A REGRESSION-BASED APPROACH TO UNDERSTAND BASELINE TOTAL NITROGEN LOADING FOR TMDL PLANNING

Gerard McMahon
U.S. Geological Survey
3916 Sunset Ridge Road
Raleigh, North Carolina 27607

Chris Roessler
North Carolina Department of Environment and Natural Resources
Raleigh, North Carolina

ABSTRACT

Requirements under Section 303(d) of the Clean Water Act have led North Carolina environmental officials to prepare a total nitrogen (TN) total maximum daily load (TMDL) plan for the entire Neuse River Basin. The SPARROW (SPAtially Referenced Regression On Watershed attributes) watershed model was used to develop baseline estimates of TN inputs and delivery and a TN budget for the Neuse River Basin and two adjoining basins, the Cape Fear and Tar-Pamlico. The model explained 94% of the variability in log-transformed stream TN flux. Estimates of stream yield typically were within 25% of the observed values at the 44 monitoring stations used to calibrate the model. The model indicates that landscape factors, such as soil drainage characteristics, and channel transport factors, such as aquatic processes in streams and reservoirs, both exert a large influence on the transport of TN at both the reach and whole-basin scale. TN losses associated with in-stream processes occurred at a rate of about 5% per kilometer in streams with a mean annual discharge less than 1.04 cubic meters per second; losses in larger streams occurred at a rate of 0.2% per kilometer.

KEYWORDS

Nutrients, watershed models, total nitrogen, geographic information system (GIS)

INTRODUCTION

The Neuse River in North Carolina has been listed as one of the 20 most threatened rivers in the United States because of nitrogen loading and nuisance algal blooms, and the river consistently appears on the State's biennial list of impaired waters (American Rivers, 1997). Requirements under Section 303(d) of the Clean Water Act have led State environmental officials to prepare a TMDL plan for the entire Neuse River Basin (U.S. Environmental Protection Agency, 2001; North Carolina Department of Environment and Natural Resources, 2001). The TMDL plan includes calculations for the maximum amount of TN that the estuary can receive and still meet water-quality standards and allocates this load among point and nonpoint sources.

Copyright ©2002 Water Environment Federation. All Rights Reserved.

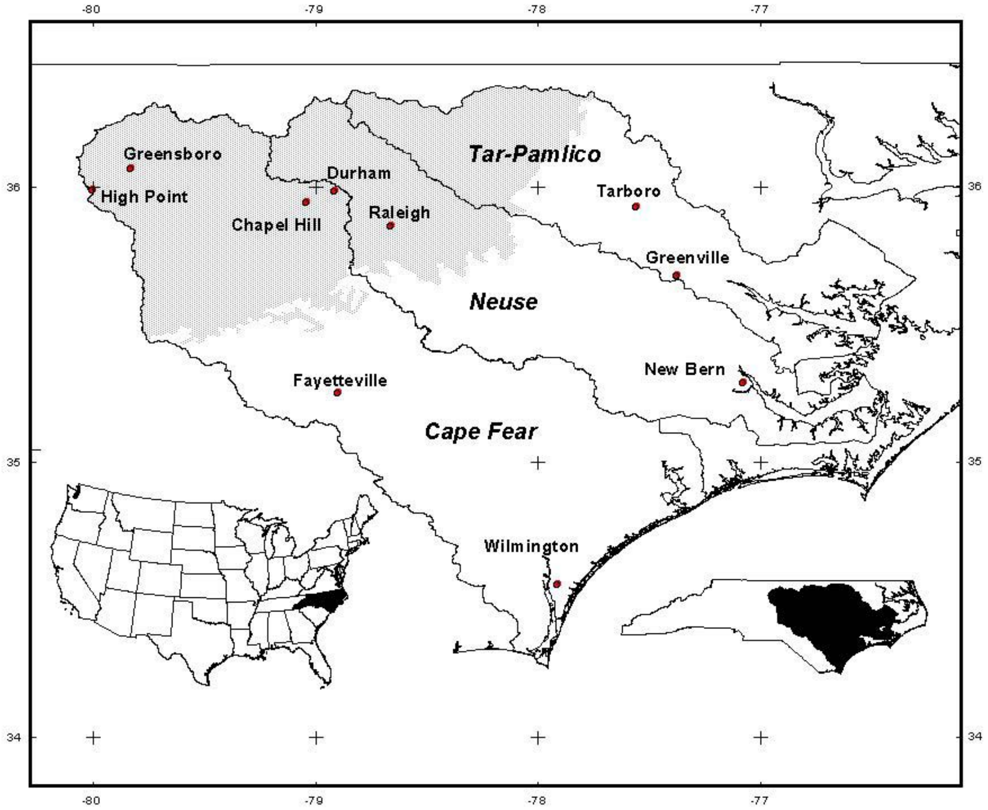
A starting point for the TMDL planning process is the estimation of baseline loading of TN to the Neuse River estuary. Baseline load refers to the TN flux at the Neuse River estuary in the early 1990's, and serves a reference point for TN management activities undertaken during the TMDL process. The State Division of Water Quality calculated a mean annual TN flux for 1991-95 at the outlet of the Neuse River Basin and apportioned this load among point and nonpoint sources (North Carolina Department of Environment and Natural Resources, 2001). Point-source loads were calculated by using data from discharge monitoring reports, and a first-order decay equation was used to simulate the loss of point-source TN between the point where it was introduced into the river and the basin outlet. The total nonpoint source load was estimated as the difference between the overall flux and the portion of point-source TN estimated to reach the outlet point. Baseline nonpoint TN loads were apportioned among four nonpoint-source categories -- urban, agriculture, forest, and open water -- by using export coefficients and land-cover data to estimate the relative proportion of TN load associated with these source categories and allocating the total nonpoint-source load with these proportions.

The initial approach used in the TMDL planning process has several shortcomings. This methodology does not provide information about the uncertainty of the load or input estimates; thus, the accuracy of the load or inputs cannot be estimated, nor can the probability of various load scenarios be predicted. Using a method of difference to estimate nonpoint source inputs at the basin outlet provides no direct information about the amount of nonpoint TN introduced into the system or about TN losses that might occur either due to landscape or aquatic processing of these nonpoint inputs (e.g., atmospheric deposition or agricultural fertilizer). A single estimate of nonpoint-source inputs at the basin outlet cannot provide guidance about where relatively large amounts of nonpoint TN inputs originate within the basin, or the relative importance of various categories of TN inputs throughout the basin. These shortcomings make it difficult to target remedial policies either to locations with relatively high nonpoint loads or to sources that are relatively important.

A nonlinear regression approach was developed to provide a more detailed baseline assessment (Smith et al., 1997; Alexander et al., in press). A SPARROW TN model was calibrated by using data from three river basins in eastern North Carolina (figure 1). The SPARROW modeling approach has three main features. First, a geographic information system (GIS) is used to manage data pertaining to TN sources, in-stream TN flux measured at monitoring sites, characteristics of the terrestrial landscape, and the location and connectivity of stream reaches. Second, the statistical basis of SPARROW provides an objective means of specifying a relation between TN flux and the sources and losses of TN within the watershed. The model specifies that in-stream TN flux is a function of a nonlinear relation between TN sources, such as point sources, atmospheric deposition, agricultural inputs, and landscape and in-stream nitrogen processing. Third, the SPARROW model makes explicit use of information that can be derived from the stream-reach network about the spatial relation among TN fluxes, sources, landscape characteristics, and stream characteristics. The SPARROW modeling approach allows the processing and delivery of nutrients to downstream water bodies to be estimated for separate sources as a function of the location, magnitude and interactions of these sources

Copyright ©2002 Water Environment Federation. All Rights Reserved.

Figure 1 -- North Carolina SPARROW modeling study area, including the Cape Fear, Neuse, and Tar-Pamlico River Basins. Shaded area in main map represents Piedmont ecoregion and unshaded area represents the Coastal Plain ecoregion.



Copyright ©2002 Water Environment Federation. All Rights Reserved.

with the terrestrial and aquatic properties of the river basin. Load estimates and the magnitude of various point and nonpoint sources can be made at the scale of an individual stream reach or for the entire basin. The regression approach allows uncertainty analysis, so that confidence intervals can be estimated for load estimates.

