Front cover: Estimated distribution of atrazine use on United States cropland during 1997. Darker shades indicate the greatest use intensity. Blue circles are streams used for model development.

Development and Application of Watershed Regressions for Pesticides (WARP) for Estimating Atrazine Concentration Distributions in Streams

By Steven J. Larson, Charles G. Crawford, and Robert J. Gilliom

---

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 03-4047

Prepared in cooperation with

U.S. Environmental Protection Agency

NATIONAL WATER-QUALITY ASSESSMENT PROGRAM

8056-84

Sacramento, California
2004

U.S. DEPARTMENT OF THE INTERIOR

GALE A. NORTON, *Secretary*

U.S. GEOLOGICAL SURVEY

Charles G. Groat, *Director*

Suggested citation: Larson, S.J., Crawford, C.G., and Gilliom, R.J., 2004, Development and application of watershed regressions for pesticides (WARP) for estimating atrazine concentration distributions in streams: U.S. Geological Survey Water-Resouces Investigations Report 03-4047, 68 p.

For additional information write to:

U.S. Geological Survey
Water Resources
Placer Hall, Suite 2012
6000 J Street
Sacramento, California 95819-6129
http://ca.water.usgs.gov

Copies of this report can be purchased from:

U.S. Geological Survey
Information Services
Building 810
Box 25286, Federal Center
Denver, CO 80225-0286

FOREWORD

The U.S. Geological Survey (USGS) is committed to serve the Nation with accurate and timely scientific information that helps enhance and protect the overall quality of life, and facilitates effective management of water, biological, energy, and mineral resources. (http://www.usgs.gov/). Information on the quality of the Nation's water resources is of critical interest to the USGS because it is so integrally linked to the long-term availability of water that is clean and safe for drinking and recreation and that is suitable for industry, irrigation, and habitat for fish and wildlife. Escalating population growth and increasing demands for the multiple water uses make water availability, now measured in terms of quantity and quality, even more critical to the long-term sustainability of our communities and ecosystems.

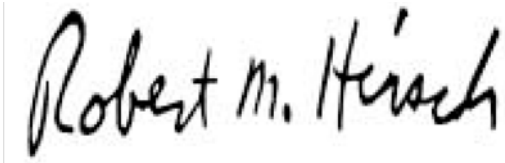
The USGS implemented the National Water-Quality Assessment (NAWQA) Program to support national, regional, and local information needs and decisions related to water-quality management and policy. (http://water.usgs.gov/nawqa).Shaped by and coordinated with ongoing efforts of other Federal, State, and local agencies, the NAWQA Program is designed to answer: What is the condition of our Nation's streams and ground water? How are the conditions changing over time? How do natural features and human activities affect the quality of streams and ground water, and where are those effects most pronounced? By combining information on water chemistry, physical characteristics, stream habitat, and aquatic life, the NAWQA Program aims to provide science-based insights for current and emerging water issues and priorities. NAWQA results can contribute to informed decisions that result in practical and effective water-resource management and strategies that protect and restore water quality.

Since 1991, the NAWQA Program has implemented interdisciplinary assessments in more than 50 of the Nation's most important river basins and aquifers, referred to as Study Units. (http://water.usgs.gov/nawqa/nawqamap.html). Collectively, these Study Units account for more than 60 percent of the overall water use and population served by public water supply, and are representative of the Nation's major hydrologic landscapes, priority ecological resources, and agricultural, urban, and natural sources of contamination.

Each assessment is guided by a nationally consistent study design and methods of sampling and analysis. The assessments thereby build local knowledge about water-quality issues and trends in a particular stream or aquifer while providing an understanding of how and why water quality varies regionally and nationally. The consistent, multi-scale approach helps to determine if certain types of water-quality issues are isolated or pervasive, and allows direct comparisons of how human activities and natural processes affect water quality and ecological health in the Nation's diverse geographic and environmental settings. Comprehensive assessments on pesticides, nutrients, vola-tile organic compounds, trace metals, and aquatic ecology are developed at the national scale through comparative analysis of the Study-Unit findings. (http://water.usgs.gov/nawqa/natsyn.html).

The USGS places high value on the communication and dissemination of credible, timely, and relevant science so that the most recent and available knowledge about water resources can be applied in management and policy decisions. We hope this NAWQA publication will provide you the needed insights and information to meet your needs, and thereby foster increased awareness and involvement in the protection and restoration of our Nation's waters.

The NAWQA Program recognizes that a national assessment by a single program cannot address all water-resource issues of interest. External coordination at all levels is critical for a fully integrated understanding of watersheds and for cost-effective management, regulation, and conservation of our Nation's water resources. The Program, therefore, depends extensively on the advice, cooperation, and information from other Federal, State, interstate, Tribal, and local agencies, non-government organizations, industry, academia, and other stakeholder groups. The assistance and suggestions of all are greatly appreciated.



Robert M. Hirsch
Associate Director for Water

CONTENTS

FIGURES

TABLES

CONVERSION FACTORS, ABBREVIATIONS, AND ACRONYMS

CONVERSION FACTORS

Multiply	By	To obtain
kilogram (kg)	2.205	pound, avoirdupois
kilometer (km)	0.6214	mile
kilogram per square kilometer (kg/km^2)	5.710	pounds per square mile
kilogram per square kilometer (kg/km^2)	0.008922	pounds per acre
square kilometer (km^2)	0.3861	square mile
square kilometer (km^2)	247.1	acre

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

$$°F=1.8 °C+32.$$

ABBREVIATIONS and ACRONYMS

µg/L, microgram per liter

µm, micrometer

L, liter

mg/L, milligram per liter

ARP, Acetochlor Registration Partnership

CI, confidence interval

CIESIN, Center for International Earth Science Information Network

DEM, Digital Elevation Model

FFDCA, Federal Food,Drug, and Cosmetic Act

FIFRA, Federal Insecticide, Fungicide, and Rodenticide Act

FQPA, Food Quality Protection Act

K-factor, soil erodibility factor

LR, log regression

OLS, ordinary least squares

pR^2, pseudo R-squared (R-squared value used for tobit regression)

NASQAN, National Stream Quality Accounting Network

NAWQA, National Water-Quality Assessment

NCFAP, National Center for Food and Agricultural Policy

NLCD, National Land Cover Data set

NOAA, National Oceanic and Atmospheric Administration

NRI, National Resources Inventory

NWQL, National Water Quality Laboratory

PI, prediction interval

PRISM, Parameter-Elevation Regressions on Independent Slopes (model)

R^2, coefficient of multiple determination

R-factor, rainfall erosivity factor

RSME, root mean square error

SAS, Statistical Analysis System

STATSGO, State Soil Geographic data base

USEPA, U.S. Environmental Protection Agency

USGS, U.S. Geological Survey

USLE, Universal Soil Loss Equation

WARP, Watershed Regressions for Pesticides

WQL, Water Quality Laboratory

Development and Application of Watershed Regressions for Pesticides (WARP) for Estimating Atrazine Concentration Distributions in Streams

By Steven J. Larson, Charles G. Crawford, and Robert J. Gilliom

ABSTRACT

Regression models were developed for predicting atrazine concentration distributions in rivers and streams, using the Watershed Regressions for Pesticides (WARP) methodology. Separate regression equations were derived for each of nine percentiles of the annual distribution of atrazine concentrations and for the annual time-weighted mean atrazine concentration. In addition, seasonal models were developed for two specific periods of the year—the high season, when the highest atrazine concentrations are expected in streams, and the low season, when concentrations are expected to be low or undetectable. Various nationally available watershed parameters were used as explanatory variables, including atrazine use intensity, soil characteristics, hydrologic parameters, climate and weather variables, land use, and agricultural management practices. Concentration data from 112 river and stream stations sampled as part of the U.S. Geological Survey's National Water-Quality Assessment and National Stream Quality Accounting Network Programs were used for computing the concentration percentiles and mean concentrations used as the response variables in regression models. Tobit regression methods, using maximum likelihood estimation, were used for developing the models because some of the concentration values used for the response variables were censored (reported as less than a detection threshold). Data from 26 stations not used for model development were used for model validation.

The annual models accounted for 62 to 77 percent of the variability in concentrations among the 112 model development stations. Atrazine use intensity (the amount of atrazine used in the watershed divided by watershed area) was the most important explanatory variable in all models, but additional watershed parameters significantly increased the amount of variability explained by the models. Predicted concentrations from all 10 models were within a factor of 10 of the observed concentrations at most model development and model validation stations. Results for the two sets of seasonal models were similar. Concentration distributions derived from the seasonal-model predictions provided additional information compared to distributions derived from the annual models.

INTRODUCTION

The U.S. Environmental Protection Agency (USEPA) regulates pesticides under two major federal statutes, the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) and the Federal Food, Drug, and Cosmetic Act (FFDCA). FIFRA prescribes labeling and other regulatory requirements to prevent adverse effects on health or the environment, and FFDCA establishes maximum legally permissible levels for pesticide residues in food. The Food Quality Protection Act of 1996 (FQPA) requires the USEPA to substantially alter how it has regulated human exposure to pesticides through food under FIFRA and FFDCA. Among other changes, the FQPA establishes a single health-based standard for pesticide residues in all types of food, replacing sometimes conflicting standards in the old law, and requires that pesticide tolerances in food take into account exposure through drinking water and all nonoccupational exposures. The requirement to incorporate drinking water into the setting of pesticide tolerances has resulted in the need for estimates of the concentrations of pesticides in drinking water. Unlike food, for which representative samples are obtained from a small number of regional distribution centers through the U.S. Department of Agriculture's Pesticide Data Program (U.S. Department of Agriculture, 1999a), drinking water comes from a variety of local sources. Pesticide concentrations in surface waters in different parts of the country can vary significantly because of differences in pesticide use, application practices and timing, and watershed characteristics. Monitoring pesticide concentrations with a sampling frequency sufficient to obtain the desired frequency distribution of pesticide concentrations at the large number of surface-water supplies in the United States would be prohibitively expensive.

