Journal of the Korean Geographical Society, Vol.37, No.5, 2002(536~550)

Spatial Variations of Nutrient Concentrations in Pennsylvania Watersheds

Heejun Chang*

펜실바니아 유역 수질의 공간적 변이에 관한 연구

장 희 준*

Abstract : This paper investigated the spatial variations of nitrogen (N) and phosphorus (P) concentrations for 38 watersheds in Pennsylvania using 7 years of hydroclimatic and water quality data. Watersheds with higher percentage of urban and agricultural land uses exhibited larger variations of nutrient concentrations than forested watersheds. N and P concentrations were strongly associated with agricultural and urban land uses, respectively. The principle component analysis identified three components - land use and topography related, hydroclimate related factors, and size. Results of partial redundancy analysis showed the joint effect of climate, land cover, and topographic variables for explaining 28.1 % of the variance of nitrogen concentrations and a pure effect of land cover for explaining 41.8 % of total variance of P concentrations. The geographical pattern of statewide nutrient concentrations demonstrated a strong spatial gradient; low concentration in northwestern PA and high concentrations in southeastern PA. This pattern is associated with combined effects of hydroclimate, land use, topography, and water quality at the regional scale.

Key Words : water quality, nitrogen, phosphorus, spatial variation, Pennsylvania, geographic information systems

요약 : 이 연구는 7년 동안의 수문, 기후, 수질자료를 이용하여 펜실바니아 38개 유역 수질의 공간적 변이 패턴을 고찰하였다. 연구기간동안 도시적, 농업적 토지이용이 많은 유역에서 농도의 변이가 심하게 나타났다. 질소의 농도는 농업적 토지이용과, 인의 농도는 도시적 토지이용과 상관관계가 높게 나타났다. 주성분분석을 통해 이들의 농도 변이를 설명하는 주요 성분 셋이 - 토지 지형관련, 기후관련, 규모 - 도출되었다. 부분 여분 분석을 통해 기후, 토지, 지형의 복합적인 영향이 질소 농도 변이의 28.1%를 설명하고 있으며, 순수한 토지 변수의 영향이 인 농도의 41.8 %를 설명하고 있음을 알 수 있었다. 주(州) 전체적인 차원에서 농도변이의 뚜렷한 지리적 패턴-북서부지역의 낮은 농도와 남동부지역의 높은 농도-이 나타나고 있으며, 이는 지역적 규모에서 수문, 기후, 토지이용, 지형 등 여러 가지 환경변수의 복합적인 상호관련성을 반영하고 있다.

주요어 : 수질, 질소, 인, 공간적 변이, 펜실바니아, 지리정보시스템

1. Introduction

Many studies have demonstrated that the water quality of watersheds is determined by physical, chemical, and biological processes within watersheds(Osborne and Wiley 1988; Peters 2000). Linking watershed characteristics and stream water quality has been a persistent theme in many hydroecological research studies(Walling and Webb 1986). Previous studies pointed out that water pollution primarily results from diverse human activities, superimposed on the watershed's natural characteristics such as geology, geomorphology, and climate(Gibbs 1970; Meybeck and Helmer 1989). Most recent studies, however, focused on a single factor analysis of water quality within one or a few basins,

---

* Assistant Professor, Department of Geography, Portland State University, Portland 97201, USA (email: changh@pdx.edu).

hard to make a general statement for explaining the spatial variations of water pollution. Thus, these studies were unable to determine the relative importance of the factors governing water quality.

There are a few recent studies that employed multiple environmental variables for assessing water quality. Ekholm et al.(2000) examined the effects of catchment characteristics and riverine processes on the concentrations of nutrients and total suspended solids for 12 tributary catchments in southwestern Finland. Three environmental variables - field percentage, the mean slope of the fields, and size of the catchments - explained most of the variations in flow-weighted mean concentrations of water quality variables, but different regression models were employed for each constituent(Total P, Total N, Total Suspended Solid, Dissolved Reactive P, Ammonium N, and nitrate + nitrite-N). In a similiar vein but more focused on aquatic ecology, Bis et al.(2000) found that the most important explanatory variables for stream conditions and nutrient dynamics were patch size, patch cover in the watershed, the length of riparian ecotone and the river dimensions. Miller et al.(1998) examined varying nutrient concentrations with relation to diverse land uses, including agricultural areas(cropland and pasture), urban areas, and forests. Among agricultural areas, streams draining areas of intense row cropping typically contained higher nitrate concentrations than those draining pastures. Streams draining forested areas typically had the lowest nutrient concentrations. In these previous studies, a variety of catchment characteristics, including land use patterns, elevation, and hydrology database were incorporated into a geographical information system(GIS) to evaluate anthropogenic impact on landscape properties. Most previous studies, however, either examined variations of water quality along one river or focused on agricultural watersheds, limiting comprehensive assessment of water quality at varying landscape characteristics at the regional scale.

The aim of this paper is to identify and describe the key variables affecting spatial variations of nutrient concentrations at the regional scale. A description of the study area is provided in the next section of this article. The third section describes the data and methods employed in the study. The fourth section reports the results of analyses focusing on spatial variations of the N and P concentrations for 38 watersheds. The final section draws conclusions of the study. The approach used in the current study is similar to the previous studies in Texas(Hudak 2000) and Sweden(Arheimer and Liden 2000), but this study differs from the previous research in terms of the methods employed. Whereas these studies used multiple regression models to explain the spatial variability of nutrient concentration, principal component analysis(PCA) and partial redundancy analysis(RDA) are used here to unravel complex interrelationships among hydroclimate, land use, and water quality.

2. Study Area

Thirty-eight Pennsylvania(PA) watersheds were employed for this study(see Figure 1). Pennsylvanian creeks and rivers exhibit a wide range of nutrient concentrations, providing an ideal laboratory to examine diverse human and physical impacts on water quality. Most of these watersheds drain to the Susquehanna River that provides almost two thirds of freshwater to Chesapeake Bay. Several watersheds located in the far east and west side of the State drain to either the Delaware River or the Ohio River.

