Impact of Nitrogen Loading from Dairy Cattle on the Susquehanna River Basin

Prepared for: Prof. Maidment
CE 394K: GIS in Water Resources

Prepared by: Austin Weidner
Masters Candidate
University of Texas
Environmental and Water Resources Engineering

December 6, 2013

Table of Contents

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The Chesapeake Bay located in the eastern United States is a large estuary well known for its biological diversity and ecological importance. Since the 1970s, the Bay has received lots of attention due to its degrading environmental condition. Enormous nutrient loadings were reaching the bay leading to eutrophication, the growth of large amounts of microorganisms such as algae. There are significant negative consequences of eutrophication ranging from an imbalance of the food chain to hypoxia in severe cases. Hypoxia, which is commonly referred to as a dead zone, is the depletion of dissolved oxygen due to increased biological activity to the point where fish and other large organisms can no longer survive.

The Clean Water Act, enacted in 1972, helped improve and protect the health of natural waters by focusing on reducing the effects of point source contamination such as municipal wastewater treatment plants and industrial dumping. However, non-point source pollution remains a critical problem that degrades the environmental health of many water bodies including the Chesapeake Bay. Non-point source pollution is contamination that cannot be attributed to a specific location but over a large area. This is normally caused from stormwater runoff from agricultural fields and urban landscapes. As the area has become increasing developed, more and more natural landscapes are converted to impervious surfaces which exasperates the quantity of stormwater and pollutants that run off into nearby waterways. Figure 1 shows the dissolved oxygen condition of the Chesapeake Bay as of August 2009. It is estimated that the area where dissolved oxygen concentrations are too low to support life, commonly known as a dead zone, makes up 20% of the entire bay.

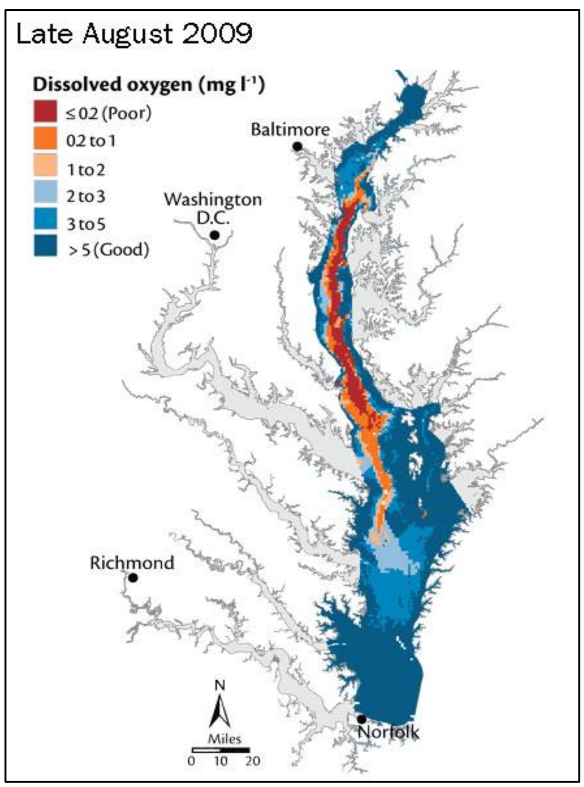


FIGURE 1: SCHEMATIC OF SOLIDS FLOW AT DAVIS WATER TREATMENT PLANT

Photo: Chesapeake Bay Foundation, 2013

Susquehanna River

Contributing 50% of the fresh water flow to the Chesapeake Bay, the Susquehanna River and its watershed, is a major source of nutrient loadings that impact the bay. The Susquehanna River, shown in Figure 2, stretches 444 miles through the Appalachian Mountains of New York, Pennsylvania, and Maryland. Its watershed drains 27,510 square miles of land (SRBC). The watershed is broken into six major subbasins as shown in Figure 3. In general, the land the watershed drains remains natural and undeveloped. Figure 4 shows a map of the land use types throughout the basin while Table 1 summarizes the land areas and the percent of the total basin area each land type covers. 62% of the basin's area is forested, located mostly in the mountainous region in the middle of the watershed. Agricultural land makes up the second biggest land type at 27.5%, while developed land makes up just over 3%. It is important to note that much of the agriculture takes place either in the northern portion of the watershed in New York and near the mouth of the river in south central PA.