Larson and Gilliom (2001) described a method of estimating pesticide concentrations in streams from watershed characteristics. The method, known as Watershed Regressions for Pesticides (WARP), is based on empirical relations between pesticide concentrations observed at monitoring stations and selected nationally available watershed characteristics such as pesticide use and soil and hydrologic characteristics. The objective of this approach is to be able to estimate the annual frequency distribution of pesticide concentrations at unmonitored streams (fig. 1).

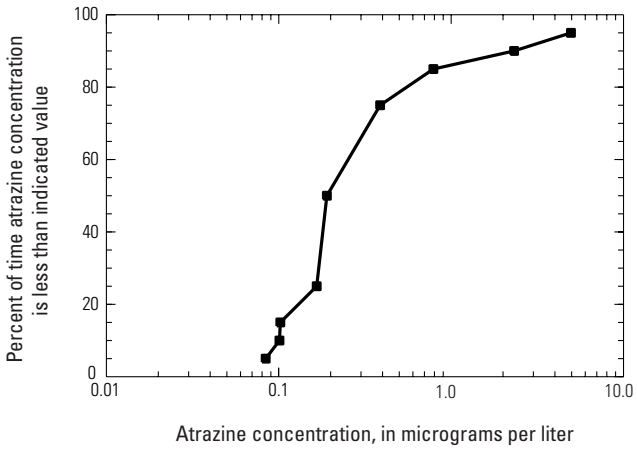


Figure 1. Example annual frequency distribution of atrazine concentrations, Sangamon River at Monticello, Illinois, for the 12-month period beginning October 1, 1996.

Purpose and Scope

This report describes new WARP models for atrazine. Improvements over previous work (Larson and Gilliom, 2001) include an increase in the number of concentration percentiles modeled; expansion of the model development data set, both in the number of stations used and the time period covered; expansion of the data set used for model validation; use of standardized criteria for inclusion of stations; and the use of more rigorous time-weighting procedures for computing concentration percentiles and annual mean concentrations. In addition, seasonal models were developed for the high season (the application period and the period of potential runoff to streams) and the low season (the remainder of the year) using the WARP method. The seasonal models provide estimates of concentration percentiles and the mean concentration for each of these two periods of the year. Separate seasonal estimates of atrazine concentration distributions may be useful for risk-assessment applications.

One purpose of this study is to provide support to the USEPA for risk assessments associated with the FQPA. Owing to the large uncertainties associated with these risk assessments, and the use of safety factors in the risk calculations, a relatively large uncertainty (for example, plus or minus 1 order of magnitude) can be tolerated in the concentrations predicted from models used for the drinking water part of the risk assessments. For this reason, comparisons between predicted and observed concentrations in this report often are in terms of the percentage of predictions within an order of magnitude of the observed values, although the actual prediction errors for the atrazine models described in this report often were substantially less than an order of magnitude.

The U.S. Geological Survey's (USGS) National Water-Quality Assessment (NAWQA) and National Stream Quality Accounting Network (NASQAN) Programs are the primary sources of atrazine concentration data that were used for this study. These data are from 112 sampling stations that represent a wide variety of environmental settings across the United States; drainage areas ranged from 17 to 2,965,000 km^2. Additional atrazine concentration data from the Water Quality Laboratory (WQL) of Heidelberg College in Tiffin, Ohio, and from the monitoring program of the Acetochlor Registration Partnership (ARP), were obtained for use in model validation.

The regression models described in this report were derived for predicting atrazine concentrations in rivers and streams of the conterminous United States. Some watershed characteristics known to significantly affect atrazine runoff and, thus, atrazine concentrations in streams, were not considered as explanatory variables in the regression models because they are not available for all watersheds across the country. For example, an indicator of the temporal correspondence of pesticide use and precipitation would be useful. However, restricting explanatory variables to those available nationally allows application of WARP-based regression models to any watershed in the nation that is reasonably represented in the range of data available for model development.

Acknowledgments

Dr. R. Peter Richards and Mr. Jack Kramer of the WQL of Heidelberg College provided pesticide data collected by the laboratory. Dr. David Gustafson of Monsanto Company provided data collected by the ARP. Naomi Nakagaki and Gail P. Thelin of the USGS provided the agricultural pesticide use estimates and watershed characteristics. The survreg procedure for the S language written by Terry M. Therneau of the Mayo Clinic was used to fit the regression models developed as part of this study. This work has benefited from discussions with members of the Intergovernmental FQPA Steering Committee Technical Working Group. Finally, funding for this study was partially provided by the USEPA's Office of Pesticide Programs.

METHODS

Methods used for this study included obtaining appropriate atrazine concentration data, estimating agricultural atrazine use and other watershed characteristics, selecting suitable statistical methods for computing annual time-weighted percentile and mean concentrations, and developing regression models to estimate atrazine concentrations.

Atrazine Data Used for Model Development

Atrazine concentration data collected as part of the NAWQA and NASQAN Programs were used for model development. The NAWQA Program collects samples from streams and rivers in about 50 study units across the United States. The study units are major hydrologic basins, which collectively encompass about two-thirds of the population and water use in the United States (Hirsch and others, 1988; U.S. Geological Survey, 1999). The NASQAN Program collects samples from major rivers of the United States (Colorado, Columbia, Mississippi, Rio Grande, and Yukon rivers) and their tributaries (Hooper and others, 1997; U.S. Geological Survey, 2002).

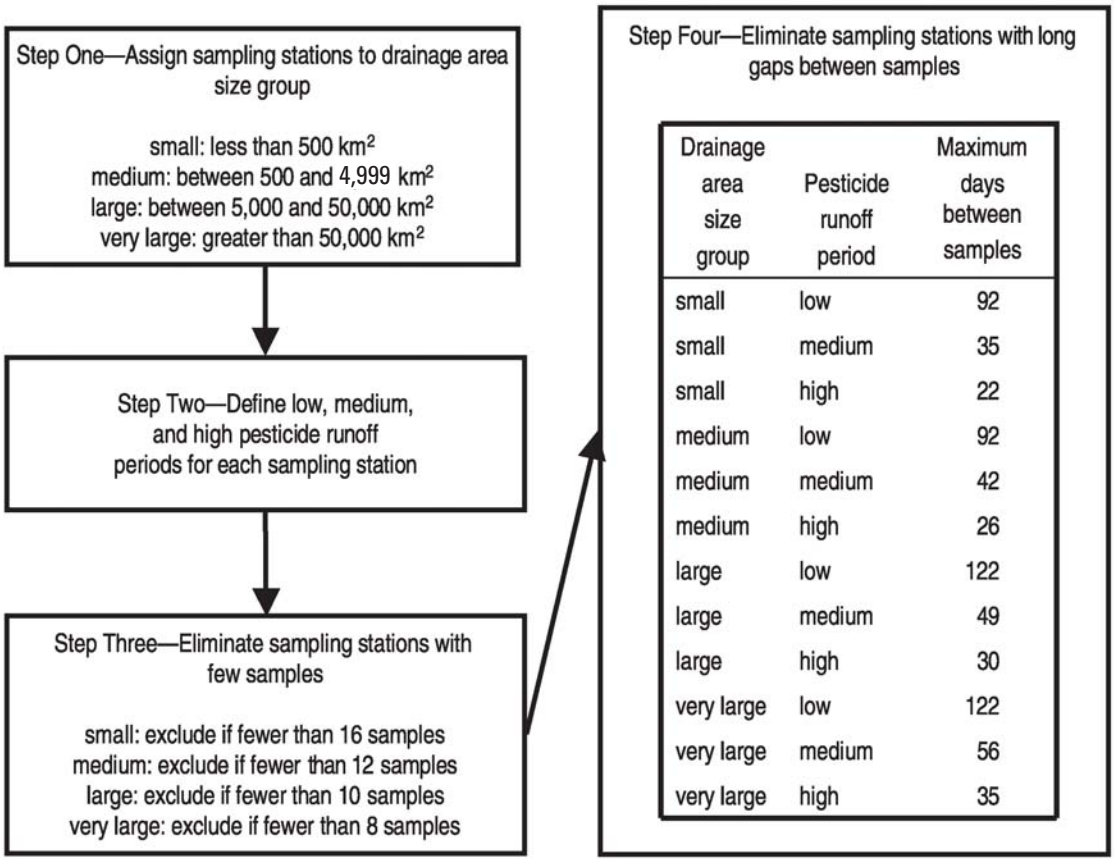
Sample Collection and Analysis Methods

Depth- and width-integrated samples were collected as part of the NAWQA Program typically every 1 to 3 weeks during the growing season, depending on the site, and about monthly during the rest of the year. The NASQAN Program collects samples about monthly with several additional samples collected during periods of high streamflow. Samples for the NAWQA Program were collected following procedures described by Shelton (1994); samples for the NASQAN Program were collected following protocols of the U.S. Geological Survey (1997 to present). Samples were typically collected using a USGS DH-81 sampler when streams were wadable or a USGS D-77 TM or modified frame sampler suspended from a small hand-operated crane when streams were not. A 3-L Teflon bottle and nozzle assembly was used with the DH-81 and D-77 TM samplers. A Teflon bag was used with the modified frame sampler.

