



Prepared in cooperation with Resources of the Future and Pennsylvania State University

# **An Initial SPARROW Model of Land Use and In-Stream Controls on Total Organic Carbon in Streams of the Conterminous United States**

By Jhih-Shyang Shih, Richard B. Alexander, Richard A. Smith, Elizabeth W. Boyer, Gregory E. Schwarz, and Susie Chung

Open-File Report 2010–1276

**U.S. Department of the Interior  
U.S. Geological Survey**

**U.S. Department of the Interior**  
KEN SALAZAR, Secretary

**U.S. Geological Survey**  
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia: 2010

For product and ordering information:  
World Wide Web: <http://www.usgs.gov/pubprod>  
Telephone: 1-888-ASK-USGS

For more information on the USGS—the Federal source for science about the Earth,  
its natural and living resources, natural hazards, and the environment:  
World Wide Web: <http://www.usgs.gov>  
Telephone: 1-888-ASK-USGS

Suggested citation:  
Shih, Jih-Shyang, Alexander, R.B., Smith, R.A., Boyer, E.W., Schwarz, G.E., and Chung, Susie, 2010,  
An initial SPARROW model of land use and in-stream controls on total organic carbon in streams of the  
conterminous United States: U.S. Geological Survey Open-File Report 2010-1276, 22 p., available at  
<http://pubs.usgs.gov/of/2010/1276>.

Any use of trade, product, or firm names is for descriptive purposes only and does not imply  
endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual  
copyright owners to reproduce any copyrighted material contained within this report.

## Contents

Conversion Factors.....	iii
Abstract .....	1
Acknowledgments .....	2
Introduction.....	2
SPARROW Model .....	3
Data Development.....	7
SPARROW Model Results .....	8
Assessment of Model Accuracy.....	12
Model Simulations .....	14
Summary and Conclusions.....	19
References Cited.....	20

## Figures

1. Model accuracy plots for total organic carbon loadings from SPARROW model .....	10
2. SPARROW estimates of in-stream, net removal rate for total organic carbon compared with rates for SPARROW nutrient models.....	11
3. Diagnostics for SPARROW model.....	12
4. Studentized residuals for 1,125 sites in SPARROW model .....	13
5. SPARROW predictions of the incremental total organic carbon yield (in kilograms per hectare per year) in approximately 62,000 reach catchments in the conterminous United States. ....	15
6. SPARROW predictions of source share contributions to total organic carbon yield in approximately 62,000 reach catchments in the conterminous United States. ....	16
7. Seven major regional drainage basins of the conterminous United States .....	17

## Tables

1. Organic carbon land-use sources evaluated in SPARROW. ....	7
2. SPARROW model statistics for total organic carbon. ....	11
3. Comparison of SPARROW and literature estimates of total organic carbon yields for major land types in the United States.....	14
4. SPARROW estimates of total organic carbon yields for major land types in the United States. ....	15
5. Total organic carbon yields and source shares delivered to coastal areas from major regional drainages in the conterminous United States. ....	18

## Conversion Factors

Multiply	By	To obtain
Flow rate		
acre-foot per day (acre-ft/d)	0.01427	cubic meter per second (m <sup>3</sup> /s)
acre-foot per year (acre-ft/yr)	1,233	cubic meter per year (m <sup>3</sup> /yr)
acre-foot per year (acre-ft/yr)	0.001233	cubic hectometer per year (hm <sup>3</sup> /yr)
foot per second (ft/s)	0.3048	meter per second (m/s)
foot per minute (ft/min)	0.3048	meter per minute (m/min)
foot per hour (ft/hr)	0.3048	meter per hour (m/hr)
foot per day (ft/d)	0.3048	meter per day (m/d)
foot per year (ft/yr)	0.3048	meter per year (m/yr)
cubic foot per second (ft <sup>3</sup> /s)	0.02832	cubic meter per second (m <sup>3</sup> /s)
cubic foot per second per square mile [(ft <sup>3</sup> /s)/mi <sup>2</sup> ]	0.01093	cubic meter per second per square kilometer [(m <sup>3</sup> /s)/km <sup>2</sup> ]
cubic foot per day (ft <sup>3</sup> /d)	0.02832	cubic meter per day (m <sup>3</sup> /d)
gallon per minute (gal/min)	0.06309	liter per second (L/s)
gallon per day (gal/d)	0.003785	cubic meter per day (m <sup>3</sup> /d)
gallon per day per square mile [(gal/d)/mi <sup>2</sup> ]	0.001461	cubic meter per day per square kilometer [(m <sup>3</sup> /d)/km <sup>2</sup> ]
million gallons per day (Mgal/d)	0.04381	cubic meter per second (m <sup>3</sup> /s)
million gallons per day per square mile [(Mgal/d)/mi <sup>2</sup> ]	1,461	cubic meter per day per square kilometer [(m <sup>3</sup> /d)/km <sup>2</sup> ]
inch per hour (in/h)	0.0254	meter per hour (m/h)
inch per year (in/yr)	25.4	millimeter per year (mm/yr)
mile per hour (mi/h)	1.609	kilometer per hour (km/h)
Mass		
ounce, avoirdupois (oz)	28.35	gram (g)
pound, avoirdupois (lb)	0.4536	kilogram (kg)
ton, short (2,000 lb)	0.9072	megagram (Mg)
ton, long (2,240 lb)	1.016	megagram (Mg)
ton per day (ton/d)	0.9072	metric ton per day
ton per day (ton/d)	0.9072	megagram per day (Mg/d)
ton per day per square mile [(ton/d)/mi <sup>2</sup> ]	0.3503	megagram per day per square kilometer [(Mg/d)/km <sup>2</sup> ]
ton per year (ton/yr)	0.9072	megagram per year (Mg/yr)
ton per year (ton/yr)	0.9072	metric ton per year

---

Application rate		
pounds per acre per year [(lb/acre)/yr]	1.121	kilograms per hectare per year [(kg/ha)/yr]

---

Concentrations of chemical constituents in water are given either in milligrams per liter (mg/L) or micrograms per liter ( $\mu\text{g/L}$ ).



# An Initial SPARROW Model of Land Use and In-Stream Controls on Total Organic Carbon in Streams of the Conterminous United States

By Jhih-Shyang Shih,<sup>1</sup> Richard B. Alexander,<sup>2</sup> Richard A. Smith,<sup>2</sup> Elizabeth W. Boyer,<sup>3</sup> Gregory E. Schwarz,<sup>2</sup> and Susie Chung<sup>1</sup>

## Abstract

Watersheds play many important roles in the carbon cycle: (1) they are sites for both terrestrial and aquatic carbon dioxide (CO<sub>2</sub>) removal through photosynthesis; (2) they transport living and decomposing organic carbon in streams and groundwater; and (3) they store organic carbon for widely varying lengths of time as a function of many biogeochemical factors. Using the U.S. Geological Survey (USGS) Spatially Referenced Regression on Watershed Attributes (SPARROW) model, along with long-term monitoring data on total organic carbon (TOC), this research quantitatively estimates the sources, transport, and fate of the long-term mean annual load of TOC in streams of the conterminous United States. The model simulations use surrogate measures of the major terrestrial and aquatic sources of organic carbon to estimate the long-term mean annual load of TOC in streams.

The estimated carbon sources in the model are associated with four land uses (urban, cultivated, forest, and wetlands) and autochthonous fixation of carbon (stream photosynthesis). Stream photosynthesis is determined by reach-level application of an empirical model of stream chlorophyll based on total phosphorus concentration, and a mechanistic model of photosynthetic rate based on chlorophyll, average daily solar irradiance, water column light attenuation, and reach dimensions. It was found that the estimate of in-stream photosynthesis is a major contributor to the mean annual TOC load per unit of drainage area (that is, yield) in large streams, with a median share of about 60 percent of the total mean annual carbon load in streams with mean flows above 500 cubic feet per second. The interquartile range of the model predictions of TOC from in-stream photosynthesis is from 0.1 to 0.4 grams (g) carbon (C) per square meter (m<sup>2</sup>) per day (day<sup>-1</sup>) for the approximately 62,000 stream reaches in the continental United States, which compares favorably with the reported literature range for net carbon fixation by phytoplankton in lakes and streams. The largest contributors per unit of drainage area to the mean annual stream TOC load among the terrestrial sources are, in descending order: wetlands, urban lands, mixed forests, agricultural lands, evergreen forests, and deciduous forests. It was found that the SPARROW model estimates of TOC contributions to streams associated with these land uses are also consistent with literature estimates. SPARROW model calibration

---

<sup>1</sup>Resources for the Future, Washington, D.C.

<sup>2</sup>U.S. Geological Survey, Reston, Va.

<sup>3</sup>Pennsylvania State University, Pa.

results are used to simulate the delivery of TOC loads to the coastal areas of seven major regional drainages. It was found that stream photosynthesis is the largest source of the TOC yields (about 50 percent) delivered to the coastal waters in two of the seven regional drainages (the Pacific Northwest and Mississippi-Atchafalaya-Red River basins ), whereas terrestrial sources are dominant (greater than 60 percent) in all other regions (North Atlantic, South Atlantic-Gulf, California, Texas-Gulf, and Great Lakes).

