

## Predictor summary table

| Variable                     | Median | Q25  | Q75  |
|------------------------------|--------|------|------|
| Max depth (m)                | 9      | 6    | 14   |
| Watershed-lake ratio         | 15     | 6    | 34   |
| Precipitation (mm/yr)        | 910    | 830  | 1000 |
| Baseflow                     | 49     | 33   | 62   |
| Fertilizer N (kg/ha)         | 55     | 33   | 90   |
| Fertilizer P (kg/ha)         | 10     | 6    | 16   |
| Manure N (kg/ha)             | 27     | 17   | 44   |
| Manure P (kg/ha)             | 7      | 5    | 12   |
| N deposition (kg/ha)         | 6      | 5    | 7    |
| Wetland potential (percent)  | 15     | 5    | 26   |
| Soil organic carbon (g C/m2) | 4000   | 2900 | 5300 |
| Buffer Ag (percent)          | 25     | 11   | 48   |
| Buffer Natural (percent)     | 41     | 23   | 59   |
| Ag (percent)                 | 42     | 25   | 63   |
| Forest (percent)             | 25     | 12   | 46   |
| Pasture (percent)            | 14     | 7    | 24   |
| Corn (percent)               | 7      | 2    | 17   |
| Soybeans (percent)           | 5      | 1    | 14   |
| Wetlands (percent)           | 3      | 1    | 8    |
| Clay (percent)               | 10     | 5    | 17   |

## Percent agriculture versus lake TP and TN

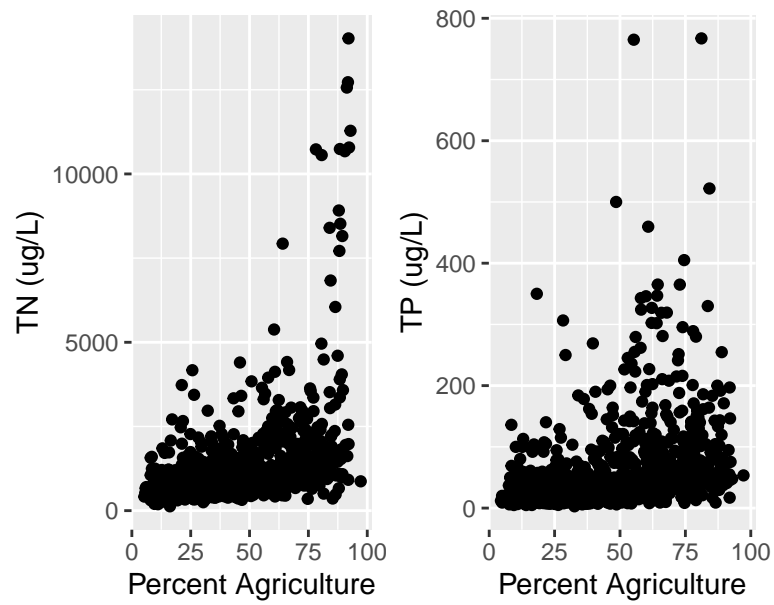


Figure 1: Lake nutrient concentrations versus percent watershed agriculture.

## Response variable maps

Some obvious data availability differences in New York and Wisconsin. Wisconsin has quite a bit fewer TN measurements.