The study described in this paper represents results of a collaboration between the Albemarle-Pamlico study unit of the U.S. Geological Survey (USGS) National Water-Quality Assessment Program (NAWQA) and the North Carolina Department of Environment and Natural Resources. The focus of this collaboration is consistent with a strong recommendation in the National Research Council evaluation of NAWQA that models such as SPARROW be used both to generalize the information developed at a necessarily limited NAWQA sampling network, and to improve the understanding of mechanisms associated with the status and trends results from the first 10-year cycle of NAWQA (National Research Council, 2002). The North Carolina SPARROW effort extends the scope of mass balance work analysis completed during an earlier cycle of NAWQA (McMahon and Woodside, 1997).

In this paper, we examine the use of the SPARROW model to estimate baseline TN loading to the Neuse River and to two adjoining river basins, the Cape Fear and Tar-Pamlico. These basins are considered together because water-quality data and streamflow data from 44 monitoring stations in the 3 basins were used to calibrate the SPARROW model. This combined data set provided a larger number of sites for the model calibration, and also provided a greater range in variability in water-quality conditions than would have been available for the Neuse Basin alone. The results from the Cape Fear and Tar-Pamlico Basins also provide a more regional context for considering the Neuse estimates. We estimate the TN inputs at each stream reach (i.e., where inputs are expressed as reach yields; that is, mass of TN input to the stream per unit area of the reach catchment), the amount of these inputs delivered to each of the three basin outlets, and reach-scale probabilities of exceeding a TN input of 1,000 kilograms per square kilometer (kg/km^2). We also estimate the share of TN loading associated with point sources and with agricultural and atmospheric nonpoint sources and the proportion of these inputs removed by landscape and aquatic processes.

METHODOLOGY

The SPARROW modeling approach uses data pertaining to streamflow, in-stream TN flux at monitoring sites, TN sources, landscape characteristics, and a digital stream-reach network (Preston and Brakebill, 1999). The stream-reach network, composed of individual, hydrologically linked stream reaches, serves as the data organization framework. Streamflow, flux, source, and landscape information are indexed to the stream-reach network. In-stream TN flux is modeled as a nonlinear function of TN sources (including point sources, atmospheric deposition, and agricultural and developed land use), land-delivery processes, and in-stream TN processing.

Copyright ©2002 Water Environment Federation. All Rights Reserved.

Model Definition

The SPARROW model can be defined in the following way:

$$Load_i = \sum_{n=1}^{N} \sum_{j \in J(i)} \beta_n S_{n,j} \exp(-\alpha Z_j) \quad H_{i,j}^{S} H_{i,j}^{R} \;\; \varepsilon_i \qquad \text{(equation 1)}$$

Detailed information about this model form, its assumptions, and applications is available elsewhere (Smith et al., 1997; Preston and Brakebill, 1999; Alexander et al., 2000; Alexander et al., in press). Model terms are described here briefly.

Load ($Load_i$) refers to the nitrogen load or flux in reach *i*, measured in metric tons for the year 1992. *J(i)* is the set of all reaches upstream and including reach *i*, except reaches at or above monitoring stations upstream of reach *i*. The source coefficients (β_n) describe the relation between the *n* TN sources, $S_{n,j}$, and in-stream TN load; together they make up the model's TN source term. Nonpoint TN sources are measured in terms of mass input units. TN sources considered in this model include atmospheric deposition, agricultural inputs, and point sources, all of which are measured in mass input units. The land-to-water delivery coefficients (α) describe the influence of landscape characteristics (Z_j) in the delivery of nonpoint sources of TN to the stream. The term $\exp(-\alpha Z_j)$ can be considered a land-to-water delivery factor. When the model is used to estimate loads, the product of the source term and the land-to-water delivery factor quantifies the yield of TN delivered to the edge of a stream. The model error term, ε_i, is a multiplicative error term assumed to be independent and identically distributed across separate subbasins defined by intervening drainage areas between monitoring stations.

The effects of stream-channel processes, $H_{i,j}^{S}$, such as sedimentation and water column and benthic processing, on the mass of TN lost during transport in the stream is quantified as a stream transport factor that is represented as follows:

$$H_{i,j}^{S} = \prod_{m} \exp(-k_m L_{i,j,m}) \qquad \text{(equation 2)}$$

where k_m is a first-order loss coefficient (units = km^{-1}), *m* is the number of discrete flow classes, and $L_{i,j,m}$ is the length of the stream channel between water bodies *j* and *i* in flow class *m*. When the model is used to estimate loads, the amount of TN reaching the edge of the stream, multiplied by the stream transport factor, quantifies the yield of TN delivered to a reach outlet. In-stream TN losses are assumed to vary as a function of stream channel length in various flow classes. These TN losses, associated with contact and exchange of the water column with the benthic environment, are assumed to decrease as the stream size increases and the exchange between the water column and stream bottom decrease (Alexander et al., 2000).

Copyright ©2002 Water Environment Federation. All Rights Reserved.

Lake and reservoir properties, $H^{R}_{i,j}$, affect the proportion of TN load in water body *j* that is delivered to water body *i* according to the following relation:

$$H^{R}_{i,j} = \prod_{l} \exp(-kq_l^{-1}) \qquad \text{(equation 3)}$$

where *k* is an estimated first-order loss rate (or settling velocity; units = meters (m) per year), q_l^{-1} is the reciprocal areal hydraulic load of lakes and reservoirs (ratio of water surface area to outflow discharge; units = year m^{-1}), and *l* is the number of lakes and reservoirs located between water bodies *j* and *i*. Nitrogen removal by lakes and reservoirs is assumed to be an inverse function of the areal hydraulic load (the ratio of reservoir outflow to surface); that is, TN losses decrease with increases in the rate of reservoir outflow or decreases in reservoir surface area (Alexander et al., in press).

Parametric coefficient estimation for the SPARROW model was performed by applying a nonlinear, least-square algorithm (PROC MODEL; SAS Institute Inc., 1999) on log transformations of the summed quantities in equation 1. Final model coefficient estimates were made by using bootstrap techniques (Smith et al., 1997). Model coefficients were estimated using sampling with replacement from the set of mean annual nutrient fluxes at the 44 monitoring stations in the three North Carolina river basins, based on separate regressions fit for a total of 200 iterations. Ninety-percent confidence intervals for load estimates were determined from the coefficient distributions for the 200 model iterations by computing the minimum range of coefficient values, such that the fraction of values inside the range equaled the confidence levels (Smith et al. 1997).

Data Sources

Ambient monitoring-derived water-quality data and information about TN sources and landscape and flow characteristics at 44 monitoring stations within the three basins were used to calibrate the SPARROW model. Water-quality data were collected by both the USGS and the State of North Carolina; streamflow data were collected by the USGS. The TN loads at the monitoring sites were estimated by using the log-linear regression model presented in Cohn et al. (1989).

The stream network for developing the North Carolina SPARROW model contains 492 reaches, and was based on an enhanced version of the U.S. Environmental Protection Agency's (EPA) River Reach File 1 (RF1), a 1:500,000-scale digital representation of stream networks (DeWald et al., 1985; Nolan et al., 2002). Drainage catchments for each stream reach, used to develop information about nonpoint source and landscape variables, were delineated by using a stream-conditioned, 30-meter resolution digital elevation model (DEM) data set and automated GIS drainage-basin delineation procedures (Hellweger and Maidment, 1997; U.S. Geological Survey, 2001). Estimates of streamflow for each reach were determined through an accumulation of average annual

Copyright ©2002 Water Environment Federation. All Rights Reserved.

runoff, based on data from the period 1951-80, for each reach catchment (Gebert et al. 1987). Stream-channel distances associated with two flow classes were used in the model to estimate in-stream TN losses. The smaller flow class was arbitrarily defined as those reaches representing the lowest 30% of the streamflow values, a value less than 1.04 cubic meters per second (m^3/s); all remaining reaches were assigned to the larger flow class.

