Adapted from

Amelia Nuding and Stephanie Hampton. March 2012. Investigating human impacts on stream ecology: locally and nationallyTeaching Issues and Experiments in Ecology, Vol. 8: Practice #1 [online]. <http://tiee.esa.org/vol/v8/issues/data_sets/nuding/abstract.html>

**Too much of a good thing? Exploring nutrient pollution in streams using bioindicators**

J. Stephen Gosnell

Department of Natural Sciences, Baruch College CUNY

PhD Program in Biology, The Graduate Center of the City University of New York

**Overview**

While pollution is a term we often use, determining what constitutes an extreme level of any substance can be difficult. Put another way, how do we know that materials that are commonly found in the environment are occurring at an unsafe level, or that some metric has dropped outside of an acceptable range? In this activity we will consider these questions by noting the relationship between levels of nitrogen and phosphorus in streams and diversity in benthic macroinvertebrates. In doing so we will explore the concepts of nutrient pollution and bioindicators while developing the tools needed to explore the distribution of data points and relationships among variables.

**Objectives**

After completing this lab, you should be able to

* Define nutrient pollution and potential impacts
* Describe drivers of nutrient pollution in the United States
* Discuss the use of bioindicators
* Predict impacts of nutrient pollution on benthic macroinvertebrates
* Filter large datasets in Google Sheets
* Develop and interpret box plots and 5-number summaries
* Plot and interpret relationships between two continuous variables, including explaining the strength of a relationship between two continuous variables using an R2 value

You need to turn in answers to each numbered question. Questions are noted throughout the text and also can be found in [Appendix 1](#bet2r82so9go).

**Introduction**

*Streams and Nutrient Pollution*

Small streams are vital to the United States' major rivers because they deliver water and provide pathways for the movement of fish and other aquatic organisms. Humans rely on both large and small waterways for drinking, irrigation, industrial uses and transportation. Thus, if the water quality of any particular stream is impacted, it affects not only *local* fish, aquatic organisms and plants, but also non-local organisms, in other parts of the state or country, and ultimately the livelihood of humans. While the term “environmental health” is a scientifically controversial term (Simberloff 1998), we may find it useful to think about stream “health” as describing the conditions that humans find desirable such as water with low levels of pollutants and pathogens, high levels of oxygen, and providing otherwise good habitat for fish and wildlife (Meyer 1997; Gordon et al. 2004). A variety of factors affect which aquatic organisms are present, such as the stream's size and morphology, geographic location, stream flow (volume and speed of water), available light, temperature, and water quality (Vannote et al. 1980; Poff 1997).

In this exercise we will examine two water quality parameters of great importance in streams: nitrogen and phosphorus. These two elements are common, naturally-occuring, and important to plant growth. They are generally found at low levels in ground and surface waters due to plant uptake, immobilization in soils, and volatilization (return to the environment)(Vitousek et al. 2002, Smith et al. 2003). However, concentrations of these elements have been increasing in streams due to human impacts. Both may be added to natural systems in the form of fertilizer applied to farm lands (and from lawns, to a lesser extent) and sewage from human or animal-based systems. Once introduced to an environment, they may be carried by ground water, rain water, or irrigation water from the land into streams (Vitousek et al. 1997; Smith 2003; Dodds 2006; EPA 2007); in some cases they also may be directly deposited via waste discharge. In addition, these nutrients may be deposited on land or water through the air after the combustion of fossil fuels or other industrial operations (Driscoll et al. 2006; Pepper et al. 2006).

These two nutrients are now among the most common pollutants in streams, lakes and coastal waters, resulting in degraded water quality (EPA 2007). When nitrogen and phosphorus levels become elevated in streams, *algae* (photosynthetic organisms) can grow extremely quickly and the waters can become cloudy, reducing the light availability in the stream. Importantly, when these algal blooms die, this large amount of dead algae fuels bacterial growth in the water. The bacteria decompose the algae and in doing so they consume much of the oxygen available in the water (Mallin et al. 2006). In many cases the waters become uninhabitable by aquatic organisms because of the lack of dissolved oxygen in the water. An extreme example of this is the “dead zone” in the Gulf of Mexico, which is an area of the ocean about as large as Connecticut, in which few aquatic organisms can survive, primarily due to the input of nutrients from the Mississippi River (USGS 2010).