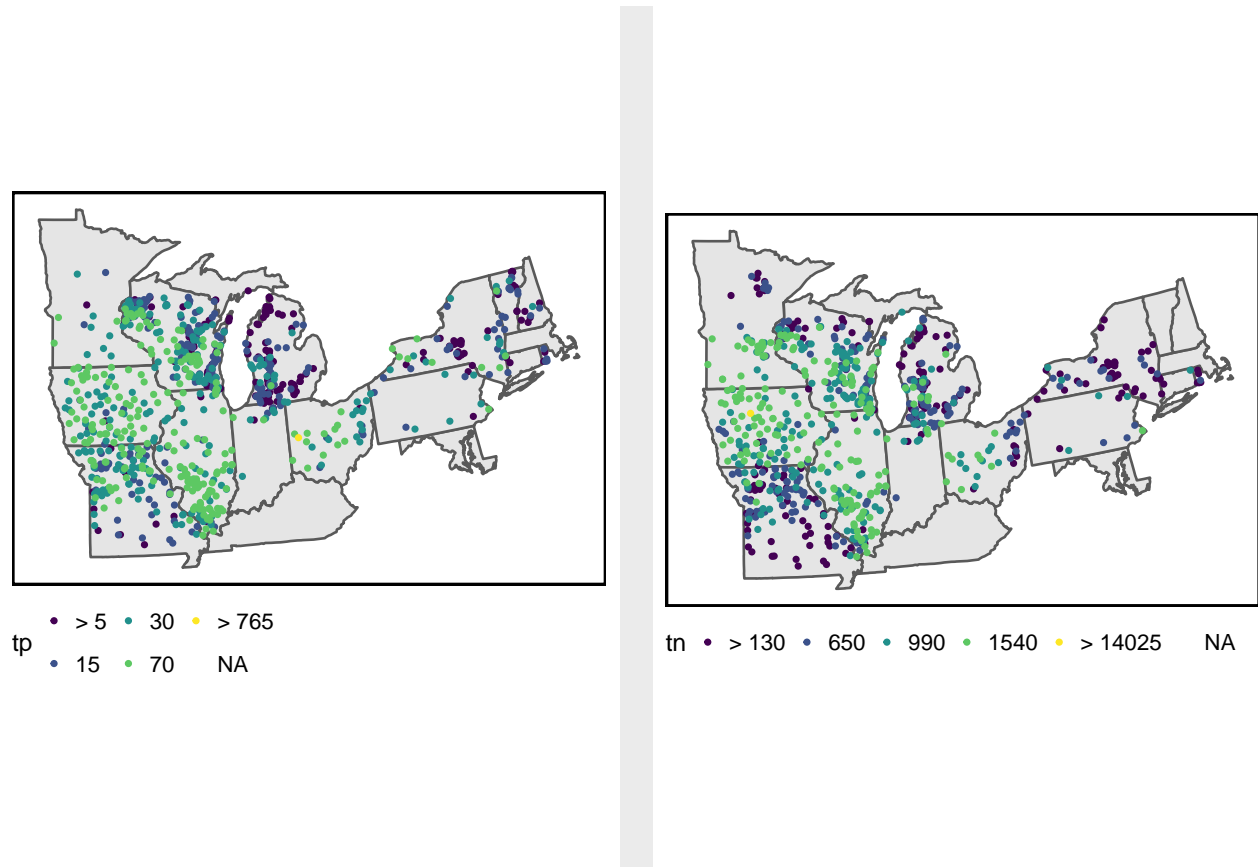


Figure 2: Quantile map of lake nutrients.

## Trophic state

Over 35 percent of our study lakes were eutrophic to hypereutrophic.

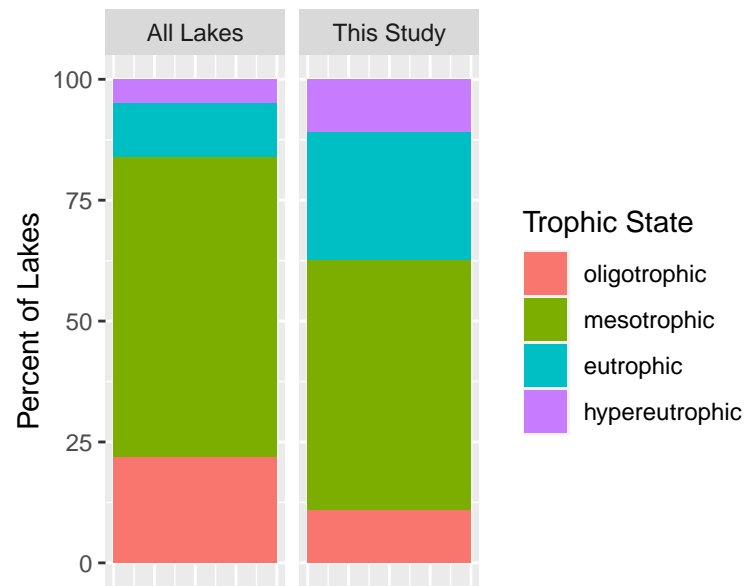


Figure 3: Lake trophic state in our study lakes versus all lakes in the study extent.