The 38 small-scale watersheds are chosen because they have both good streamflow and nutrient concentration data at the same location for a relatively long time period. Basin sizes range from 164.94 km^2(Wisahickon) to 2552.07 km^2(Pine Creek). They spread over three physiographic regions, the Allegheny Plateau(Northwestern PA), the Ridge and Valley(Central PA), and the Piedmont(Southeastern PA). With diverse geomorphic and geologic charac-

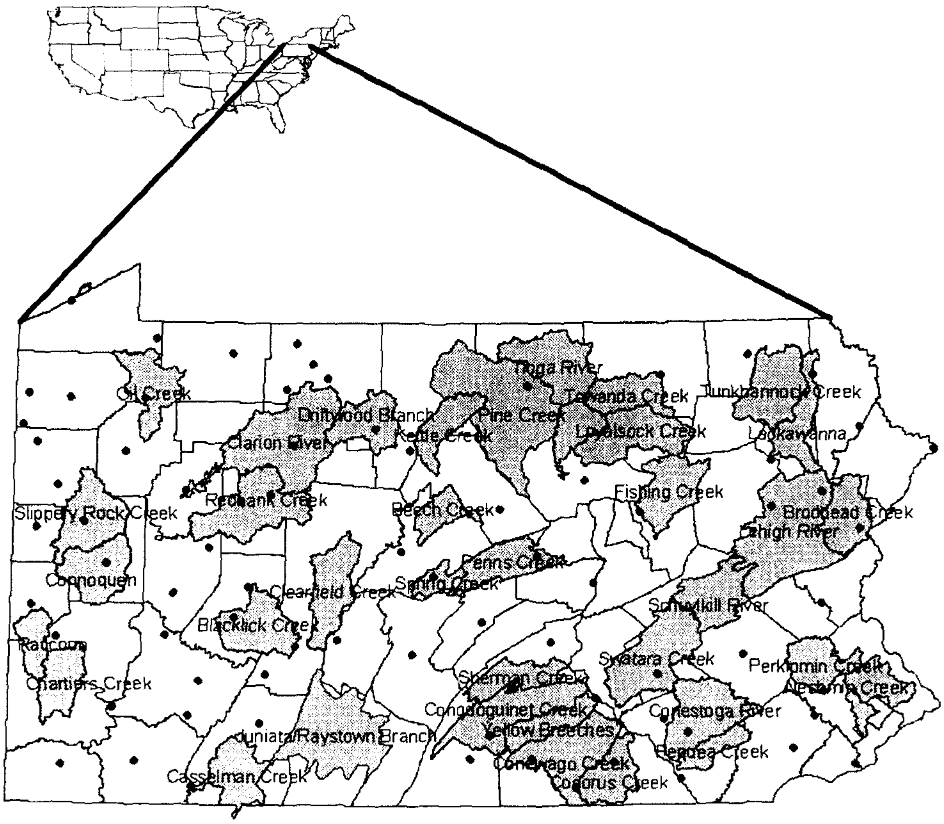

Figure 1. Locations of 38 study watersheds in Pennsylvania. Dots represent weather stations employed for estimating watershed climates.

teristics, these physiographic regions have different hydrochemical and human settlement patterns. With rolling hills and sedimentary rocks, much of the Allegeny Plateau is covered with forests except for major metropolitan areas such as Pittsburgh. The Ridge and Valley has a folded terrain with diverse soil conditions, which influences human use of land and population. The Piedmont contains the principal agricultural areas in PA, with denser population.

Although the climate of PA is broadly categorized as having similar climate and relatively even distribution of precipitation throughout the year, there are some regional variations. Latitude, elevation, and topographic factors create local variations of precipitation. Temperatures are also controlled by latitude, elevation, and proximity to a water body. Mean annual temperatures increase southeastward where the Piedmont region is located. Temperatures are lowest in northwestern PA because of the increasing elevation and continentality. For further description of the regionalization of PA climate, see Yarnal(1995).

3. Data and Methods

Table 1 describes all environmental variables for 38 watersheds analyzed. For the spatial variations of N and P concentration, data for 1990 - 1996 were used to describe the current status of water quality with relation to various hydroclimatic and land use characteristics. Hydroclimatic data originally collected on a daily time scale were aggregated into monthly data. Daily precipitation and temperature data were obtained from the National Climate Data

Table 1. Watershed environmental variables for 38 watersheds in PA.

Watershed	Agricul - ture (%)	Forest (%)	Urban (%)	Size (km^2)	Mean Slope	Standard Deviation of slope	Mean Elevation (m)	Runoff (cm)	Tempera- ture(°C)	Precipita- tion(cm)
Schuylkill	14.6	74.8	3.8	907.6	4.96	4.32	315	78.6	8.6	141.0
Wisahickon	16.1	40.7	42.3	164.9	1.06	1.39	85	58.4	10.1	112.3
Perkiomen	42.7	51.6	4.9	721.5	1.81	1.99	141	60.6	10.1	112.3
Neshaminy	40.4	37.8	20.2	539.4	1.04	1.21	91	55.2	10.1	112.3
Lehigh	10.7	83.3	2.4	2284.5	3.34	3.64	426	79.8	6.8	123.6
Broadhead	7.8	86.5	4.2	778.6	3.43	3.32	330	79.9	10.9	128.5
Pequea	71.0	26.3	2.3	397.3	2.25	2.11	153	44.3	11.2	113.7
Codorus	62.2	27.2	9.0	719.0	2.94	2.25	201	41.2	11.7	117.5
West Conewago	63.8	32.4	2.7	1322.3	1.95	2.41	188	52.4	10.4	119.8
Swatara	48.8	43.8	5.7	1479.3	2.78	3.81	199	62.5	10.4	107.1
Yellow Breeches	36.8	56.4	6.1	565.9	3.24	3.57	245	55.4	11.0	115.1
Conodoquinet	61.3	32.8	5.0	1312.7	2.88	2.70	357	48.9	11.5	110.4
Juniata/Raystown	33.4	64.6	1.2	1862.2	5.54	4.90	460	47.8	10.0	104.7
Penns	28.7	70.3	0.3	803.5	5.75	5.01	425	60.7	10.5	114.7
Conestoga	63.9	25.4	9.6	1214.7	1.69	2.28	145	52.7	11.2	113.7
Sherman	30.0	69.1	0.2	633.3	5.54	4.88	290	56.8	10.7	109.5
Fishing	30.2	68.5	0.6	942.9	5.26	5.15	372	67.1	8.1	90.6
Lackawanna	10.2	72.2	13.4	693.3	4.14	3.79	471	52.9	10.1	96.3
Tunkhonnock	28.6	68.0	1.5	1071.5	4.52	3.75	397	49.9	6.5	122.3
Towanda	30.9	68.4	0.3	717.7	5.31	4.98	474	53.8	8.9	89.0
Tioga	34.0	64.3	0.4	1146.0	5.30	4.35	518	48.9	7.2	91.6
Loyalsock	10.1	88.6	0.2	1131.5	5.70	5.90	505	69.7	9.7	97.9
Lycoming	13.4	85.5	0.5	558.1	7.78	7.24	494	70.8	10.5	106.9
Pine	10.5	88.5	0.2	2552.1	7.25	6.60	533	59.8	7.2	91.6
Spring	49.5	44.0	6.1	297.9	2.94	3.76	391	43.3	9.6	109.3
Driftwood	2.5	96.6	0.4	770.9	7.93	6.12	521	66.1	8.8	111.7
Clearfield	13.9	80.6	1.0	976.7	3.72	2.98	517	61.7	8.9	125.7
Beech	3.9	90.1	0.6	444.0	4.94	4.66	482	65.7	9.9	107.5
Kettle	3.2	95.9	0.1	638.1	8.55	6.74	536	63.7	8.6	115.3
Casselman	34.8	61.2	1.6	828.8	3.50	3.31	692	65.5	9.5	115.8
Blacklick	20.7	73.9	2.5	999.4	3.47	3.22	501	67.5	10.1	120.8
Redbank	24.9	70.9	2.2	1376.8	3.80	3.06	471	67.6	8.4	116.0
Clarion	5.6	91.8	1.1	2134.7	3.76	3.59	534	70.2	7.5	113.6
Oil	21.2	77.4	1.0	827.8	2.79	2.81	456	69.2	8.1	122.1
Raccoon	26.3	67.7	3.5	463.7	3.30	3.07	336	39.2	11.2	98.2
Connoquenessing	38.2	56.3	4.6	921.8	2.48	4.27	221	46.6	9.4	106.7
Chartiers	32.6	48.9	17.4	711.8	3.33	2.93	338	40.2	11.1	99.7
Slippery Rock	38.7	57.2	1.5	1052.7	2.20	2.28	395	54.8	8.6	114.0

