Regional interpretation of water-quality monitoring data

Richard A. Smith, Gregory E. Schwarz, and Richard B. Alexander
U.S. Geological Survey, Reston, Virginia

Abstract. We describe a method for using spatially referenced regressions of contaminant transport on watershed attributes (SPARROW) in regional water-quality assessment. The method is designed to reduce the problems of data interpretation caused by sparse sampling, network bias, and basin heterogeneity. The regression equation relates measured transport rates in streams to spatially referenced descriptors of pollution sources and land-surface and stream-channel characteristics. Regression models of total phosphorus (TP) and total nitrogen (TN) transport are constructed for a region defined as the nontidal conterminous United States. Observed TN and TP transport rates are derived from water-quality records for 414 stations in the National Stream Quality Accounting Network. Nutrient sources identified in the equations include point sources, applied fertilizer, livestock waste, nonagricultural land, and atmospheric deposition (TN only). Surface characteristics found to be significant predictors of land-water delivery include soil permeability, stream density, and temperature (TN only). Estimated instream decay coefficients for the two contaminants decrease monotonically with increasing stream size. TP transport is found to be significantly reduced by reservoir retention. Spatial referencing of basin attributes in relation to the stream channel network greatly increases their statistical significance and model accuracy. The method is used to estimate the proportion of watersheds in the conterminous United States (i.e., hydrologic cataloging units) with outflow TP concentrations less than the criterion of 0.1 mg/L, and to classify cataloging units according to local TN yield (kg/km^2/yr).

1. Introduction

The objectives of regional water-quality assessments are to describe spatial and temporal patterns in water quality and identify the factors and processes that influence those conditions [*National Research Council*, 1994; *Mueller et al*., 1997]. Some regional assessments have the specific purpose of relating water quality to legal standards. Since its enactment in 1972, the Federal Clean Water Act (Public Law 92-500) has required state governments, river basin commissions, and the federal government to estimate biennially the proportion of surface waters that meet accepted quality standards. These assessments, and numerous others called for in subsequent amendments to the act [*Knopman and Smith*, 1993], influence a myriad of regulatory decisions and the expenditure of billions of dollars annually [*U.S. Environmental Protection Agency*, 1990b].

Historically, a broad combination of state and federal monitoring networks and programs has served as the principal source of data for assessments. Efforts to compile data from multiple sampling stations for assessing water quality at the state [*Dole and Wesbrook*, 1907; *Van Winkle and Eaton*, 1910], river basin [*Leighton and Holister*, 1904; *Barrows and Whipple*, 1907], and national [*Dole*, 1909] levels can be traced to the early 20th century. Since that time, network sampling programs have increased in number, size, and complexity [*National Research Council*, 1990] and now support regional water-quality assessment activities at spatial scales ranging from local to global [e.g., *Hirsch et al*., 1988; *Meybeck*, 1982].

Despite widespread and long-standing interest in the use of data from monitoring networks, certain commonly encountered problems make it difficult to interpret point-level water-quality data in areal terms and thus meet the objectives of regional water-quality assessments. Even when their objectives are clearly established and sampling programs are well planned, regional water-quality assessments are often complicated by (1) sparseness of sampling locations due to cost constraints, (2) spatial biases in the sampling network due to the need to target sampling toward specific pollution sources, and (3) drainage basin heterogeneity. These complications impede data interpretation in distinct ways by limiting sample sizes, reducing the regional representativeness of the sampling network, and limiting the ability to relate in-stream conditions to specific pollution sources.

In this paper, we describe a method for interpreting monitoring data that reduces the commonly encountered problems of network sparseness, bias, and basin heterogeneity. The method involves construction of a statistical model relating water-quality observations to spatially referenced (and potentially temporally referenced) data on basin attributes. The introduction of an explanatory model and ancillary data on basin attributes expands the information base for interpreting water-quality measurements, enhancing the ability both to describe regional conditions and identify causative factors.

Current methods for interpreting data from sampling networks are based on relatively simple models of limited usefulness in assessments because of the complications of network sampling. For example, estimates of the proportion of a water resource meeting specified quality standards are generally based on frequency-related interpretations of monitoring data; but such estimates are invalid when spatial or temporal sampling biases are large. Accurate estimates of proportions are currently achievable only through randomized sampling

This paper is not subject to U.S. copyright. Published in 1997 by the American Geophysical Union.

Paper number 97WR02171.

[*Messer et al*., 1991] or when an existing sampling network can be shown through ex post analysis [*Smith et al*., 1993a] to provide approximate estimates. In practice, however, strong incentives exist for conducting nonrandomized (i.e., targeted) sampling to determine the causes of poor water quality [*Knopman and Smith*, 1993] or to take advantage of established monitoring systems such as streamflow measuring networks. Thus, despite the important role that proportions play in legally mandated water-quality assessments, rigorous estimates of proportions are rarely made.

Spatial-analytic techniques such as kriging [e.g., *Clark*, 1979; *Hughes and Lettenmaier*, 1981] are designed to interpret spatial gradients and account for spatial dependencies in data from sampling networks. In theory, these techniques would provide a basis for drawing maps of water-quality conditions and making unbiased estimates of proportions. However, these interpolation techniques are not well suited to the dendritic nature of stream networks, and their usefulness for understanding the causes of contaminated conditions is limited by the lack of true explanatory variables.

A variety of statistical methods has been used to relate network-derived water-quality data to explanatory factors. One simple approach uses box plots to compare the distributions of water-quality measurements for groups of sampling stations differentiated by land use or other basin attributes [*Mueller et al*., 1995; *Smith et al*., 1993a; see also *Omernick*, 1976]. Although causal links have been identified, basin heterogeneity usually limits the ability to clearly differentiate the effects of specific attributes, and the predictive power of the technique is low. Beginning with attempts to determine the factors influencing suspended sediment yields [*Branson and Owen*, 1970; *Flaxman*, 1972; *Hindall*, 1976], a variety of regression techniques have been used to relate water quality to basin attributes [*Steele and Jennings*, 1972; *Lystrom et al*., 1978; *Steele*, 1983; *Peters*, 1984; *Osborne and Wiley*, 1988; *Driver and Tasker*, 1990; *Hainly and Kahn*, 1996; *Mueller et al*., 1997]. However, these models treat basin attributes as homogeneously distributed in the watershed, an assumption that fails to account for important relationships among the attributes and between the attributes and water-quality measurements on the dependent side of the regression equation. A recent paper by *Cressie and Majure* [1997] combines kriging techniques with a regression model that recognizes a finite zone of influence within the watershed, thus avoiding the strict homogeneous assumption.

The primary distinction between our method and previous regression approaches is our explicit inclusion of the spatial dimension in the underlying model. In the model developed here, contaminant transport is described as a function of spatially referenced land-surface and stream-channel characteristics. An important, testable hypothesis of the present study is that spatial referencing of basin attributes increases their correlation with water-quality measurements. A second, though less easily tested, hypothesis of this study is that spatial referencing facilitates the interpretation of model coefficients in terms of the sources and processes involved in contaminant transport through watersheds. As discussed and illustrated in section 4, this enhanced capability allows the model to be applied to a variety of assessment-related problems not possible with earlier approaches. We refer to the method described here as Spatially Referenced Regressions On Watershed attributes (SPARROW).

The SPARROW method described here includes several modifications and refinements of a prototype method for spatially referenced regressions of water-quality measurements on basin attributes previously described by *Smith et al*. [1993b]. The principal differences between the earlier study and this one are the consideration of more varied sources of pollution, additional processes affecting the delivery of these sources, a better statistical treatment of the decay process, and the application of the model to a larger region with a more diverse range of basin sizes and characteristics. The present model is based on data from approximately 400 monitoring stations in the National Stream Quality Accounting Network [*Alexander et al*., 1996; *Ficke and Hawkinson*, 1975] and is applied to nontidal stream reaches in the conterminous United States.

To demonstrate use of the model in the context of regional water-quality assessment, we undertake two applications. Both focus on important estimation problems that are difficult to resolve with network-derived monitoring data alone. The first is that of estimating the proportion of streams in regions of the conterminous United States in which the concentration of a contaminant exceeds an established criterion or threshold concentration. As discussed above, accurate estimates of such proportions are an important goal of federally mandated assessments under the Clean Water Act but are difficult to obtain when spatial biases exist in the sampling network. We show that the model developed here can be used to overcome the effects of spatial sampling biases in estimating the proportion of U.S. watersheds (i.e., hydrologic cataloging units [*Seaber et al*., 1987]) having phosphorus concentrations at the watershed outflow that exceed the commonly accepted criterion of 0.1 mg/L.

The second application illustrates a method for prioritizing watersheds for nonpoint-source pollution controls. Currently, there is considerable interest in increasing the efficiency of pollution control programs, especially nonpoint-source control programs, by focusing control efforts on watersheds where they are most needed. Stream monitoring alone does not provide sufficient information to compare watersheds because the effects of local pollution sources on in-stream water quality cannot be separated from the effects of contaminants originating in upstream watersheds. We show that the spatially referenced model developed here can be used to compare the hydrologic cataloging units on the basis of local nutrient yields. Through the use of “bootstrap” methods [*Efron*, 1982], we obtain robust estimates of the accuracy of these predictions.

The paper is organized into five sections. Following the introduction, a methods section includes a mathematical development of the underlying model and a description of the data sets and statistical procedures used to build the model. Section 3 presents the results of the regression and error analyses and includes a description of the bootstrap procedures used in error analysis. Section 4 presents the two model applications, and section 5 contains a summary discussion.

2. Methods

2.1. Overview of the Method

This section describes construction of a statistical model relating in-stream water-quality measurements to spatially referenced watershed attributes. Spatial referencing of land-based and water-based variables is accomplished via superposition of a set of contiguous land-surface polygons on a digitized network of stream reaches that define surface-water flow paths for the region of interest. Water-quality measurements are available from monitoring stations located in a sub-

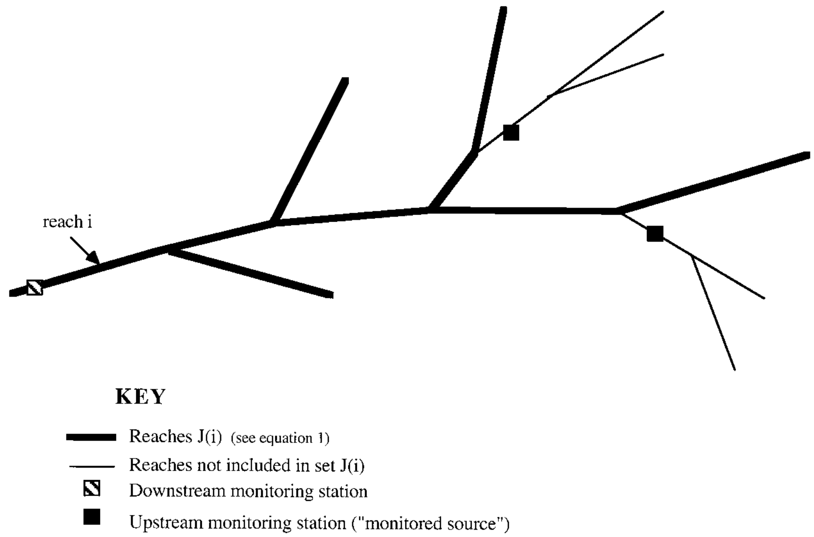

Figure 1. Schematic diagram of stream reaches in relation to monitoring stations (reaches extend from one tributary junction to another). In calibrating the model, reach i refers to any reach containing a monitoring station. In applying the model, reach i refers to any reach where a prediction is made. $J(i)$ represents the set of reaches that includes reach i and all upstream reaches except those that either contain or are located upstream of monitoring stations upstream of i.

set of the stream reaches. Water-quality predictors in the model are developed as functions of both reach and land-surface attributes and include quantities describing contaminant sources (point and nonpoint) as well as factors associated with rates of material transport through the watershed (such as soil permeability and stream velocity). Predictor formulae describe the transport of contaminant mass from specific sources to the downstream end of a specific reach. Loss of contaminant mass occurs during both overland and in-stream transport. In calibrating the model, measured rates of contaminant transport are regressed on predicted transport rates at the locations of the monitoring stations, giving rise to a set of estimated linear and nonlinear coefficients from the predictor formulae. Once calibrated, the model is used to estimate contaminant transport and concentration in all stream reaches. A variety of regional characterizations of water-quality conditions are then possible based on statistical summarization of reach-level estimates.