Shortly after collection, samples were filtered through a baked 0.7 μm glass-fiber filter. Water was forced through the filter by a Teflon diaphragm pump through Teflon tubing. Equipment was cleaned with phosphate-free detergent and rinsed using deionized water and methanol, followed by native water before sample collection and processing. After field processing, samples were chilled until analyzed by gas chromatography/mass spectrometry (Zaugg and others, 1995; Lindley and others, 1996) for atrazine at the USGS National Water Quality Laboratory (NWQL) in Denver, Colo. This method requires solid-phase extraction of samples, which was either done in the field or by the NWQL within 96 hours of sample collection. Samples were kept chilled until extracted. The quality-control protocol used for NAWQA sampling is described by Mueller and others (1997). Results from analysis of NAWQA quality assurance samples are reported by Martin (1999) and Martin and others (1999). Analytical recovery of atrazine is monitored for the NWQL method through the use of spiked laboratory samples. For the period 1992 through 1996, mean recovery of atrazine in 1,002 spiked samples was 93.4 percent, with a relative standard deviation of 20.0 percent (Martin, 1999).

Sampling Station Selection

Atrazine data from selected NAWQA and NASQAN sampling stations were used for this study. Stations that had few samples or significant gaps between samples were excluded. More samples were required at small streams than at large streams. The maximum gap allowed between samples was longer for large streams than for small streams, and for periods when atrazine concentrations were expected to be lower and less variable (for example, during the winter in northern states when atrazine use and runoff is low). The criteria used for selecting sampling stations were subjectively chosen with the goal of increasing the accuracy of annual percentile and mean atrazine concentrations while retaining a large number of sampling locations. The procedure used to select sampling stations is shown in figure 2. Not all NASQAN stations that met the sampling frequency requirements were used because this program collects data from multiple stations on large rivers. To minimize correlation between stations, only one station each on the upper and lower parts of the Ohio and Mississippi Rivers and one station on the Missouri River were used.



Drainage area size group	Pesticide runoff period	Maximum days between samples
small	low	92
small	medium	35
small	high	22
medium	low	92
medium	medium	42
medium	high	26
large	low	122
large	medium	49
large	high	30
very large	low	122
very large	medium	56
very large	high	35

Figure 2. Procedure used to select sampling stations for use in Watershed Regressions for Pesticides (WARP) model development.

km^2, square kilometer.

The pesticide runoff periods (table 1) used to screen stations were determined in consultation with USGS personnel familiar with different regions of the country. Each sampling station was assigned to one of seven pesticide runoff groups (A–G in table 1). The same group was used for all stations in a NAWQA study unit. Low, medium, and high runoff periods were chosen to qualitatively reflect the likelihood of pesticides being applied and transported by streams during the period in a given area of the country. The purpose was to identify periods of the year when fewer samples were needed to characterize pesticide concentrations. For example, in a midwestern state such as Indiana (group F), the low period extended from October through March; the medium period included April, August, and September; and the high period extended from May through July. In southern states, such as Georgia and Florida (group A), where the growing season is longer and autumn and winter temperatures are warmer than in the Midwest, the low period included November and December, the medium period included January and October, and the high period extended from February through September. No attempt was made to correlate periods across regions. Thus, pesticide concentrations in a medium or low period from a high pesticide use region of the country may be greater than concentrations during a high period in a low use region of the country.

Data from 103 NAWQA and 9 NASQAN river and stream stations met the screening criteria and were used for development of the regression models (fig. 3 and table 2). The selected NAWQA sampling stations are not distributed evenly across the country: more stations are in the eastern United States and on the west coast than in the Great Plains. The watershed areas of the 112 stations ranged from 17 to 2,965,000 km^2; the median drainage area was 1,244 km^2. Twenty-five percent of the watersheds were smaller than 144 km^2, and 25 percent were larger than 9,500 km^2. Sampling stations for the NAWQA Program were not randomly chosen. Watersheds of NAWQA stations are either relatively homogeneous, representing a particular combination of land use, geomorphic, and geologic features in a NAWQA study unit, or include a mixture of the land uses and natural features in a study unit. The latter type of stations are larger than the former and are usually near the mouths of rivers. Watersheds of the NASQAN stations used are much larger than those of NAWQA stations. The median watershed area of the 9 selected NASQAN stations was nearly 300 times the median watershed area of the 103 selected NAWQA stations.

The number of samples collected per year at the 103 NAWQA stations ranged from 11 to 55; median number was 24. The number of samples collected at the 9 NASQAN stations ranged from 13 to 21; median number was 18. More samples generally were collected at stations that have smaller watersheds. The median number of samples collected, by drainage area, was as follows: less than 500 km^2, 30; 500 to 5,000 km^2, 24; and greater than 5,000 km^2, 19. The median time between samples during spring and summer was 10 days at NAWQA sampling stations and 17 days at NASQAN stations. During autumn and winter, the median time between samples was 26 days at NAWQA stations and 27 days at NASQAN stations.

Table 1. Pesticide runoff periods used to screen sampling stations.

Pesticide runoff group	Months in low period	Months in medium period	Months in high period
A	November–December	January; October	February–September
B	October–December	January; September	February–August
C	January; December	February; November	March–October
D	January–February; October–December	September	March–August
E	January–March; October–December	August–September	April–July
F	January–March; October–December	April; August–September	May–July
G	January–April; September–December	none	May–August

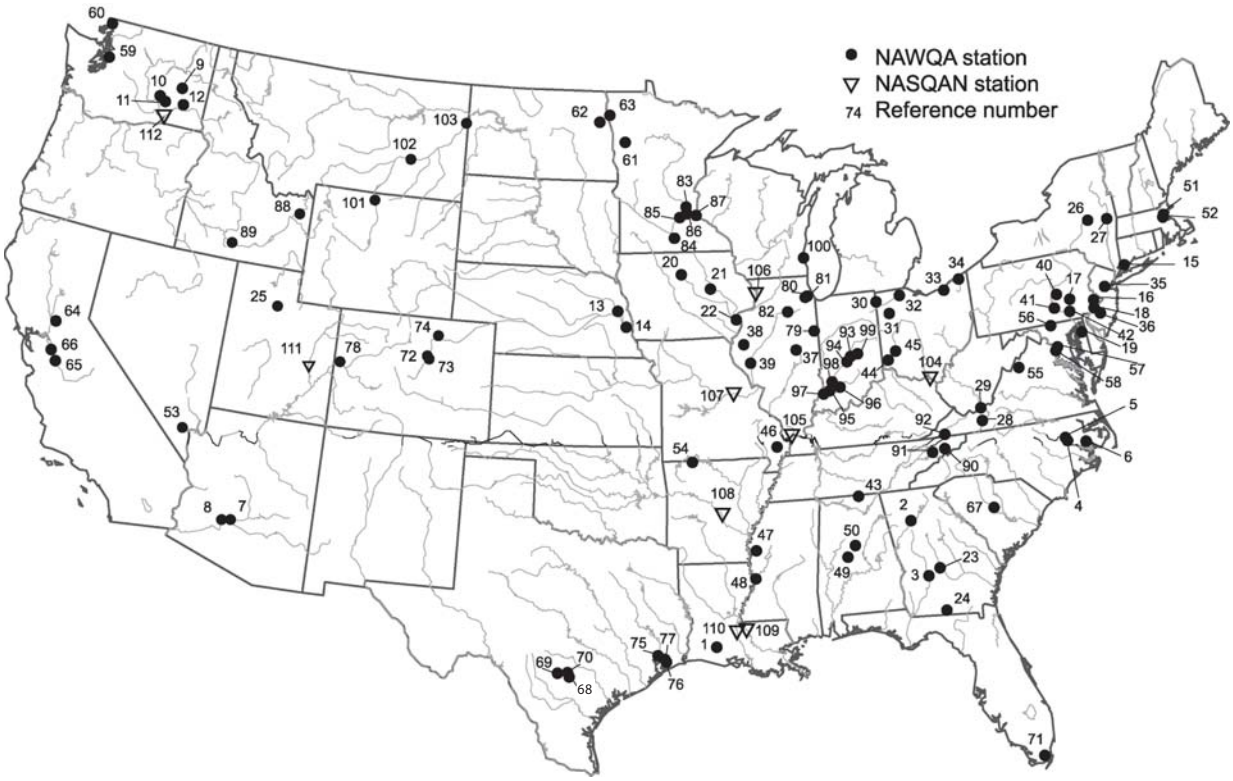


Figure 3. Location of sampling stations used for WARP model development.

WARP, Watershed Regressions for Pesticides; NAWQA, National Water-Quality Assessment; NASQAN, National Stream Quality Accounting Network.

Atrazine Data Used for Model Validation

A relatively small set of stations was available for model validation (fig. 4). The validation data set consists of 9 stream stations sampled by the Heidelberg College WQL, 10 stream stations from the ARP sampling program, and 7 USGS (mainly NASQAN) stations on rivers and streams not used as part of the model development data set. The WQL at Heidelberg College has been collecting pesticide data at several streams that drain to Lake Erie since the early 1980s. This pesticide monitoring program is among the longest running and most intensive pesticide data collection programs in the United States. Information about the WQL monitoring program and some findings based on its data are given in Richards and Baker (1993) and Richards and others (1996).

