GEOG777 Project 1 Report: Spatial Analysis of Nitrate Concentration and Cancer Rates in Wisconsin

Author: Morgan Grundy

Course: GEOG 777

Introduction

Elevated nitrate concentrations in groundwater have become an increasing concern due to their potential health risks, particularly their emerging association with certain types of cancer. The objective of this project is to explore the spatial relationship between nitrate levels in well water and cancer incidence rates in Wisconsin. This was accomplished by building a browser-based GIS application that visualizes spatial patterns in both phenomena and supports interactive analysis.

This project specifically aimed to:
- Interpolate nitrate concentrations using Inverse Distance Weighting (IDW).
- Aggregate interpolated nitrate values to census tracts.
- Compare these values to population-adjusted cancer incidence rates.
- Allow dynamic exploration of the results using an interactive web interface.

Implementation Plan

Technology Stack:
- Frontend Framework: Vite (Vanilla JavaScript)
- Mapping Library: Leaflet.js
- Spatial Analysis: Turf.js
- Hosting: GitHub Pages
- Data Sources:
  - Nitrate well sample data (GeoJSON, projected: EPSG:4269)
  - Cancer rate data by census tract (GeoJSON, projected: EPSG:4269)

Data Preprocessing and Projection:
All spatial data was initially projected in EPSG:4269 (NAD83). Since GeoJSON and most browser-based spatial libraries operate in EPSG:4326 (WGS84), we converted the data to WGS84 for compatibility using GIS software. While this made integration with Leaflet and Turf.js more straightforward, it introduces slight distortions in distance calculations, especially over large geographic areas. This has minor implications for the IDW interpolation, which relies on accurate spatial distances, but was deemed acceptable for the goals of the project.

Nitrate Interpolation via IDW:
Nitrate values were interpolated using a custom implementation of the IDW method, executed on a hexagonal grid created with turf.hexGrid(). Each hexagon was assigned a nitrate value based on a weighted average of nearby well measurements:

Interpolated Value = Σ (nitrate_i / distance_i^k) / Σ (1 / distance_i^k)

The user can adjust the decay exponent k using a slider (range: 1.1 to 5.0) to explore how the spatial smoothness of the surface changes. Increasing k emphasizes local influence, while lower values produce a more generalized surface.

Aggregating to Census Tracts:
The interpolated nitrate values at hex centroids were aggregated to census tracts by averaging the values of centroids falling within each tract boundary (using turf.booleanPointInPolygon()). The result was stored in a new field, avg_nitrate, within the tract GeoJSON features.

The cancer rate field canrate was already normalized in the provided dataset. It represents population-weighted cancer incidence (i.e., total cancer cases divided by tract population), allowing meaningful comparisons between tracts of varying sizes.

Statistical Analysis: Linear Regression:
Linear regression was used to assess the relationship between avg_nitrate and canrate. We employed the regression-js library to compute the linear model and display the regression line and R^2 value dynamically. These values are recalculated each time the user adjusts the IDW exponent k, making the statistical model responsive to the underlying spatial interpolation method.

User Experience:
The application includes the following features:
- Map interaction: Zoom/pan across Wisconsin
- Dynamic IDW: Slider to control the decay exponent k
- Color toggle: Radio buttons to switch between coloring tracts by avg_nitrate or canrate
- Popups: Tract-specific nitrate and cancer values
- Regression box: Live display of the regression equation and R^2
- Legend: Adaptive color scale based on selected variable

The map defaults to a Wisconsin-wide view centered at latitude 44.5 and longitude -89.5 with zoom level 6.

Findings and Analysis

The web-based GIS application allowed users to visually explore the spatial distribution of nitrate and cancer rates, and to investigate possible statistical relationships between the two.

Visual Patterns:
- Areas with higher interpolated nitrate levels appeared in multiple parts of the state, particularly where well sampling was dense.
- Some tracts with higher cancer rates corresponded spatially with elevated nitrate levels, though this was not consistent across the entire map.

Regression Results:
- The regression model showed a weak to moderate correlation, with R^2 values typically

between 0.1 and 0.4 depending on the chosen k.
- This suggests that nitrate concentration alone is not a strong predictor of cancer incidence, although a weak relationship may exist.

Effect of Projection Choice:
- Because Turf.js performs distance calculations using WGS84 coordinates (degrees), and IDW relies on accurate distance measurement, using unprojected data may reduce accuracy.
- For large-scale public health studies, a projected coordinate system in meters (e.g., EPSG:3071) would yield more accurate results, especially in interpolation.
- However, for the purposes of this exploratory and educational tool, the tradeoff was considered acceptable.

Conclusion

This project demonstrates how a lightweight web application can combine spatial interpolation, data aggregation, and statistical modeling to analyze the relationship between environmental and public health data. By allowing users to interactively change parameters and compare results, the tool supports exploratory spatial analysis and encourages critical thinking about data quality, scale, and methodology.

While the correlation observed was relatively weak, the framework built here can be extended to test other hypotheses, incorporate additional data sources, and refine spatial accuracy through projection-aware methods.

The full application is published at:
https://sojourner1066.github.io/project-1/