Controls on water quality in the western U.S.

Bella Oleksy

Last compiled on 03 September, 2021

Contents

Project motivation **1**
Impetus for this work . . . 1
Research questions . . . 2
Hypothesized patterns . . . 2
Some papers relevant to this work . . . 3
Data we will use or still need . . . 4
Expertise needed . . . 4
Document structure . . . 4

1 Summary stats on LAGOS NETS **5**
1.1 Inventory of spatial data . . . 5
1.2 Patterns in western US hydrology/networks . . . 10

2 Spatial patterns in TP and NO3 concentrations **12**
2.1 Elevation and connectivity classs . . . 12
2.2 Upstream waterbodies . . . 13
2.3 Hydrology . . . 15
2.4 Climate . . . 18

Project motivation

Impetus for this work

- Climate change is not only a threat to water **quantity** in the western United States, but likely also water *quality*.
  - It is estimated that ~60 million people rely on snowmelt for drinking water and irrigation in the region
  - Recent climate change and drought-related dust deposition events are implicated in the decline in snowpack in the western U.S., and resultant declines in streamflow
- While there has been a large emphasis on how climate change is impacting and will continue to impact water quantity in the western U.S., there has been less attention on how climate and land use change will impact water *quality* in lakes and reservoirs.
  - Most of what we know about the controls on water quality come from a legacy of research in temperate regions
  - However, the western U.S. is climatologically very different than the northeastern United States. Some general characteristics include:
    * Snowmelt driven hydrology
    * Extremely high rates of evaporation and evapotranspiration in the terrestrial landscape
    * Drought and fire are common and are increasing in their frequency and intensity
- In the West, water is highly managed and controlled because it is scarce.

  - Reservoirs make up a high proportion of water bodies in the western U.S.
  - We have many reasons to believe that the controls on water quality in these different water bodies differs in systematic ways, and that natural vs. man-made water bodies respond differently to climate change (Hayes et al. 2017)
- Broad scale motivation
  - Given the ongoing and anticipated future stress on water resources (e.g., drought, heat waves, wildfires) it is critical to understand the local and regional controls on water quality.
  - A deeper understanding will allow for better prediction of vulnerabilities in our water resource systems.
  - A focus on hydrologic connectivity may help us understand macroscale patterns in water quality

Research questions

1. What are the landscape, climatic, and hydrologic controls on total P concentrations?*
2. Do the controls vary by ecoregion?
3. Are the controls on water chemistry different in natural vs. man-made?

*See data summary below—TP is where we have the most data, but we can also add in nitrate without losing too many sites.

Hypothesized patterns

1. Total P will vary based on landscape or network position;
   a. Specifically, I hypothesize that in hydrologically connected lake networks, TP concentrations will be generally lower in lakes down network of other lakes due to processes that promote burial of sediment/organic matter (longer residence times).
   b. Further, because the elevational gradients are greater in the western US than the north-east/midwestern US, I hypothesize that we will see a greater effect of elevation on TP concentrations. For similar mechanisms as above, in drainage lakes connected to upstream lakes, TP concentrations will be lower in low elevation lakes compared to higher elevation lakes.
   c. Total P in hydrologically isolated lakes or drainage lakes lacking connections to upstream lakes will be lower than other lake types due to smaller watershed contributing areas
2. Total P will vary by climate and land use
   a. Like a lot of macroscale studies, I hypothesize that we will see a relationship of higher TP concentrations in areas with high % of cropland or pasture
   b. Unlike other macroscale studies, in this region I don't think we will see a strong relationship between forest cover and lower TP concentration. My reasoning is that there is a very specific band where we see forest cover in the western U.S. and it's mostly in the mountainous areas. My guess is that a hydrologic signal or winter climatic signal will overpower the effect of forest cover to control TP concentrations in those regions.
   c. I hypothesize that TP concentrations will be relatively low in the maritime climate regions of the Pacific Northwest and in areas that receive a large amount of their precipitation in winter as snow.
3. Total P will be higher in reservoirs than natural lakes
   a. This is a pattern that we see in the EPA NLA dataset and I can only imagine that if we expand our sample size this pattern would follow (see Figure 1).
   b. Why is this the case? It could be because of the tendency for reservoirs to be in larger watersheds relative to lakes of similar sizes. Therefore, compared to a natural lake of similar size, the WSA:LA is generally higher in a reservoir resulting in higher mass inputs than natural lakes.
   c. Another reason could be greater lake-level fluctuations in reservoirs, which could indirectly impact water chemistry via changes in lake stratification/mixing

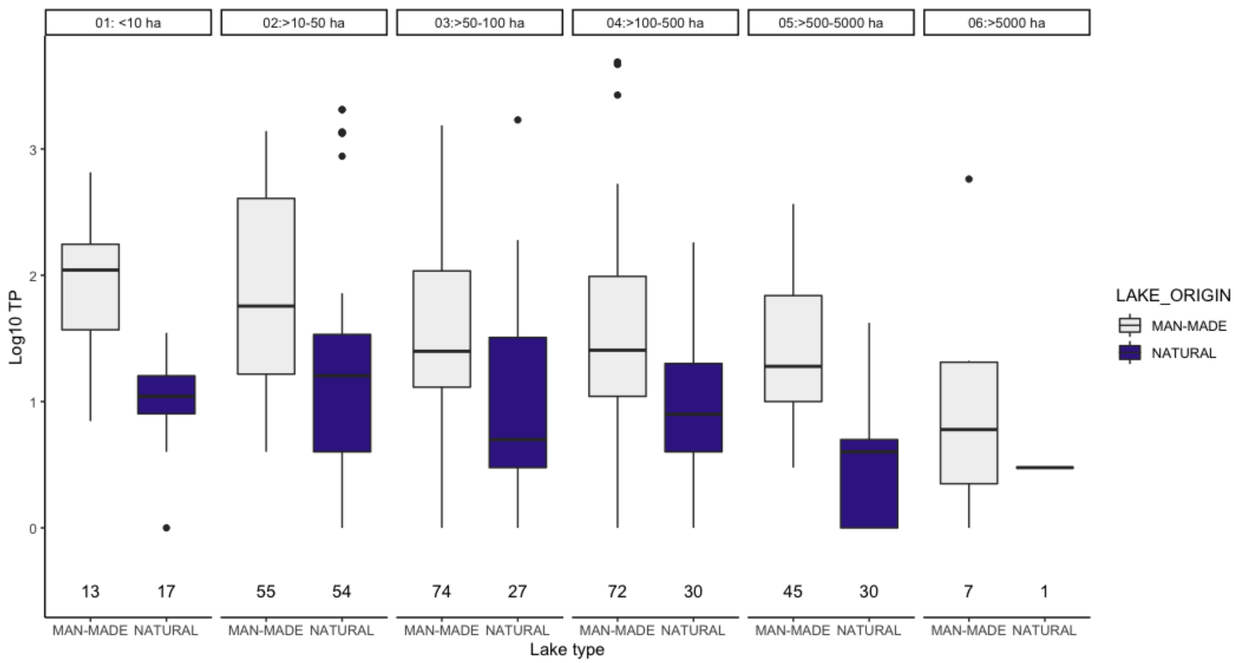


Figure 1: *Total P concentrations by lake size and lake type. Source: EPA NLA 2007*

Some papers relevant to this work

Fergus et al (2020). Lake Water Levels and Associated Hydrologic Characteristics in the Conterminous U.S.

- Lakes, but especially reservoirs, in the western mountain ecoregion experience much greater drawdown than other regions of the U.S. – this of course will have implications beyond water balance and will impact lake physics, habitat, and ultimately chemistry.

Mosley LM. 2015. Drought impacts on the water quality of freshwater systems; review and integration.

- Prolonged (or severe) droughts can lead to longer water residence times, lower dilution potential, and increasing nutrient levels in lakes

Hayes et al. (2017). Key differences between lakes and reservoirs modify climate signals: A case for a new conceptual model.

- This paper was one of my initial motivators for this work. As far as I can tell, there hasn't been any follow-up work explicitly looking at differences in water quality in natural lakes versus reservoirs, or testing any of the climatic sensitivity hypotheses that they pose in this paper.

Gardner et al. (2019). The Abundance, Size, and Spacing of Lakes and Reservoirs Connected to River Networks.

- They characterize lake-reservoir-river network in the conterminous United States into distinct typologies. Might give us some food-for-thought relating to how to characterize some of the networks in our study, or how to interpret results
- I like this bit: *"By coupling a new understanding of river network topology with geospatial data sets and network modeling, we can examine how different types of river networks with lakes/reservoirs transport sediment, propagate geomorphic adjustment process carbon and nutrients and disperse species."*

McCullough et al. (2019). Do lakes feel the burn? Ecological consequences of increasing exposure of lakes to fire in the continental United States.

- Wildfires are ubiquitous enough in the western U.S. that we may see a signal in the water quality data. It would be worth revisiting the conceptual model in this paper.

Oleksy et al. (2020). The role of warm, dry summers and variation in snowpack on phytoplankton dynamics in mountain lakes.