Because elevated levels of nitrogen and phosphorus may make waters uninhabitable for some species, measuring local diversity presents an opportunity to monitor stream health. Aquatic *benthic* *macroinvertebrates* are insects and other small invertebrates (like crustaceans, mollusks and aquatic worms) that live in streams and other aquatic habitats. “Benthic” refers to the lowest level of a water body, which in this case is the stream bed, and this is where these animals reside. “Macro” means that these animals are large enough to be seen with the naked eye, and “invertebrates” means that they have no spine. Many common flying insects have larval stages in streams and lakes, such as mayflies, stoneflies and dragonflies. These stream macroinvertebrates play very important roles in the stream ecosystem. For example, by shredding leaves and other detritus that falls into streams, they convert terrestrial carbon and other nutrients into forms available to other stream organisms (Vannote et al. 1980; Wallace and Webster 1996). Some macroinvertebrates eat algae, and others are predators on small invertebrates. Most macroinvertebrates eventually become an important food resource to fish and birds (Vannote et al. 1980; Wallace and Webster 1996).

Benthic macroinvertebrates display a range of responses to environmental changes. Some macroinvertebrates require high levels of oxygen and low pollutant levels, while others can tolerate low oxygen and high pollutant levels. Because of this, the presence or absence of certain species can be used as an indicator of the health of a stream (EPA 2007). Since these species are not highly mobile (especially compared with fish), the presence, size, and stages of local populations can be used to infer the history of past pollution events or the overall presence of pollutants in the water or local sediments. Species used to assess the quality of the environment or changes over time are called   
bioindicators (Holt 2010).

In this activity you will be looking at values of “total nitrogen” and “total phosphorus” in streams. These are two very commonly measured water quality parameters. Total nitrogen includes all organic and inorganic nitrogen-containing compounds in the water. Inorganic forms are nitrate (NO3-), nitrite (NO2-), ammonia (NH3), and ammonium (NH4+). Organic forms include proteins, peptides, nucleic acids, urea and synthetic organic materials (Pepper et al. 2006). Total phosphorus, similarly, includes all phosphorus-containing compounds in the water, which includes orthophosphate (PO4) and organically bound phosphate. You will also be considering how these levels relate to a multi-metric index that focuses on the taxonomic and functional diversity of benthic macroinvertebrates (Stoddard et al. 2008).

*Data Set Description*

## 

In 2002, the US Environmental Protection Agency (EPA) set out to characterize the health of all the waterways throughout the continental US. The EPA is required by law as set forth in the Clean Water Act to report to Congress on the health of the nation’s waters. This survey of “wadeable” streams – streams shallow enough to sample without a boat – is the EPA’s largest effort to make a scientifically and statistically defensible claim about how healthy, or unhealthy, the nation’s waters are (EPA 2007).

A statistically sound sampling design was necessary for the EPA to be able to detect major trends in stream quality across the nation. Ideally the EPA would take samples from every waterway in the US, but that is totally infeasible; it would be a huge, expensive, and very time-consuming endeavor. Therefore, the EPA devised a sampling regime of wadeable streams. Wadeable streams provide a strong link between land use and water quality, and they contribute to larger river systems, so they are a good indicator of the health of waters throughout the entire US (EPA 2007). Even though wadeable streams are relatively small and shallow, they comprise about 90% of the length of all perennial waterways.