2.2. Model Development

The mathematical core of the SPARROW method is a relation that expresses the in-stream load (i.e., transport rate) of a contaminant at the downstream end of a given reach as the sum of monitored and unmonitored contributions to the load at that location from all upstream sources.

Let $J(i)$ represent the set of reaches j that includes reach i and all upstream reaches except those that either contain or are located upstream of monitoring stations upstream of i (Figure 1). We define

$$L_i = \sum_{n=1}^{N} S_{n,i} \tag{1}$$

where L_i is contaminant transport in reach i, $S_{n,i}$ is the contaminant load from source n delivered to reach i from all reaches in the subbasin delineated by $J(i)$.

In general, the N sources referred to in (1) include all types of point and nonpoint sources of the contaminant in the region of interest. We assume that monitoring stations are located near the downstream end of the reaches that contain them. Thus in-stream contaminant loads at the location of monitoring points upstream of i enter $J(i)$ and are included as one of the N sources. We refer to these entering in-stream loads as "monitored sources" (see Figure 1).

The source terms $S_{n,i}$ includes the effects of a two-stage delivery process operating on the contaminant mass as it moves through the watershed. The first stage dictates the proportion of contaminant mass that is delivered from the land surface to the channel network at reach j. For nonpoint sources the first-stage delivery process is mediated by a vector of reach-specific land-surface characteristics Z_j, which influence land-water delivery. Point sources and monitored sources enter reach j directly and are not influenced by Z_j. The second stage of the delivery process, applicable to point sources, nonpoint sources, and "monitored sources" (see above), determines the proportion of contaminant present in the channel in reach j that is transported to reach i. This in-stream delivery process is assumed to result from first-order decay of contaminant mass expressed as a function of a vector of channel characteristics $T_{i,j}$. These channel characteristics, evaluated over the entire channel length from reach j to reach i, include time-of-travel and channel size. Thus the source terms $S_{n,i}$ are expressed as

$$S_{n,i} = \sum_{j \in J(i)} s_{n,j} D_n(Z_j) K(T_{i,j}) \tag{2}$$

where $s_{n,j}$ is a measure of the contaminant mass from source n that is present in the drainage to reach j, $D_n(Z_j)$ is the proportion of $s_{n,j}$ that is delivered to reach j as a function of land-surface characteristics Z_j, and $K(T_{i,j})$ is the proportion of contaminant mass present in reach j that is transported to reach i as a function of channel characteristics $T_{i,j}$.

In distributing the land-surface contaminant source data to stream reaches, the quantities $s_{n,j}$ are obtained as the reported quantity of contaminant from source n present in the land-surface polygon containing reach j times the ratio of the length of reach j to the total length of all reaches in the polygon. That is, if $P(j)$ is the set of all reaches in the polygon containing reach j,

$$s_{n,j} = s_n\left(l_j \Big/ \sum_{j \in P(j)} l_j\right) \tag{3}$$

where s_n is the reported quantity of contaminant from source n present in the land-surface polygon containing reach j and l_j is the length of reach j.

Empirical implementation of the model given in (1) and (2) requires an explicit functional form for the delivery factors $D_n(Z_j)$ and $K(T_{i,j})$. The land-water delivery factor is parameterized as

$$\begin{aligned} D_n(Z_j) &= \beta_n \exp(-\alpha' Z_j) \\ D_n(Z_j) &= \beta_n \\ D_n(Z_j) &= 1 \end{aligned} \tag{4}$$

for nonpoint sources, point sources, and upstream monitored loads, respectively, where β_n is a source-specific coefficient and α is a vector of delivery coefficients associated with the land-surface characteristics Z_j.

Because reported contaminant-source data may include surrogate measures or may understate or overstate the true availability of contaminant mass, the coefficients β_n may deviate from unity. Also, the land-surface characteristics Z_j may be either positively or negatively related to contaminant delivery. If a land-surface characteristic is thought to be positively related to delivery (stream density, for example), its element of Z_j is taken to be the reciprocal of the characteristic, and if a land-surface characteristic is thought to be negatively related to delivery (soil permeability, for example), its element of Z_j is taken to be the characteristic itself. In either case, the land-water delivery coefficients α are expected to be positive.

The polygon-based land-surface characteristics Z_j are distributed to stream reaches by assuming the land-surface characteristic for reach j to be equal to the sum of the length-weighted land-surface characteristic of all polygons associated with reach j. Thus, if $P(x)$ is the set of all polygons containing reach j,

$$Z_j = \sum_{j \in P(x)} (z_{x,j}\, l_{x,j}/l_j) \tag{5}$$

where $z_{x,j}$ is the land-surface characteristic of polygon x associated with reach j, $l_{x,j}$ is the length of the portion of reach j associated with polygon x, and l_j is the total length of reach j.

The in-stream delivery terms $K(T_{i,j})$ are parameterized as

$$K(T_{i,j}) = \exp(-\delta' T_{i,j}) \tag{6}$$

where δ is a vector of first-order decay coefficients associated with the flow path characteristics $T_{i,j}$. Details of the specific form of the $T_{i,j}$ are discussed in section 2.5.

2.3. Estimation of the Model

An estimable version of the model given by (1), (2), (4), and (6) is obtained by introducing a multiplicative error term e^{ε}, where ε_i is assumed to be independent and identically distributed across nonoverlapping subbasins. Applying a logarithmic transformation of (1), we obtain the estimable model

$$L_i = \ln\left(\sum_{n=1}^{N} S_{n,i}\right) + \varepsilon_i \tag{1'}$$

where L_i is the natural logarithm of transport. The delivered sources $S_{n,i}$ appearing in (1′) continue to be given by (2), (4), and (6). Coefficient estimation was performed using the model procedure by *SAS Institute* [1993].

The assumption that the error term ε_i is independent across observations implies there is no correlation in the errors among the monitored basins. This assumption is consistent with the way the model treats nested basins. When one monitored basin contains another monitored basin, the model uses the monitored transport from the upstream basin (rather than the model-estimated transport) to represent contaminant sources entering the lower basin. Thus prediction errors that occur at the upstream monitored site do not cascade down to the lower monitored site and do not induce correlation across the subbasin error terms.

The following sections describe the contaminant sources, land-surface characteristics, and channel characteristics in the model, and their anticipated statistical relationship to contaminant transport. The data sources used in model calibration are also described.

2.4. River Reach Network

A 1:500,000-scale digital stream network for the conterminous United States (River Reach File 1-RF1 [*DeWald et al.*, 1985] defines surface water flow paths during model calibration and prediction. The network consists of approximately 60,000 reaches representing approximately 1 million km of total channel length (mean reach length is 17 km). Reach attributes include estimates of mean streamflow and water velocity [*DeWald et al.*, 1985]. Average stream velocity was estimated for each reach using regression equations relating velocity to long-term mean streamflow and stream order. The regressions were calibrated using U.S. Geological Survey time-of-travel studies. The reaches associated with 2100 large reservoirs (normal capacity greater than 5000 ac ft (1 ac ft = 1233 m^3) [see *Ruddy and Hitt*, 1990]) were also designated in the reach attribute file. Finally, the stream network contains numerous instances of diversions and stream braiding. These features were incorporated into the model with the assumption that contaminant concentration is uniformly distributed in the channel cross section. Thus, if a diversion diverts 20% of streamflow, it is also assumed to divert 20% of the load predicted at the point of the diversion.

2.5. Channel Transport Characteristics

In-stream losses of contaminant mass occur as a function of three variables: the travel time, streamflow (serving as a surrogate for channel depth), and whether or not the reach is part of a reservoir. Travel time is computed as the ratio of reach length over stream velocity. Because the major processes involved in in-stream loss of total phosphorus (TP) and total nitrogen (TN) (sedimentation and denitrification) operate at the channel bottom, deeper streams are expected to have lower rates of decay [*Howarth et al.*, 1997]. To incorporate this effect, we divide streams into three flow classes: <28.3 m^3/s (1000 ft^3/s), 28.3–283 m^3/s and >283 m^3/s (10,000 ft^3/s). We then separate the time of travel between reach i and reach j into the total time of travel over each of the three classes of streams. Thus the flow path characteristics vector T_{ij} consists of three variables representing the travel times associated with transport in the three stream classes. A network climbing algorithm [*White et al.*, 1992] performs the required accumulation. In estimating the decay coefficients (δ in (6)), we expect monotonically decreasing values with increasing streamflow (i.e., increasing channel depth).

TP retention in reservoirs may be expected to differ from TP decay in streams because of differences in settling rates of sediment-bound phosphorus in the two environments. To estimate this effect, we define a fourth stream class in the TP model consisting of reaches that are classified as reservoirs. Accordingly, a fourth element is included in the flow path characteristics vector T_{ij}, consisting of the time of travel in reservoir reaches between reach i and reach j. Reservoir reaches are excluded when computing time of travel for the three flow class variables. Exploratory regressions showed reservoir retention is not a significant factor in channel decay of TN.

2.6. Dependent Variable Data

Data from the U.S. Geological Survey's (USGS) National Stream Quality Accounting Network (NASQAN) were used to develop observational data for model calibration. Commencing during the period 1973–1978, NASQAN records consist of quarterly to monthly water column measurements from approximately 400 stations located near the outlets of selected U.S. hydrologic cataloging units [*Alexander et al.*, 1996; *Seaber et al.*, 1987]. Stations with indeterminant drainage area and stations with significant Mexican or Canadian drainage were not used in the analysis. We estimated long-term mean annual transport for TP and TN (sum of dissolved nitrate + nitrite and total Kjeldahl nitrogen measurements) at NASQAN stations using bias-corrected, log-linear regression-based load-estimation techniques [*Cohn et al.*, 1989, 1992; *Gilroy et al.*, 1990]. We modeled the total nutrients TN and TP rather than their component species because source data are generally unavailable for the latter.

Periodic instantaneous measurements of nutrient transport for the period 1974–1989 (or period of record for stations with shorter records; the number of observations ranged from 60 to 120) were regressed on a set of up to five explanatory variables. The full model is of the form

$$\ln(L) = \lambda_0 + \lambda_1 t + \lambda_2 \sin(2\pi t) + \lambda_3 \cos(2\pi t) + \lambda_4 \ln(q) + \lambda_5 [\ln(q)]^2 + e \qquad (7)$$

where L is the instantaneous nutrient transport, t is decimal time, q is instantaneous discharge, e is the sampling and model error assumed to be independent and identically distributed, ln is the natural logarithm, $\sin(2\pi t)$ and $\cos(2\pi t)$ are trigonometric functions that jointly approximate seasonal variations in transport, and the λ are regression coefficients. The model with the minimum prediction error sum of squares (PRESS [*Montgomery and Peck*, 1982]) was selected as the "best" fit from among the 15 possible models (seasonal terms enter and exit as a pair).

Using the regression parameters from (7), we estimated the annual transport that would have occurred at each station in 1987 if streamflow had corresponded to average conditions over the period 1970–1988. These estimates of mean annual transport for the base year 1987 (L) become the dependent variable in the SPARROW calibrations described below, and are obtained as

$$L = \frac{1}{365}\sum_{d=1}^{365} \exp[\lambda_0 + \lambda_1 t'_d + \lambda_2 \sin(2\pi t'_d) + \lambda_3 \cos(2\pi t'_d) + \lambda_4 \ln q_d + \lambda_5(\ln q_d)^2] g_{m,d} \qquad (8)$$

where q_d is the mean of daily streamflow values for the dth day of the year over the 1970–1988 period, t'_d is decimal time for the dth day in 1987, and $g_{m,d}$ is the minimum variance bias-correction factor of *Bradu and Mundlak* [1970; see *Cohn et al.*, 1989; *Gilroy et al.*, 1990], where m is the degrees of freedom of the best fit regression based on (7).

To reduce the effects of measurement error in the SPARROW calibrations, we excluded stations where the standard error of transport estimation was larger than 20% of the mean estimated transport. TN calibrations were based on a total of 414 stations, and TP calibrations were based on 381 stations.

2.7. Contaminant Source Data

Five specific contaminant sources are included in the TN and TP models calibrated in this study: point sources, applied fertilizers, livestock wastes, runoff from nonagricultural land, and atmospherically deposited nitrogen. (The atmosphere is assumed to be a negligible source of total phosphorus.) As described above, measurements of TN and TP transport at upstream monitoring sites are included in the model as instream sources of contaminant for downstream sites. This section describes the sources in greater detail and explains how they appear in the model.