The ARP was formed by manufacturers of the herbicide acetochlor as part of the registration of this herbicide with the USEPA in 1994 (Acetochlor Registration Partnership, 2002; U.S. Environmental Protection Agency, 2002a). As part of the agreement, a monitoring program was established to collect water samples from streams, lakes, and reservoirs in the acetochlor use area and analyze them for acetochlor and several other herbicides, including atrazine. Sampling in the ARP program was targeted at water bodies used as sources of drinking water, and many of the analyses were done on water after treatment at a water utility. In some cases, untreated source water also was sampled. The ARP data included in the validation data set are from analyses of untreated water only. Stations from the ARP program are not identified by name in this report because locations of drinking water sources are confidential.

Table 2. Sampling stations used for Watershed Regressions for Pesticides (WARP) model development.

[km^2, square kilometer; ID, identification; NAWQA, National Water-Quality Assessment; NASQAN, National Stream Quality Accounting Network]

Sampling station number (fig. 3)	USGS Station ID	Station name	Beginning date of 12-month data period	Drainage area (km^2)	Pesticide runoff group (from table 1)
NAWQA stations					
1	08012150	Mermentau River at Mermentau, Louisiana	11/01/1998	3,580	D
2	02335870	Sope Creek near Marietta, Georgia	03/01/1993	79	A
3	02350080	Lime Creek at County Road near Cobb, Georgia	03/01/1993	161	A
4	02083500	Tar River at Tarboro, North Carolina	03/01/1993	5,750	D
5	02083833	Pete Mitchell Swamp near Penny Hill, North Carolina	03/01/1993	44	D
6	02084558	Albemarle Canal near Swindell, North Carolina	03/01/1993	191	D
7	09514000	Buckeye Canal near Avondale, Arizona	12/01/1996	117,000	C
8	09517000	Hassayampa River near Arlington, Arizona	02/01/1997	3,970	C
9	12464770	Crab Creek near Ritzville, Washington	04/01/1993	1,190	E
10	12472380	Crab Creek Lateral near Othello, Washington	04/01/1993	145	E
11	12473740	El 68 D Wasteway near Othello, Washington	04/01/1993	377	E
12	13351000	Palouse River at Hooper, Washington	04/01/1993	6,380	E
13	06800000	Maple Creek near Nickerson, Nebraska	05/01/1998	955	F
14	06805500	Platte River at Louisville, Nebraska	05/01/1997	221,000	F
15	01209710	Norwalk River at Winnipauk, Connecticut	03/01/1993	85	F
16	01464907	Little Neshaminy Creek near Neshaminy, Pennsylvania	01/01/1999	72	F
17	01470779	Tulpehocken Creek near Bernville, Pennsylvania	12/01/1998	184	F
18	01474500	Schuylkill River at Philadelphia, Pennsylvania	01/01/1999	4,900	F
19	01493112	Chesterville Branch near Crumpton, Maryland	02/01/1999	17	F
20	05449500	Iowa River near Rowan, Iowa	03/01/1997	1,080	F
21	05464220	Wolf Creek near Dysart, Iowa	03/01/1997	775	F
22	05465500	Iowa River at Wapello, Iowa	03/01/1997	32,400	F
23	02215100	Tucsawhatchee Creek near Hawkinsville, Georgia	03/01/1993	420	A
24	02318500	Withlacoochee River near Quitman, Georgia	03/01/1993	3,860	A
25	10171000	Jordan River at Salt Lake City, Utah	01/01/1999	9,100	F
26	01349150	Canajoharie Creek near Canajoharie, New York	03/01/1998	154	F
27	01357500	Mohawk River at Cohoes, New York	03/01/1998	9,110	F
28	03167000	Reed Creek at Grahams Forge, Virginia	01/01/1997	669	F
29	03176500	New River at Glen Lyn, Virginia	01/01/1997	9,780	F
30	04178000	St. Joseph River near Newville, Indiana	03/01/1997	1,600	F
31	04186500	Auglaize River near Fort Jennings, Ohio	03/01/1997	858	F
32	04193500	Maumee River at Waterville, Ohio	03/01/1997	16,400	F
33	04208504	Cuyahoga River at Cleveland, Ohio	04/01/1997	2,040	F
34	04211820	Grand River at Harpersfield, Ohio	04/01/1997	1,430	F
35	01403900	Bound Brook at Middlesex, New Jersey	05/01/1996	125	F
36	01410784	Great Egg Harbor River near Sicklerville, New Jersey	05/01/1996	39	F
37	05572000	Sangamon River at Monticello, Illinois	10/01/1996	1,430	F

Table 2. Sampling stations used for Watershed Regressions for Pesticides (WARP) model development. *(Continued)*

Sampling station number (fig. 3)	USGS Station ID	Station name	Beginning date of 12-month data period	Drainage area (km^2)	Pesticide runoff group (from table 1)
38	05584500	La Moine River at Colmar, Illinois	04/01/1997	1,700	F
39	05586100	Illinois River at Valley City, Illinois	05/01/1997	69,200	F
40	01555400	East Mahantango Creek at Klingerstown, Pennsylvania	04/01/1993	115	F
41	01571490	Cedar Run at Eberlys Mill, Pennsylvania	04/01/1993	32	F
42	01576540	Mill Creek near Lyndon, Pennsylvania	03/01/1993	140	F
43	0357479650	Hester Creek near Plevena, Alabama	02/01/1999	76	F
44	03274000	Great Miami River at Hamilton, Ohio	03/01/1999	9,400	F
45	393944084120700	Holes Creek at Kettering, Ohio	04/01/1999	51	F
46	07043500	Little River Ditch no. 1 near Morehouse, Missouri	02/01/1996	1,140	A
47	07288650	Bogue Phalia near Leland, Mississippi	02/01/1997	1,300	A
48	07288955	Yazoo River near Long Lake, Mississippi	02/01/1997	34,800	A
49	0242354750	Cahaba Valley Creek at Pelham, Alabama	02/01/1999	65	C
50	02424000	Cahaba River at Centreville, Alabama	02/01/1999	2,660	C
51	01102500	Aberjona River at Winchester, Massachusetts	04/01/1999	59	F
52	01104615	Charles River at Watertown, Massachusetts	05/01/1999	694	F
53	094196783	Las Vegas Wash near Las Vegas, Nevada	06/01/1993	2,640	C
54	07053250	Yocum Creek near Oak Grove, Arkansas	02/01/1994	134	G
55	01621050	Muddy Creek at Mount Clinton, Virginia	04/01/1993	37	F
56	01639000	Monocacy River at Bridgeport, Maryland	06/01/1994	456	F
57	01646580	Potomac River at Washington, D.C.	06/01/1999	30,000	F
58	01654000	Accotink Creek near Annandale, Virginia	03/01/1994	60	F
59	12128000	Thornton Creek near Seattle, Washington	03/01/1996	29	E
60	12213140	Nooksack River at Brennan, Washington	03/01/1996	2,020	E
61	05062500	Wild Rice River at Twin Valley, Minnesota	04/01/1993	2,410	G
62	05082625	Turtle River near Arvilla, North Dakota	03/01/1993	658	G
63	05085900	Snake River above Alvarado, Minnesota	05/01/1993	565	G
64	11447360	Arcade Creek near Del Paso Heights, California	12/01/1996	81	B
65	11273500	Merced River near Newman, California	02/01/1993	3,620	B
66	11303500	San Joaquin River near Vernalis, California	11/01/1992	19,000	B
67	02169570	Gills Creek at Columbia, South Carolina	02/01/1996	154	D
68	08178800	Salado Creek at San Antonio, Texas	02/01/1997	505	F
69	08180640	Medina River at La Coste, Texas	02/01/1997	2,100	F
70	08181800	San Antonio River near Elmendorf, Texas	02/01/1997	4,530	F
71	252414080333200	C-111 Canal near Homestead, Florida	10/01/1996	132	A
72	06713500	Cherry Creek at Denver, Colorado	03/01/1993	61	F
73	06714000	South Platte River at Denver, Colorado	12/01/1993	10,000	F
74	06753990	Lonetree Creek near Greeley, Colorado	04/01/1993	1,470	F
75	294349094345999	East Fork Double Bayou near Anahuac, Texas	03/01/1994	111	F
76	295001094384699	Whites Bayou near Anahuac, Texas	03/01/1994	27	F

Table 2. Sampling stations used for Watershed Regressions for Pesticides (WARP) model development. *(Continued)*