A total of 1,392 sites were sampled in the 48 states. The type of sampling design selected for any ecological study or experiment is key to making more general assertions about the status of waterways throughout the nation. The sampling was designed to ensure that the site selection was representative and random. More information about the sampling design can be found in Chapter 1in the section of the WSA Report entitled “How Were Sampling Sites Chosen?” (pgs. 15 – 17, <http://www.epa.gov/owow/streamsurvey/pdf/WSA_Assessment_May2007.pdf>).

**Exploring Data Sets**

*Land Cover, Use, and Impact on Nutrient Levels*

Since land use may be related to the input of excess nitrogen and phosphorus to streams, we’ll first consider land use across the United States. This also aids in considering management decisions, which may be set by region. For example, the EPA divides the contiguous United States (also known as the lower 48) into 10 regions.

Many agencies and other organizations that have a lot of ecological data, like the U.S. Geological Survey, provide those data online with some tools to explore the data. Look at the map of NLCD land cover ([Supplement 1](http://tiee.esa.org/vol/v8/issues/data_sets/nuding/resources/supplement1.pdf)), or on a website provided by your instructor, to answer the following questions. You may find it helpful to print out a color copy of [Supplement 1](http://tiee.esa.org/vol/v8/issues/data_sets/nuding/resources/supplement1.pdf), or view it on your screen at a larger size (e.g., 200%) to see more detail.

# Choose 2 focal EPA Regions. You may want to consider areas in different parts of the country (east vs west), with different traits (urban vs rural), or other characteristics, such as land cover, which is noted in [Supplement 1](http://tiee.esa.org/vol/v8/issues/data_sets/nuding/resources/supplement1.pdf). What are the dominant types of land cover for your two focal regions?

# Based on your knowledge of sources of nutrients, which of these 2 two focal EPA Regions do you predict will have the highest and lowest nutrient values in streams? Why? If you are using an online tool to determine land cover of vegetation types you can refer to [Supplement 2](http://tiee.esa.org/vol/v8/issues/data_sets/nuding/resources/supplement2.pdf) for a map of the EPA Regions.

# Overall, which EPA Region(s) do you predict will have the highest and lowest nutrient values in streams? Why?

# 

*Nutrients*

Land cover may be a useful predictor of nutrient levels, but other factors may also influence nutrient levels in streams.

# Now look at the table of mean (average) values of total nitrogen (NTL) and total phosphorus (PTL) in each EPA region, located in [Supplement 3](http://tiee.esa.org/vol/v8/issues/data_sets/nuding/resources/supplement3.pdf). Which regions have the lowest and highest mean nutrients, and how does this compare with your prediction?

# Are the mean values useful at characterizing streams in a region? Why or why not? Offer some suggestions about what additional information would be useful to more fully characterize nutrient levels within an EPA region.

Besides the average, or mean value, understanding the spread of data may be important. For example, a relatively low average level of N might hide a few extremely polluted waterways. For these reasons scientists also consider the distribution of the data. A common first step in analyzing the distribution of data points is through the use of a five-number summary.

Imagine ranking, or ordering, the data from smallest to largest. The smallest point in the dataset is the minimum. The lower quartile, Q1, is the point below which 25% of the data is contained. The middle data point is the 2nd quartile (Q2), which is also called the median. It turns out that the median is another measure of central tendency like the mean, but the median is less impacted by the mean by outliers, or points that are far away from the average value. Why?

Consider this - find the stream in your regions with the highest N value. If you doubled that value (making it even more of an outlier!) the mean would change, but the median would not!

You can also define Q1 as the median of the data points that lie between your lowest value and Q2.

The upper quartile, Q3, is the point below which 75% of the data is contained. To find Q3, you take the median of the data points that lie between your highest value and Q2. The highest value in a dataset is the maximum.

