UPDATE ON C, N, AND P STUDY

Utah Lake Water Quality Study
Science Panel Call
2020-12-14

LITERATURE REVIEW ON UTAH LAKE CNP

- hydrologic and atmospheric inputs to the lake, (tributaries, groundwater, drains, direct precipitation, and atmospheric deposition)
- hydrologic and atmospheric outputs from the lake (tributaries, groundwater, and evaporation)
- water column transformations of elements
- sediment transformations of elements
- elemental fluxes between the water column and sediment
- elemental standing stocks in various pools in the water column and sediment

LITERATURE COMPILATION

Peer reviewed studies, theses, reports, letters from

- Files obtained through ULWQS project
- Web of Science
- Google Scholar
- Direct contact with scientists

LITERATURE COMPILATION

- 79 total documents
- 38 have data on C, N, and/or P in Utah Lake
- 30 have data on Utah Lake but not C, N, or P (e.g., fish biomass)
- 5 are relevant studies in other systems
- 9 pertain to atmospheric deposition (6 without data on rates)

DATA COMPILATION

All studies w/ CNP data reviewed → metadata and data recorded

FileName	Author(s)	Year	Title	ArticleType	Process or Pool	Fraction
Abu Hmeidan 2017 thesis	Abu Hmeidan	2017	Characteri	Thesis	Sediment P	TP
Abu Hmeidan 2017 thesis	Abu Hmeidan	2017	Characteri	Thesis	Sediment P	TP
Abu Hmeidan 2017 thesis	Abu Hmeidan	2017	Characteri	Thesis	Sediment P	TP
Abu Hmeidan 2017 thesis	Abu Hmeidan	2017	Characteri	Thesis	Sediment P	TP

Rate or Amou	Units	Aggregation	Year	Month	Location	Measured or Estimated?	Approach	Notes
280	ppm	Minimum	2015-2016		Lakewide	Measured	Microwave Diges	Same resu
465	ppm	Minimum	2015-2016		Provo Bay	Measured	Microwave Diges	Same resu
604	ppm	Mean	2016		Lakewide	Measured	Microwave Diges	Same resu
631	ppm	Mean	2015-2016		Middle Upper Quad	Measured	Microwave Diges	Same resu

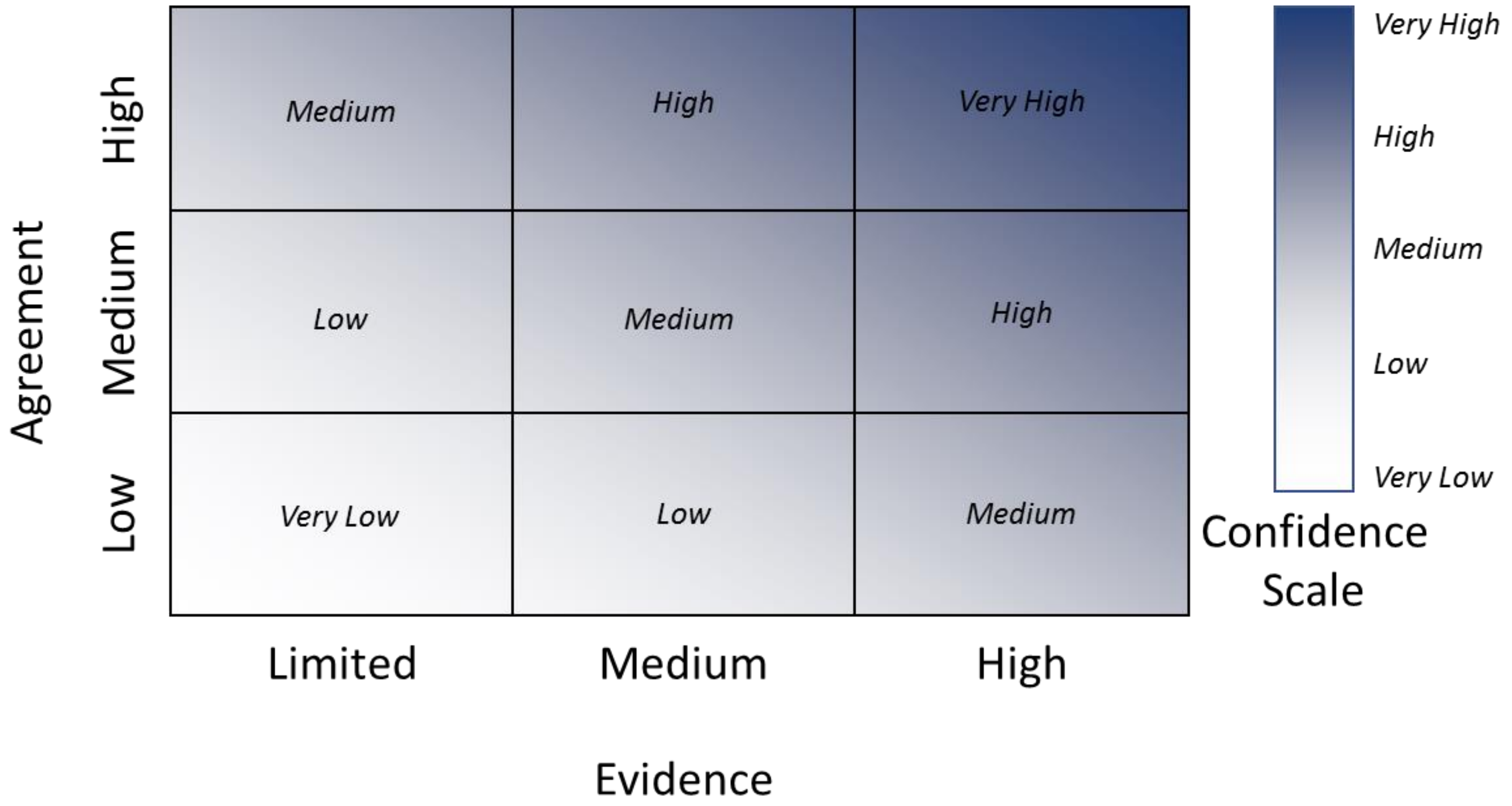
UNCERTAINTY ASSESSMENT

For data pertaining to a given elemental form + process/stock, uncertainty was qualified according to the Uncertainty Guidance document

- Evidence
  - Quality: standard/direct methods > nonstandard/estimated methods
  - Quantity: # of studies
  - Relevance: all studies are in Utah Lake → High
- Agreement
  - Sometimes low agreement is due to spatial/temporal variability

High	< 10% difference in values
Medium-High	< 25% difference in values
Medium	< 50% difference in values
Medium-Low	>50% difference, <500% difference in values
Low	>500% difference in values

UNCERTAINTY ASSESSMENT: CONFIDENCE



QUESTIONS FOR SP

1. Are there studies on Utah Lake C, N, and/or P that are not included?
2. Feedback on application of uncertainty assessment?