## Model results when lake are restricted to $> 40\%$ watershed agriculture

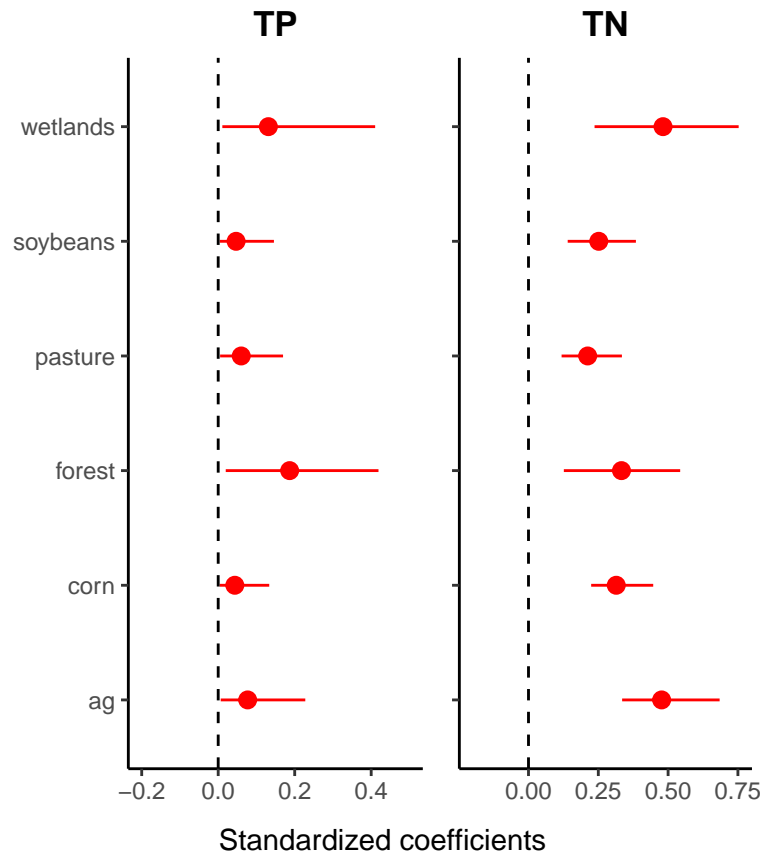


Figure 4: Population level effect of watershed land-use cover on lake TP and TN from six candidate models for lakes with greater than 40% watershed agricultural land-use. Values shown are posterior medians (filled circles) and 95% credible intervals (solid lines). Also shown is a comparison to a zero effect (dashed line).