Sampling station number (fig. 3)	USGS Station ID	Station name	Beginning date of 12-month data period	Drainage area (km^2)	Pesticide runoff group (from table 1)
77	295740094542399	West Prong Old River near Dayton, Texas	03/01/1994	75	F
78	09153290	Reed Wash near Mack, Colorado	11/01/1996	35	F
79	05525500	Sugar Creek at Milford, Illinois	05/01/1999	1,160	F
80	05531500	Salt Creek at Western Springs, Illinois	03/01/1999	290	F
81	05532500	Des Plaines River at Riverside, Illinois	03/01/1999	1,630	F
82	05553500	Illinois River at Ottawa, Illinois	10/01/1997	28,300	F
83	05288705	Shingle Creek Minneapolis, Minnesota	05/01/1997	73	F
84	05320270	Little Cobb River near Beauford, Minnesota	05/01/1997	336	F
85	05330000	Minnesota River near Jordan, Minnesota	05/01/1997	42,000	F
86	05330902	Nine Mile Creek at Bloomington, Minnesota	01/01/1997	115	F
87	05331580	Mississippi River at Hastings, Minnesota	04/01/1997	96,000	F
88	13055000	Teton River near St. Anthony, Idaho	05/01/1993	2,290	F
89	13092747	Rock Creek at Twin Falls, Idaho	04/01/1993	623	F
90	03466208	Big Limestone Creek near Limestone, Tennessee	05/01/1996	205	F
91	03467609	Nolichucky River near Lowland, Tennessee	03/01/1996	4,370	F
92	03526000	Copper Creek near Gate City, Virginia	05/01/1996	276	F
93	03353637	Little Buck Creek near Indianapolis, Indiana	05/01/1992	44	F
94	03354000	White River near Centerton, Indiana	04/01/1994	6,320	F
95	03360895	Kessinger Ditch near Monroe City, Indiana	04/01/1993	145	F
96	03373500	East Fork White River at Shoals, Indiana	04/01/1994	12,800	F
97	03374100	White River at Hazleton, Indiana	05/01/1996	29,300	F
98	385234087071801	White River near Elnora, Indiana	04/01/1994	12,400	F
99	394340085524601	Sugar Creek near New Palestine, Indiana	05/01/1992	246	F
100	04087000	Milwaukee River at Milwaukee, Wisconsin	04/01/1993	1,800	F
101	06279500	Bighorn River at Kane, Wyoming	01/01/1999	40,800	G
102	06295000	Yellowstone River at Forsyth, Montana	02/01/1999	102,000	G
103	06329500	Yellowstone River near Sidney, Montana	02/01/1999	177,000	G
NASQAN stations					
104	03216600	Ohio River at Greenup Dam, Kentucky	12/01/1996	159,000	F
105	03612500	Ohio River near Grand Chain, Illinois	02/01/1997	527,000	F
106	05420500	Mississippi River at Clinton, Iowa	03/01/1996	239,000	F
107	06934500	Missouri River at Hermann, Missouri	11/01/1995	1,350,000	F
108	07263620	Arkansas River at David Terry Dam, Arkansas	10/01/1995	409,000	F
109	07373420	Mississippi River at St. Francisville, Louisiana	10/01/1995	2,965,000	F
110	07381495	Atchafalaya River at Melville, Louisiana	10/01/1998	241,000[1]	F
111	09315000	Green River at Green River, Utah	06/01/1998	106,000	F
112	13353200	Snake River at Burbank, Washington	11/01/1995	279,000	F

[1] Atchafalaya River receives approximately 30 percent of the Mississippi River discharge; watershed area shown represents only the part of the watershed distinct from that of the Mississippi River.

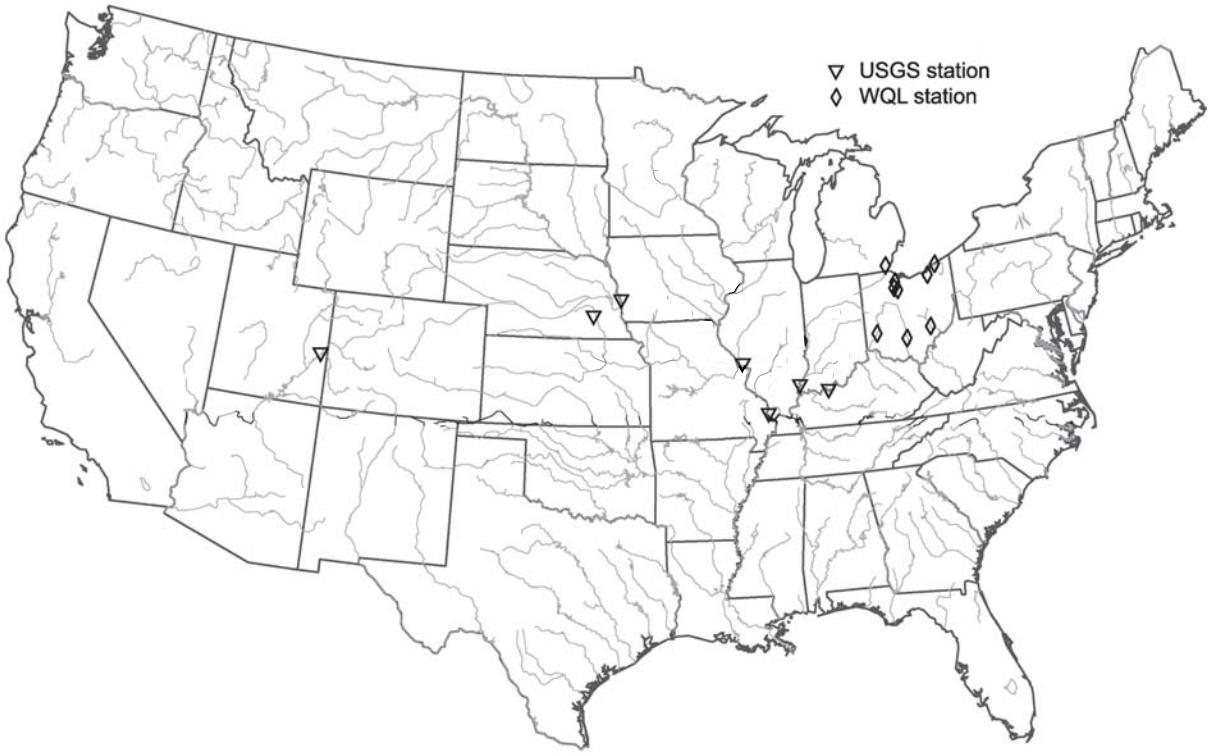


Figure 4. Location of sampling stations used for WARP model validation.

WARP, Watershed Regressions for Pesticides; USGS, U.S. Geological Survey; WQL, Water Quality Laboratory of Heidelberg College.

The sampling frequency at the validation stations sampled by Heidelberg College and the USGS satisfied the criteria shown in figure 2. For ARP stations, somewhat less stringent criteria were used. At the 10 ARP stations used for validation, 14 to 15 samples were collected during the 1-year period. Samples were collected at ARP stations about once every 2 weeks during the growing season and once every 2 months during the rest of the year. Work by Charles Crawford (U.S. Geological Survey, unpub. data, 2002) has shown that sampling programs targeted at the growing season provide nearly as much information about the annual distribution of concentrations as sampling throughout the year. This is especially true for compounds like atrazine, which have a strong seasonal concentration pattern.

Data also were available from 16 ARP stations on lakes and reservoirs. These stations were not included in the validation data set because concentrations at these stations would not be directly comparable to estimates from models developed using concentration data from streams. The models were applied to the 16 lake and reservoir stations to illustrate the systematic differences compared to streams.

Watershed Characteristics Used as Explanatory Variables

WARP models are empirical regression models. A large number of variables that could reasonably affect or indicate an influence on pesticide transport and runoff are considered as potential explanatory variables for estimating atrazine concentrations. Statistical procedures are used to select from among the potential variables on the basis of their ability to reproduce observed pesticide concentrations.

Estimation of Atrazine Use

Annual agricultural atrazine use in a watershed was estimated by combining county-level estimates of applications of atrazine to agricultural crops with information on the spatial distribution of agricultural land within the watershed. Total atrazine use in a watershed is determined by (1) computing areal weights of specific types of agricultural land for each county intersected by the watershed; (2) multiplying the weights by county atrazine use on each of the agricultural land types; and (3) summing the apportioned county estimates.

The amount of atrazine applied at the county level was estimated using methods developed by Thelin and Gianessi (2000), which integrate state-level information on pesticide applications to individual crops with county-level crop acreage information. State-level atrazine use rates were obtained from the National Center for Food and Agricultural Policy (NCFAP) and county-level crop data were obtained from the Census of Agriculture. The source for the spatial distribution of agricultural land was the circa 1992, 30-meter resolution, National Land Cover Data set (NLCD; Vogelmann and others, 2001). "Agricultural land" in the calculation of watershed use consists of the NLCD classifications pasture/hay and orchards/vineyards/other, and a combined category consisting of three individual NLCD classifications: row crops, small grains, and fallow.

Water samples used for this study were collected between 1992 and 2000. For stations with samples collected before 1995, estimates of atrazine use are based on the 1992 Census of Agriculture (U.S. Department of Commerce, 1995) and NCFAP state use estimates representing the 1991–1994 period (Gianessi and Anderson, 1995). For stations with pesticide samples collected after 1994, estimates of atrazine use are based on the 1997 Census of Agriculture (U.S. Department of Agriculture, 1999b) and NCFAP data for the period 1994–1998 (Gianessi and Marcelli, 2000).

The estimates of atrazine use obtained by these methods described here are for agricultural use only. There currently are no nationally available estimates for nonagricultural use of atrazine or other pesticides. The primary use of atrazine is application to agricultural crops, and the lack of data on nonagricultural use probably has a minimal effect on the regression analysis. Because of the wide range in drainage areas among stations used for model development and validation, the variable used to represent atrazine use in the regression models was watershed use intensity (amount of atrazine used in the watershed divided by watershed area) rather than simply the amount of atrazine used in the watershed.