Contaminant-source data were obtained for years corresponding as close as possible to 1987, the year specified in (8) as the base year for transport estimation at monitoring stations. Data for land-distributed sources were obtained as polygonal files representing either counties, cataloging units, physiographic regions, or contoured surfaces derived through spatial interpolation of point data.

County-based estimates of the nitrogen and phosphorus content of applied fertilizers [W. Battaglin and D. Goolsby, U.S. Geological Survey, written communication, 1996; *U.S. Environmental Protection Agency*, 1990a] and livestock wastes were obtained for 1987. The nitrogen and phosphorus content of livestock wastes was estimated using 1987 Census of Agriculture farm-animal population counts [*U.S. Bureau of Census*, 1989] and estimates of per-animal nutrient production rates [*U.S. Soil Conservation Service*, 1992].

Estimates of atmospheric wet deposition of nitrate were determined through a linear spatial interpolation of 1987 mean deposition at 188 wet-fall monitoring stations in the National Atmospheric Deposition Program [*National Atmospheric Deposition Program*, 1988, written communication, 1995]. Because wet nitrate deposition represents only a fraction of total nitrogen deposition (recent data suggest that total deposition may be 3–4 times wet nitrate deposition [*Stensland et al.*, 1986]), the source coefficient for atmospheric deposition is expected to exceed unity.

In addition to those identified above, we include an additional nonpoint source representing contaminants contained in the runoff from all types of nonagricultural land. The magnitude of this source is assumed to be proportional to the total area of urban, forest, and range land in the basin as given by the National Resources Inventory county estimates of land cover for 1987 [*U.S. Soil Conservation Service*, 1989]. Because the data used to represent this source are not expressed in mass units, the source coefficient is not expected to approximate unity.

County-based estimates of point-source discharges of total phosphorus and Kjeldahl nitrogen [*Gianessi and Peskin*, 1984] were obtained from an inventory of 32,000 point-source facilities, including industries, municipal wastewater treatment plants, and small sanitary waste discharges for the years 1977–1981. This inventory represents the most recent nationally comprehensive compilation of point-source discharges of nitrogen and phosphorus. Reach-level point-source discharge estimates were obtained by disaggregating county-level estimates to reaches in proportion to the population density in the vicinity of the reaches. Reach-level estimates of 1980 population density [*U.S. Bureau of Census*, 1983] were determined by assigning the census-unit population to the nearest reach within 10 km and located in the same cataloging unit. Contaminant loads from point sources are delivered directly to streams, and the land-water delivery factor, $e^{(-\alpha Z_j)}$, is not applied. If the point-source data used in the regressions were perfectly accurate, we would expect the source-specific coefficient β_n to equal unity. However, available national-level data on point-source discharges of total nitrogen and total phosphorus [*Gianessi and Peskin*, 1984] refer to 1977–1981 conditions and, in many cases, are estimated from information on type of facility. Thus we do not restrict the source-specific coefficient for point sources.

Because the "monitored sources" (see above) emanating from upstream basins are, effectively, in-stream sources, their land-water delivery factor is constrained to unity. However, these loads are "decayed" according to the same in-stream processes assumed for all other sources. The fact that monitored transport rates appear on the predictive (as well as the response) variable side of the regression equation makes it important that transport be measured accurately. The accuracy filter applied to the transport estimates (see above) assures that this is so.

2.8. Land Surface Characteristics

Eight land surface characteristics are included as potential land-water delivery factors (Z_j in (4)) in SPARROW calibrations: soil permeability, slope, stream density, fraction of total area classified as wetlands, long-term average ambient air temperature, long-term mean precipitation, fraction of cropland that is irrigated, and inches of applied irrigation.

Areas with highly permeable soils are expected to "absorb" contaminants readily and divert them to the subsurface. Thus high rates of soil permeability are expected to decrease the delivery of contaminants to streams. Cataloging unit-based estimates of permeability were determined from digital data in the State Soil Geographic (STATSGO) database [*U.S. Soil Conservation Service*, 1994; *Schwarz and Alexander*, 1995]. Average depth-weighted permeability was computed for approximately 78,000 STATSGO soil geographic units and was summarized by cataloging unit.

Land surface slope is included in calibrations in the reciprocal form, implying an expected positive effect on contaminant delivery. This assumption is consistent with the view that higher velocities associated with steeper slopes decrease the effects of time-dependent decay processes. We note, however, that steeper slopes may also produce more subsurface flow [*Dunne and Leopold*, 1978], thereby reducing the delivery of contaminants. The source and methods of compilation of land-surface slope data are the same as for soil permeability [*Schwarz and Alexander*, 1995].

Stream density, defined as the ratio of channel length to drainage area, is included in the reciprocal form, indicating a positive effect on land-water delivery. A greater stream density implies land-surface contaminants travel shorter distances on average to reach streams. Estimates of stream density are computed directly from the length and area attributes of the stream network coverage.

Wetlands are widely recognized as effective filters for removing nutrients and are especially effective in removing nitrogen [*Howarth et al.*, 1996; *Seitzinger*, 1988]. To account for this potential effect, we include the fraction of the land surface classified as wetlands as an element of the delivery vector Z_j. County-level estimates of wetland area are taken from the 1987 National Resources Inventory [*U.S. Soil Conservation Service*, 1989].

Higher temperatures increase rates of denitrification [*Seitzinger*, 1988] and are expected therefore to decrease the delivery of TN to streams. Higher temperatures would also be expected to increase rates of nitrogen fixation, but because natural fixation is a relatively minor source of new nitrogen in most watersheds compared to cultural sources, and because denitrification is by far the most important sink for TN, we expect an overall negative effect of temperature on TN delivery. Because the most important processes affecting the transport of TP are physical rather than biochemical, we expect temperature to have an insignificant effect on TP delivery. Nevertheless, temperature is included as a potential delivery factor in the exploratory regressions for both TN and TP. Average ambient air temperature was obtained for 350 climate divisions in the conterminous United States for the period 1951–1980 [*U.S. Geological Survey*, 1986].

High rates of precipitation were expected to speed the delivery of contaminants to the stream network, resulting in higher transport rates. In this case, precipitation would be positively associated with delivery and should be included in the delivery vector Z_j in reciprocal form. However, preliminary analyses showed complications in using simple precipitation as an explanatory variable. Arid areas of the country with high rates of irrigation and low levels of precipitation, in some cases, have relatively high contaminant loads due, perhaps, to the efficiency with which irrigation runoff is returned to the stream. In any case, precipitation appeared to be inversely related to transport, and irrigation rate was poorly correlated with transport.

In an attempt to better isolate the separate effects of irrigation and precipitation, we formed two interaction variables. The first variable consisted of the reciprocal of applied irrigation (in inches per year) times the share of cropland that is irrigated. The second variable was the reciprocal of precipitation times the share of cropland that is not irrigated. Simultaneous inclusion of both variables might be expected to capture the separate effects of irrigation and precipitation on delivery. However, in many areas, irrigation is used as a supplement for precipitation. In such areas, the amount of runoff from fields is relatively constant, regardless of the level of irrigation. To allow for this effect, we included as a third variable the share of cropland that is irrigated.

The source for mean annual precipitation data is the same as that for temperature data [*U.S. Geological Survey*, 1986]. County estimates of annual irrigation water use were obtained for 1985 from U.S. Geological Survey data files [*Smith et al.*,

Table 1. Results of Parametric and Bootstrap Regressions of Total Nitrogen at 414 NASQAN Stations on Basin Attributes

		Exploratory Model		Final Model					
Model Parameters	Coefficient Units[a]	Parametric Coefficient	p	Parametric Coefficient	Parametric p	Bootstrap Coefficient	Lower 90% CI[b]	Upper 90% CI[b]	Bootstrap p
Nitrogen source β									
Point sources	dimensionless	0.4112	0.0004	0.3464	0.0049	0.4401	0.0864	0.8173	<0.005
Fertilizer application	dimensionless	2.798	0.0154	1.278	0.0022	1.433	0.6149	2.373	<0.005
Livestock waste production	dimensionless	1.340	0.1553	0.9723	0.0629	1.058	0.0859	1.919	0.005
Atmospheric deposition	dimensionless	3.334	0.2513	6.465	0.0033	6.555	3.270	9.323	<0.005
Nonagricultural land	kg/ha/yr	38.49	0.0154	14.67	0.0005	16.65	7.130	29.89	<0.005
Land to water delivery α									
Temperature	$°F^{-1}$	0.0228	0.0001	0.0196	0.0001	0.0198	0.0117	0.0259	<0.005
Slope[c]	%	0.2034	0.2187						
Soil permeability	h/cm	0.0295	0.0022	0.0442	0.0001	0.0447	0.0334	0.0572	<0.005
Stream density[c]	km^{-1}	0.0205	0.0124	0.0215	0.0095	0.0243	−0.0003	0.0450	0.025
Wetland[d]	dimensionless	0.7177	0.2962						
Irrigated land[e]	dimensionless	1.101	0.0001						
Precipitation[f]	cm	38.52	0.0057						
Irrigated water use[g]	cm	0.0772	0.3117						
In-stream decay[h] δ									
δ_1 ($Q < 28.3$ m^3/s)	d^{-1}	0.2917	0.0001	0.3758	0.0001	0.3842	0.2981	0.4768	<0.005
δ_2 (28.3 $m^3/s < Q <$ 283 m^3/s)	d^{-1}	0.1099	0.0001	0.1233	0.0001	0.1227	0.0621	0.1710	<0.005
δ_3 ($Q > 283$ m^3/s)	d^{-1}	0.0352	0.1794	0.0406	0.1321	0.0408	0.0176	0.0685	0.015
R^2		0.8822		0.8743					
Mean square error		0.4310		0.4543					
Number of observations		414		414					

[a]Dependent variable (nitrogen transport) in kilograms per year.
[b]Minimum bootstrap confidence intervals (CI).
[c]Variable enters the model in reciprocal form.
[d]Ratio of wetland area to total land area.
[e]Ratio of irrigated land area to total cropland area.
[f]Product of reciprocal precipitation and one minus the ratio of irrigated land area to total cropland area.
[g]Ratio of irrigated land area to irrigated water use.
[h]Decay coefficients fit separately for stream reaches with mean streamflow (Q) corresponding to indicated intervals. The streamflow interval breakpoints of 28.3 and 283 m^3/s correspond to 1000 and 10,000 ft^3/s, respectively.

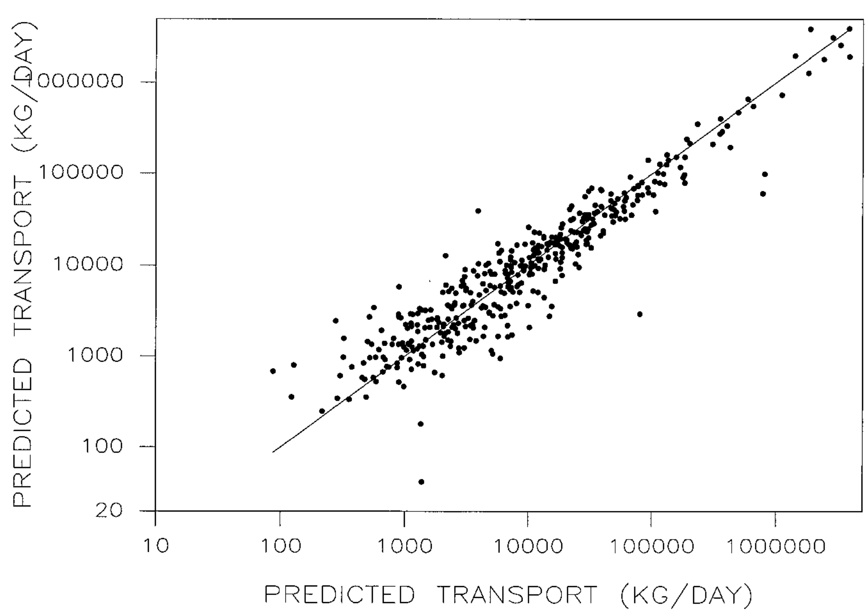

Figure 2. Predicted versus observed transport of total nitrogen.