MODELING EFFORTS

Quantifying conceptual models

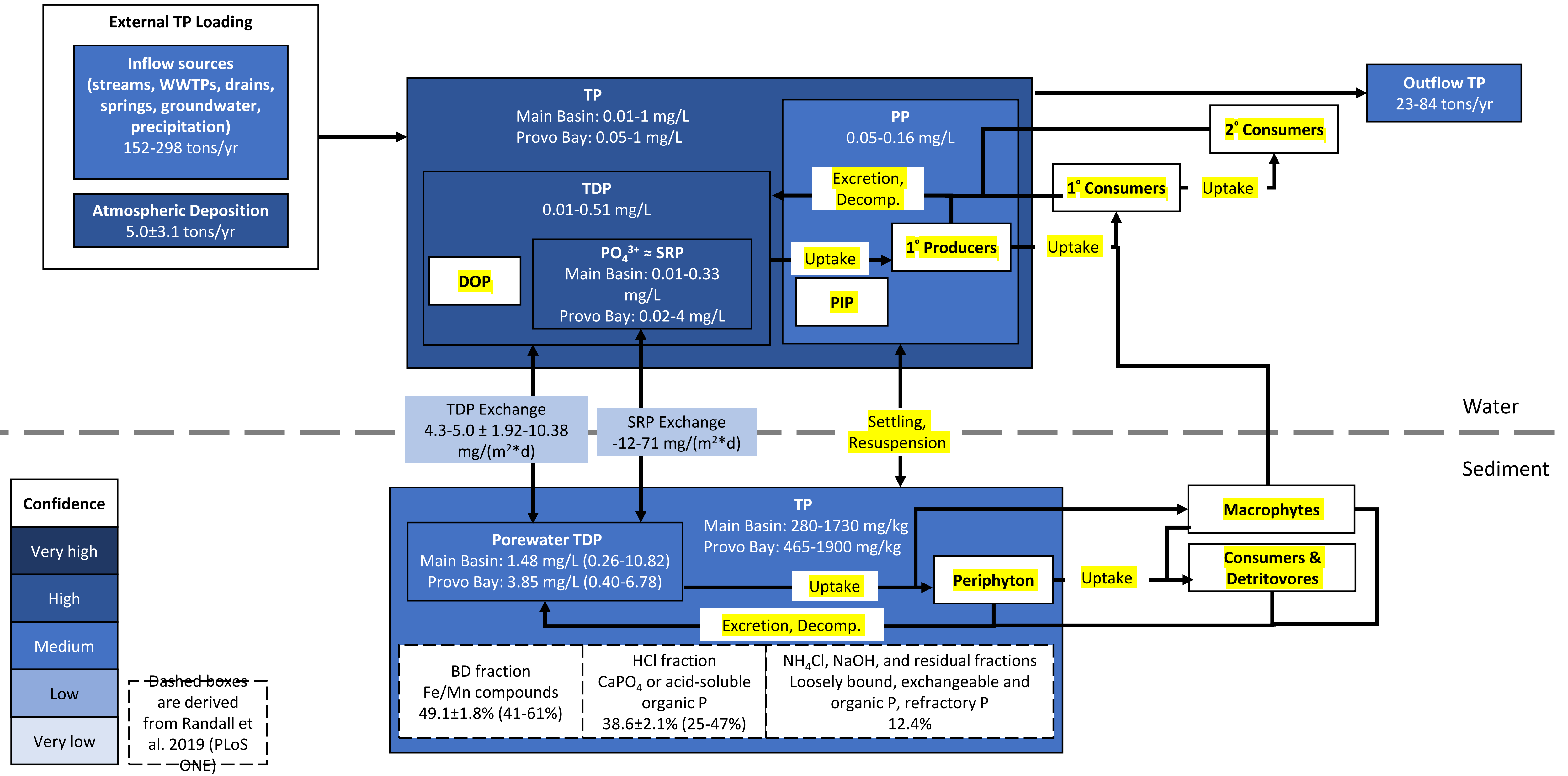
SedFlux model

External mass balance model

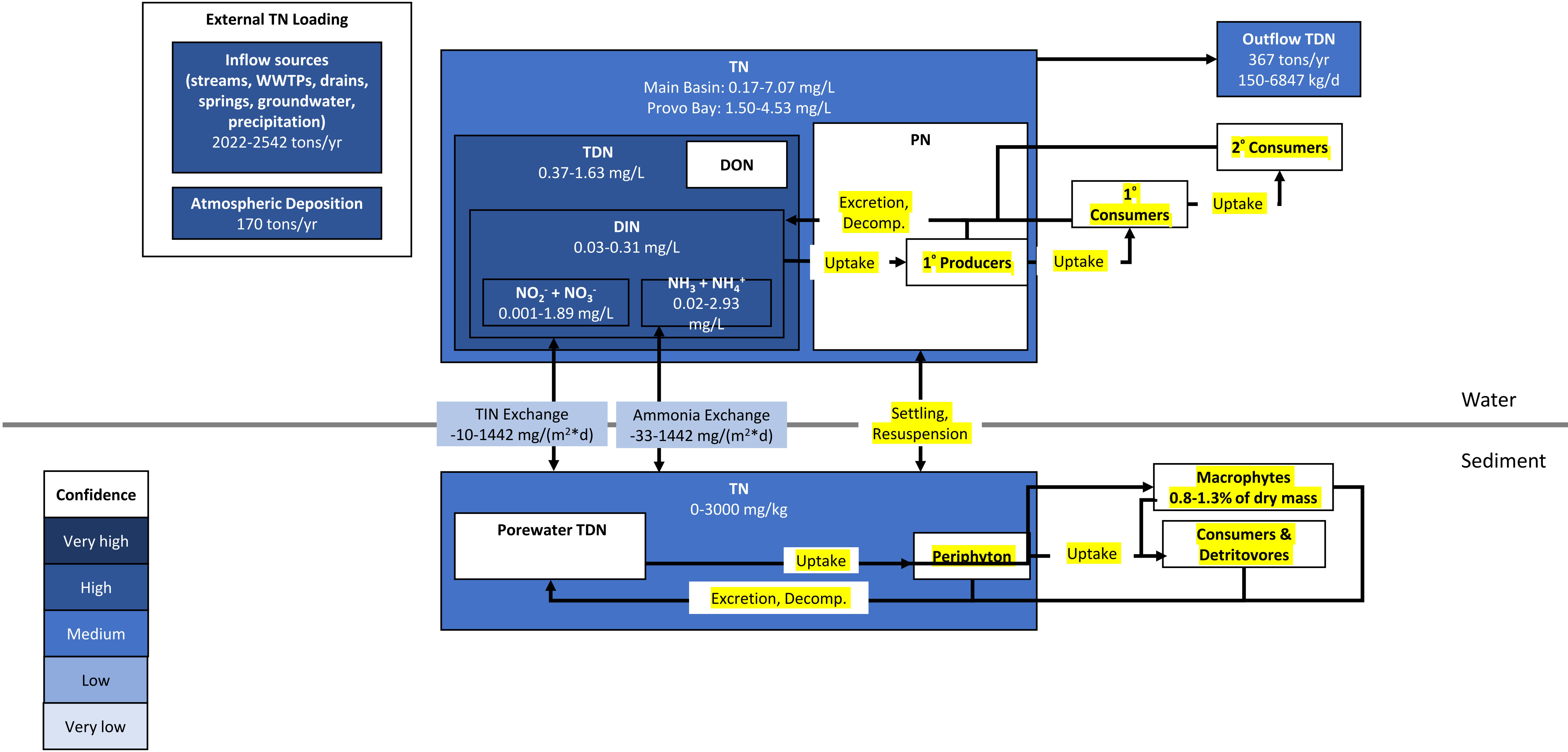
CONCEPTUAL MODELING

1. Apply measured values to stocks/processes
2. When DWQ monitoring data available, compare w/ studies from lit review
3. Assign confidence
4. For unmeasured values, consult literature
   a. Known total stock, need to apply elemental ratio (e.g., phytoplankton, zooplankton, fish biomass)
   b. Chemical fraction unknown (e.g., DON and DOP)
   c. Rate unknown but common across systems (e.g., phytoplankton nutrient uptake)

CONCEPTUAL MODEL: P



CONCEPTUAL MODEL: N



SEDFLUX MODEL

- Sediment diagenesis model
- Predicts sediment-water fluxes
- Inputs (water column conditions)
  - Water depth
  - Water temperature
  - DO concentration
  - NH_3 concentration
  - NO_3^- concentration
  - SRP (PO_4^{3-}) concentration
  - Settling of C, N, and P to sediments

EXTERNAL MASS BALANCE

- Build on work from PSOMAS and SWCA 2007, Merritt and Miller 2016, Brett 2019
- Hydrologic and nutrient budget (inputs and outputs)
- Timeframe: 2015-2020, monthly

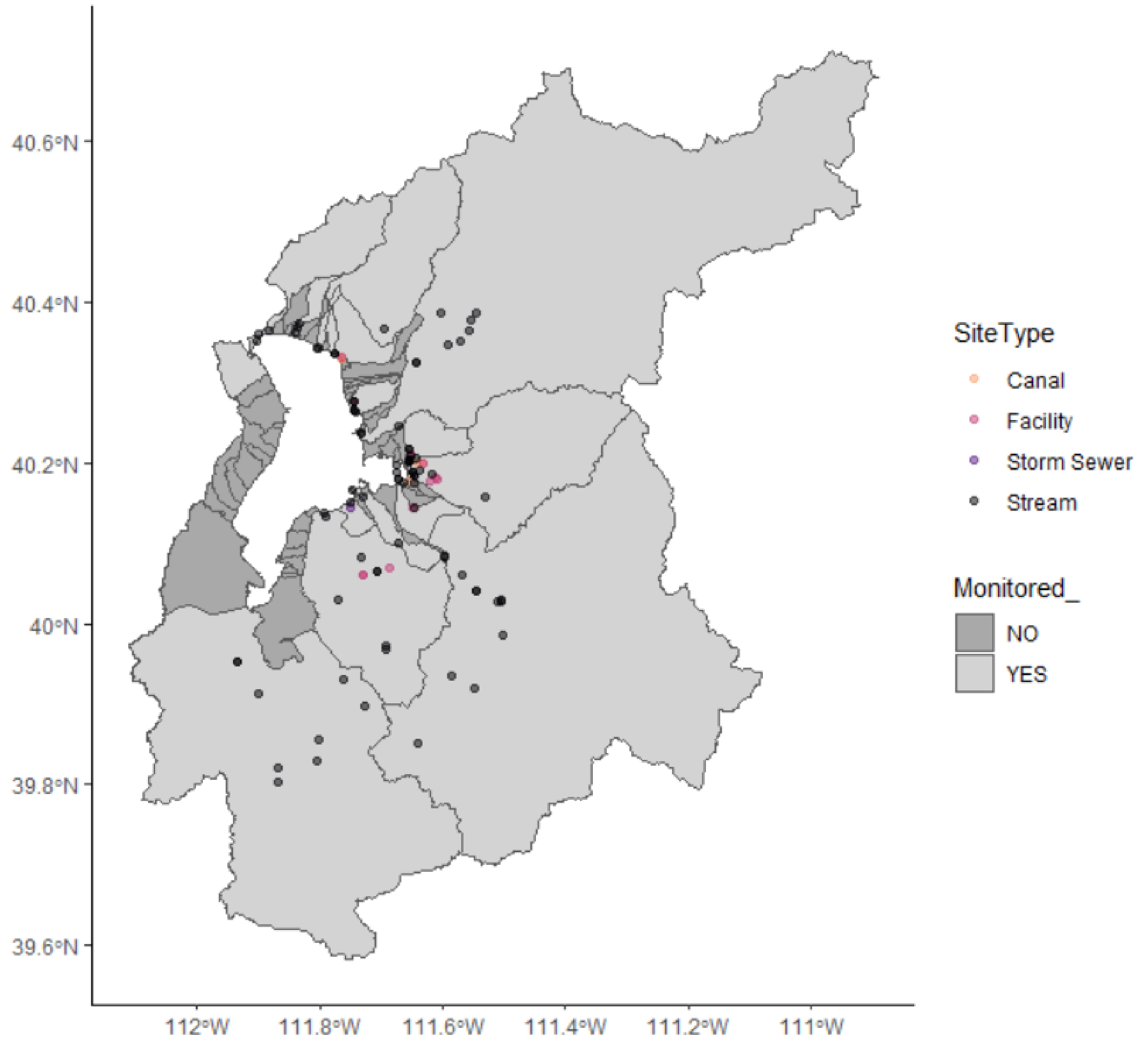
EXTERNAL MASS BALANCE APPROACH

Work to date

1. Processing tributary and facility site data
   a. Select downstream sites within watershed
   b. Analyze DEQ and WFWQC site data when both available in a watershed
   c. Spatial context: downstream of diversion, confluence, hydrology (backwater, intermittent, effluent dominated), flow measurement notes
2. Creating time series of chemistry data
3. Creating time series of flow data