Other Watershed Characteristics

Many watershed characteristics were evaluated as potential explanatory variables (table 3). In addition to atrazine use intensity, variables representing land use and population, agricultural management practices, soil properties, physical watershed characteristics, weather and climate characteristics, and hydrologic properties were considered. All variables considered are available for the entire conterminous United States. Land-use data were obtained from the National Land Cover Data set (Vogelmann and others, 2001). Georeferenced population data based on the U.S. Census were obtained from the Center for International Earth Science Information Network (1996). Agricultural practices data were obtained from the National Resources Inventory (NRI; Natural Resources Conservation Service, 2000). Soil properties data were obtained from the U.S. Department of Agriculture's State Soil Geographic database (Natural Resources Conservation Service, 1994). Mean basin elevation and slope were computed from USGS digital elevation model data (U.S. Geological Survey, 2000). Mean annual precipitation and temperature are from Owensby and Ezell (1992). The mean annual number of consecutive wet and dry days and precipitation intensity are from Hughes and others (1992). Mean annual runoff is from Gebert and others (1987). Estimates of overland flow and average subsurface contact time were obtained using a national scale application of the TOPMODEL rainfall-runoff model (Wolock, 1993).

Average R-factor values for the basins were estimated from a 2.5-minute resolution grid of mean-annual (1971–2000) R-factor values. The grid was derived from R-factor values estimated by the Illinois State Water Survey at 1,842 meteorological stations. These station data were interpolated to a grid by the Spatial Climate Analysis Service at Oregon State University using the Parameter-elevation Regressions on Independent Slopes Model (PRISM) methodology (Greg Johnson, U.S. Department of Agriculture, written commun., 5/12/2003).

Table 3. Watershed characteristics considered as explanatory variables for WARP models.

[WARP, Watershed Regressions for Pesticides; CIESIN, Center for International Earth Science Information Network; DEM, Digital Elevation Model; NCFAP, National Center for Food and Agriculture Policy; NLCD, National Land Cover data set; NOAA, National Oceanic and Atmospheric Administration NRI, National Resources Inventory; STATSGO, State Soil Geographic data base; USLE, Universal Soil Loss Equation. cm, centimeter; d, day; kg, kilogram km, kilometer; km^2, square kilometer; m, meter; mm, millimeter; yr, year]

Abbreviation	Description
Pesticide use	
useint	Atrazine agricultural use intensity [(sum of NCFAP data for applications to row crops, orchards and vineyards, and pasture and hay crops extrapolated to the basin scale) / basin area], kg/km^2
Land use and population	
ag	Percent of basin with agricultural land use [(sum of NLCD categories row crops, small grains, pasture/hay, and orchards/vineyards/other at 30-m cell resolution) x100 / basin area]
forest	Percent of basin with forest land use [(sum of NLCD categories deciduous forest, evergreen forest, and mixed forest at 30-m cell resolution) × 100 / basin area]
urban	Percent of basin with urban land use [(sum of NLCD categories low intensity residential, high intensity residential, commercial/industrial/ transportation, and urban/recreational grasses at 30-m cell resolution) × 100 / basin area]
popden	Mean 1990 population density in basin (people/km^2) (CIESIN 1-km grid)
Agricultural management practices	
artdrn	Percent of basin that is artificially drained [(aggregated representation of conservation practice categories 606, 607, and 608 from the 1997 NRI polygons converted to 1-km cells) x100 / basin area]
contill	Percent of basin under conservation tillage [(1997 NRI conservation practice 329 polygons converted to 1-km cells) × 100 / basin area]
irri	Percent of basin that is irrigated [(aggregated representation of irrigation type categories 01, 02, and 03 from the 1997 NRI polygons converted to 1-km cells) × 100 / basin area]
tile	Percent of the basin that is drained by tiles [(conservation practice category 606 from the 1997 NRI polygons converted to 1-km cells) × 100 / basin area]
Soil properties	
awc	Mean available water capacity (fraction) (STATSGO polygons converted to 1-km cells) in basin
clay	Mean percent clay in basin soils (STATSGO polygons converted to 1-km cells)
hgab	Mean percent of basin soils classified as hydrologic groups A and B (sum of soil hydrologic groups A and B from STATSGO polygons converted to 1-km cells)
hgcd	Mean percent of basin soils classified as hydrologic groups C, D, and C/D (sum of soil hydrologic groups C, D, and C/D from STATSGO polygons converted to 1-km cells)
kfact	Mean soil erodibility (K-factor for USLE) (STATSGO polygons converted to 1-km cells) in basin
orgm	Mean percent organic matter in basin soils (STATSGO polygons converted to 1-km cells)
perm	Mean soil permeability in cm/hr (STATSGO polygons converted to 1-km cells)
sand	Mean percent sand in basin soils (STATSGO polygons converted to 1-km cells)
silt	Mean percent silt in basin soils (STATSGO polygons converted to 1-km cells)
Physical watershed characteristics	
darea	Basin drainage area in km^2
elev	Mean basin elevation in m (DEM at 1-km cell resolution)
latc	Latitude of basin centroid in decimal degrees
lonc	Longitude of basin centroid in decimal degrees
slope	Mean percent slope in basin (DEM at 1-km cell resolution)

Table 3. Watershed characteristics considered as explanatory variables for WARP models. (*Continued*) .

Abbre-viation	Description
Weather/climate characteristics	
adry	Mean annual number of consecutive dry days (NOAA data interpolated & converted to 1-km cells)
appt	Mean annual 1961–90 precipitation in cm/yr (NOAA data interpolated & converted to 1-km cells)
appti	Mean annual precipitation intensity in mm/d (NOAA data interpolated & converted to 1-km cells)
atemp	Mean annual 1961–90 temperature in °C (NOAA data interpolated & converted to 1-km cells)
awet	Mean annual number of consecutive wet days (NOAA data interpolated & converted to 1-km cells)
rfact	Mean annual 1971–2000 rainfall erosivity (R-factor for USLE) (NOAA station data analyzed & interpolated to 1-km cells)
Hydrologic properties	
contact	Mean subsurface contact time in days (estimated by means of TOPMODEL [Wolock, 1993] hydrologic model)
perdun	Percent of basin streamflow contributed by Dunne overland flow (estimated by means of TOPMODEL hydrologic model)
perhor	Percent of basin streamflow contributed by Horton overland flow (estimated by means of TOPMODEL hydrologic model)
pet	Mean potential evapotranspiration in cm (estimated using temperature data derived from the Parameter-Elevation Regressions on Independent Slopes Model (PRISM; Daly and others, 1997) and the Hamon PET equation)
roff	Mean annual 1951–80 runoff in cm/yr (USGS data interpolated & converted to 1-km cells)

Values for characteristics of very large watersheds included in the model development data set (for example, the Mississippi, the Ohio, and the Missouri Rivers) represent a broad average of conditions in the watershed. In the most extreme case, values of parameters for the watershed represented by the station about 200 miles from the mouth of the Mississippi River (St. Francisville, La.) are averages of conditions in diverse regions stretching from the Rocky Mountains to the Appalachian Mountains. Thus, differences in the watershed characteristics for the very large watersheds are more generalized than for smaller watersheds, for which the watershed characteristics are more representative of the watershed as a whole. Despite the averaging effect for the very large watersheds, however, variability in the values of specific characteristics among these watersheds was still substantial, indicating that the variables were contributing useful information about these watersheds in the regression analysis. The potential for added uncertainty caused by inclusion of the very large watersheds was regarded as less important than the benefit of inclusion of a wide range of watershed sizes in the regression analysis.

The representativeness of the values used for watershed characteristics of large watersheds also is affected by differences in discharge from major tributaries in the watershed. For example, water flowing in the lower Mississippi River comes from three main sources: the Missouri River, the Ohio River, and the upper Mississippi River (the main stem river upstream of confluences with the Ohio and the Missouri Rivers). Average annual discharge from the Ohio River accounted for 57 percent of the discharge in the lower Mississippi River from 1980 to 1996, with the upper Mississippi contributing about 25 percent and the Missouri River about 18 percent (Coupe and Goolsby, 1999). This implies that land use and watershed characteristics in the Ohio River watershed may influence water quality in the lower Mississippi River more than those of the upper Mississippi River or Missouri River watershed. However, this is not reflected in the values of watershed parameters used for the St. Francisville, La., sampling station. All parts of the watershed are weighted equally in the calculation of the values of watershed parameters used for a sampling station. This problem exists to some extent for all of the watersheds used in the regression analysis, but may be more important for the larger watersheds where watershed characteristics can vary considerably among different areas of the watershed. The Mississippi River watershed is the extreme example of this problem, as many of the watershed parameters listed in table 3 exhibit an east-to-west gradient, and the eastern and western parts of the watershed are separately drained by large tributaries with substantially different discharges. No attempt was made in this study to address this problem, in part because detailed data on discharges of tributaries are not available for all watersheds and because discharge data may not be available at all for unmonitored streams and tributaries to which the models may be applied.

Statistical Analysis

Statistical methods include those for estimating annual percentile and mean pesticide concentrations from the sample data, estimating the coefficients of the regression equations, and selecting and transforming variables as needed to be included in the regression models.

Calculation of Annual Percentile and Mean Concentrations

Annual percentiles and mean concentrations of atrazine were computed from the time series of data collected at individual sampling stations. An example time series is shown in figure 5. The number of years of sampling data available from individual stations varied, but most stations had only 1 year of intensively collected data. To avoid giving undue weight to the small number of stations with more than 1 year of intensive data, a single year of data from each station was used for model development (table 2). This was generally the year with the most atrazine data. In most cases, the 1-year period did not correspond to a calendar year, but rather the 12-month period following the onset of sampling at the station.