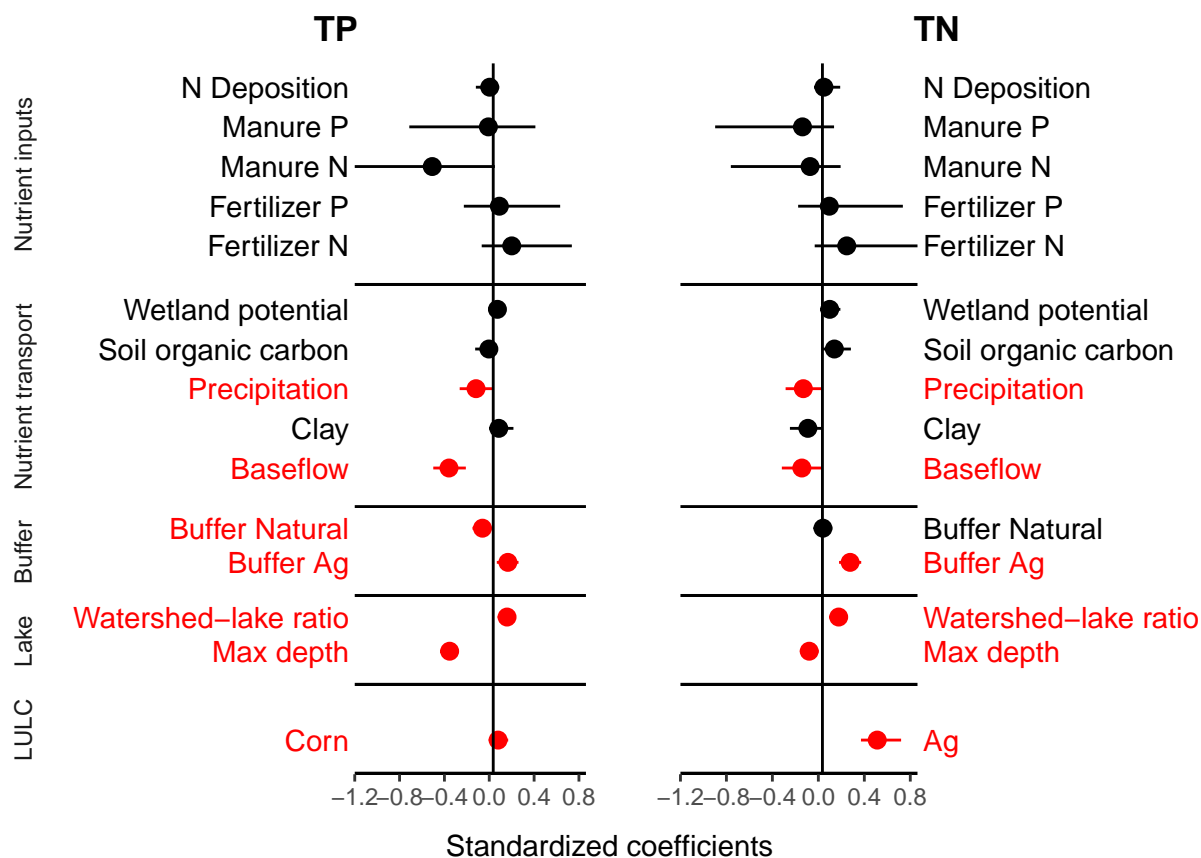
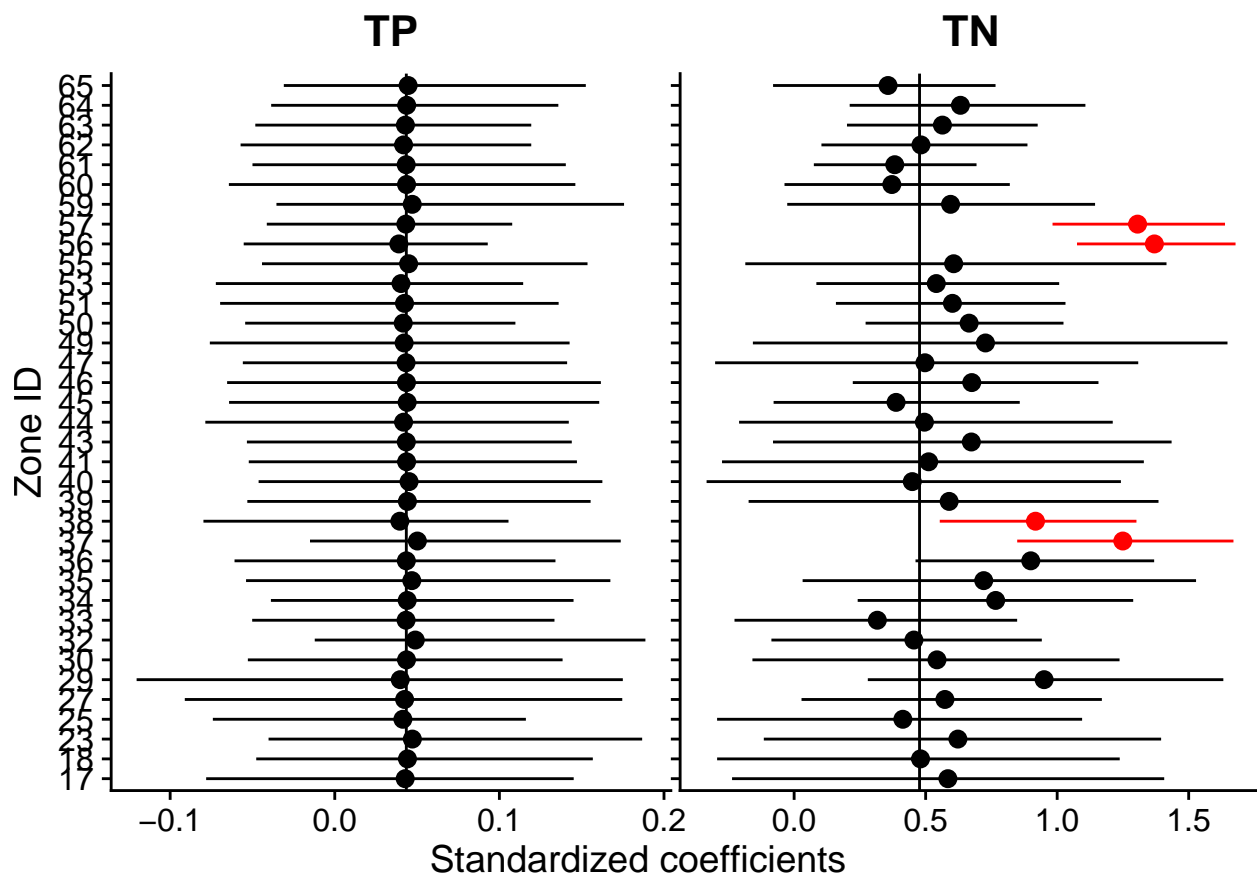


Figure 5: Global (fixed effect) coefficient values and credible intervals for best-fit lake TP and TN models for lakes with greater than 40% watershed agricultural land-use. Values that do not overlap zero are shaded in red. Horizontal bars separate coefficients in distinct predictor categories. Coefficient estimates are reported relative to standardized predictor variables centered at zero with unit variance.