Center. As shown in Figure 1, weather stations in and around each watershed were used for the analysis. When there were more than one stations in the watershed, arithmetic averages were employed. When there were no weather stations within the basin the nearest weather station from the center of the watershed was chosen as an alternative. Daily streamflow data were measured at the outlets of each watershed from the United States Geological Survey flow gauges.

Water quality data for N and P concentrations were obtained from the EPA's STORET CD(Earth Info 1998), a composite data from various data collection agencies. The CD contains two types of data; monthly grab samples from a fixed station Water Quality Network(WQN) and intensive samples(15 to 30 minutes interval) at certain sites for a particular time period; Only monthly WQN data were employed for the current study. Monthly grab samples were randomly taken at the same locations where the streamflow data were collected. There may be some uncertainties in using these monthly random samples. First, samples may not necessarily represent mean conditions of each watershed because of the sporadic sampling frequency during the period of data collection. In addition to different sampling frequencies, because the samples were taken on different days with diverse weather and hydrologic conditions across the watersheds, direct comparisons among the watersheds' water pollution may be biased. However, samples were collected at different times of the year, making it possible to identify generalized seasonal patterns of changes in nutrient concentrations for each watershed. For land use data, Landsat thematic mapper(TM) data from 1987 to 1994 were employed(Loveland and Shaw 1996).

Descriptive statistics, principal component analysis(PCA), redundancy analysis(RDA), and geographic information systems(GIS) were employed for this analysis. Descriptive statistics including the mean, median, and standard deviation were used to describe water quality variations across watersheds. Multiple regression analysis has been commonly used for empirical analysis of spatial variability in nutrient concentrations(Arheimer and Liden 2000). Because there are high collinearities among independent variables, it is necessary to reorganize such variables. PCA, the ordination technique that constructs a theoretical variable that minimizes the total sum of squares after fitting straight lines to the environmental data(Jongman et al. 1998), has been used to reduce the dimensionality of the data set. The original n variables correlated with each other are transformed into uncorrelated(orthogonal) variables. The orthogonal variables called principal components(PCs) are weighed linear combinations of the original values. The associated variances of PCs are measured by eigenvalues, and the sum of eigenvalues is the total number of variables. PCA thus makes it possible to sort out variables and classify observations(Vidal et al. 2000). In hydroecology, PCA has been used to identify major landcover types within the watershed(Aspinall and Pearson 2000), to separate human-induced variations in water pollution from natural sources of variation in hydrochemistry for various constituents(Pardo 1994; Vega et al. 1998; Perona et al. 1999). In the current study, PCA was applied to the normalized data to assess association between environmental variables.

The partial redundancy analysis(RDA) was used to estimate the total variation of nutrient concentration by three sets of independent variables - land cover, hydroclimate, and topographic characteristics. Similiar to PCA, RDA is the technique selecting linear combination of environmental variables that gives the smallest total residual sum of squares (Jongman et al. 1997). RDA is able to partition the influence of each set; it decomposes variations by the pure effect of each set as well as the joint effects. Although becoming popular in ecological studies(e.g., Borcard et al. 1992; Anderson and Gribble 1998), RDA has not been applied to explain variations in nutrient concentrations in the literature. To run the partial RDA in this study, nutrient concentrations were employed as response, and other three sets were used as either environmental (exploratory) variables or covariables(variables whose influence is partialled out of the analysis). In a sense, the partial RDA is a residual analysis where the relationship between the response and explanatory variable is analyzed once the effect of the covariables has been removed(Liu 1997). RDA analyses were undertaken using FORTRAN program CANOCO Version 4.0(ter Braak and Šmilauer 1998).

ARCVIEW GIS and Spatial Analyst are used to estimate the proportion of land use types, slope, standard deviation of slope, and the mean elevation of each watershed. Digital Elevation Model was used to derive all topographic variables. Land cover is determined by generating and clustering mosaics of multiple scenes of TM data and labeling spectral classes into 15 land cover categories at a spatial resolution of 30m × 30m. For the detailed methods of land use classification, refer to Vogelmann et al.(1998). The original 15 categories are then collapsed into 5 categories: agricultural, urban, forested, disturbed(quarry & mining), water using ARCVIEW GIS. Soils and Geological information were not considered in the current analysis because their control on N and P concentration were considered relatively minimal unlike other solutes such as conductance, magnesium and dissolved SiO_2.