- This is most definitely not a macroscale study, but at least in high-elevation lakes in the southern Rockies we see that snowpack and summer drought impact chlorophyll concentrations more so than land cover, lake position, etc. Researchers like Steve Sadro and his colleagues have seen similar things in the Sierra Nevada range. It would be cool to test the conceptual framework that we pose at the end of the paper.

Data we will use or still need

Table 1: Data to use in this project

dataset	status
LAGOS-US LOCUS	Available internally
LAGOS-US GEO	Release sometime Sept 2021
LAGOS-US LIMNO	Available internally
LAGOS-US NETS	Published!
LAGOS-US RSVR	Hopefully available soon?
LAGOS-US DEPTH	Supposedly available internally, but I don't see it on Dropbox
LakeCat	Published

Expertise needed

In addition to me and Sarah's knowledge on the controls on water quality, it would be great to have some people on board who think a lot about the role of hydrology and connectivity in structuring limnological patterns across scales. Emi Fergus and K. Webster would be perfect for this. Once this project has some stronger legs, we can open it up to the Modelscape aquatic folks to bring in some river knowledge.

As a first approach, I think it would be cool to use sparse modeling to narrow down our predictor set. However, the limitation of something like SuSiE or Lasso is that as far as I can tell the model can identify interaction (at least I haven't figured out how to do this). So, I think the impact of the paper would probably be more powerful if we could pair with either with some sort of regression tree/machine learning algorithms (e.g., boosted regression trees, RF) or Bayesian hierarchical modeling to account for all the different scales we are trying to compare across. Someone more knowledgable than me on spatial statistics would also be valuable to the team.

Document structure

In the remainder of the document, I'll give the reader an idea of data availability for TP as well as NO3, our second most abundant measurement in the western U.S. subset of lakes. We'll have to give some thought as to whether we want to focus the analysis strictly on TP, or also look at N. There are some key differences between TP and NO3 which are interesting to think about.

1 Summary stats on LAGOS NETS

1.1 Inventory of spatial data

Table 2: Table summarizing the 10 top most abundant water quality parameters in the LAGOS-US database where 'count' refers to the number of unique lakes

parameter	count
tp_ugl_median	1205
no2no3n_ugl_median	1184
nh4n_ugl_median	992
spcond_uscm_median	747
secchi_m_median	637
chla_ugl_median	551
srpp_ugl_median	551
mg_mgl_median	494
tn_ugl_median	388
turb_ntu_median	337

The first thing I want to do is taken an inventory of the water quality data that we have so far. Some caveats:

- Currently the summary data I'm showing from LAGOS-US uses preliminary nutrient data from Dropbox.
- I filtered out the "western US" by state, including: CA, UT, NV, WA, OR, ID, MT, WY, CO, NM, and AZ. We will eventually want to be more thoughtful about how we break this up geographically in a way that is relevant to hydrology/limnology (e.g., by HU4 or major drainage basin boundaries).

What are the top 10 most abundant measured parameters, and how many lakes do we have for each of those parameters? The table above shows the breakdown.

How many lakes have both TP and NO3?

```
## [1] 855
```

How many lakes have TP, NO3, and NH4?

```
## [1] 743
```

How many lakes have TP, NO3, and secchi?

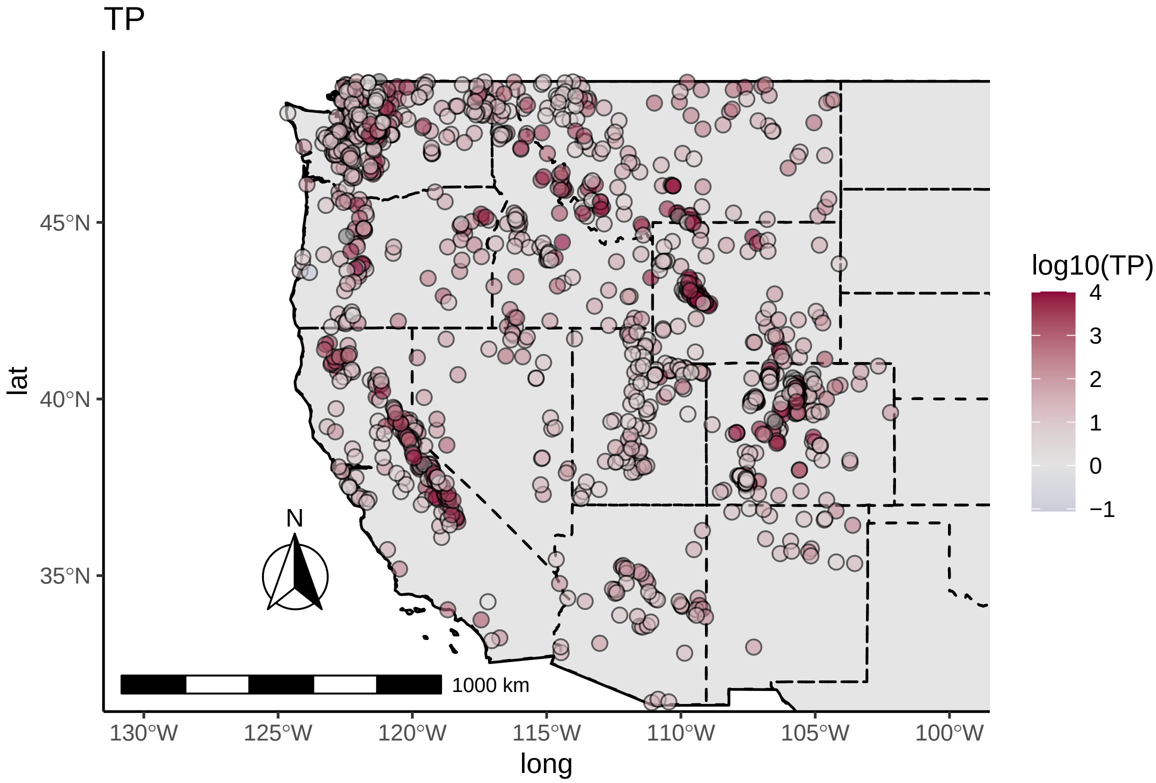
```
## [1] 334
```

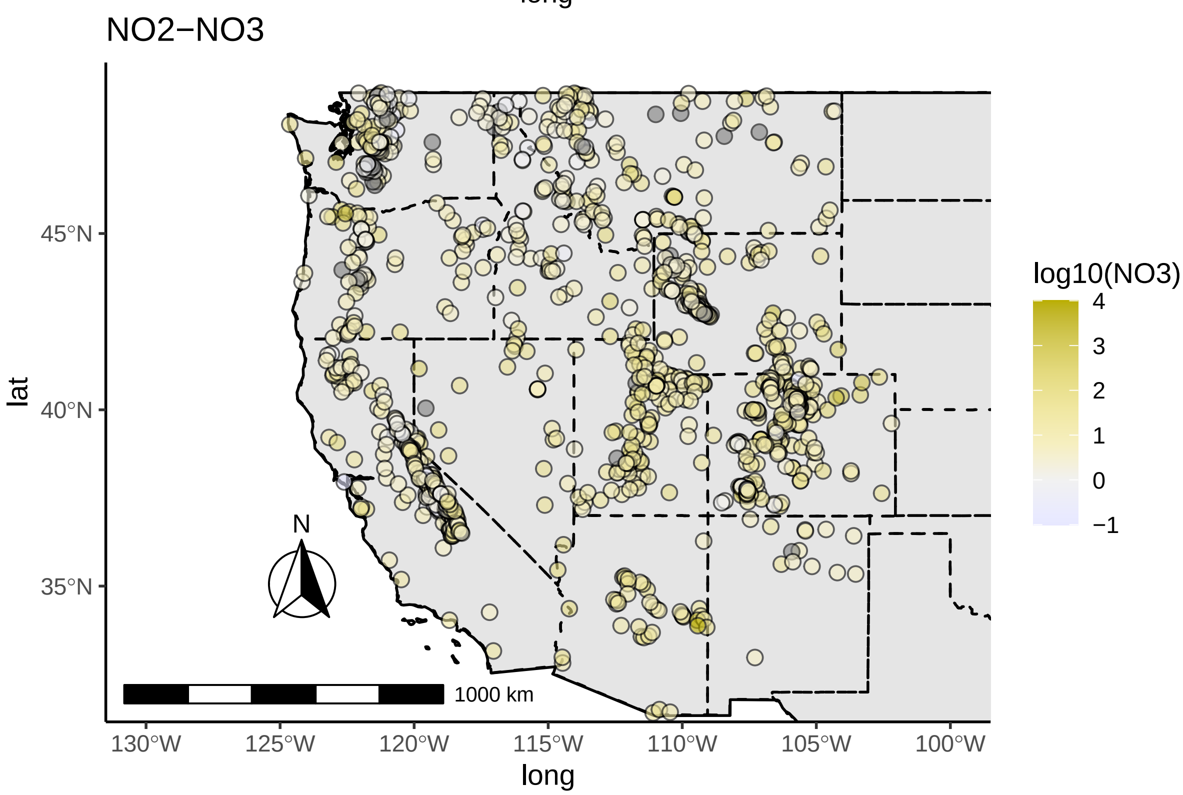
As you can see, once we go beyond the top two parameters (TP and NO3) our sample size of lakes decreases quite rapidly (assuming we want to study the same population of lakes).