Table 2. Results of Parametric and Bootstrap Regressions of Total Phosphorus Transport at 381 NASQAN Stations on Basin Attributes

Model Parameters	Coefficient Units[a]	Exploratory Model: Parametric Coefficient	Exploratory Model: p	Final Model: Parametric Coefficient	Final Model: Parametric p	Final Model: Bootstrap Coefficient	Final Model: Lower 90% CI[b]	Final Model: Upper 90% CI[b]	Final Model: Bootstrap p
Phosphorus source β									
Point sources	dimensionless	0.2507	0.0010	0.2972	0.0003	0.3033	0.0914	0.4651	<0.005
Fertilizer application	dimensionless	0.1242	0.0847	0.1267	0.0035	0.1332	0.0398	0.2042	<0.005
Livestock waste production	dimensionless	0.1640	0.0976	0.1973	0.0003	0.1884	0.1019	0.2821	<0.005
Nonagricultural land	kg/hectare/yr	0.3505	0.0461	0.4092	0.0001	0.4236	0.2769	0.5937	<0.005
Land to water delivery α									
Temperature	degrees °F.$^{-1}$	−0.0058	0.4471						
Slope[c]	percent	−0.0243	0.8338						
Soil permeability	hour/cm	0.0348	0.0068	0.0441	0.0001	0.0501	0.0262	0.0736	<0.005
Stream density[c]	km^{-1}	0.0467	0.0050	0.0401	0.0119	0.0290	−0.0196	0.0630	0.185
Wetland[d]	dimensionless	0.9229	0.3222						
Irrigated land[e]	dimensionless	0.0704	0.8034						
Precipitation[f]	cm	17.63	0.3481						
Irrigated water use[g]	cm	0.0716	0.4347						
In-stream decay[f] δ									
δ_1 ($Q < 28.3$ m^3/s)	day^{-1}	0.2391	0.0001	0.2584	0.0001	0.2680	0.1885	0.3497	<0.005
δ_2 (28.3 m^3/s $< Q <$ 283 m^3/s)	day^{-1}	0.0905	0.0342	0.0947	0.0271	0.0956	0.0156	0.1834	0.010
δ_3 ($Q > 283$ m^3/s)	day^{-1}	−0.0276	0.5819						
δ_4 (reservoir retention)[i]	day^{-1}	0.3378	0.0001	0.3377	0.0001	0.3586	0.2263	0.4697	<0.005
R^2		0.8179		0.8143					
Mean square error		0.7070		0.7074					
Number of observations		381		381					

[a]Dependent variable (phosphorus transport) in kilograms per year.
[b]Minimum bootstrap confidence intervals (CI).
[c]Variable enters the model in reciprocal form.
[d]Ratio of wetland area to total land area.
[e]Ratio of irrigated land area to total cropland area.
[f]Product of reciprocal precipitation and one minus the ratio of irrigated land area to total cropland area.
[g]Ratio of irrigated land area to irrigated water use.
[h]Decay coefficients fit separately for stream reaches with mean streamflow (Q) corresponding to indicated intervals. The streamflow interval breakpoints of 28.3 and 283 m^3/s correspond to 1000 and 10,000 ft^3/s, respectively. Channel time-of-travel excludes that for reaches associated with reservoirs.
[i]Decay coefficient based on channel time-of-travel for reaches associated with reservoirs. In nested F tests the coefficient is statistically separable (p values less than 0.01) from the reach decay coefficients.

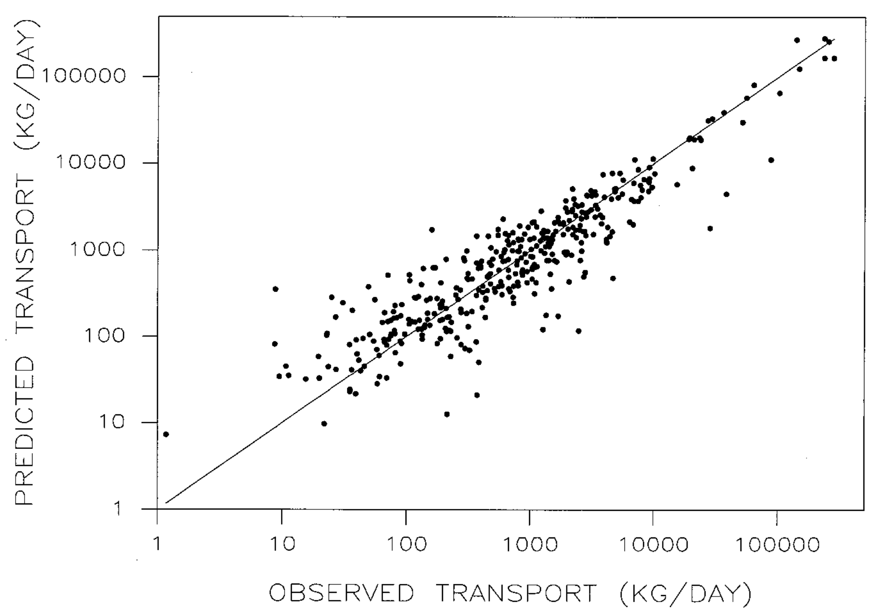

Figure 3. Predicted versus observed transport of total phosphorus.

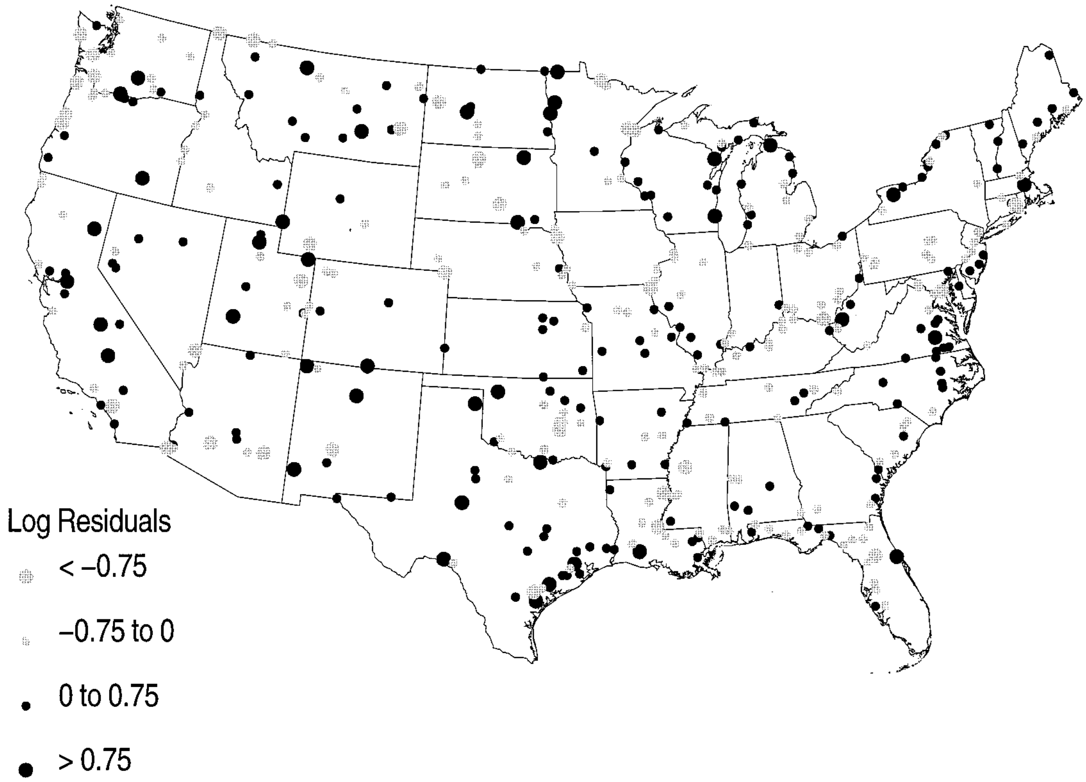

Figure 4. Total nitrogen residuals (predicted minus observed values) for 414 NASQAN stream monitoring locations in the conterminous United States.

1993a]. County-level estimates of the share of irrigated cropland are from the 1987 National Resources Inventory [*U.S. Soil Conservation Service*, 1989].

3. Results

The results of nonlinear least squares estimation of the parameters of the TN and TP models are presented in Tables 1 and 2, respectively. Coefficient estimates are presented in three groups: source-specific coefficients (β_n in (4)), land-water delivery coefficients (α in (4)), and in-stream decay coefficients (δ in (6)). Results for two alternative model specifications are included in the tables under the column headings Exploratory Model and Final Model. The exploratory models for both TN and TP contain the full set of estimated coefficients described in the Methods section. Note that a reservoir retention coefficient is not included in the TN model, and an atmospheric deposition coefficient is not included in the TP model. Preliminary regression of the TN model showed that there was no significant difference in nitrogen transport between reservoirs and streams. The final models were developed through elimination of coefficients from the exploratory models, primarily on the basis of statistical significance. The final models for both TN and TP perform nearly as well as the more highly parameterized alternatives in terms of prediction error and are used in the model applications described in the next section.

To assess the robustness of the parameter estimates, the final models undergo a bootstrap analysis [*Efron*, 1982]. The bootstrap procedure involves randomly selecting with replacement M monitored loads and their associated predictor variables from among the M observations in the original calibration data set (M is the number of monitored reaches in the reach network). In cases where a sampled observation has an upstream monitored load as one of its predictors, the monitored value is used, regardless of whether the upstream station appears in the bootstrap sample. A set of coefficient values is then estimated from the bootstrap sample. The bootstrap process is repeated 200 times, resulting in 200 estimates of the coefficients. From these estimates it is possible to determine the mean coefficient value (called the bootstrap estimate), a minimum confidence interval (evaluated as the minimum range of the bootstrap coefficient estimates such that the proportion of estimates lying inside the range equals the confidence level), and the probability that the estimated coefficient has the wrong sign (the p value or proportion of bootstrap coefficient estimates with the wrong sign).

The final models are analyzed graphically in this section as follows: plots of predicted versus observed transport are shown in Figures 2 and 3; the geographic distributions of regression residuals are mapped in Figures 4 and 5.

3.1. Total Nitrogen Model

Values of R^2 for the exploratory and final TN models are 0.88 and 0.87, and values for mean square error are 0.431 and 0.454, respectively. The two models differ only in the specification of the land-to-water delivery term. Coefficient estimates in the exploratory model for land slope, percent wetland, and irrigation water use are statistically weak and are eliminated from the delivery term with little effect on model performance. Elimination of precipitation and irrigation ratio from the exploratory model is based on three considerations. (1) Prelim-

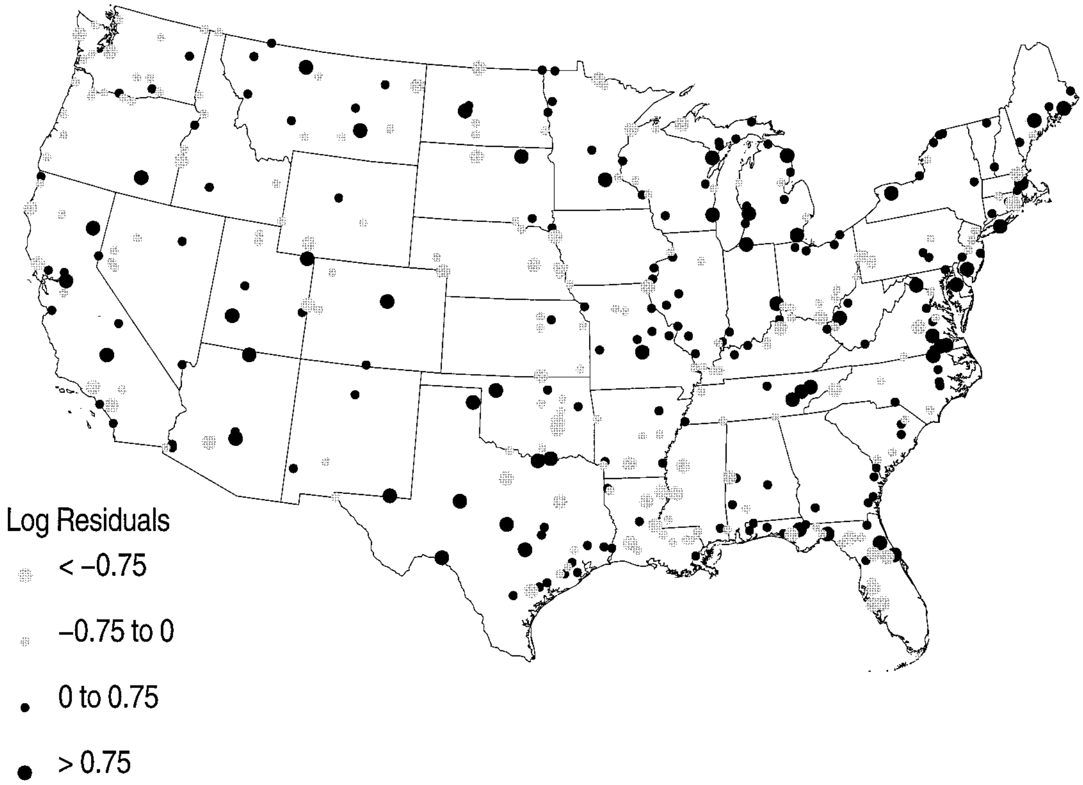

Figure 5. Total phosphorus residuals (predicted minus observed values) for 381 NASQAN stream monitoring locations in the conterminous United States.

inary regressions indicated that neither precipitation nor irrigation ratio, alone, is a significant predictor of TN transport. (2) Preliminary regressions also indicated that inclusion of precipitation, with or without irrigation ratio, interferes with successful estimation of the source coefficient for atmospheric deposition. We believe that precipitation may act as a surrogate for deposition because precipitation data are spatially and temporally more extensive than the deposition data and are correlated with deposition. (3) The results of TP model calibration (see below) show no significant predictive effect of either precipitation or irrigation ratio (the joint significance of these variables, as measured by a likelihood ratio test, is 0.24). The effect of these deletions is to greatly increase the significance levels of all four nonpoint-source coefficients with little loss of prediction accuracy. Of the five source coefficients in the model (that is, four nonpoint-source coefficients plus one point-source coefficient), four are highly significant ($p < 0.005$), and the fifth, livestock waste production, is moderately so ($p = 0.063$). The contributions of the five TN sources to predicted transport rate vary with location and are addressed in the model application section that follows.

