Three-Dimensional Map of Nitrate Concentration in Groundwater in Wisconsin With Spatially Adjusted Random Forests

Xindi Lin‡, Christopher Zahasky†, Hyunseung Kang‡

†Department of Geoscience, University of Wisconsin--Madison, ‡ Department of Statistics, University of Wisconsin--Madison

Overview

- Nitrate remains the most widespread groundwater contaminant in Wisconsin.
- Millions of dollars are spent to meet the 10mg/L public health standard.
- The project aims to create ***a three-dimensional map of nitrate concentration across Wisconsin.***
- Our approach involves:
  - Collect publicly available **groundwater nitrate measurements from 2014 to 2024** in Wisconsin.
  - Collect **environmental predictive variables** that are hypothesized to predict nitrate contamination, including land use, precipitation, soil drainage, concentrated animal feeding operations (CAFOs), static water levels, and well depth.
  - Use **Random Forest** to assess the influence of the predictive variables on nitrate contamination and apply **kriging** to capture the spatial dependency of nitrate levels in groundwater.

Model Validation

- Use 80% of the data (Total:76604 measurement) as the **training set,** and 20% of the data (Total:19151 measurement) as the **validation set**.
- **Root median squared test prediction error**: 0.515 mg/L.

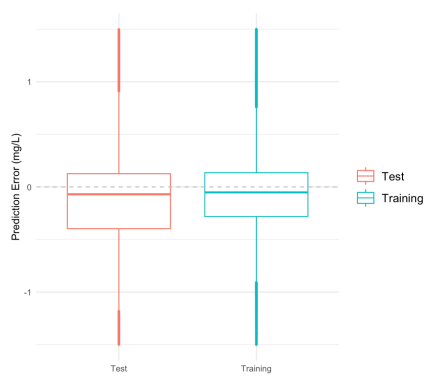


Figure: Prediction error on training set v.s. test set.

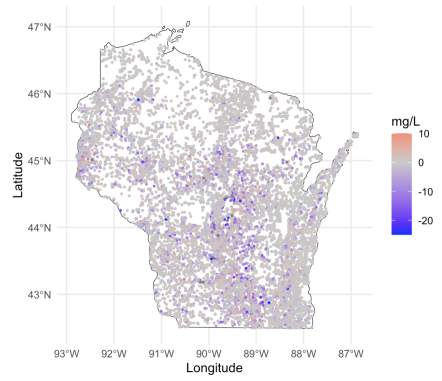
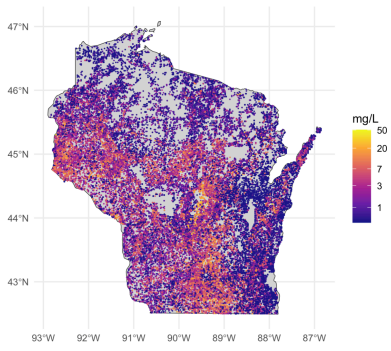


Figure: Map of test prediction errors.

Data Visualization

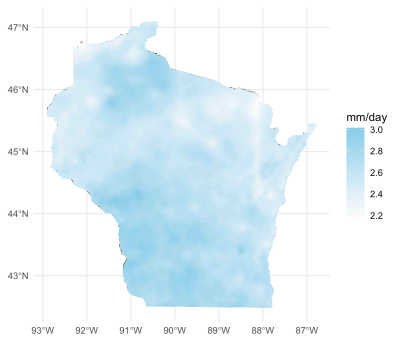
Nitrate

- **Nitrate measurements**[1] taken between 2014 and 2024 in Wisconsin.
- Data is measured and integrated based on the **Public Land Survey System (PLSS)** grid system in Wisconsin.



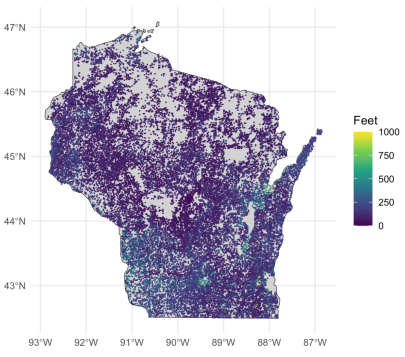
Precipitation

- **Daily average precipitation**[3] at each site from 2014 to 2024 (Total: 18,880 sites)
- **Bilinear interpolation** of the average precipitation.



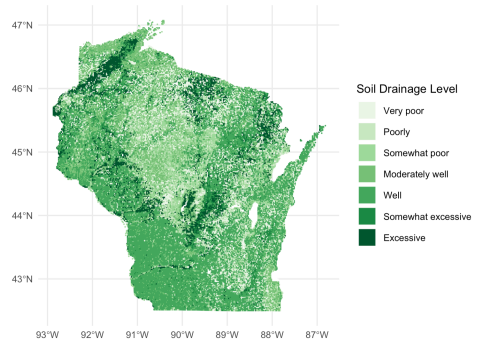
Well Depth

- **Well depth**[1] in Wisconsin wells.
- $\text{Well Depth} = \text{Casing Depth} + \frac{\text{Well Bottom} - \text{Casing Depth}}{2}$



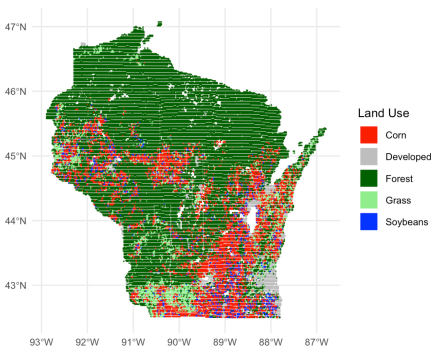
Soil Drainage

- **Seven levels** of soil drainage[4].