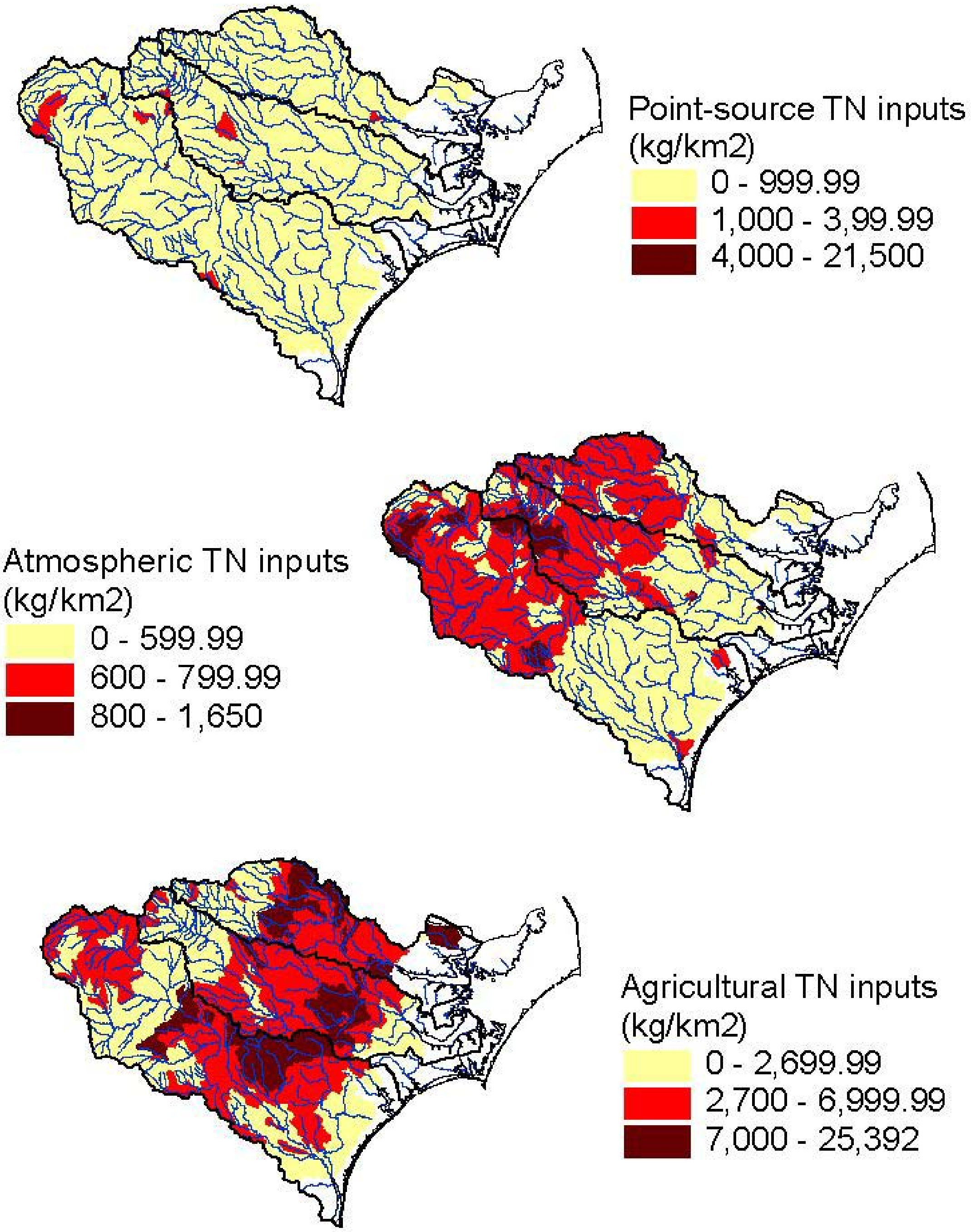
Estimates of 1992 TN discharged by 249 permitted point-source dischargers in the three river basins were made by using data derived from Discharge Monitoring Reports provided by the North Carolina Division of Water Quality (figure 2; Chris Roessler, North Carolina Department of Environment and Natural Resources, written commun., 2001). Load information from individual facilities was allocated to individual reach catchments by using GIS analytical processes. Although most large dischargers are located on the RF1 stream network, many smaller dischargers are not, so that the effects of stream decay on the point sources will be somewhat underestimated.

Estimates of atmospheric TN deposition were developed by using early 1990's wet nitrate and ammonium nitrogen deposition data collected in the National Atmospheric Deposition Program (NADP) (National Atmospheric Deposition Program, 2002), which were interpolated into national atmospheric deposition maps (figure 2; U.S. Geological Survey, 2000). Contributions of TN from atmospheric deposition were calculated as the sum of nitrate nitrogen and ammonium nitrogen wet deposition, nitrate nitrogen dry deposition, and developed-area wet and dry nitrate deposition, following procedures described in Sisterson (1990) and McMahon and Woodside (1997). Spatial data sets for atmospheric wet deposition of nitrate and ammonium nitrogen were overlain with reach catchment boundaries, and individual categories of nitrogen deposition were calculated and summed for each reach.

Estimates of 1992 TN agricultural inputs were derived by using a national land-cover data base developed with remotely sensed data from the early 1990's (Loveland and Shaw, 1996; Vogelmann et al., 2001; U.S. Geological Survey, 2002). Areas of agricultural land in each reach catchment were estimated by using GIS overlay analysis. Estimates of mass TN inputs arising from agriculture within each reach catchment were based on an allocation of county-level agricultural statistics for crops, nitrogen application rates, livestock inventories, animal-waste TN content, and crop harvest TN removal rates (figure 2; College of Agriculture and Life Sciences, 1991; Zublena, 1991; McMahon and Woodside, 1997; U.S. Department of Agriculture, 2001; Deanna Osmond, North Carolina State University, written commun., 2001). County crop and livestock data were allocated according to the proportion of agricultural land in each county within each reach catchment, and were used to calculate reach-level TN input estimates, expressed in mass units.

Copyright ©2002 Water Environment Federation. All Rights Reserved.

Figure 2 -- Point-source, atmospheric, and agricultural total nitrogen (TN) inputs used in calibrating the North Carolina SPARROW model.



Copyright ©2002 Water Environment Federation. All Rights Reserved.

Estimates of fertilizer use were made by multiplying recommended application rates by the planted acres of each crop and summing over all crop types. Estimates of animal waste-related TN were made by multiplying the inventories of each animal type by an annual waste-generation factor for each animal type and summing over all crop types. Estimates of crop harvest-related TN removal were made by multiplying county-level crop harvest data for each crop produced in the county by TN nutrient removal factors associated with the harvested crop, including vegetative matter.

For this analysis, fertilizer application rates for all counties associated with the Neuse River Basin were derived from a North Carolina Agricultural Extension Service survey of nitrogen application rates for major crops (Deanna Osmond, North Carolina State University, written commun., 2001). The extension service survey of recommended rates were used for crops in all other counties. Agricultural TN inputs were calculated as the larger of the crop-fertilizer or animal-waste, minus TN removed by crop harvest. This procedure reflects the assumption provided by the Neuse River TMDL stakeholders group that agricultural nitrogen inputs within a catchment would not exceed the larger of either the amount of animal waste-related nitrogen or the agronomic nitrogen requirements of the crops grown within that catchment (Deanna Osmond, North Carolina State University and Anne Coan, North Carolina Farm Bureau, oral commun., 2001). Two other agricultural input scenarios were considered during model calibration. Inputs were equal to the combined amount of estimated fertilizer and animal waste nitrogen, with nitrogen lost to crop harvest considered in one scenario but not the other. The scenario used in this model, the larger of crop fertilizer or animal waste minus harvest nitrogen removal, was chosen based on a superior statistical fit.

Average values of several soil characteristics were developed for the reach catchments in this study by using data sets associated with the State Soils (STATSGO) Geographic Database (U.S. Department of Agriculture, 1994; Miller and White, 1998; Shirazi et al., 2001a, 2001b). For this model an area-weighted average soil hydrologic group variable was calculated for each reach catchment; lower values refer to less well-drained soils.

MODEL CALIBRATION

The model has three source variables, a single landscape delivery variable, two in-stream loss coefficients, and one reservoir decay rate coefficient and explains approximately 94% of the variation in the natural logarithm of mean annual TN flux, with a mean square error (MSE) of 0.19 (table 1). The median value of the difference between observed and predicted loads is 1.3% and the interquartile range is from -23% to 22%. The plot of predicted and observed yields (figure 3) indicates a tendency for the model to over-predict at smaller yield stations and under-predict at larger yield sites. Model residuals were examined for normality, constant variance, and nonlinear patterns to determine if regression assumptions were satisfied. The model residuals suggest a tendency to under predict in smaller basins, primarily in the Coastal Plain, indicating that the model specification could be improved. The R-squared and MSE of this model are similar to

Copyright ©2002 Water Environment Federation. All Rights Reserved.

