Support of Total Maximum Daily Load Programs Using Spatially Referenced Regression Models

Gerard McMahon[1]; Richard B. Alexander[2]; and Song Qian[3]

Abstract: The spatially referenced regressions on watershed attributes modeling approach, as applied to predictions of total nitrogen flux in three North Carolina river basins, addresses several information needs identified by a National Research Council evaluation of the total maximum daily load program. The model provides reach-level predictions of the probability of exceeding water-quality criteria, and estimates of total nitrogen budgets. Model estimates of point- and diffuse-source contributions and nitrogen loss rates in streams and reservoirs compared moderately well with literature estimates. Maps of reach-level predictions of nutrient inputs and delivery provide an intuitive and spatially detailed summary of the origins and fate of nutrients within a basin.

DOI: 10.1061/(ASCE)0733-9496(2003)129:4(315)

CE Database subject headings: Nitrogen; Regression models; North Carolina; River basins; Nutrients.

Introduction

Despite the expenditure of hundreds of billions of dollars over the last 30 years, the 1972 Clean Water Act goals of fishable and swimmable waters have not been achieved, largely because contaminants from diffuse sources have not been controlled successfully (Knopman and Smith 1993; National Research Council 2001). Almost 500,000 river and shore kilometers (km) and 2 million lake hectares (ha) have been identified as impaired in terms of water-quality criteria intended to assess compliance with designated uses (National Research Council 2001).

Under requirements of Section 303d of the Clean Water Act, the states must identify and list impaired water bodies and implement a total maximum daily load (TMDL) program for water bodies where a problem contaminant can be identified (National Research Council 2001). The objective of a TMDL program is to improve water quality through control of both point- and diffuse-contaminant sources. Initially, a model is used to determine the total daily pollutant load that can exist and still meet water-quality standards for a problem contaminant, such as nitrogen. The maximum load (i.e., TMDL) is then allocated for point and diffuse sources of nitrogen and an estimated margin of safety (U.S. Environmental Protection Agency 2001). A margin-of-safety estimate is required to assure that an impaired water body will meet its designated uses once the prescribed load reductions are realized; the estimate is based on uncertainties in the model, observed data, and natural variability. Successful implementation of a TMDL program for a water body requires that point- and diffuse-source loading reductions be sufficient to bring the water body into compliance with water-quality standards.

The U.S. Environmental Protection Agency (USEPA) estimates that approximately 4,000 TMDLs must be developed per year over the next 8–13 years to meet current legal deadlines (National Research Council 2001). Many, if not most, states claim that they lack adequate resources to assess water-quality conditions statewide, identify waters not meeting standards, and develop TMDLs. These claims were supported by a General Accounting Office (2000) report describing a pervasive lack of data at the state level to complete these tasks.

Because of concern about the inability of states to comply with Clean Water Act requirements related to TMDL programs, Congress asked the National Research Council (NRC) in 2001 to assess the scientific basis of TMDL programs with respect to the availability and reliability of information required to identify contaminant sources and allocate reduction in contaminant loadings. An eight-member NRC committee provided Congress with multiple recommendations related to the use of models to identify impaired waters and determine contaminant loads that with an adequate margin of safety would not violate state water-quality standards (National Research Council 2001). The NRC committee recommended that models associated with TMDL development incorporate the theoretical characteristics of mechanistic models while including the simpler calibration and use characteristic of empirical models. Modeling efforts should be accompanied by uncertainty analysis, replacing the ad hoc and even arbitrary safety margins of many previous TMDLs with margins of safety based on the explicit analysis of model and data uncertainty and natural variability. Finally, the committee recommended an expanded role for models in specific TMDL activities, such as predicting impairment probabilities and designing monitoring programs. For example, models could be used to develop preliminary impairment estimates for all water bodies in a basin or state, allowing the state to focus limited monitoring resources in areas where water-quality conditions are more uncertain.

In this paper, each of the modeling recommendations by the NRC committee is addressed using a North Carolina example. A spatially referenced regression model (SPARROW—spatially ref-

[1]Geographer, U.S. Geological Survey, 3916 Sunset Ridge Rd., Raleigh, NC 27607. E-mail: gmcmahon@usgs.gov

[2]Hydrologist, U.S. Geological Survey, MS 413, 12201 Sunrise Valley Dr., Reston, VA 20192.

[3]The Cadmus Group, 6330 Quadrangle Dr., Suite 180, Chapel Hill, NC 27517.

Note. Discussion open until December 1, 2003. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on June 4, 2002; approved on August 2, 2002. This paper is part of the ***Journal of Water Resources Planning and Management***, Vol. 129, No. 4, July 1, 2003. ©ASCE, ISSN 0733-9496/2003/4-315–329/$18.00.

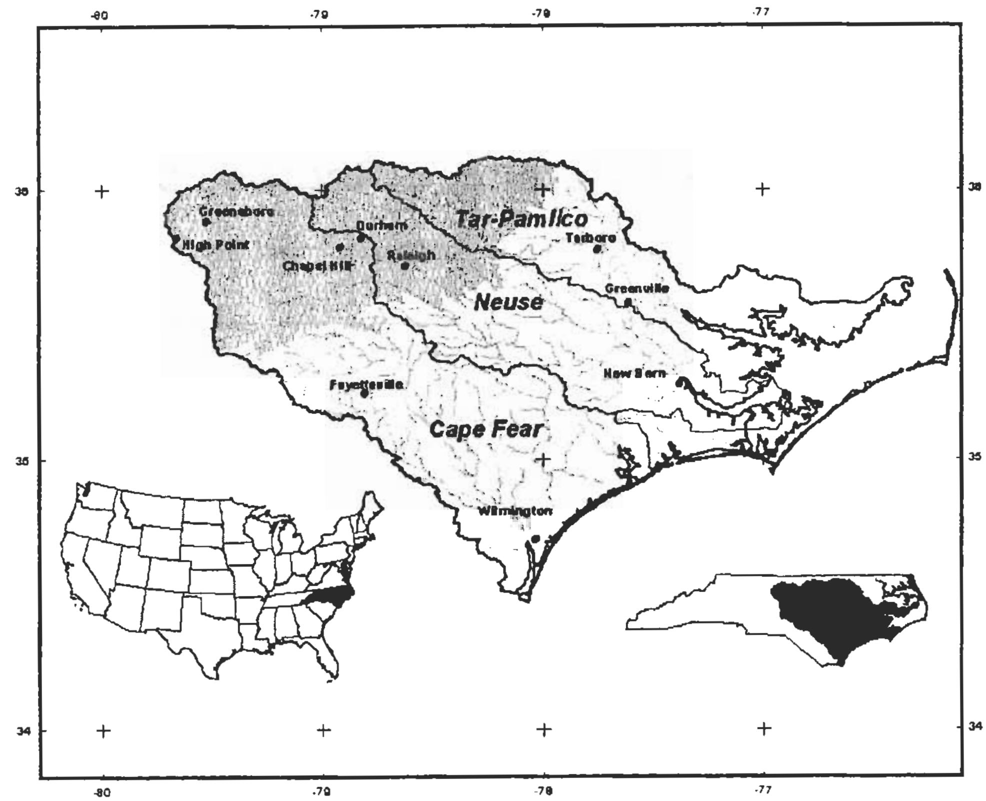

Fig. 1. North Carolina SPARROW modeling study area. Shaded area in main study area map represents Piedmont ecoregion and unshaded area represents Coastal Plain ecoregion.

erenced regressions on watershed attributes; Smith et al. 1997) that contains process-oriented terms for nitrogen transport can be used to predict the probabilities for concentrations of total nitrogen exceeding an environmental guideline (e.g., 1.5 mg/L), assess the movement and delivery of nitrogen to a watershed outlet, and develop a total nitrogen budget that quantifies aquatic losses and the contributions of specific nitrogen sources to streams. A SPARROW total nitrogen model was calibrated using data from three large river basins in eastern North Carolina—the Cape Fear, Neuse, and Tar-Pamlico (Fig. 1). The development and use of digital spatial and tabular data sets needed to support the model are described, and results of the model calibration are presented and discussed. Three model applications are described. Parameter and model uncertainty information are used to predict the proportion of stream reaches (i.e., relatively short, averaging 10 km in length in this study, stream or reservoir segments) in the three basins with an average annual total nitrogen concentration greater than the hypothetical total nitrogen criterion of 1.5 mg/L. Prediction uncertainty information is used to identify locations where additional water-quality monitoring may be useful. The relative importance of individual total nitrogen sources and the effects of landscape and aquatic processes are presented.

Spatially Reinforced Regressions on Watershed Methodology

The SPARROW modeling approach has several important features (Preston and Brakebill 1999). Data for streamflow, in-stream nitrogen flux at monitoring sites, nitrogen sources, and landscape characteristics are spatially linked to a digital stream-reach network. The stream-reach network (Nolan et al. 2002), composed of individual, hydrologically linked stream reaches, provides a spatial framework for organizing and relating streamflow, flux, source, and landscape information. This network allows the processing and delivery of nutrients to downstream water bodies to be simultaneously estimated for separate sources as a function of the location, magnitude, and interactions of these sources with the terrestrial and aquatic properties of the river basin. The statistical basis for calibrating SPARROW models provides an objective means of empirically estimating the relation between in-stream measurements of nitrogen flux and the sources and losses of nitrogen within the watershed. In-stream nitrogen flux is modeled as a nonlinear function of nitrogen sources (including point sources, atmospheric deposition, and agricultural and developed land use), land-delivery processes, and in-stream nitrogen processing.