1.1.1 Maps of spatial data

1.1.1.1 Concentrations across space

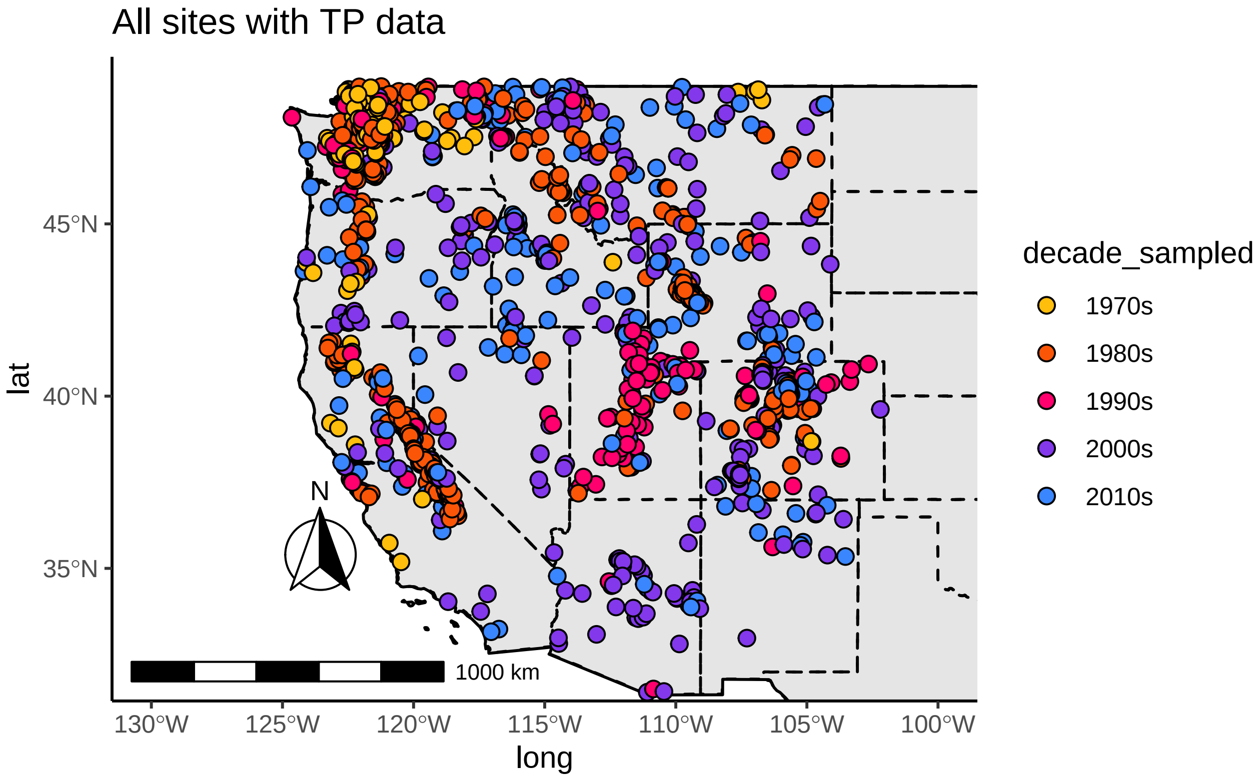
Total phosphorus and NO3 concentrations across space

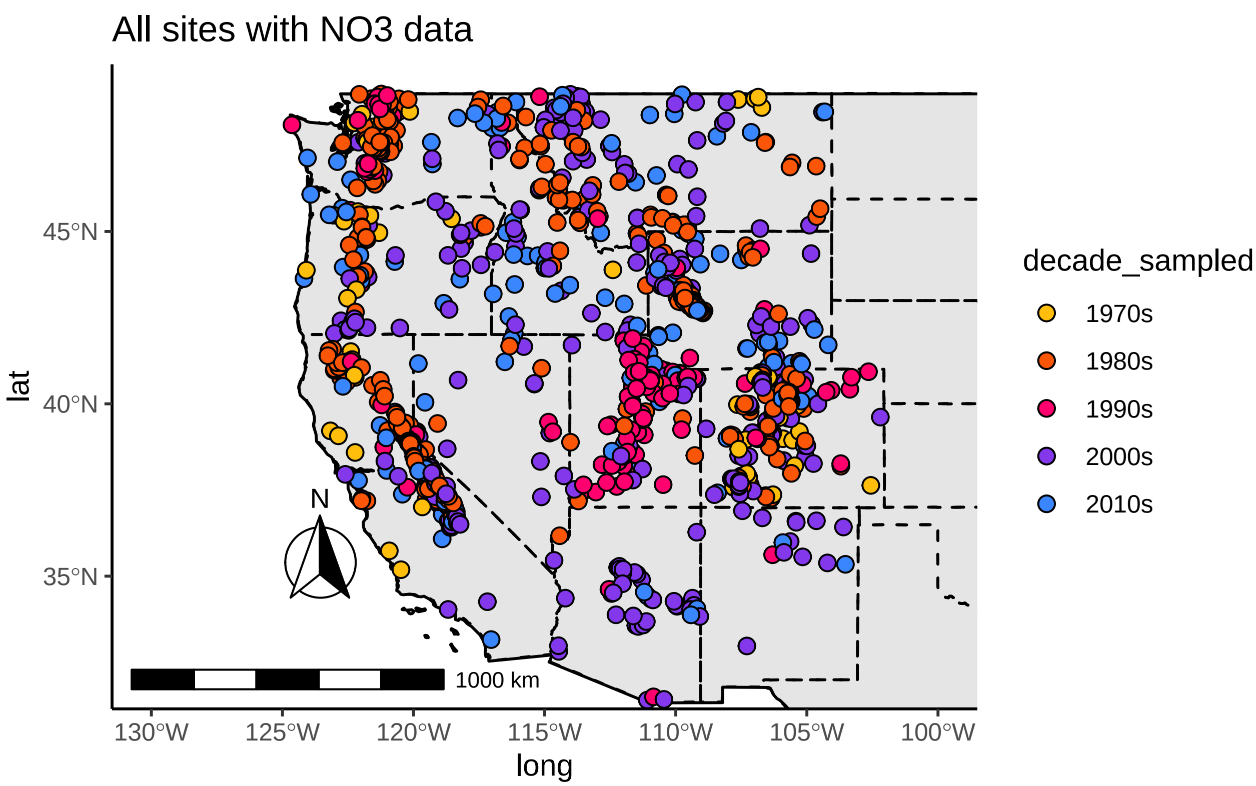




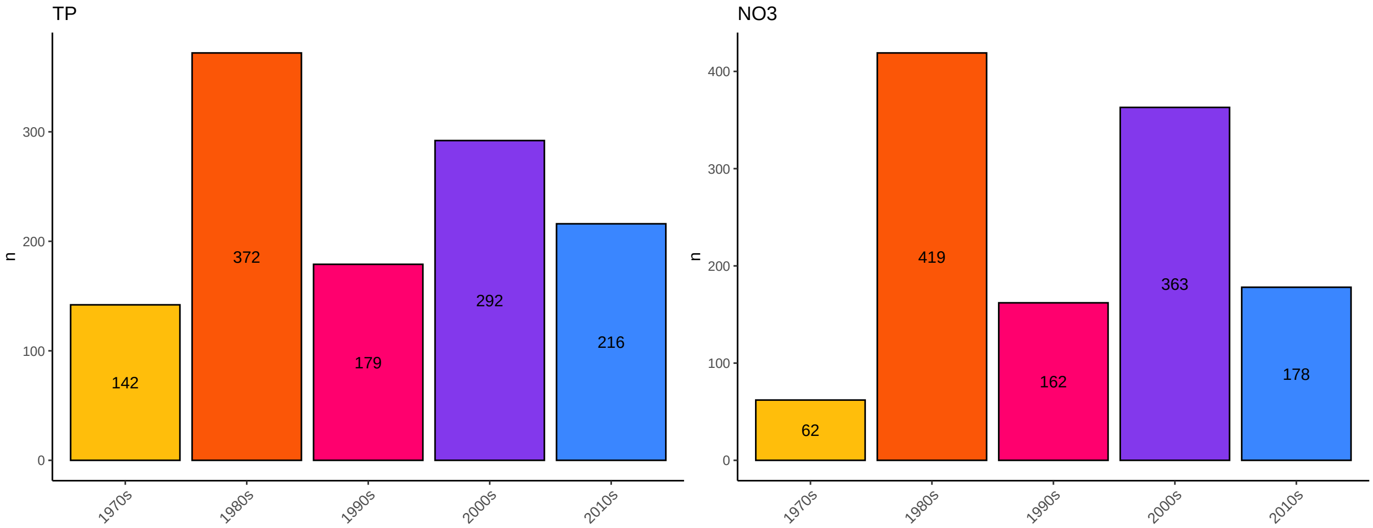
1.1.1.2 Timing of sampling

For the lakes with TP and NO3 data, when were most of the samples obtained? (median of all years sampled, grouped by decade)