## Acknowledgments

We thank the Resources for the Future for providing software support, and John Brakebill for providing a bridge file, which includes water-quality monitoring station and stream-reach identification numbers, to link water-quality and reach attribute databases. We would also like to thank Molly K. Macauley (Resources for the Future), and Donna Myers, Jamie Shanley, and David Wolock of the USGS, for providing valuable comments on the manuscript.

## Introduction

Carbon cycling describes the complex re-circulatory processes by which carbon is transformed from inorganic to organic forms then back again. Watersheds play many important roles in the carbon cycle: (1) they are sites for both terrestrial and aquatic carbon dioxide (CO<sub>2</sub>) removal through photosynthesis; (2) they transport living and decomposing organic carbon in streams and groundwater; and (3) they store organic carbon for widely varying lengths of time as a function of many hydrological and biogeochemical factors. Here, an initial model framework (SPATIally Referenced Regression on Watershed attributes; SPARROW; Schwarz and others, 2006) is developed for quantifying flows of total organic carbon (TOC) from landscapes to river networks throughout the conterminous United States, and the associated delivery to estuarine and coastal waters. The model structure is developed based on a mass balance budget approach (Schwarz and others, 2006) and used to explore the sources and fate of TOC in watersheds. This tool is useful for quantifying net loadings to surface waters (on long-term average timescales) and the dominant terrestrial and aquatic factors controlling them.

In surface waters, TOC is a critical water quality characteristic, given that it drives the energy balance and food chains in aquatic ecosystems, is important in the mobilization and transport of contaminants along flow paths, and is associated with the formation of known carcinogens in drinking water supplies. The importance of organic carbon on water quality has long been recognized, but challenges remain in: (1) quantifying fluxes of organic carbon in surface waters at regional scales; (2) partitioning how much of the organic carbon that is stored in lakes, rivers and streams comes from allochthonous sources (produced in the terrestrial landscape) versus autochthonous sources (produced in-stream by primary production) ; and (3) understanding the importance of surface water stores of carbon to regional carbon balances. The initial modeling framework presented here for TOC addresses all of these areas quantitatively. Future work will refine these initial estimates, will quantify delivery of other forms of carbon to the coastal zone, and will explore how variability in climate and land use would affect lateral fluxes of carbon in surface waters. Results from the SPARROW (SPATIally Referenced Regression on Watershed attributes) TOC models are useful to advance understanding of the role of fresh waters in contributing to regional carbon budgets, given that the recent literature suggests that freshwater ecosystems may play a more important role than previously recognized (Wolf-Gladrow and others, 1999; Allen, 2007). Further, recent studies propose that freshwater

ecosystems process a large amount of terrestrially derived organic matter that may alter the balance between carbon sequestration and net CO<sub>2</sub> release (Cole and others, 2007).

It is well documented that flow paths and residence times of water, including the timing and magnitude of precipitation events and associated runoff, are among the primary controls on the delivery of organic matter to streams (Boyer and others, 2000). Watershed studies highlight the importance of the coupled nature of the hydrological and biogeochemical processes, with examples from all major land-use types, including wetlands (Hinton and others, 1998), forests (McGlynn and McDonnell, 2003), agricultural settings (Dalzell and others, 2007; Morel and others, 2009), and urban lands (Hook and Yeakley, 2005). Applications at multiple spatial and temporal scales have shown that information on land use (as a proxy for carbon sources), climate (as related to runoff processes), temperature (affecting rates of biogeochemical processes), and watershed characteristics (mediating fate and transport) is important for understanding concentrations and loadings of organic matter in streams and rivers (Arvola, and others, 2004, Mulholland, 2003; Sebestyen and others, 2009). The SPARROW model provides a modeling framework to integrate such information at regional scales (Smith and others, 1997), and is well suited to represent the importance of runoff and the coupled hydrological and biogeochemical processes that affect TOC fluxes (Alexander and others, 2009).

SPARROW has been used successfully to predict surface-water quality and to understand sources of solutes in streamflow in large watersheds under long-term conditions (Smith and others, 1997; Alexander and others, 2008). SPARROW plays a central role in the U.S. Geological Survey (USGS) National Water-Quality Assessment (NAWQA) Program in the conterminous United States, integrating complex information to advance understanding of the effects of nutrient sources and to transport processes over large spatial scales in watersheds (Smith and others, 1997). Previous model applications have considered various water-quality indicators, including phosphorus, nitrogen, organic carbon, suspended sediments, pesticides, and fecal coliform bacteria. Federal and State stakeholders have used SPARROW to characterize water quality in unmonitored watersheds, to track the quantities of contaminants entering estuaries, and to establish the relative importance of different pollution sources.

In this report, the SPARROW modeling framework is adapted to quantitatively link TOC in surface waters to major terrestrial and aquatic sources of organic carbon. First, a detailed description of the SPARROW model specification for TOC is presented. This is followed by a description of the spatial and temporal data that are required for calibrating the TOC model. An initial set of model parameters and diagnostics are then presented for the estimated model. The results of model simulations based on the use of this initial modeling framework are then used to generate hypotheses about TOC sources and to provide a first approximation of TOC mass loads to surface waters throughout the conterminous United States. Finally, concluding remarks and a brief discussion of future work are presented.

## **SPARROW Model**

An initial total organic carbon (TOC) model was developed, following the conventional SPARROW water-quality modeling construct that has been previously applied to other nutrients (see Schwarz and others, 2006). The model uses a statistical method to explain stream-water quality in relation to upstream sources and watershed properties, such as soil characteristics, precipitation, and land cover, which influence the transport of constituents to streams and their delivery to receiving water bodies. Though statistical in nature, the SPARROW modeling approach uses mechanistic formulations (for example, stream network flow paths and first-order

loss functions), imposes mass-balance constraints, and provides a formal-parameter estimation structure to statistically estimate sources and fate of organic carbon in terrestrial and aquatic ecosystems. Model predictions include estimates (plus uncertainty measures) of concentrations and loadings of nutrients in individual stream reaches, characterizing the delivery of TOC from source areas to streams and their transport and fate as they move downstream within the stream network.

The SPARROW model utilizes a nonlinear regression in which the long-term, mean annual stream load of TOC is related to upstream land-use sources, land-surface characteristics, and in-stream and reservoir transport properties. Spatial referencing is accomplished by linking land-use sources and land-surface characteristics, and then loading information to a spatially detailed river-reach dataset that serves as a network for relating upstream and downstream loads. TOC inputs to each river reach include loading from individual land-use sources within the watershed that drain to the reach, as well as loading from the upstream reach. Land-surface characteristics that affect the delivery of TOC to the reach are included by linking the relative amount of the specific characteristic in the contributing drainage area to that reach. All of the response and explanatory variables are spatially defined by georeferenced polygons that are related to the stream network, which defines connectivity within watersheds and allows predictions to be presented in a spatial context.

Once the SPARROW model is calibrated and evaluated for accuracy, it can be used to predict the long-term mean annual flux in streams across broad spatial scales ranging from small watersheds to large river drainages. Simulation results can be used to identify sources of constituents that affect water quality over large spatial scales or to estimate the origin and fate of constituents in streams and receiving waters. The model estimates can be illustrated through detailed maps that provide information about constituent loading at multiple scales for specific watersheds or geographic areas. Further details of the methodology are presented below, and a complete description can be found in Schwarz and others (2006).

The SPARROW statistical model includes three types of parameters: land-use source, land-to-water delivery, and in-stream loss parameters. The basic form of the statistical model is:

$$F_i^* = \left\{ \left[ \sum_{j \in J(i)} F_j' \right] A(Z_i^S, Z_i^R; \theta_S, \theta_R) + \left[ \sum_{n=1}^{N_s} S_{n,i} \alpha_n D_n(Z_i^D; \theta_D) \right] A'(Z_i^S, Z_i^R; \theta_S, \theta_R) \right\} \varepsilon_i. \quad (1)$$

The first summation term represents the amount of flux that leaves upstream reaches and is delivered downstream to reach  $i$ , where  $F_j'$  equals measured flux,  $F_j^M$ , if upstream reach  $j$  is monitored or, if it is not, is given by the model-estimated flux  $F_j^*$ .  $A(\cdot)$  is the stream delivery function representing loss processes acting on flux as it travels along the reach pathway. This function defines the fraction of flux entering reach  $i$  at the upstream node that is delivered to the reach's downstream node. The factor is a function of measured stream and reservoir characteristics, denoted by the vectors  $Z^S$  and  $Z^R$ , with corresponding coefficient vectors  $\theta_S$  and  $\theta_R$ . If reach  $i$  is a stream, then only the  $Z^S$  and  $\theta_S$  terms determine the value of  $A(\cdot)$ ; conversely, if reach  $i$  is a reservoir, then the terms that determine  $A(\cdot)$  consist of  $Z^R$  and  $\theta_R$ .

The second summation term represents the amount of flux introduced to the stream network at reach  $i$ . This term is composed of the flux originating in specific sources, indexed by  $n=1, \dots, N_s$ . Associated with each source is a source variable, denoted as  $S_n$ , and its associated

source-specific coefficient,  $\alpha_n$ . This coefficient retains the units that convert the source-variable units to flux units. The function  $D_n(\cdot)$  represents the land-to-water delivery factor. For sources associated with the landscape, this function, along with the source-specific coefficient, represents the rate at which the source variable is converted to organic carbon mass that is delivered to streams. The land-to-water delivery factor is a source-specific function of a vector of delivery variables, denoted by  $Z_i^D$ , and an associated vector of coefficients  $\theta_D$ .