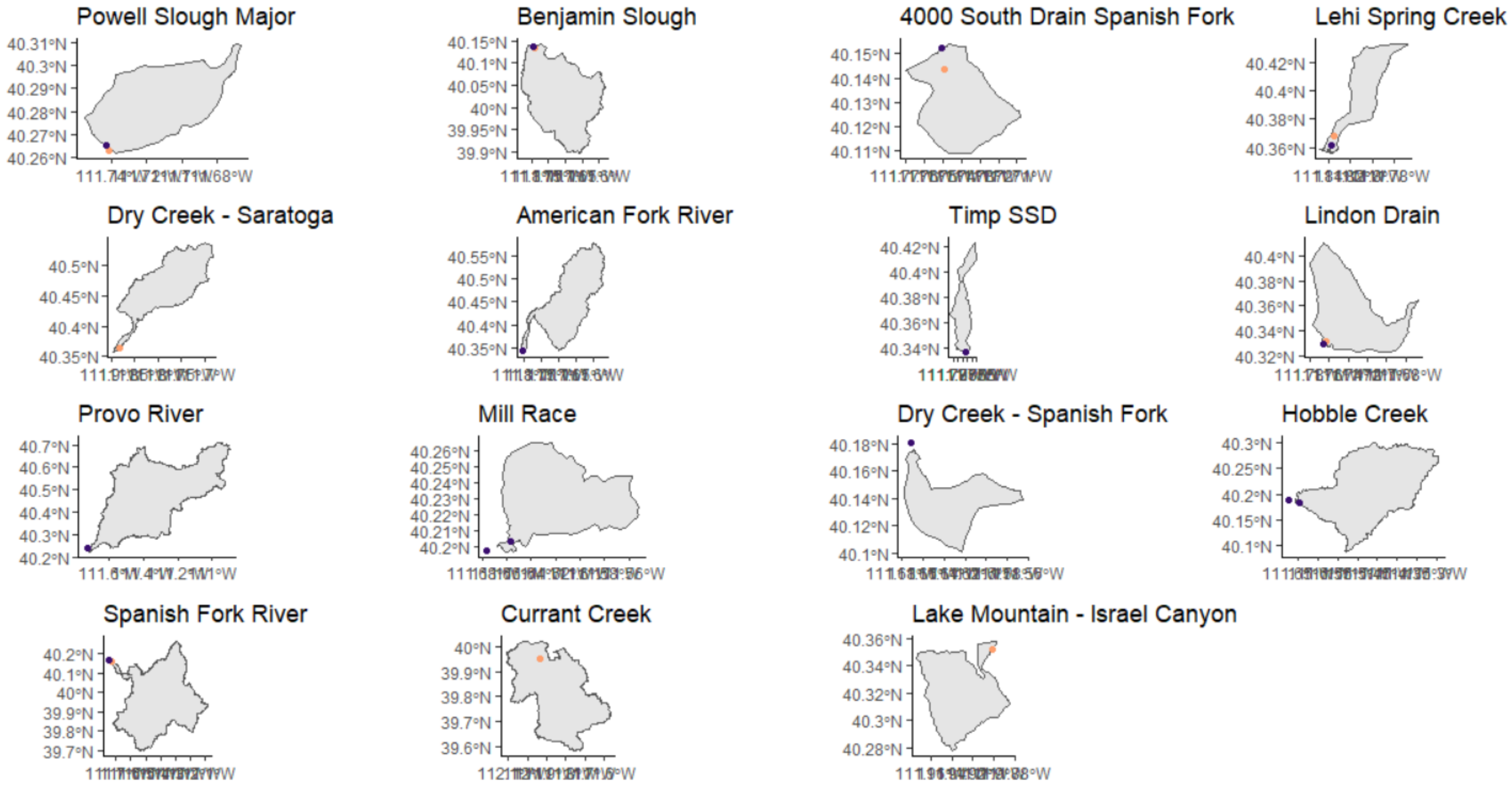
TRIBUTARY AND FACILITY SITES

- Monitored watersheds: downstream monitoring sites → loading estimates
- Unmonitored watersheds: paired watershed approach
  - Area
  - LULC
  - Slope
  - Type of inflow



DOWNSTREAM SITES IN EACH WATERSHED

Note: graph is preliminary, further processing needed (intended as illustration of approach)

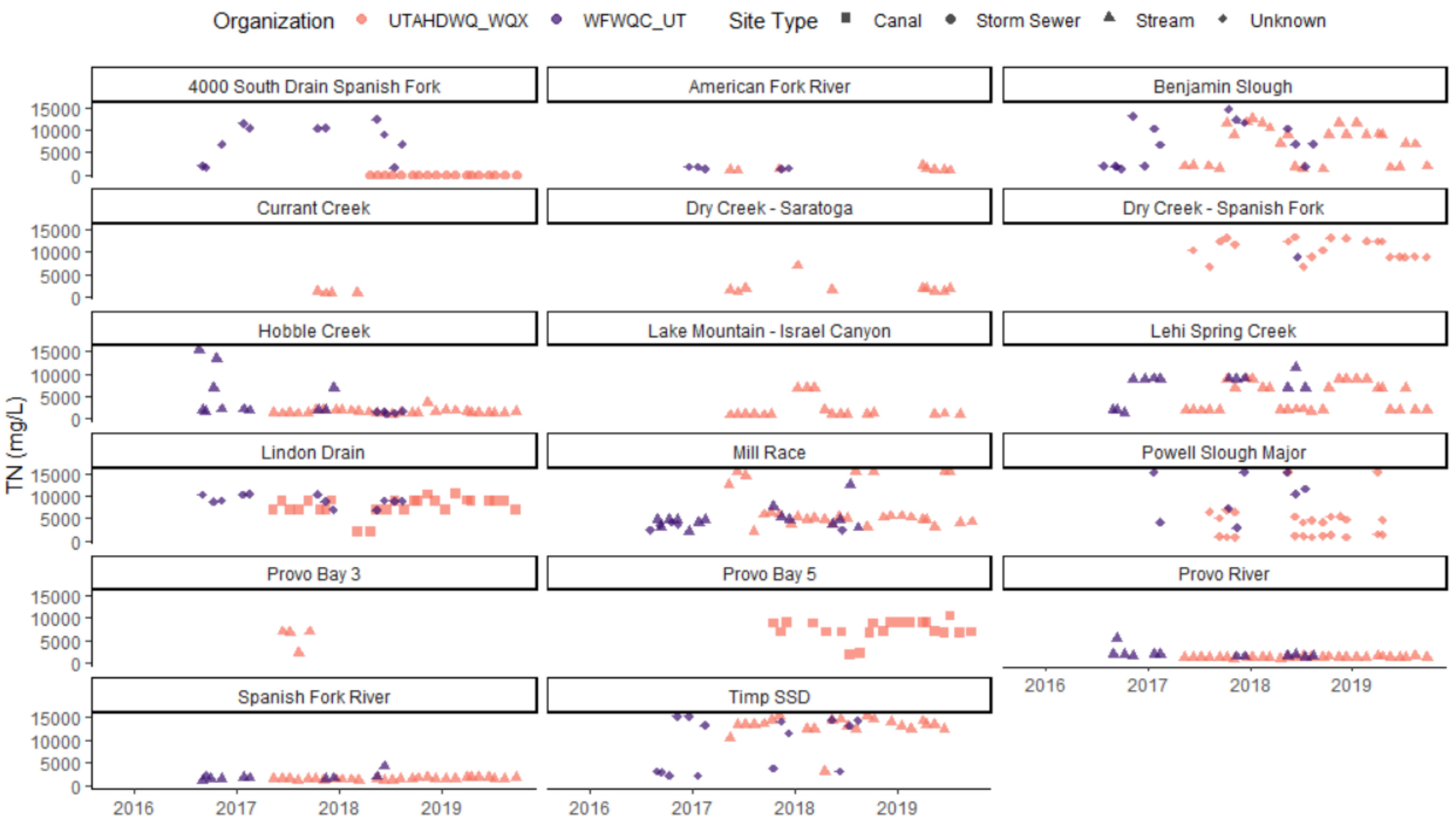


TIME SERIES OF CHEMISTRY

Constituents to be analyzed:

- TN
- TDN
- TP
- SRP
- TOC
- DOC

Note: graph is preliminary, further processing needed (intended as illustration of approach)



EXTERNAL MASS BALANCE APPROACH

Next steps

1. Generate monthly time series
2. Process loading data from WWTP facilities
3. Process precipitation, evaporation data (EFDC/WASP output)
4. Estimate flow and loading from unmonitored watersheds (paired watershed approach)

A mass balance based analysis of internal loading and a prediction of Utah Lake's temporal recovery from external nutrient load reduction

A mass balance based analysis of internal loading and a prediction of Utah Lake's temporal recovery from external nutrient load reduction

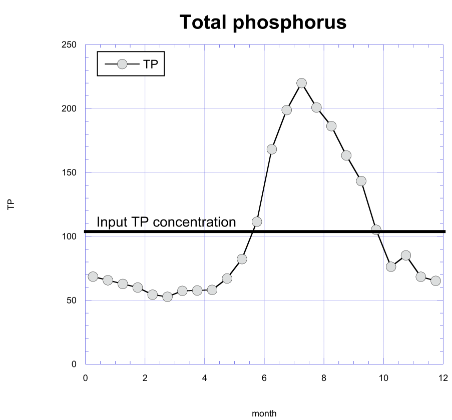
Goals:

1) Define and quantify "internal loading" for both TN and TP

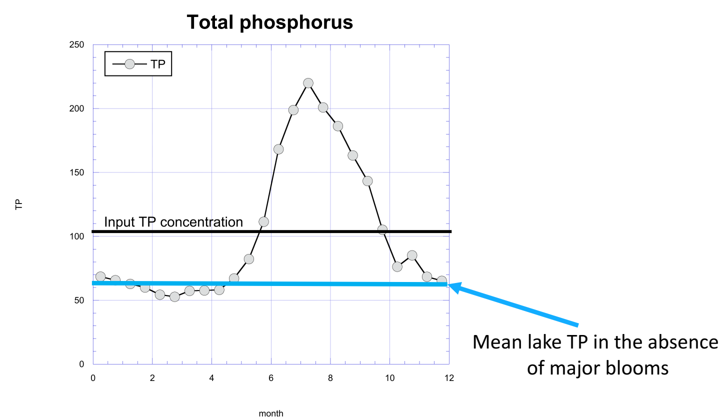
2) Quantify nutrient retention

3) Predict the time frame for recovery from external nutrient inputs

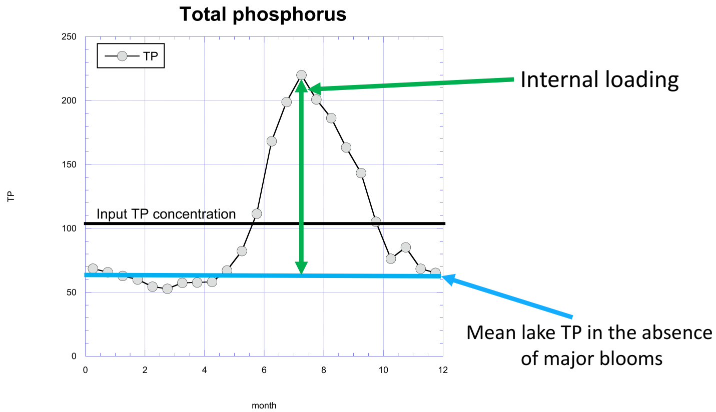
Annual average TP concentrations in Upper Klamath Lake



Annual average TP concentrations in Upper Klamath Lake



Annual average TP concentrations in Upper Klamath Lake



$$Accumulation\ (or\ dilution) = Input - Advection \pm Removal\ (or\ production)$$

Mathematically this can be expressed as:

$$V\frac{dC_{out}}{dt} = QC_{in} - QC_{out} \pm rV$$

where V represents the volume of the reactor (or lake in this example),
C_{out} represents the constituent (or nutrient) concentration in the outflow,
Q_{in} represents the inflow to the reactor,
C_{in} represents the input concentration,
Q_{out} represents the outflow, and
r represents a reaction rate for the removal or production of the constituent in the system.
C_{out} also equals the concentration in the reactor or C_{lake} because the reactor (or natural system) is assumed to be perfectly mixed.

The reaction rate can be due to zero-order, first-order, second-order or even intermediate order processes.

If this Equation is expressed for the case of first-order removal (i.e., removal is proportional to concentration – the most common case), and it is expressed solely in terms of removal (i.e., production is neglected), this equation becomes:

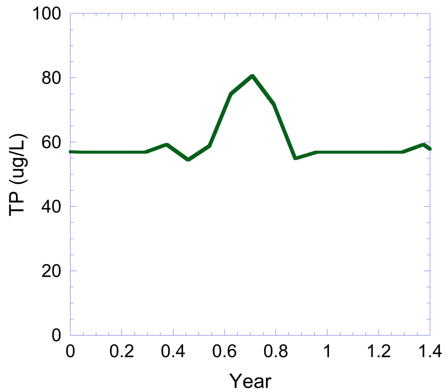
$$V\frac{dC_{out}}{dt} = QC_{in} - QC_{out} - \sigma VC_{out}$$

where σ represents the first-order rate constant for losses of the constituent from the reactor volume.

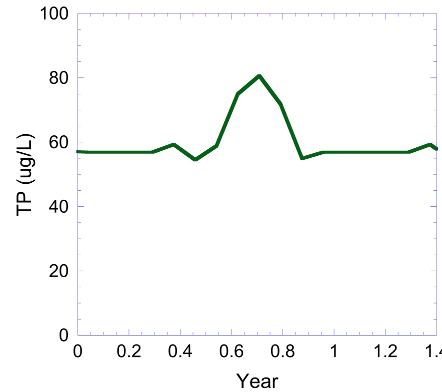
The data used to calculate internal loading:

1. Only the last five years,
2. Utah State Department of Environmental Quality,
3. Samples collected from the main basin of Utah Lake.

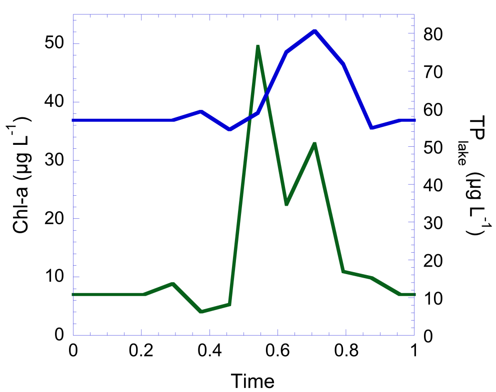
Intra-annual variability in mean TP concentration



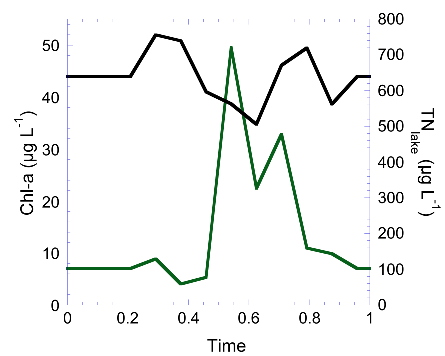
Intra-annual variability in mean TP concentration



Internal loading = +24 µg L^{-1}
or
+42% above baseline
concentrations

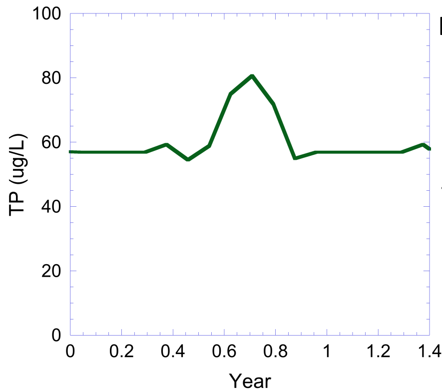


The peak in TP **follows**
the peak in Chl-a



The intra-annual time-series
for TN does not have
an obvious pattern

Intra-annual variability in mean TP concentration



Internal loading = +24 µg L^{-1}
or
+42% above baseline concentrations

To predict recovery we can assume these values will represent future internal loading, but most likely internal loading will slowly decline as external inputs decline

The overall TP removal in Utah Lake for steady-state conditions can be calculated accordingly:

$$Q_{in} * TP_{in} - Q_{out} * TP_{out} = \sigma * V * TP_{out}$$

Because Utah Lake has high evaporative loses:

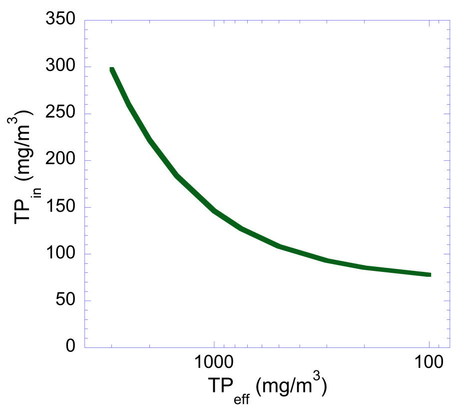
$$Q_{out} \approx 0.5 * Q_{in}$$

At steady-state the first-order rate constant (σ) for TP losses in Utah Lake can be calculated by rearranging the above equations accordingly:

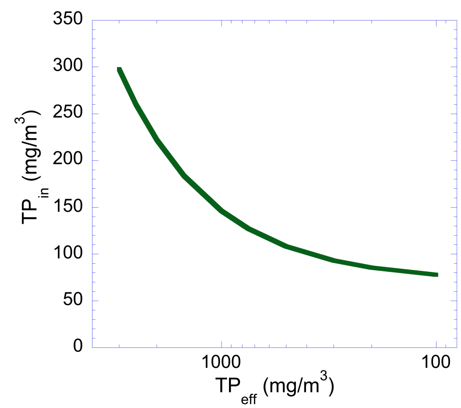
$$\sigma = \frac{\left(\frac{TP_{in}}{TP_{out}} - 0.5\right)}{\theta}$$

Where θ represents V/Q_{out}, the water residence time for Utah Lake.

WWTP effluent TP concentrations vs TP_{in}

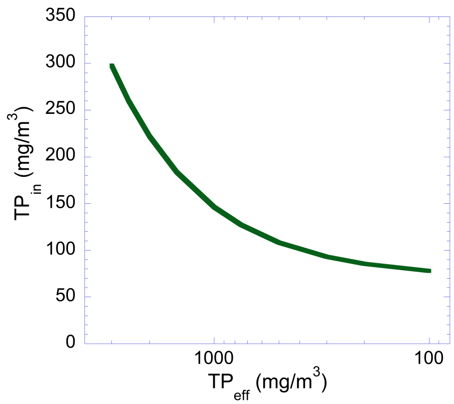


WWTP effluent TP concentrations vs TP_{in}



WWTPs are ≈ 4% of inflows to Utah Lake

WWTP effluent TP concentrations vs TP_{in}



WWTPs are ≈ 4% of inflows to Utah Lake

$$TP_{in} = \frac{\Sigma(Q * C)}{\Sigma Q} = \frac{\Sigma(Q*C_{tributaries} + Q*C_{drains} + Q*C_{precipitation} + Q*C_{WWTP})}{\Sigma(Q_{tributaries} + Q_{drains} + Q_{precipitation} + Q_{WWTP})}$$

Non-steady state mass balance (predicting recovery)

$$C_t = C_\infty + (C_o - C_\infty) * e^{-(\sigma+\rho)*t}$$

or

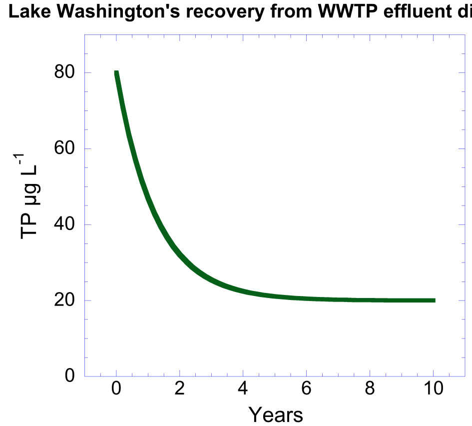
$$TP_t = TP_\infty + (TP_o - TP_\infty) * e^{-(\sigma+\rho)*t}$$

where TP_t represents the TP concentration at some point in time after new discharge concentration takes effect,
TP_∞ represents the new steady-state concentration when the system has reached equilibrium relative to the new discharge condition,
TP_o represents the initial TP_{lake} concentration,
ρ represents θ^{-1} or the lake's flushing rate, and
t represents the time since the change in the discharge condition.