4. Results and Discussion

1) Variations of nutrient concentrations for selected watershed: a synopsis

Figures 2 and 3 illustrate box-whisker plots of N and P concentration for selected watersheds with different land uses. The box-whisker plots provide brief synopsis of the distribution of nutrient concentrations for each watershed. The plot shows which watersheds have a greater range of nutrient concentration over the study period than the other watersheds. Eliminating the outliers and extreme values, the Conoquenessing watershed has the greatest range of N concentrations extending from 1.5 to 15.3 mg/L, followed by the Conestoga River watershed(3.3 - 11.4 mg/L). Although most median N concentrations are less than 10 mg/L in most cases, there are some instances that the concentrations exceed 10 mg/L, EPA maximum contaminant level in drinking water(USEPA, 2002). Decreases in median N concentration and the range of concentration coincide with increases in forested areas within the watershed. Pine

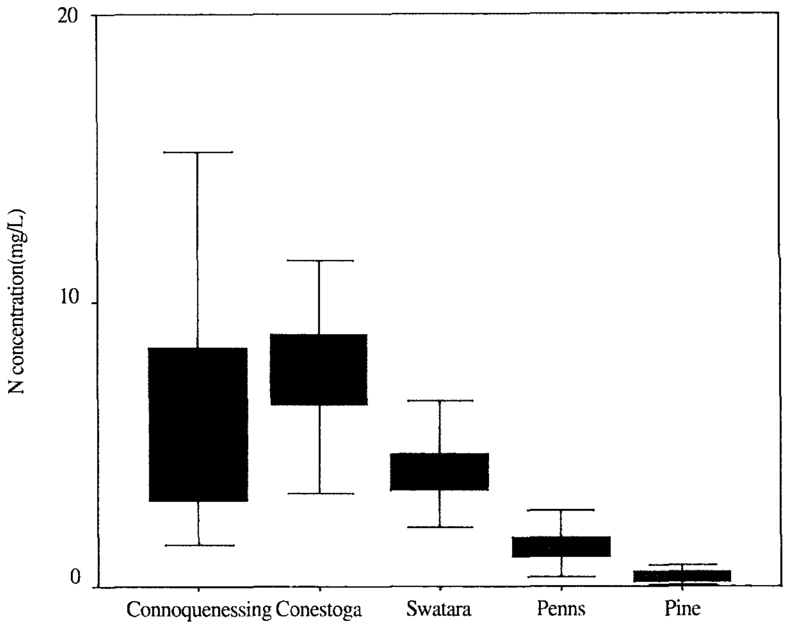

Figure 2. N concentrations for 5 selected watersheds in Pennsylvania, 1990-1996.

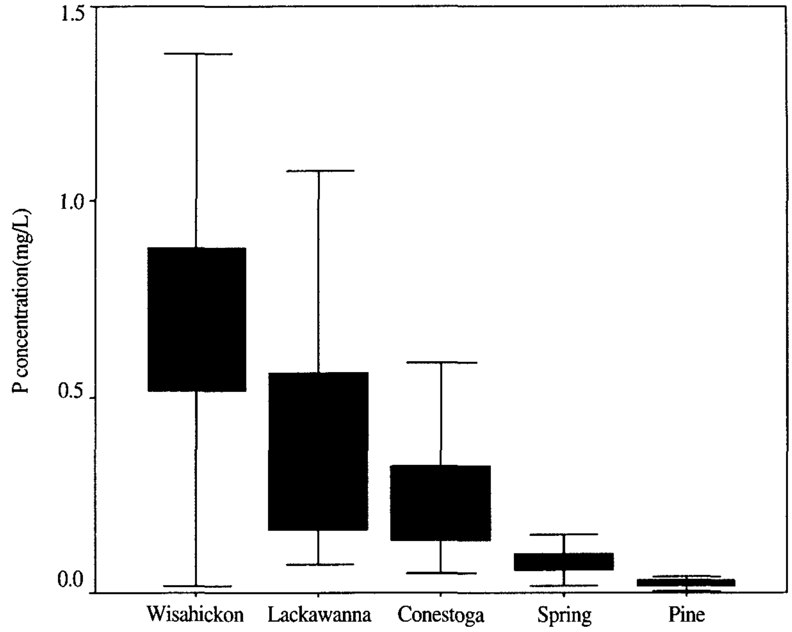

Figure 3. P concentrations for 5 selected watersheds in Pennsylvania, 1990-1996.

Creek, the most forested watershed among the 5 watershed displayed in the box-whisker plot, exhibits the lowest median N concentration and a reduced range N concentration. This suggests that landcover is a reasonable indicator of N concentration at the regional scale.

The range of P concentrations has different patterns across the watersheds. The Wisahickon basin has the highest range of P concentrations(0.05 - 2.34 mg/L), followed by the Lackawanna basin(0.02 - 1.38 mg/L). P concentrations of these two watersheds in most cases exceed 0.1 mg/L, EPA's recommendation to control eutrophication. These two watersheds have a relatively high percentage of urban areas; Wisahickon has 42.3 % and Lackawanna has 13.4 %, respectively. The mean of percent urban area for the entire watersheds is 4.8 %. The next watersheds with a high P range are Conestoga and Pequea watersheds both of which have agricultural land uses. Urban and agricultural watersheds typically exhibit high P concentrations due to fertilizer applied to lawns (Waschbusch et al. 1999) and farmlands; the consistent high P concentrations are also attributed to cow manures widely applied for grass and flowers at most agricultural areas(Hainly and Loper 1997). Once again Pine Creek has the lowest median and range of P concentrations among the five watershed examined.

Not surprisingly, watersheds with higher average nutrient concentrations have a greater range of nutrient concentration. The pattern of variations differs by nutrients, implying that different environmental factors are involved in explaining the spatial variations of N and P concentrations, respectively. In other words, different amounts of natural and human inputs of N and P are becoming important in explaining the different spatial patterns of N and P concentrations. While untreated sewage and fertilizers are common causes of enhanced N and P concentrations, N is additionally derived from animal wastes, N-fixing plants, and atmospheric depositions. The major difference from N is that P is not nearly as soluble. A large amount of P is normally attached to particles of sediment(Davie, 2002).

2) Relationship between nutrient concentrations and environmental variables

Table 2 shows Pearson's correlations of each watershed variable with N and P concentrations, respectively. Land use variables are generally strongly associated with nutrient concentrations. Percentage of wooded land is negatively correlated with both N and P concentration, with a stronger correlation for N concentration than for P concentration. Percentage of developed land is also strongly correlated with nutrient concentration, with a stronger positive correlation for P concentration. Percentage of agricultural land is the second most important variable related to N concentration, while there is a weak correlation between agricultural land and P concentration. Topographic variables - mean slope, standard deviation of slope, mean elevation - are all negatively correlated with N and P concentration. Steeper slopes are typically found in mountainous and hilly watersheds where there is a reduced soil nutrient availability due to lack of human activities including fertilizer applications. Hydroclimatic variables such as runoff, temperature, and runoff ratio are only significantly related to N concentration.

The bivariate analysis suggests that land use and topographic variables are generally important variables for explaining spatial variations of nutrient concentration. As suggested in the literature, watersheds with intensive agricultural land use often have nutrient-rich, drained, tiled and regularly fertilized soils and, thus, increased nutrient leaching(e.g. Andersson, 1986). Seasonal lack of vegetation cover also prohibits the plant uptake of nutrients. For the same reason, in urban areas vegetation uptake and denitrification are less during summer because of the increased built environment. Elevation and slope are negatively associated with N concentrations, indicating that lower hillslopes typically in wet soils contribute to increased N concentration through nitrification(Welsh et al, 2001).