The last term in the equation, the function  $A(\cdot)$ , represents the fraction of flux originating in and delivered to reach  $i$  that is transported to the reach's downstream node and is similar in form to the stream-delivery factor defined in the first summation term of the equation. If reach  $i$  is classified as a stream (as opposed to a reservoir reach), the organic carbon introduced to the reach from its incremental drainage area receives the square root of the reach's full in-stream delivery. This assumption is consistent with the notion that contaminants are introduced to the reach network at the midpoint of reach  $i$  and thus are subjected to only half the reach's time of travel. Alternatively, for reaches classified as reservoirs, we assume that the nutrient mass receives the full attenuation defined for the reach. The multiplicative error term in equation (1),  $\varepsilon_i$ , is applicable in cases where reach  $i$  is a monitored reach. The error is assumed to be independent and identically distributed across independent subbasins in the intervening drainage between stream-monitoring sites.

It is recognized that organic carbon can be both produced and lost via autochthonous processes in aquatic systems, within the stream and river network. Here, a net loss of organic carbon (that could be positive or negative upon calibration) is represented according to a first-order decay process in which the fraction of the organic carbon mass originating from the upstream node and transported along reach  $i$  to its downstream node is estimated as a continuous function of the mean water time of travel ( $T_i^S$ ; units of time); mean water depth,  $D_i$ , in reach  $i$ , such that

$$A\left(Z_i^S, Z_i^R; \theta_S, \theta_R\right) = \exp\left(-\theta_S \frac{T_i^S}{D_i}\right), \quad (2)$$

where  $\theta_S$  is an estimated mass-transfer flux-rate coefficient in units of length time<sup>-1</sup>. The rate coefficient is independent of the properties of the water column that are proportional to water volume, such as stream flow and depth. The rate can be re-expressed as a reaction rate coefficient (time<sup>-1</sup>) that is dependent on water-column depth by dividing by the mean water depth. The in-stream transport of organic carbon may reflect the effects of the long-term storage and oxidation losses of organic carbon, which may be important for relating upstream sources to downstream fluxes.

Net organic carbon loss in lakes and reservoirs is simulated according to a first-order process in which the fraction of the nutrient mass originating from the upstream reach node and transported through the reservoir segment of reach  $i$  to its downstream node is estimated as a function of the reciprocal of the areal hydraulic load ( $q_i^R$ )<sup>-1</sup> (units of time length<sup>-1</sup>) for the reservoir associated with reach  $i$  and an apparent settling velocity coefficient ( $\theta_R$ ; units of length time<sup>-1</sup>), such that

$$A(Z_i^S, Z_i^R; \theta_S, \theta_R) = \frac{1}{1 + \theta_R (q_i^R)^{-1}}. \quad (3)$$

The areal hydraulic load is estimated as the quotient of the outflow discharge to the surface area of the impoundment.

The carbon source proxies considered (see table 1) include seven land-use types, as well as photosynthesis in streams and reservoirs. The carbon source parameters  $\beta_n$  are included to determine the significance of different land-use types and stream photosynthesis in explaining the variation of TOC loads among reaches. Seven land-use sources were evaluated that were in preliminary versions of the model: agricultural land, deciduous forest, evergreen forest mixed forest urban land, range grasslands, and wetlands. Some of these sources are aggregations of several more specific land-use types, defined in the table. Data on land use are obtained from the 1992 National Land Cover Data Set (Vogelmann and others, 2001).<sup>4</sup> The quantity of organic carbon production via photosynthesis in reservoirs is assumed to be proportional to the surface area of these waters, which was previously reported for the river network (Nolan and others, 2002).

In-stream photosynthesis was calculated for each reach on the basis of the following input data: total phosphorus concentration, depth, width, length, and average daily solar irradiance. Depth and width were estimated as a function of stream flow using the relations of Leopold and Maddock (1953). An average value for non-chlorophyll light attenuation at USGS NASQAN stations (Soballe and Kimmel, 1987) was assumed to apply to all reaches. Total phosphorus concentrations were taken from a previously calibrated SPARROW model (Alexander and others, 2008), and were used to estimate algal chlorophyll concentrations on the basis of an empirically derived equation (Van Nieuwenhuysse and Jones, 1996). The chlorophyll estimates together with the value for non-chlorophyll light attenuation were used to calculate a total light attenuation coefficient for each reach. Net in-stream photosynthetic rate ( $\text{g C m}^{-2}\text{d}^{-1}$ ) was calculated from chlorophyll, irradiance, and light attenuation using the model of Smith (1980). The model accounts for the opposing processes of light trapping and self-shading by chlorophyll in the water column. The in-stream TOC source in the SPARROW model was calculated as the product of photosynthetic rate and reach surface area. In contrast to terrestrial sources, the in-stream TOC source term is not multiplied by the land-to-water delivery fraction; in calibration, the estimated in-stream source coefficient is expected to be approximately 1.0 if the models for stream chlorophyll (Van Nieuwenhuysse and Jones, 1996) and photosynthetic rate (Smith, 1980) are approximately accurate.

---

<sup>4</sup>Many researchers have studied the accuracy of the National Land Cover Data (NLCD) set. For example, Smith and others (2003) evaluated the effects of patch size and land-cover heterogeneity on classification accuracy. Logistic regression models were used to quantify the relations between classification accuracy and these landscape variables for each land-cover class at both the Anderson Levels I and II classification schemes employed in the NLCD. An assessment of 1992 NLCD accuracy can also be found in Wickham and others (2004).

**Table 1.** Organic carbon land-use sources<sup>1</sup> evaluated in SPARROW.

<b>Carbon source</b>	<b>Definition</b>
Agricultural land	Area of row crops; small grains; fallow; pasture; and orchards-vineyards-other.
Deciduous forest	Area of deciduous forest.
Evergreen forest	Area of evergreen forest.
Mixed forest	Area of mixed forest.
Range and grass lands	Area of shrub lands and herbaceous grass lands.
Urban land	Area of low-intensity residential, high-intensity residential, and commercial-industrial-transportation land; urban-residential grasses.
Wetlands	Area of woody wetlands; emergent herbaceous wetlands.
Photosynthesis in streams	Based on total phosphorus concentration, solar irradiance, and channel dimensions.
Photosynthesis in reservoirs	Estimated surface area. <sup>2</sup>

<sup>1</sup>Vogelmann and others, 2001.

<sup>2</sup>Nolan and others, 2002.

The land-to-water delivery parameters determine the significance of different climatic and landscape properties that increase or decrease the transport of TOC from the land surface to the stream reach. For example, relatively large percentages of impermeable surface area might be expected to increase delivery from the land surface to stream reaches. Land-surface characteristics that were considered include permeability, precipitation, percentage of area that is tile drained, stream density, type of conservation tillage, and slope.

Model parameters in the national model were statistically estimated by applying a nonlinear least-squares algorithm to equation (1) above. The estimation procedure is informed by spatial correlations between total organic carbon flux estimated at monitoring stations with watershed data on sources, and landscape and surface-water properties that affect transport. The error term in the model is assumed to be multiplicative, and the estimation algorithm was applied after both sides of the equation were converted to logarithmic form.

## Data Development

The organic carbon data used to calibrate the SPARROW model were retrieved from the U.S. Environmental Protection Agency's (USEPA) STORET system and the USGS National Water Information System (NWIS), which includes USGS national network sites (NAWQA, NASQAN, and BENCHMARK) and USGS Water Science Center cooperative program sites operating in cooperation with State and local agencies. A total of 2,307 STORET sites and 2,491 NWIS sites were evaluated for possible use in modeling.

The following three types of organic carbon data were extracted for the period from October 1, 1969, through December 31, 2008: dissolved organic carbon (parameter code = 00681), suspended organic carbon (parameter code = 00689), and total organic carbon (parameter code = 00680). Missing values of total organic carbon were filled in for days and times when measurements of dissolved and suspended organic carbon were available, which were summed.

The long-term mean annual stream load, the response variable of the SPARROW model, was estimated for each monitoring site prior to SPARROW modeling based on the application of load estimation procedures (Schwarz and others, 2006). These procedures use the periodically collected measurements of organic carbon and the daily streamflow data collected at numerous USGS monitoring sites. Because water quality is rarely measured with the frequency necessary

to reliably estimate the long-term mean load, the load estimation method uses a linear regression model (that quantifies the correlation between the observed pairs of water-quality and flow measurements) together with the more complete set of daily flows for the period-of-record to obtain an estimate of the long-term mean annual load. This estimate is more accurate (i.e., lower variance) than one based exclusively on the infrequently sampled pairs of water quality and streamflow (Schwarz and others, 2006).

The following criteria were used to identify the monitoring sites with sufficient data and acceptable estimates of the long-term mean annual load for use in SPARROW modeling: (1) >15 observations of total organic carbon; (2) 3 consecutive years of daily flow record; (3) standard error of the mean annual load <50 percent; (4) ratio of the water-quality monitoring site drainage area to the streamflow gage area >0.5 and <2.0; (5) foreign (Canadian or Mexican) drainage areas <10 percent; and (6) a single monitoring station is selected for each stream reach. Using these criteria produced a total of 1,125 monitoring sites for use in estimating the SPARROW organic carbon model.