The SPARROW model expresses mean annual nitrogen load at monitoring Station X (Fig. 2), located at the downstream end of water-body reach i, as a nonlinear function of monitored [e.g., the load estimated at the upstream monitoring station (Station Y)] and unmonitored nitrogen sources [point and diffuse nitrogen sources introduced between the upstream monitoring location (Station Y and monitoring Station X)] and the attenuation associated with landscape and aquatic processes. Unmonitored sources are introduced into reach i (or, in the case of diffuse sources, into the drainage area or catchment of reach i) or one of the set of reaches $J(i)$ (or the associated reach catchments) that are upstream from reach i and that exclude reaches draining into upstream monitoring stations (Smith et al. 1997; Alexander et al. 2002). Diffuse-source nitrogen is attenuated as it moves across the landscape to the stream's edge. Additional attenuation occurs as nitrogen

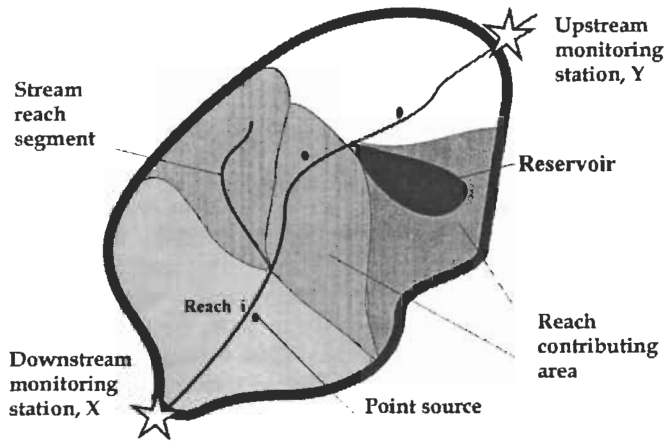

Fig. 2. Hypothetical set of five stream reaches that make up set of reaches $J(i)$ and associated reach catchments. These reaches are located between upstream monitoring station and downstream monitoring station.

moves through the reservoir and stream network to a downstream monitoring station.

The model can be defined as follows:

$$\text{Load}_i=\sum_{n=1}^{N}\sum_{j\in J(i)}\beta_n S_{n,j}e^{(-\alpha Z_j)}H_{i,j}^{S}H_{i,j}^{R}\varepsilon_i \tag{1}$$

where Load_i = nitrogen load or flux in reach i, measured in metric tons for the year 1992; n, N = source index where N is the total number of individual n sources; $J(i)$ = set of all reaches upstream and including reach i, except reaches at or above monitoring stations upstream from reach i; β_n = estimated source coefficient for source n; $S_{n,j}$ = nitrogen mass from source n in drainage to reach j; α = estimated vector of land-to-water delivery coefficients; Z_j = land-surface characteristics associated with drainage to reach j; $H_{i,j}^{S}$ = fraction of nutrient mass present in water body j transported to water body i as a function of first order loss processes associated with stream channels; $H_{i,j}^{R}$ = fraction of nutrient mass present in water body j transported to water body i as a function of first order loss processes associated with lakes and reservoirs; and ε_i = multiplicative error term assumed to be independent and identically distributed across separate subbasins defined by intervening drainage areas between monitoring stations.

Details about the form of this model, its assumptions, and applications are available elsewhere (Smith et al. 1997; Preston and Brakebill 1999; Alexander et al. 2000; Alexander et al. 2002). Model terms are described briefly here. The source coefficients (β_n) describe the relation between nitrogen sources S_{nj} and in-stream nitrogen load. Nitrogen sources considered in the North Carolina SPARROW model include atmospheric deposition, developed and agricultural land use area (i.e., which represent diffuse sources such as fertilizer and animal waste), and point sources. The land-to-water delivery coefficients (α) describe the influence of landscape characteristics in the delivery of diffuse sources of nitrogen to the stream. Because these factors are assumed to decrease the delivery of nutrients to the stream (i.e., α must be negative), the reciprocal value is used in calibrating the model for any landscape features thought to increase the delivery of nutrients from the landscape to the stream (e.g., slope; Smith et al. 1997). Landscape-delivery features considered in the North Carolina model include topographic and soil characteristics. Because point sources discharge directly to streams and are assumed to be unaffected by landscape-delivery factors, the landscape-delivery coefficient for point sources is set equal to one. Nitrogen loads measured at monitoring stations upstream from the monitoring station at reach i also have a landscape-delivery coefficient constrained to one. Thus, nitrogen loads associated with the monitoring station at reach i include nitrogen loads measured at upstream monitoring stations and nitrogen associated with sources located in the catchments of all reaches between the upstream monitoring stations and reach i. Nitrogen delivered to the edges of streams from diffuse sources within these catchments is reduced by the landscape-delivery coefficient, a single coefficient that applies to all diffuse sources.

A detailed, mechanistic specification of nitrogen-loss processes associated with streams, lakes, and reservoirs in the North Carolina SPARROW model cannot be accomplished because of the large land area to be modeled and the difficulty in accurately characterizing these processes over a large area using limited data. A simpler approach is used here that is amenable to an empirical treatment yet reflects the influence of important mechanisms of nitrogen processing. In-stream nitrogen losses are assumed to vary as a function of stream channel length in various flow classes. In-stream nitrogen 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 decreases (Alexander et al. 2000). Stream size is assumed to be associated with stream flow; streams with a low-mean annual flow are assumed to have greater nitrogen loss than streams with larger flows.

The effects of stream-channel processes, such as sedimentation and water column and benthic processing, on the mass of nitrogen lost from the location where the load enters a water body (whether from upstream monitoring stations or from intervening stream reaches) until the load reaches a downstream monitoring location is quantified as

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

where k_m = first-order loss coefficient (km^{-1}); m = number of discrete flow classes; and $L_{i,j,m}$ = length of the stream channel between water bodies j and i in flow class m. A k value of 0.08, for example, indicates that nitrogen is removed at a rate of about 8% per kilometer of channel length. Two stream-size classes were used in the final North Carolina SPARROW model, with stream-size class affiliation for all stream reaches being determined by whether mean annual flow is above or below 1.04 m^3/s. This rate was arbitrarily chosen to define the break between small and large streams; approximately one-third of all stream reaches have a mean annual discharge less than this rate.

Lake and reservoir properties affect the proportion of nitrogen load in water body j that is delivered to water body i according to the relation

$$H_{i,j}^{R}=\prod_{l}\exp(-kq_l^{-1}) \tag{3}$$

where k = estimated first-order loss rate (or settling velocity; units = m/year); q_l^{-1} = reciprocal areal hydraulic load of lake or reservoir (ratio of water-surface area to outflow discharge; units = year/m) for each of the lakes and reservoirs (l) 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

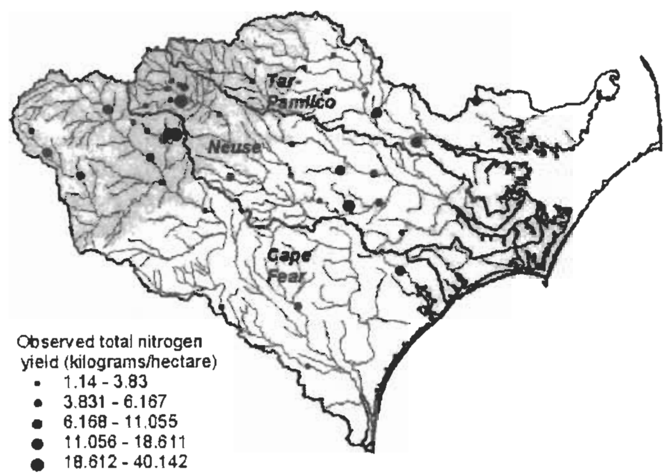

Fig. 3. Monitoring stations used in calibrating North Carolina SPARROW model. Shaded area in main study area map represents Piedmont ecoregion and unshaded area represents Coastal Plain ecoregion.

ratio of reservoir outflow to surface); that is, nitrogen losses decrease with increases in the rate of reservoir outflow or decreases in reservoir surface area (Alexander et al. 2002). The reservoir/lake loss rate is assumed to be a net rate, accounting for inputs from nitrogen fixation.

Coefficient estimation for the SPARROW model was performed by applying a nonlinear, least-squares algorithm (PROC MODEL; SAS Institute Inc. 1999) to Eq. (1). The calibration used a reach-specific data set containing values for all source and landscape terms associated with the catchment of each reach in the stream network of the three North Carolina river basins, along with stream (class and length) and reservoir (areal hydraulic load) data for each reach with the stream network. Model fitting was performed on the log transforms (natural log) of the summed quantities in Eq. (1). Model residuals were examined for normality, constant variance, and nonlinear patterns to determine if regression assumptions were satisfied. Final model coefficient estimates were made using bootstrap techniques (Smith et al. 1997). Model coefficients were estimated with 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 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). The upper and lower values of this range established the confidence intervals for the coefficient.

Data Sets

Data from three contiguous North Carolina river basins were used in this study. A general west-east trend occurs in the size of streamflow and nitrogen loads in the mainstem river reaches. Nitrogen yields (Fig. 3), however, do not show this same spatial pattern and are often largest at monitoring stations downstream from large point sources (especially in the Piedmont) or intense agricultural activity (in the Coastal Plain). Ambient monitoring-derived water-quality data and information about nitrogen sources and landscape and flow characteristics at 44 monitoring stations within the basin were used to fit the SPARROW model (Fig. 3). Source, landscape, and streamflow data were used to apply the model throughout the three basins.

Calculation of Nitrogen Flux at Monitoring Stations

Mean annual nitrogen flux, the dependent variable in the SPARROW model, was estimated for 44 long-term monitoring sites based on water-column measurements of total nitrogen and continuous-flow measurements (Fig. 3). The length of the flow record varies from station to station. Most stations have flow measurements from the mid-1970s to 2000. Some stations have flow measurements dated in the 1920s. Nitrogen concentration values were available for most of the 44 stations for the period of 1990–1995. Two stations have nitrogen data from 1993 to 1995. The nitrogen loads at the monitoring sites were estimated by using the log-linear regression model presented in Cohn et al. (1989).

Streams and Reservoirs

Spatial data for stream reaches and for the catchments, or drainage areas, associated with each reach served as the primary framework for organizing source, landscape, and stream-size information. The stream network for developing the North Carolina SPARROW model was based primarily on an enhanced version of the USEPA's River Reach File 1 (RF1), a 1:500,000-scale digital representation of stream networks (DeWald et al. 1985; Nolan et al. 2002). Enhancements included (1) the addition of reaches necessary to include several monitoring stations based on streams defined in the U.S. Geological Survey (USGS) National Hydrography Dataset, a 1:100,000-scale stream network representation (U.S. Geological Survey 2001a); and (2) recoding of several reaches of the national scale RF1 to add locally developed information about reaches that were part of reservoirs not indicated in the national data set (North Carolina Department of Environment, Health, and Natural Resources 1992).

