

A Machine Learning Approach to Estimate Plume Discharge from Electrical Geophysical Measurements

Neil Terry

Frederick Day-Lewis

John W. Lane

Jared Trost (USGS Mounds View)

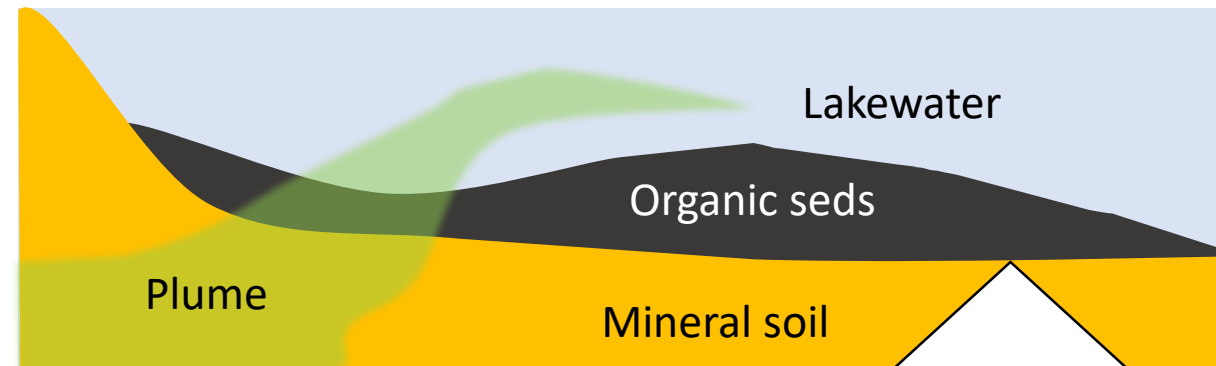
Isabelle Cozzarelli (USGS Reston)

Barbara Bekins (USGS Menlo Park)

Motivation for Machine Learning (ML)

- Geophysics can give indirect information about contaminants
- Inversion and image thresholding used to map ranges of values to features of interest

True conditions



Geophysical inverse results



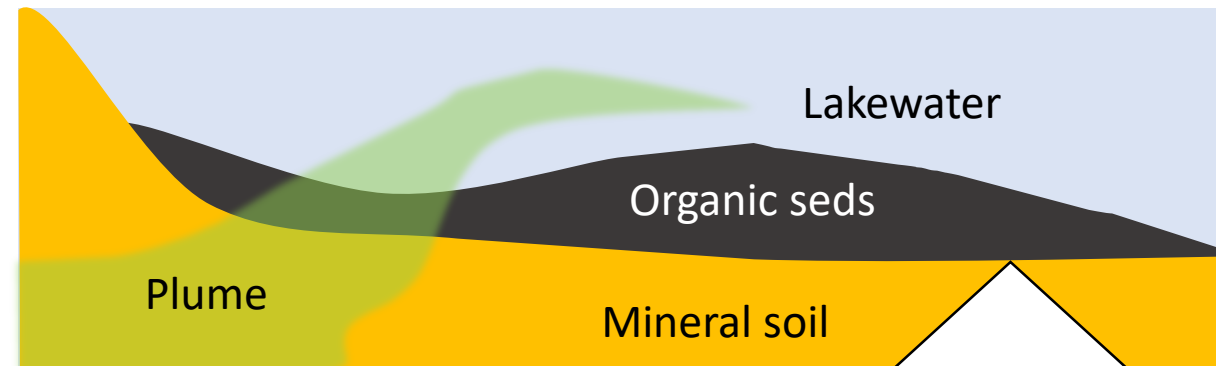
Image thresholding



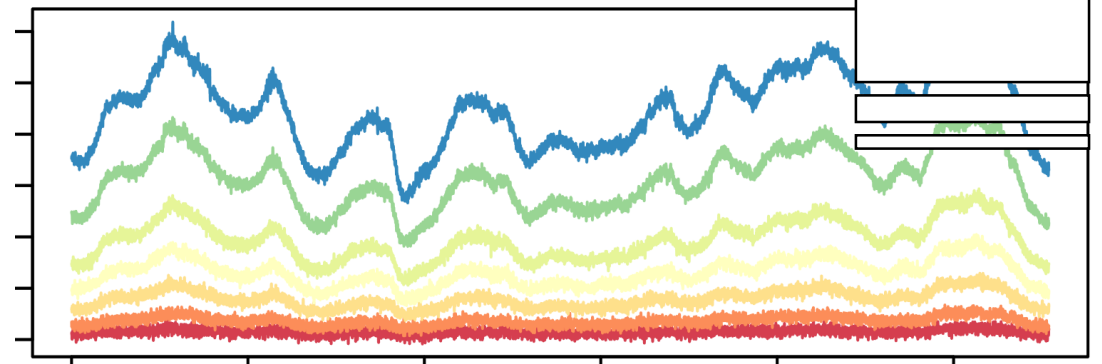
Motivation for Machine Learning (ML)

- Alternatively, ML can be used to develop direct relationship between contamination and available data