SPARROW model fits for the Chesapeake Bay (R^2=0.96 and MSE=0.17, calibrated using 79 stations; Preston and Brakebill, 1999) and the Waikato River Basin in New Zealand (R2=0.97 and MSE=0.14, calibrated using 37 stations; Alexander et al., 2002) and better than a national SPARROW model (R^2=0.88 and MSE=0.43, calibrated using 414 stations across the U.S.; Smith et al., 1997).

The three source coefficients account for the influence of point and nonpoint TN sources on in-stream TN flux. A value of one is expected for the point-source coefficient, since point-source TN is discharged directly to streams and the model accounts for in-stream losses. Although the confidence interval for the point-source coefficient includes one, the fact that the coefficient is less than one suggests that the model representation of TN inputs and attenuation could be improved. The values of the two nonpoint source variables indicate that atmospheric sources contribute a larger share of inputs than do agricultural sources, when statistically controlling for the nitrogen transport losses associated with soils and the stream and reservoir network. Atmospheric inputs and forested land area are highly correlated; neither variable could be included in the final model with signs that made physical sense. In addition, values of the atmospheric deposition variable are higher in urban areas, accounting for the contribution of dry and wet nitrate deposition. Thus, it is probable that the magnitude of the atmospheric coefficient reflects the combined influence of both atmospheric deposition and land cover-related nitrogen sources associated with forested (i.e., especially organic nitrogen) and developed land areas. The source and fate of these atmospheric inputs is further considered in the TN budget discussion below.

TN loads are inversely related to soil drainage characteristics, a finding consistent with other SPARROW models (Smith et al., 1997). Some of the areas with poorly drained soils also have tile drains that might expedite the movement of nitrogen from the land to the edge of the stream. The small and large stream coefficients indicate the first-order rate of TN loss per kilometer of stream length. Losses in small streams (about 5% per kilometer) are more than an order of magnitude larger than losses in large stream (0.2% per kilometer). This inverse relation between in-stream TN loss and stream size is consistent with other SPARROW models. The reservoir settling velocity term of 18.8 meters per year (m/yr) describes the mean water column length from which TN is removed annually.

The 200 sets of bootstrap coefficients were used for error analysis; the bootstrap coefficients in table 1 represent the mean bootstrap coefficient value. Although similar to the parametric estimates, the mean bootstrap coefficients for the two nonpoint-source variables and the landscape delivery variable were slightly smaller than their parametric counterparts, whereas the point source and reservoir terms were somewhat larger. The 90% confidence interval for the two in-stream loss variables includes zero; these variables had very weak statistical significance in the parametric model.

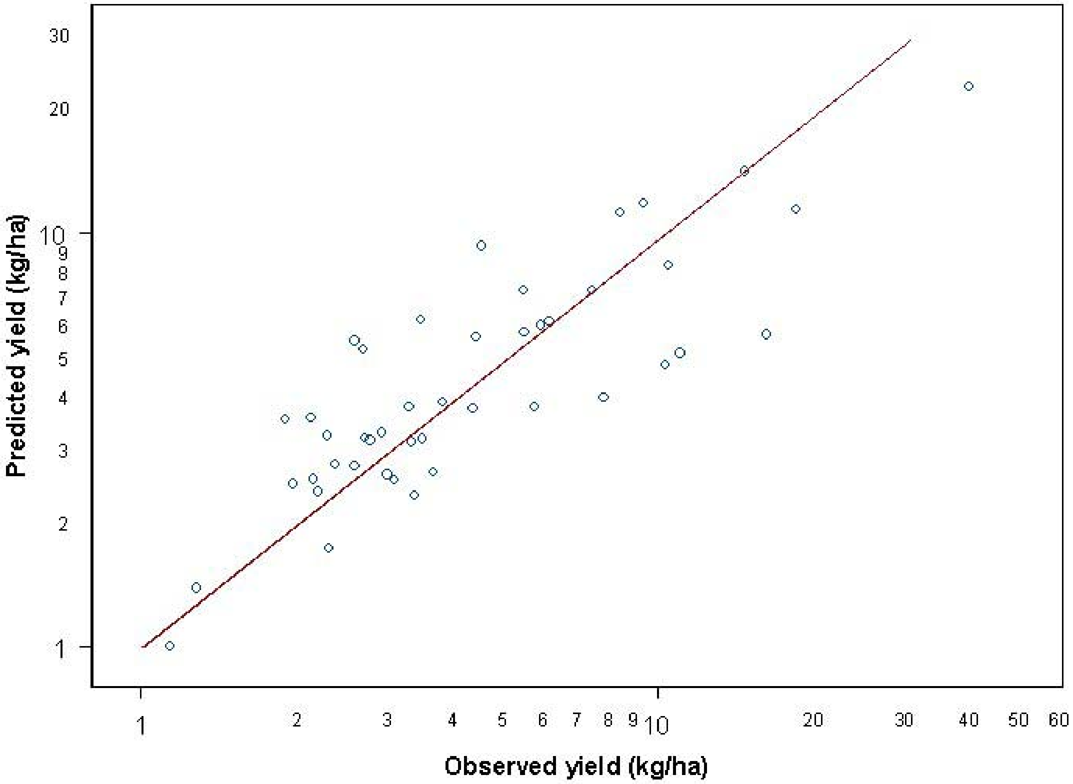
Copyright ©2002 Water Environment Federation. All Rights Reserved.

Table 1 -- Parametric and bootstrap coefficients for North Carolina SPARROW total nitrogen regression model (MSE - mean square error; CI - confidence interval)

	parametric model coefficient estimate	p-value	bootstrap coefficient estimate	lower 90% CI	upper 90% CI
R2	0.94				
MSE	0.19				
TN sources					
Point sources	0.690	0.008	0.702	0.056	1.020
Agricultural inputs	0.430	0.080	0.408	0.060	0.658
Atmospheric inputs	2.800	0.060	2.765	0.001	4.979
Land delivery variable					
Soil hydrologic group	4.500	0.0002	4.072	1.439	5.693
Aquatic loss					
Small stream (km^{-1})	0.050	0.110	0.051	-0.012	0.094
Large stream (km^{-1})	0.002	0.320	0.003	-0.001	0.006
Reservoir (m/yr)	18.800	0.001	20.259	9.196	31.330

Copyright ©2002 Water Environment Federation. All Rights Reserved.

Figure 3 -- Observed and predicted total nitrogen yield (kilograms per hectare, kg/ha) at 44 monitoring sites included in the North Carolina SPARROW model.



Copyright ©2002 Water Environment Federation. All Rights Reserved.

ESTIMATING INPUTS AND TRANSPORT OF TOTAL NITROGEN

Models can be used in developing the detailed baseline assessment information needed to manage impaired waters by providing estimates of where potential contaminant inputs may occur and by estimating the relative importance of different contaminant sources (Smith et al., 1997; National Research Council, 2001). The SPARROW model provides answers to several questions related to the estimation of baseline TN loading, including: Where do TN inputs occur throughout the basin? How much of these inputs make it to the basin outlet? What is the probability that a reach contributes a high level of inputs to the basin? What is the relative importance of different source types? What factors distinguish low and high input basins?

