Feature Importance — Permutation Methods



- **Water Temperature** & **Antecedent Velocity (1-month Avg.)** rank consistently **top 2** environmental **variables** across all models.
- **Antecedent Flow (1-month Avg.)** and **Specific Conductivity** are also **important**, especially for a few models, including LightGBM and CatBoost.
- **Turbidity (FNU)** and **nutrients** show **lower importance**, but relevant for ANN-Attention.

SHAP Feature Importance — Cross-Model Comparison



Water Temperature: highest SHAP in ANN-Attention (0.29)
— neural networks capture non-linear thermal effects

Antecedent Velocity (1-month Avg.): dominant in LightGBM (0.23) and CatBoost (0.24) — flow suppresses blooms

Specific Conductivity: consistently ~0.16 – 0.23 across tree-based and ANN models

Turbidity (FNU) elevated for ANN-Attention (0.22) — light limitation captured by attention mechanism

Key Findings — Phase 2 Work 2026

1. Model Performance

- ✓ Test Combined Scores: 0.76 – 0.83
(8 ML models; CatBoost leading)
- ✓ Recall (Mandatory Management): 0.67 – 0.96
(operationally most critical)

2. Top Drivers of Mandatory HABs

- ✓ Water Temperature and
 - ✓ Antecedent Velocity (1-month Avg.)
- } most important

3. Permutation Importance Analysis

- ✓ Water temperature, Antecedent velocity & flow (1-month Avg.), and Specific conductivity } more important
- ✓ Turbidity and Nutrients ← less important

4. SHAP Reveals Model-Specific Patterns

- ✓ ANN-Attention captures non-linear thermal interactions (Water Temperature SHAP = 0.29).
- ✓ Tree-based models emphasize antecedent flow (1-month Avg.) suppression.

5. Class Imbalance Challenge

- ✓ Mandatory Management : ~15% of samples.
- ✓ Combined score and class probabilities threshold optimization improved Recall(Mandatory) significantly.

6. Ensemble & Uncertainty

- ✓ Top 4-models ensemble show consistent predictions.
- ✓ Prediction Uncertainty on Test Samples:
 - ❖ LOW: 107
 - ❖ MEDIUM: 87
 - ❖ HIGH: 108

Summary

What We Did

- ✓ 8 ML models trained for reclassified binary MVI
- ✓ Optimized combined score and class separation threshold
- ✓ Ensembled top 4-models and estimated uncertainty
- ✓ Completed feature importance and SHAP analysis

Future Work

- Cluster and associated MVI cross-tabulation
- PCA and UMAP with error overlay visualisation
- Deep diving to Misclassification cluster analysis
- MVI dashboard development
- MVI modeling Annual Report 2026-2027 preparation
- MVI modeling manuscript preparation

References

1. Flynn, T., Lehman, P., Lesmeister, S., and Waller, S. 2022. A Visual Scale for Microcystis Bloom Severity. <https://doi.org/10.6084/m9.figshare.19239882.v1>
2. Bouma-Gregson K, Bosworth D H, Flynn T M, Maguire A, Rinde J, and Hartman R. 2024. Delta Blue(green)s: The Effect of Drought and Drought-Management Actions on Microcystis in the Sacramento–San Joaquin Delta. San Francisco Estuary and Watershed Science. <https://doi.org/10.15447/sfews.2024v22iss1art2>

Acknowledgements

DWR

☐ Modeling Support Office

☐ Ellen Preece

☐ Rosemary Hartman

☐ Shaun Philippart

☐ Silvia Angles

☐ Emily Richardson

MWQI

☐ Leslie Palencia

Thank You!

Contact Us

☐ Gourab.Saha@water.ca.gov

Questions?