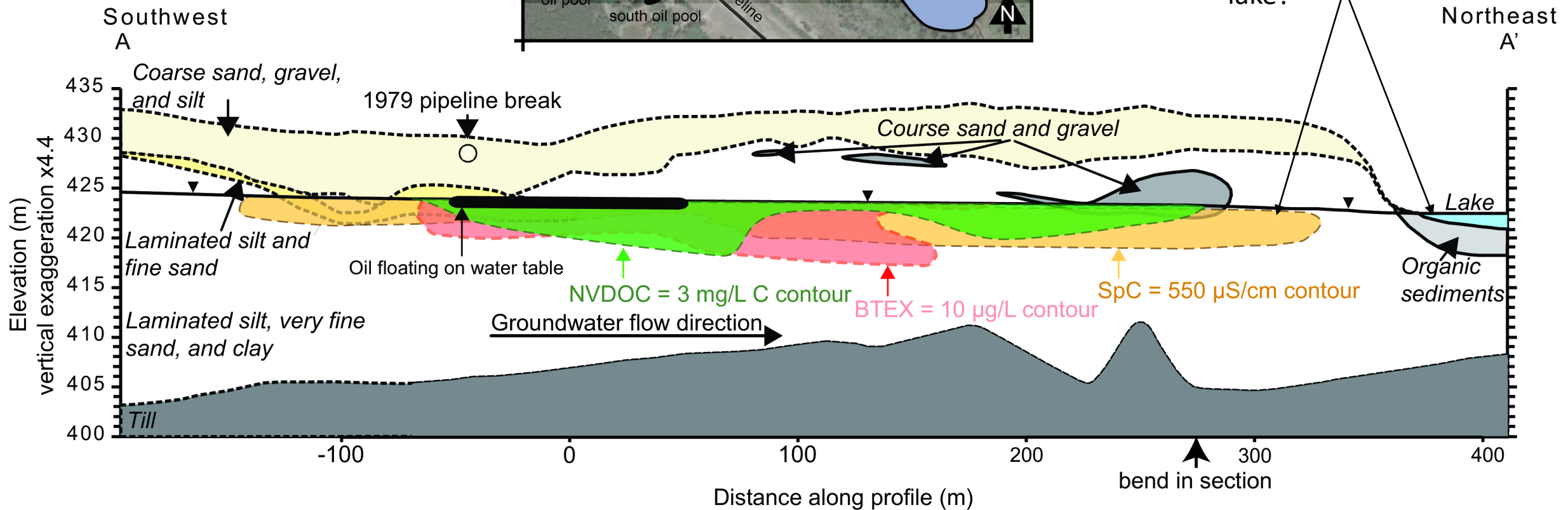
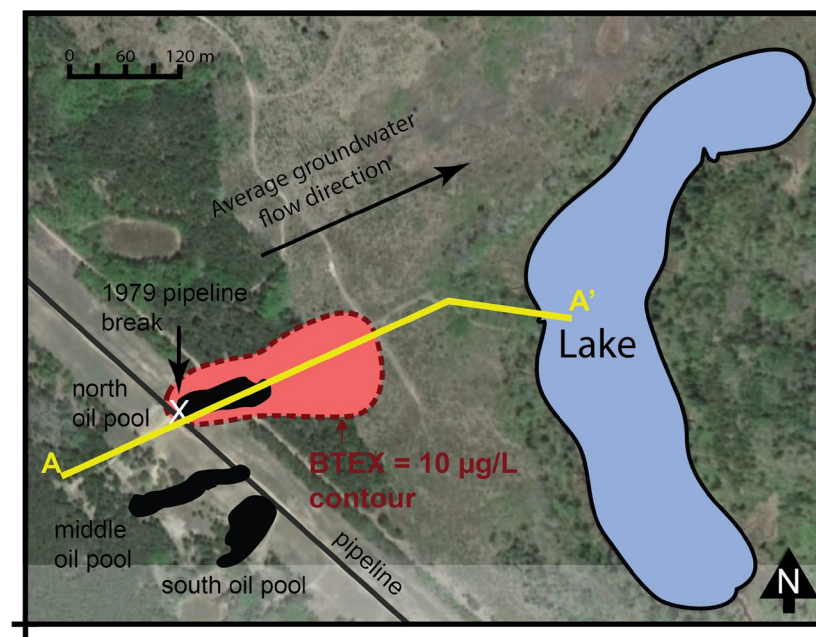
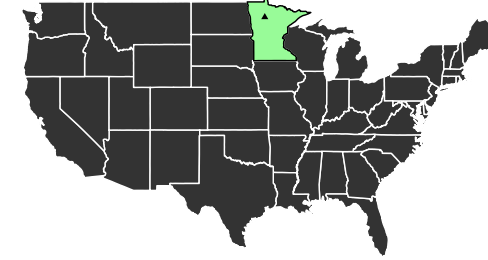
True conditions