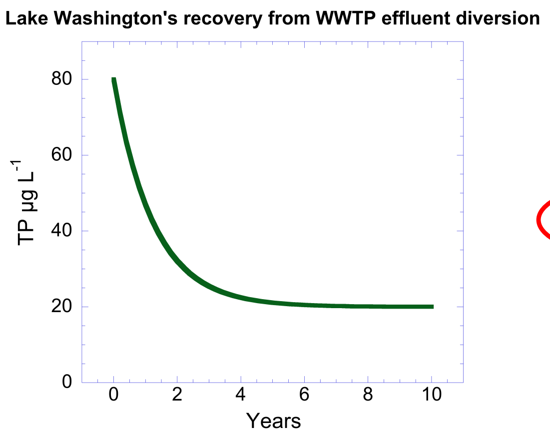
Because of the high evaporative loses for Utah Lake, in this case TP_∞ is calculated accordingly:

$$TP_\infty = \frac{TP_{in}}{0.5+\sigma\theta}$$

Lake Washington's recovery from WWTP effluent diversion



TP_o = 80 µg L^{-1}
TP_∞ = 20 µg L^{-1}
ρ = 0.40 yr^{-1}
σ = 0.40 yr^{-1}



TP_o = 80 µg L^{-1}
TP_∞ = 20 µg L^{-1}
ρ = 0.40 yr^{-1}
σ = 0.40 yr^{-1}

Next steps:

1. Update the estimated TP and TN inputs for non-WWTP sources

2. Repeat the internal loading calculations for Provo Bay

3. Complete the memo describing these analyses and the estimated recovery time for Utah Lake for different WWTP effluent TP and TN concentrations

4. Repeat these calculations at a monthly time step for Q_{in}, Q_{out}, and lake volume

UPDATES ON ANALYSIS REPORT: PHYTOPLANKTON MODELING

Utah Lake Water Quality Study
Science Panel Call
2020-12-14

UTAH DEPARTMENT of ENVIRONMENTAL QUALITY

RECAP & RATIONALE

1. Analysis report has section on phytoplankton temporal and spatial modeling
   a. Test for a relationship between nutrient concentrations and HAB abundances
   b. Test for a relationship between lake level and HAB abundances
   c. Test for a relationship between temperature, stratification and HAB abundances
2. Steering Committee question on management goals: relationship between cyanobacteria and nutrients
   a. Bivariate relationship
   b. Covariates hypothesized as mediating factors
   c. Abundance metrics (cell count, biovolume)

→ Combine analyses as a comprehensive phytoplankton analysis

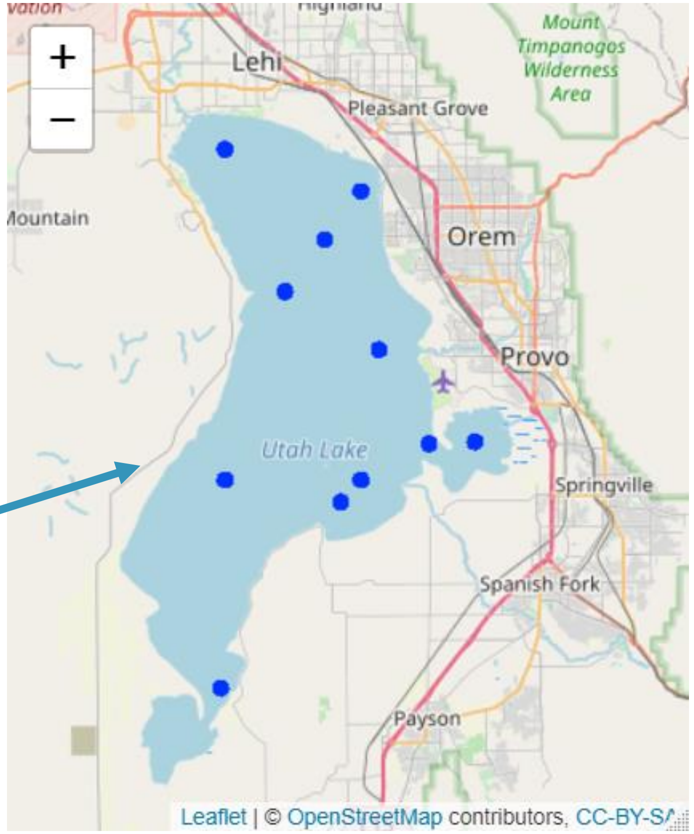
DATASET OPPORTUNITIES AND LIMITATIONS

HAB Advisory Program dataset

- Cell count, biomass, *toxins*
- Includes composite surface and surface "scum" samples
- Beach and marina samples are abundant due to sampling scheme

Routine Monitoring Program dataset

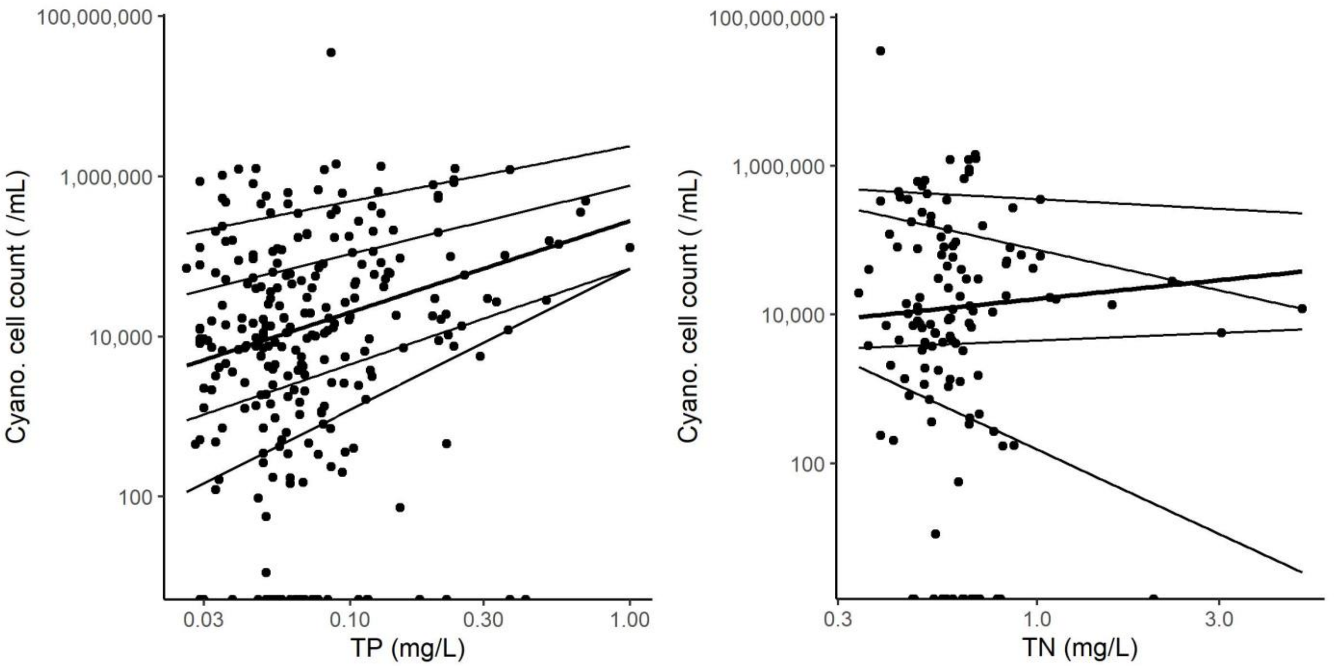
- Cell count, biomass, *chemistry*
- Includes only composite surface samples
- Paired phyto-chem samples are open water samples
- Water depth not available at all sites → overall lake elevation vs. water depth at a given site



RECAP OF MANAGEMENT GOALS QUESTIONS

Recall: bivariate relationship of TP or TN with cyanobacterial cell count has a great deal of scatter