The stream network was composed of 492 stream reaches having an average length of 14.6 km (median = 11.2 km, with an interquartile range of 6.0–19.6 km). Drainage catchments for each stream reach were delineated by using a stream-conditioned, 30-m resolution digital elevation model (DEM) data set and automated geographic information system (GIS) drainage-basin delineation procedures (Hellweger and Maidmont 1997; U.S. Geological Survey 2001b). The areas of the reach catchments ranged from 0.05 to 759 km^2, with a median of 53.5 km^2 and interquartile range of 23.2–129.8 km^2. Estimates of streamflow for each reach were determined through an accumulation of average annual runoff, based on data from the period 1951–1980, for each reach catchment (Gebert et al. 1987). Average annual runoff associated with each reach catchment was calculated; runoff values for individual catchments were summed, using networking algorithms, to estimate average annual streamflow at monitoring stations. The estimated mean streamflows for river reaches ranged from 0.04 to 291 m^3/s, with a median of 2.47 m^3/s and an interquartile range of 0.88–11.8 m^3/s. Estimated streamflows typically were smaller than actual reported annual discharges at gauging sites, with a median difference of 7%.

Stream-channel distances for two flow classes were used in the model to estimate in-stream nitrogen losses. Channel distances were computed by using a Fortran-based network climbing algorithm (White et al. 1992) and information about the stream-reach lengths and connectivity produced by GIS analysis. Reservoir surface area and outflow data measured at the most downstream

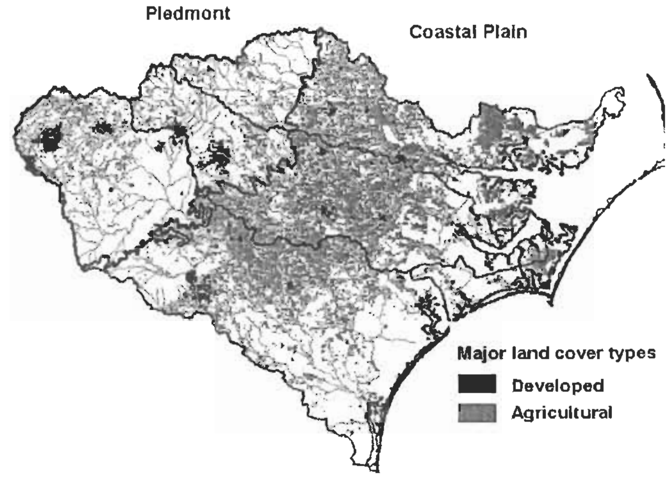

Fig. 4. Agricultural and developed land cover in Piedmont and Coastal Plain of North Carolina SPARROW model study area (land-cover data have been generalized to 1 km^2 pixels)

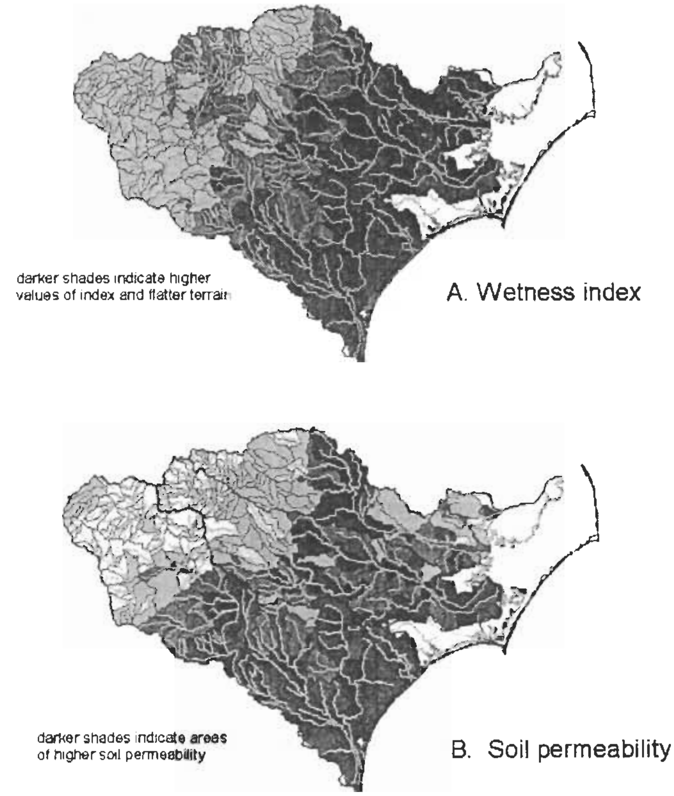

Fig. 5. (A) Wetness index and (B) soil permeability land-delivery variables used in North Carolina SPARROW model

reservoir reach were used to estimate areal hydraulic load associated with 14 reservoirs and lakes in the three river basins (North Carolina Department of Environment, Health, and Natural Resources 1992). Estimates of areal hydraulic load were calculated as the ratio of reservoir outflow discharge to the water-surface area of the reservoir.

Nutrient Sources

Estimates of total nitrogen discharged by 249 permitted point-source dischargers were made for the year 1992 in the three river basins by using data derived from Discharge Monitoring Reports (Chris Roessler, North Carolina Division of Water Quality, written communication 2001). The median load from these dischargers was 1,056 kg of total nitrogen, with an interquartile range of 150–11,860 kg. Load information from individual facilities was allocated to individual reach catchments by using GIS overlay processes.

Estimates of atmospheric deposition of nitrogen were developed by using early 1990s wet nitrate and ammonium nitrogen deposition data collected in the National Atmospheric Deposition Program (NADP 2002) that were interpolated into national atmospheric deposition maps (U.S. Geological Survey 2000). Total nitrogen contribution from atmospheric deposition was 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 in a GIS, and individual categories of nitrogen deposition were calculated and summed for each reach.

Estimates of 1992 nitrogen inputs associated with both agriculture and nonagricultural areas were derived by using a national land-cover data base developed with remotely sensed data from the early 1990s (Fig. 4; Loveland and Shaw 1996; Vogelman et al. 2001; U.S. Geological Survey 2002a). Areas of agricultural and nonagricultural (i.e., combined developed and forested land cover) land in each reach catchment were estimated by using GIS overlay analysis. An estimate of agricultural nitrogen loads was made by using agricultural land area as a source term in Eq. (1); nonagricultural nitrogen inputs were estimated using the combined developed/forested land area as another source term (units $=km^2$). The resulting coefficients represent nitrogen export coefficients for these land-cover types, indicating the rate of nitrogen export (e.g., metric tons per square kilometer) associated with the land-cover area. Using land-cover area for estimating diffuse nitrogen inputs has three advantages—it is conceptually simple to understand and interpret; the resulting export coefficients can be compared with export coefficients reported in the literature; and the land-cover area approach minimizes the introduction of error and uncertainty associated with the apportionment of tabular county agricultural statistics to the relatively small reach catchments.

Landscape Characteristics

Climatic and landscape features can influence delivery of diffuse sources of nitrogen from within reach catchments to the edges of streams. The significance of several of these features, including climate, topography, and soil, were tested using the SPARROW modeling approach.

Topography helps control the movement of water through a river basin and, thus, may play an important role in the movement of nitrogen within the basin. A DEM spatial data set (U.S. Geological Survey 2001b) was used to estimate several topographic measures for each reach catchment, including slope and a wetness index measuring the relative susceptibility of the reach catchment to saturation due to movement of precipitation through surface and subsurface flow paths (Wolock 1993). Because higher values of slope are positively associated with runoff and nutrient delivery, reciprocal values of slope were used in the model. The higher

Table 1. Regression Models Considered for North Carolina Spatially Referenced Regression on Watershed Implementation.

Variable	Model run	1	2	3	4	5	6	7	8
	Model R^2	0.93	0.94	0.94	0.93	0.94	0.95	0.95	0.96
	Mean square error	0.22	0.22	0.21	0.25	0.2	0.18	0.19	0.13
Variable	Model category								
Slope	Landscape delivery							−00.33 (.02)	
Wetness index	Landscape delivery								−1.12 (0.0001)
Permeability	Landscape delivery					−0.05 (0.30)			
Soil organic carbon	Landscape delivery					−0.04 (0.05)	−0.04 (0.01)		
Hydrologic soil group	Landscape delivery	4.1 (0.001)	4.24 (0.001)	2.94 (0.05)	5.04 (0.0004)		2.32 (0.08)	0.82 (0.65)	0.6 (0.62)
Small stream	Stream decay	0.08 (0.02)	0.08 (0.01)	0.06 (0.06)	0.09 (0.007)	0.03 (0.30)	0.05 (0.11)	0.06 (0.05)	0.05 (0.06)
Large stream	Stream decay	0.002 (0.35)	0.002 (0.38)	0.002 (0.34)	0.002 (0.38)	0.003 (0.25)	0.002 (0.49)	0.002 (0.46)	0.002 (0.36)
Reservoir	Reservoir decay	16.3 (0.008)	16.7 (0.007)	16.6 (0.006)	16.7 (0.01)	14.48 (0.01)	15.1 (0.007)	14.9 (0.008)	18.3 (0.0003)
Atmospheric	Nitrogen source		2.22 (0.16)	3.7 (0.11)		0.09 (0.31)	0.25 (0.49)	0.2 (0.48)	0.00005 (0.69)
Point source	Nitrogen source	0.85 (0.006)	0.52 (0.13)	0.54 (0.1)	0.48 (0.2)	0.65 (0.03)	0.52 (0.07)	0.59 (0.05)	0.51 (0.02)
Agricultural land	Nitrogen source	5.9 (0.09)	5.7 (0.12)	1.9 (0.49)	11.1 (0.07)	0.52 (0.06)	2.32 (0.14)	1.12 (0.28)	0.00005 (0.72)
Developed land	Nitrogen source		5.9 (0.38)	0.04 (0.99)	13.9 (0.18)	1.57 (0.04)	4.64 (0.12)	2.92 (0.15)	0.0004 (0.66)
Forested land	Nitrogen source			−1.83 (0.15)	1.9 (0.29)				
Nonagricultural land	Nitrogen source	1.8 (0.08)							

Note: Values for each variable include model coefficient value and p value (in parentheses).

the wetness index, the flatter the land surface. Generally, there are pronounced west-to-east gradients in these topographic variables (Fig. 5).