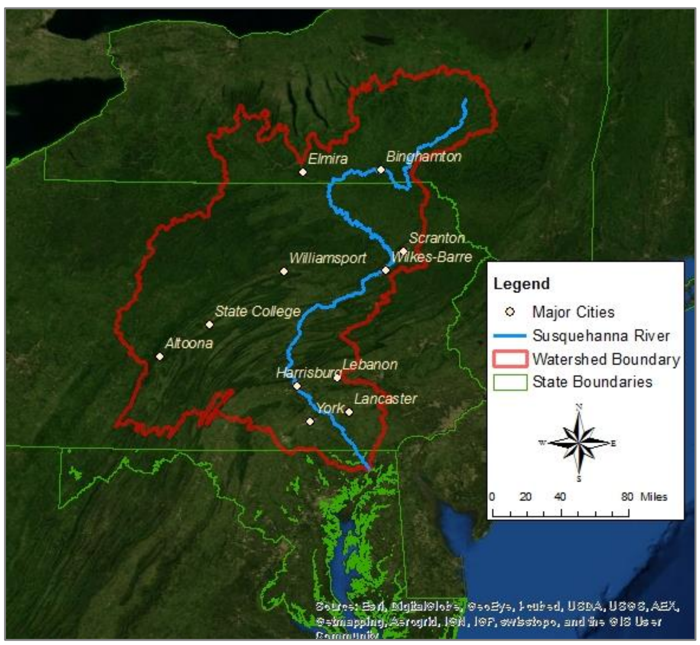


FIGURE 2: MAP OF SUSQUEHANNA RIVER LOCATION AND WATERSHED

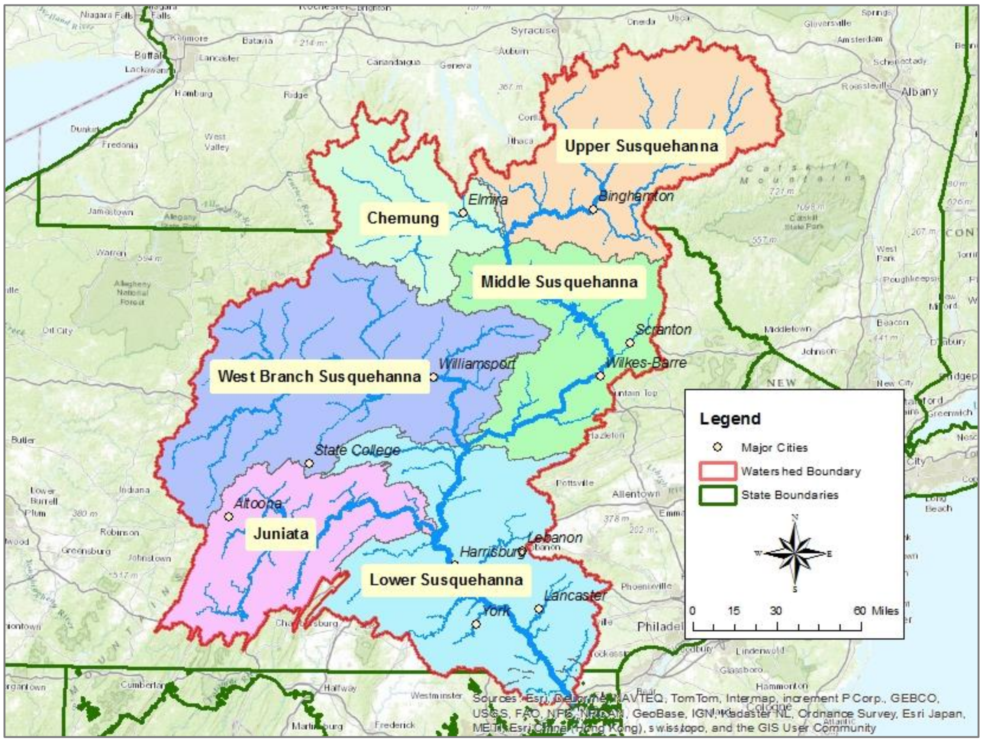


FIGURE 3: MAJOR SUBBASINS OF THE SUSQUEHANNA RIVER

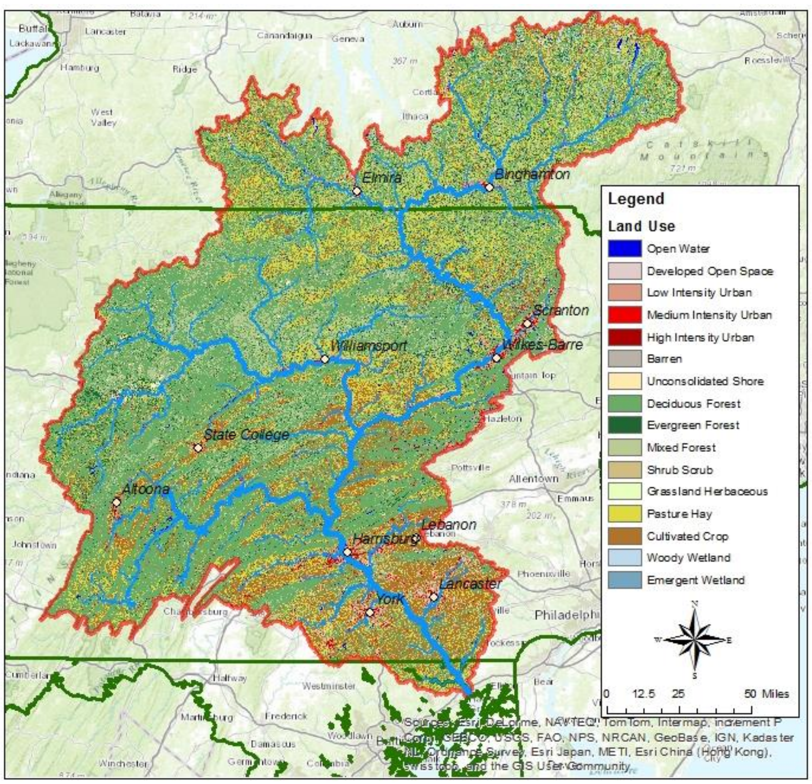


FIGURE 4: LAND USE MAP FOR THE SUSQUEHANNA RIVER WATERSHED

TABLE 1: SUMMARY OF LAND USE AREAS FOR SUSQUEHANNA RIVER BASIN

Land Use Type	Area (km^2)	Percent of Total Area
Open Water	794.8	1.1%
Developed Open Space	492.5	0.7%
Low Intensity Urban	1637.6	2.3%
Medium Intensity Urban	590.2	0.8%
High Intensity Urban	292.2	0.4%
Barren	246.3	0.3%
Unconsolidated Shore	5.8	0.0%
Deciduous Forest	32998.5	46.3%
Evergreen Forest	3500.9	4.9%
Mixed Forest	8068.2	11.3%
Shrub Scrub	1603.2	2.3%
Grassland Herbaceous	481.4	0.7%
Pasture Hay	9529.7	13.4%
Cultivated Crop	10039.4	14.1%
Woody Wetland	803.8	1.1%
Emergent Wetland	142.7	0.2%
Total	71227.1	100%

Despite a largely natural landscape, the Susquehanna River basin still contributes a significant portion of contaminants to the Chesapeake Bay. The Chesapeake Bay Foundation estimates that 40% of the total nitrogen load to the Chesapeake comes from the Susquehanna River. This is mostly due to agricultural runoff in the form of fertilizer application and livestock manure. In fact, the foundation reports that of the 40% of the nitrogen that runs off from the watershed, 50% can be attributed to livestock manure (CBF, 2006). Table 2 shows the amount of nitrogen excreted over the course of a year for each class of livestock. By far, the dairy cattle leads all other types of livestock, producing 121 kg of N per cow per year (Boyer et al., 2002).