The delivery variables that remain in the final TN model include temperature, soil permeability, and stream density. The temperature and soil permeability coefficients are highly significant ($p < 0.0001$). Although the t statistic for stream density is highly significant ($p = 0.009$), the bootstrap results show the coefficient is imprecisely estimated and only moderately significant.

In-stream decay rate coefficients in Table 1 are estimated for three stream size classes defined as a discrete function of streamflow: <28.3 m^3/s (1000 ft^3/s), 28.3–283 m^3/s, and >283 m^3/s (10,000 ft^3/s). Exploratory regressions indicated that estimated decay rate decreases monotonically with increasing stream size. The above classification is a somewhat arbitrary representation of this relationship. Two of the three decay rate coefficients are highly significant ($p < 0.0001$), whereas the coefficient for large streams is only weakly so (although it is highly significant in the bootstrap analysis). The insignificance of this coefficient stems from the relatively small value for decay in high flow reaches.

The plot of predicted versus observed transport for the TN Model (Figure 2) shows a tendency toward overprediction for sites with transport rates less than about 1000 kg/d. The basins in question tend to be relatively small and rural and receive a large portion of their predicted TN transport from nonagricultural land. Thus the bias appears to stem, in part, from the fact that no distinction is made in the model between various types of nonagricultural nonpoint sources. This might well be a shortcoming because forests tend to retain nitrogen to varying degrees [*Johnson*, 1992], whereas rangeland fixes nitrogen to varying degrees [*Rychert et al*., 1978], and urban runoff is frequently rich in nitrogen. Nevertheless, efforts to estimate specific source coefficients for forest and urban land were not successful in preliminary regressions.

The TN residuals map (Figure 4) shows little evidence of major regional biases in TN prediction. There is an apparent tendency for both positive and negative residuals to be larger west of the Mississippi River, in part, because transport measurement errors (dependent variable) are larger in western basins. At more local scales there are several examples of small regional clusters of predominantly positive or negative residuals, including a region of frequent positive residuals (that is, overprediction) along the Mid-Atlantic Coastal Plain and a region of underprediction in the far northwest. Interpretation of the underlying causes of such patterns raises the possibility of future improvements in model performance.

3.2. Total Phosphorus Model

Values of R^2 for the exploratory and final TP models are 0.82 and 0.81, and values for mean square error are 0.7070 and 0.7074, respectively. The parameters eliminated from the exploratory model include six from the land-to-water delivery term (temperature, land slope, percent wetland, precipitation, irrigation ratio, and irrigation water use) and the in-stream decay coefficient for streams >283 m^3/s. All eliminated parameters are statistically weak, and their loss has little effect on model performance. The final TP model has a similar structure to the final TN model with three exceptions. First, temperature is not correlated with land-to-water loss of phosphorus, presumably because microbiological processes do not remove phosphorus during this stage of total phosphorus transport. Second, reservoir retention is shown to be an important aspect of in-stream phosphorus transport. Finally, as previously noted, atmospheric deposition is assumed to be a negligible source of phosphorus and is not included in either the exploratory or final model. The four remaining source coefficients (point sources, fertilizer application, livestock waste production, and nonagricultural land) are all highly significant in the final TP model. As in the TN model, the soil permeability coefficient is highly significant, implying a strong negative correlation with phosphorus delivery, and stream density is positively linked to phosphorus delivery. Although the t statistic for stream density is moderately significant ($p = 0.0119$), the bootstrap results show the coefficient is imprecisely estimated and of questionable significance.

The plot of predicted versus observed transport for the TP model (Figure 3) shows a tendency toward overprediction for sites with transport rates less than about 100 kg/d. The pattern is similar to the pattern of prediction biases observed for the TN model (Figure 2) and involves many of the same sites. There is also evidence in Figure 3 of a tendency toward underprediction for TP transport rates greater than about 1000 kg/d.

The TP residuals map (Figure 5) shows little evidence of major regional patterns in either the sign or magnitude of TP prediction errors. As with TN errors (see Figure 4), however, there are regional clusters of predominantly positive or negative residuals, including several that appear to correspond approximately in location and sign to clusters of TN residuals. Nationwide, the Spearman rank correlation coefficient between TN and TP residuals is 0.61 ($N = 368$), suggesting significant commonality in the sources of prediction error.

3.3. Effects of Spatial Referencing

An important hypothesis of the present study is that spatial referencing of basin attributes increases the accuracy of regression-based predictions of water quality over that of nonspatially referenced models. We can test this hypothesis with respect to the SPARROW model by constructing nonspatially referenced models from the TN and TP data sets used here and comparing their error characteristics to those reported above. Spatial referencing exists in the SPARROW model through the integration of two spatial structures: (1) the set of land-surface polygons on which are mapped the nonpoint contaminant sources and the land-water delivery variables; and (2) the network of stream reaches on which are mapped the point sources, the channel transport characteristics, and the measured transport rates.

Table 3. Effect of Spatial Referencing on Measures of Regression Model Performance

	Mean Square Error		R^2	
Model Components	TN	TP	TN	TP
Includes full spatial referencing (SPARROW)[a]	0.4544	0.7074	0.8743	0.8143
Excludes in-stream decay and reservoir retention	0.9659[b]	1.1185[b]	0.7307	0.7041

TN, total nitrogen; TP, total phosphorus.
[a]Includes final model terms as specified in Tables 1 and 2.
[b]In a likelihood ratio test, the sum of squares of error of the spatially restricted model is significantly ($p \ll 0.001$) larger than that of the fully spatially referenced model (SPARROW).

The major effect of spatial referencing in SPARROW can be examined by eliminating the channel decay coefficients (including the reservoir-TP retention coefficient) from the model and estimating a new model containing only the contaminant sources and land-water delivery variables contained in the original model. As summarized in Table 3, loss of the spatial structure provided by the reach network increases the mean square error of the TN and TP models by 113 and 58%, respectively; R^2 values decrease by 14 and 11 points, respectively. A secondary effect of the loss of the reach network is an increase in the percent standard error of a majority of the source and delivery coefficient estimates (not shown in Table 3). This effect is especially pronounced in the nonspatially referenced TN model, where the coefficients for livestock wastes, nonagricultural land, and soil permeability are not statistically significant. The implication of these results is clear: knowledge of the location of contaminant sources and other basin attributes in terms of the approximate travel time and size of channel connecting them to specific stream locations contributes significantly to the accuracy of transport predictions and to the ability to determine the influence of specific sources and processes on transport.

4. Model Applications

One of the challenges of water-quality assessment programs is to describe regional water quality in useful and valid ways on the basis of data from sampling networks. Because of the problems of station sparseness, spatial bias, and basin heterogeneity (see above), network-derived water-quality data frequently are inadequate for regional characterization when they are interpreted simply as a sample of regional conditions. In the method presented here, network-derived data are interpreted in the context of a model linking in-stream water quality to spatially referenced information on contaminant sources and other watershed attributes relevant to contaminant transport. Once calibrated, the model serves two functions in regional assessment. First, the model is used to predict contaminant concentrations and transport rates at an appropriate set of stream locations for characterizing regional water quality. Second, the model is used to gain insight into the factors that influence water quality at those locations and, by inference, in the region in general.

Figure 6. Water resource regions in the conterminous United States.

In this section we present example applications of the TN and TP models developed above. In both applications, the region of interest is defined as nontidal watersheds in the conterminous United States. In characterizing water quality in this region, we have chosen to focus on the set of 2057 stream locations corresponding to the outflows of the eight-digit hydrologic "cataloging" units located in the 18 water resource regions of the conterminous United States [*Seaber et al*., 1987] (see Figure 6). These locations are a logical choice for national-level water-quality characterization in several respects: (1) the cataloging units represent a systematically developed and widely recognized delineation of U.S. watersheds; (2) the criteria for establishing unit boundaries [*Seaber et al*., 1987] were based on drainage area and other hydrologic considerations rather than cultural factors and provide a spatially representative view of water-quality conditions; (3) the accuracy of streamflow estimates for these locations is relatively high because gauging stations are located at or near the outflow of many of the units. Streamflow estimates are an important component of model predictions of contaminant concentrations and play an important role in the example application of the total phosphorus model presented below.

4.1. Total Phosphorus Model Application

One of the common objectives of water-quality assessment programs is to determine the proportion of water resources that meet specified quality criteria. Despite long-standing legal requirements (Section 305b of the Clean Water Act, Public Law 92-500) that states, river basin commissions, and the federal government regularly conduct such assessments, no nationally consistent method has emerged for determining criteria-based proportions. The problem stems, in part, from a frequently encountered dilemma in assessment strategy: there are strong incentives for targeting sampling to specific locations in order to determine the causes of poor water quality; but such nonrandomized sampling increases the difficulty of statistically characterizing regionwide water quality. The result has been a practice of dividing the state or region of interest into "assessed" and "unassessed" areas [*U.S. Environmental Protection Agency*, 1994] and a general inability to assign confidence intervals to estimated proportions.

As an illustration of the use of the method described here for interpreting data from nonrandomized sampling networks, we use the TP model to estimate regional proportions of the nation's cataloging units with TP concentrations meeting the widely accepted criterion of 0.1 mg/L [*U.S. Environmental Protection Agency*, 1976].

To fully incorporate model prediction error, the results presented below rely on bootstrap simulations. Simpler parametric approximations of the method are possible. In a two-step procedure, we make 200 stochastic predictions of TP transport for each cataloging unit outflow (if a cataloging unit has multiple outflows, the outflow with the largest streamflow is chosen). First, each iteration b of the procedure predicts transport at every cataloging unit outflow using a unique set of model coefficients drawn sequentially from the $B = 200$ sets developed during the previously described bootstrap coefficient estimation process (see Results section). The resulting point predictions reflect all of the covariances inherent in the coefficient estimates, as well as the covariances that arise from basing predictions at multiple cataloging units on a common set of coefficients. The second step of the procedure incorporates the effect of model error (the ε term in (1′)) by making use of the set of M residuals estimated from the M observations that went into the calibration of the bth set of coefficients. The M estimated residuals are first transformed by the exponential function. We then multiply each cataloging unit point prediction by one of the M randomly selected exponential errors to obtain a stochastic realization of transport. Use of this second step rests on the assumption that the model error is homoscedastic and independent across cataloging units.