Soil characteristics also influence the movement of nitrogen within the river basins. In an earlier study, North Carolina Coastal Plain streams in agricultural drainage basins with relatively poorly drained soils had higher-ambient nitrogen concentrations than better drained soils, perhaps reflecting higher surface-runoff rates (McMahon and Harned 1998). 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), including permeability (Shirazi et al. 2001a,b), soil hydrologic groups (Miller and White 1998; Earth Systems Science Center 2001), and soil carbon (Sharon Waltman, U.S. Department of Agriculture, written communication 2001). An area-weighted average soil hydrologic group variable was calculated for each reach catchment, with a value ranging from 1 (relatively well drained) to 4 (relatively poorly drained). Because higher values for this variable are expected to be associated with greater nutrient delivery, the reciprocal value of this variable is used in the model. Values of permeability and soil carbon have a pronounced west-east spatial trend, reflecting large-scale processes associated with soil formation (Fig. 5).

Model Calibration Results

A number of models were considered as the basis for implementing the North Carolina SPARROW effort (Table 1). All models had one or more landscape-delivery variables representing factors that influence the movement of nitrogen over the landscape to the stream edge. Landscape-delivery factors were based either on topographic (e.g., slope and wetness index) or soil (e.g., permeability, soil organic carbon, and soil hydrologic group) characteristics. Three aquatic attenuation variables were considered, representing small and large streams and reservoirs. Small streams were defined as those with flows less than 1.04 m^3/s; 30% of the stream reaches had average flows less than this value. The source variables for point sources and atmospheric deposition defined nitrogen inputs in terms of mass (metric tons) for 1992. Other diffuse-source variables (e.g., agricultural, developed, forested, and combined developed/forested) represented nitrogen inputs by land-use area.

The eight regression models presented in Table 1 each have a relatively good fit [the coefficient of determination, or R-squared (R^2), ranges from 0.93 to 0.96; mean square error (MSE) ranges from 0.13 to 0.25]. Models are listed with the most parsimonious first (#1), followed by models with more complex source variables (#2, #3, #4) and then with additional delivery factors (#5, #6, #7, #8). Although model fit improves as more landscape-

Table 2. Comparison of Parametric and Bootstrap Coefficients for North Carolina Spatially Referenced Regressions on Watershed Total Nitrogen Regression Model

	Parametric model coefficient estimate (Model 1)	p-value	Bootstrap coefficient estimate	Lower 90% confidence level	Upper 90% confidence level
R^2	0.93				
Mean square error (MSE)	0.22				
Total nitrogen sources					
Point sources (ton/1992)	0.85	0.006	0.858	0.147	1.315
Agricultural land area (ton/km^2)	5.9	0.09	6.628	0.393	11.685
Nonagricultural land area (ton/km^2)	1.79	0.08	1.986	−0.102	3.563
Land-delivery variable					
Soil hydrologic group	4.13	0.001	3.868	1.275	6.215
Aquatic loss					
Small stream (km^{-1})	0.08	0.02	0.086	0.009	0.159
Large stream (km^{-1})	0.002	0.35	0.003	−0.001	0.006
Reservoir (m/year)	16.4	0.008	14.919	9.144	19.055

delivery factors are included (i.e., MSE drops), the delivery terms enter with the wrong sign (i.e., indicating that the landscape-delivery variable has a positive relation with load) and the significance of the source variables declines.

The model selected for use in this analysis (Model 1) has three source variables (point sources, agricultural sources, and nonagricultural sources), a landscape-delivery term based on soil hydrologic group, two in-stream loss-rate coefficients, and one reservoir decay-rate coefficient. Model 1 has the most statistically significant set of source, landscape-delivery, and aquatic-decay variables among the various models that were compared. Model 1 also corresponds to a physically interpretable description of landscape delivery, based on the statistical significance and sign of the hydrologic soil-group coefficient. Another strength is that the point-source term in this model is close to unity relative to the other models. Because point sources are discharged directly to streams without any land-delivery attenuation, it is expected that the fitted coefficient should be one. The closeness of the point-source coefficient to a value of 1 provides evidence that the model is relatively accurately specified.

Model 1 explains approximately 93% of the variation in the natural logarithms of mean annual total nitrogen flux, with a MSE of 0.22 (Tables 1 and 2). Plots of observed and predicted nitrogen load and yield indicate relatively good agreement between observed and predicted values (Fig. 6). The median value of the difference between observed and predicted yields is 6% and the interquartile range is from −23 to 35%. Observed and predicted loads agree most closely at sites that have larger loads, which tend to be associated with sites that have longer discharge and water-quality records. Observed and predicted nitrogen yields, which adjust for the influence of drainage area on nitrogen loads, also indicate relatively good agreement ($R^2=0.62$). The single large outlier is associated with a monitoring station on Ellerbe Creek in the Neuse River Basin and is located downstream from a relatively large wastewater treatment plant. The smaller than expected point-source coefficient contributes to the underestimation at this site. The R^2 and MSE of the North Carolina model suggest a fit slightly inferior to SPARROW models fit 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 ($R^2=0.97$ and MSE$=0.14$, calibrated using 37 stations; Alexander et al. 2002) but a somewhat better fit than a nationwide SPARROW model ($R^2=0.88$ and MSE$=0.43$, calibrated using 414 stations across the United States; Smith et al. 1997).

A final set of model coefficient estimates was developed using 200 iterations of a bootstrapping procedure; bootstrap coefficients in Table 2 represent the average bootstrap coefficient values. Although the average values of the bootstrap coefficients generally were similar to the parametric estimates, bootstrap coefficients for the two land-area source variables were larger than their parametric counterparts, and the land-delivery and reservoir attenuation factors were both somewhat smaller. The 90% confidence interval for the nonagricultural land and large stream-attenuation coefficients includes zero. Both coefficients were retained in the final model because of interest in these coefficients on the part of state officials and stakeholders in the Neuse River TMDL stakeholder group.

A spatial trend is evident in the residuals from the parametric model. This trend indicates a tendency for over prediction in the Piedmont portion of the three basins and under prediction in the Coastal Plain portion (Fig. 7). This spatial pattern suggests that the model is not adequately specified; that is, it does not include one or more variables that would account for the spatial pattern associated with variations in the load variable. This spatial pattern may be related to the previously noted spatial trend evident in a number of the land-delivery variables. Another potential reason for this trend is that agricultural cultivation tends to be more intense in the Coastal Plain than in the Piedmont, not just in terms of area but also in terms of crop types and the intensity of fertilizer use (U.S. Department of Agriculture 2001). Using agricultural land area to represent agricultural nitrogen inputs in the Coastal Plain may underrepresent the actual agricultural inputs in areas of intense agricultural cultivation. Accounting for spatial differences in crop type, fertilizer use, and livestock inventories in future models may provide a more accurate specification of agricultural nitrogen sources.

Three source coefficients are included in the final model—one represents point sources and the other two represent diffuse nitrogen sources associated with agricultural and nonagricultural land. The fact that the point-source coefficient is less than the expected value of 1 suggests that the model representation of sources and attenuation could be improved. The diffuse-source coefficients can be interpreted as nitrogen export coefficients [model coefficient units, metric tons per square kilometer (ton/km^2), are transformed to more conventional export coefficient units, kilograms per hectare (kg/ha), by multiplying the coefficients by 10]. The values for agricultural (59 kg/ha) and nonagricultural (18 kg/ha) land are within the interquartile range of values for these land-use

(A) Observed and predicted total nitrogen loads

(B) Observed and predicted total nitrogen yields

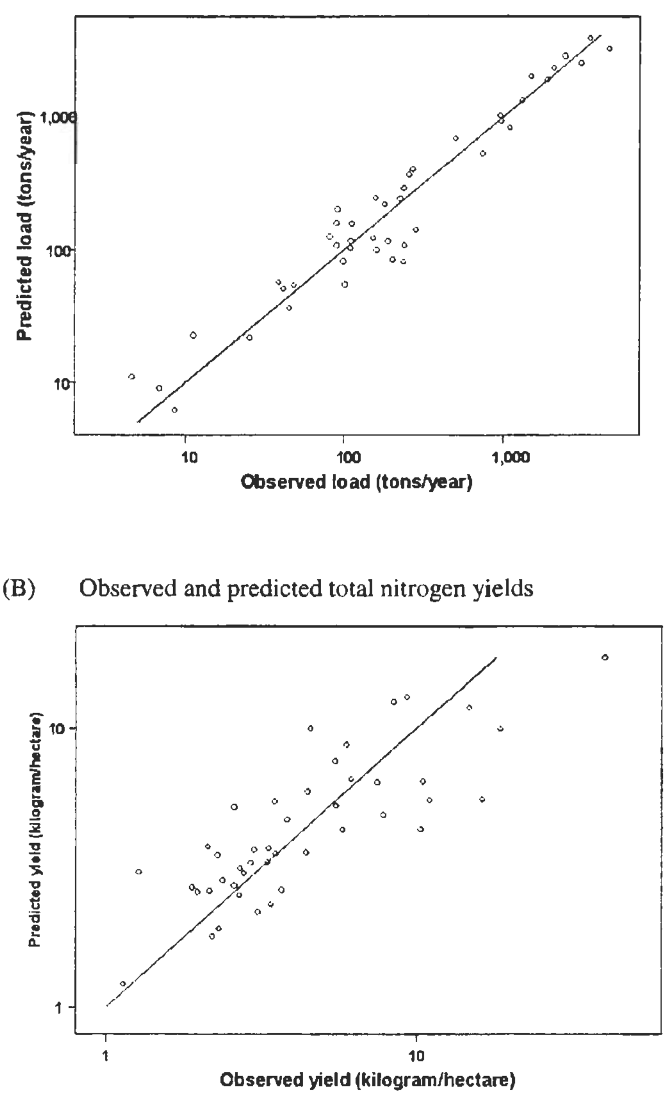

Fig. 6. Observed and predicted total nitrogen (A) loads and (B) yields at 44 monitoring sites included in North Carolina SPARROW model

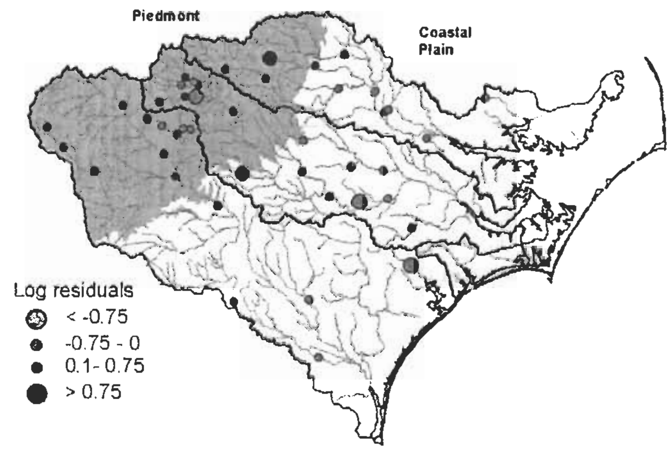

Fig. 7. Log residuals (predicted minus observed values) for 44 stream monitoring stations used to calibrate North Carolina SPARROW model. Shaded area in main study area map represents Piedmont ecoregion and unshaded area represents Coastal Plain ecoregion.

types reported in Beaulac and Reckhow (1982), but at the high end of the range. Typically, export coefficients estimate the loading of nutrients over a short distance, such as at the edge of an agricultural field or near a developed land-cover feature. The diffuse-source coefficients reported in Table 2 have been estimated empirically to explain variations in in-stream nutrient loads associated with a water-quality monitoring site at some distance from most land-cover patches within the drainage basin. Because the model statistically controls for attenuation associated with landscape and aquatic processes, the diffuse-source parameters are larger than they would be if the attenuation terms were not included in the model. For example, if the landscape-delivery and aquatic terms are excluded from the parametric model reported in Table 2, the value of the agricultural (7.4 kg/ha) and nonagricultural (2.5 kg/ha) coefficients decline by almost an order of magnitude ($R^2=0.89$ and MSE$=0.34$ for this reduced model).