TABLE 2: NITROGEN LOADS FROM LIVESTOCK EXCRETION FOR LIVESTOCK CLASSES IDENTIFIED IN CENSUS OF AGRICULTURE (SOURCE: BOYER ET AL., 2002)

Animal Type	Excretion N Input Estimates (kg N $animal^{-1}$ $year^{-1}$)
Beef Cattle	58.51
Dairy Cattle	121.00
Pigs and Hogs	5.84
Sheep	5.00
Goats	5.00
Horses	40.00
Chickens	0.07-0.55
Turkeys	0.39

Project Objectives

This project will focus on the impact that dairy cattle within the Susquehanna River have on the Chesapeake Bay. Dairy cattle were selected as the focus because the Susquehanna River is home to a large dairy cattle population and because they contribute more nitrogen than all other livestock types combined, thus dairy cattle are expected to have the biggest impact on downstream water quality. Gaining an understanding of the impacts from dairy cattle could lead to the development of management strategies which would help reduce non-point source pollution from dairy farms from running off and contaminating the Chesapeake Bay. In addition, observed water quality data will be used to determine if a correlation exists between actual nitrogen concentrations of streams and dairy cattle population.

METHODS

Data Sources

Susquehanna River Basin Data

The Susquehanna River Basin Commission (SRBC), founded in 1970, was established to coordinate the water resource efforts between all the different stakeholders within the watershed (SRBC). As part of their mission to educate the public to protect the water resources of the basin, the SRBC maintains a massive GIS program and make most of its data available to the public. On their website I collected land use data, precipitation, and basin and subbasin delineation maps (SRBC).

In addition I needed other physical attributes of the watershed such as flowlines and state and county boundaries. Flows lines were extracted from the National Hydrologic Dataset Plus version 2 (NHD Plus v2), which were made available through ArcGIS services. County shape files were collected for each of the three states individually.

Dairy Cattle Data

The United States Department of Agriculture (USDA) conducts a massive census of agriculture every five years. Examples of data that can be found within the Census include, acreage of crops planted, economic value of crop production, population of all types of livestock, farm size and count, etc. All of these data can be summarized at different geographic levels: county, state, region. In addition they calculate the U.S. rank for each data set within each geographic level. The USDA Census of Agriculture data for 2007, 2002, and 1997 can be easily queried and downloaded from the USDA website. For this data source I collected dairy cattle populations per county for NY, PA, and MD for the three most recent census years (USDA).

Water Quality Measurements

The Environmental Protection Agency (EPA) manages a water quality data warehouse called STORET, which is short for STOrage and RETrieval. Any organization can upload data into the warehouse including federal and state agencies, private organizations, universities, and even individuals. Once uploaded the data is made publically available without and processing by the EPA. Types of water quality data that are available include chemical, physical, biological, and habitat assessment data (US EPA). Because the EPA STORET data warehouse contains over 24 million data entries, only specific parameters that related to this project were queried. Total Nitrogen data within the 19 HUC8 basins within the Susquehanna River

Watershed was downloaded. To reduce the data further, only data measurements taken over within the time frame of 2004-2007 were used, to match the time frame of the most recent available dairy cattle data from 2007.

Despite over 5000 data points meeting the above search criteria, many problems remain within the data that are important to understand as they relate to the future analysis. First, since the data is taken by many different entities, there may be difference in sampling techniques and analytical procedures that could make comparing data difficult. The sampling time period is not taken uniformly nor instantaneously. That is, one sampling location may only have 5 data point over five different days in the autumn of 2006 while another point might have a monthly sample for the entire 2007 year. Unfortunately this means that seasonal variability or yearly trends are difficult to analyze, because no direct comparison in time and entity can be made through space. Another motivation for using a time span of four years, was to achieve more of an average data set that could to some degree cancel out seasonal, entity, and sampling time period variability.

Data Processing

County and Basin intersections and development

State boundaries were drawn by dissolving each county dataset shapefile. Since the county shapefiles were downloads for each individual state, the first step was to merge all the counties together to create one large feature class of all counties. The combined counties feature class was intersected with the watershed boundary to select just those counties that have any area within the watershed. Finally in order to do future analysis, the dataset was project into the North America Albers Equal Area Conic projected coordinate system.

Dairy Cow Calculations

Once creating the feature class of counties within the watershed, USDA Census dairy cow population data could be joined to the counties data. Since there are a few counties that share the same name with counties in the other three states, the join could not be done just by using the county name. To solve this problem a new field was created specifically for the join, in which the state ID was concatenated with the county ID. Once the two tables were joined, the field calculator in ArcGIS was used to calculate the density of dairy cows per square kilometer (Equation 1) and the nitrogen loading from dairy cow excrement per year (Equation 2).

$$Cow\ Density = \frac{Number\ of\ Cows\ per\ County}{County\ Area\ (km^2)} \quad (1)$$

$$\frac{Total\ Nitrogen\ Loading}{year \times County} = \frac{\#\ of\ cows}{County} \times \frac{121\ kg - N}{cow \times year} \quad (2)$$

Unfortunately spatial detail on the county level is not very helpful to determine the contribution of non-point source pollution into small stream segments. Instead, it is only possible to analyze contributions over much larger catchments, for example the six subbasins. This way multiple data points (in this case, counties) can be included in the analysis. To spread the county level data into six subbasins, we first had to use the intersect tool, to find the areas of each county that fall within the each subbasin. Then the intersected county areas were multiplied by the original county cattle densities to determine the number of cattle within the parts of each county that fall within the watershed. This assumes that the dairy cows are evenly distributed across the entire county, which obviously is not an accurate assumption, but is the best we can do with the limited spatial detail of the data. Finally the summary statistics tool was used to sum the dairy cows and the nitrogen loadings within each subbasin.

Flow Lines

The National Hydrography Dataset (NHD) does not allow the extraction of more than 17,500 stream segments, which means all the stream segments within the Susquehanna River Basin cannot be extracted at one time. Instead, the stream lines were extracted for each of the six subbasins and merged together. Using the select by attributes feature, the high order stream flowlines were exported to map only major tributaries. Also, the segments with the name Susquehanna River were selected and exported so the total river length could be measured.