## Model results when land-use predictors are excluded

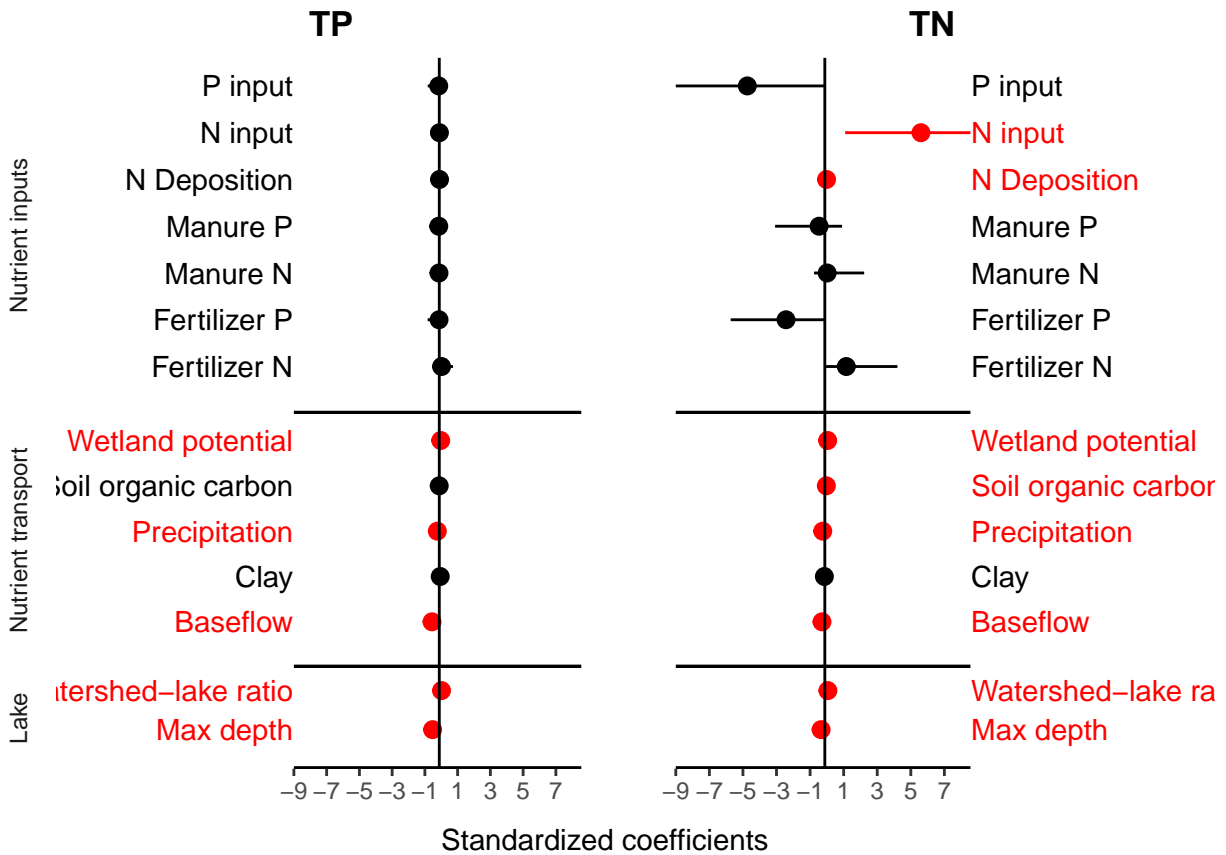


Figure 6: Global (fixed effect) coefficient values and credible intervals for best-fit lake TP and TN models without a land-use predictor. Values that do not overlap zero are shaded in red. Horizontal bars separate coefficients in distinct predictor categories. Coefficient estimates are reported relative to standardized predictor variables centered at zero with unit variance.



## Bivariate relationships among predictors and lake nutrients

We examined bivariate relationships between lake nutrient concentrations and predictors to determine, at least to first-approximation, the likely drivers of lake nutrients (Figure 2). We found large differences between the drivers of lake phosphorus and lake nitrogen concentrations. For example, max depth and baseflow had a much larger effect on P than N. Pearson’s correlation coefficients for these relationships were XX and YY respectively. In contrast, watershed land-use corn cover, fertilizer inputs, and riparian buffer configuration appear to have a larger effect on N than P Pearson’s correlation coefficients for these relationships were XX, YY, and ZZ respectively.