Table 4. Proportion of Hydrologic Cataloging Units With Predicted Total Phosphorus (TP) Concentrations Not Exceeding the TP Criterion of 0.1 mg/L

Region[a]	Number of Cataloging Units	Proportion	Lower 90% CI[b]	Upper 90% CI[b]
United States (conterminous)	2048	0.394	0.373	0.416
New England	52	0.838	0.750	0.904
Mid-Atlantic	88	0.598	0.534	0.671
South Atlantic Gulf	191	0.580	0.524	0.623
Great Lakes	106	0.562	0.500	0.604
Ohio	120	0.511	0.450	0.575
Tennessee	32	0.709	0.563	0.781
Upper Mississippi	131	0.185	0.153	0.229
Lower Mississippi	82	0.471	0.390	0.549
Souris-Red-Rainy	42	0.217	0.143	0.286
Missouri	302	0.180	0.146	0.219
Arkansas-White-Red	171	0.189	0.140	0.240
Texas-Gulf	117	0.212	0.154	0.256
Rio Grande	67	0.344	0.254	0.418
Upper Colorado	62	0.339	0.258	0.419
Lower Colorado	75	0.108	0.067	0.160
Great Basin	61	0.241	0.115	0.361
Pacific Northwest	217	0.673	0.627	0.705
California	132	0.453	0.409	0.523

[a]Water resource regions of the United States shown in Figure 6.
[b]Minimum confidence intervals (CI).

The proportion of cataloging units with TP concentrations meeting the 0.1 mg/L criterion is estimated as follows. The 200 transport predictions made for the outflow of each cataloging unit are divided by estimated mean streamflow to obtain 200 estimates of flow-weighted mean concentration. Note that concentration estimates do not account for error in estimated streamflow at the cataloging unit outlets. For each bootstrap iteration b, we determine the proportion P_b of cataloging units that have estimated concentrations meeting the 0.1 mg/L criterion. Finally, we average P_b over all bootstrap iterations to determine the bootstrap estimated proportion of cataloging units meeting the criterion. A confidence interval for this proportion is computed by finding the minimum range over the 200 proportion estimates such that the fraction of estimates lying inside the range equals the confidence level.

The results of the TP model application are presented in Tables 4 and 5 and Figure 7. Table 4 shows the estimated proportions of cataloging units with "low" TP concentrations (that is, meeting the 0.1 mg/L TP criterion) for the 18 water resource regions of the conterminous United States. The values vary widely from region to region, ranging from 0.84 in New England to 0.11 in the Lower Colorado. The proportion of watersheds that meet the criterion is consistently less than 0.25 throughout the midcontinent region (Upper Mississippi, Missouri, and Arkansas regions) where agricultural sources of phosphorus are high. Nationally, the proportion of cataloging units meeting the criterion is estimated as 0.39. Figure 7 shows the locations of cataloging units classified according to their likelihood of meeting the TP criterion. It is clear from model input that the occurrence of predicted high TP concentrations throughout the arid West is more a reflection of low average streamflow than of high TP sources.

The most comparable previous characterization of TP concentrations in U.S. streams during the mid-1980s (*Smith et al.* [1993a]; see especially Figure 42a) is based on data from 410 monitoring stations and shows a similar geographic pattern of concentrations exceeding 0.1 mg/L with high frequency in the agricultural areas of the Midwest, northern and southern plains, and arid West. Nationally, the proportion of stations meeting the TP criterion in that study is 0.52. Because the sampling locations in the study do not represent a statistical sample of a precisely defined population, no confidence estimates accompany the estimate.

The sizes of the confidence intervals surrounding the estimated proportions in Table 4 vary geographically and are generally smaller for regions with larger numbers of cataloging units. The 90% confidence interval for the smallest region, the Tennessee (32 units), is 0.56–0.78, whereas that for the largest, the Missouri (302 units), is 0.15–0.22. Nationally, the 90% confidence interval surrounding the nonexceedance proportion for the total 2048 cataloging units is 0.37–0.41.

One of the advantages of combining model building with data collection in water-quality assessment programs is that models provide a potential link between the descriptive and explanatory aspects of assessment. Table 5 presents model estimates of several variables pertaining to TP sources and transport in watersheds classified according to their probability of exceeding the criterion. The first three columns in the table describe the shares (in percent) of total phosphorus transport contributed by the four major TP sources for the watersheds in the two probability classes. The percentages refer to the sizes

Table 5. Sources and Transport Factors Related to Predicted Total Phosphorus (TP) in Hydrologic Cataloging Units of the United States

Source	Share of TP Transport, %			Land-Water Delivery Factor[a]			Channel Transport Factor		
	10th	Mean	90th	10th	Mean	90th	10th	Mean	90th
Hydrologic Units With Concentrations Not Exceeding TP Criterion (n = 797)									
Point sources	0.0	9.1	28.0	0.297	0.297	0.297	0.503	0.731	0.924
Fertilizer application	1.5	15.7	34.9	0.038	0.067	0.088	0.513	0.703	0.878
Livestock waste production	6.0	26.0	48.0	0.059	0.104	0.136	0.507	0.697	0.874
Nonagricultural land	17.0	49.3	85.1	···	···	···	0.525	0.694	0.875
Hydrologic Units With Concentrations Exceeding TP Criterion (n = 1251)									
Point sources	0.0	7.9	22.4	0.297	0.297	0.297	0.120	0.491	0.831
Fertilizer application	0.9	21.0	46.0	0.048	0.072	0.090	0.169	0.462	0.751
Livestock waste production	10.2	37.7	63.8	0.075	0.111	0.140	0.166	0.452	0.743
Nonagricultural land	3.3	33.4	78.6	···	···	···	0.159	0.452	0.748

[a]Land-water delivery cannot be estimated for nonagricultural land because the source is not expressed in mass units.

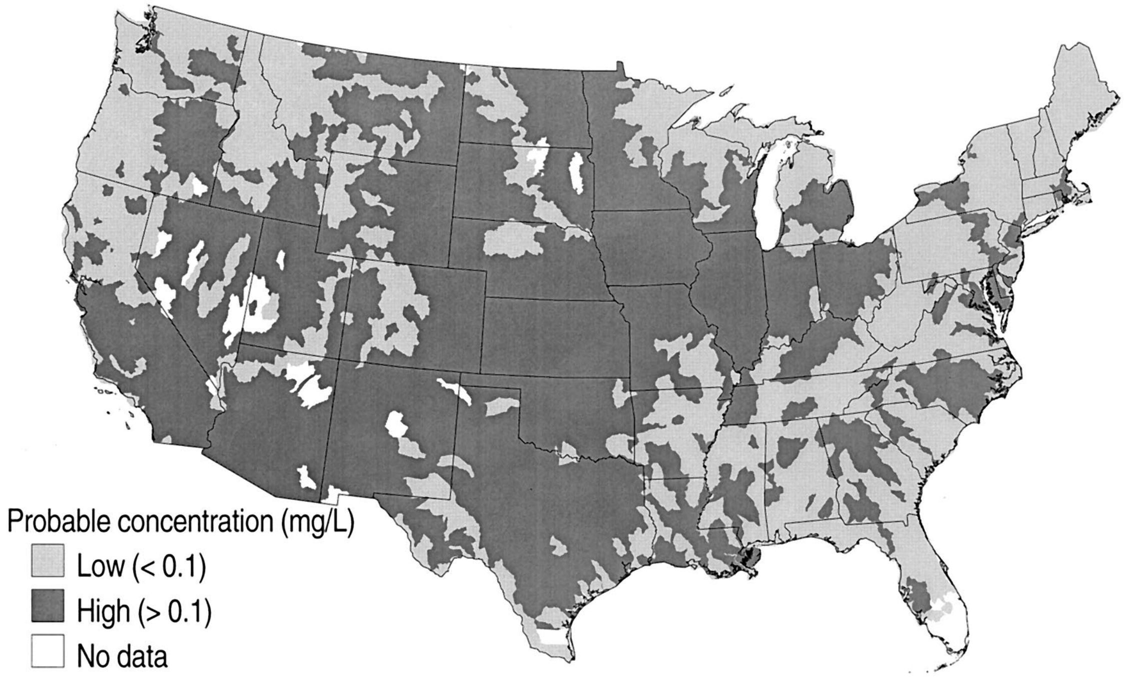

Figure 7. Classification of predicted total phosphorus concentrations in hydrologic cataloging units of the conterminous United States.

of the shares at the outflows of the cataloging units after accounting for losses associated with land-water delivery and channel deposition. Variation among watersheds is summarized in terms of the mean, 10th percentile, and 90th percentile values for the source shares. All results are averages of 200 bootstrap estimates (although bootstrap errors are not reported). Note, however, that whereas the estimates incorporate the effects of coefficient error, they do not include effects from model error.

The second and third groups of columns in Table 5 give the fraction of the contaminant mass from each source that is transported during land-water delivery and channel transport, respectively. The product of the mean values in each row gives the fraction of the source that is transported over the total path from its origin to the outflow of the cataloging unit. For example, an estimated 6.7% of the phosphorus applied as fertilizer in an average low-TP cataloging unit is delivered to stream channels, and an estimated 70.3% of that is ultimately transported to the unit outflow. Point sources in the model are not under the influence of any of the specified land-water delivery variables (such as soil permeability); the fact that the regression estimate of the point-source coefficient (0.297) deviates from unity is likely the result, in part, of error in the point-source inputs. One documented source of such error is a reduction in the magnitude of point-source discharges between the late 1970s when the data were compiled and 1987, the base year of the simulation (total phosphorus concentrations in primary and secondary effluent fell by as much as 50% from the 1970s to the late 1980s [*Gianessi and Peskin*, 1984; *National Research Council*, 1993]).

One clear pattern in Table 5 is that agricultural sources (applied fertilizer plus livestock waste) contribute a larger share of TP in watersheds with a high probability of exceeding the 0.1 mg/L criterion than in those with a low exceedance probability (59% versus 42% on average). The reverse pattern applies to the shares from nonagricultural nonpoint sources (34% in high probability units versus 49% in low probability units), whereas point sources contribute approximately the same share (8% versus 9%) to TP transport in both categories of watersheds. The association between exceedance probability and dominance of agricultural sources is not surprising, given the geographic distribution of high-TP watersheds (Figure 7). Nevertheless, the indication in Table 5 that livestock waste contributes more than applied fertilizer to TP transport in general is an interesting result in the context of water-quality assessment. The land-water delivery factors in Table 5 are nearly identical for the high-TP and low-TP cataloging units, whereas channel transport factors are larger among the low-TP units. The latter pattern appears to result from the fact that reservoir retention of phosphorus is especially important in the midcontinent region where TP is high in spite of their effect. Thus it seems that differences in transport processes among cataloging units do not, in general, explain differences in the probability of exceeding the TP criterion.

4.2. Total Nitrogen Model Application

Currently, there is considerable interest in increasing the efficiency of pollution control programs, especially nonpoint-source control programs, by focusing control efforts on watersheds where they will have the most effect. Stream monitoring alone does not provide sufficient information to prioritize watersheds because the effects of local pollution sources on in-stream water quality cannot be separated from the effects of contaminants originating in upstream watersheds. Similarly, information on the size of pollution sources in watersheds is not sufficient for directing controls because transport processes heavily influence in-stream quality.

In this application, we use the TN model to classify cataloging units on the basis of their nitrogen yield (transport per unit area) from local sources alone, independent of upstream units.

Table 6. Proportion of Hydrologic Cataloging Units With Predicted Total Nitrogen Yield <500 and <1000 kg/km^2/yr

Region[a]	Number of Cataloging Units	Yield <500 kg/km^2/yr Proportion	Lower 90% CI[b]	Upper 90% CI[b]	Yield <1000 kg/km^2/yr Proportion	Lower 90% CI[b]	Upper 90% CI[b]
United States (conterminous)	2057	0.603	0.568	0.634	0.819	0.781	0.843
New England	52	0.430	0.308	0.539	0.787	0.712	0.904
Mid-Atlantic	88	0.290	0.205	0.352	0.647	0.568	0.727
South Atlantic Gulf	191	0.565	0.492	0.634	0.871	0.822	0.911
Great Lakes	106	0.374	0.311	0.434	0.625	0.547	0.689
Ohio	120	0.206	0.142	0.267	0.563	0.467	0.633
Tennessee	32	0.292	0.156	0.406	0.724	0.594	0.844
Upper Mississippi	131	0.235	0.176	0.290	0.513	0.420	0.595
Lower Mississippi	82	0.392	0.305	0.476	0.713	0.610	0.793
Souris-Red-Rainy	42	0.704	0.571	0.810	0.893	0.810	0.952
Missouri	302	0.773	0.738	0.801	0.905	0.874	0.934
Arkansas-White-Red	171	0.694	0.620	0.749	0.903	0.860	0.947
Texas-Gulf	117	0.724	0.658	0.786	0.912	0.863	0.949
Rio Grande	67	0.917	0.866	0.955	0.968	0.940	0.985
Upper Colorado	62	0.925	0.871	0.968	0.982	0.951	1.000
Lower Colorado	83	0.975	0.952	1.000	0.993	0.976	1.000
Great Basin	62	0.925	0.887	0.968	0.978	0.952	1.000
Pacific Northwest	217	0.690	0.590	0.783	0.895	0.848	0.945
California	132	0.559	0.470	0.621	0.775	0.689	0.833

[a]Water resources regions of the United States shown in Figure 6.
[b]Minimum confidence intervals (CI).