Drainage density is a watershed characteristic used to understand how dense the river network is in a watershed (Equation 3). The total stream length for each subbasin was found from the NHD Plus v2 Flowlines using the summary statistics tool. Total subbasin area was calculated by ArcGIS's shape area feature.

$$Drainage\ Density\ = \frac{Total\ Stream\ Length\ (km)}{Total\ Basin\ Area\ (km^2)} \quad \textbf{(3)}$$

Water Quality

Typically in water quality analysis it is more important to analyze mass loadings instead of concentrations because mass is a conserved term and not affected by the discharge of each stream as is the case with concentration. However the STORET dataset does not contain discharge measurements, so no each way to convert the concentrations to loadings was possible. Therefore, all the water quality analysis used only measured concentration data. It is important to keep this assumption in mind when interpreting the results. Figure 5 shows the map of measurement locations across the watershed. It can be seen that the density of data points is greater in the northern and southern part of the watershed, which is beneficial because these areas are also determined from land use imaging (Figure 4) to be high in agricultural activity and thus pose potential risk to contaminating water bodies. The data was interpolated spatial in ArcGIS using four different interpolation techniques: spline, natural neighbor, IDW, and Kriging.

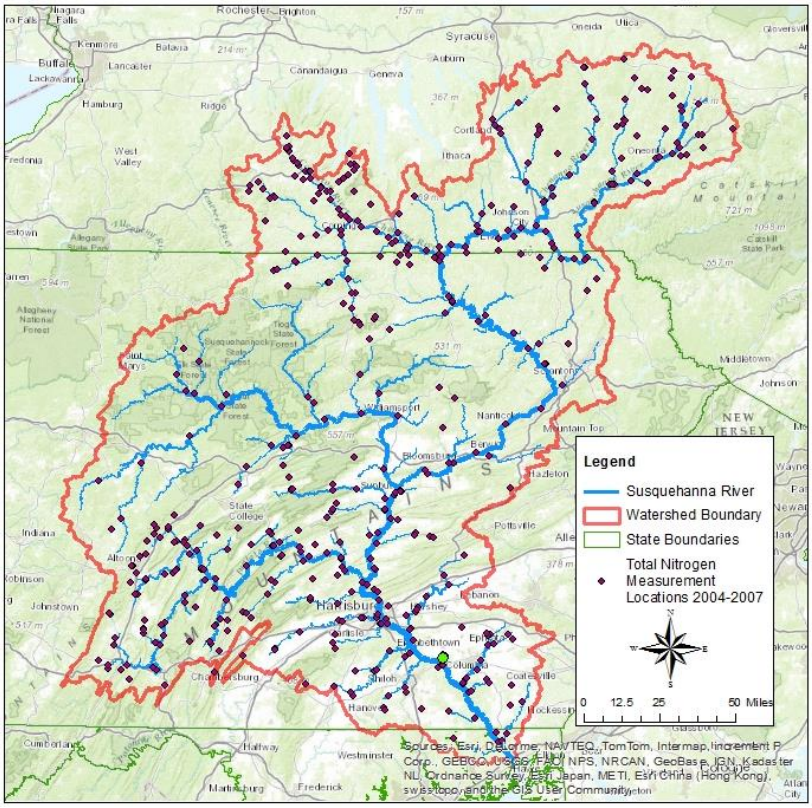


FIGURE 5: MAP OF TOTAL NITROGEN SAMPLING LOCATIONS FROM 2004-2007

RESULTS

Basin Wide Analysis

Dairy Cattle in these three states represent one of the largest dairy regions in the United States. New York and Pennsylvania rank 3rd and 4th respectively in the US in terms of economic production from dairy products, while Maryland ranks 27th. Figure 6 shows the population of dairy cattle within each state for the three years that the census data is available. Interestingly, the overall trend of dairy cattle for each of these states is decreasing. This is most likely attributed to increase in development of farmland to residential and commercial areas over the past 10 years.

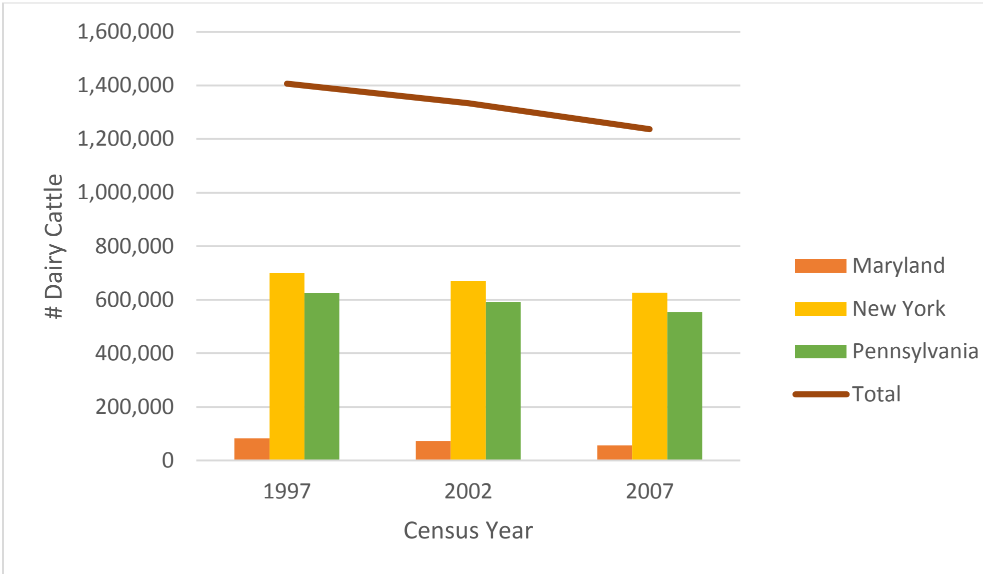
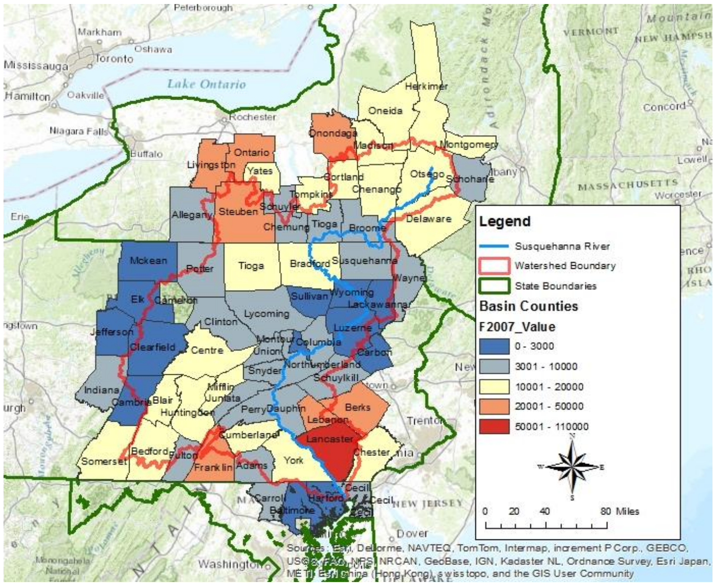