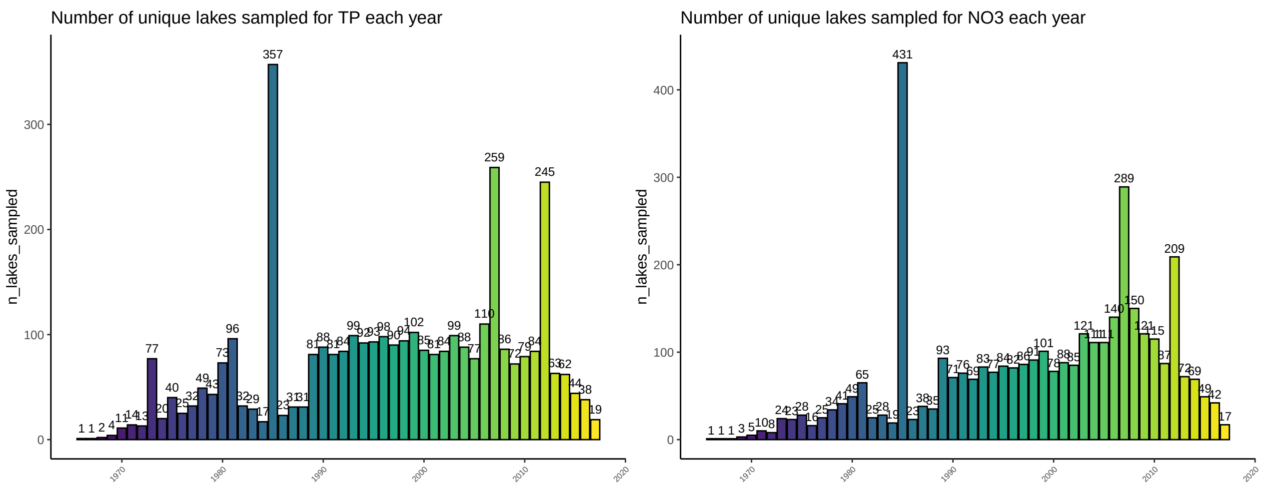




The plots below shows the counts of all the lakes sampled for TP and NO3, grouped by when most of their samples were taken

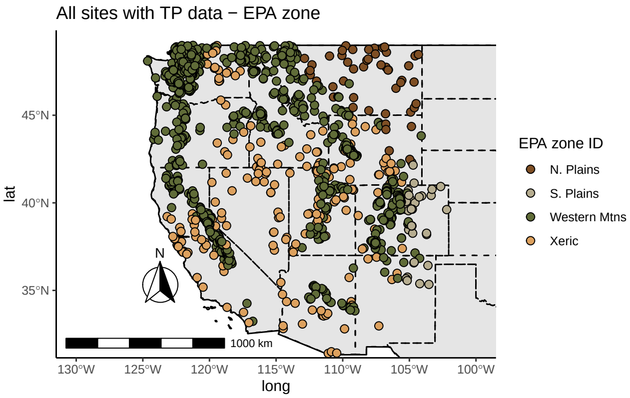


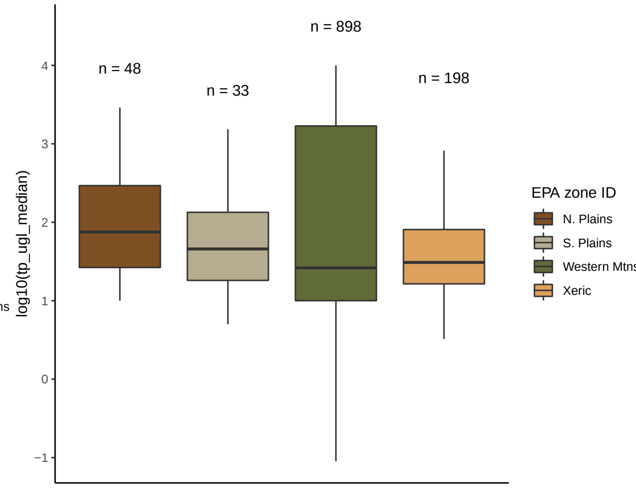
In both TP and NO3, you can see the influence of the Western Lakes Survey in the 80s. Many of the lakes included in the 2000s and 2010s were part of the EPA National Lakes Assessment. Spatially, I notice that Utah had a pretty aggressive sampling campaign in the 1990s. My guess is that for very high elevation lakes in each state, most of those lakes were sampled during the WLS. Another way to visualize this:

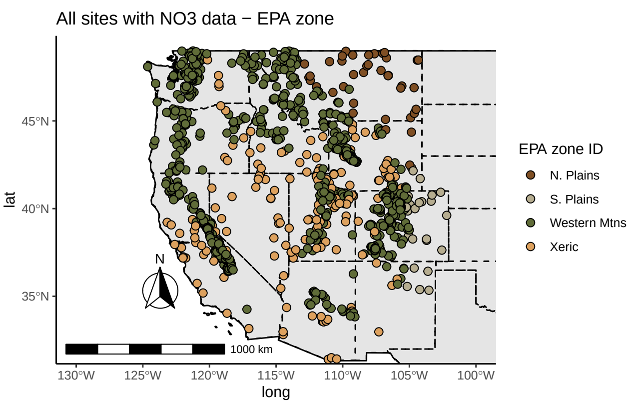


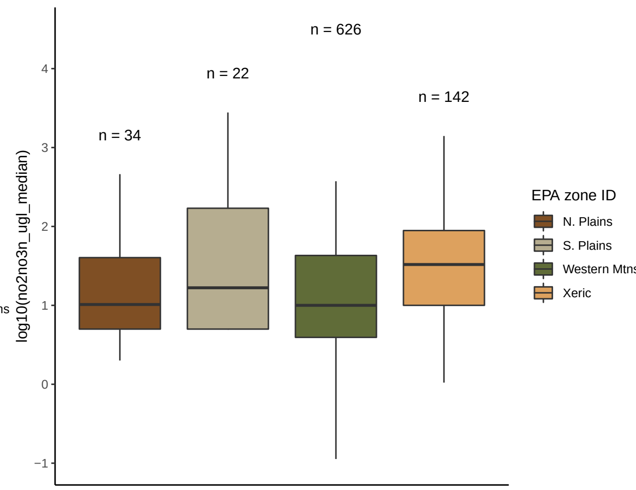
Pretty depressing just how FEW lakes in the western US are sampled every year for TP or NO3, and in recent years those numbers seem to actually be *decreasing*!

1.1.1.3 Observations by EPA nutrient region Let's take a look at how the sites are distributed by nutrient region.







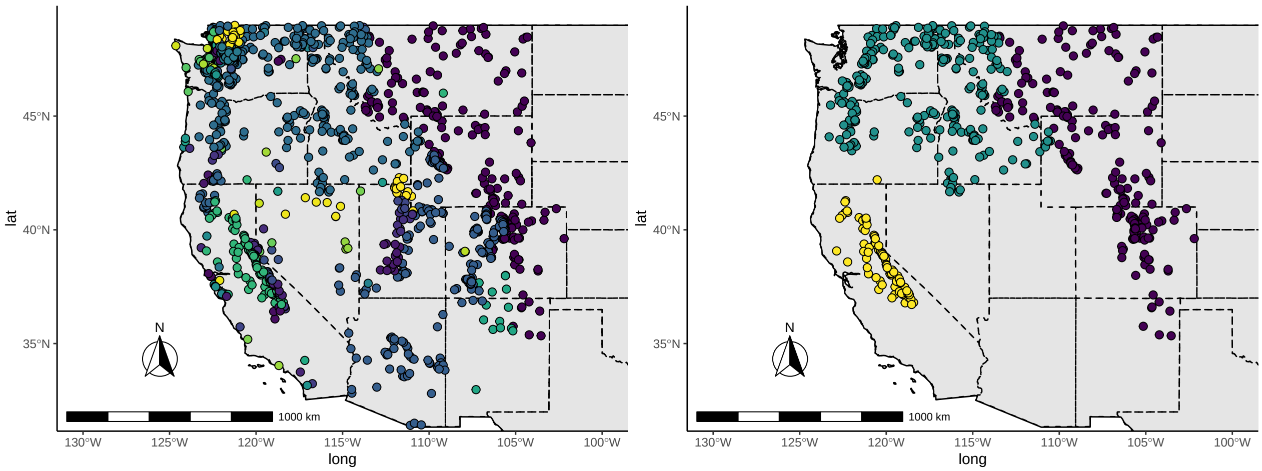


Most of the sitse are in the "western mountains" region. There are LOTS of different ways we can classify these sites, including categories that are included in the LAGOS-US database (like Bailey, WWF, Omernik3, MLRA zone IDs). Eventually we will have to think about a sensible geographic cut off for the sites because currently I just filtered out all the lakes in CA, UT, NV, WA, OR, ID, MT, WY, CO, NM, and AZ.

And of course, beyond categorical variables we can look at how the underlying climate (which will influence things like lake inflows, evaporation, vegetation of surroundings, lake drawdown, etc) influences water chemistry. See section 2.4 later in the document.

1.2 Patterns in western US hydrology/networks

In the entire western U.S. region, there are a total of 208 lake networks, according to the LAGOS-US NETS database. Just narrowing it down to the sites where we have Total P data, those lakes fall into 89 distinct networks.