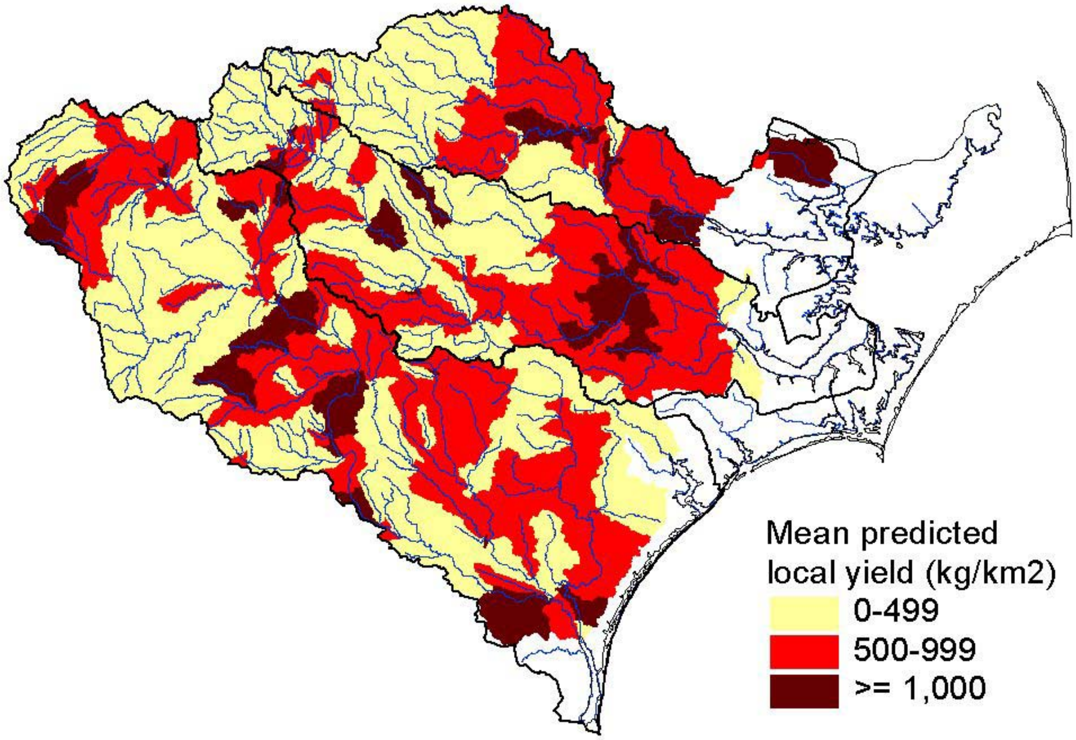
TN Inputs and Delivery

The allocation strategy in the current version of the Neuse River TMDL plan (North Carolina Department of Environment and Natural Resources, 2001) depends on a single baseline estimate of TN load or flux at the mouth of the estuary, an estimate of point-source inputs that is based on discharge records, the use of a first-order loss term to route these inputs to the basin outlet, and an estimate of nonpoint TN sources that is based on subtracting the point-source inputs from the total flux. The SPARROW model provides additional insight into the baseline situation. Because the model is spatially referenced and because it was developed by using stream-reach-specific information, model estimates can be mapped at the reach scale and compiled by geographic regions of interest.

Estimates of local TN inputs were made for each reach catchment (figure 4). Local inputs, expressed as yield (input mass per unit area of the reach catchment), refer to the amount of TN generated within the catchment and delivered to the stream reach, independent of upstream TN inputs. These inputs reflect the land-delivery attenuation associated with the soil hydrologic group characteristics of the catchment. Estimated yields are divided into three classes: low (less than 500 kg/km^2), medium (500-999.99 kg/km^2), and high (greater than or equal to 1,000 kg/km^2) yield reaches. The spatial distribution of local inputs of TN is consistent with the spatial distribution of the three source types (figure 2); areas with lower estimated local yields than might be expected, given the reach source values, reflect the influence of the soil-drainage characteristics of the reach.

Copyright ©2002 Water Environment Federation. All Rights Reserved.

Figure 4 -- Classification of estimated 1992 local total nitrogen inputs in stream reaches of the Cape Fear, Neuse, and Tar-Pamlico River Basins (North Carolina), expressed as local yields.



Copyright ©2002 Water Environment Federation. All Rights Reserved.

A tabular summary of these estimates, including error analysis, provides additional basin-specific information about TN generation (table 2). All three basins have a similar proportion of reaches in the low-yield category (average proportion = 0.58); this proportion is comparable to the proportion of hydrologic units in the southeastern United States estimated to have this same yield (proportion for the southeastern U.S. = 0.57; Smith et al., 1997). The estimated proportion of reaches in this class is more uncertain in the Neuse Basin than in the other two basins. The average proportion of reaches with less than 1,000 kg/km^2 local TN yield across the three basins (0.89) is also comparable to the proportion of hydrologic units in the southeastern U.S. with a similarly sized local TN yield (0.87; Smith et al., 1997). On average, slightly greater than 10% of the stream reaches across the three basins are in the high-yield class; the Tar-Pamlico Basin has the largest proportion of reaches in this class.

Bootstrap-derived coefficients can be used to estimate reach-specific probabilities associated with the high-yield class (figure 5). In the Piedmont ecoregion, reaches with relatively high probabilities of exceeding an estimated local yield of 1,000 kg/km^2 are primarily associated with reach catchments having point-source discharges. In the Coastal Plain, high probability reaches are associated with agricultural inputs.

Delivered-yield estimates indicate the amount of locally generated TN inputs (figure 4) that are delivered to the basin outlet (figure 6). Delivered yields generally are less than input yields, reflecting the influence of TN processing that occurs on the landscape and within streams and reservoirs. Delivered yields tend to be higher for reach catchments with higher inputs and for reaches located near the basin outlet point or along large streams. In each case, TN inputs are exposed to less aquatic processing.

Total Nitrogen Budget

Understanding the baseline TN load in the Neuse Basin requires not only an estimate of the overall inputs and delivery of TN to the basin, but also an allocation of these inputs among the sources of the TN. SPARROW separates the contributions of each of the source terms on both the reach and whole-basin scale. SPARROW also enables an assessment of the effects of land-to-water delivery and channel transport processes by source category.

The spatial distribution of the input shares (reflecting inputs at the edge of the stream) illustrates the relative importance, within the three basins, of each source as an input into the stream system (figure 7). Point sources represent a relatively large share of TN inputs in Piedmont urban areas. Atmospheric sources also are important in urban areas, such as Greensboro, Fayetteville, and Raleigh, because of increased wet and dry nitrate

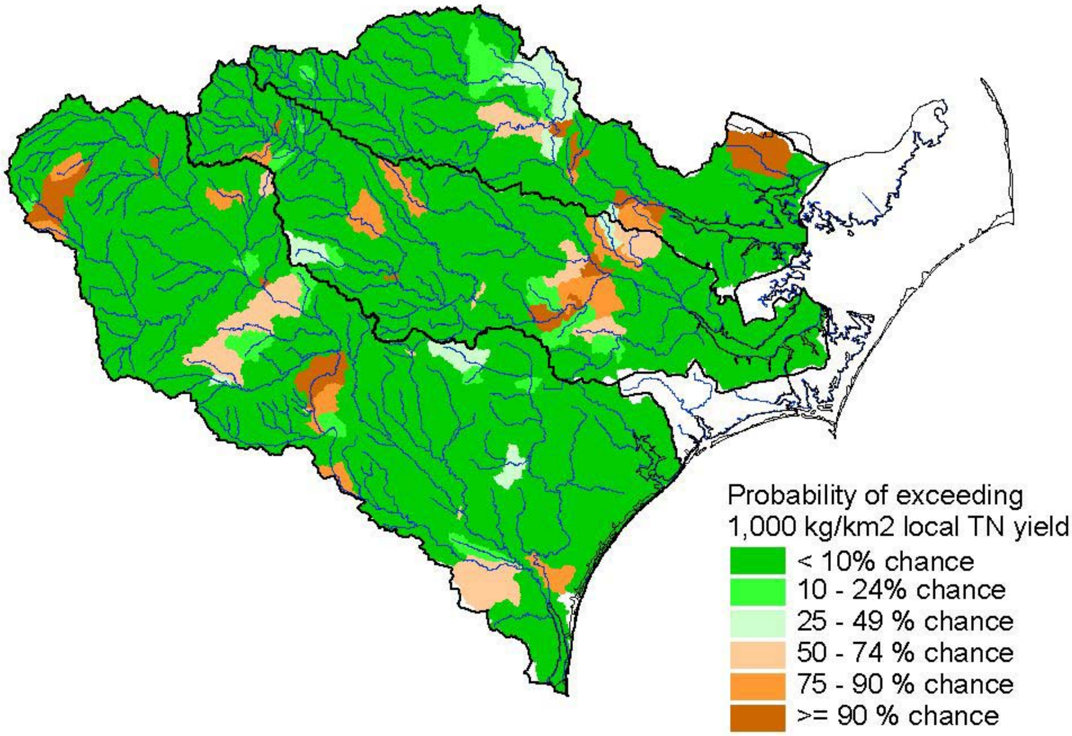
Copyright ©2002 Water Environment Federation. All Rights Reserved.