Open the WSA data table. The data is available online (WSA\_data\_for\_students.csv, accessible from: [http://knb.ecoinformatics.org/knb/metacat/nuding.9.4/knb](http://knb.ecoinformatics.org/knb/metacat?action=read&qformat=knb&sessionid=0&docid=nuding.9.4)), or you can download it from the class site. After opening the file find the minimum, first quartile, median, third quartile, and maximum values of total nitrogen and total phosphorus for your two focal regions. **YOU MAY FIND IT USEFUL TO CREATE A COPY OF THE DATA SHEET AND FILTER OUT THE DATA BY REGION AS DEMONSTRATED IN CLASS.** The median and quartiles can be found using the following functions in Google Sheets. The parentheses should contain the column or row of data you wish to analyze. Here we use cells E1:E70, just as an example.

To calculate the median, type (without the quotation marks) “=MEDIAN(E1:E70)”

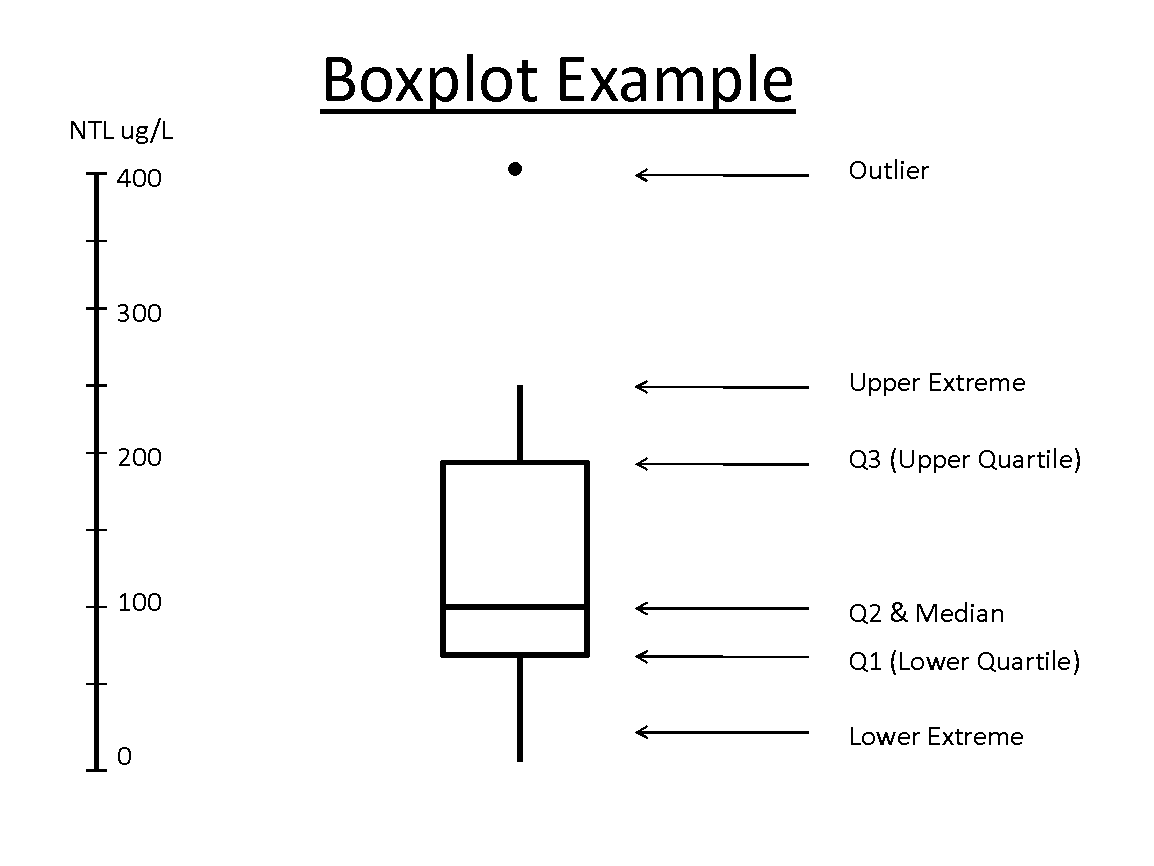
To calculate the Q1, type “=QUARTILE(E1:E70,1)”