In the WARP method, specific percentiles (5th, 10th, 15th, 25th, 50th, 75th, 85th, 90th, and 95th) of the annual frequency distribution, and the annual mean concentration, are determined for each stream. These concentrations were calculated by weighting each concentration according to the amount of time it was used to represent the atrazine concentration in the stream. Specifically, the weights were computed as the amount of time extending from one-half the time interval between a value and the preceding value and one-half the time interval extending from the value to the subsequent value, divided by the total time in 1 year. An example is shown in figure 6. The annual mean concentration is simply the sum of the sample weights times the sample concentrations. Conceptually, percentiles of the concentration distribution at a given station can be obtained by first ranking the concentrations measured during the 1-year period from low to high. The weights assigned to each concentration then are cumulatively summed. The *p*th percentile corresponds to the concentration for which the summed weight is equal to *p*. For example, the 95th percentile concentration corresponds to the concentration for which the summed weights equal 95 (fig. 7). Percentiles computed in this manner can be thought of as representing the percentage of time during the year in which the concentration in the stream is less than the corresponding concentration. For example, a 95th percentile concentration of 0.4 µg/L implies that concentrations in the stream were less than 0.4 µg/L for 95 percent of the year, or 347 days.

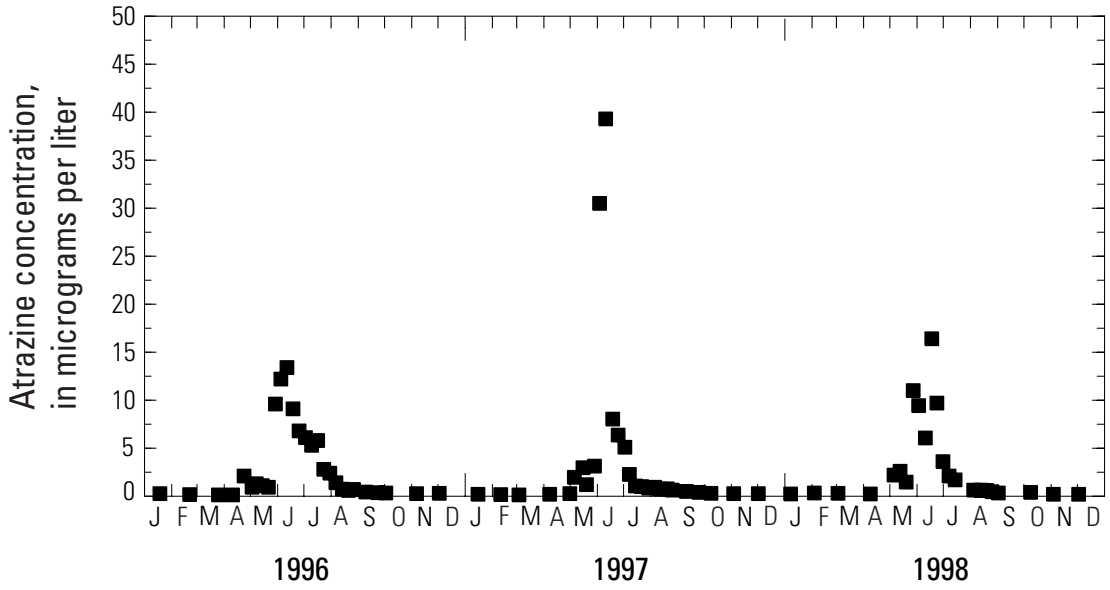


Figure 5. Atrazine concentrations in the White River at Hazleton, Indiana, from January 1996 through December 1998.

The Statistical Analysis System (SAS) Proc Univariate procedure was used to compute the means and percentiles from the weighted values (SAS Institute, 1999). The presence of censored observations (concentrations reported as less than a specified value by the laboratory) complicated the computation of the mean and percentile concentrations. Most of the censoring thresholds were near the long-term reporting limit for the analytical method. However, owing to matrix interferences, sample volume differences, refinement of the analytical methods used over time, or other factors, some censoring thresholds were higher than the long-term reporting level, sometimes substantially so.

For computing annual mean concentrations, if less than 10 percent of the weighted data for a station and sampling year were censored, censored values were replaced by one-half the censoring threshold reported by the laboratory (U.S. Environmental Protection Agency, 2000). If more than 10 percent of the weighted data were censored, and there were 20 or more observations, at least 10 of which were uncensored, and at least 33 percent of the sample weights were represented by uncensored observations, then the log regression (LR) method (Gilliom and Helsel, 1986; Helsel and Gilliom, 1986), implemented as a SAS macro, was used to approximate the mean concentration. Otherwise, the mean returned by the Univariate procedure was considered censored. (For example, the mean of 3, 4, 5, and <4 would be considered to be <4.)

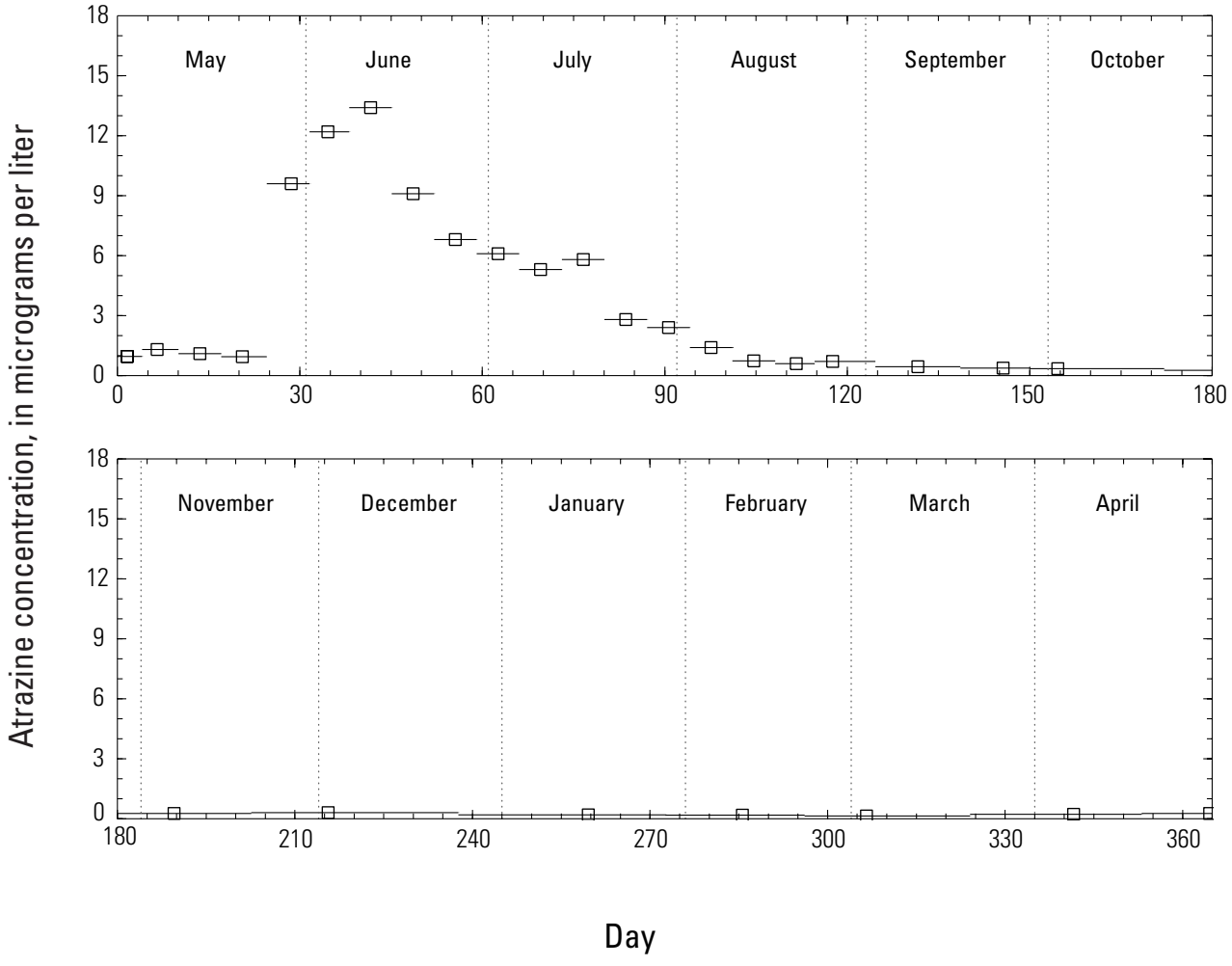


Figure 6. The period of time represented by each sample used in computing time-weighted percentiles of atrazine, White River at Hazleton, Indiana, for the 12-month period beginning May 1, 1996.

Period of time is shown by horizontal lines. Sample weights are equal to this period divided by 365.