## SPARROW Model Results

The statistics for the final version of the SPARROW model of total organic carbon, estimated with 1,125 monitoring sites, are shown in table 2. The variables and estimated coefficients are presented for the three major model components: (1) organic carbon sources; (2) land-to-water delivery factors; and (3) stream or river transport.

The adjusted R-squared value of the model is 0.928 for the log-transformed mean annual load (see fig. 1a). Correcting for the intrinsic effect of drainage area on the prediction of load in SPARROW gives a yield R-squared value of 0.77 (fig. 1b); the yield R-squared value generally gives a more informative measure of the explanatory power of the model. The Root Mean Square Error (RMSE) gives an approximate estimate of the mean error of a reach-level prediction by the model; thus, the long-term mean annual TOC load can be predicted to within  $\pm 54$  percent in a reach based on the estimated model. This level of prediction error is generally similar to that observed for a prior national SPARROW model of total nitrogen (Alexander and others, 2008).

Our final model (table 2) includes seven of the nine evaluated sources (land uses and aquatic photosynthesis), five land-to-water delivery factors, and an in-stream carbon removal term; all coefficients are highly statistically significant (level of statistical significance or “p” < 0.002) except for the land slope variable. Two sources, range and grasslands and reservoir photosynthesis, were eliminated from the final model because they were estimated to be statistically indistinguishable from zero, suggesting that their net organic carbon contributions to downstream waters are zero—that is, organic carbon production and loss are approximately balanced; see discussion below on reservoir loss). The source coefficients are standardized for spatial variability in the land-to-water delivery factors and expressed relative to the mean of these factors. Therefore, all of the source coefficients (except for in-stream photosynthesis) provide a measure of the stream TOC load generated in response to a unit areal contribution (square kilometers) from the source variable, and can be compared directly to one another to obtain a sense of their relative national importance in the model. Among these sources with comparable units, it is found that wetlands make the largest mass contribution per unit area (or yield) to the stream organic carbon load, followed in declining order by urban lands, mixed forests, agricultural lands, evergreen forests, and deciduous forests. These results are generally consistent with the magnitude and relative ordering of organic carbon exports for small catchments in these land uses (Mulholland, 2003; also see discussion in the “Assessment of

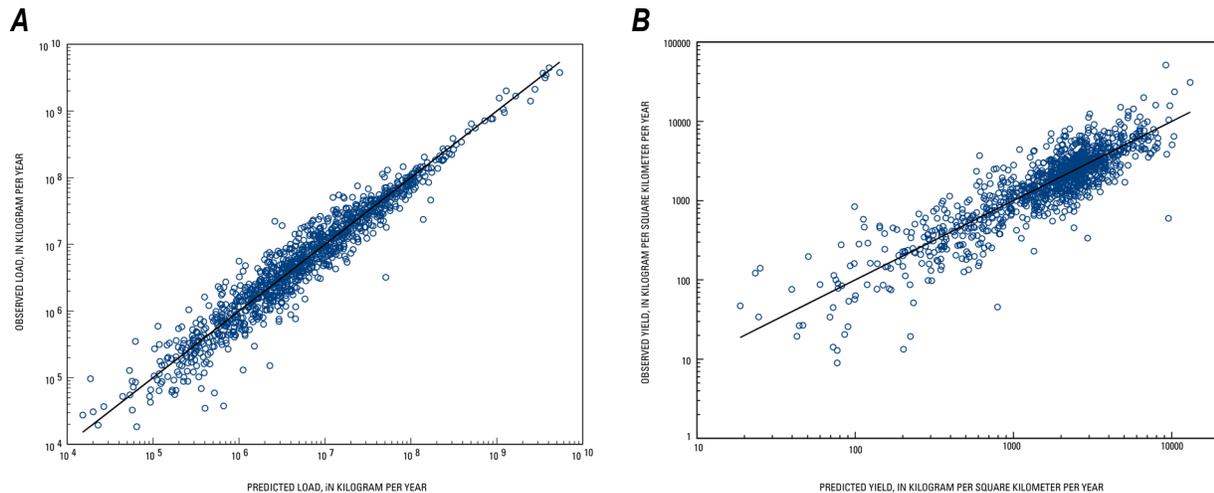
Model Accuracy” section below). The “mixed forest” TOC export coefficient is notably larger than the TOC exports for deciduous and evergreen forests by a factor of about 3, a result that could be related to the effects of the particular climatic and biogeochemical conditions of these watersheds, such as their location in more humid regions of the country and in areas where wetlands are potentially hidden under homogeneous forest canopy (that is, “cryptic” wetlands; Creed and others, 2003; see discussion below and in the “Model Simulations” section below .

It was also found that the estimates of stream photosynthesis compare favorably with carbon fixation rates for net aquatic photosynthesis reported in the literature. For example, mean annual phytoplankton photosynthesis rates are frequently reported in the range of 0.1 to 1.0 grams (g) carbon (C) per square meter ( $\text{m}^{-2}$ ) per day ( $\text{day}^{-1}$ ) for lakes (Allen, 2007) and 0.04 to 2  $\text{gC m}^{-2} \text{day}^{-1}$  for streams and rivers (0.38 to 2.1 for European rivers, Billen and others, 1995; 0.23 to 1.25 for eastern U.S. rivers, Webster and others, 1995; and 0.04 to 1.81 for western U.S. rivers, Fisher, 1995). These values compare well with the interquartile range of TOC SPARROW model predictions of from about 0.1 to 0.4  $\text{gC m}^{-2} \text{day}^{-1}$  for the approximately 62,000 reaches, based on reach-level estimates of total phosphorus concentration, solar irradiance, and channel morphology. The fact that the estimated coefficient for in-stream photosynthesis (1.10 in table 2) is close to the expected value of 1.0 suggests that the models used to estimate autochthonous carbon production (stream chlorophyll and photosynthetic rate) are generally accurate. An important feature of the model predictions of stream photosynthesis is that they vary greatly with stream size. At streamflows less than  $10 \text{ ft}^3 \text{ second}^{-1}$  ( $0.357 \text{ m}^3 \text{ second}^{-1}$ ), for example, the median photosynthetic rate is about  $0.1 \text{ g C m}^{-2} \text{ day}^{-1}$ , but at flows above 500 cubic feet per second, the median photosynthetic rate is about  $0.8 \text{ g C m}^{-2} \text{ day}^{-1}$ . The share of the annual TOC load contributed by stream photosynthesis varies in accordance with this pattern, with stream photosynthesis in the lower streamflow class contributing a median of about 2 percent of total annual carbon load, whereas stream photosynthesis in the upper streamflow class contribute a median of about 60 percent of total carbon load.

The land-to-water delivery coefficients estimated in the model (table 2) may reflect the effects of various hydrological and biogeochemical processes that are associated with these factors. It was found that land-to-water delivery of TOC is enhanced by precipitation, drainage density, and artificial drainage (the latter affects only transport from agricultural lands), which are all positively correlated with stream TOC loads. Precipitation would be expected to generally increase runoff and TOC delivery to streams, which was confirmed by the model and consistent with the widely reported (Mulholland, 2003) response of the riverine export of organic carbon in watersheds. Artificial drainage (subsurface tiles and ditches used predominantly in agriculture) is a practice that rapidly removes excess water from the land surface and shallow subsurface of soil to enhance the productivity of the soil for cultivation. One would expect that a high percentage of land area with artificial drainage would increase water delivery to the stream and, therefore, have a positive impact on the in-stream load of organic carbon, as evidenced in the estimated SPARROW model. By contrast, soil permeability and land slope were found to be negatively correlated with stream TOC loads in the model. Areas with relatively less permeable soils and low slope might be expected to deliver more TOC, on average, to streams. For example, reach drainages with relatively impermeable soils may tend to route water and organic carbon overland to streams, whereas drainages with relatively high permeability may generally have less runoff and lower TOC load, corresponding to the increased opportunities for oxidation of TOC in soils and the subsurface. Drainages with low average slopes, reflecting possible catchments with saturated soils, are more common, which can enhance subsurface and surface hydrological

connectivity to streams and delivery of TOC. The generally lower rates of anaerobic decomposition in the saturated soils of topographically flat drainages could also account for enhanced delivery of TOC to streams.

It was found that the net rates of TOC removal in streams decrease with increases in water depth (fig. 2); by contrast, the net removal rate for reservoirs was estimated to be zero and was not statistically significant. The fact that the terms for both production and loss of TOC in reservoirs were not found to be statistically significant in the model suggests that the two processes may be approximately balanced on average, based on the current model specifications. The estimated TOC mass-transfer coefficient for streams was  $0.034 \text{ m day}^{-1}$  (table 2;  $12.4 \text{ m yr}^{-1}$ ), which corresponds to the series of reaction rate coefficients (units of per day) as shown in figure 2 for a range of water depths in streams in the river network. The inverse relation between TOC removal and stream-water depth is consistent with that reported for prior SPARROW models for total nitrogen and total phosphorus (for example, Alexander and others, 2008; fig. 2), and with current understanding of the hydrologic and biogeochemical processes (particulate settling, oxidation, depth, and water-volume-to-surface-area ratios) responsible for organic carbon removal in natural waters (Hope and others, 1994; Naiman and others, 1987).