Table 2 -- Proportion of stream reaches in Cape Fear, Neuse, and Tar-Pamlico River Basins (North Carolina) with estimated 1992 total nitrogen yields of less than 500, less than 1,000, and greater than or equal to 1,000 kg/km^2. Yield reflects nitrogen delivered to the edge of the stream (CI - confidence interval)

	All reaches	Cape Fear	Neuse	Tar-Pamlico
Number of stream reaches	492	226	162	104
Yield < 500 kg/km^2				
Lower 90% CI	0.48	0.46	0.47	0.52
Mean Proportion	0.58	0.56	0.59	0.59
Upper 90% CI	0.69	0.69	0.74	0.65
Yield < 1000 kg/km^2				
Lower 90% CI	0.85	0.87	0.85	0.79
Mean Proportion	0.89	0.91	0.90	0.84
Upper 90% CI	0.93	0.94	0.93	0.89
Yield >= 1000 kg/km^2				
Lower 90% CI	0.07	0.06	0.07	0.11
Mean Proportion	0.11	0.09	0.10	0.16
Upper 90% CI	0.15	0.13	0.15	0.21

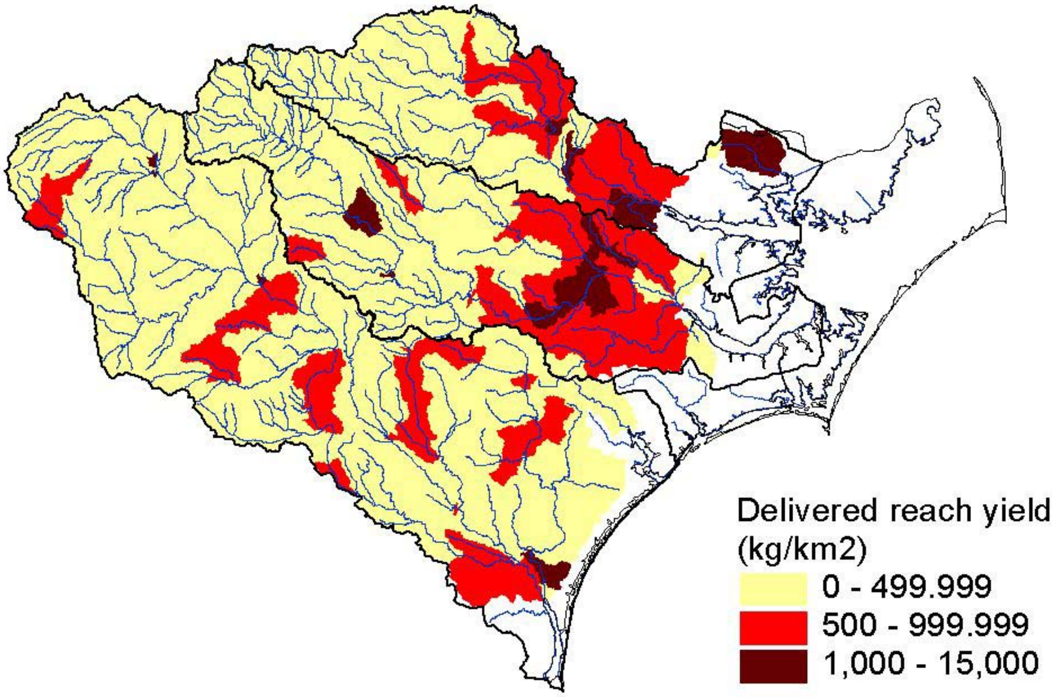
Copyright ©2002 Water Environment Federation. All Rights Reserved.

Figure 5 -- Probability of estimated 1992 local total nitrogen (TN) yield, at edge of stream, exceeding 1,000 kilograms per square kilometer (kg/km^2) sources in the Cape Fear, Neuse, and Tar-Pamlico River, based on analysis of 200 bootstrap estimates of estimated local TN load.



Copyright ©2002 Water Environment Federation. All Rights Reserved.

Figure 6 -- Delivered yield of total nitrogen (TN) in the Neuse, Cape Fear, and Tar-Pamlico River Basins during 1992. Delivered yield is the amount of TN that is generated locally for each stream weighted by the amount of in-stream loss that occurs during transport from the reach outlet to the river basin outlet.



Copyright ©2002 Water Environment Federation. All Rights Reserved.

deposition in urban areas. Generally, atmospheric contributions are larger in the western part of the study area, whereas agricultural shares are higher in the Coastal Plain ecoregion.

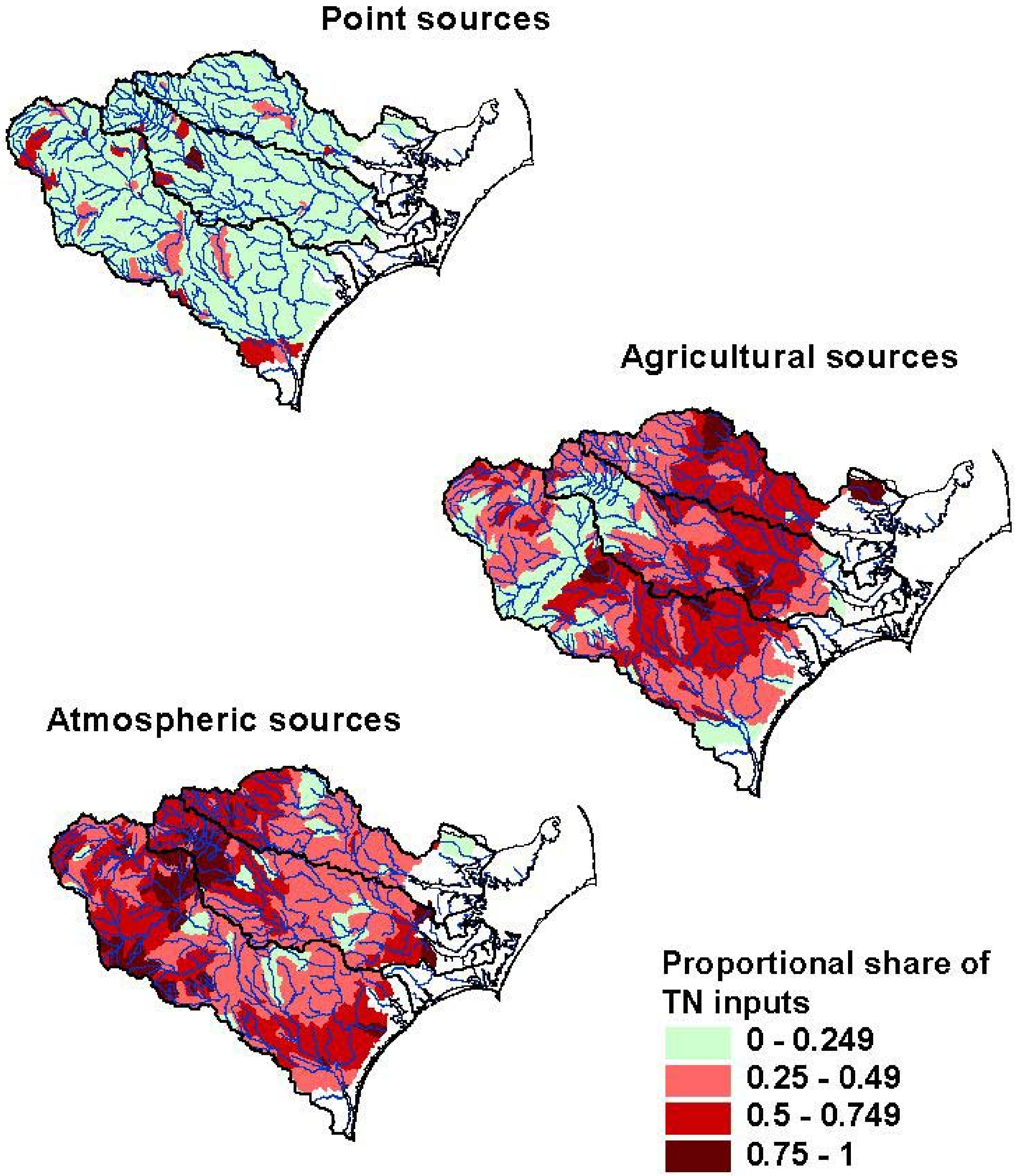
The TN budget also can provide insight on the factors that distinguish reaches with low and high TN yields (figure 8). Atmospheric inputs represent the largest share of TN inputs in all yield classes. The share of point sources changes dramatically from low (1%) to high (20%) yield reaches, indicating the impact of point-source discharges on determining whether a reach will have high local TN yields

Table 3 provides additional information about the sources and transport of nitrogen in the three basins. The first row of Table 3 describes the model estimates of the average share (in percent) of nitrogen inputs at the edge of the stream that can be attributed to each input type across all reaches in the three basins. The second and third rows in Table 3 present the fraction of the nitrogen mass from each source that is transported during land-water delivery and channel transport, respectively. The product of the values of the land-delivery and channel transport factors in each row gives the fraction of each source that is transported over the path from origin on the landscape in a reach catchment to the outlet of the reach catchment. For example, on average an estimated 7% of the agricultural nitrogen inputs in high yield reach catchments is delivered to the edge of the stream, and an estimated 94% of that is transported to the outlet of the reach catchment.