The landscape-delivery variable used in the final model had the correct sign and strong statistical significance. Although some of the landscape-delivery terms other than soil hydrologic group had very strong statistical significance (e.g., the wetness index in Model 4), the landscape-delivery terms other than the soil hydrologic group entered the model with a negative sign, which contradicts the logic expressed in Eq. (1). Landscape-delivery variables should diminish the amount of nitrogen delivered to a stream; a negative value for α in Eq. (1) results in an increased load. Soil hydrologic group is the only landscape-delivery variable that consistently enters the model with the expected sign. It is worth noting that each of the other landscape-delivery variables has a relatively strong west-to-east spatial trend, which corresponds to a general west-to-east trend in streamflow and nitrogen load (see Fig. 4 for examples). It may be that the model calibration results are influenced by these spatial trends so that, for example, higher values of soil permeability are associated with higher in-stream nutrient loads (e.g., Model 5 in Table 1) rather than the expected opposite relation.

The sign of the landscape-delivery variable in Model 1 suggests an inverse relation between soil drainage characteristics and nitrogen loading, which is consistent with the findings of other SPARROW models (Smith et al. 1997; Preston and Brakebill 1999; Alexander et al. 2002). The product of this coefficient and the soil-drainage characteristics of any reach catchment creates an index of nitrogen attenuation in soils for this catchment (Alexander et al. 2002). The product of this index and the diffuse-source components—estimated coefficient and agricultural and nonagricultural land area in the catchment—of the model [see Eq. (1)] quantifies the amount of nutrients delivered to a stream reach by a particular source.

The stream and reservoir decay loss coefficients indicate that nutrient removal by small streams occurs at a rate that is an order of magnitude greater than in large streams. Nitrogen is removed from small streams at a rate of about 8% per kilometer of stream length, compared with a rate of less than 1% per kilometer for large streams. The inverse relation between nutrient-loss rate and stream size is consistent with findings in other SPARROW models (Smith et al. 1997; Preston and Brakebill 1999; Alexander et al. 2002). We reexpressed the estimates of in-stream loss per kilometer as loss rates per unit of time based on estimates of mean water velocity for the RF1 stream reaches (Alexander et al. 1999) in the North Carolina watersheds in each of the two flow classes. The North Carolina SPARROW-estimated loss rate for

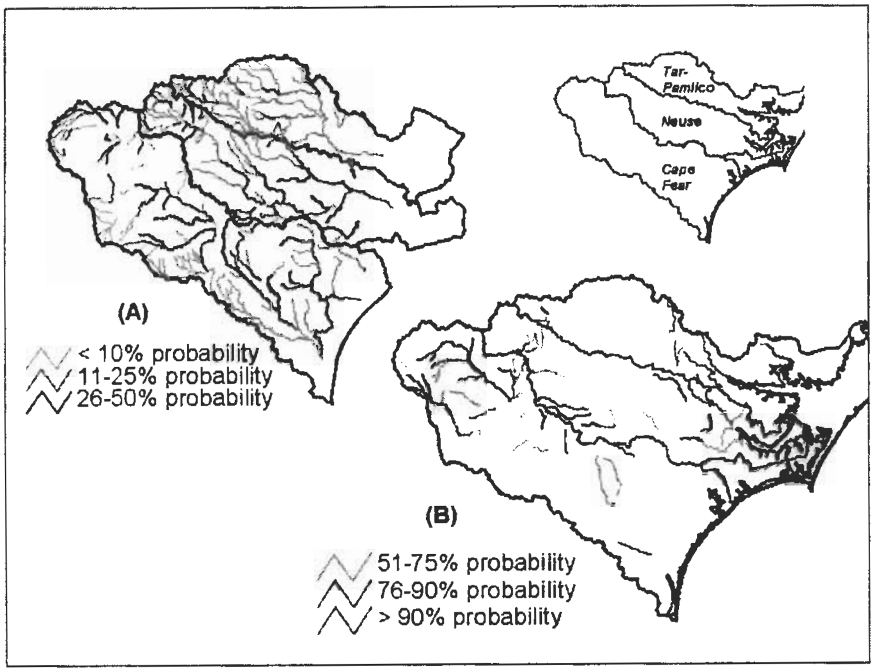

Fig. 8. Stream reaches in North Carolina SPARROW model study area that have (A) less than 50% probability and (B) greater than 50% probability of exceeding total nitrogen concentration of 1.5 mg/L.

small streams [0.99 per day; estimated velocity of 12.3 km/day] is about one and one-half to two times as large as the estimates obtained in the national and Chesapeake Bay SPARROW models for similar size streams (i.e., 0.45 and 0.76 per day, respectively), whereas the North Carolina loss rate for large streams (0.06 per day; estimated velocity of 30.4 km/day) is within the range of estimates obtained in the other SPARROW models (Alexander et al. 2000).

The estimated SPARROW total nitrogen settling velocity in lakes and reservoirs is the most statistically significant of the aquatic-loss parameters and is within the range of values in New Zealand, Europe, and North America reported in Alexander et al. (2002). The mean settling velocity of 16.4 m/year is an intermediate value between that reported for lakes with a settling velocity of less than 10 m/year, where nitrogen removal is predominantly by denitrification, and lakes with a settling velocity greater than 25 m/year, where biological uptake and sedimentation are the dominant removal processes (Alexander et al. 2002). The fraction of total nitrogen removed from lakes and reservoirs was calculated as the product of the settling velocity rates from the 200 bootstrap coefficient estimates and the inverse of areal hydraulic load for each reservoir [Eq. (3)]. The range of hydraulic loads for the 14 lakes and reservoirs in the three study basins reflects the range of lake-surface area and outflows associated with the disparate lake and reservoir types in the three basins. Nutrient removal fractions range from a low of 15% at the Tar River Reservoir to a high of 94% at Harris Lake (median=54%; the interquartile range is from 32 to 63%).

Potential Total Maximum Daily Load Applications of Spatially Referenced Regressions on Watershed Model

The TMDL process, as outlined by the NRC assessment committee, is iterative (National Research Council 2001). Information gathering and management activities cycle among (1) preliminary assessment of a stream's ability to meet a designated use; (2) additional monitoring, when necessary, to make a more complete quantitative assessment of compliance with a designated use; and (3) development and implementation of management practices that apportion responsibilities for improving ambient water-quality conditions among contributors of point and diffuse sources of nutrients.

We present several applications of the SPARROW model that address information needs associated with TMDL activities. Among these information needs are the identification of a spatially explicit probability distribution of impaired reaches, the refinement of a monitoring strategy based on impairment probabilities and model-prediction uncertainty, and the allocation of total nitrogen loads among sources.

Identifying the Location of Ambient Water-Quality Problems

The design of management activities to improve water quality is predicated on the ability to assess the population of water bodies (e.g., streams, lakes, and reservoirs) that have water-quality problems. Because states have an average of 70,000 stream reaches (National Research Council 2001) and limited resources for monitoring ambient water quality, this is not a trivial task. The NRC committee recommended an increase in the use of statistical models, calibrated with ambient monitoring data, to characterize the water quality at unmonitored locations.

The SPARROW model was used to predict the location of impaired reaches in terms of a hypothetical quantitative water-quality criterion for total nitrogen. Although a total nitrogen criterion has not been established nationally or in North Carolina, concentrations of total nitrogen greater than 1.5 mg/L have been associated with eutrophication in streams and rivers (Dodds et al. 1998; U.S. Environmental Protection Agency 2000). Model pre-