Most of those lakes fall into 3 river networks (Network 34 = 279 lakes, network 1 = 246 lakes, and network 3 = 231 lakes), which I believe are the Mississipii, Columbia, and Sacramento–San Joaquin River (?) river basins.

We can also look at the lakes by "connectivity class." As a reminder, this is how those classes are defined:

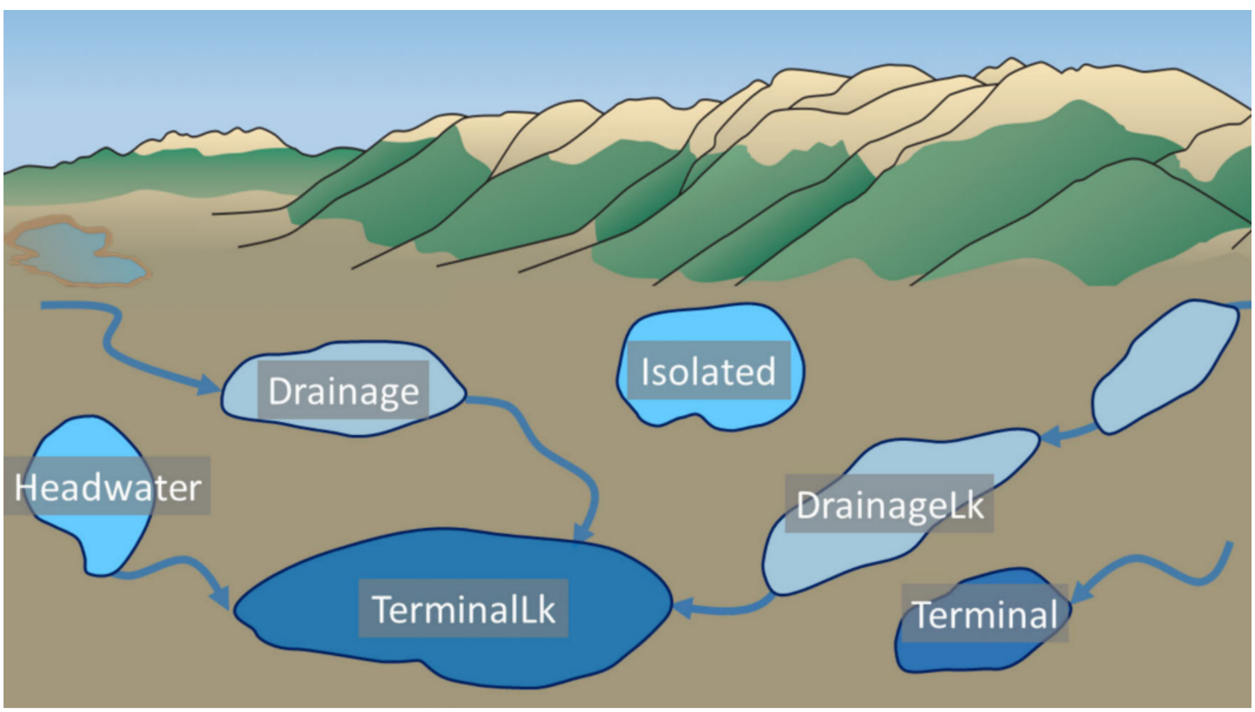
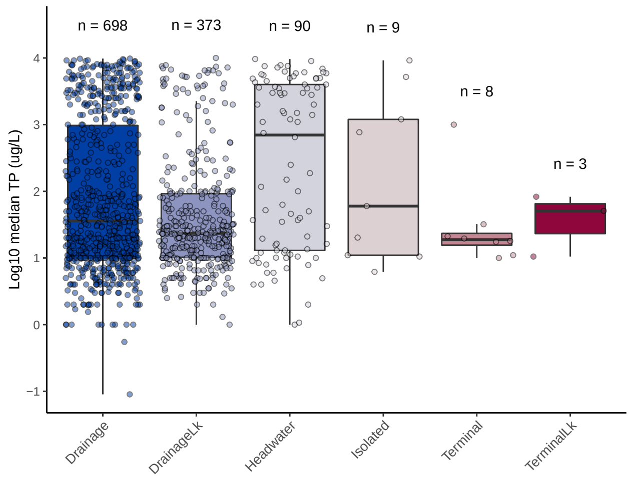
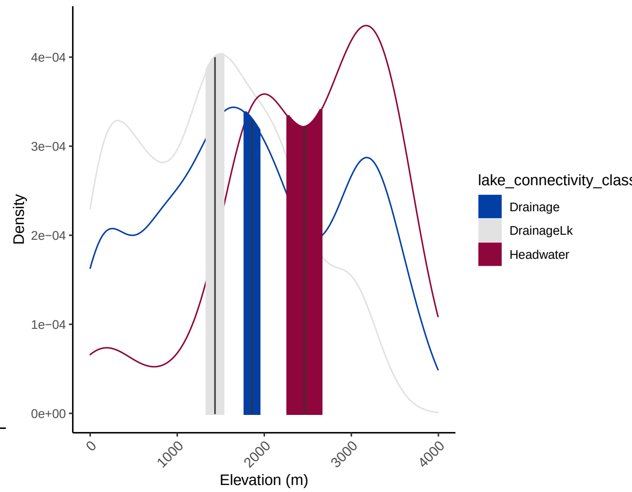


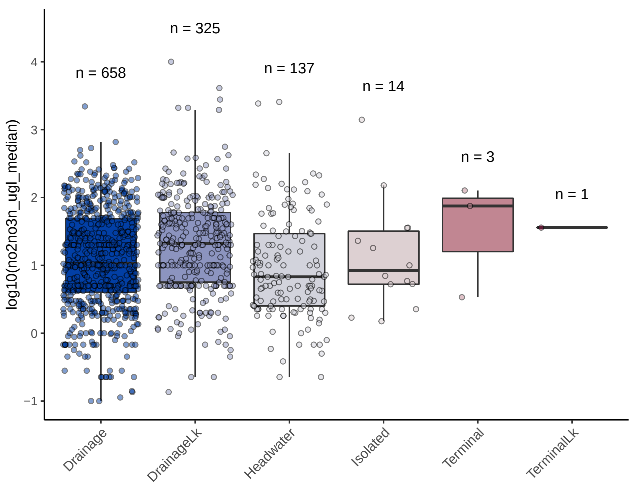
Figure 2: *Visual characterization of lake connectivity class, from Cheruvelil et al. 2021*

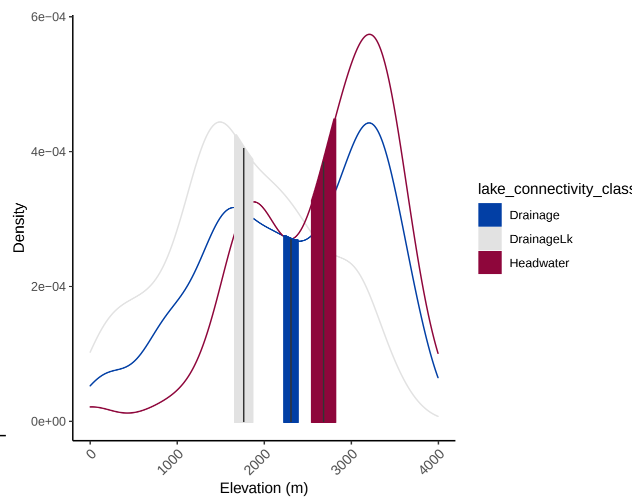
How many lakes fall into each category? And how do TP concentrations vary by lake connectivity class?





Most lakes with TP data in the west are drainage lakes or drainage lakes with a lake upstream of it ("DrainageLk"). About 7.6% of the lakes are headwater lakes with a very small number of isolated or terminal lakes. Headwater lakes tend to have higher TP concentrations than drainage lakes, and drainages lakes with a lake located upstream tend to have slightly lower concentrations of TP on compared to drainage lakes without an upstream lake. Headwater lakes tend to be the highest elevation unsurprisingly, but drainage lakes span basically the entire elevation range in the dataset.However, if you look at the density graphs and the scattered points around each boxplot, you can see that there is some clustering of points. For example, there is a very *wide* spread in the TP values of headwater lakes, and I would venture to guess that the lower TP concentration headwater lakes are located way up in the mountains. We see a similar number of lakes by size class with NO3 data, but as you'll see the spatial and connectivity patterns are different.