To calculate Q3, type “=QUARTILE(E1:E70,3)”

# Fill in the following table for each of your regions.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Minimum | 1st quartile | Median | 3rd quartile | Maximum |
| NTL |  |  |  |  |  |
| PTL |  |  |  |  |  |

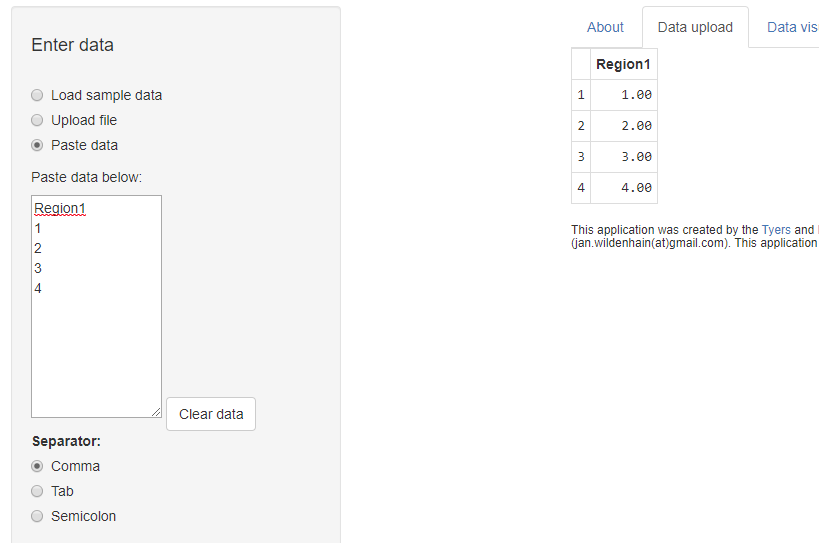
The five-number summary can also be displayed graphically. An example of how a boxplot can be used to plot this data is shown below.



**Figure 1**. An example of a box plot table for total nitrogen (NTL).

Boxplots plot the five-number summary we noted above. In some cases (but not always!) boxplots point out how extreme outliers are by considering how far they are from the median. For example, in some programs if points lie more than 1.5\*(Q3-Q1) above or below the median, those extreme points may be plotted as “dots” outside the “whiskers” of the boxplot. Other times the whiskers extend fully to the maximum and minimum.

# Produce box plots of NTL and PTL for each region using the BoxPlotR shiny app <http://shiny.chemgrid.org/boxplotr/>. You can paste the data directly into the shiny app or upload a file. You can also add the region number in the first line (it will have a weird period if you use any spaces, but that’s ok for our purposes) of the pasted or copied data. See Figure 2. Copy the plots here and add appropriate captions (Which region? Which nutrient?). You can also put in more than one dataset at a time if you prefer (e.g., compare nitrogen for both of your regions in one graph)!



**Figure 2**. Example of data entry in the shiny app.

Examining the plots can help you determine the distribution of the data. Plots with long upper whiskers or multiple high-value outliers are *right-skewed*, while those with longer bottom whiskers or multiple low-value outliers are *left-skewed*.

# 

# Describe the distribution of the data outside the box, e.g. are the points relatively close or spread widely, where are outliers, etc. Why might this be important to consider for the region?

# Given what you have learned so far, would you expect an “outlier below Q1” to have more or less algae than other streams? Why?