Interestingly, the land-delivery factor for atmospheric sources is approximately 6 times larger than for agricultural sources across all three classes. This difference suggests that the spatial distribution of agricultural inputs may be more closely associated with the spatial distribution of relatively well-drained soils than the atmospheric inputs. The model indicates that such well-drained soils are inversely related to nitrogen reaching the edge of the stream; thus, the loss of nitrogen from agricultural due to land-delivery processes is greater. The channel transport values reflect the spatial distribution of the two nonpoint nitrogen sources vis-à-vis the stream and reservoir network. Channel transport (transport is inversely related to nitrogen losses due to sedimentation and other stream and reservoir loss processes) is smaller for atmospheric sources than for agricultural sources, suggesting that aquatic-processing losses are larger for atmospheric than agricultural sources. Atmospheric inputs are relatively greater in the Piedmont than in the Coastal Plain (figure 2). Because there are a number of larger reservoirs located at or near the division between the Piedmont and Coastal Plain, it might be hypothesized that the larger channel losses associated with atmospheric inputs are due to nitrogen losses in the reservoirs.

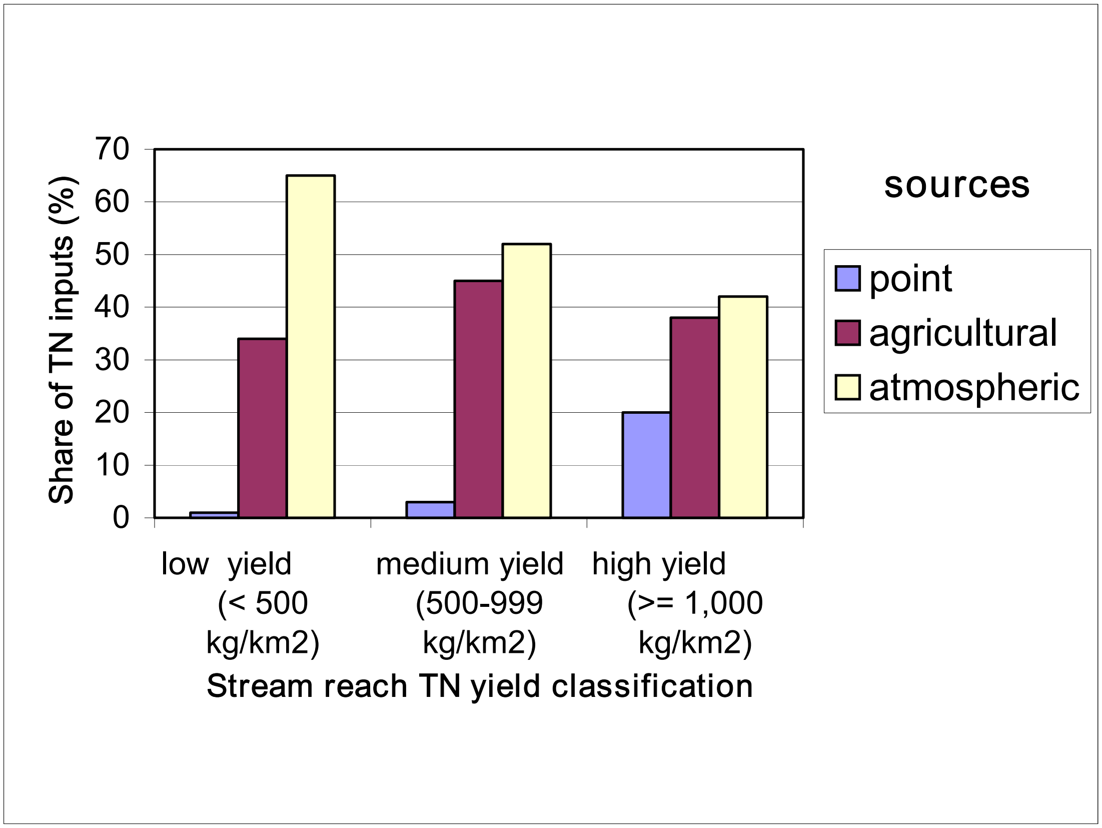
Copyright ©2002 Water Environment Federation. All Rights Reserved.

Figure 7 -- Proportional share of estimated 1992 total nitrogen (TN) inputs, at edge of stream, attributable to point sources, agricultural sources, and atmospheric sources in the Cape Fear, Neuse, and Tar-Pamlico River Basins in North Carolina.



Copyright ©2002 Water Environment Federation. All Rights Reserved.

Figure 8 -- Total nitrogen (TN) budget, by source, for delivery to stream reaches in the Cape Fear, Neuse, and Tar-Pamlico River Basins. Stream reaches classified as low-yield reaches have local TN contributions < 500 kg/km^2 of reach catchment area; medium-yield reaches between 500-999 kg/km^2, and high-yield reaches greater than or equal to 1,000 kg/km^2.



Copyright ©2002 Water Environment Federation. All Rights Reserved.

Table 3 -- Sources and transport factors related to estimated 1992 total nitrogen in the Cape Fear, Neuse, and Tar-Pamlico River Basins in North Carolina

	Point sources	Agricultural Sources	Atmospheric Sources
Share of Total Inputs (%)	5%	39%	56%
Land-Water Delivery Factor	0.70	0.07	0.44
Channel Transport Factor	0.95	0.94	0.86

Copyright ©2002 Water Environment Federation. All Rights Reserved.

CONCLUSIONS

The SPARROW modeling approach provides an objective method for addressing important TMDL-related information needs associated with baseline TN loading. The model provides relatively accurate estimates of annual in-stream loads for the Cape Fear, Neuse, and Tar River Basins. The model and its applications indicate that soil-drainage characteristics and aquatic processes in streams and reservoirs exert a large influence on the transport of TN at both the reach and whole-basin scale.

The location of TN inputs, regardless of the size of the inputs, is an important consideration in assessing baseline TN loads for two reasons. The influence of the land-to-water delivery factor varies spatially; the amount of nonpoint TN introduced in a reach catchment that actually reaches the stream varies accordingly. Because of differences between small and large streams with respect to in-stream TN processing and the direct association between in-stream processing and channel length, the amount of TN inputs in any reach catchment that are transported to the basin outlet point is closely tied to the location of the reach catchment in the stream network.

Several issues are related to the use of this model for baseline TN assessment in the Neuse River TMDL. First, this is one of many SPARROW models that can be specified, and it is likely that results, such as estimated delivered yield, will differ from one model to another. Expert judgment and statistical approaches should be used to continue to improve model specification and to develop procedures to pool information provided by estimates from multiple models. Next, the model has a spatial trend in the residuals, indicating a pattern of under-prediction in the Coastal Plain. Improvements must be made in model specification, perhaps to include a dummy variable to distinguish whether a reach catchment is located in the Coastal Plain or Piedmont ecoregions. Third, the model is calibrated using stream-monitoring data on TN, which was the nitrogen form for which the greatest amount of monitoring information was available. It is possible that some of the stream reaches in forested or swampy areas that have elevated TN concentrations are reaches where the TN concentration might be comprised of relatively high levels of organic nitrogen. Organic nitrogen is not as biologically available as other nitrogen species (e.g., nitrate), so that estimating what may be high levels of organic nitrogen may not be as useful to water-quality managers as estimates for nitrate. Calibrating a SPARROW model for nitrate in North Carolina would require a great deal more monitoring data about nitrate concentrations than is currently available.