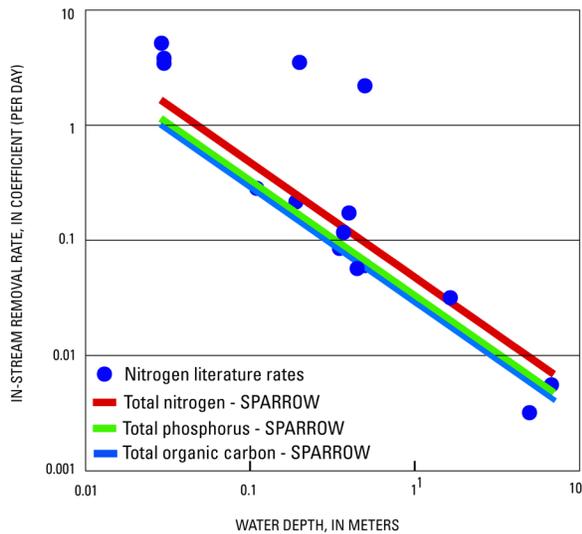


**Figure 1.** Model accuracy plots for total organic carbon loadings from SPARROW model: A, observed and predicted load (kilograms per year), and B, observed and predicted yield (kilograms per square kilometer per year).

**Table 2.** SPARROW model statistics for total organic carbon.

Parameter	Coefficient units <sup>1</sup>	Estimate	Standard error	t value	Level of statistical significance, p
Agriculture	kg km <sup>-2</sup> yr <sup>-1</sup>	1,454	167	8.65	0.000
Forest, deciduous	kg km <sup>-2</sup> yr <sup>-1</sup>	1,061	191	5.54	0.002
Forest, evergreen	kg km <sup>-2</sup> yr <sup>-1</sup>	1,378	167	8.21	0.000
Forest, mixed	kg km <sup>-2</sup> yr <sup>-1</sup>	2,568	627	4.09	0.001
Urban	kg km <sup>-2</sup> yr <sup>-1</sup>	4,777	778	6.14	0.000
Wetlands	kg km <sup>-2</sup> yr <sup>-1</sup>	25,008	2,529	9.89	0.000
In-stream photosynthesis	dimensionless	1.10	0.13	8.67	0.000
Soil permeability	log (cm hr <sup>-1</sup> )	-0.1407	0.0368	-3.82	0.000
Precipitation	cm	0.0047	0.0006	7.51	0.000
Artificial drainage	percent area	0.0116	0.0031	3.82	0.001
Drainage density	log (km <sup>-1</sup> )	0.4407	0.0545	8.08	0.000
Land slope	log (percent)	-0.0023	0.0040	-0.58	0.5620
In-stream carbon removal	m day <sup>-1</sup>	0.0338	0.0036	9.31	0.000
Log root mean square error		0.540			
Number of observations		1,125			
Adjusted R-squared		0.928			
Yield R-squared		0.77			

<sup>1</sup>kg = kilogram; km = kilometers; yr = year; cm = centimeters; hr = hour; m = meters.

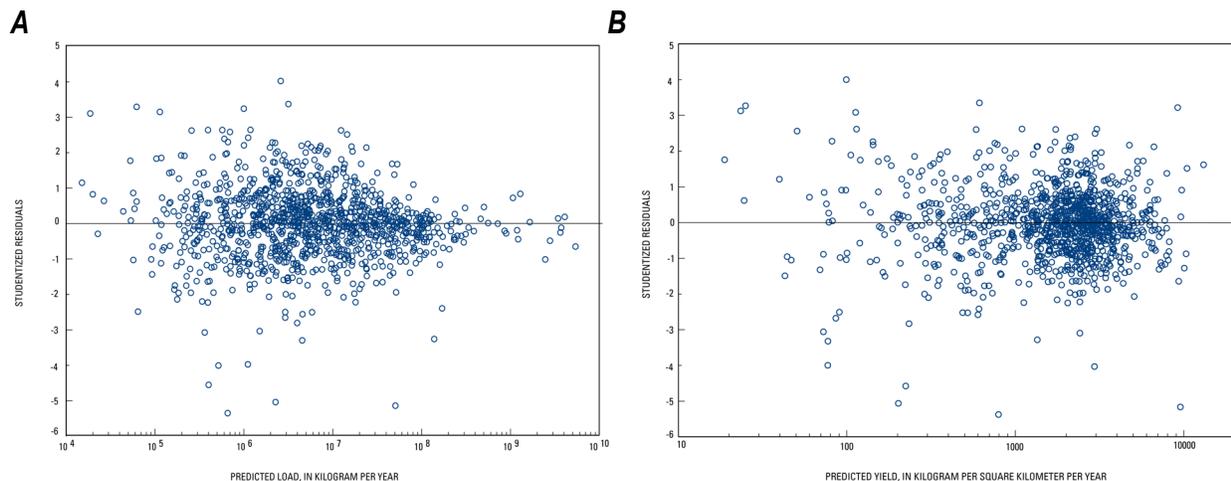


**Figure 2.** SPARROW estimates of in-stream, net removal rate for total organic carbon compared with rates for SPARROW nutrient models. [Modified from Alexander and others (2008).]

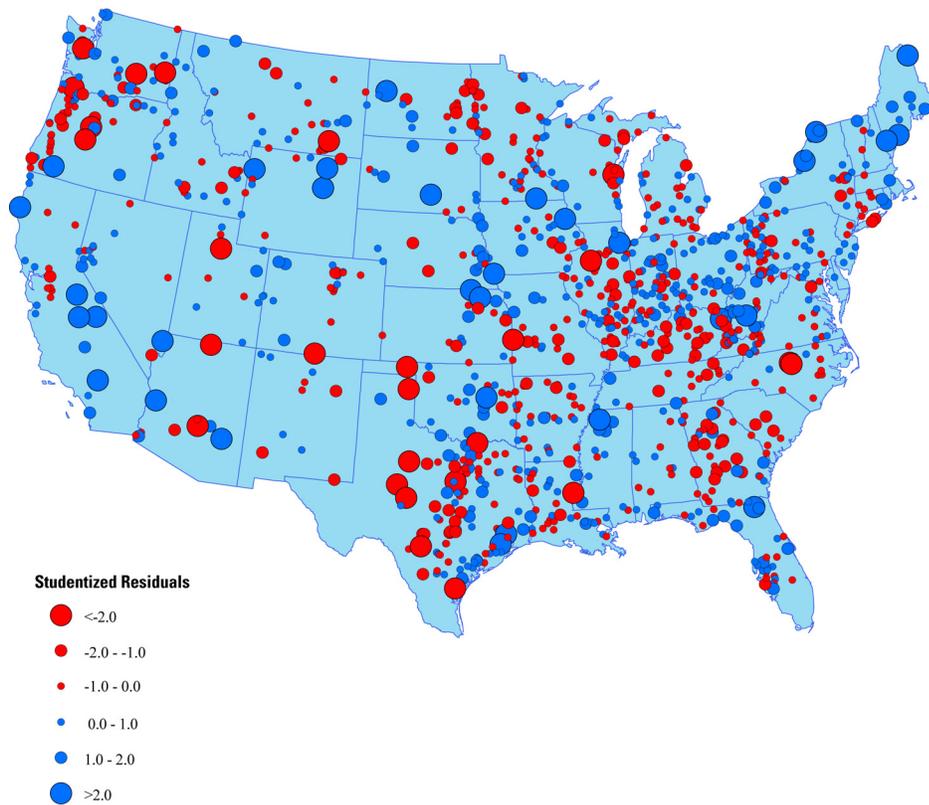
## Assessment of Model Accuracy

Plots of the model predictions and residuals of the final models are shown in figure 3 and provide information on how well the model errors satisfy the assumptions of constant variance. Each residual is expressed as a studentized value, which is computed by standardizing the observed residual in relation to the residual standard error that excludes the effect of the observed residual value. For a correctly specified model, the residuals would be expected to be distributed as a t-distribution with  $n-p-1$  degrees of freedom, where  $n$  is the number of observations and  $p$  is the number of estimated model parameters. Overall, the residuals of the models show generally constant variance for both load (fig. 3a) and yield (fig. 3b). Notable exceptions are observed for the variance of the 20 largest load values for monitoring stations located on large rivers where model accuracy is relatively high.

The final model was evaluated for evidence of regional biases (overprediction and underprediction) in the model predictions, based on visual inspections of maps of the station residuals (fig. 4). Overall, the model shows some evidence of prediction biases in selected regional watersheds, including a tendency for overprediction (studentized residual  $< -1.0$ ) at sites in areas of the Pacific Northwest, western Texas, Ohio basin, and the Southeast. Areas of underprediction are apparent in southern California, central United States, and the extreme Northeast. Some of the prediction biases may be related to temporal differences in the environmental conditions reflected by the period of record covered by the various monitoring stations. For example, a general tendency was found for the model to underpredict at stations where the TOC records ended prior to 1983, and over-predict at stations where TOC records began after 1990. The nature of the apparent bias in the SPARROW predictions may be somewhat inconsistent with what might be expected based on recent trends in dissolved organic carbon found in the surface waters of undeveloped catchments in eastern North America (Monteith and others, 2007); more work will be needed to address this issue in the next version of the model. Additional evaluations of these biases are warranted in future models, and improvements made in the models to account for these biases will likely lead to enhanced accuracy and interpretability.