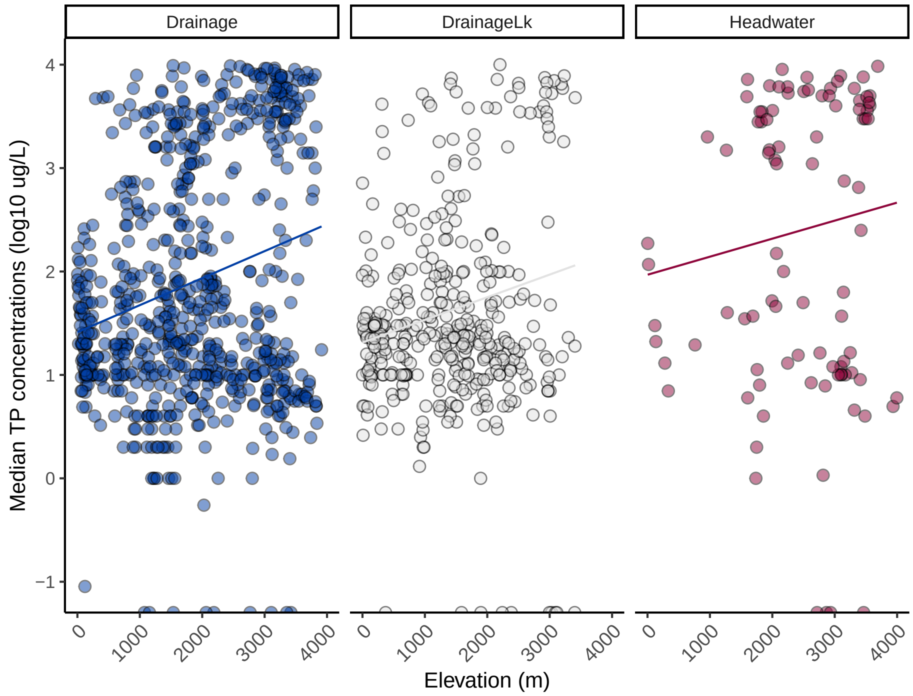


Because most of the lakes are Drainage, DrainageLk, or Headwater, I filtered out the other types for sake of visualization in the rest of the plots

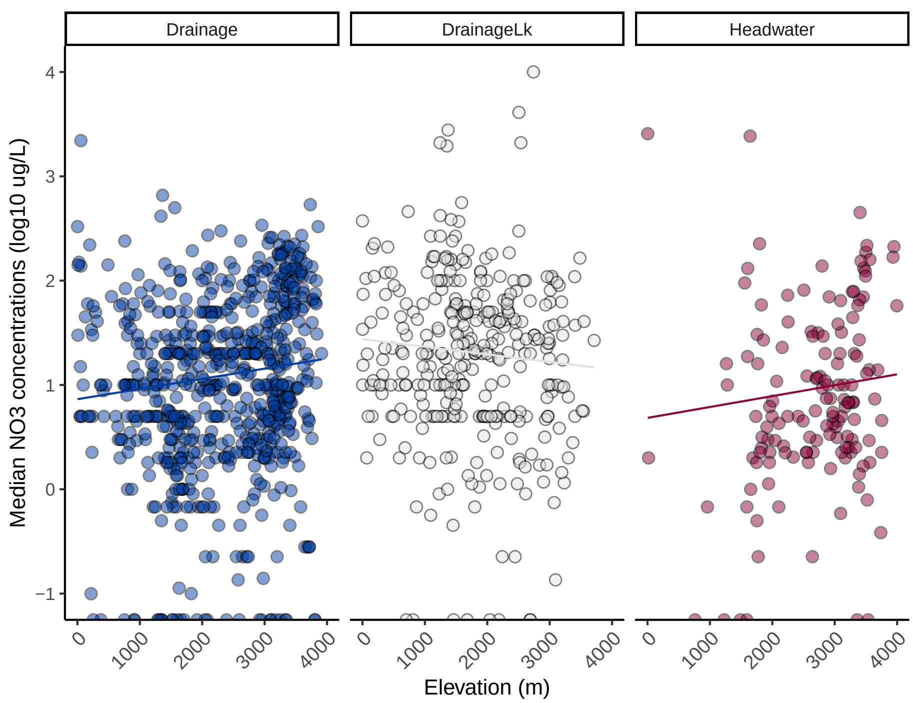
2 Spatial patterns in TP and NO3 concentrations

2.1 Elevation and connectivity classs

I was curious whether or not we see any consistent patterns of TP or NO3 with elevation, and whether those patterns vary depending on the hydrologic connectivity of the waterbody.



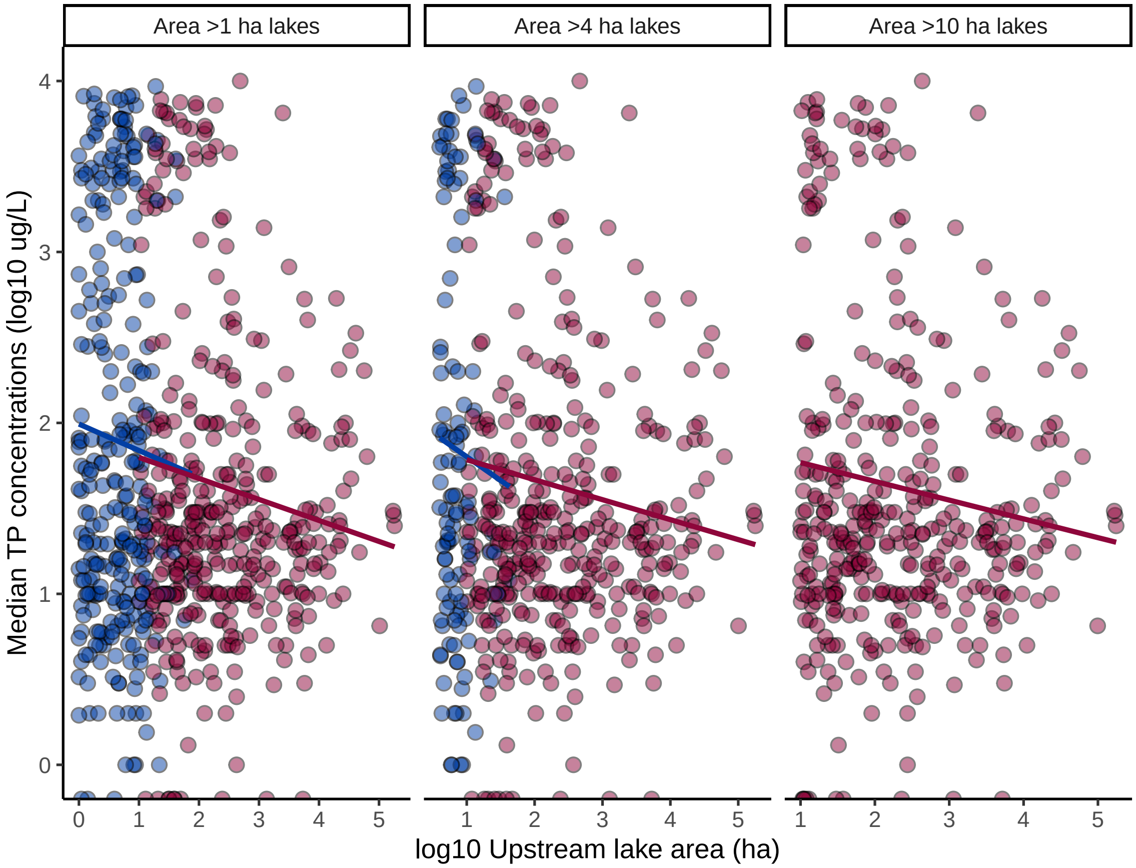
TP concentrations increase as you move up in elevation in the three most abundant connectivity classes



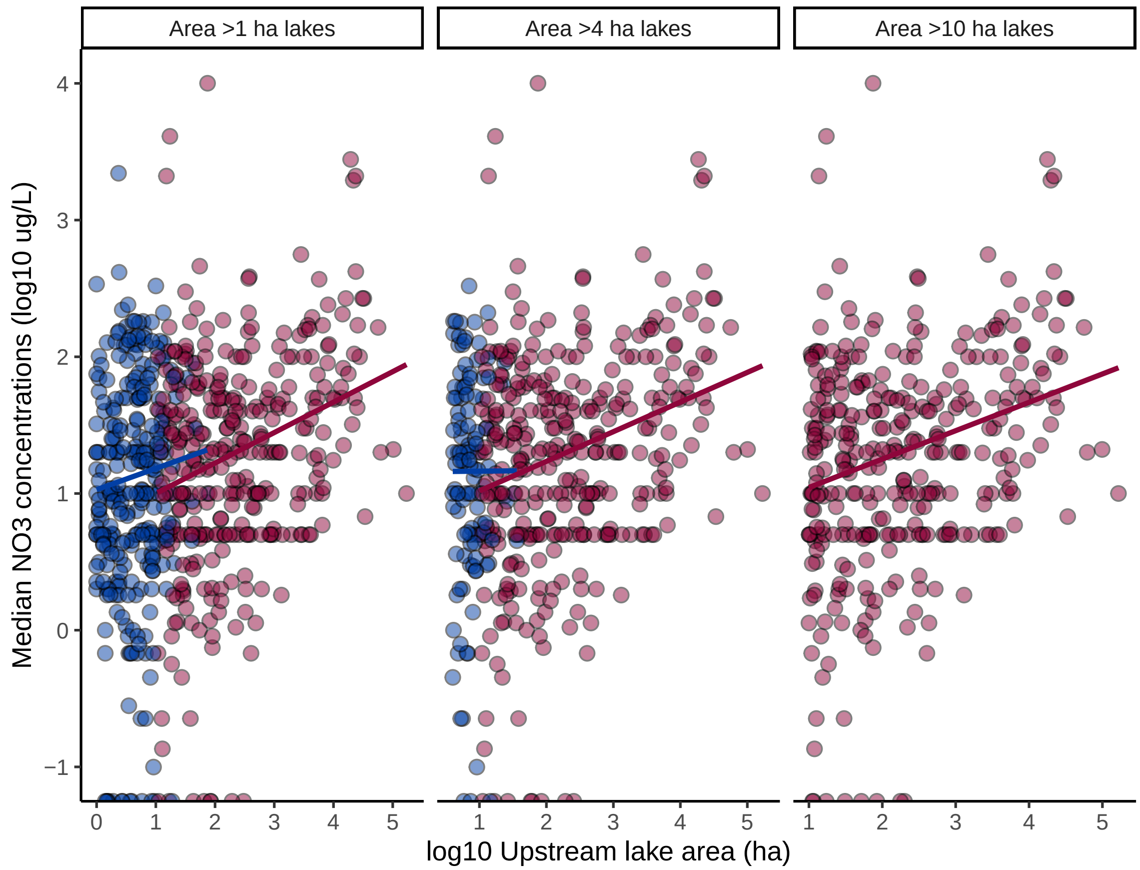
NO3 concentrations, on the other hand, do not show clear elevational patterns like TP does.