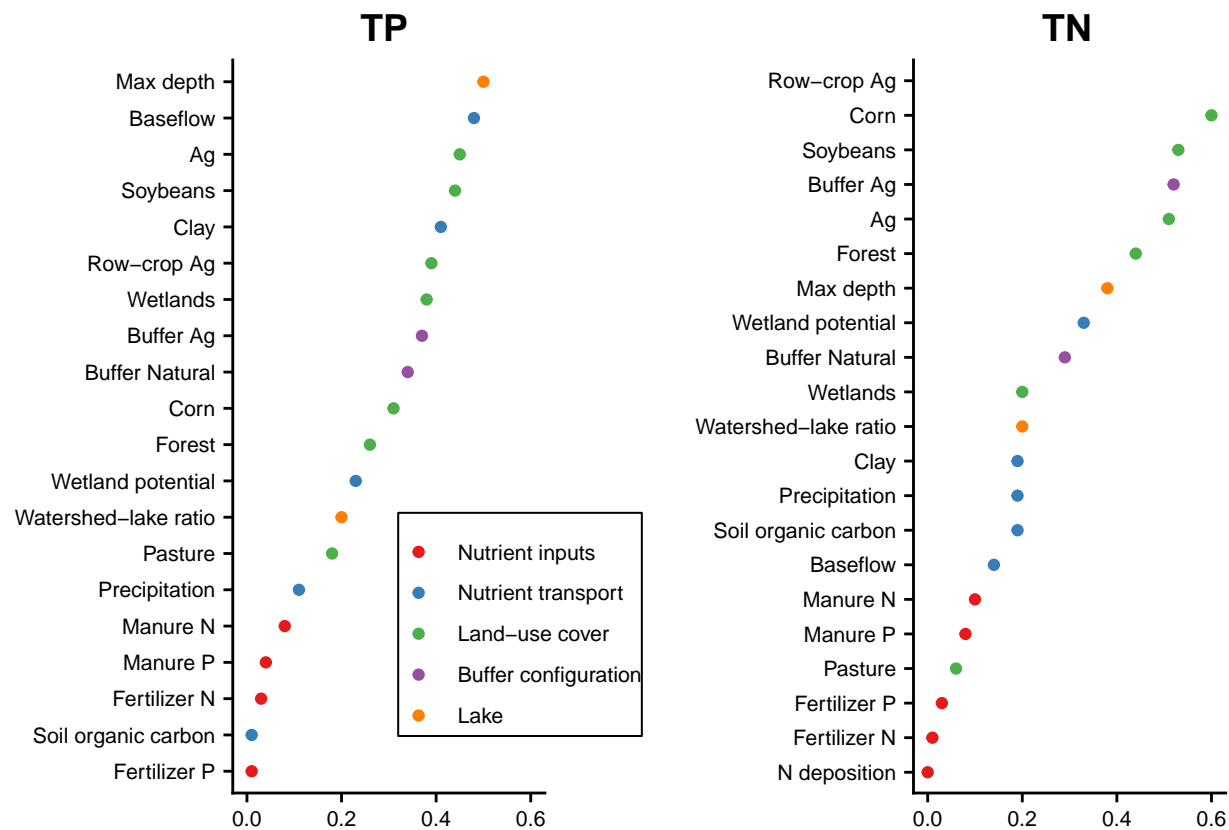


Figure 7: Dotplots showing absolute value Pearson’s correlation coefficients. Dots are colored according to predictor categories.

## Correlations among predictor variables

There is clearly a cluster of related predictor variables that includes most nutrient input, Ag land-use, and buffer configuration measures.