**Figure 3.** Diagnostics for SPARROW model: A, studentized residuals and predicted load, and B, studentized residuals and predicted yield.



**Figure 4.** SPARROW estimates of total organic carbon yields and source share percentages for the incremental watershed associated with stream reaches in the conterminous United States.

SPARROW predictions of stream TOC yields for catchments dominated by different types of land uses compare reasonably well with TOC yields reported in the literature (see table 3). These comparisons provide partial confirmation of the accuracy of the models. Stream organic carbon yields from the literature are reported for watersheds with relatively uniform land use but include the variable influences of climate, soils, wetlands, vegetation, catchment size, and in-stream processes that can complicate efforts to make precise comparisons between model predictions and literature yields. In general, we find that watersheds with large areas of wetlands have the largest SPARROW predictions of stream organic carbon yield, whereas those for watersheds predominantly in forest and cropland are (as much as an order of magnitude) smaller, which is generally consistent with the relative differences in yields reported in the literature for these land uses. In addition, somewhat larger yields are predicted by SPARROW for evergreen forest than for deciduous forest, a result that is also generally consistent with the literature. SPARROW estimates of organic carbon from watersheds predominantly in range and grassland include low-level “background” organic carbon inputs from natural sources and are an order of magnitude lower than the yields predicted for predominantly agricultural and forested watersheds. Although the SPARROW yield predictions for watersheds predominantly in agriculture and wetlands overlap with the range of yields reported in the literature for these land types, the SPARROW predictions are somewhat larger than those in the literature. Given the high level of sensitivity of TOC to runoff and precipitation (Mulholland, 2003), differences in hydrologic conditions may explain some of the differences between SPARROW and the literature yields. However, because the literature yields for wetlands cover a range of runoff

conditions, some generally positive bias (tendency to overpredict) may exist with the model for wetlands and will need to be evaluated in future work.

**Table 3.** Comparison of SPARROW and literature estimates of total organic carbon yields for major land types in the United States.

[SPARROW estimates reflect the yields for watersheds of the specified predominant land-use type as determined from NLCD Anderson Level II classifications and associated with individual stream reaches as defined by the digital river network for the conterminous United States. Kg, kilogram; ha, hectare; yr, year]

Watershed land-cover type	Percentiles of total organic carbon yield exported from SPARROW watersheds <sup>1</sup> (kg ha <sup>-1</sup> yr <sup>-1</sup> )					Literature yields <sup>2</sup> (kg ha <sup>-1</sup> yr <sup>-1</sup> )	
	No. of watersheds	10 <sup>th</sup>	25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	90 <sup>th</sup>	Range of values
Agriculture	1,841	15.1	19.1	24.5	34.7	59.9	14.1–19.5 <sup>2</sup>
Forest	70	12.1	14.0	16.2	21.4	34.0	4–80 <sup>3</sup>
Deciduous							
Forest	248	13.6	16.1	20.3	28.6	56.0	14–500 <sup>3</sup>
Evergreen							
Range	3,203	0.1	0.3	0.8	2.3	8.2	4–13 <sup>3</sup>
Urban	143	36.3	48.3	73.3	108.6	329.3	19–146 <sup>3</sup>
Wetlands	191	160.3	276.4	476.3	801.6	2180.0	50–220 <sup>4</sup>

<sup>1</sup>The land-cover types represent the following percentages of the land area in SPARROW watersheds: agricultural land (>90 percent), forest (>95%), urban (>90%), wetlands (>95%), and range (95%).

<sup>2</sup>Dalzell and others, 2007.

<sup>3</sup>Hope and others, 1994; North America, New Zealand, and Russia (total organic carbon).

<sup>4</sup>Mulholland, 2003 (dissolved organic carbon).

## Model Simulations

Application of the model to the approximately 62,000 reach catchments gives predictions of the TOC yields and source share contributions for the seven sources described by the model. The average results are shown in table 4. The stream photosynthesis source is the largest overall source (expressed as a mean of the reach-level source share percentages), followed by (in order) wetlands, agriculture, evergreen forest, mixed forest, deciduous forest, and urban land.

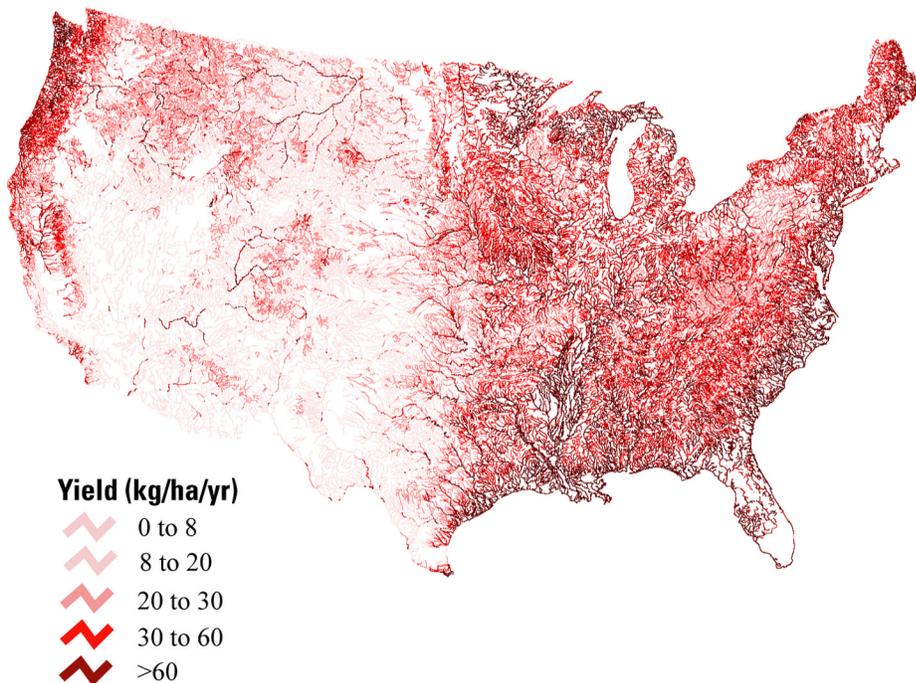
The geography of the incremental yields and source contributions are shown in figures 5 and 6, respectively. TOC yields are generally larger in the eastern United States than in the montane West, reflecting the effects of precipitation. Wetlands are an important source of organic carbon in the Southeast and northern portions of Minnesota (see fig. 6B). Stream photosynthesis is estimated to be a major contributor of organic carbon in streams among the seven terrestrial and aquatic sources considered (see fig. 6A). Stream photosynthesis is especially important (contributing about 50 percent of the organic carbon to streams) in watersheds of the Pacific Northwest and Mississippi-Atchafalaya-Red River; stream photosynthesis also is a major source (>60 percent) in many large rivers in the eastern United States. Mixed forests account for a major source of TOC in streams (>60 percent) in areas of the southern Mississippi valley, Southeast uplands, and watersheds of the Northeast (see fig. 6G); humid conditions and hydric soils are common in these regions and may partially explain the larger threefold SPARROW

estimate of the areal export for mixed forests (table 2) compared with the areal export estimated for deciduous and evergreen forests.