Raw datasets

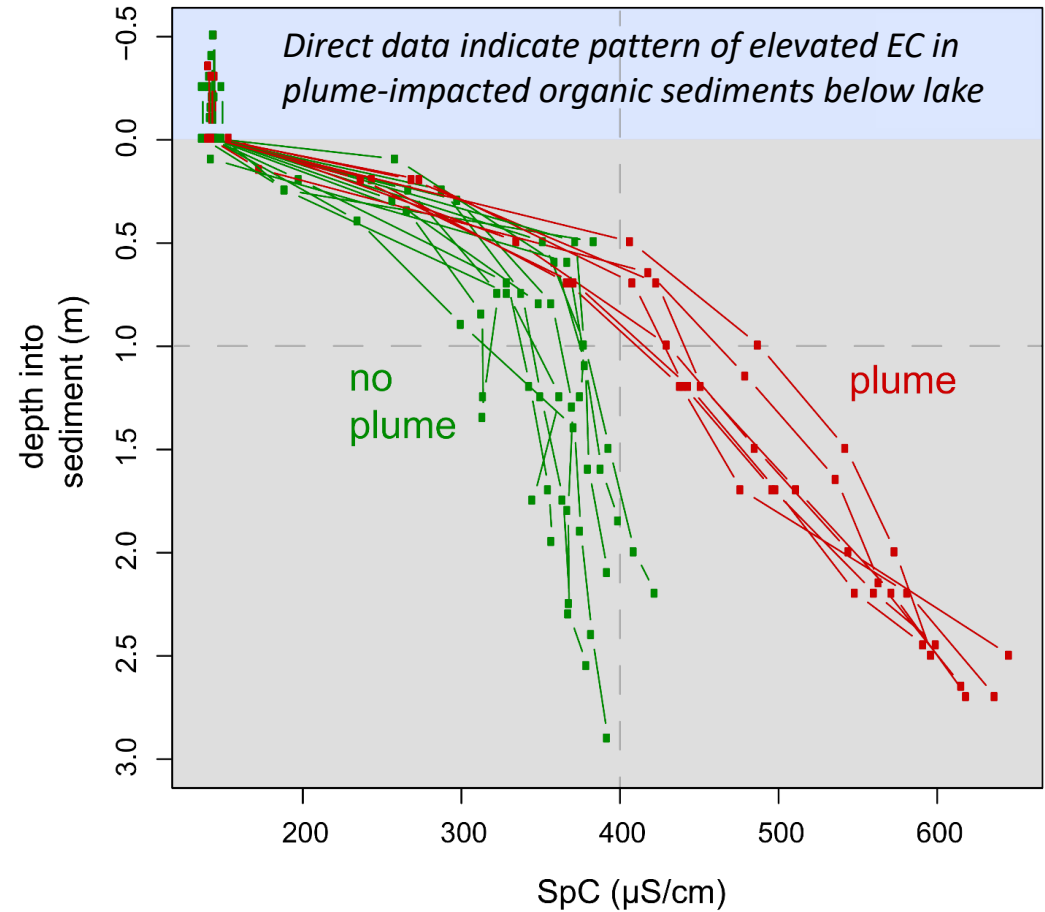
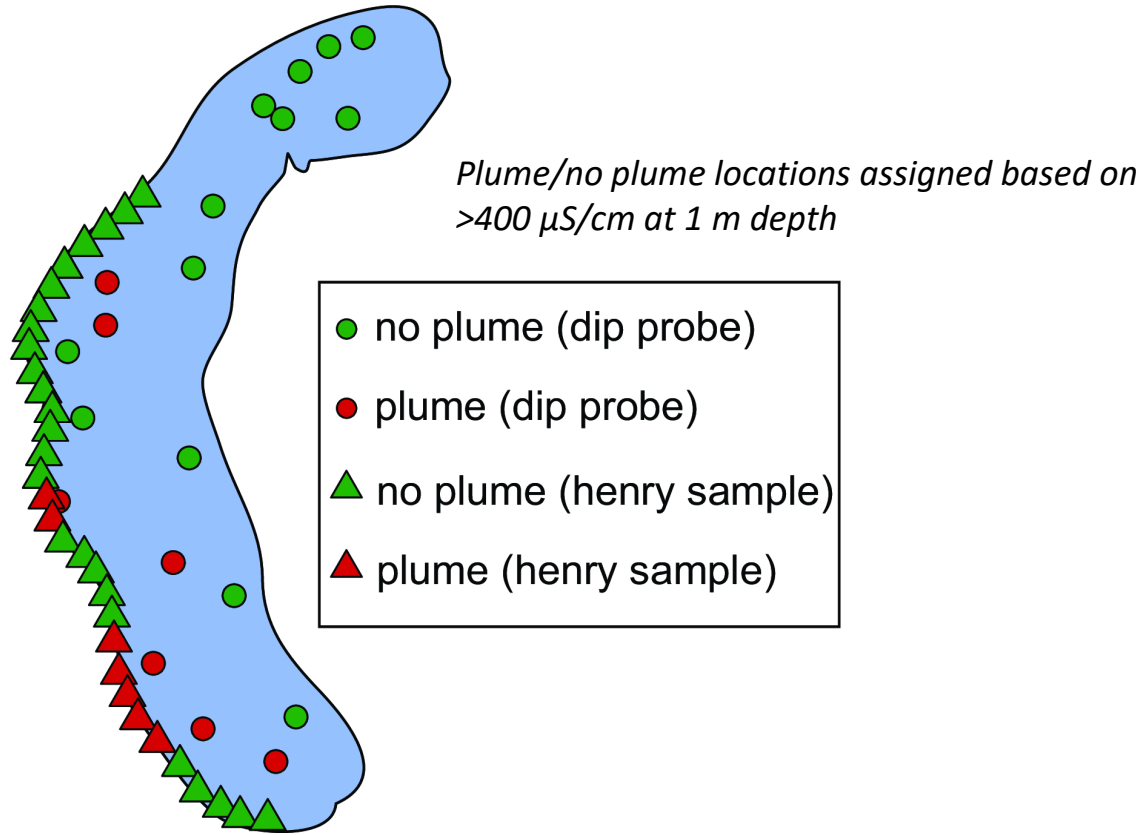


Bemidji, MN



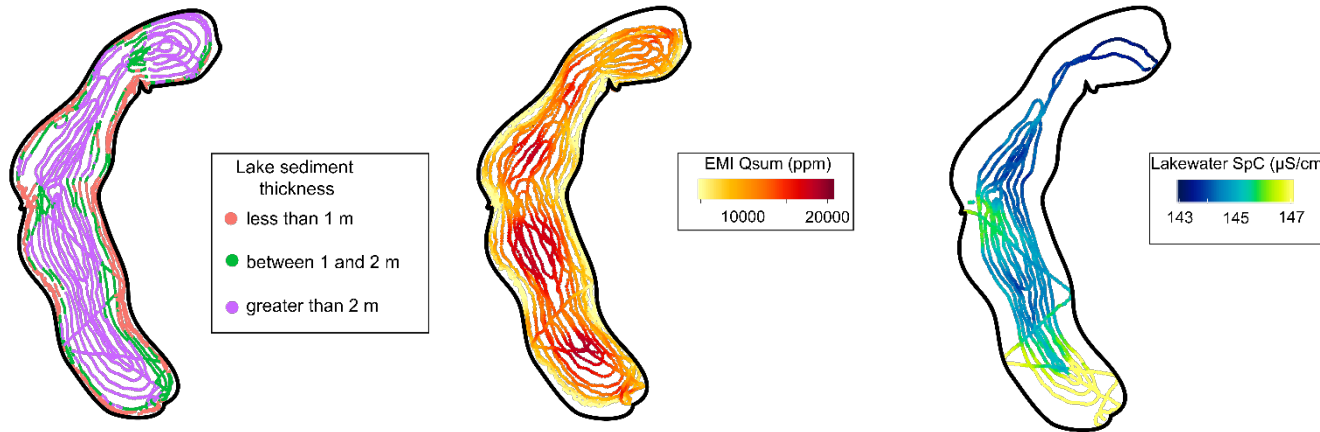
Evidence of SpC plume discharging to lake?