Such a classification is useful for comparing watersheds in terms of their contribution to the quantity of contaminant present in a region in general. Again, our procedure uses bootstrap methods to account for prediction error. Because we are interested in local yields, however, we do not allow predicted loads to move beyond cataloging unit boundaries and do not use monitored loads to make load predictions. If a cataloging unit has multiple outflows, the predicted loads from all its outflows are summed.

The results of the TN model application are presented in Tables 6 and 7 and Figure 8. One of the striking features of the geographic pattern of TN yields in Table 6 and Figure 8 is that it differs markedly from the regional pattern of TP concentrations (compare with Table 4 and Figure 7). TN yields, like TP concentrations, are high in the agricultural Midwest but are low in the Plains and Southwest and high in the North Atlantic drainage, the inverse of the TP concentration pattern. There are two reasons for the differences. First, the variable of interest in the TN application, yield, is not increased by the lower streamflow conditions of the western drainages, as concentration is in the TP application. Second, TN yields are increased significantly in the North Atlantic drainage by atmospheric sources of nitrogen, which do not apply to TP. According to Table 7, atmospheric deposition contributes up to 33% (90th percentile) or more of the nitrogen leaving the high-yield cataloging units.

Table 7 provides information on several factors that appear to distinguish high-yield from low-yield watersheds. The importance of applied fertilizer is most obvious, averaging 48% of TN yield in high-yield units and exceeding a 72% share in 10% of those units. Point sources also contribute a significantly higher share of yield in high-yield units than in low-yield units (mean of 10.6% versus 2.4%). Livestock wastes contribute only slightly more to yield in high-yield than low-yield units (mean of 15.4% versus 12.8%) and contribute significantly less to

Table 7. Sources and Transport Factors Related to Predicted Total Nitrogen (TN) in Hydrologic Cataloging Units of the United States

Source	Share of TN Transport, % 10th	Mean	90th	Land-Water Delivery Factor[a] 10th	Mean	90th	Channel-Transport Factor 10th	Mean	90th
Hydrologic Units With Yield <500 kg/km^2/yr (n = 1253)									
Point sources	0.0	2.4	4.8	0.347	0.347	0.347	0.059	0.444	0.841
Fertilizer application	1.0	20.3	50.6	0.166	0.277	0.394	0.102	0.403	0.746
Livestock waste production	2.7	12.8	24.6	0.127	0.211	0.300	0.100	0.395	0.744
Atmospheric deposition	6.8	18.3	32.3	0.843	1.40	1.99	0.103	0.392	0.729
Nonagricultural land	13.1	46.3	78.0	···	···	···	0.099	0.390	0.737
Hydrologic Units With Yield >1000 kg/km^2/yr (n = 271)									
Point sources	0.0	10.6	39.1	0.347	0.347	0.347	0.410	0.695	0.896
Fertilizer application	12.1	47.9	71.7	0.255	0.320	0.373	0.409	0.642	0.843
Livestock waste production	3.0	15.4	27.7	0.194	0.244	0.284	0.400	0.634	0.828
Atmospheric deposition	7.5	18.0	32.6	1.29	1.62	1.89	0.429	0.645	0.828
Nonagricultural land	1.8	8.2	18.0	···	···	···	0.449	0.650	0.832

[a]Land-water delivery cannot be estimated for nonagricultural land because the source is not expressed in mass units.

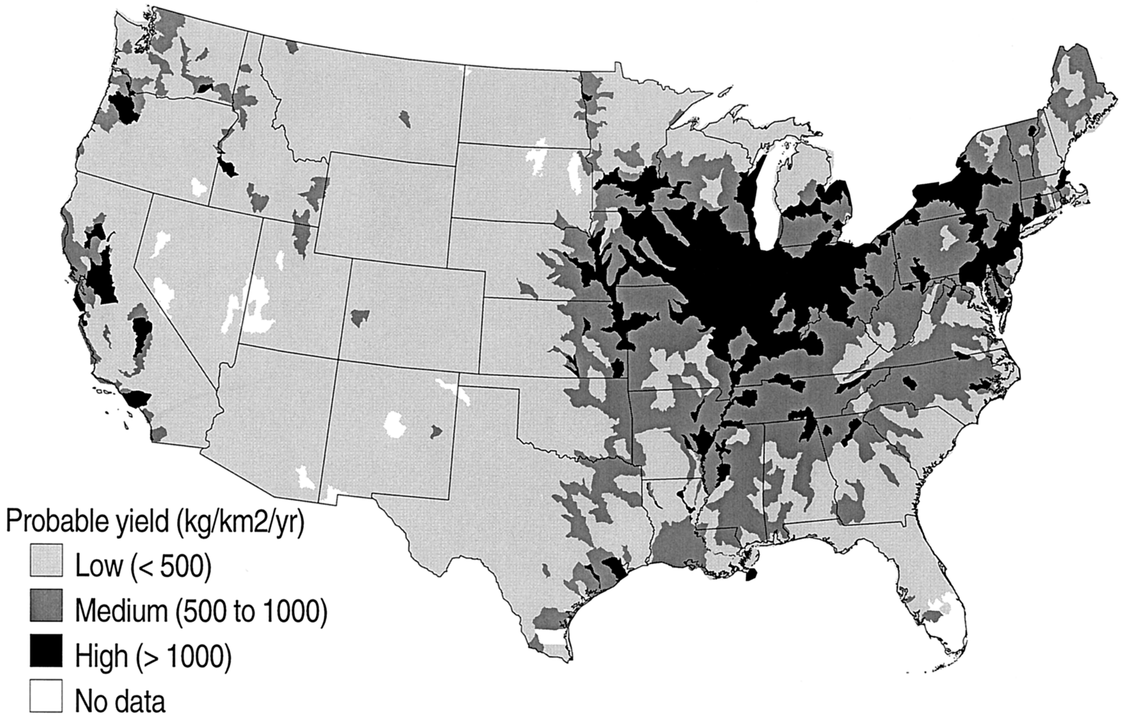

Figure 8. Classification of predicted local total nitrogen yield in hydrologic cataloging units of the conterminous United States. Local yield refers to transport per unit area at the outflow of the unit due to nitrogen sources within the unit, independent of upstream sources.

yield than applied fertilizer in both yield classes, a notable departure from the pattern observed for TP (Table 5).

Finally, there are several noteworthy results pertaining to the transport factors in Table 7. As for TP (Table 5), land-water delivery factors for TN are similar for the two yield classes. TN delivery factors for agricultural sources, however, are much higher than those for TP, differing by more than a factor of 4 in the case of applied fertilizer. The difference is consistent with the well-established greater mobility of TN compared to TP in soils and watersheds in general, stemming from the adsorptive and sedimentary properties of TP. Delivery factors in Table 7 for atmospheric deposition exceed unity, implying an upward bias in estimating the source coefficient. It seems likely that much of the apparent bias stems from the fact that atmospheric inputs in the TN model are based on measured wet nitrate deposition and ignore the nonnitrate and dry components, which commonly contribute more than 60% of total nitrogen deposition [*Stensland et al.*, 1986].

5. Discussion

5.1. Model Reliability

The bootstrap simulations of prediction errors included in both the model construction and application phases of this study were designed to provide coefficient estimates, model predictions, and confidence intervals that are robust with respect to the characteristics of both model and sampling errors. Some discussion is in order, however, of certain aspects of error that are not fully accounted for in the results. As previously noted, there is evidence in Figure 4 of larger errors in TN transport west of the Mississippi River than in eastern basins. The bootstrap analysis fully accounts for the observed magnitude of prediction errors but does not recognize regional patterns in error. Thus we expect that the true confidence intervals surrounding the TN yield proportions for eastern regions are somewhat narrower than those stated in Table 6 and are somewhat wider than those estimated for western regions. Such geographic (or other) patterns in model precision will have no effect on the estimated proportions per se. By contrast, prediction biases do potentially affect the estimated proportions. For example, the tendency for the TP and TN models to overpredict loads in certain small basins is accounted for in the bootstrap analysis in terms of error magnitude but not in terms of direction. If there were large differences in the numbers of affected basins among the water resource regions, the estimated regional proportions would be biased. The lack of evidence of major regional biases in either the TN or TP residuals maps (Figures 4 and 5) suggests that the effect is generally small. The localized clusters of same-sign residuals that are evident in the maps do provide evidence of spatial dependence of errors on a subregional scale that could affect the results for certain of the smaller water resource regions. We also hasten to note that success in interpreting the nature of these spatial dependencies would provide a means for significantly reducing model bias and increasing precision.

In addition to predicting nutrient concentrations and yields, the models developed in this study are designed to provide some insight into the important sources and processes affecting nutrients in watersheds. The reliability of estimates of source shares, land-water delivery factors, and in-stream transport factors (Tables 5 and 7) can be evaluated to a limited extent in terms of the estimation error of model coefficients and the overall accuracy of model predictions. The generally high precision of coefficient estimates and moderately high R^2

values obtained in this study therefore encourage tentative use of the models in process interpretation. However, the ultimate limitation on use of this method in the explanatory aspects of water-quality assessment lies in the necessary oversimplification of transport processes that is inherent in the model structure.

One of the important variables influencing the precision of the SPARROW method in estimating proportions and other regional water-quality descriptors is the number of independent point estimates of transport that are aggregated in the regional estimates. The widths of the confidence intervals shown in Tables 4 and 6 decrease as the number of cataloging units per region increases. The full range of this effect becomes evident in comparing the precision of individual point predictions of transport, which are commonly in error by 50% or more, with the precision of the estimated national TN and TP proportions (based on the total 2046 units), which are estimated to be in error by only a few percentage points. In general, the most useful applications of the SPARROW method are likely to involve analysis of aggregate properties of regional watersheds. For a model with given error characteristics, the effective minimum region size for useful analysis becomes largely dictated by the number of statistically independent locations that are available for prediction. At the other extreme, the accuracy of aggregate statistics for large regions with many independent prediction points is ultimately limited by the estimation error of the model coefficients.

5.2. Concluding Remarks

Interpretive models are an important adjunct to data collection in regional water-quality assessment programs as a means of addressing the problems imposed by limited sampling resources, network bias, and basin heterogeneity. Models relating water-quality measurements to contaminant sources and other watershed attributes provide a means of linking the descriptive and explanatory functions of water-quality assessment. The results of this study demonstrate that spatial referencing of in-stream measurements in relation to basin attributes greatly increases the predictive accuracy and explanatory potential of regression-based water-quality models. Spatial referencing also facilitates application of the model to a variety of assessment-related problems, which are not amenable to analysis with nonspatially referenced models.

The SPARROW method can support water-quality assessment objectives in several ways. The TP and TN model applications presented here are examples of two important categories of potential applications. The first, estimation of TP exceedance proportions, represents a format for regional assessment that is repeatable at regular intervals for purposes of public information. We point out that although the model functions according to watershed boundaries in the region of interest, the results may be interpreted and displayed according to any geographic scheme, including political units. Thus analyses of state-level influences on regional water-quality conditions are possible, provided the confidence intervals associated with analyses at this scale are acceptably narrow.

The TN model application presented here is a simple example of use of the method in support of pollution control design. Other applications in this category would logically involve simulation of the effects of alternative designs. Of particular value is the capacity of the model to project the effects of contaminant sources originating in one location on water quality elsewhere. Indeed, an important result of coefficient estimation for both the TP and TN models is the clear indication that in-stream transport becomes increasingly conservative as channel size increases. An important consequence of this is that once delivered to large rivers, nutrient loads travel long distances and exert their effects far from their point of origin. The implications of this for nonpoint-source control of nitrogen are especially large because its ecological effects are thought to be greatest in coastal areas.