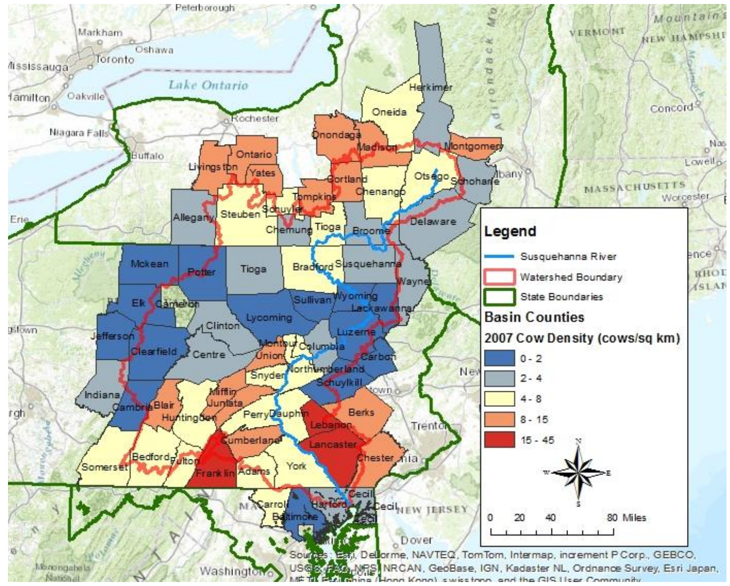
FIGURE 6: TOTAL STATE DAIRY COW POPULATION

Dairy cattle populations are mapped spatially across the watershed in Figure 7 A. Lancaster county is southern Pennsylvania stands out right away since it has by far the largest cow population, at nearly 110,000 cows, compared to ever other county in the basin. Lancaster County is one of the most productive counties in the entire US in terms of agricultural value (USDA). Cow density (Figure 7 B) is also important to examine because it better reflects the spatial distribution of cows across counties with drastically different areas. Again, Lancaster County stands out as the largest cow density, but we can also see that smaller counties in central Pennsylvania and northern New York now have greater impact due to their small size. The areas of high cattle population and density, including southern PA and north central NY match well with the areas of high agricultural land uses as seen in Figure 4. Moreover, areas with lower dairy cattle densities are located in the mountainous western and northern regions of PA, which is predominated by forest land type. Finally, the nitrogen loadings per year were calculated and mapped per county in Figure 7 C.

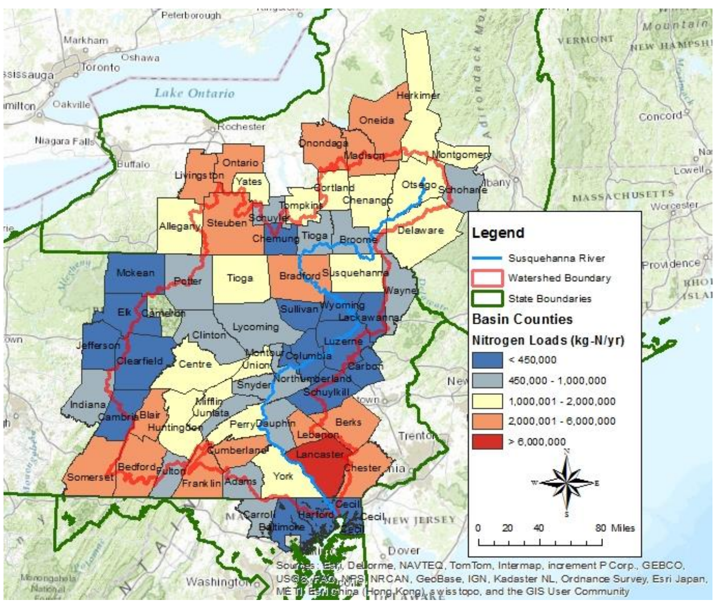
To understand the cumulative effects of each counties dairy cattle population on the overall watershed, the data was summed across the six major subbasins within the Susquehanna River Watershed. Combining the data in ArcGIS also had the benefit of only studying the area of each county that falls within the watershed. For example, Berks County in southern PA has a large cow population but because only a very small area of the county is contained in the watershed, its impact on the overall river health is greatly diminished. The maps shown in Figure 8 show the results of the data summarization for total cow population (A) and total nitrogen load per year (B). The summary of these values can be seen in Table 3. Not surprisingly, the Lower Susquehanna subbasin has the largest dairy cow population, mainly because it contains Lancaster County, which makes up over half of the total number of cows. This analysis shows that almost 450,000 dairy cattle reside within the Susquehanna River basin and produce 54,100,000 kg of nitrogen yearly just in excrement.