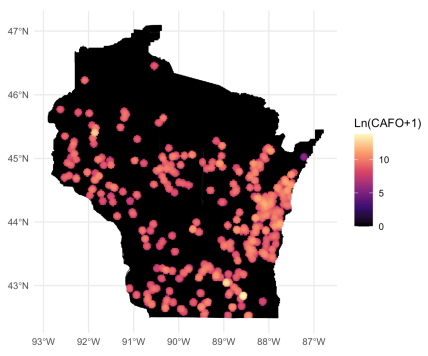
Land Cover

- **Land cover**[2] in Wisconsin.
- Original categories are combined into **Grass, Forest, Corn, Soybeans**, and **Developed** urban areas.



CAFO

- Concentrated animal feeding operations[5]
- $\text{CAFO} = \text{Scaled factor} \times \text{Animal Unit} \times \left(1 - \left(\frac{\text{Distance}}{\text{Max distance}}\right)^2\right)$
- The **maximum distance** is 10 kilometers.
- **Scaled factor**[6] for animals:
  - Source: USDA report in references.
  - Dairy (5.1), turkey (18.0), swine (4.0), sheep (1.1), beef (4.4), chicken (17.1), duck (9.8)



Spatially Adjusted Random Forest

Statistical Model

Let Y be the nitrate concentration, X be the predictive variables, s be the locations. We assume the following statistical model

- $\log(Y) = f(X) + U(s) + \epsilon$.
- $f(X)$ models the effect of the predictive variables.
- $U(s)$ models the spatial dependence of nitrates.
- ϵ models measurement error.

Estimation

- Estimate $f(X)$ by **Random Forest**:
  - Captures higher-order interactions of predictive variables.
  - Cross-validation to select hyper-parameters and to prevent overfitting.
- Estimate $U(s)$ by **kriging**:
  - Interpolate $U(s)$ at unobserved locations with Gaussian process regression.

Three-Dimensional Map of Nitrate Concentration

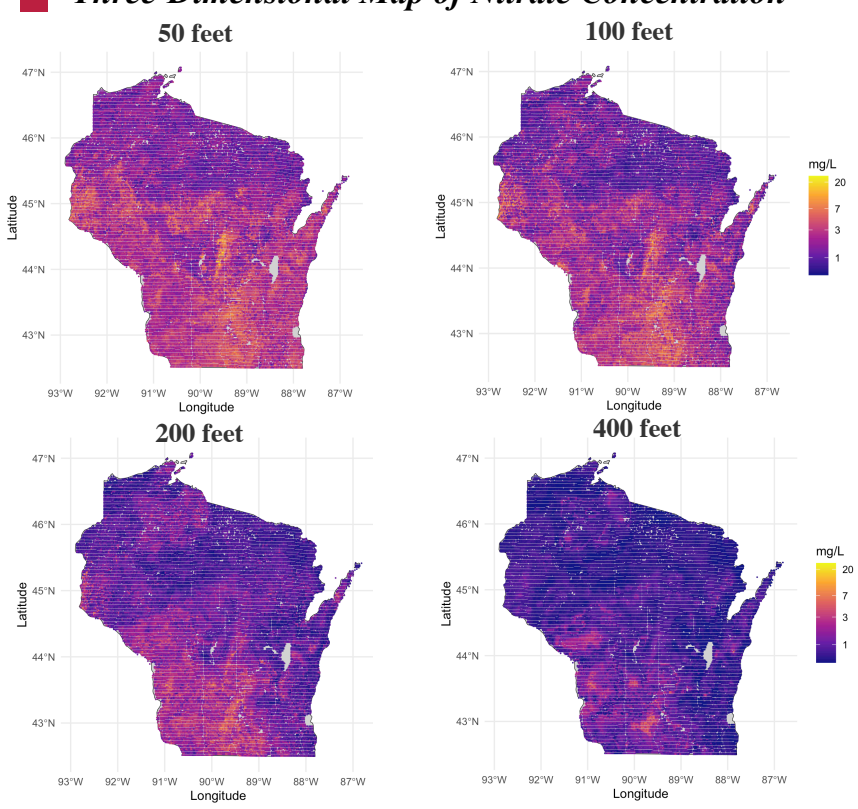


Figure: Top Left: nitrate map at 50 feet depth. Top Right: nitrate map at 100 feet depth. Bottom Left: nitrate map at 200 feet depth. Bottom Right: nitrate map at 400 feet depth.

Acknowledgement and References

- References: [1] Groundwater Retrieval Network (GRN). [2] Cropland Data Layer (CDL). [3] United States Geological Survey (USGS). [4] Soil Survey Geographic Database (SSURGO). [5] Wisconsin Pollutant Discharge Elimination System (WPDES). [6] Estimates of Recoverable and Non-Recoverable Manure Nutrients Based on the Census of Agriculture, USDA.
- Funding: Advanced statistical models for large datasets of pesticide and emerging contaminants in groundwater, Wisconsin Groundwater Coordinating Council, Department of Agriculture, Trade, and Consumer Protection.