SpC measurements

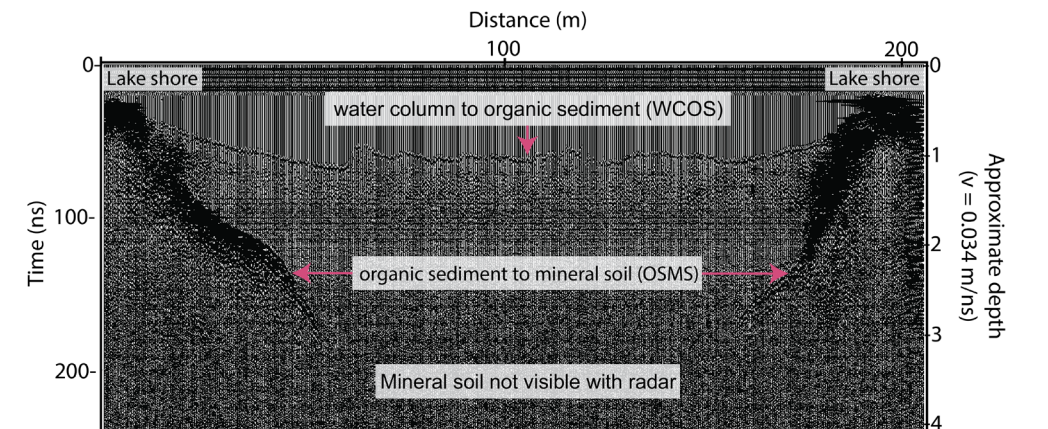
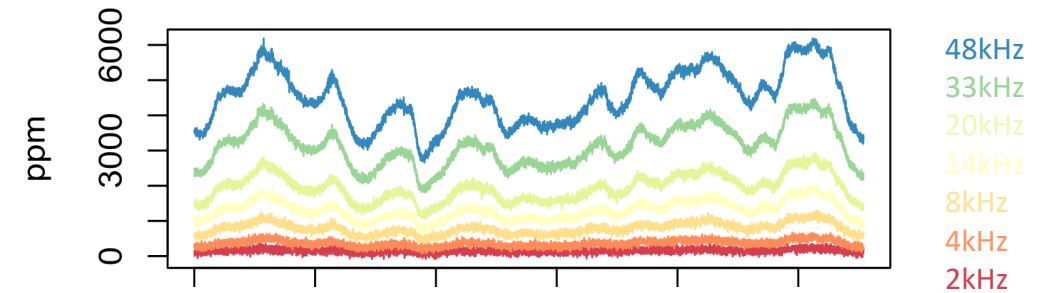


Geophysical datasets

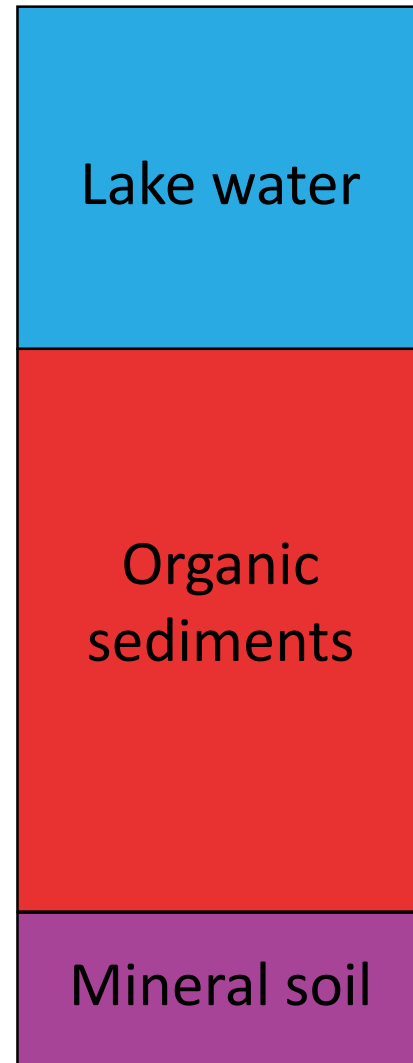
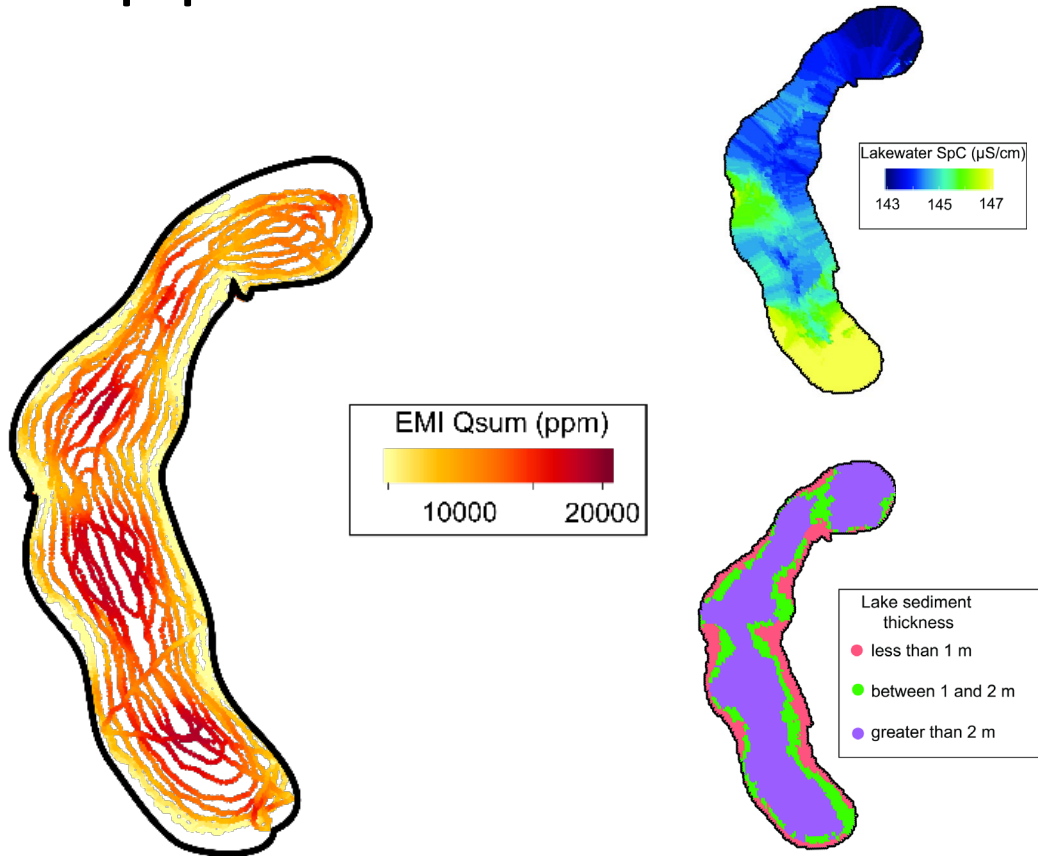
- Spatially distributed multifrequency EM induction, GPR, and lakewater SpC collected via kayak over the lake