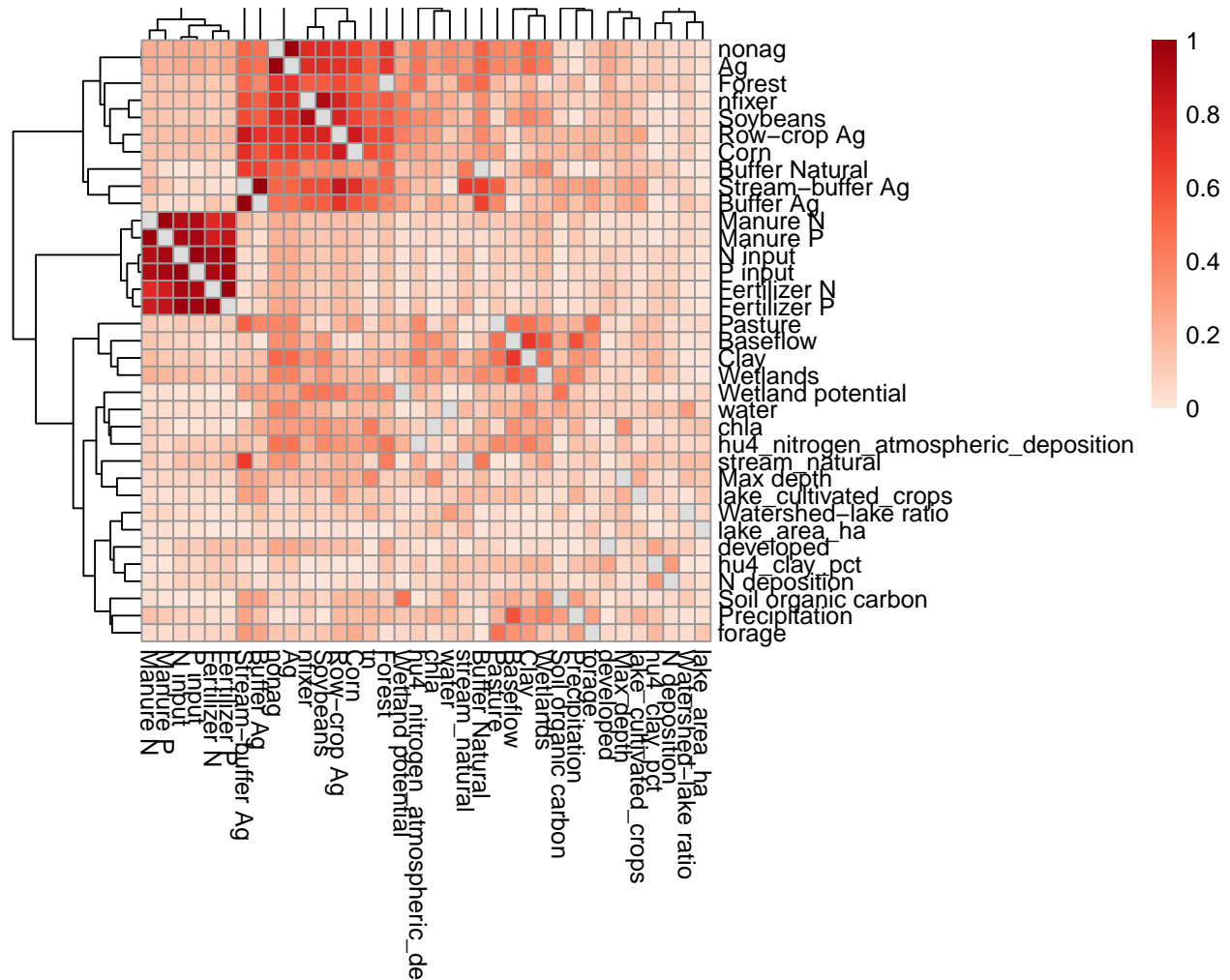


Figure 8: Heatmap showing absolute value correlation coefficients.

## The Cropland Data Layer

The National land cover database (NLCD) is more typically used in macroecology studies. In our paper, we wanted to go beyond the coarse categories in the NLCD and test relationships with individual crops. The Cropland Data Layer (CDL) is a product that allowed us to do this. Because the CDL is more granular than the NLCD we could not directly compare them. Instead, we aggregated all the CDL categories associated with agriculture (also pasture) and compared these values for each watershed with the corresponding NLCD estimates.

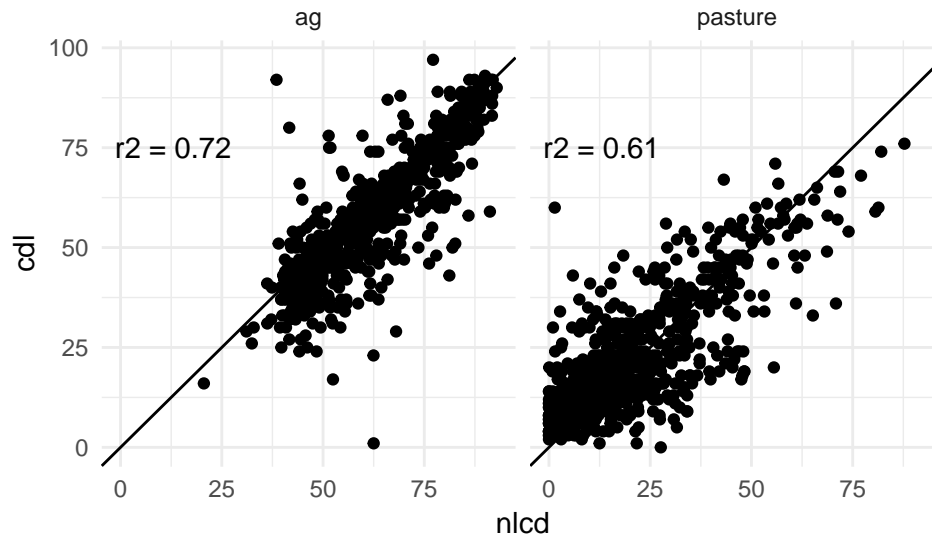


Figure 9: Comparison between NLCD (2011) derived and CDL (2010) derived total watershed ag and watershed pasture cover. Line is 1:1 not a regression best fit.