In a separate nutrient study by the EPA (EPA 2001), the lower quartile (Q1) value from water samples was recommended as the level below which nutrients in streams should be maintained in each *ecoregion* - note this is not the same as an *EPA Region*, as you can see in the map provided in [Supplement 4](http://tiee.esa.org/vol/v8/issues/data_sets/nuding/resources/supplement4.pdf). For example, below are the **recommended total nitrogen values** for the ecoregions within EPA Regions 7 and 10. The full data set and a map of the ecoregions are available in [Supplement 4](http://tiee.esa.org/vol/v8/issues/data_sets/nuding/resources/supplement4.pdf). **NOTE THE NITROGEN RECOMMENDATIONS ARE GIVEN IN MG/L, NOT UG!**

EPA REGION 7:

Ecoregion IV (Great Plains Grass and Shrublands) – 560 ug/L

Ecoregion V (South Central Cultivated Great Plains) – 880 ug/L

Ecoregion VI (Corn Belt and Northern Great Plains) – 2180 ug/L

EPA REGION 10:

Ecoregion II (Western Forested Mountain) – 120 ug/L

Ecoregion III (Xeric West) – 380 ug/L

Compare the information you generated in your table with the EPA recommended nutrient criteria for your regions.

### 

# How do the values you calculated for your regions compare to EPA’s recommendations for various ecosystem types, or ecoregions, within these two regions?

# 

# Based on the data and maps you’ve looked at, can you assert whether the waters in your regions are “clean” or “polluted”? Explain your answer using data from parts I and II. If you feel more information is needed to make an assertion, state what information you would need to have. Note: there is not one “right” answer here, so what is important is how well you explain your answer.

*Macroinvertebrates*

Streams often contain a large number and a wide variety of benthic macroinvertebrates. Stream ecologists often look for the pollution-sensitive insects, because they serve as bioindicators of stream health. The following taxonomic orders of pollutant-intolerant insects are collectively referred to as “EPT” because they are such a useful grouping of insects that react strongly to pollution: Ephemeroptera (mayflies), Plecoptera (stoneflies) and Trichoptera (caddisflies). These insects begin their lives in the water, and later emerge as adults to live on the land.

Macroinvertebrates can be sampled by shuffling a net along the bottom of the stream bed and counting the number and types of insects brought up in the net. Ecologists have developed sophisticated methods of characterizing these benthic macroinvertebrate communities – rather than just counting the number of insects, we consider things like the diversity of the species present, and how many of those species are known to be pollution intolerant (e.g., EPT).

The EPA developed a method of characterization that worked best for the Wadeable Stream Assessment and called it “MMI” for “Multi Metric Index” (EPA 2007). It includes the following six categories that consider both taxonomic and functional diversity:

* Taxonomic measures
  + Taxonomic richness – the number of distinct taxa (e.g., species)
  + Taxonomic composition – a measure of the abundance of the ecologically important taxa in the sample
  + Taxonomic diversity – the distribution of the numbers of organisms in different taxa
* Functional measures
  + Feeding groups – the distribution of macroinvertebrates that have different feeding habits (e.g. leaf shredders vs. algal feeders)
  + Habits –the distribution of macroinvertebrates that burrow, cling, crawl and/or swim
  + Pollution tolerance – a measure of how many taxa present are pollution tolerant and intolerant

More information about these groups can be found in the WSA report (Ch 2, pgs. 27 – 29). In all cases, a higher value indicates a greater diversity of organisms and is considered to be indicative of a healthier stream. A high MMI score (max = 100) tends to indicate a healthy stream, and low score (min = 0) tends to indicate an impaired stream.

# Make a Prediction. If we plotted the total nitrogen values (for the entire nation) on the x-axis, and the benthic macroinvertebrate index values on the y-axis, what do you think the graph would look like? Why?

# 

# Using actual WSA data, create a scatter plot of MMI vs. NTL and MMI vs. PTL, with a trend line for each region. Does the direction of the relationship between nutrient concentrations and MMI match your predictions? How does it appear that the macroinvertebrate community changes at low and high levels of total nitrogen and phosphorus?

# Besides considering the direction of the relationship, we can also consider how closely connected the two variables are. Another value, R2, can tell you how well the trend line fits the data (how strong is the relationship between the x and y variables). Display the R2 on each chart and past them here.