Geophysical datasets collected over the unnamed lake



Geophysical inversion approach

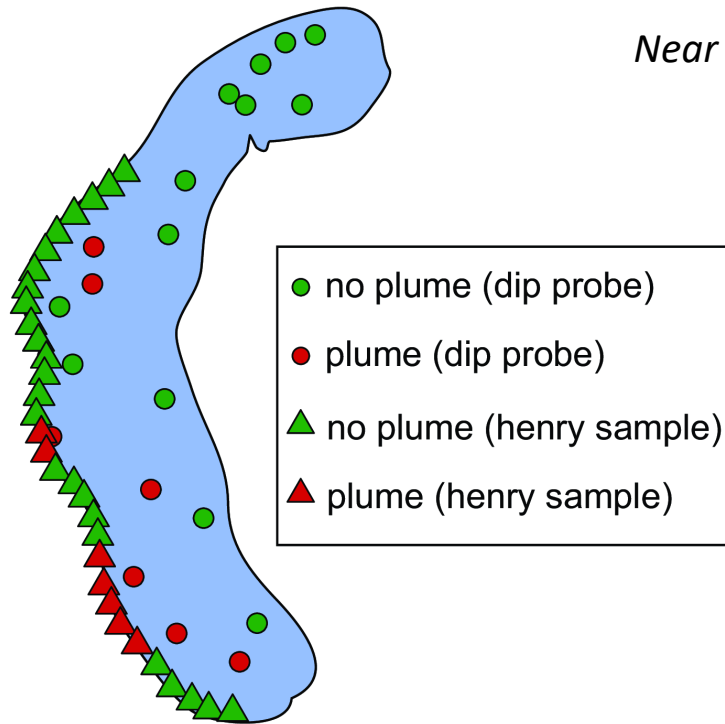


Depth and SpC of lakewater accurately known: hard constraint on the inversion

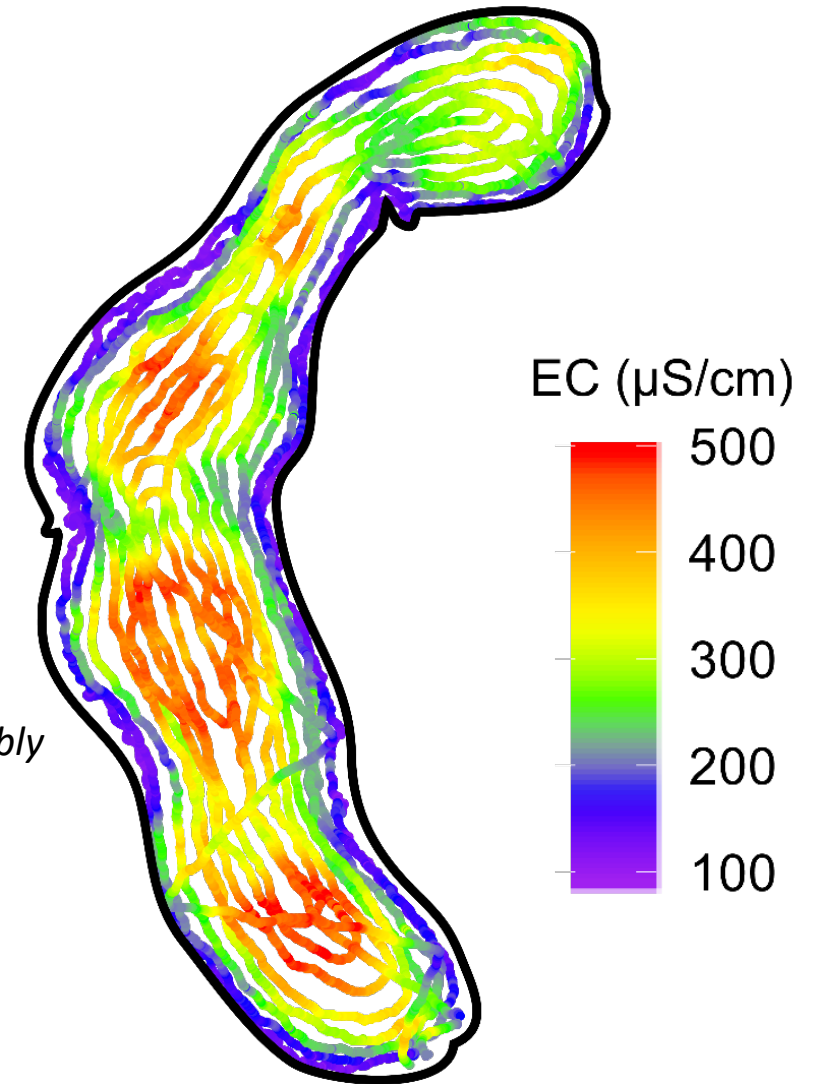
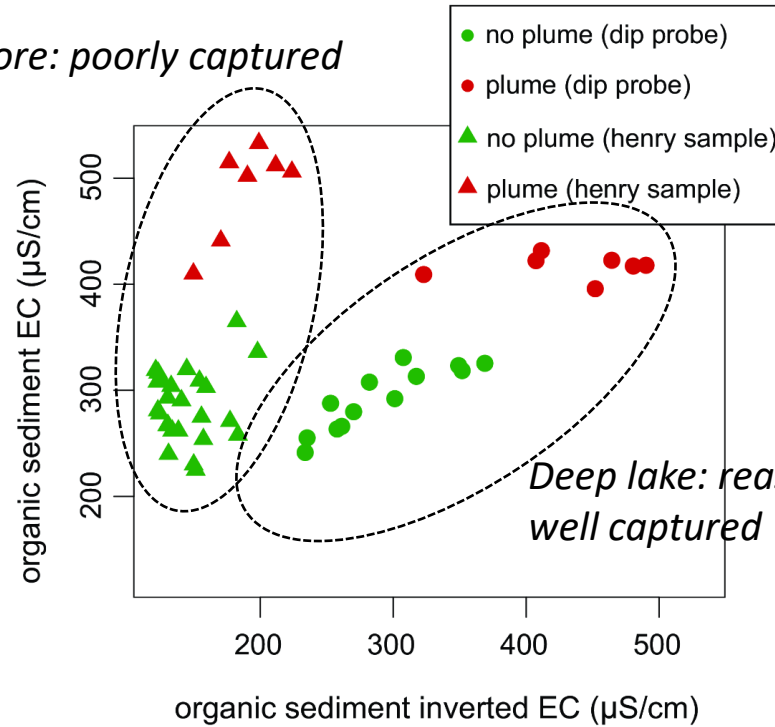
Approximate range of EC and thickness of organic sediments available: soft constraint on the inversion