→ hypothesized that covariates may explain variability

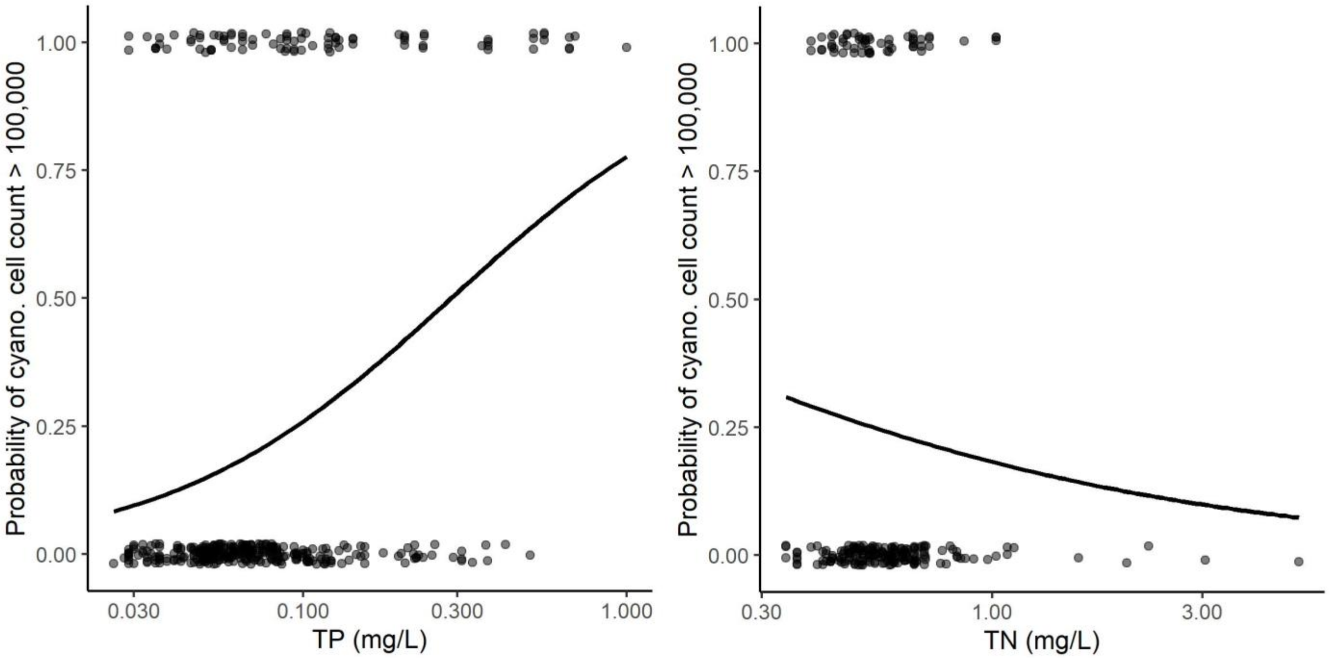


RECAP OF MANAGEMENT GOALS QUESTIONS

Logistic regression enables assessment of risk

Covariates can be examined here as well

Important: need to set a response variable threshold before analysis



MIXED EFFECTS MODELING

Response variables

- Cyanobacteria cell count
- Total phytoplankton cell count
- Cyanobacteria biovolume
- Total phytoplankton biovolume

Predictor variables – matched by site and date

- TP or TN
- Turbidity
- Lake elevation
- Month (categorical)
- Water temperature → collinear with month → removed from model

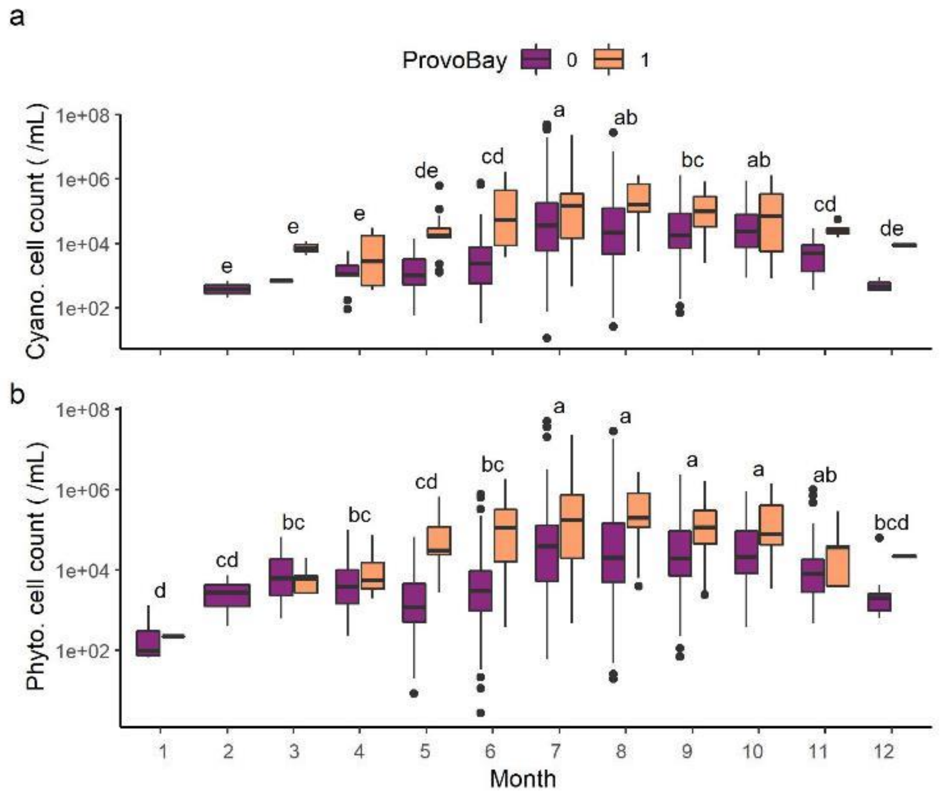
Random effect: site

SPATIAL CONSIDERATIONS

Site as random effect → characterizes spatial variability across the lake, does *not* predict conditions at a given site

Recall that Provo Bay (orange) typically has higher cell counts and biovolumes than the rest of the lake → is a separate model specifically for Provo Bay desirable? Key assumptions:

- Interested in specifically modeling conditions in Provo Bay
- Provo Bay operates biogeochemically differently than main basin



STATISTICAL APPROACH

Theron Miller suggestions:

- Poisson regression
- Negative binomial regression
- Truncated Poisson or negative binomial regression
- Machine learning methods such as Random Forest

Approach taken so far:

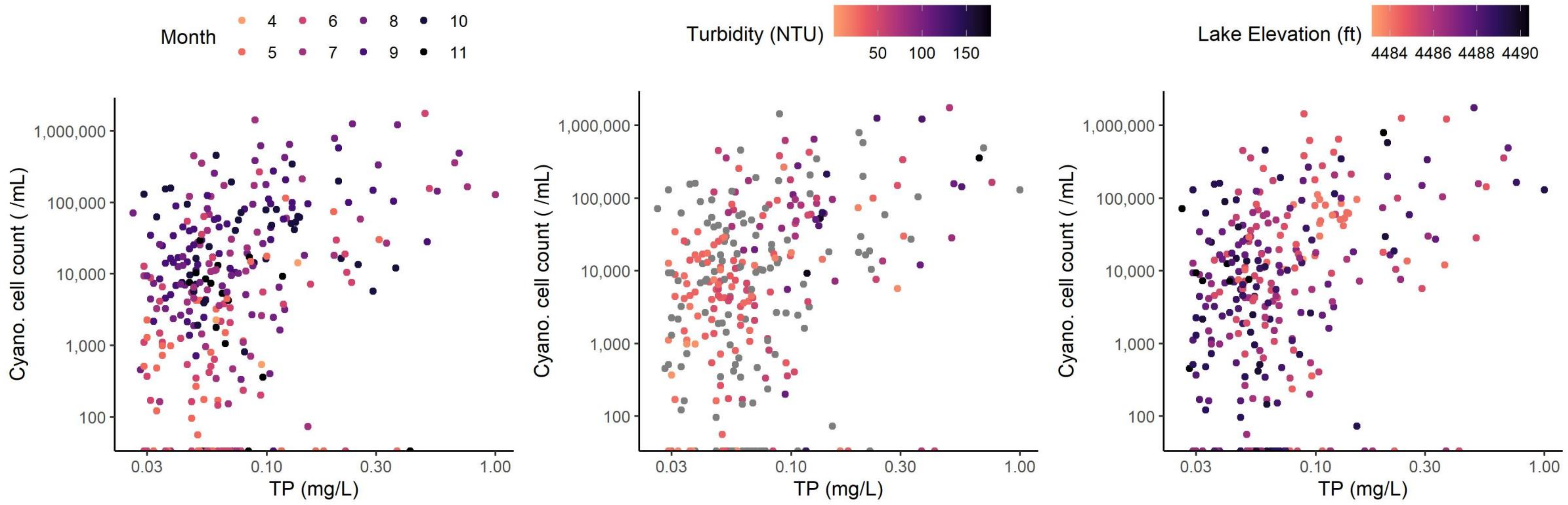
- Multiple regression, with fixed and random effects
- Anticipate that inclusion of covariates negates need for quantile regression (covariates explain “wedge” response)
- Variables log-transformed when appropriate (assumptions of log-normality confirmed)
- Exploring negative binomial distribution for cell count data (zero-inflated for cyano cell count)
- We are interested in modeling response variables by a linear combination of coefficients → Random Forest helps to assign variable importance but generating posterior estimates are not straightforward

PHYTO ~ TP MODELS

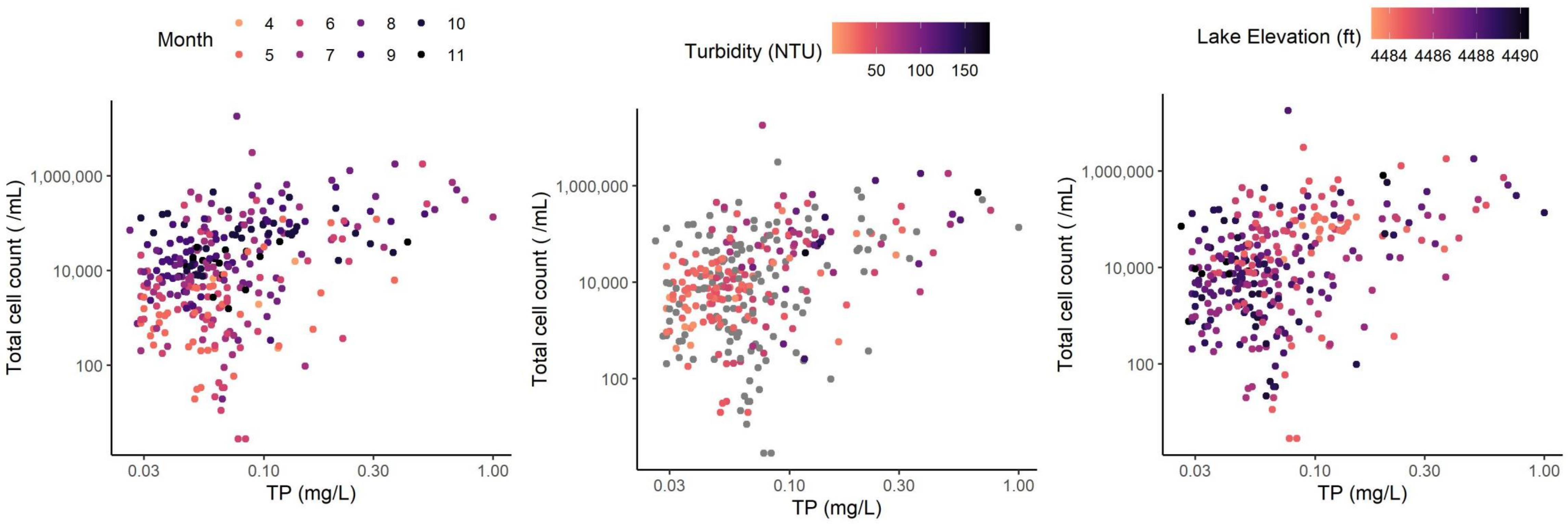
- TP always positive
- Turbidity negatively associated with cell count (cyano and total)
- Lake elevation negatively associated with all except cyano biovolume
- Month is a significant categorical predictor for all
- Spatial random effect increases explanatory power 12-16%