(A)

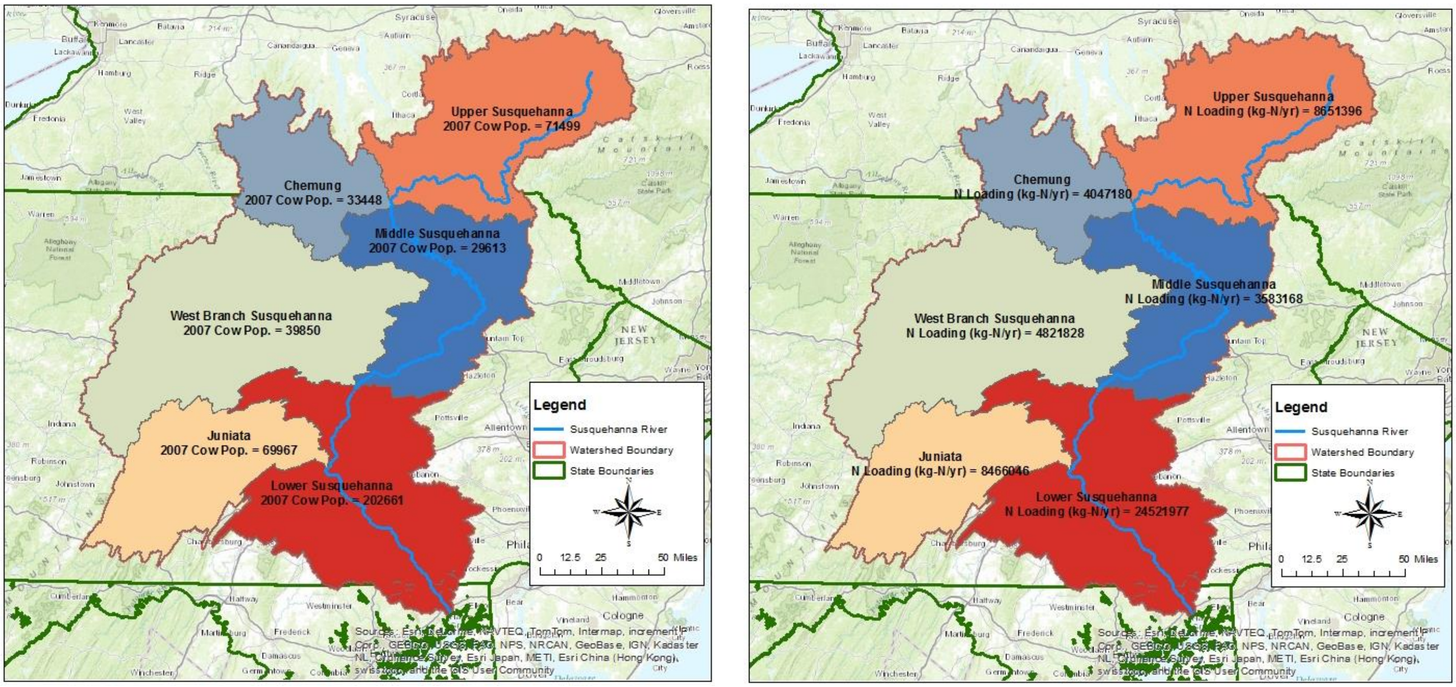


(B)



(C)

FIGURE 7: SPATIAL DISTRIBUTION ACROSS COUNTIES OF
A) DAIRY CATTLE POPULATION, B) DAIRY CATTLE DENSITY (COWS/KM^2) , AND C) NITROGEN LOAD PER YEAR



(A) **(B)**

FIGURE 8: A) DAIRY CATTLE POPULATION AND B) NITROGEN LOADING/YEAR FOR EACH OF THE MAJOR SUBBASINS OF THE SUSQUEHANNA RIVER

TABLE 3: SUMMARY OF SUBBASIN DAIRY COW POPULATION AND NITROGEN LOADING

Subbasin Name	2007 Cow Population	N from Manure (kg/year)
Upper Susquehanna	71,499	8,651,379
Chemung	33,448	4,047,208
Middle Susquehanna	29,613	3,583,173
West Branch Susquehanna	39,850	4,821,850
Juniata	69,967	8,466,007
Lower Susquehanna	202,661	24,521,981
TOTALS	447,038	54,091,598

Water Quality Analysis

The EPA has set recommended water quality parameters for fresh water streams and lakes to promote healthy aquatic ecosystems. According to the criteria for the defined ecoregions that fall within the Susquehanna River basin, the recommended total nitrogen concentration should be between 0.30 to 0.70 mg/L (EPA, 2012). The collected water quality data were screened using a much larger TN concentration of 4 mg/L as a threshold. The larger concentration above the EPA recommended value was selected to identify only those sampling locations that exhibit major nutrient contamination. The locations that observed any measurements that exceed the threshold are shown in Figure 9 D as larger points than those that fall below the threshold. The measurement locations are superimposed over the county nitrogen loading per county map as shown earlier. This allows us to see that almost all of the stations reporting TN concentrations above 4 mg/L are located in counties that have high nitrogen loadings from dairy cattle.

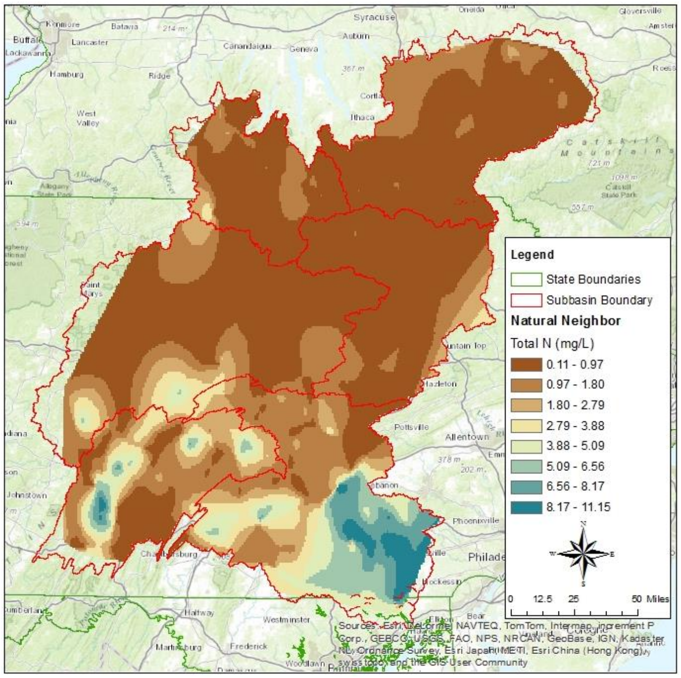
Another visualization tool, interpolation, was used to see all of the water quality data points averaged out spatially across the entire watershed. Interpolation allows us to identify areas within the watershed that exhibit large observed nitrogen concentrations as potential risk areas, but with the spatial, and time variables washed out. Maps of the three best interpolation techniques are shown in Figure 9 A, B, and C. The spline interpolation was thrown out because it contained drastically positive and negative numbers that skewed the interpolation. As can been seen, all three of the interpolation techniques shown in Figure 9 (Kriging, IDW, and Natural Neighbor) show similar results. However, Kriging was selected as the best of all the strategies because it is overall smoother, which means it reduces the effect of the time and entity variables in the data.