Mineral soil is infinite layer with unconstrained EC

The problem with inversion



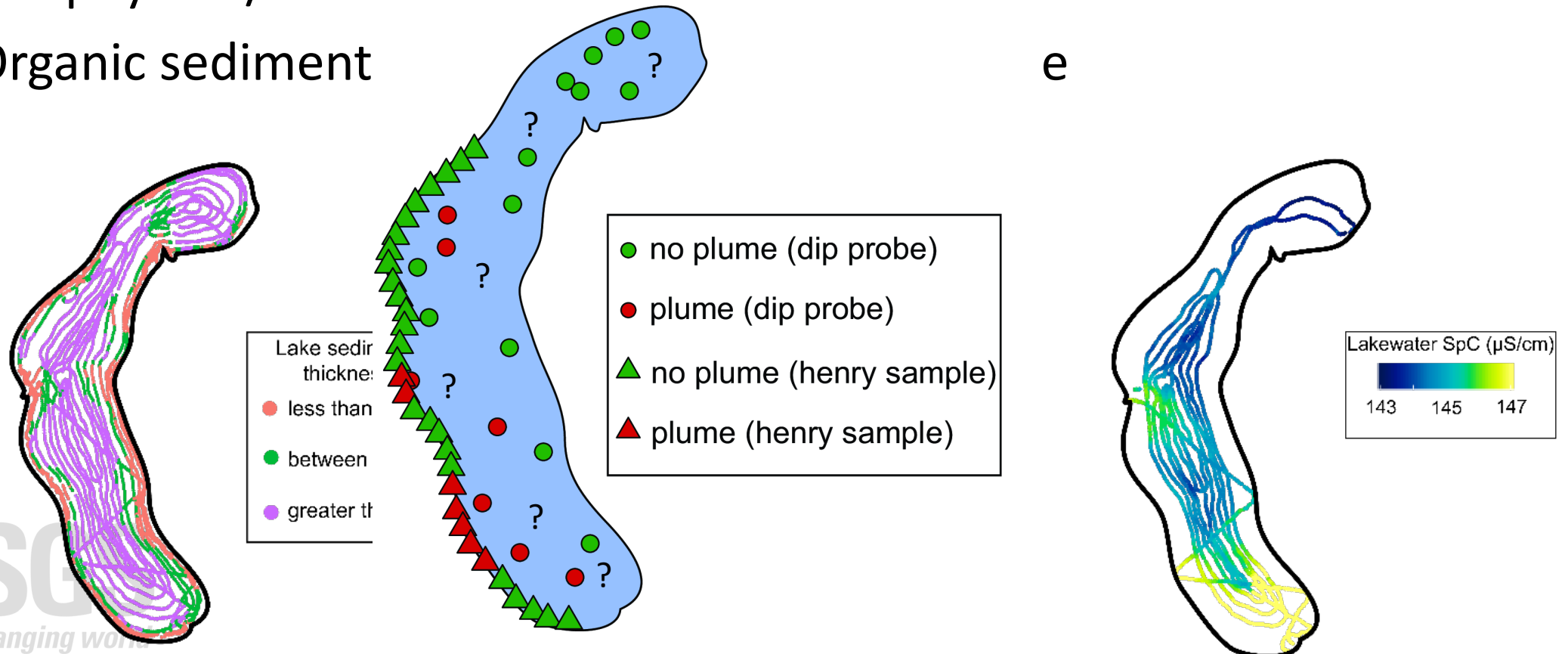
Near shore: poorly captured



Organic sediment inverted EC

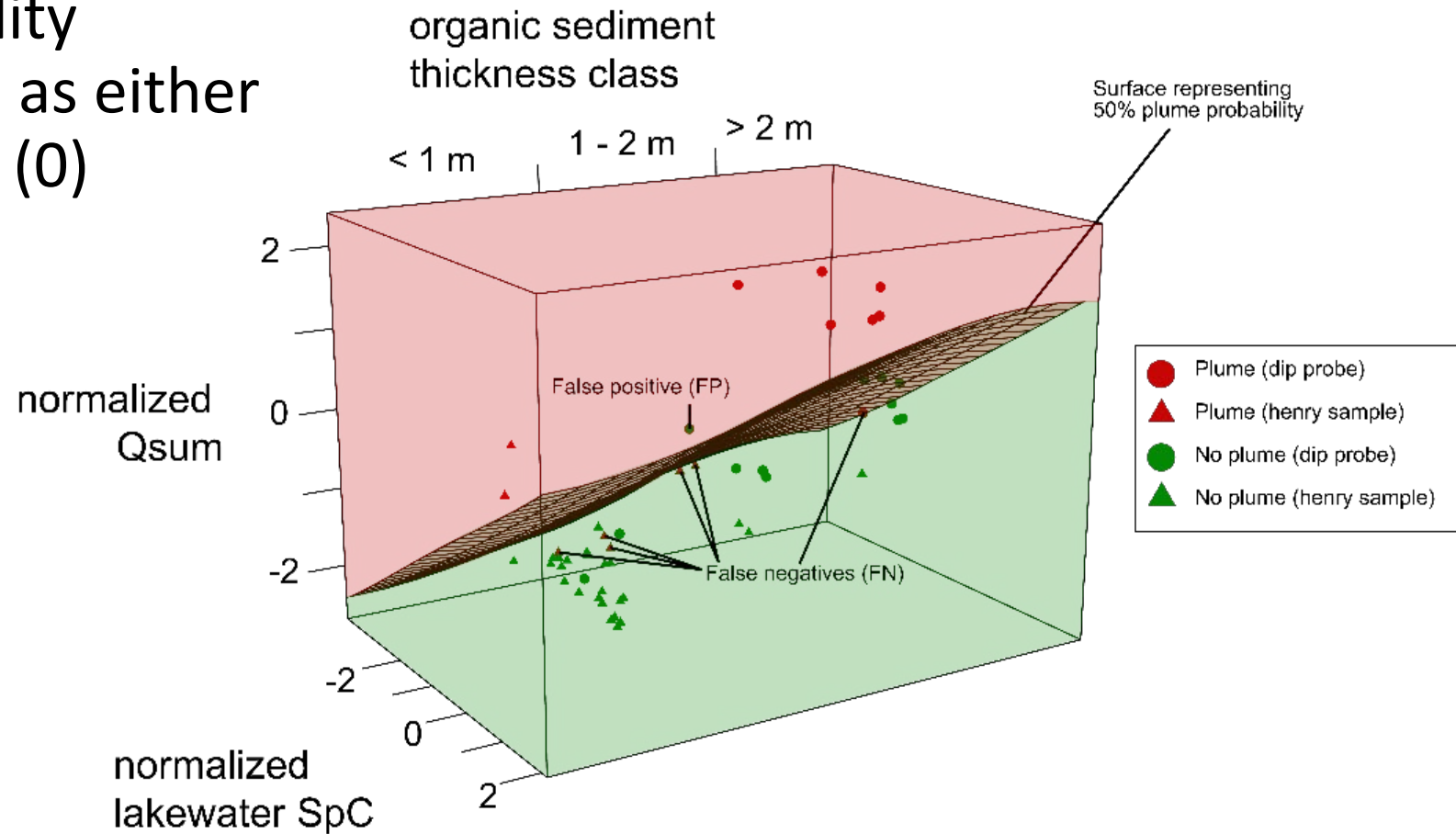
Machine learning alternative