Finally, the development of baseline information needed for implementing a TMDL effort can be aided by the use of models, such as SPARROW, that allow information produced by a relatively sparse water-quality monitoring network to be generalized over a larger spatial area of interest. The modeling exercise is complicated, however, by uncertainty inherent in model calibration (e.g., model and parameter uncertainty) and even in model specification. The SPARROW approach is particularly useful in this context. It allows statistical uncertainty to be explicitly incorporated into estimates,

Copyright ©2002 Water Environment Federation. All Rights Reserved.

which allows applications such as mapping the spatial probability distribution of stream reaches with high total nitrogen yields. Multiple model specifications can be explored in a reasonable amount of time, allowing a variety of expert perspectives to be brought to bear on understanding a complex terrestrial-aquatic system. Some measure of uncertainty also exists in estimating all source variables related to land-cover classification accuracy (Yang and others, 2001; Smith et al., 2002), the accuracy of tabular county-level agricultural data, the attribution of these tabular data to specific reach catchments, and the estimation of point-source fluxes using periodic, self-reported data collected from point-source dischargers. These uncertainties are not currently considered in calibrating and applying the SPARROW model. Because the uncertainties associated with estimating reach-level values for these source variables may be large in some cases, it would be useful to include these errors/uncertainties in the calibration and use of future SPARROW models.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the insights and guidance provided by several colleagues, including Richard Alexander, Sandy Cooper, Becky Deckard, Doug Harned, and Tim Spruill.

REFERENCES

Alexander, R.B., Smith, R.A., and Schwarz, G.E. (2000). Effect of Stream Channel Size on the Delivery of Nitrogen to the Gulf of Mexico. *Nature*, v. 403, pp. 758-761.

Alexander, R.B., Elliott, A.H., Shankar, U. and McBride, G.B. (*in press*). Estimating the Sources and Transport of Nutrients in the Waikato River Basin, New Zealand. *Water Resources Research*.

American Rivers (1997). *Most Endangered and Threatened Rivers of North American*, Washington, D.C.

Cohn, T.A., DeLong, L.L., Gilroy, E.J., Hirsch, R.M., and Wells, D.K. (1989). Estimating Constituent Loads. *Water Resources Research*, v. 25, pp. 937-942.

College of Agriculture and Life Sciences. (1991). Chapter 10: Fertilizer use, In *1991 North Carolina Agricultural Chemicals Manual*, College of Agriculture and Life Sciences, North Carolina State University, Raleigh, North Carolina.

DeWald, T., Horn, R., Greenspun, R., Taylor, P., Manning, L., and Montalbano, A. (1985). *STORET Reach Retrieval Documentation*. U.S. Environmental Protection Agency, Washington, D.C.

Gebert, W.A., Graczyk, D.J., and Krug, W.R. (1987). Average Annual Runoff in the United States, 1951-80. *U.S. Geological Survey Hydrologic Investigations Atlas*, HA-710, Reston, Va.

Copyright ©2002 Water Environment Federation. All Rights Reserved.

Hellweger, Ferdi and Maidment, David. (1997). AGREE-DEM Surface Reconditioning System. University of Texas. http://www.crwr.utexas.edu/gis/gishyd98/quality/agree/agree.htm, (August 14, 2001).

Loveland, T.R. and Shaw, D.M. (1996). Multi-Resolution Land Characterization: Building Collaborative Partnerships, In Scott, et al., (eds). *GAP Analysis: A Landscape Approach to Biodiversity Planning*. American Society for Photogrammetry and Remote Sensing, pp. 83-90.

McMahon, G. and Woodside, M.D. (1997). Nutrient Mass Balance for the Albemarle-Pamlico Drainage Basin, North Carolina and Virginia, 1990. *Journal of the American Water Resources Association*, v. 33 No. 3, pp. 573-590.

Miller, D.A. and White, R.A. (1998). A Conterminous United States Multi-Layer Soil Characteristics Data Set for Regional Climate and Hydrology Modeling. *Earth Interactions*, v. 2. http://EarthInteractions.org (February 6, 2002).

National Atmospheric Deposition Program (2002). *National Atmospheric Deposition Program Data Networks*. http://nadp.sws.uiuc.edu/ (February 2, 2002).

National Research Council (2001). *Assessing the TMDL Approach to Water-quality Management*. National Academy Press, Washington, D.C.

National Research Council (2002). *Opportunities to Improve the U.S. Geological Survey National Water-quality Assessment Program*. National Academy Press, Washington, D.C.

Nolan, J.V., Brakebill, J.W., Alexander, R.B., and Schwarz, G.E. (2002). ERF1-2 – Enhanced River Reach File 2. *U. S. Geological Survey Open-File Report 02-40*, Reston, Va. http://water.usgs.gov/GIS/metadata/usgswrd/erf1-2.html (May 3, 2002).

North Carolina Department of Environment and Natural Resources (2001). *Phase II of the total maximum daily load fortotal nitrogen to the Neuse River Estuary, North Carolina*. Division of Water-quality, Raleigh, N.C. December 2001.

Preston, S.D. and Brakebill, J.W. (1999). Application of Spatially Referenced Regression Modeling for the Evaluation of Total Nitrogen Loading in the Chesapeake Bay Watershed. *U.S. Geological Survey Water-Resources Investigations Report 99-4054*, Reston, Va.

SAS Institute, Inc. (1999). *SAS/ETS User's Guide, Version 8*. Cary, N.C.

Shirazi, M.A., Boersman, L., Haggerty, P.K., and Johnson, C.B. (2001a). Predicting Physical and Chemical Water Properties from Relationships with Watershed Soil Characteristics. *Journal of Environmental Quality,* v. 30, pp. 112-120.

Copyright ©2002 Water Environment Federation. All Rights Reserved.

Shirazi, M.A., Boersman, L., Haggerty, P.K., and Johnson, C.B. (2001b). Spatial Extrapolation of Soil Characteristics Using Whole-Soil Particle Size Distributions. *Journal of Environmental Quality,* v. 30, pp. 101-111.

Sisterson, D.L. (1990). Appendix A: Detailed SOx-S and NOx-N Mass Budgets for the United States and Canada. In Venkatram, A., McNaughton, D., Karamchandani, P.K., Shannon, J., Sisterson, D.L., and Fernau, M., *Acidic Deposition: State of Science and Technology, Report 8, Relationships between Atmospheric Emissions and Deposition/Air Quality*, National Acid Precipitation Assessment Program, pp. 8A-1-10.

Smith, J.H., Wickham, J.D., Stehman, S.V., and Yang, L. (2002). Impacts of Patch Size and Land-Cover Heterogeneity on Thematic Image Classification Accuracy. *Photogrammetric Engineering and Remote Sensing*, v. 68 No. 1, pp. 65-70.

Smith, R.A., Schwarz, G.E., and Alexander, R.B. (1997). Regional Interpretation of Water-Quality Monitoring Data. *Water Resources Research*, v. 33, pp. 2781-2798.

U.S. Department of Agriculture (1994). State Soil Geographic (STATSGO) Data Base Data Use Information. *U.S. Department of Agriculture-Natural Resources Conservation Service Miscellaneous Publication 1492*.

U.S. Department of Agriculture (2001). *National Agricultural Statistics Service*. http://www.usda.gov/nass/ (February 1, 2002).

U.S. Environmental Protection Agency (2001). *Total Maximum Daily Load (TMDL) Program*. http://www.epa.gov/OWOW/tmdl/index.html (February 4, 2002).

U.S. Geological Survey (2000). *Digital Data Used to Relate Nutrient Inputs to Water Quality in the Chesapeake Bay Watershed, Version 2.0--Atmospheric Deposition*. http://md.water.usgs.gov/gis/chesbay/sparrow2/doc/atdep92.htm. (February 1, 2002).

U.S. Geological Survey (2001). *The National Elevation Dataset*. http://edcnts12.cr.usgs.gov/ned/factsheet.htm (February 1, 2002).

U.S. Geological Survey (2002). *National Land Cover Characterization Project*. http://landcover.usgs.gov/nationallandcover.html (February 1, 2002).

Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., Van Driel, N. (2001). Completion of the 1990s National Land Cover Data Set for the Conterminous United States from Landsat Thematic Mapper Data and Ancillary Data Sources. *Photogrammetric Engineering and Remote Sensing*, v. 67, pp. 650-652.

Copyright ©2002 Water Environment Federation. All Rights Reserved.

Yang, L., Stehman, S.V., Smith, J.H., and Wickham, J.D. (2001). Thematic Accuracy of MRLC Land Cover for the Eastern United States. *Remote Sensing of the Environment*, v. 76, pp. 418-422.

Zublena, J.P. (1991). *Soil Facts: Nutrient removal by Crops in North Carolina, AG-439-16*, North Carolina Cooperative Extension Service, Raleigh, North Carolina.

Copyright ©2002 Water Environment Federation. All Rights Reserved.