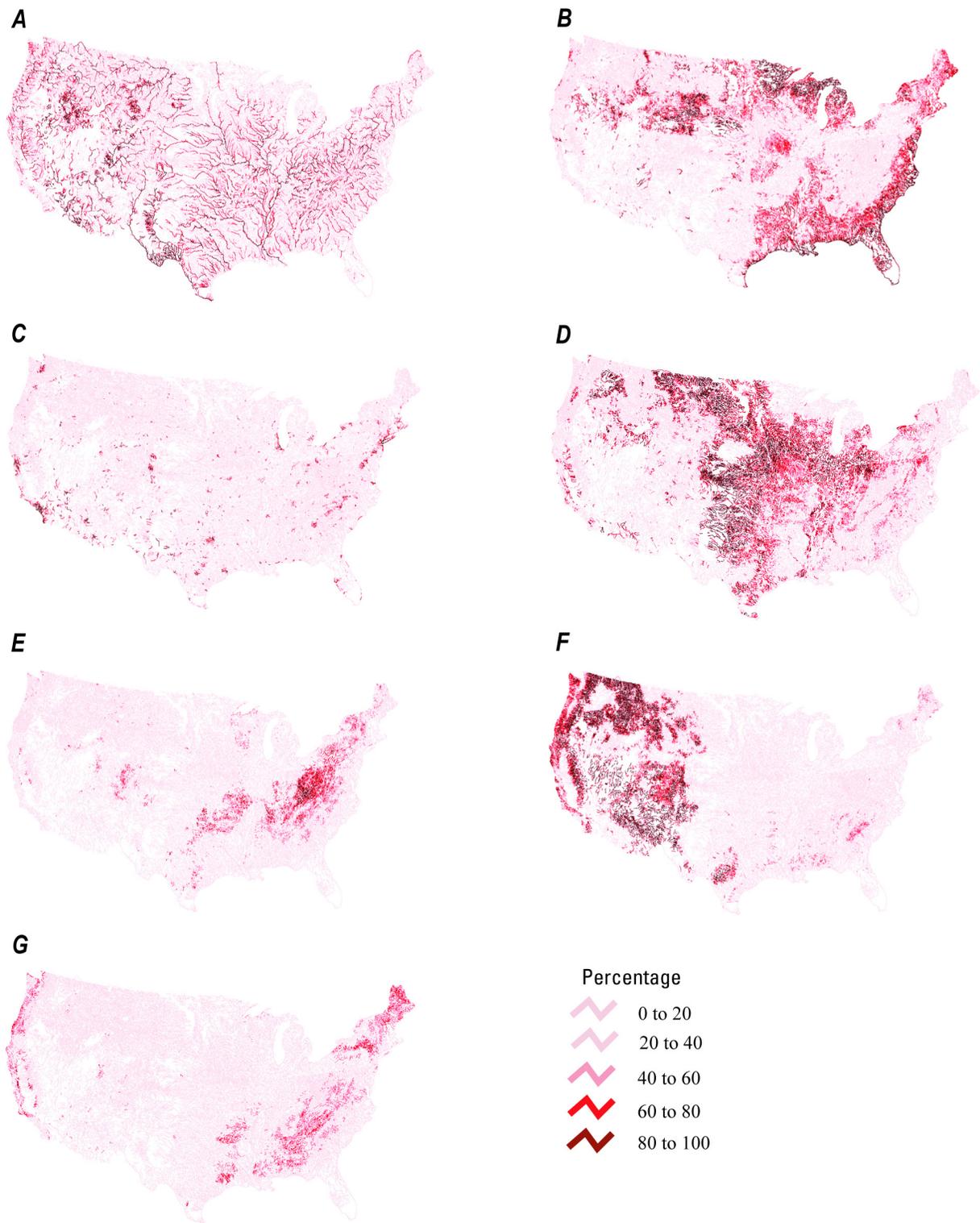
**Table 4.** SPARROW estimates of total organic carbon yields and source share percentages for the incremental watersheds associated with stream reaches in the conterminous United States.

Variable	Number of watersheds	Mean	Standard deviation
Incremental yield (kg ha <sup>-1</sup> yr <sup>-1</sup> )	61,121	72.1	495
Agriculture source share (%)	62,303	19.9	27.2
Forest, deciduous source share (%)	62,303	7.5	13.3
Forest, evergreen source share (%)	62,303	19.7	30.4
Forest, mixed source share (%)	62,303	6.4	10.8
Urban source share (%)	62,303	4.3	11.5
Wetlands source share (%)	62,303	19.8	27.1
In-stream photosynthesis source share (%)	62,303	22.4	29.2

<sup>1</sup>Kilograms per hectare per year.



**Figure 5.** SPARROW predictions of the incremental total organic carbon yield (in kilograms per hectare per year) in approximately 62,000 reach catchments in the conterminous United States.



**Figure 6.** SPARROW predictions of source share contributions to total organic carbon yield in approximately 62,000 reach catchments in the conterminous United States. *A*, In-stream photosynthesis; *B*, Wetlands; *C*, Urban; *D*, Agriculture; *E*, Deciduous forest; *F*, Evergreen forest; and *G*, Mixed forest.

The SPARROW model was also used to predict the sources and yield of TOC for seven major regional drainage basins of the conterminous United States (see fig. 7 and table 5). The regional summary in table 5 provides information about the relative importance of terrestrial and aquatic sources of organic carbon exported from the continental United States and delivered to coastal waters. The reported TOC “delivered” yields (table 5) reflect the net long-term mean delivery of TOC mass to coastal waters per unit of drainage area, and account for the mean effects of in-stream losses (for example, settling and oxidation) on TOC during transport in streams and rivers. It was found that the largest delivered yields occur in regional drainages of the eastern United States, including the South Atlantic Gulf region and the Great Lakes. The Texas-Gulf region and drainages of the western United States (Pacific Northwest and California) are among the smallest.



**Figure 7.** Seven major regional drainage basins of the conterminous United States. The total organic carbon loads issuing from these basins, shares delivered from various sources, and total delivered yield are shown in table 5.

**Table 5.** Total organic carbon yields and source shares delivered to coastal areas from major regional drainages in the conterminous United States.

Region	Drainage area (km <sup>2</sup> ) <sup>1</sup>	Delivered yield (kg ha <sup>-1</sup> yr <sup>-1</sup> )	Source share percentage							
			Agriculture	Forest			Urban	Wetland	Stream photosynthesis	All terrestrial <sup>3</sup>
				Deciduous	Evergreen	Mixed				
North Atlantic	446,500	34	7	10	5	14	10	32	22	78
SouthAtlantic- Gulf	730,000	59	5	3	5	5	3	58	21	79
Miss.-Atch.-Red <sup>4</sup>	3,248,700	32	14	4	1	2	2	22	54	46
Texas-Gulf	925,400	11	14	3	4	5	5	34	36	64
Pacific Northwest	713,700	17	6	2	31	6	3	5	48	52
California	234,200	15	8	2	29	8	10	7	37	63
Great Lakes	313,100	38	15	6	1	4	4	60	10	90

<sup>1</sup>Square kilometer.

<sup>2</sup>Kilograms per hectare per year.

<sup>3</sup>Sum of agriculture, forest, urban, and wetland contributions.

<sup>4</sup>Mississippi-Atchafalaya-Red Valley.

Stream photosynthesis is the major source of organic carbon in coastal waters in two of the seven regions, the Pacific Northwest and Mississippi-Atchafalaya-Red River basins ( about 50 percent), whereas terrestrial sources are dominant in all other regions (>63 percent in table 5). In regions where terrestrial sources are dominant, wetlands and evergreen forests constitute an appreciable portion (one-quarter to one-half) of the terrestrial sources of organic carbon that are exported to coastal waters. Wetland sources are highest in the South Atlantic-Gulf and Great Lakes regions (58 percent and 60 percent, respectively). Forest contributions to the organic carbon export to coastal waters are generally somewhat larger or similar to those originating from agricultural lands.

## Summary and Conclusions

A SPARROW model was developed to quantitatively link total organic carbon in streams and reservoirs with major terrestrial and aquatic sources of carbon. Using the USGS SPARROW model and total organic carbon (TOC) data from USGS NWIS and USEPA STORET databases, TOC yields for seven major land types in the conterminous United States were quantitatively estimated. The simulations use surrogate measures of the major terrestrial and aquatic sources of organic carbon to estimate the long-term mean annual load of TOC in streams. The estimated carbon sources in the model are associated with four land uses (urban, cultivated, forest, and wetlands) and autochthonous fixation of carbon (stream photosynthesis). Stream photosynthesis is determined by reach-level application of an empirical model of stream chlorophyll based on total phosphorus concentration, and a mechanistic model of photosynthetic rate based on chlorophyll, average daily solar irradiance, water column light attenuation, and reach dimensions. It was found that the model estimate of carbon originating from in-stream photosynthesis is a major contributor to the mean annual TOC load in large streams, with a median share of about 60 percent of the total mean annual carbon load in streams with mean flows above 500 ft<sup>3</sup> second<sup>-1</sup> (17.85 m<sup>3</sup> second<sup>-1</sup>). The largest areal contributors to stream TOC load among the terrestrial sources are, in descending order, wetlands, urban lands, mixed forests, agricultural lands, evergreen forests, and deciduous forests . The SPARROW estimates of the TOC contributions to streams from stream photosynthesis and the various land uses are generally consistent with literature estimates.

SPARROW model calibration results are used to simulate TOC yields and source shares for all streams in the conterminous United States and to simulate TOC delivery to coastal areas from seven major regional drainages. SPARROW predictions of the source share contributions to total organic carbon yield indicate that stream photosynthesis is the largest source of TOC load ( about 50 percent) in coastal waters in two of the seven regions, the Pacific Northwest and Mississippi-Atchafalaya-Red River , whereas terrestrial sources are dominant (>63 percent) in all other regions.

Model performance and accuracy were evaluated through statistical assessments of the model fit (residual plots and maps) and evaluations of the physical interpretability of the model predictions. Overall, the model shows some evidence of prediction biases in selected regional watersheds, including a tendency for overprediction at sites in areas of the Pacific Northwest, western Texas, Ohio basin, and the Southeast. Areas of underprediction are apparent in southern California, the central United States, and the extreme Northeast. Some of the prediction biases may be related to temporal differences in the environmental conditions reflected by the period of record covered by the various monitoring stations. For example, we found a general tendency for the model to underpredict at stations where the TOC records ended prior to 1983, and to

overpredict at stations where the records began after 1990. The nature of the apparent bias in the SPARROW predictions may be somewhat inconsistent with what might be expected based on recent trends in dissolved organic carbon in the surface waters of undeveloped catchments in eastern North America (Monteith and others, 2007) and more work will be needed to address this issue in the next version of the model. Additional evaluations of these biases are warranted in future models; improvements in the models to account for these biases will likely lead to enhanced model accuracy and interpretability.