One area for application of SPARROW models that should not be overlooked is water-quality sampling and network design. A logical format for applications in this area is to use the model to simulate the effects of changes in sampling location and frequency on the reliability of water-quality predictions based on the model. The merits of alternative designs may then be explored in terms of accuracy improvements for various types of model output. For example, the accuracy of predictions related to contaminant sources and processes is heavily dependent on coefficient error, which can be reduced by sampling the widest possible range of the predictor variables. Alternatively, predictions of regional-aggregate water quality may be improved more through spatially representative sampling.

Acknowledgments. We would like to thank Gary Tasker, David Wolock, and Robert Hirsch of the USGS for reviewing early drafts of the paper and making helpful suggestions. We would also like to thank Graham McBride of the National Institute for Water and Atmospheric Research, Hamilton, New Zealand, and an anonymous referee for additional improvements to the manuscript.

References

Alexander, R. B., A. S. Ludtke, K. K. Fitzgerald, and T. L. Schertz, Data from selected U.S. Geological Survey water-quality monitoring networks (WQN) on CD-ROM, *U.S. Geol. Surv. Open File Rep.*, *96-337*, 79 pp., 1996.

Barrows, H. K., and G. C. Whipple, Water resources of the Kennebec River Basin, Maine, *U.S. Geol. Surv. Water Supply Pap.*, *198*, 235 pp., 1907.

Bradu, D., and Y. Mundlak, Estimation in lognormal linear models, *J. Am. Stat. Assoc.*, *65*, 198–211, 1970.

Branson, F. A., and J. R. Owen, Plant cover, runoff, and sediment yield relationships on Mancos shales in western Colorado, *Water Resour. Res.*, *6*, 783–790, 1970.

Clark, I., *Practical Geostatistics*, Appl. Sci., London, 1979.

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

Cohn, T. A., D. L. Caulder, E. J. Gilroy, L. D. Zynjuk, and R. M. Summers, The validity of a simple statistical model for estimating fluvial constituent loads: An empirical study involving nutrient loads entering Chesapeake Bay, *Water Resour. Res.*, *28*, 2353–2363, 1992.

Cressie, N., and J. Majure, Spatio-temporal statistical modeling of livestock wastes in streams, *J. Agric. Biol. Environ. Stat.*, *2*, 24–47, 1997.

DeWald, T., R. Horn, R. Greenspun, P. Taylor, L. Manning, and A. Montalbano, STORET Reach Retrieval Documentation, U.S. Environ. Prot. Agency, Washington, D. C., 1985.

Dole, R. B., The quality of surface waters in the United States, Part I—Analyses of waters east of the one hundredth meridian, *U.S. Geol. Surv. Water Supply Pap.*, *236*, 123 pp., 1909.

Dole, R. B., and F. F. Wesbrook, The quality of surface waters in Minnesota, *U.S. Geol. Surv. Water Supply Pap.*, *193*, 171 pp., 1907.

Driver, N. E., and G. D. Tasker, Techniques for estimation of storm-runoff loads, volumes, and selected constituent concentrations in urban watersheds in the United States, *U.S. Geol. Surv. Water Supply Pap.*, *2363*, 44 pp., 1990.

Dunne, T., and L. B. Leopold, *Water in Environmental Planning*, 818 pp., W. H. Freeman, New York, 1978.

Efron, B., The jackknife, the bootstrap and other resampling plans, in *CBMS-NSF Regional Conference Series in Applied Mathematics*, Soc. for Ind. and Appl. Math., Philadelphia, Pa., 1982.

Ficke, J. F., and R. O. Hawkinson, The national stream quality accounting network (NASQAN)—Some questions and answers, *U.S. Geol. Surv. Circ.*, *719*, 23 pp., 1975.

Flaxman, E. M., Predicting sediment yield in western United States, *J. Hydraul. Div.*, *Am. Soc. Civil Eng.*, *98*, 2073–2085, 1972.

Gianessi, L. P., and H. M. Peskin, *An Overview of the RFF Environmental Data Inventory: Methods, Sources and Preliminary Results*, Vol. 1, 111 pp., Resour. for the Future, Washington, D. C., 1984.

Gilroy, E. J., R. M. Hirsch, and T. A. Cohn, Mean square error of regression-based constituent transport estimates, *Water Resour. Res.*, *26*, 2069–2077, 1990.

Hainly, R. A., and J. M. Kahn, Factors affecting herbicide yields in the Chesapeake Bay watershed, June 1994, *Water Resour. Bull.*, *32*, 965–984, 1996.

Hindall, S. M., Measurement and prediction of sediment yields in Wisconsin streams, *U.S. Geol. Surv. Water-Resour. Invest. Rep.*, *54-75*, 27 pp., 1976.

Hirsch, R. M., W. M. Alley, and W. G. Wilber, Concepts for a national water-quality assessment program, *U.S. Geol. Surv. Circ.*, *1021*, 1988.

Howarth, R. W., et al., Regional nitrogen budgets and riverine N and P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences, *Biogeochemistry*, *35*, 75–139, 1996.

Hughes, J. P., and D. P. Lettenmaier, Data requirements for kriging: Estimation and network design, *Water Resour. Res.*, *17*, 1641–1650, 1981.

Johnson, D. W., Nitrogen retention in forest soils, *J. Environ. Qual.*, *21*, 1–12, 1992.

Knopman, D. S., and R. A. Smith, Twenty years of the clean water act: Has U.S. water quality improved?, *Environment*, *35*, 16–20, 34–41, 1993.

Leighton, M. O., and G. B. Holister, Quality of water in the Susquehanna River drainage basin, *U.S. Geol. Surv. Water Supply Pap.*, *108*, 76 pp., 1904.

Lystrom, D. J., F. A. Rinella, D. A. Rickert, and L. Zimmermann, Regional analysis of the effects of land use on stream-water quality, methodology and application in the Susquehanna River Basin, Pennsylvania and New York, *U.S. Geol. Surv. Water Resour. Invest.*, *78-12*, 60 pp., 1978.

Messer, J. J., R. A. Linthurst, and W. S. Overton, An EPA program for monitoring ecological status and trends, *Environ. Monit. Assess.*, *17*, 67–78, 1991.

Meybeck, M., Carbon, nitrogen, and phosphorus transport by world rivers, *Am. J. Sci.*, *232*, 401–450, 1982.

Montgomery, D. C., and E. A. Peck, *Introduction to Linear Regression Analysis*, John Wiley, New York, 1982.

Mueller, D. K., P. A. Hamilton, D. R. Helsel, K. J. Hitt, and B. C. Ruddy, Nutrients in ground water and surface water of the United States—An analysis of data through 1992, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, *95-4031*, 74 pp., 1995.

Mueller, D. K., B. C. Ruddy, and W. A. Battaglin, Logistic model of nitrate in streams of the upper-midwestern United States, *J. Environ. Qual.*, *26*, 1223–1230, 1997.

National Atmospheric Deposition Program, NADP/NTN Annual data summary: Precipitation chemistry in the United States 1987, Nat. Resour. Ecol. Lab., Colo. State Univ., Fort Collins, 1988.

National Research Council, Managing troubled waters: The role of marine environmental monitoring, Committee on a systems assessment of marine environmental monitoring, Natl. Acad. Press, Washington, D. C., 1990.

National Research Council, Managing wastewater in coastal urban areas, Committee on wastewater management for coastal urban areas, Natl. Acad. Press, Washington, D. C., 1993.

National Research Council, National water quality assessment program: The challenge of national synthesis, Committee on U.S. Geological Survey water resources research, Natl. Acad. Press, Washington, D. C., 1994.

Omernick, J. M., The influence of land use on stream nutrient levels, *Ecol. Res. Ser. Rep. EPA-600/3-76-014*, U.S. Environ. Prot. Agency, Washington, D. C., 1976.

Osborne, L. L., and M. J. Wiley, Empirical relationships between land use/cover and stream water quality in an agricultural watershed, *J. Environ. Manage.*, *26*, 9–27, 1988.

Peters, N. E., Evaluation of environmental factors affecting yields of major dissolved ions of streams in the United States, *U.S. Geol. Surv. Water Supply Pap.*, *2228*, 39 pp., 1984.

Ruddy, B. C., and K. J. Hitt, Summary of selected characteristics of large reservoirs in the United States and Puerto Rico 1988, *U.S. Geol. Surv. Open File Rep.*, *90-163*, 1990.

Rychert, R., D. Sorensen, and D. Porcella, Nitrogen fixation by lichens and free-living micro-organisms, in *Nitrogen in Desert Ecosystems*, *US/IBP Syn. Ser.*, vol. 9, edited by N. E. West and J. Skujins, pp. 22–30, Van Nostrand Reinhold, New York, 1978.

SAS Institute Inc., *SAS/ETS User's Guide*, *Version 6*, 2nd ed., Cary, N. C., 1993.

Schwarz, G. E., and R. B. Alexander, State soil geographic (STATSGO) database for the conterminous United States, *U.S. Geol. Surv. Open File Rep.*, *95-449*, 1995.

Seaber, P. R., F. P. Kapinos, and G. L. Knapp, Hydrologic units maps, *U.S. Geol. Surv. Water Supply Pap.*, *2294*, 63 pp., 1987.

Seitzinger, S. P., Denitrification in freshwater and coastal marine ecosystems: Ecological and geochemical significance, *Limnol. Oceanogr.*, *33*, 702–724, 1988.

Smith, R. A., R. B. Alexander, and K. J. Lanfear, Stream water quality in the conterminous United States—Status and trends of selected indicators during the 1980's, *U.S. Geol. Surv. Water Supply Pap.*, *2400*, 590 pp., 1993a.

Smith, R. A., R. B. Alexander, G. D. Tasker, C. V. Price, K. W. Robinson, and D. A. White, Statistical modeling of water quality in regional watersheds, paper presented at Watershed '93, A national conference on watershed management, Alexandria, Va., March 21–24, 1993b.

Steele, T. D., Dissolved loads of rivers and surface water quantity/quality relationships, Proceedings of a symposium held during the XVIII General Assembly of the International Union of Geodesy and Geophysics, Hamburg, West Germany, Aug. 15–27, 1983, *IAHS Publ.*, *141*, 131–143, 1983.

Steele, T. D., and M. E. Jennings, Regional analysis of streamflow chemical quality in Texas, *U.S. Geol. Surv. Water Resour. Invest. Res.*, *8*, 460–477, 1972.

Stensland, G. J., D. M. Whelpdale, and G. Oehlert, Precipitation chemistry, in *Acid Deposition: Long-Term Trends*, pp. 128–199, Natl. Res. Counc., Natl. Acad. Press, Washington, D. C., 1986.

U.S. Bureau of Census, Census of population and housing 1980: Master area reference file (MARF) 2, machine-readable data file, Washington, D. C., 1983.

U.S. Bureau of Census, 1987 Census of agriculture, Final county file, Washington, D. C., 1989.

U.S. Environmental Protection Agency, Quality criteria for water, Washington, D. C., 1976.

U.S. Environmental Protection Agency, County-level fertilizer sales data, *Rep. PM-221*, Off. of Policy, Plann., and Eval., Washington, D. C., 1990a.

U.S. Environmental Protection Agency, Environmental investments—The cost of a clean environment, *U.S. EPA-230/11-90-083*, Washington, D. C., 1990b.

U.S. Environmental Protection Agency, National water quality inventory: 1992 report to Congress, *EPA 841-R-94-001*, Washington, D. C., 1994.

U.S. Geological Survey, National water summary 1985—Hydrologic events and surface-water resources, *U.S. Geol. Surv. Water Supply Pap.*, *2300*, 1986.

U.S. Soil Conservation Service, Summary report 1987—National resources inventory, *Stat. Bull.*, *790*, U.S. Soil Conserv. Serv., Washington, D. C., 1989.

U.S. Soil Conservation Service, Agricultural waste management field handbook, in *National Engineering Handbook*, chap. 4, Washington, D. C., 1992.

U.S. Soil Conservation Service, State soil geographic (STATSGO) database, *Publ. 1492*, Natl. Soil Surv. Cent., Washington, D. C., 1994.

Van Winkle, W., and F. M. Eaton, The quality of the surface waters of California, *U.S. Geol. Surv. Water Supply Pap.*, *237*, 142 pp., 1910.

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

R. B. Alexander, G. E. Schwarz, and R. A. Smith, U.S. Geological Survey, MS 413, Reston, VA 20192. (e-mail: rsmith1@usgs.gov)

(Received December 31, 1996; revised July 16, 1997; accepted July 30, 1997.)