Table 2. Pearson's correlations between watershed variables and nutrient concentrations

Variables measured as	N concentration	P concentration
Developed(%)	0.394*	0.879**
Wooded(%)	-0.768**	-0.420**
Agricultural(%)	0.697**	0.094
Size(km^2)	-0.225	-0.342*
Mean slope(%)	-0.538**	-0.481**
Standard deviation of slope(%)	-0.416**	-0.483**
Mean elevation(m)	-0.674**	-0.514**
Runoff(mm)	-0.519**	-0.218
Temperature(°C)	0.509**	0.305
Precipitation(mm)	-0.003	-0.039
Runoff Ratio(%)	-0.549**	-0.217

** Pearson Correlation is significant at the 0.01 level(2-tailed).
* Pearson Correlation is significant at the 0.05 level(2-tailed).

However, P concentration is not significantly correlated with agricultural land use. P concentration is strongly related to soil erodibility and thus to soil texture, or to in-stream processes, which has not been examined in this study. High P concentrations, typically found in the upper layer of grassland soil, generally decline with increasing soil depth (Heathwaite and Dils 2000). Size, however, becomes an important explanatory variable for P concentration. As the size of the watershed increases, P concentration decreases. It may be possible that bigger watersheds exhibit varying degrees of soil P availablity. Because the source areas of P are often spatially confined into small areas, generally reflecting P status and inputs(Pionke et al., 1997; Gburek and Sharpley, 1998), P delivery per unit area might decrease with increasing watershed size(Prairie and Kalff 1986; Weld et al. 2001).

Table 3 shows the correlation matrix of the 10 variables used for this analysis. Some relationships can be inferred: positive correlation between wooded land and mean elevation(0.744), runoff and wooded land(0.645), wooded land and standard deviation of slope(0.697), wooded land and mean slope(0.727), mean slope and standard deviation of slope(0.944), mean slope and mean elevation(0.661); negative correlation between agricultural land and wooded land(-0.924), temperature and wooded

Table 3. Pearson's correlations(r) of watershed's environmental variables

	Wooded (%)	Agricul-ture(%)	Size (km^2)	Mean Slope	SD of Slope	Elevation (mm)	Runoff (mm)	Temp (°C)	Prcp (mm)
Developed	-.463**	.096	-.339*	-.513**	-.524**	-.560**	-.228	.308	-.001
Wooded		-.924**	.229	.727**	.697**	.744**	.644**	-.540**	-.012
Agriculture			-.114	-.580**	-.539**	-.597**	-.637**	.477**	-.018
Size				.142	.185	.293	.257	-.496**	.038
Mean slope					.944**	.661**	.288	-.325*	-.265
SD Slope						.599**	.303	.307	-.347*
Elevation							.346*	-.502**	-.158
Runoff								-.382*	.408*
Temp									.003

** Correlation is significant at the 0.01 level(2-tailed).

* Correlation is significant at the 0.05 level(2-tailed).

land(-0.544), agricultural land and runoff(-0.637). Because there are multiple colinearities among variables, estimating the direct influence of each variable on N and P concentrations is difficult. To resolve this problem, other statistics are required.

Table 4. Explained variance and cumulative variance by PCA

Component	Eigenvalue	% of variance	Cumulative %
1	4.862	48.618	48.618
2	1.633	16.332	64.950
3	1.234	12.342	77.291

3) Unraveling complex relationships among climate, land use, topography, and water quality

How do these environmental variables affect nutrient concentrations as a group? By reducing or categorizing many variables into a smaller number of meaningful group variables, one can uncover the influence of group variables on N and P concentrations. To answer this question, two different approaches are employed in the current study. The first approach is to sort out relevant variables through Principal component analysis, and the second is to use predefined group variables to identify the joint effects of each group.

(1) Results of Principle component analysis(PCA)

Table 4 summarizes the result of PCA, showing the eigenvalues and percentages of variance explained by the principal components. PCA analysis distinguished three main components that account for variances greater than 1. The three components explain about 77 % of the total variance; PCA axis 1 explains 48.6 % of the data variation, followed by axis 2(16.3 %) and axis 3(12.3 %). PCA axis 1 includes many variables, preventing further understanding of contributing variables. Likewise, PCA axis 2 cannot be easily interpreted. To resolve this problem and achieve a more meaningful representation of the underlying factors, the axes of the principal components were rotated. By rotating these axes, a new set of factors was produced with a minimum overlap between original variables. As a result, the variance explained by each factor was modified, although rotation does not influence the goodness of fit of the principal components(Vega et al. 1998).

The result of the rotated factor matrix is illustrated in Table 5. The Oblimin rotation and Kaiser normalization method was used for this since the extracted PCs might be correlated with each other. PCA analysis shows that principle component(PC) 1 now explains 45.5 % of data variation, followed by component 2(16.4 %), component 3(26.4 %). PC 1(absolute values greater than 1) includes wooded land(+), agricultural land(-), mean slope(+), standard

Table 5. Loadings of environmental variables on three rotated significant principal components for 38 watersheds, PA(Bold indicates values over 0.65.)

	Component 1	Component 2	Component 3
Wooded land	**0.907**	0.211	-0.095
Agricultural land	**-0.901**	-0.354	-0.141
Mean slope	**0.877**	-0.305	-0.033
Standard deviation of slope	**0.840**	-0.354	-0.067
Mean elevation	**0.663**	-0.088	-0.326
Precipitation	-0.182	**0.818**	-0.024
Runoff	0.502	**0.690**	-0.123
Size	-0.209	0.052	**-0.936**
Temperature	-0.233	-0.216	0.632
Developed	-0.249	0.207	0.595
Total variance	4.55	1.64	2.64

Rotation method: Oblimin rotation and Kaiser Normalization

deviation of slope(+), and mean elevation(+), and thus represents the watershed land use and topography. Agricultural land is least likely in areas with high elevation, steep and variable slope, and thus these areas are likely to remain in the forest.

PC 2 exhibits hydroclimate variables, including precipitation(+) and runoff(+). A strong linear negative correlation exists between temperature and elevation; Places in higher elevation tend to be forested rather than agricultural. PC 3 has a negative load of size(-0.936). By employing PCA, it is possible to regroup highly-correlated variables in the same PC, reducing the size of the variables in the further analysis. Note that the first two components explain a lesser amount of variance(61.9 %) than the original two PCA components(64.9 %).