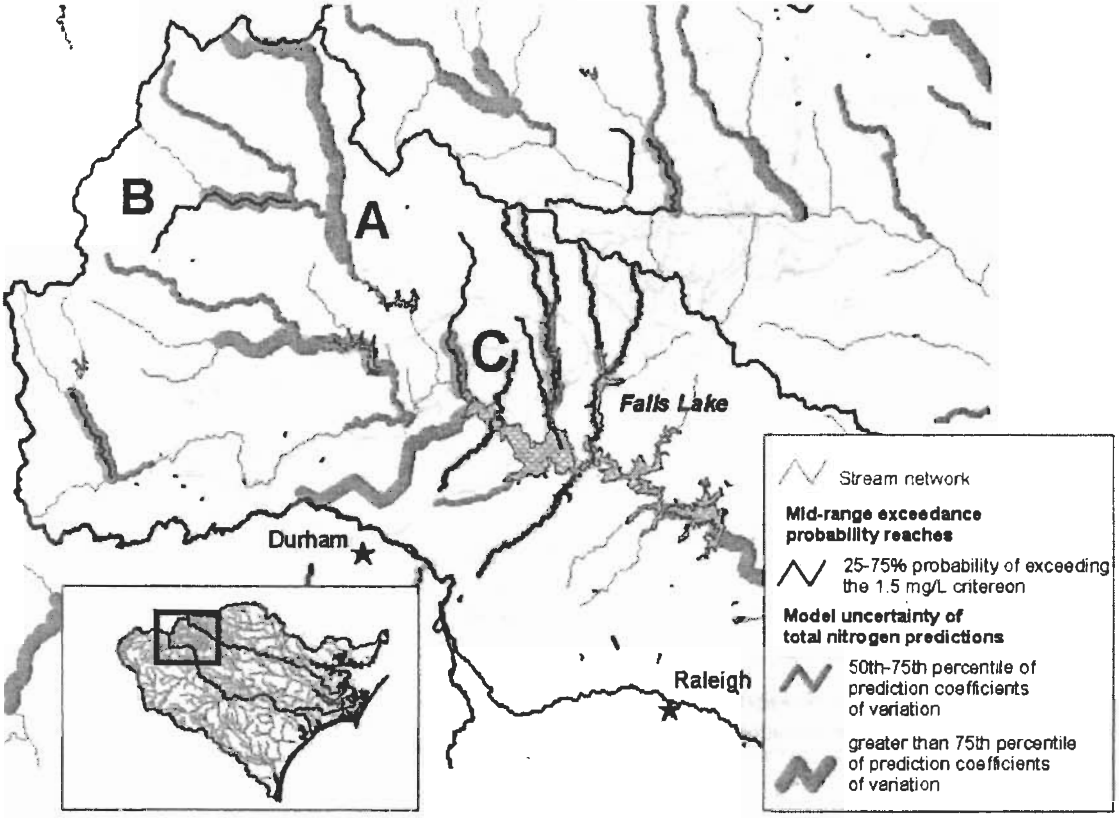

Fig. 9. Locations of stream reaches in vicinity of Falls Lake, N.C., with (A) relatively high coefficient of variation (CV) in total nitrogen predictions, (B) midrange total nitrogen exceedance probability predictions (25–75% probability of exceeding 1.5 mg/L hypothetical criteria), and (C) both high uncertainty and midrange exceedance probability

dictions of annual total nitrogen concentrations were used to develop an impairment probability distribution for reaches with an average annual concentration exceeding 1.5 mg/L. For each of the 492 reaches in the North Carolina watersheds, we developed a distribution of 200 flow-weighted mean annual total nitrogen concentrations, based on the set of 200 bootstrap model coefficients and reach-specific information about nitrogen sources, soil hydrologic group, and stream and reservoir properties. Estimates were then made of the percentage of the 200 predicted total nitrogen concentration values that exceeded 1.5 mg/L.

The results indicate a clear spatial pattern (Fig. 8). The map in Fig. 8(A) identifies reaches that have less than a 50% probability of exceeding 1.5 mg/L total nitrogen concentration; streams with the lightest shading have less than a 10% probability of exceeding this criterion. The Tar-Pamlico Basin has a preponderance of reaches in this lowest probability class. These reaches are located upstream from intensive agricultural areas in the lower part of the basin. The map in Fig. 8(B) identifies reaches with a relatively high probability of exceeding the criterion; reaches with the darkest shading have a greater than 90% probability of having total nitrogen concentrations that exceed 1.5 mg/L. Many of the reaches in this highest-probability class are located in the Piedmont area of the three basins and generally are associated with developed areas and locations of wastewater-treatment plants. Other reaches with a high probability of exceeding this criterion are located in Coastal Plain areas with high levels of agricultural activity.

Using Model Predictions to Improve Monitoring Strategies

Although the above application makes use of a hypothetical impairment standard, it illustrates how information on the spatial distribution of impairment probabilities can be used to locate stream reaches with water-quality problems. Reaches identified as having a relatively high-impairment probability may correspond to what the NRC committee calls a "preliminary list" of impaired water bodies to be considered for further TMDL actions (National Research Council 2001). Further TMDL actions would be contingent on results of additional monitoring.

Where should additional water-quality data-collection efforts be focused? Several guidelines should be considered in making this decision. First, it probably does not make sense to expend scarce monitoring resources to collect additional water-quality information about reaches with exceedance probabilities at the extremes of the impairment distributions. For example, additional monitoring information probably is not needed for reaches with less than a 25% probability or greater than a 75% probability of exceeding the hypothetical nitrogen criterion. Second, limited monitoring resources probably should be focused on areas where predicted water-quality conditions are most uncertain (National Research Council 2001).

Based on the NRC guidelines, we developed two measures of impairment uncertainty to identify reaches that have both a midrange probability of exceeding a specific criterion and where

Table 3. Total Nitrogen Budget for North Carolina Spatially Referenced Regression on Watershed Model Study Area.

	(A) Budget for Three North Carolina Basins			(B) Budget for Reaches with Low and High Total Nitrogen Concentrations	
	Cape Fear	Neuse	Tar-Pamlico	Reaches with mean total nitrogen concentration less than or equal to 1.5 mg/L ($n=358$)	Reaches with mean total nitrogen concentration greater than 1.5 mg/L ($n=135$)
Landscape yield (kg/ha)	6.7	7.0	5.7	5.1	37.7
Sources of landscape yield (%)					
Point sources	21	14	9	6	20
Agricultural land cover	41	55	58	51	48
Nonagricultural land cover	38	31	33	44	32
Stream and reservoir loss (%)	42	33	25	32	57
Watershed yield (kg/ha)	3.9	4.7	4.3	3.5	16.1
Sources of watershed yield (%)					
Point sources	19	12	11	6	20
Agricultural land cover	41	59	58	52	48
Nonagricultural land cover	39	30	31	42	32

Note: Basin budgets are estimated at final reach upstream from basin outlet. Reach budgets represent average across all reaches with concentrations above or below designated level. TN=total nitrogen and n=number of reaches in concentration class.

the water-quality predictions are relatively uncertain. Midrange probability reaches are those with between 25 and 75% probability of exceeding the hypothetical nutrient criterion of 1.5 mg/L. For this application, prediction uncertainty was characterized by using the 200 bootstrap predictions for each reach. A coefficient of variation (CV) was calculated by using the 200 predictions for each reach, and the CV quartiles across all reaches were calculated. Reaches with a CV in the highest 50% of all reach CVs were considered to have relatively high prediction uncertainty.

An application of the approach was developed for streams reaches in the vicinity of the Raleigh-Chapel Hill-Durham, N.C., area (Fig. 9). Falls Lake, shown in the center of Fig. 9, is the water-supply source for the city of Raleigh. Because of the importance of this water body as a drinking-water supply and the potential increase of water-treatment costs that may be associated with eutrophication, it is worth using these proposed measures of impairment uncertainty to determine whether additional monitoring within the lake's drainage area is justified. Stream reaches in the watershed represent a variety of combinations of exceedance probability and prediction uncertainty. Some reaches (e.g., Reach A in Fig. 9) with high-prediction uncertainty have either a relatively high probability (greater than 75%) or relatively low probability (less than 25%) of exceeding the criterion and, thus, may not necessarily have a high priority for additional monitoring. There also are reaches (e.g., Reach B in Fig. 9) with probabilities of exceeding the nitrogen criterion that fall within the desired mid-range but with predictions that are relatively certain. Finally, there are reaches (e.g., Reach C in Fig. 9) that fall within the desired exceedance range and have model predictions that are relatively uncertain. Such reaches may be candidates for additional water-quality monitoring.

Understanding the Sources and Movement of Nitrogen

Once a set of reaches (e.g., the Falls Lake watershed) or even an entire basin has been designated for full implementation of a TMDL program, developing and implementing a maximum loading-management strategy requires a constituent budget. A budget describes the total amount of a constituent entering and leaving a reach, watershed, or basin (e.g., the load of total nitrogen that arrives at the mouth of the Neuse River over a year) and allocates these loads among the sources of the constituent (e.g., point and diffuse sources; North Carolina Department of Environment and Natural Resources 1999). This final application describes the development of a total nitrogen budget and illustrates the spatial pattern of delivery of nitrogen from point and diffuse sources.

The nitrogen budget depicts the quantities of nitrogen entering the stream and reservoir systems (expressed per unit of drainage area and referred to as landscape yield), the amount leaving these systems (either at the mouth of a river basin or at the scale of the individual reach outlet and referred to as watershed yield), and the percentage of the inputs that are processed by the aquatic systems (Table 3). Two nitrogen budgets are presented—one for the three major river basins included in the study and one for the stream reaches representing two classes of water-quality conditions. Yield estimates in the latter case represent the average yield over all reaches that have a predicted concentration less than or equal to 1.5 mg/L or greater than this concentration. Yields in both budget scenarios also were apportioned among the three source categories considered in the model, including point sources and two diffuse sources.

Landscape yields for the three basins indicate that both the Cape Fear and Neuse River Basins have higher-nitrogen input rates than the Tar-Pamlico Basin, and agricultural inputs are the dominant source of nitrogen introduced into the stream network in all three basins. The Cape Fear and Neuse Basins have more developed land and larger point-source discharges of total nitrogen than the Tar-Pamlico Basin. Agricultural contributions, as a percentage of all nitrogen entering the basin's stream and reservoir systems, are approximately equal in the Neuse and Tar-Pamlico Basins (Fig. 10). Although agricultural contributions in the Cape Fear Basin are almost 20% less than in the other basins, they still represent the dominant source of nitrogen. Point-source contributions in the Cape Fear Basin, which occur predominantly in the upper part of the basin, are approximately twice the percentage of point-source contributions in the other basins.