- Goal is to identify where SpC is daylighting: binary variable
- Geophysical/lake level
- Organic sediment

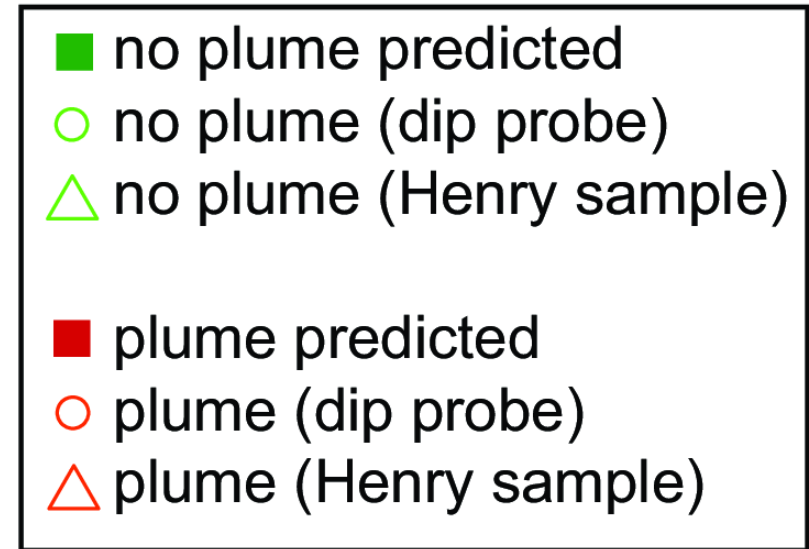
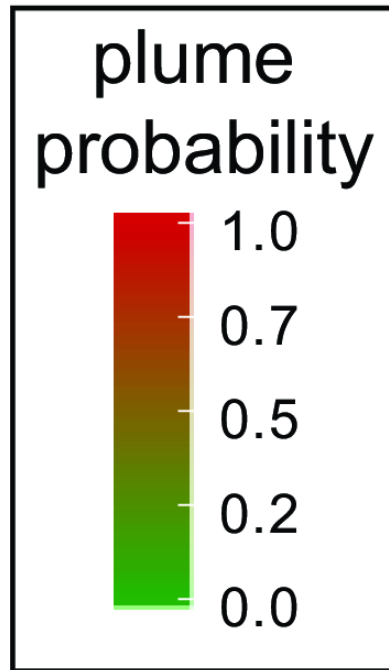
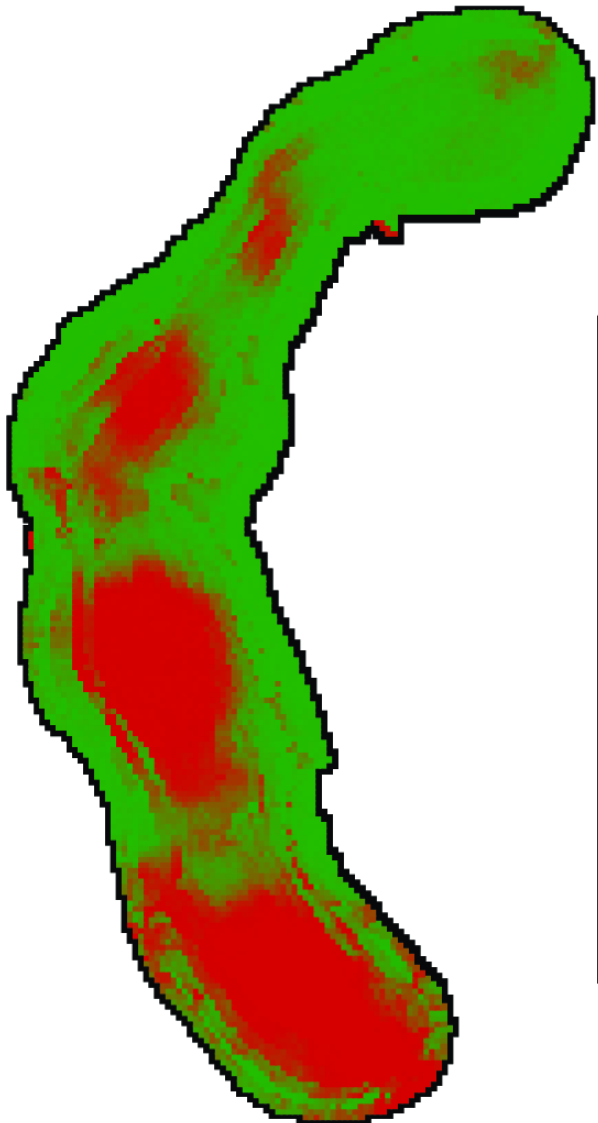