2.2 Upstream waterbodies

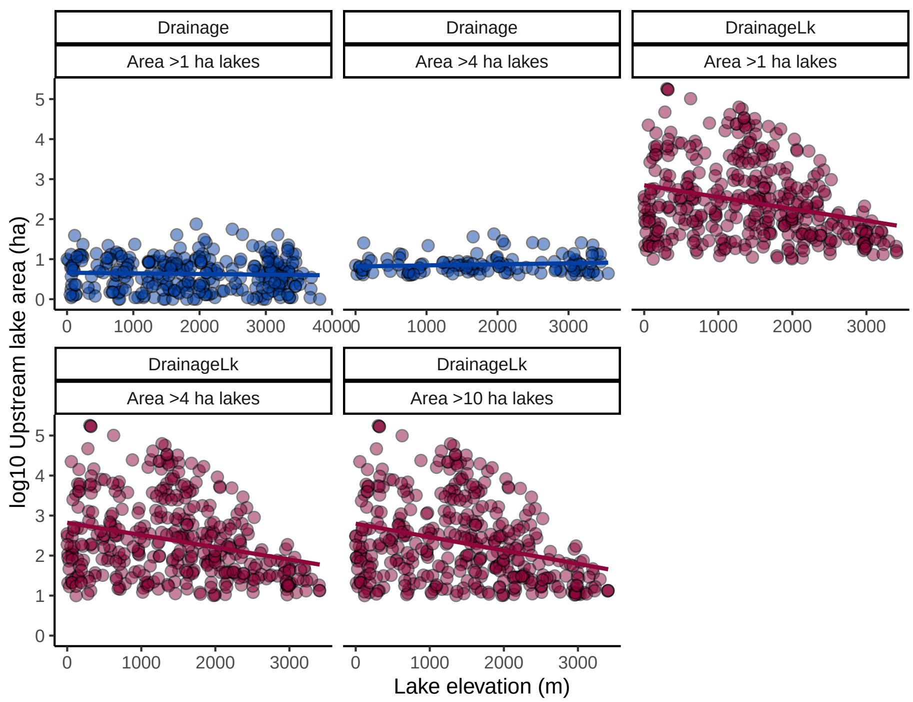
Another feature of LAGOS both a count and total area of upstream waterbodies. Do upstream waterbodies impact TP and NO3 concentrations?

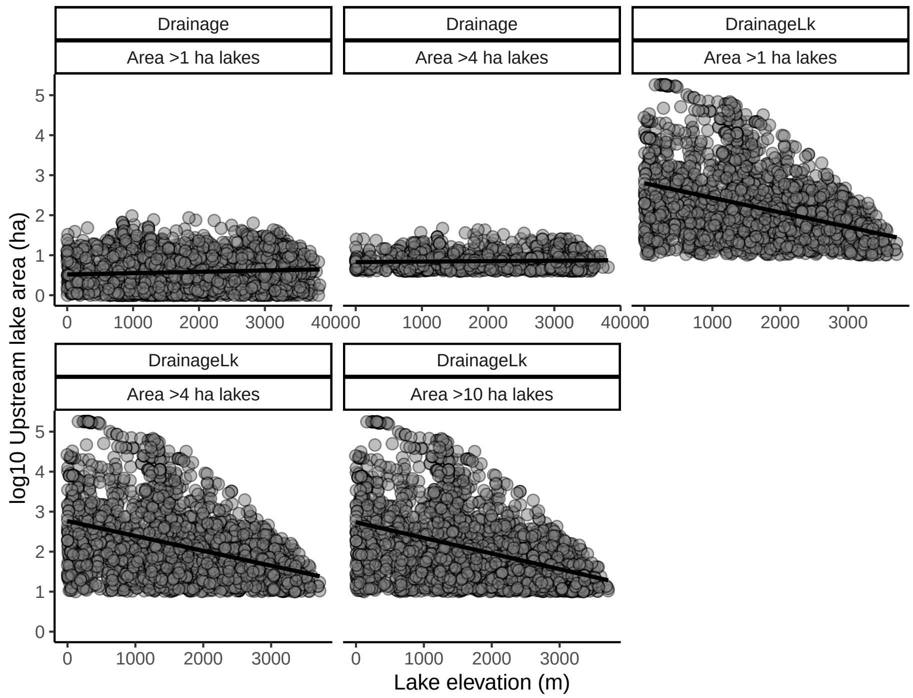


Fairly intuitive, but the greater the upstream lake area, the lower the TP concentration of the lake.



In DrainageLks, NO3 concentrations increase with great upstream lake area. The same is not true for Drainage.





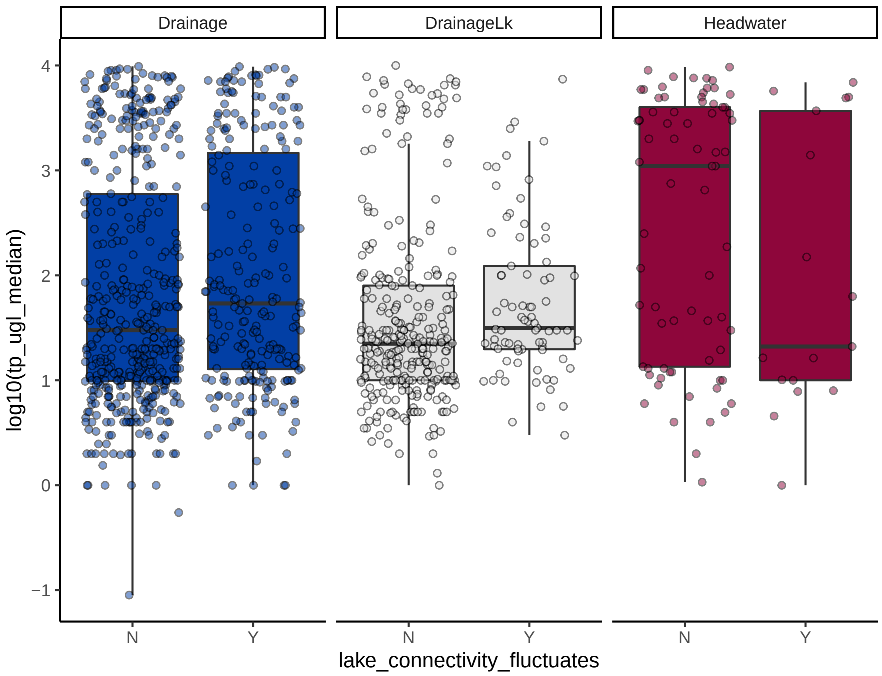
The above plots show patterns of upstream lake area by elevation in lakes with TP data (colored plots) and the greater population of lakes (b&w plots).

So in summary, the general patterns are:

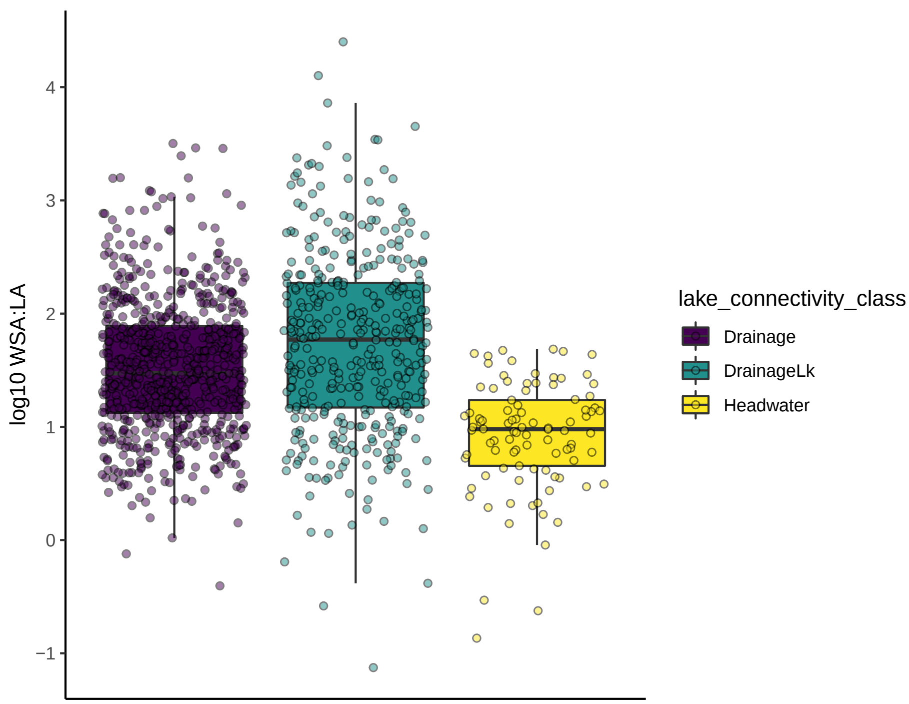
- TP decreases with increasing upstream lake area
- TP increases with increasing lake elevation
- In DrainageLks (with upstream lakes >10 ha) you see decreasing upstream lake area with increasing elevation.
  - That's pretty intuitive to me, but interesting that there is *no* such pattern in Drainage lakes (no lake upstream lakes)
- N03 shows the opposite pattern as TP, where concentrations **increase** with increasing upstream lake area
- NO3 doesn't show a distinct pattern with elevation, unlike TP.