# 

# What are the similarities and differences between the two graphs? Is this what you expected?

# Because the nitrogen data have such a wide range of values (over three orders of magnitude) and many data points are clustered at one end, we may want to use the logarithm (log) of the nitrogen data to plot on the x-axis instead. Transforming the data this way reduces the right-skew of the data. Create this plot with MMI on the y-axis and Log NTL on the x-axis for each region and add a trend line.

# Comment on the results. Explain why the numbers on the x-axis changed as they did - please give a sample calculation. How did the log transformation change the distribution of the points and the fit of the trend line (the R2 value)? Would you also perform this operation on the total phosphorus data? Why or why not?

# Given your results, how would you advise natural resource managers to prioritize stream protection and restoration? What other studies would you need to help your decision? This question is much smaller but similar to the grant proposal you will be producing for class!

**Literature Cited**

Dodds, W.K. 2006. Eutrophication and trophic state in rivers and streams. Limnology and Oceanography 51:671-680.

Driscoll, C., D. Whitall, J. Aber, E. Boyer, M. Castro, C. Cronan, C. Goodale, et al. 2003. Nitrogen pollution in the northeastern United States: sources, effects, and management options. BioScience 53:357-374.

Gordon, D.N., T.A. McMahon, B.L. Finlayson, C.J. Gippel, R.J. Nathan. 2004. Stream hydrology: an introduction for ecologists. Wiley, New York, New York, USA.

Holt, E. A., Miller, S. W. (2010) Bioindicators: Using organisms to measure environmental impacts. Nature Education Knowledge 3(10):8.

Mallin, M., V. Johnson, S. Ensign, and T. MacPherson. 2006. Factors contributing to hypoxia in rivers, lakes, and streams. Limnology and Oceanography 51:690-701.

Meyer, J.L. 1997. Stream health: Incorporating the human dimension to advance stream ecology. Journal of the North American Benthological Society 16:439-447.

Pepper, I.A., C.P. Gerba, M.L. Brusseau. 2006. Environmental and Pollution Science. Second edition. Academic Press, Elsevier, Burlington, MA, USA.

Poff, N. 1997. Landscape filters and species traits: Towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16:391-409.

Simberloff, D. 1998. Flagships, umbrellas, and keystones: is single-species management passé in the landscape era? Biological Conservation 83: 247-257.

Smith, V.H. 2003. Eutrophication of freshwater and coastal marine ecosystems a global problem. Environ Science and Pollution Research 10:126-139.

Stoddard, J.L., Herlihy, A.T., Peck, D.V., Hughes, R.M., Whittier, T.R. Tarquinio, E.. 2008. A process for creating multimetric indices for large-scale aquatic surveys.” Journal of the North American Benthological Society 27 (4): 878–91.

U.S. Environmental Protection Agency. 2006. Wadeable Streams Assessment: a collaborative survey of the nation's streams. (<http://www.epa.gov/owow/streamsurvey/>) Accessed August 17, 2010.

U.S. Geological Survey 2010. The Gulf of Mexico hypoxic zone. (<http://toxics.usgs.gov/hypoxia/hypoxic_zone.html>). Accessed August 17, 2010.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, C.E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 11:130–137.

Vitousek, P. M., Hättenschwiler, S., Olander, L., Allison, S.. 2002. Nitrogen and nature. AMBIO: A Journal of the Human Environment 31 (2): 97–101.

Wallace, J. and J. Webster. 1996. The role of macroinvertebrates in stream ecosystem function. Annual Review of Entomology 41:115-139.