The interpolation maps correlate well with the previous maps showing dairy cattle distribution and nitrogen loading from dairy cattle within the watershed. The streams with the highest nitrogen concentrations can be found in the area of Lancaster County, with a second large pocket of high TN concentrations in the southwest corner of the watershed in Cumberland, Bedford and Blair Counties. Interestingly, the Kriging interpolation does not show very high concentrations of TN for streams in New York, even though we know the population of dairy cattle is high in this area. The discontinuity could be due to factors including farming techniques and policies used between NY and PA or other sources of nitrogen that might be used in the region besides dairy cow manure.

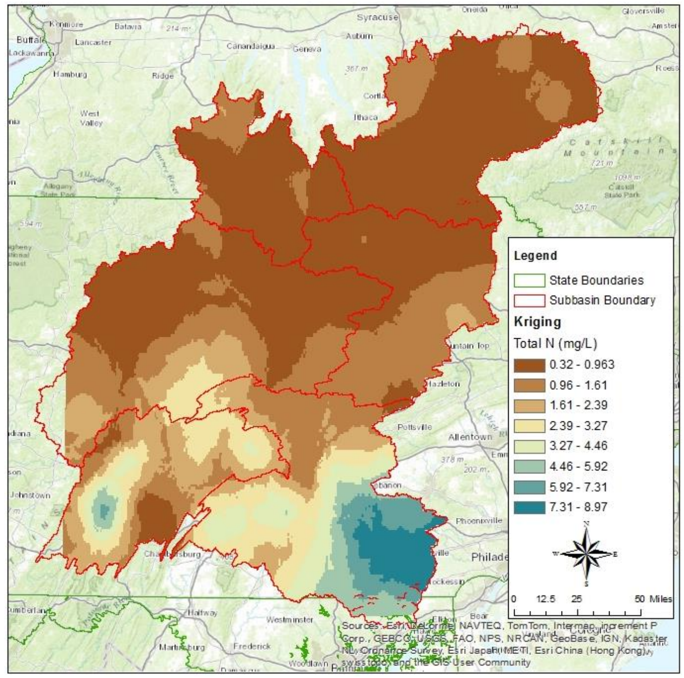
A comparison of the land use types (Table 4) between the northern two subbasins and the Lower Susquehanna Subbasin reveals a possible explanation for the different between the two areas. The Lower Susquehanna Basin has nearly twice as much percent of land covered by crops, while the upper two basins have a higher percentage of land devoted to pasture than the lower subbasin. Knowing that fertilizer is also a major source of nitrogen it could be theorized that the high TN concentrations in the Lower Susquehanna Subbasin is due to crop agriculture instead of livestock manure. Likewise, the lack of high concentrations in the streams of the upper basins regardless of the large dairy cattle population, is because there is much less crop farming in this region and thus less fertilizer use that could potentially runoff.

TABLE 4: LAND USE COMPARISONS BETWEEN THREE SUBBASINS

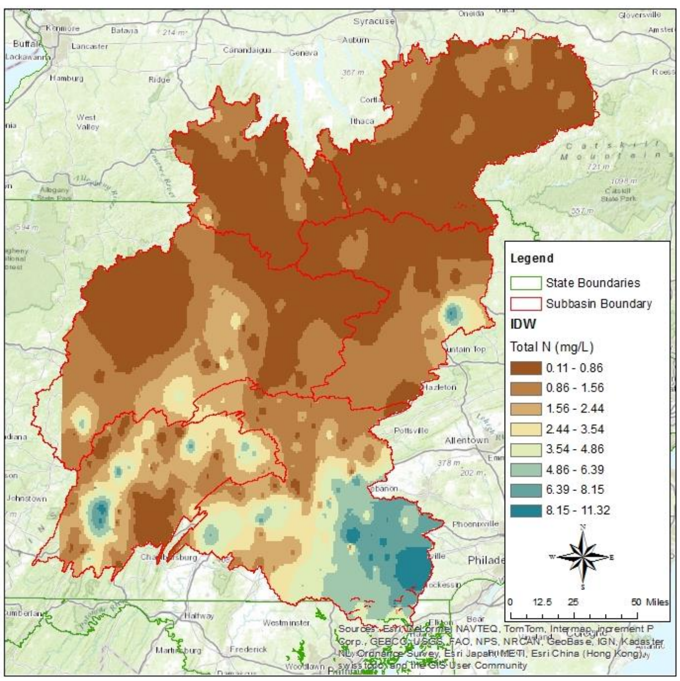
Land Use Type	Upper Sus. Basin		Chemung Basin		Lower Sus. Basin	
	Area (km2)	% of Total	Area (km2)	% of Total	Area (km2)	% of Total
Open Water	131.3	0.64%	41.9	0.39%	295.9	1.31%
Developed	326.7	1.59%	172.7	1.60%	1386.4	6.12%
Forest	7561.6	36.90%	3963.0	36.70%	6363.2	28.08%
Shrub Scrub	354.7	1.73%	238.7	2.21%	390.1	1.72%
Pasture Hay	2492.7	12.17%	1370.2	12.69%	1951.7	8.61%
Cultivated Crop	1180.2	5.76%	778.7	7.21%	4417.4	19.49%
Wetland	554.4	2.71%	98.6	0.91%	104.9	0.46%