Like other statistical procedures, PCA helps to understand and interpret complex environmental phenomena. The results of PCA suggest that land use and topographic factors are more important than hydroclimate variables in explaining spatial variations of N and P concentrations in PA watersheds. This is due to the fact that climate is relatively uniform(homogeneous) across the watersheds. As Johnes(1999) noted, climate variability plays an important role in the explanation of nutrient export in the UK catchments where mean annual rainfall varies from 640 to 1280 mm. In PA, regional precipitation variation is much smaller, varying less than 350 mm annually.

(2) Results of Partial Redundancy Analysis(RDA)

Figure 4 schematically shows how the variation in nutrient concentrations is partitioned by the three groups of variables representing land cover(L), climate(C) and topography(T). Variations explained by partial RDA are the sum of the joint effects of L, C, and T(LCT), partial joint effects(LC, LT, CT) and unique variations by each environmental group(L, C,

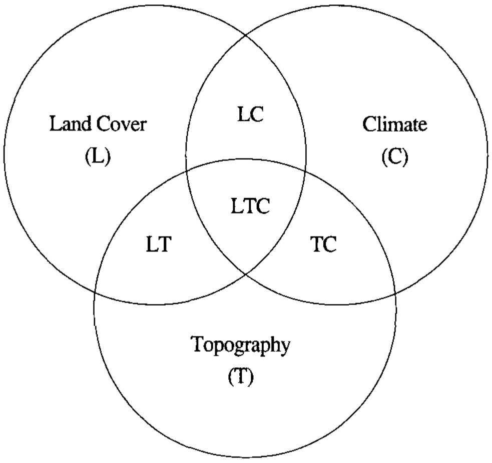

Figure 4. Schematic representation of variation partitioning by land cover, climate, and topography

T). The unexplained variance is calculated by substracting the nutrient variation explained by variation partitioning from the total nutrient variance.

Table 6 illustrates the variation partitioning. First, all three sets - land cover, climate, and topography - are used as explanatory variables to obtain the total amount of variations explained by them. There is no covariable in this analysis. Second, one set is used as an explanatory variable and the other two are used as covariables, and the other way around. The RDA is run four times for each set of combination; total 12 runs of RDA are performed. The second step identifies the pure effects of each set as well as some of the joint effects partitioned from the total variation.

The total amount of variations explained by RDA differs by nutrients. About 62.6 percent of the variations in N concentrations is explained by the land cover, climate and topography, while over 81 % of P variations is accounted for by the three environmental variables. The pure effect of land explains much of the variations in P(41.8 %), but it explains only 5.5 % of N variations. The pure effect of climate is minimal for explaining variations in N concentration(0.9 %), while the pure effect of topography(1.7 %) is minimal in explaining P concentration. The pure effect of land and climate is also greater in P(0.465) than in N(0.172). The pure effect of climate and topography, however, is greater in N(0.033) than in P(0.015). Using the following equations, the total variation is portioned into seven components, corresponding to each section of the Venn diagram shown in Figure 4.

Total explained variance(N) = 0.281(LCT) + 0.013(CT) + 0.149(LT) + 0.108(LC) + 0.011(T) + 0.009(C) + 0.055(L) = 0.626

Total explained variance(P) = 0.135(LCT) 0.043(CT) + 0.241(LT) + 0.006(LC) + 0.017(T) + 0.041(C) + 0.418(L) = 0.815

The main source of N variation is the joint effect of all three groups(LCT), as indicated by the highest value 0.281(45 % of the 0.626). In contrast, the joint effect of all three explains only 13.5 % of the varia-

Table 6. Computational procedure for variation partitioning

Run	Explanatory variable	Covariable	Eigenvalue	
			N	P
Total effect	Topography & Land & Climate	No	0.626	0.815
Combination 1	Topography	Land & Climate	0.011	0.017
	Land & Climate	No	0.615	0.798
Topography and	Land & Climate	Topography	0.172	0.465
Land & Climate	Topography	No	0.454	0.351
Joint effect:Topography ↔	Land & Climate	= 0.615 - 0.172 = 0.454 - 0.111 = 0.443 (N)		
		= 0.798 - 0.465 = 0.351 - 0.017 = 0.334 (P)		
Combination 2	Climate	Land & Topo	0.009	0.041
	Land & Topo	No	0.617	0.774
Climate and	Land & Topo	Climate	0.215	0.676
Land & Topography	Climate	No	0.411	0.139
Joint effect: Climate ↔	Land & Topography	= 0.617 - 0.215 = 0.411 - 0.009 = 0.402 (N)		
		= 0.774 - 0.676 = 0.139 - 0.041 = 0.058 (P)		
Combination 3	Land	Climate & Topo	0.055	0.418
	Climate & Topo	No	0.571	0.397
Land and	Climate & Topo	Land	0.033	0.015
Climate & Topography	Land	No	0.593	0.800
Joint effect: Land ↔	Climate & Topography	= 0.571 - 0.033 = 0.593 - 0.055 = 0.538 (N)		
		= 0.397 - 0.015 = 0.800 - 0.418 = 0.382 (P)		

tions in P concentration. The results of RDA confirm the major role of land cover in explaining the spatial variations in nutrient concentration. Climate has a joint effect with others rather than having its own effect. No joint effect was found between climate and topography in the case of P concentration. While land cover is the main determinant to nutrient concentration, climate and topography accelerate nutrient concentrations through the movement and cycling of nutrients in watersheds.

The unexplained variance is 0.374 for N and 0.185 for P. The greater joint effects in explaining variations in N concentration suggest that the spatial pattern of N concentration is more complicated than that of P. To further explain the spatial variations of N concentration, other environmental factors such as soils and atmospheric depositions need to be incorporated in the analysis.

4) Geographic patterns of N and P concentrations

Results of PCA and RDA suggest that environmental variables employed in the analyses are interrelated with each other, creating complex patterns of nutrient concentrations. The geographic pattern of nutrient concentrations, however, can be identified corresponding to the variations of such environmental variables. As shown in Figures 5 and 6, there is a marked gradient from northwest to southeast, which reflects statewide differences in land cover, topography, and hydroclimatic conditions. Watersheds with similar characteristics with land cover, topography, and hydroclimate show similar levels of nutrient concentrations, indicating a possibility of regionalization of water pollution in PA. Watersheds with higher elevations are hilly and have less agricultural activities, thus lowering nutrient concentrations. Such watersheds generally have a higher ratio of runoff to precipitation, because higher elevation and slope causes respectively orographic precipitation and greater runoff. Higher runoff might also dilute nutrient concentrations in streams. In contrast, watersheds with lower elevations are flat and have more agricultural activities, which result in higher nutrient concentrations. These watersheds have less precipitation and a lower ratio of runoff to precipitation, contributing to increases in nutrient concentration. Statewide patterns of nutrient concentrations can be explained in more detail by analyzing individual watersheds characteristics. Each watershed's soil type, hydrogeology, farm type, and agricultural practices will further account for the spatial variations of nutrient con-

