COA 616 Geostatistics in Environmental Sciences
Homework 2
Assigned: September 20, 2016
Due: September 27, 2016

In this homework, you will examine data on atmospheric deposition in the upper Midwest and northeastern United States.
The data you will be analyzing was downloaded from the National Acid Deposition Program (NADP) website. The background of the acid rain and the data can be found at the website. Note all deposition values are in units of kg/ha – that is, kilograms of the chemical per hectare of land area. A hectare is $10,000 \text{ m}^2$.

Bonus: Project the data to NAD83 UTM zone 18 before carrying out the following analyses. You may proceed without projection but keep in mind projected data will be preferred in the practice.

1. Is it reasonable to assume that H^+("H" in the data) deposition in the region is normally distributed? Would you suggest using a transformation to normalize the H+ deposition data? Are there any H^+ deposition values that you would flag as possible outliers?

2. Note that two sites - NY98 and PY29 – are missing H^+ deposition data. Given the causes of acid deposition, one might expect that H^+ deposition is closely related to SO_4^{2-} deposition. Plot the relationship between H^+ deposition and SO_4^{2-} deposition and develop a regression equation. Use the regression equation to predict the two missing values. Answer the following question:
1) Is the regression relationship statistically significant? Is it strong?
2) How do you feel about the quality of the predicted H^+ deposition values? Would you feel good about including them in the data set alongside the values from other sites? Why or why not.

3. Regardless of your answers to problem 2, add your predicted H^+ deposition values for NY98 and PY29 to the dataset. Perform an analysis of spatial variation in H^+ deposition. You need:
1) Post the data. Are there any obvious patterns to the data? Any regions of high or low values? Any broad regional trends?
2) Explore anisotropy. Based on the prevailing wind direction, try θ=60° and θ=150° to plot sample semivariogram. Is the spatial variation isotropic or anisotropic? If it is anisotropic, what is the nature of the anisotropy (geometric or zonal)? Because of the small amount of data, you'll want to use a fairly wide angle tolerance, perhaps $\Delta\theta = \pm 30°$.