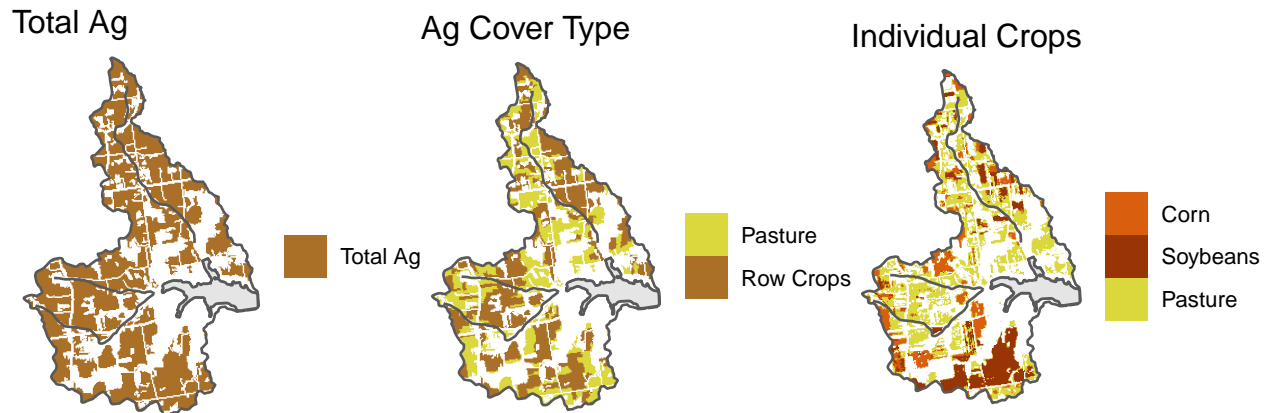


Figure 10: Example of increasing granularity for total Ag to Ag versus pasture, to pasture versus specific crops. For illustration in this example, only corn, soybeans, and pasture are shown. Rather than all CDL land-use categories.

## Demo of watersheds and buffer configuration

Here are two lakes each with the same percent Ag in their watersheds (red outlines). However, the lake on the left (Lake Carleton) has a riparian buffer (teal outlines) with a much higher percentage of Ag relative to the lake on the right (Argyle Lake).

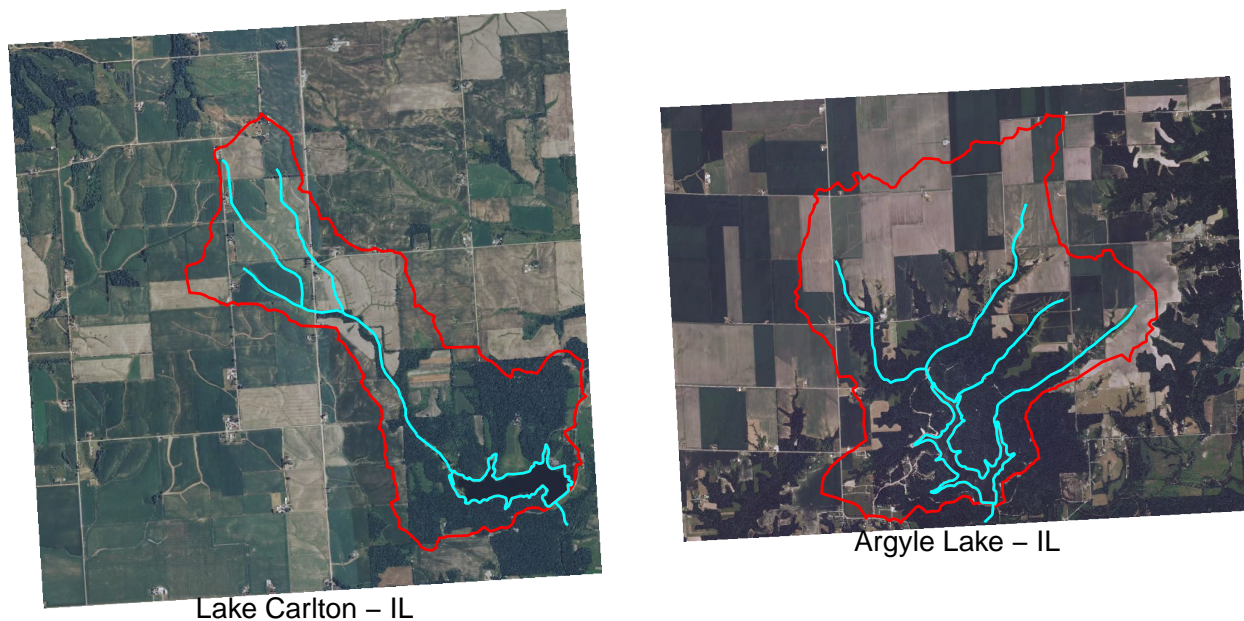


Figure 11: Stream buffer Ag compared to IWS Ag.