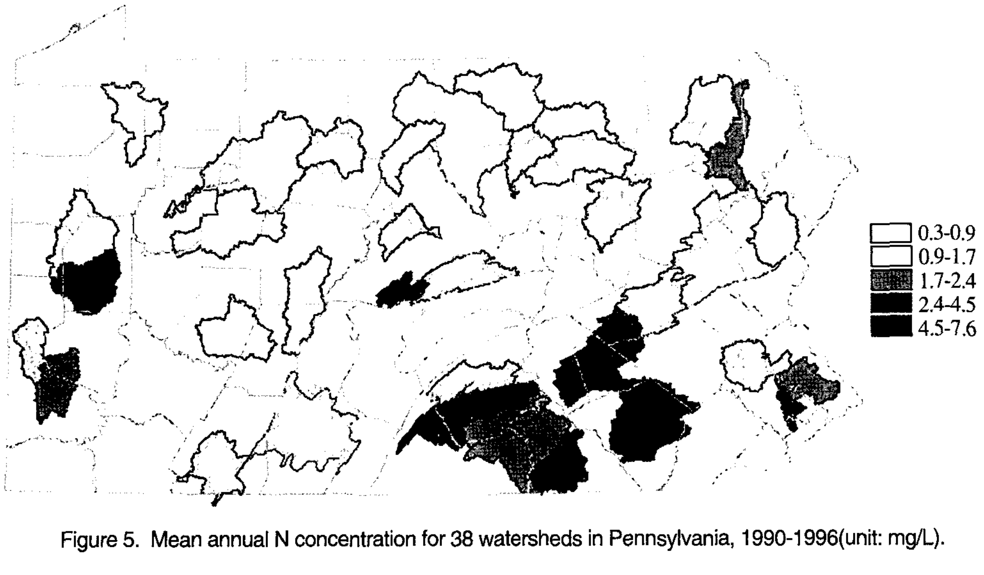

Figure 5. Mean annual N concentration for 38 watersheds in Pennsylvania, 1990-1996(unit: mg/L).

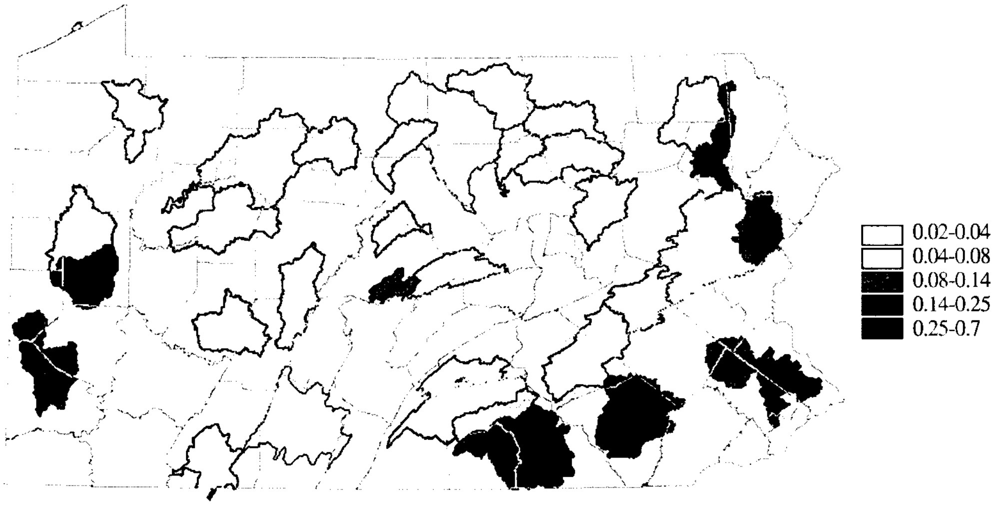

Figure 6. Mean annual P concentration for 38 watersheds in Pennsylvania, 1990-1996(Unit: mg/L).

centrations. In addition, within watersheds differences should be noted to identify any potential local variations of nutrient concentrations in different parts of each watershed.

5. Conclusions

This study examined the spatial variations of nutrient concentrations in 38 watersheds in PA using 7 years (1990-1996) of N and P concentration data. N concentrations are strongly positively correlated with the percentage of agricultural land use(r = 0.7), while P concentrations are more closely associated with the percentage of urban land use in the basin(r = 0.88). This reflects different sources of nutrients; a large proportion of P is derived from industrial and domestic waste discharges as well as agricultural inputs, while intensification of agriculture mostly contributes to enhanced N concentrations.

Watersheds with higher levels of nutrient concentrations also generally have a greater range of variations in concentrations, suggesting that these watersheds are more vulnerable to changes in the environment. It should be noted that there is multicollinearity among the variables used in this study. Low elevation watersheds have generally higher temperature, less precipitation and agricultural or urban land use, and show higher concentrations of nutrients. In contrast, watersheds with higher elevations are mostly likely covered with forests with lower temperature and greater precipitation. PCA analysis identified 3 major components - the land use and topography related, hydroclimate related factors, and size. Results of the partial redundancy analysis show strong joint effects of climate, land cover, and topographic variables for explaining the variations of N concentrations(28.1 %) and P concentrations(13.5 %). A pure effect of land cover explains 41.8 % of P concentrations.

This paper underscores that the spatial variations of N and P concentrations at statewide and regional scales reveal more complex patterns of variation compared with a coarser scale analysis. Marked geographical patterns of N and P concentrations were identified, suggesting that watershed vulnerability to future environmental change will also vary over space as the scale of analysis becomes finer. Finer temporal scale analysis may also help identify dynamics of nutrient concentrations as they are associated with the watersheds' environmental variables. Namely, seasonal or monthly separation may help identify complex relationships among hydroclimate, land use, and water quality in more detail. Clearly, further research is needed to unravel the complex relationship between climate and nutrient concentrations with finer time and spatial scales and different stages of hydrograph(e.g. falling or increasing stage). In addition, the influence of changes in land use/cover along with changes in management practices should also be taken into account for the long-term assessment of nutrient concentrations and loads at the regional scale.

References

Anderson, M. J., & Gribble, N. A., 1998, Partitioning the variation among spatial, temporal, and environmental components in a multivariate data set. *Australian Journal of Ecology*, 23(2), 158-167.

Andersson, R., 1986, *Forluster av kvave och fosfor fran akermark i Sverige.* Unpublished Dissertation, Swedish University of Agricultural Sciences, Uppsala.