Variable	cyano. cell count	total cell count	cyano. biovol.	total biovol
log(TP)	+	+	+	+
log(Turbidity)	-	-		
Lake elevation	-	-		-
Month	+/-	+/-	+/-	+/-
1|Site	Random	Random	Random	Random
Marginal R^2	0.30	0.36	0.29	0.40
Conditional R^2	0.44	0.48	0.42	0.46

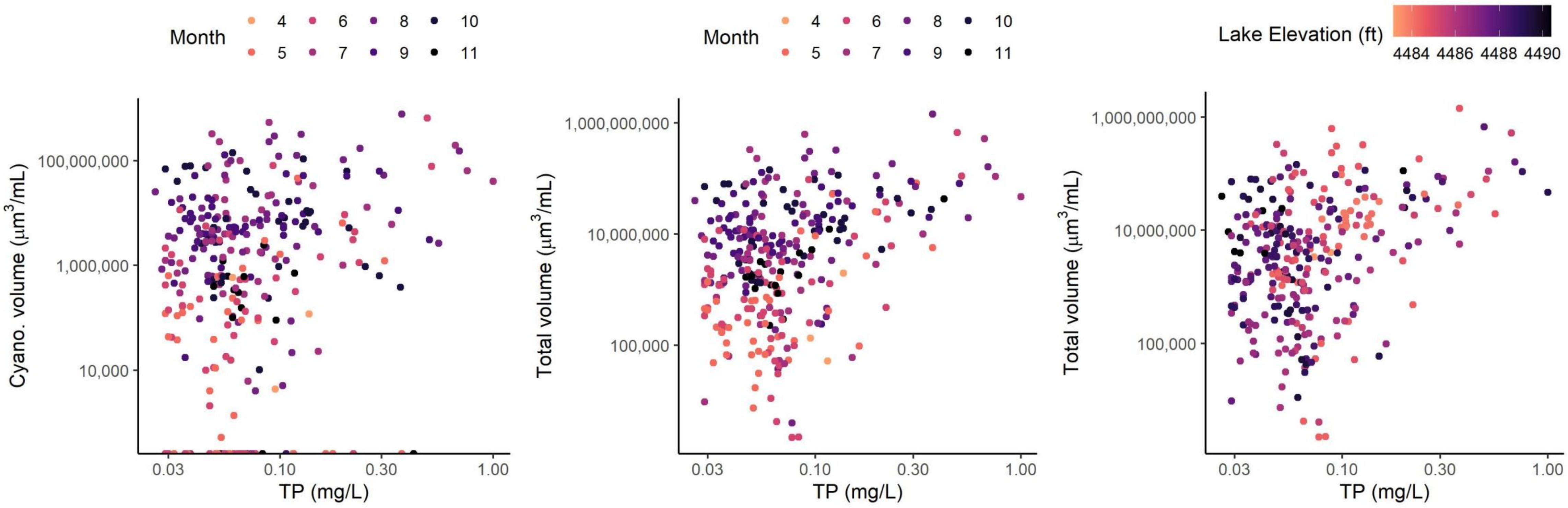
COVARIATES FOR CYANO CELL COUNT ~ TP



COVARIATES FOR TOTAL CELL COUNT ~ TP



COVARIATES FOR CYANO/TOTAL BIOVOLUME~ TP

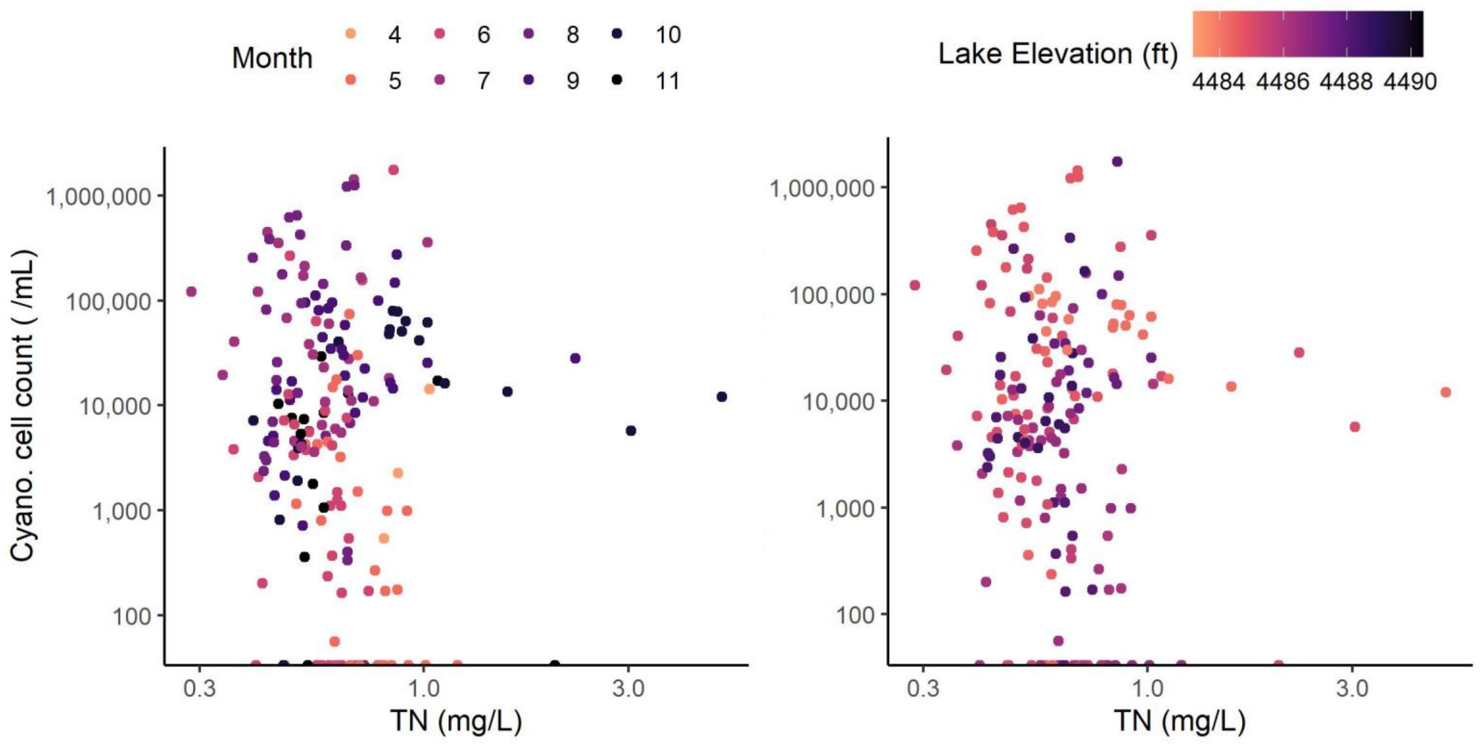


PHYTO ~ TN MODELS

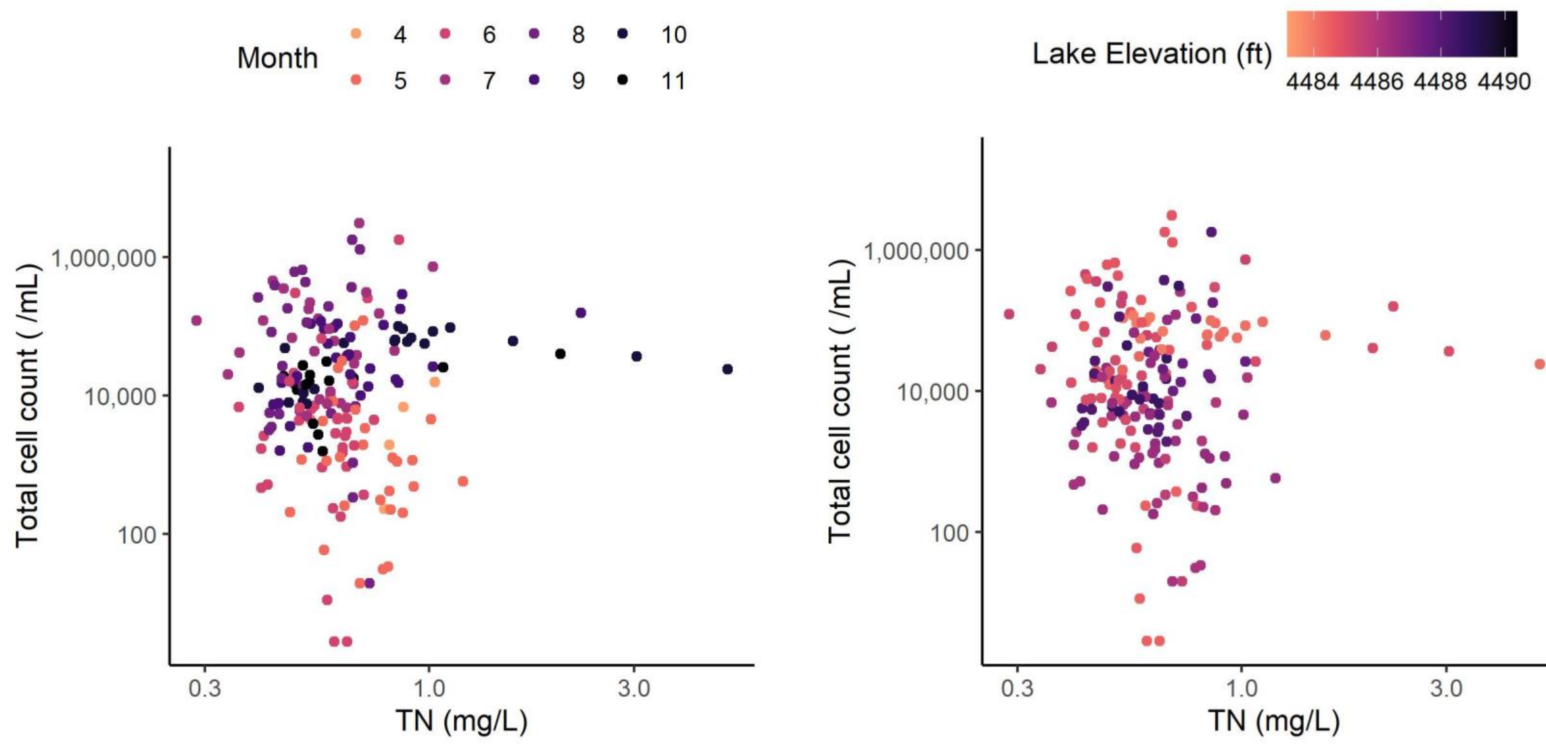
- TN always negative
- Turbidity not a significant predictor
- Lake elevation negatively associated with all
- Month is a significant categorical predictor for all
- Spatial random effect increases explanatory power 11-31%