Watershed yields in the three basins indicate the importance of stream and reservoir nitrogen processing and the magnitude and locations of nitrogen sources (Table 3; Fig. 11). The general decrease between landscape and watershed yields is a result of stream and reservoir nitrogen processing. The effects of these processes appear to be especially important in the Cape Fear

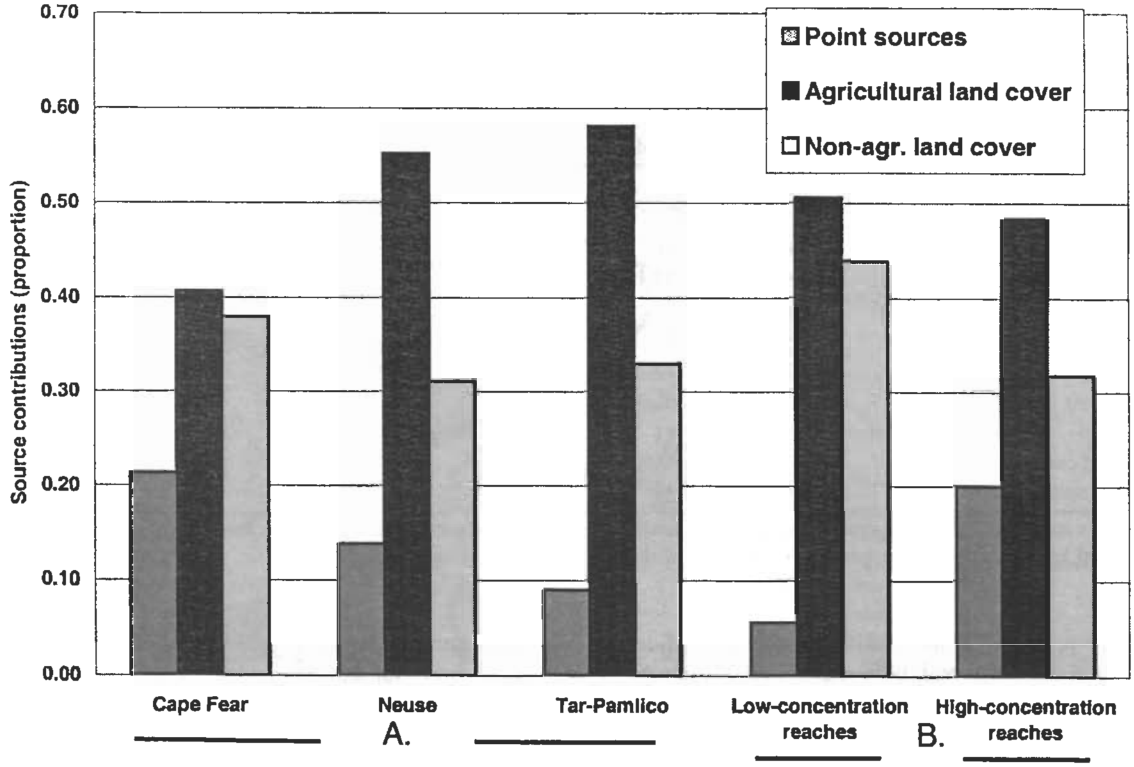

Fig. 10. Total nitrogen budget for nitrogen sources delivered to (A) three river basins used in North Carolina SPARROW model, and (B) two classes of stream reaches in these basins, where reaches are identified as having either low (less than or equal to 1.5 mg/L) or high (greater than 1.5 mg/L) total nitrogen concentration

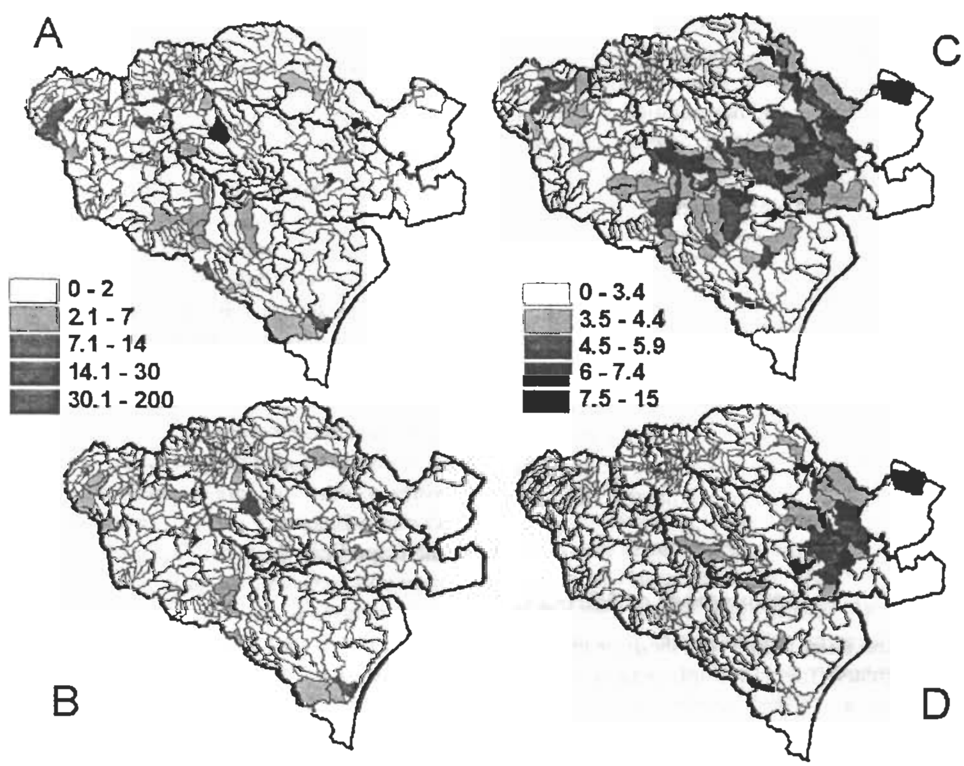

Fig. 11. Loading of total nitrogen (kilograms per hectare) associated with point (A) and agricultural (C) sources and delivery of total nitrogen (kilograms per hectare) to watershed outlet associated with point (B) and agricultural (D) sources

River Basin. The large losses (42%) in the Cape Fear Basin are a function of the fact that a great deal of the point-source inputs, which represent a relatively important source of nitrogen, are located in the upper and middle parts of the basin and travel a long distance to the outlet reach where watershed yield is measured (Fig. 11, Part A). The Cape Fear Basin has the most river miles among the three basins, which lead to greater cumulative losses of nitrogen by denitrification and sedimentation (Fig. 11, Part B). The Tar-Pamlico Basin has the smallest amount of nitrogen removed by streams and reservoirs (25%). Most of the nitrogen in this basin is introduced by agricultural activities in the lower part of the basin (Fig. 11, Part C); thus, the opportunity for nitrogen processing by the stream network in this basin is small relative to the other basins. Much of the agricultural nitrogen introduced in the lower part of the basin is delivered to the estuary (Fig. 11, Part D). The intermediate stream and reservoir-related loss in the Neuse Basin (33%) reflects the spatial distribution of agricultural inputs. Inputs in the middle part of the basin may be transformed by in-stream processing, while inputs closer to the estuary are subjected to less processing.

Landscape yields for reaches associated with either low or high-nitrogen concentrations present a dramatic contrast (Table 3; Fig. 10). All reaches in the North Carolina SPARROW model study area were assigned to either a low- or high-concentration class, depending on whether the average total nitrogen concentration per reach was less than or equal to 1.5 mg/L or greater than this concentration. Reach-level landscape (and watershed) yields were averaged across all reaches in each class. Reaches with high-total nitrogen concentrations had a landscape yield seven times higher than reaches with low-nitrogen concentration and five times higher than the landscape yield for the basin (Neuse) with the highest-basin landscape yield (Table 3). A similar procedure was applied to average reach-level source contributions. The source percentages indicate that agricultural sources contribute approximately equal percentages of landscape yield for both high- and low-concentration classes. The percentage contributions of point sources, however, are more than three times higher in stream reaches with high-total nitrogen concentrations.

Discussion and Conclusions

The SPARROW modeling approach, as applied to measurements of total nitrogen flux in three North Carolina river basins, addresses several information needs identified by a National Research Council-sponsored evaluation of the TMDL program. Water-quality and stream-discharge data collected by federal and state agencies were used effectively in this spatially referenced empirical water-quality model with information about nitrogen sources and landscape and aquatic features that affect nitrogen processing. The model of 492 river reaches and 14 reservoirs provided relatively accurate predictions of in-stream total nitrogen loads for reaches in the three basins. Model estimates of point and diffuse sources and estimates of nitrogen-loss rates in streams and reservoirs compared reasonably well with literature estimates. The model provides information that may be useful for water-quality managers who must implement TMDLs, including reach-level predictions of the probability of exceeding quantitative water-quality criteria and estimates of source contributions and nitrogen losses in streams based on total nitrogen budgets. Because SPARROW estimation procedures accounted for the movement and transformation of nitrogen across the landscape and through the stream and reservoir network, reach-level predictions of nutrient inputs and delivery could be mapped, providing an intuitive and spatially detailed summary of both the origins and fate of nutrients within a basin.

There are several concerns about the use of the SPARROW modeling approach for supporting TMDL information needs, including the use of SPARROW in conjunction with other monitoring and modeling efforts; the characterization of uncertainty in calculations of nitrogen inputs; possible enhancements to the SPARROW methodology to address spatial autocorrelation, and the adequacy of reliance on a single model to support TMDL efforts.

The NRC assessment committee suggested that models, such as SPARROW, that are capable of determining a preliminary probability distribution of impairment can be used in a Bayesian framework to reduce the quantity of additional water-quality monitoring data that are needed to make TMDL listing decisions at a given level of confidence (National Research Council 2001). SPARROW predictions can provide a prior estimate of impairment probability. Use of this prior information in a Bayesian analysis can allow final decisions about impairment probabilities to be made with less additional sampling than would be required without the prior information. The committee also proposed that a model such as SPARROW could be used in conjunction with other models. For example, SPARROW nutrient loading predictions and associated estimates of uncertainty could be used as inputs in an estuary water-quality model to predict biological endpoints, such as the number of fish kills (Borsuk 2001). Although neither of these approaches for SPARROW predictions has been attempted, both seem conceptually reasonable and merit additional assessment.

Diffuse-source nitrogen inputs cannot be estimated exactly in regional-scale studies. The application reported here relied on estimates of agricultural and nonagricultural land-cover area developed by overlaying reach catchment boundaries on mapped land-cover data derived from satellite data. Land-cover classifications are interpretations containing errors. Satellite-based land-cover classifications contain errors of commission (an image pixel may be assigned to a class to which it does not belong) and omission (a pixel is not assigned to its appropriate class; Jensen 1986). Although the land-cover data used in this study have an overall accuracy of 80.5% across the eastern United States (Yang et al. 2001), classification accuracy tends to diminish in the heterogeneous landscapes typical of the three study basins, where an intermixing of relatively small patches of various land-cover types occurs (Smith et al. 2002; Fig. 4). Estimates of mass nitrogen inputs associated with agriculture would have to be derived by estimating agriculture inputs as a function of both land area and nutrient input rates associated with various crop types. Even if the land-cover data are entirely accurate, attributing tabular agricultural data to reach catchments may introduce substantial error in agricultural nutrient estimates. This would occur if the assumption is incorrect that agricultural activities are distributed evenly across the county or if the tabular data are incorrect. Similarly, the accuracy of mass nitrogen input estimates associated with developed land-cover types (e.g., urban fertilizer) depends on the accuracy of land-cover data and input rates (e.g., nitrogen fertilizer used per square kilometer of developed land-cover) used with the area data.