2.3 Hydrology

Whether or not lake levels fluctuate may be indicative of whether or not the lake is "natural" or manmade. Eventually LAGOS will publish a dataset that confirms this, but for now it may serve as a proxy. At first glance it would appear that in drainage lakes, lake level flucuations are associated with higher TP, though not dramatically so.

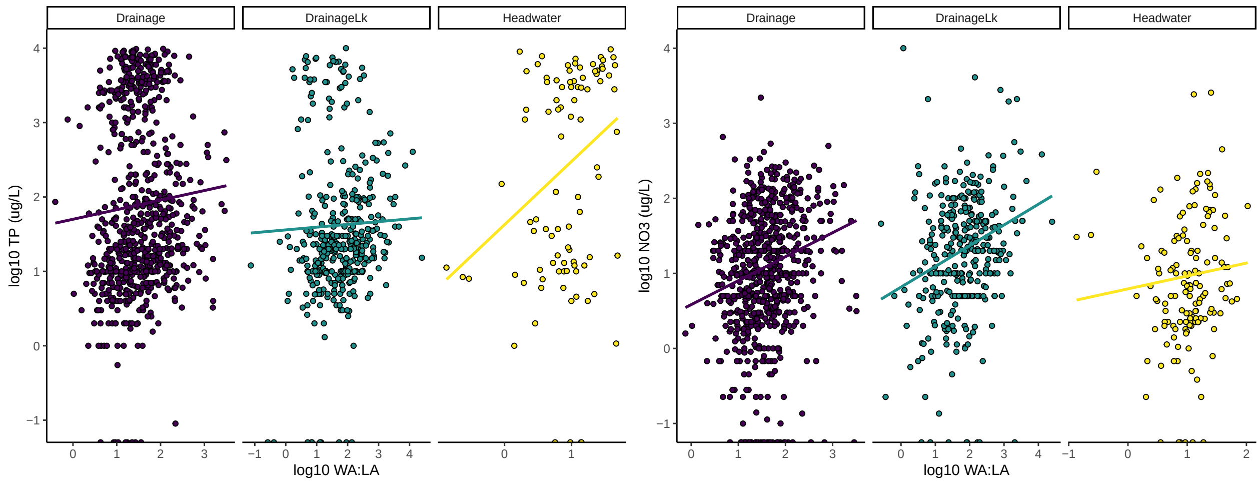


Watershed area to lake area is a decent proxy for hydrologic residence time. How does that vary with connectivity class?



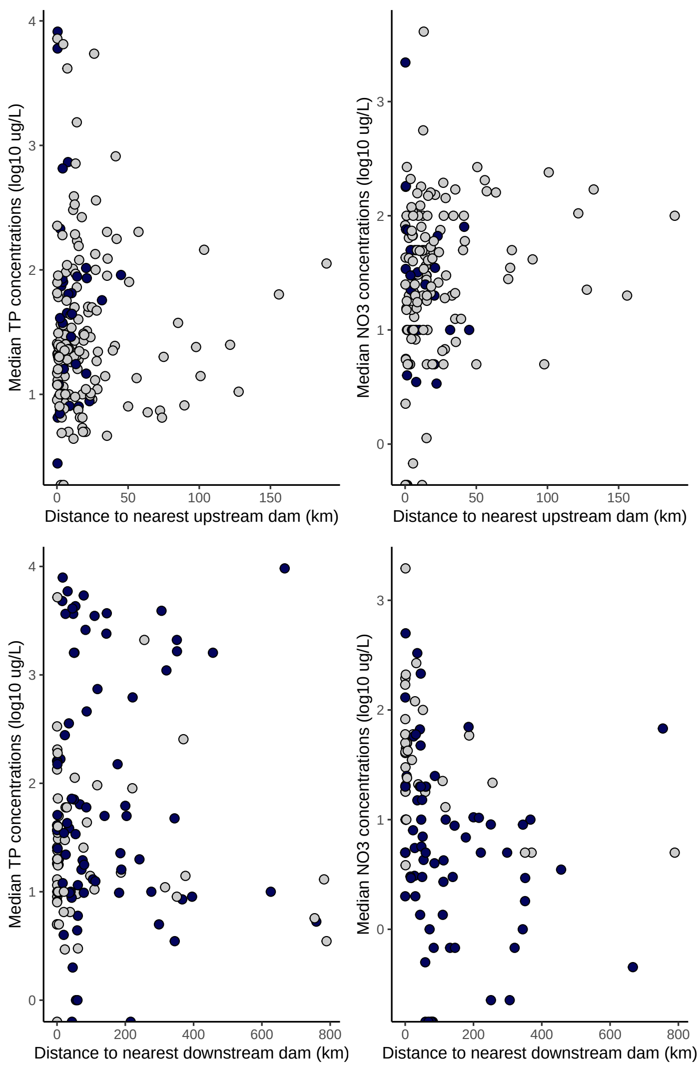
This also makes sense. We would expect Headwater lakes to generally have smaller WSA:LA and longer residence times. It also makes sense that DrainageLk's would have generally higher WSA:LA than Drainage.

Is there a correlation between TP or NO3 and WSA:LA?



Hard to say with any certainty for TP, but there appars to be a fairly strong relationship of increasing NO3 concentration with increasing WSA:LA. I wouldn't put to much stock in the patterns we see with the Headwater lakes just because the range of WSA:LA values is so low.

2.3.1 Dams

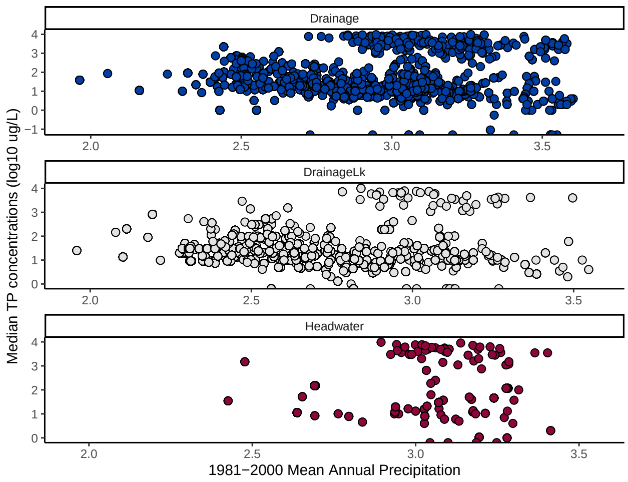


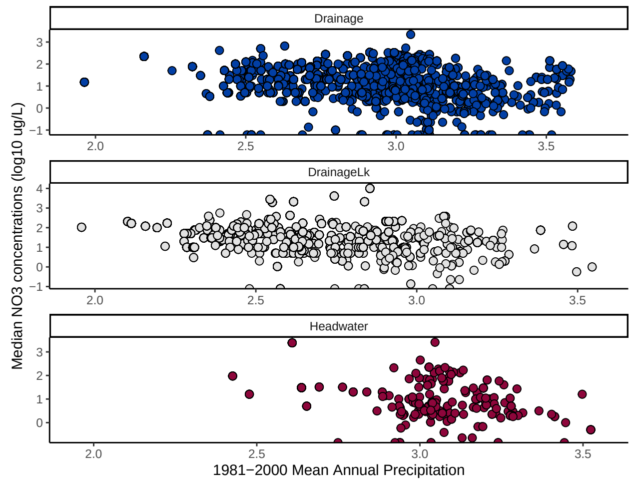
There don't appear to be an obvious relationship between TP nor NO3 concentrations and distance to the nearest upstream nor downstream dam, but there is probably a better way to look at this.

2.4 Climate

2.4.1 Mean annual precipitation

At first glance it doesn't appear to be a strong relationship between TP and 1981-2000 mean precipitation, but it would be interesting to eventually look at finer-scales such as winter precipitation or summer temperatures. For NO3, on the other hand, if you squint it does seem like wetter areas tend to high lower NO3 concentrations on average.





2.4.2 Mean annual temperature

Similar story with temperature and TP… not seeing any immediate patterns popping out. But with NO3, in some areas or lake-types, you see higher concentrations in warmer climates.