Logistic regression classifier

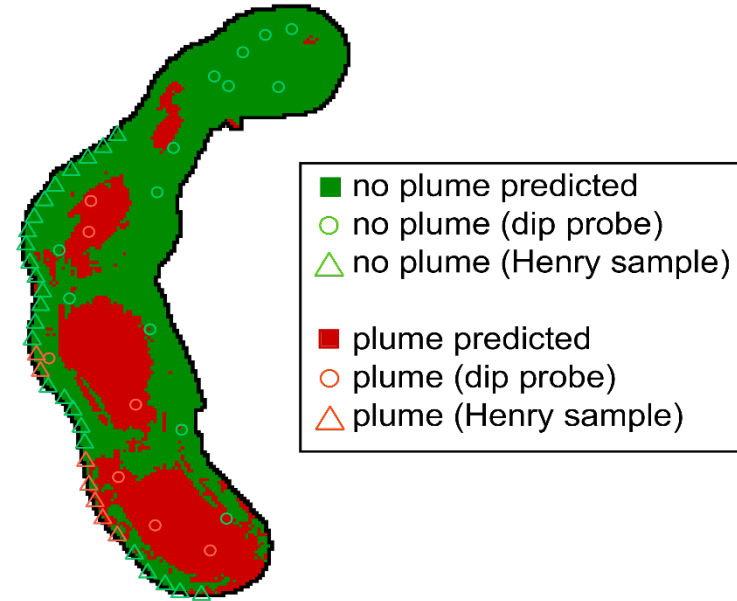
- Develops 50% probability surface to classify data as either plume (1) or no plume (0)



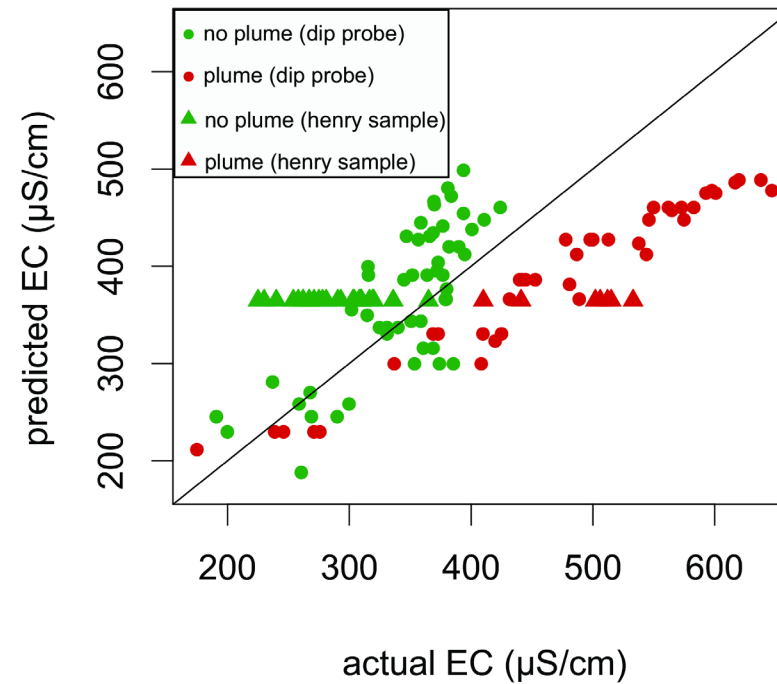
Classification from logistic regression



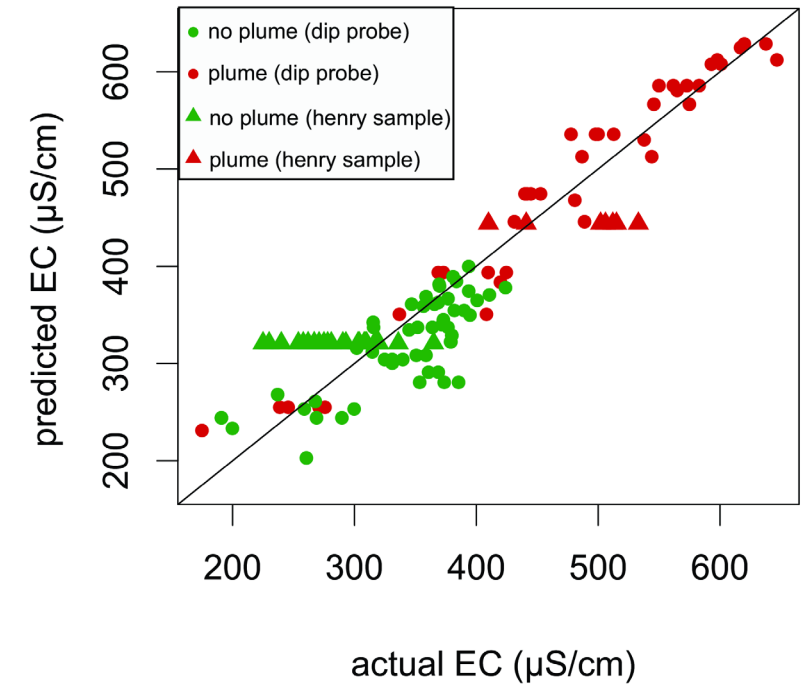
Organic sediment EC prediction



(a) Multivariate regression only



(b) Logistic + multivariate regression



Summary

- Logistic regression classifier provides direct estimate of plume discharge
- Secondary regression step better predicts EC with depth
- Site-specific relationships but potentially transferrable methodology