Appendix 1: List of Student Questions

1. [Choose 2 focal EPA Regions. You may want to consider areas in different parts of the country (east vs west), with different traits (urban vs rural), or other characteristics, such as land cover, which is noted in Supplement 1. What are the dominant types of land cover for your two focal regions?](#_2cquzvez8vpt)
2. [Based on your knowledge of sources of nutrients, which of these 2 two focal EPA Regions do you predict will have the highest and lowest nutrient values in streams? Why? If you are using an online tool to determine land cover of vegetation types you can refer to Supplement 2 for a map of the EPA Regions.](#_2gyufe3r9vrh)
3. [Overall, which EPA Region(s) do you predict will have the highest and lowest nutrient values in streams? Why?](#_m6rq7u1xdtg6)
4. [Now look at the table of mean (average) values of total nitrogen (NTL) and total phosphorus (PTL) in each EPA region, located in Supplement 3. Which regions have the lowest and highest mean nutrients, and how does this compare with your prediction?](#_64zu8md98jl2)
5. [Are the mean values useful at characterizing streams in a region? Why or why not? Offer some suggestions about what additional information would be useful to more fully characterize nutrient levels within an EPA region.](#_5kkhz6kez4g3)
6. [Fill in the following table for each of your regions.](#_m4tqdh50gjxh)
7. [Produce box plots of NTL and PTL for each region using the BoxPlotR shiny app http://shiny.chemgrid.org/boxplotr/. You can paste the data directly into the shiny app or upload a file. You can also add the region number in the first line (it will have a weird period if you use any spaces, but that’s ok for our purposes) of the pasted or copied data. See Figure 2. Copy the plots here and add appropriate captions (Which region? Which nutrient?). You can also put in more than one dataset at a time if you prefer (e.g., compare nitrogen for both of your regions in one graph)!](#_u2ief3hf3isp)
8. [Describe the distribution of the data outside the box, e.g. are the points relatively close or spread widely, where are outliers, etc. Why might this be important to consider for the region?](#_1wnggan9vboy)
9. [Given what you have learned so far, would you expect an “outlier below Q1” to have more or less algae than other streams? Why?](#_1wnggan9vboy)
10. [How do the values you calculated for your regions compare to EPA’s recommendations for various ecosystem types, or ecoregions, within these two regions?](#_nxop4zxb2vp4)
11. [Based on the data and maps you’ve looked at, can you assert whether the waters in your regions are “clean” or “polluted”? Explain your answer using data from parts I and II. If you feel more information is needed to make an assertion, state what information you would need to have. Note: there is not one “right” answer here, so what is important is how well you explain your answer.](#_nxop4zxb2vp4)
12. [Make a Prediction. If we plotted the total nitrogen values (for the entire nation) on the x-axis, and the benthic macroinvertebrate index values on the y-axis, what do you think the graph would look like? Why?](#_i3stihtasbcb)
13. [Using actual WSA data, create a scatter plot of MMI vs. NTL and MMI vs. PTL, with a trend line for each region. Does the direction of the relationship between nutrient concentrations and MMI match your predictions? How does it appear that the macroinvertebrate community changes at low and high levels of total nitrogen and phosphorus?](#_n6ajwg2q8sv)
14. [Besides considering the direction of the relationship, we can also consider how closely connected the two variables are. Another value, R2, can tell you how well the trend line fits the data (how strong is the relationship between the x and y variables). Display the R2 on each chart and past them here.](#_dgf05ezi8ik2)
15. [What are the similarities and differences between the two graphs? Is this what you expected?](#_42nqyi3zw3yk)
16. [Because the nitrogen data have such a wide range of values (over three orders of magnitude) and many data points are clustered at one end, we may want to use the logarithm (log) of the nitrogen data to plot on the x-axis instead. Transforming the data this way reduces the right-skew of the data. Create this plot with MMI on the y-axis and Log NTL on the x-axis for each region and add a trend line.](#_xzftlytplaoy)
17. [Comment on the results. Explain why the numbers on the x-axis changed as they did - please give a sample calculation. How did the log transformation change the distribution of the points and the fit of the trend line (the R2 value)? Would you also perform this operation on the total phosphorus data? Why or why not?](#_st7dcgkbov8y)
18. [Given your results, how would you advise natural resource managers to prioritize stream protection and restoration? What other studies would you need to help your decision? This question is much smaller but similar to the grant proposal you will be producing for class!](#_xzftlytplaoy)