While standard methods have been developed to characterize classification uncertainty (Yang et al. 2001) and the uncertainty associated with fertilizer application (Deanna Osmond, North Carolina State University, written communication, October 2001), additional research is needed regarding use of quantitative descriptors of data quality (e.g., classification error rate or variability in application rates) in models to estimate nutrient loading (Hess and Bay 1997). In particular, Monte Carlo or other statistical methods might be used in future SPARROW model applications to assess the effects of explanatory variable measurement errors on model-parameter estimates and predictions.

Two areas of enhancement of the SPARROW methodology will be considered for addressing spatial autocorrelation. First, as indicated by the spatial pattern in the residuals (Fig. 7), it is reasonable to believe that different subwatersheds may have slightly different coefficient values (e.g., the nutrient-delivery rates). This can be achieved by using a mixed-effect modeling approach (McCulloch and Searle 2001) or a Bayesian hierarchical modeling approach (e.g., Congdon 2001). The mixed-effect model introduces a random effect to the model coefficient for each subwatershed. The Bayesian hierarchical model applies a distributional structure to the model coefficient. Each model coefficient is assumed to have an overall (or super) distribution that is the parent distribution of subwatershed coefficient distributions. This change can be used to examine spatial variations in model coefficients.

The second improvement is to model the spatial autocorrelation in the data that is not explained by the stream network. The sources of spatial autocorrelation may include similarity in soil and land-use characteristics and physical interactions among neighboring subwatersheds. The Bayesian conditional autoregressive modeling approach (Besag and Kooperberg 1995) can be used for this purpose. A new computational strategy is necessary for both improvements. One such strategy is the Markov chain Monte Carlo simulation (Gilks et al. 1996), where model parameter estimation and uncertainty analysis can be done simultaneously.

Finally, a variety of reasonable models can be formulated for predicting the loads that are created and transformed in the three North Carolina basins (Table 1), including models that represent agricultural and nonagricultural diffuse inputs using land-cover area and models where estimates of mass nitrogen inputs from agriculture or developed areas are based on land-cover area and nitrogen application and use rates. The effects of using different specifications on SPARROW predictions of impairment probabilities and nutrient budgets have not been examined for the three North Carolina river basins. It can be anticipated, however, that different conceptual approaches to characterizing nutrient inputs (e.g., area based as opposed to mass-input based); landscape-delivery factors (e.g., use of topographic terms instead of, or in addition to, soil terms); and aquatic systems (e.g., using five stream-size classes based on the quantiles of reach-flow characteristics) may well result in different spatial characterizations of impairment probabilities and yields of nitrogen at the basin outlets. Future SPARROW investigations should examine the sensitivity of SPARROW outputs to different model specifications and explore methods for combining information resulting from different conceptual descriptions of the ways nutrients enter and move through a basin.

Acknowledgments

This paper was prepared as part of the Albemarle-Pamlico project of the U.S. Geological Survey's (USGS) National Water Quality Assessment Program (U.S. Geological Survey 2002b). The writers benefited substantially from guidance provided by John Brakebill (USGS), Anne Coan (North Carolina Farm Bureau), Becky Deckard (USGS), Doug Harned (USGS), Richard Moore (USGS), Deanna Osmond (North Carolina State University), Chris Roessler (North Carolina Department of Environment and Natural Resources), Dick Smith (USGS), Steve Preston (USGS), and three anonymous reviewers.

References

Alexander, R. B., Brakebill, J. W., Brew, R. E., and Smith, R. A. (1999). "ERF1—Enhanced river reach file 1.2." *U.S. Geological Survey Open-File Rep. 99–457*, Reston, Va., ⟨http://water.usgs.gov/GIS/metadata/usgswrd/erfl.html⟩ (May 6, 2002).

Alexander, R. B., Elliott, A. H., Shankar, U., and McBride, G. B. (2002). "Estimating the sources and transport of nutrients in the Waikato River Basin, New Zealand." *Water Resour. Res.*, 38, 1268–1290.

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 (London)*, 403, 758–761.

Beaulac, M. N., and Reckhow, K. H. (1982). "An examination of land use–nutrient export relationships." *Water Resour. Bull.*, 18, 1013–1024.

Besag, J., and Kooperberg, C. L. (1995). "On conditional and intrinsic autoregressions." *Biometrika*, 82, 733–746.

Borsuk, M. E. (2001). "A graphical probability network model to support water quality decision making for the Neuse River Estuary, North Carolina." PhD dissertation, Duke Univ., Durham, N.C.

Cohn, T. A., DeLong, L. L., Gilroy, E. J., Hirsch, R. M., and Wells, D. K. (1989). "Estimating constituent loads." *Water Resour. Res.*, 25, 937–942.

Congdon, P. (2001). *Bayesian statistical modeling*, Wiley, Chichester, England.

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.

Dodds, W. K., Jones, J. R., and Welch, E. B. (1998). "Suggested classification of stream trophic state: Distributions of temperate stream types by chlorophyll, total nitrogen, and phosphorus." *Water Res.*, 32(5), 1455–1462.

Earth Systems Science Center. (2001). "Soil information for environmental modeling and ecosystem management: Soil datasets." Pennsylvania State Univ., State College, Pa., ⟨http://www.essc.psu.edu/soil_info/index.cgi?soil_data&conus&data_cov⟩ (February 6, 2002).

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.

General Accounting Office. (2000). "Water quality—Key EPA and state decisions limited by inconsistent and incomplete data." *Rep. No. GAO/RCED-00-54*, Washington, D.C.

Gilks, W. R., Richardson, S., and Spiegelalter, D. J., eds. (1996). *Markov chain Monte Carlo in practice*, Chapman and Hall, New York.

Hellweger, F., and Maidmont, D. (1997). "AGREE-DEM surface reconditioning system." Univ. of Texas, Austin, Tex., ⟨http://www.crwr.utexas.edu/gis/gishyd98/quality/agree/agree.htm⟩ (August 14, 2001).

Hess, G. R., and Bay, J. M. (1997). "Generating confidence intervals for composition-based landscape indexes." *Landscape Ecology*, 12, 309–320.

Jensen, J. R. (1986). *Introductory digital image processing*, Prentice-Hall, Englewood Cliffs, N.J.

Knopman, D. S., and Smith, R. A. (1993). "20 years of the Clean Water Act." *Environment*, 35(1), 17–41.

Loveland, T. R., and Shaw, D. M. (1996). "Multi-resolution land characterization: Building collaborative partnerships." *GAP analysis: A landscape approach to biodiversity planning*, J. M. Scott, T. Tear, and F. Davis, eds., American Society for Photogrammetry and Remote Sensing, Charlotte, N.C., 83–90.

McCulloch, C. E., and Searle, S. R. (2001). *Generalized, linear, and mixed models*, Wiley, New York.

McMahon, G., and Harned, D. A. (1998). "Effect of environmental setting on sediment, nitrogen, and phosphorus concentrations in the Albemarle-Pamlico Drainage Basin." *Environ. Manage.*, 22(6), 887–903.

McMahon, G., and Woodside, M. D. (1997). "Nutrient mass balance for the Albemarle-Pamlico Drainage Basin, North Carolina and Virginia, 1990." *J. Am. Water Res. Assoc.*, 33(3), 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*, ⟨http://EarthInteractions.org⟩ (February 6, 2002).

National Atmospheric Deposition Program (NADP). (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.

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 Rep. 02–40*, Reston, Va., ⟨http://water.usgs.gov/GIS/metadata/usgswrd/erf1-2.html⟩ (May 3, 2002).

North Carolina Department of Environment, Health, and Natural Resources. (1992). "North Carolina lake assessment report." *Rep. No. 92-02*, Division of Environmental Management, Raleigh, N.C.

North Carolina Department of Environment and Natural Resources. (1999). *Total maximum daily load for nitrogen to the Neuse River Estuary*, Division of Water Quality, Raleigh, N.C.

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 Rep. 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." *J. Environ. Qual.*, 30, 112–120.

Shirazi, M. A., Boersman, L., Haggerty, P. K., and Johnson, C. B. (2001b). "Spatial extrapolation of soil characteristics using whole-soil particle size distributions." *J. Environ. Qual.*, 30, 101–111.

Sisterson, D. L. (1990). "Appendix A: Detailed SOx-S and NOx-N mass budgets for the United States and Canada." *Acidic deposition: State of science and technology, Rep. 8, Relationships between atmospheric emissions and deposition/air quality*, A. Venkatram, D. McNaughton, P. K. Karamchandani, J. Shannon, D. L. Sisterson, and M. Fernau, eds., National Acid Precipitation Assessment Program, 8A1-8A10.

Smith, R. A., Schwarz, G. E., and Alexander, R. B. (1997). "Regional interpretation of water-quality monitoring data." *Water Resour. Res.*, 33, 2781–2798.

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." *Photogramm. Eng. Remote Sens.*, 68(1), 65–70.

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*, Ft. Worth, Tex.

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

U.S. Environmental Protection Agency. (2000). "Nutrient criteria technical guidance manual, rivers and streams." *EPA-822-B-00-002*, Office of Water, Washington, D.C.

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. (2001a). "NHD data." ⟨http://nhd.usgs.gov/data.html⟩ (September 7, 2001).

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

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

U.S. Geological Survey. (2002b). National Water-Quality Assessment Program, ⟨http://water.usgs.gov/nawqa/⟩ (October 9, 2002).

Vogelmann, J. E., Howard, S. M., Yang, L., Larson, C. R., Wylie, B. K., and Van Driel, N. (2001). "Completion of the 1990s National Land Cover Dataset for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources." *Photogramm. Eng. Remote Sens.*, 67, 650–652.

White, D. A., Smith, R. A., Price, C. V., Alexander, R. B., and Robinson, K. W. (1992). "A spatial model to aggregate point source and diffuse-source water-quality data over large areas." *Comput. Geosci.*, 18(8), 1055–1073.

Wolock, D. M. (1993). "Simulating the variable source-area concept of streamflow generation with the watershed model TOPMODEL." *U.S. Geological Survey Water-Resources Investigations Rep. No. 93-4124*, Lawrence, Kan.

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 Sens. Environ.*, 76, 418–422.