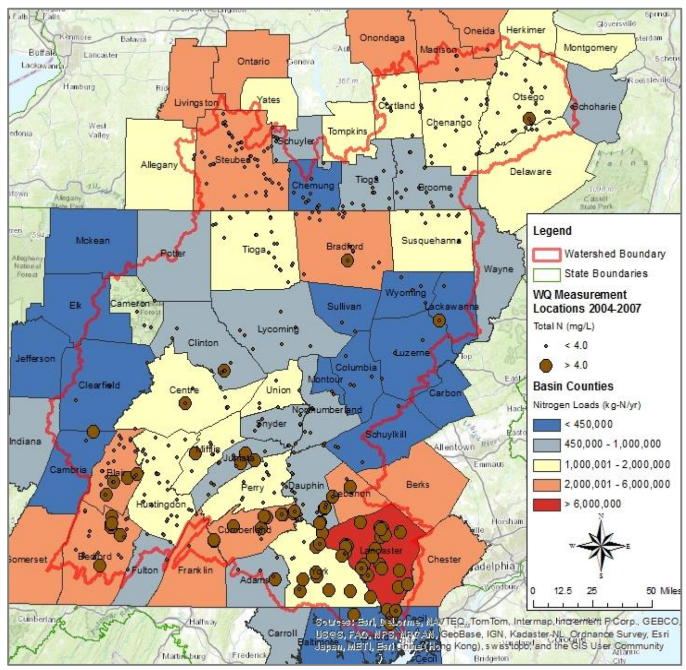
(A)



(B)



(C)



(D)

FIGURE 9: OBSERVED TOTAL NITROGEN CONCENTRATIONS VISUALIZED BY A) NATURAL NEIGHBOR INTERPOLATION, B) KRIGING INTERPOLATION, C) IDW INTERPOLATION, AND D) TN CONC. EXCEEDING 4MG/L COMPARED TO COUNTY WIDE N LOADS FROM DAIRY CATTLE

The final analysis that was conducted was to investigate if there was any correlation between drainage density and nitrogen loading. The hypothesis was that higher drainage densities would mean that more streams would be closer to non-point pollution sources and thus result in a higher observed concentration of nutrients in the streams. Table 5 shows the drainage densities that were calculated for the six major subbasins within the watershed. Except for the Chemung subbasin, which has a drainage density of 0.62, all the other basins have very comparable densities between 0.73 and 0.78.

TABLE 5: DRAINAGE DENSITY OF SUSQUEHANNA RIVER SUBBAINS

Subbasin	Total Stream Length km	Total Area km2	Drainage Density 1/km
Upper Susquehanna	9981.2	12807.3	0.779
Chemung	4156.4	6719.8	0.619
Middle Susquehanna	7126.9	9766.4	0.730
West Branch Susquehanna	13097.9	18073.3	0.725
Juniata	6780.1	8815.2	0.769
Lower Susquehanna	10907.0	15044.8	0.725

Without a major difference in drainage densities it is hard to determine whether this watershed characteristic has any effect on runoff. For example, the Lower Susquehanna subbasin has the most nitrogen from manure at 24.5 mil kg/year while the Middle Susquehanna only produces 3.6 mil kg/year (Table 3). In addition the Kriging interpolation map shows high observed TN concentrations in streams for the Lower subbasin than the Middle subbasin. However, the drainage densities for these two subbasins are almost identical near 0.73. Unfortunately this leads us to conclude that drainage density cannot be used to correlate expected loadings from dairy cattle to observed concentrations.

CONCLUSIONS

As the largest tributary to the Chesapeake Bay, the Susquehanna River is also a major source of nutrients that harm the bay’s health. A dairy cow population of nearly 450,000 contribute significant nitrogen to the basin in the form of manure. Major areas of cattle population are found in the southern and northern part of the watershed. Observed Total Nitrogen concentrations in streams within the watershed strongly correlate to the expected nitrogen loadings due to dairy cow populations. However, it is not quite clear if the high observed concentrations are a result of dairy cow manure directly or from other nutrient runoff sources like crop fertilizer, urban areas, and air deposition.

Future work will include better quantifying the different sources of nitrogen besides livestock manure. NANI, New Anthropogenic Nitrogen Inputs, is a GIS toolbox that can be used to quantify nitrogen inputs from all anthropogenic source. Other similar toolsets are specific for regions and states, such as TX-ANB for Texas (Meyer, 2006). Once more specific data is collected, the data can be spread to smaller spatial areas and then modelled with a river network to determine streams that exhibit high concentrations.

REFERENCES

Chesapeake Bay Foundation. (2006). *Waters at Risk*. Retrieved from: http://www.cbf.org/document.doc?id=197.

Chesapeake Bay Program. (2013). *The Conowingo Dam and Chesapeake Bay Water Quality.* Retrieved from: http://www.chesapeakebay.net/documents/Backgrounder_-_Conowingo_Dam_1_14_13_FINAL.pdf.

Boyer et al. (2002). Anthropogenic nitrogen sources and relationships to riverine nitrogen export in the northeastern U.S.A. *Biochemistry,* 57/58, 137-169.

Meyer, L. (2012). *Quantifying the Role of Agriculture and Urbanization in the Nitrogen Cycle across Texas.* (Master's Thesis). University of Texas, Austin, TX.

National Hydrography Dataset. (2012). *NHDPlusV2 Flow Lines* [Shape File]. Retrieved from: http://landscape1.arcgis.com/arcgis/services.

Susquehanna River Basin Commission. (n.d.). *Susquehanna River Basin Commission Overview.* Retrieved from http://www.srbc.net/about/geninfo.htm.

Susquehanna River Basin Commission. (n.d.). *Basinwide Maps & GIS Data* [Shape File]. Retrieved from: http://www.srbc.net/atlas/bwmg.asp.

United States Department of Agriculture. (2007). *Census of Agriculture* [Data File]. Available from: http://quickstats.nass.usda.gov/.

United States Environmental Protection Agency. (2013). *STORET Data Warehouse* [Data File]. Available from: http://www.epa.gov/storet/dw_home.html.

United States Environmental Protection Agency. (2012). *Ecoregional Criteria Documents.* Retrieved from: http://www2.epa.gov/nutrient-policy-data/ecoregional-criteria-documents.