Arheimer, B., & Liden, R., 2000, Nitrogen and phosphorus concentrations from agricultural catchments - influence of spatial and temporal variables. *Journal of Hydrology*, 227, 140-159.

Aspinall, R., & Pearson, D., 2000, Integrated geographical assessment of environmental condition in water catchments: Linking landscape ecology, environmental modelling and GIS. *Journal of Environmental Management*, 59, 299-319.

Bis, B., Zdanowicz, A., & Zalewski, M., 2000, Effects of catchment properties on hydrochemistry, habitat complexity and invertebrate community structure in a lowland river. *Hydrobiologia*, 422, 369-387.

Borcard, D., Legendre, P., & Drapeau, P., 1992, Partialling out the Spatial Component of Ecological Variation. *Ecology*, 73, 1045-1055.

Davie, T. 2002, *Fundamentals of Hydrology*, Routledge, London and New York.

EarthInfo Inc. 1998, *EPA STORET 1998 Region 3:1*, Earth Info Inc., Boulder, CO.

Ekholm, P., Kallio, K., Salo, S. Pietilainen,O-P., Rekolainen, S., Lanie, Y., & Joukola, M., 2000, Relationship between catchment characteristics and nutrient concentrations in an agricultural river system. *Water Research*, 34, 3709-3716.

Gibbs, R. J., 1970, Mechanism controlling world water chemistry. *Science*, 170, 1088-1090.

Gburek, W. J. & Sharpley, A. N., 1998, Hydrologic controls on phosphorus loss from upland agricultural watersheds. *Journal of Environmental Quality*, 27, 267-277.

Hainly, R. A. and Loper, C. A. 1997, *Water Quality Assessment of the Lower Susquehanna River Basin, Pennsylvania and Maryland - Sources, Characteristics, Analysis, and Limitations of Nutrient and Suspended-Sediment Data, 1975-90*(U.S. Geological Survey Water-Resources Investigations Report 97-4209).

Heathwaite, A. L., & Dils, R. M., 2000, Characterising phosphorus loss in surface and subsurface hydrological pathways. *The Science of the Total Environment*, 251/252, 523-538.

Hudak, P. F., 2000, Regional trends in nitrate content of Texas groundwater. *Journal of Hydrology*, 228, 37-47.

Johnes, P. J., 1999, Understanding lake and catchment history as a tool for integrated lake management. *Hydrobiologia*, 396, 41-60.

Jongman, R. H. G., ter Braak, C. J. F., and van Tongeren, O. F. R.(eds), 1997, *Data Analysis in Community and Landscape Ecology*, New York: Cambridge University Press.

Liu, Q. H., 1997, Variation partitioning by partial redundancy analysis(RDA). *Environmetrics*, 8(2), 75-85.

Loveland, T. R., & Shaw, D. M., 1996, *Multiresolution land characterization: Building collaborative partnerships.* Paper presented at the Gap Analysis: A Landscape Approach to Biodiversity

Planning, Charlotte, North Carolina.

Meybeck, M., & Helmer, R., 1989, The quality of rivers: from pristine stage to global pollution. *Paleogeography, Paleoclimatology, and Paleoecology*, 75, 283-309.

Miller, C. V. Denis, J. M., Ator, S. W., & Brakebill, J. W., 1997, Nutrients in streams during baseflow in selected environmental settings of the Potomac River Basin, *Journal of American Water Resources Association*, 33, 1155-1171.

Osborne, L. L., & Wiley, M. J., 1988, Empirical relationships between land use/cover and stream water quality in an agricultural watershed. *Journal of Environmental Management*, 26, 9-27.

Pardo, I., 1994, Comparative water quality characterization by PCA of an unperturbed and a polluted stream. *Arch Hydrobiol*, 132, 95-114.

Perona, E., Bonilla, I., & P., M., 1999, Spatial and temporal changes in water quality in a Spanish river. *The Science of the Total Environment*, 241, 75-90.

Peters, N. E., & Meybeck, M., 2000, Water quality degradation effects on freshwater availability: Impacts to human activities. *Water International*, 25(2), 185-193.

Pionke, H. B., Gburek, W. J., Sharpley, A. N., & Zollweg, J. A., 1997, Hydrologic and chemical controls on phosphorus loss from soil to water In H. Tunney(ed.), *Phosphorus Loss from Soil to Water*. Cambridge, England, CAB Int. Press.

Prairie, Y. T., & Kalff, J., 1986, Effect of catchment size on phosphorus export. *Water Research Bulletin*, 22(3), 465-470.

ter Braak, C. J. F., & Šmilauer, P., 1998, *CANOCO Reference Manual and User's Guide to Canoco for Windows. Software for Canonical Community Ordination(version 4)*. Ithaca NY: Microcomputer Power.

United States Environmental Protection Agency, 2002, Drinking water regulations and health advisories, Office of Water, available at http://www.epa.gov/waterscience/drinking/standards/summary.html accessed on Nov 28th, 2002.

Vega, M., Pardo, R., Barrado, E., & Deban, L., 1998, Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research*, 32(12), 3581-3592.

Vidal, M., Melgar, J., Lopez, A. E., & Santoalla, M. C., 2000, Spatial and temporal hydrochemical changes in groundwater under the contaminating effects of fertilizers and wastewater. *Journal of Environmental Management*, 60, 215-225.

Vogelmann, J. E., Sohl, T., & Howard, S. M., 1998, Regional characterization of land cover using multiple sources of data. *Photogrammetric Engineering and Remote Sensing*, 64, 45-57.

Walling, D. E., & Webb, B. W., 1986, Solutes in river systems. In S. T. Trudgill(Ed.), *Solute Processes*(pp. 251-327). Chichester: John Wiley & Sons.

Waschbusch, R. J., Selbig, W. R., & Bannerman, R. T., 1999, *Sources of Phosphorus in Stormwater and Street Dirt from Two Urban Residential Basins in Madison, Wisconsin, 1994-95*, U.S. Geological Survey.

Welsch, D. L., Kroll, C. N., McDonnell, J. J., Burns, D. A., 2001, Topographic controls on the chemistry of subsurface stormflow. *Hydrological Processes*, 15, 1925-1938.

Weld, J., Sharpley, A. N., Beegle, D. B., & Gburek, W. J., 2001, Identifying critical sources of phosphorus export from agricultural watersheds. *Nutrient Cycling in Agroecosystems*, 59(1), 29-38.

Yarnal, B., 1995, Climate of Pennsylvania. In E. W. Miller(Ed.), *A Geography of Pennsylvania*(pp. 44-55). University Park: Pennsylvania State University Press.

(*Received* November 11, 2002)