Variable	cyano. cell count	total cell count	cyano. biovol.	total biovol
log(TN)	-	-	-	-
log(Turbidity)				
Lake elevation	-	-	-	-
Month	+/-	+/-	+/-	+/-
1|Site	Random	Random	Random	Random
Marginal R^2	0.29	0.23	0.29	0.27
Conditional R^2	0.44	0.54	0.40	0.53

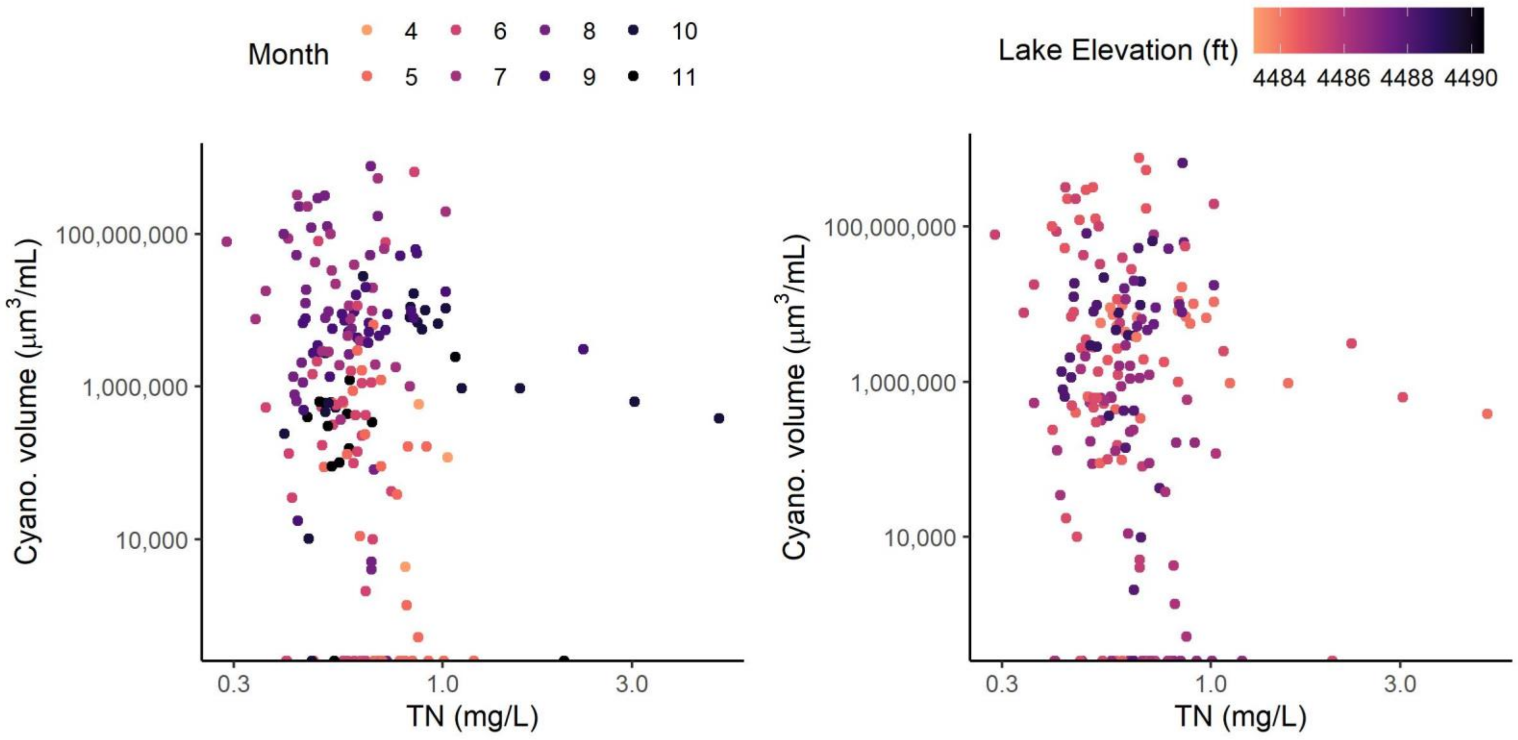
COVARIATES FOR CYANO CELL COUNT ~ TN



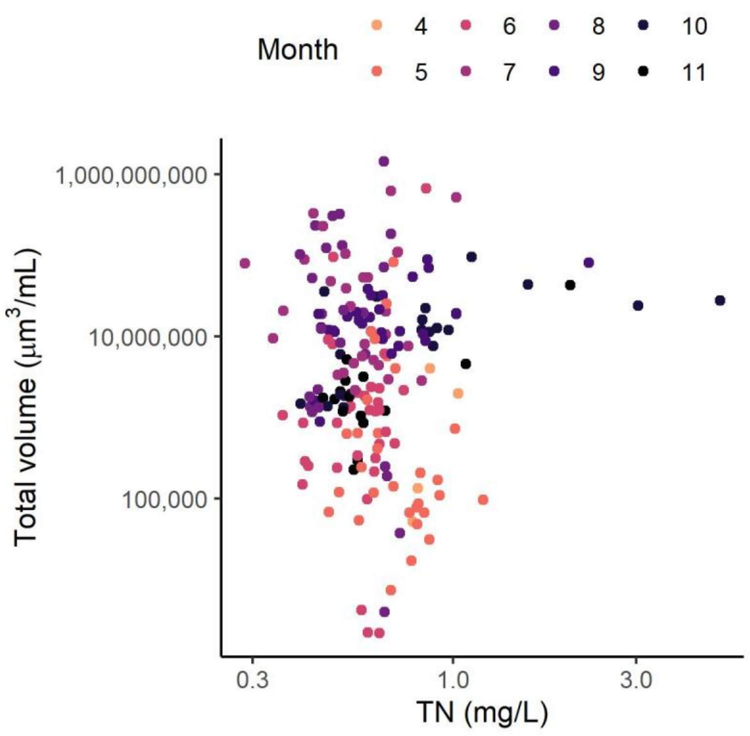
COVARIATES FOR TOTAL CELL COUNT ~ TN

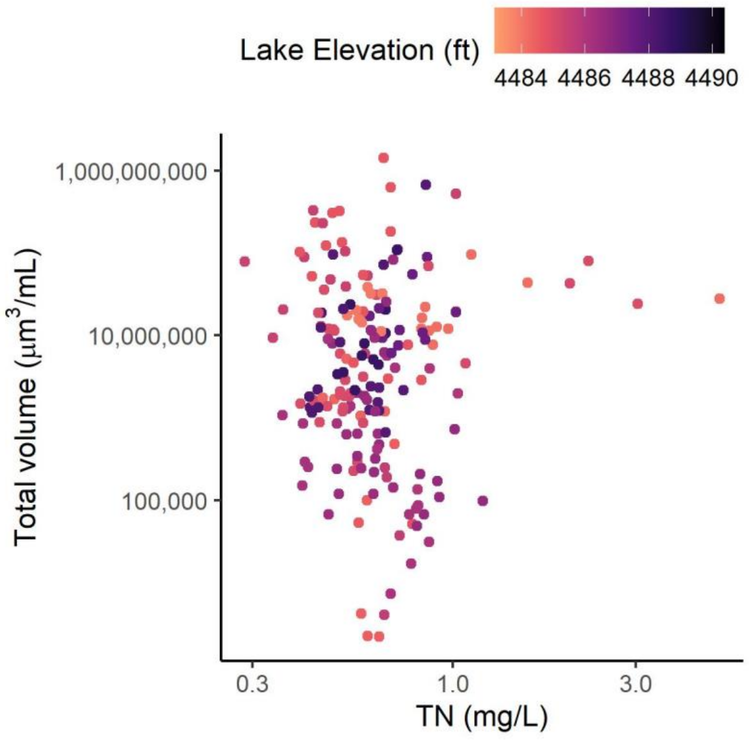


COVARIATES FOR CYANO BIOVOLUME~ TN



COVARIATES FOR TOTAL BIOVOLUME ~ TN





FURTHER ANALYSIS

Are we interested in other aggregations and/or antecedent conditions?

- Seasonal means
- Subset of months
- Spring nutrients → summer phytoplankton
- Additional antecedent predictors e.g., precipitation, evaporation (see analysis report)

WRAP-UP DISCUSSION

Do these models get us answers to the questions we have? If not, how can models be changed/improved?

Considerations:

- Limitation of datasets → which aspects of interest can reasonably be explored with existing data?
- Statistical approaches → does mixed effects GLM (log-normal and negative binomial distributions) answer the relevant questions?
- Accounting for spatial component → what aspect of spatial variability are we interested in accounting for?
- Additional aggregations or antecedent conditions?