Future work will further refine the SPARROW model structure along these lines, explore precursor sources of organic carbon (in addition to and as a replacement for the area-based carbon source proxies in the current model), and consider responses to climatic variability and change. Replacements for the area-based carbon sources in the present model might logically be based on estimates of terrestrial net primary production derived from remote (for example, satellite-based) sensors. Finally, further investigations into the role of reservoir and lake processes in TOC transport are also warranted based on our findings that both production and loss of TOC in reservoirs are not statistically significant in the model. This result suggests that the two processes may be approximately balanced on average in the current model; the inclusion of additional reservoir data on carbon processes (for example, sedimentation rates) or more complex photosynthetic production terms in the model may improve the likelihood of detecting a statistically significant effect of reservoirs processes on aquatic TOC loads.

## References Cited

- Alexander, R.B., Smith, R.A., Schwarz, G.E., Boyer, E.W., Nolan, J.V., and Brakebill, J.W., 2008, Differences in phosphorus and nitrogen delivery to the Gulf of Mexico from the Mississippi River Basin: *Environmental Science & Technology*, v. 42, no. 3, p. 822–830, accessed January 11, 2011, at <http://dx.doi.org/10.1021/es0716103>).
- Alexander, R.B., Böhlke, J.F., Boyer, E.W., David, M., Harvey, J.W., Mulholland, P.J., Seitzinger, S.P., Tobias, C.R., Tonitto, C., and Wollheim, W., 2009, Dynamic modeling of nitrogen losses in river networks unravels the coupled effects of hydrological and biogeochemical processes: *Biogeochemistry*, v. 93, nos. 1–2, p. 91–116, accessed January 11, 2011, at <http://dx.doi.org/10.1007/s10533-008-9274-8>).
- Allen, H. L., 2007, Production and Utilization of dissolved organic carbon during in situ phytoplankton photosynthesis measurements: *Internationale Revue der gesamten Hydrobiologie und Hydrographie*, v. 58, no. 6, p. 843–849.
- Arvola, L., Raike, A., Kortelainen, P., and Jarvinen, M., 2004, The effect of climate and land use on TOC concentrations and loads in Finnish rivers: *Boreal Environmental Research*: v. 9, p. 381–387.
- Billen, G., Decamps, H., Garnier, J., Boet, P., Meybeck, M., and Servais, P., 1995, Atlantic river systems of Europe, in Cushing, C.E., Cummins, K.W., and Minshall, G.W., eds., *Ecosystems of the world; river and stream ecosystems*: Amsterdam, Elsevier, 817 p.
- Boyer, E.W., Hornberger, G.M., Bencala, K.E., and McKnight, D.M., 2000, Effects of asynchronous snowmelt on flushing of dissolved organic carbon: a mixing model approach: *Hydrological Processes*, v. 14, p. 3291–3308.
- Cole, J.J., Prairie, Y.T., Caraco, N.F., McDowell, W.H., Tranvik, L.J., Striegl, R.G., Duarte, C.M., Kortelainen, P., Downing, J.A., Middelburg, J.J., and Melack, J., 2007, Plumbing the global carbon cycle—integrating inland waters into the terrestrial carbon budget: *Ecosystems* 10, p. 171–184.

- Creed, I.F., Sanford, S.E., Beall, F.D., Molot, L.A., and Dillon, P.J., 2003, Cryptic wetlands: Integrating hidden wetlands in regression models of the export of dissolved organic carbon from forest landscapes: *Hydrological Processes*, v. 17, p. 3629–3648.
- Dalzell, B.J., Filley, T.R., and Harbor, J.M., 2007, The role of hydrology in annual organic carbon loads and terrestrial organic matter export from a midwestern agricultural watershed: *Geochimica et Cosmochimica Acta*, v. 71, no. 6, p. 1448–1462.
- Fisher, S.G., 1995, Stream ecosystems of the western United States, *in* Cushing, C.E., Cummins, K.W., and Minshall, G.W., eds., *Ecosystems of the world; river and stream ecosystems*: Amsterdam, Elsevier, 817 p.
- Hinton, M.J., Schiff, S., and English, M., 1998, Sources and flowpaths of dissolved organic carbon during storms in two Precambrian shield catchments: *Biogeochemistry*, v. 41, p. 175–197.
- Hook, A.M., and Yeakley, J.A., 2005, Stormflow dynamics of dissolved organic carbon and total dissolved nitrogen in a small urban watershed: *Biogeochemistry*, v. 75, no. 3, p. 409–431, accessed January 11, 2011, at <http://dx.doi.org/10.1007/s10533-005-1860-4>.
- Hope, D., Billett, M.F., and Cresser, M.S., 1994, A review of the export of carbon in river water—fluxes and processes: *Environmental Pollution*, v. 84, p. 301–324.
- Leopold, L.B., and Maddock, T., 1953, The hydraulic geometry of stream channels and some physiographic implications: U.S. Geological Survey Professional Paper 252.
- McGlynn, B.L., and McDonnell, J.J., 2003, Role of discrete landscape units in controlling catchment dissolved organic carbon dynamics: *Water Resources Research*, v. 39, no. 4, p. 1090, accessed January 11, 2011, at <http://dx.doi.org/10.1029/2002WR001525>.
- Monteith, D.T., Stoddard, J.L., Evans, C.D., de Wit, H.A., Forsius, M., Høgåsen, T., Wilander, A., Skjelkvåle, B.L., Jeffries, D.S., Vuorenmaa, J., Keller, B., Kopáček, J., and Vesely, J., 2007, Dissolved organic carbon trends resulting from changes in atmospheric deposition chemistry: *Nature*, v. 450, p. 537–54.
- Morel, B., Durand, P., Jaffrezic, A., Gruau, G., and Molenat, J., 2009, Sources of dissolved organic carbon during stormflow in a headwater agricultural catchment: *Hydrological Processes*, v. 23, no. 20, p. 2888–2901.
- Mulholland, P.J., 2003, Large scale patterns in DOC concentration, flux, and sources *in* Findlay, S., and Sinsabaugh, R., eds., *Aquatic ecosystems—interactivity of dissolved organic matter*: Elsevier Science, p. 139–159.
- Naiman, R.J., Melillo, J.M., Lock, M.A., Ford, T.E., and Reice, S.R., 1987, Longitudinal Patterns of ecosystem processes and community structure in a subarctic river continuum: *Ecology*, v. 68, no. 5, p. 1139–1156.
- Nolan, J.V., Brakebill, J.W., Alexander, R.B., and Schwarz, G.E., 2002, Enhanced river reach file 2: U.S. Geological Survey Open-File Report 02–40. (Also available at <http://pubs.er.usgs.gov/usgspubs/ofr/ofr0240>.)
- Schwarz, G.E., Hoos, A.B., Alexander, R.B., and Smith, R.A., 2006, The SPARROW surface water-quality model: theory, application and user documentation: U.S. Geological Survey Techniques and Methods Report, book 6, chap. B3, accessed January 11, 2011, at <http://pubs.usgs.gov/tm/2006/tm6b3/>.

- Sebestyen, S.D., Boyer, E.W., and Shanley, J.B., 2009, Responses of stream nitrate and dissolved organic carbon loadings to hydrological forcing and climate change in an upland forest of the northeastern United States: *Journal of Geophysical Research*, v. 114, G02002, accessed January 11, 2011, at <http://dx.doi.org/10.1029/2008JG000778>).
- Smith, R.A. 1980, The theoretical basis for estimating phytoplankton production and specific growth rate from chlorophyll, light, and temperature data: *Ecological Modeling*, v. 10: p. 243–264.
- Smith, J.H., Stehman, S.V., Wickham, J.D., and Yang, L., 2003, Effects of landscape characteristics on land-cover class accuracy: *Remote Sensing of Environment*, v. 84, p. 342–349.
- Smith, R.A., Schwarz, G.E., and Alexander, R.B., 1997, Regional interpretation of water-quality monitoring data: *Water Resources Research*, v. 33, no. 12, p. 2781–2798.
- Soballe, D.M., and Kimmel, B.L., 1987, A large-scale comparison of factors influencing phytoplankton abundance in rivers, lakes, and impoundments: *Ecology*, v. 68, p. 1943–1954.
- Van Nieuwenhuysen, E.E., and Jones, J.R., 1996, Phosphorus-chlorophyll relationship in temperate streams and its variation with stream catchment area: *Canadian Journal of Fisheries and Aquatic Sciences*, v. 53, p. 99–105.
- Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., and Van Driel, N., 2001, Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources: *Photogrammetric Engineering and Remote Sensing*, v. 67, p. 650–652.
- Webster, J.R., Wallace, J.B., and Benfield, E.F., 1995, Organic processes in streams of the eastern United States, *in* Cushing, C.E., Cummins, K.W., and Minshall, G.W., eds., *Ecosystems of the world; river and stream ecosystems*: Amsterdam, Elsevier, 817 p.
- Wickham, J.D., Stehman, S.V., Smith, J.H., and Yang, L., 2004, Thematic accuracy of MRLC-NLCD land cover for the western United States: *Remote Sensing of Environment*, v. 91, p. 452–468.
- Wolf-Gladrow, D.A., Riebesell, U., Burkhardt, S., and Bijma, J., 1999, Direct effects of CO<sub>2</sub> concentration on growth and isotopic composition of marine plankton: *Tellus*, v. 51B, p. 461–476.