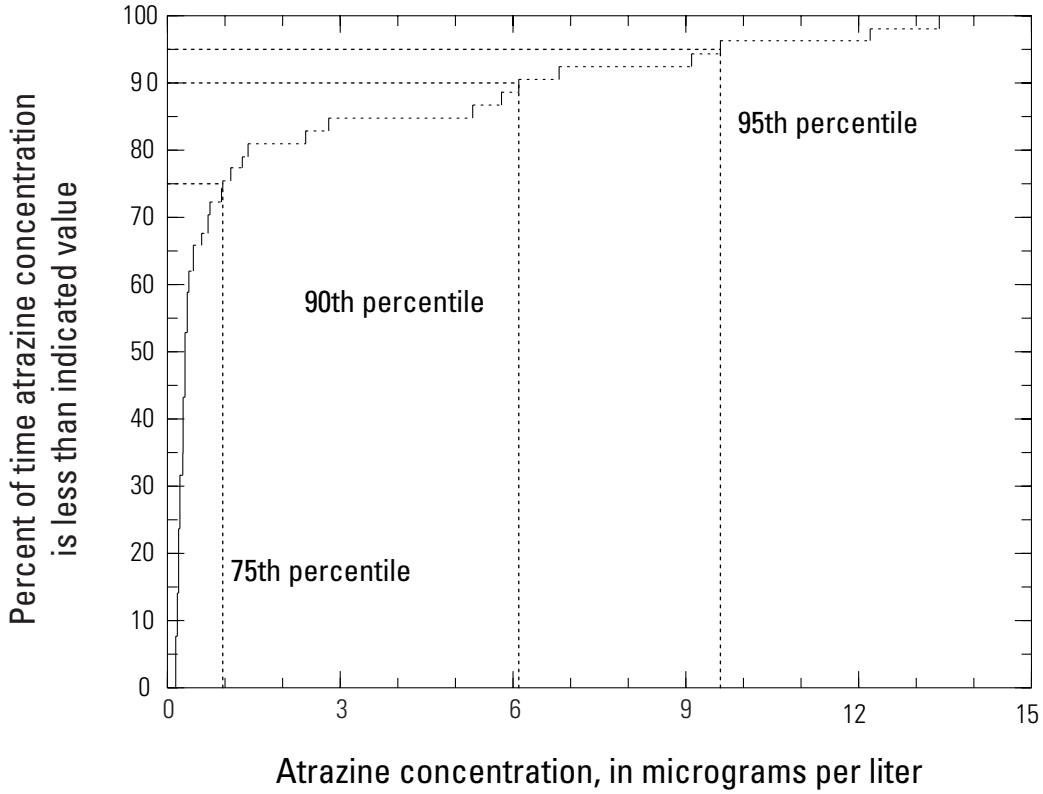


Figure 7. Example computation of annual time-weighted percentile concentrations of atrazine, White River at Hazleton, Indiana, for the 12-month period beginning May 1, 1996.

Each solid vertical line corresponds to a sample concentration. The length of the vertical lines corresponds to the sample weight.

For computing percentiles, if more than *p* percent of the weighted sample data were censored, and at least 20 observations in the sample with at least 10 uncensored observations and at least 33 percent of the sample weights were represented by uncensored observations, then the LR procedure was used to approximate the *p*th percentile. Otherwise, the percentile computed by the Univariate procedure was used and considered censored at that value.

Samples associated with the highest 0.5 percent of the censoring thresholds were deleted from the analysis so that an unusually high censoring threshold for one sample did not obscure detected concentrations in other samples. Any annual time-weighted percentile or mean concentration estimated as less than 0.001 µg/L was considered censored at the 0.001 µg/L level.

At some sampling stations, atrazine concentrations only approximately conform to the assumptions of the LR method. However, the criteria set for use of the LR method limited its use to approximating only the mean and lower percentile concentrations. The method was not used to estimate extremely high (greater than the 95th) or low (less than the 5th) percentiles where departures from the assumptions would be most critical. The advantage of having fewer censored observations in the data set available for model development was deemed to offset potential errors in approximated concentrations due to deviations from the distributional assumptions required by the LR method.

The time-weighting procedure was used to estimate annual percentile and mean pesticide concentrations because it minimized the need for assumptions about the shape of the probability distribution of the data. No assumptions were required for computing the upper percentile concentrations. Alternative approaches, such as fitting a parametric probability distribution to the sample data (Myers and others, 2000), require distributional assumptions that may not be reasonable at a given sampling station. However, the approach used in this study constrained the number of percentiles that could be estimated at the extremes of the distribution. The limited sample sizes available at most stations did not support estimation of percentiles lower than the 5th or higher than the 95th.

The accuracy of the annual percentile and mean concentrations computed using these methods is a function of the number of samples collected and the time of sample collection. Uncertainty in estimated percentiles generally is higher for stations where relatively few samples were collected, although the uncertainty also is a function of stream size and hydrologic parameters of the watershed. For example, uncertainty generally is higher for small streams where discharge and pesticide concentrations can change rapidly in response to rainfall in the watershed, and the samples may not reflect the true variation in concentrations in the stream. Samples collected during the period when atrazine runoff is highest contribute more to accuracy of the percentiles than samples collected at other times. Simulation studies suggest that upper percentiles of the annual distribution of concentrations, in particular the 95th percentile concentration, obtained from the smaller stations used for this study may be underestimated (Charles Crawford, U.S. Geological Survey, unpub. data, 2002).

Regression Methods

Many of the annual percentile and mean concentrations computed for this study were less than a censoring threshold. In the context of an explanatory model, a censored observation is one in which the value of the response variable was not observable. Conventional least-squares methods for estimating parameters of this model, using either the entire sample or the subsample of complete observations, yield biased and inconsistent estimates (Judge and others, 1985, p. 780). One way of expressing the regression model in the presence of censored observations is

$$Y_i = \begin{cases} f(X_i) + \varepsilon_i & \text{if } Y_i > CT_i \\ CT_i & \text{otherwise} \end{cases} \quad (1)$$

where

Y_i is the response variable,
$f(X_i)$ is a function of one or more explanatory variables,
CT_i is the censoring threshold for the ith observation, and
ε_i is the residual error.

This model is often called the tobit model (Judge and others, 1985, p. 780) after James Tobin who first explored the problem (Tobin, 1958).

A number of alternative methods for obtaining parameter estimates of models when the available data contain censored data have been proposed (Tobin, 1958; Miller, 1976; Buckley and James, 1979; Miller and Halpern, 1982; Amemiya, 1984; Powell, 1984; Chatterjee and McLeish, 1986; Duncan, 1986; Horowitz, 1986; Schneider and Weissfeld, 1986). When the regression residual errors are independent, identically and normally distributed, with mean zero and variance σ_ε^2, the maximum likelihood method can be used to obtain parameter estimates of a censored linear model (Maddala, 1983). Maximum likelihood estimates are obtained by minimizing ln Lk, the log-likelihood function (Amemiya, 1984; Haas and Jacangelo, 1993):

$$\ln Lk = -n \ln \sigma_\varepsilon - \frac{1}{2} \sum_{i=1}^{n} \left(\frac{(C_i - \hat{C}_i)}{\sigma_\varepsilon} \right)^2 + \sum_{i=n+1}^{n+m} \ln \Phi \left(\frac{CT_i - \hat{C}_i}{\sigma_\varepsilon} \right)^2 \quad (2)$$

where

C_i is the ith observed concentration,
$\hat{C}_i$ is the concentration estimated from the regression model for the ith observation,
ln is the natural logarithm,
σ_ε is the standard deviation of the residual error,
n is sample data based on observable (uncensored) concentrations,
m is sample data based on concentrations less than a censoring threshold (censored), and
Φ is the cumulative normal integral function.

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{(x^2)}{2}\right) dx \quad (3)$$

Equation 2 cannot be solved analytically, and numerical methods must be used to obtain a solution. Maximum likelihood methods implemented in the survreg procedure (Therneau, 1999) in the statistical analysis program S-PLUS (Insightful Corporation, 1999) were used to estimate the parameters of the regression models.

Measures of goodness of fit, such as the standard deviation of the residual error (commonly referred to as the root mean square error in conventional regression analysis) or the coefficient of multiple determination (R^2), used for conventional least squares regression analysis, cannot be computed for the tobit regression model. The standard deviation of the residual error (σ_ε) is alternatively referred to as the scale parameter in maximum likelihood estimation. Estimates of the scale parameter from the maximum likelihood procedure provide only asymptotically unbiased estimates of the standard deviation of the residual error when estimated from sample data (Aitkin, 1981). These estimates, on average, underestimate the true standard deviation. The bias is a function of the sample size and the degree of censoring. In this report, biased estimates of the standard deviation of the residual error are referred to as scale in figures and tables. Several pseudo R^2 (pR^2) measures suitable for use with the tobit regression model have been proposed in the literature as alternatives to R^2. For this study, pR^2 was calculated using the method of Laitila (1993). As with conventional R^2, the pR^2 computed by the method of Laitila (1993) ranges from 0 to 1 and is an estimate of the proportion of the variation in the response variable explained by the regression model (0 indicates no variation is explained; 1 indicates all variation is explained).

Transformation of Response and Explanatory Variables

The maximum likelihood methods used for estimating the parameters of the regression models require several assumptions. The relation between the variables must be linear in the parameters, and the residual error must be identically and normally distributed. Departures from these assumptions can result in estimates of model coefficients that are considerably in error. Nonidentically distributed residual error (inhomogeneity of residual variance across observations) can seriously bias the estimates of model coefficients obtained from maximum likelihood methods (Greene, 2000, p. 912). Residual error not normally distributed can result in inconsistent estimates of the model coefficients (that is, as the sample size tends to infinity, the estimated coefficients do not converge to the true coefficients). One means of addressing departures from model assumptions is through transformations of the response or explanatory variables, or both (Neter and others, 1985, p. 132).

Various transformations were considered to minimize departures from the assumptions of the maximum likelihood methods used. For example, the relation between agricultural use intensity and atrazine concentration was observed to be nonlinear (fig. 8). Further, the variance of atrazine concentrations tended to increase with increasing use, violating the assumption of identically distributed residual error. To address these problems, several transformations of both the explanatory and response variables were tried. Examples of transformations tried for atrazine concentration and use intensity are shown in figures 9 through 11. The relation between the logarithm of atrazine concentration and the fourth root of use intensity was approximately linear and has approximately identically distributed residual errors (fig. 11). These transformations of concentration and agricultural use intensity produced similar results (a linear relation between the transformed variables and approximately identically distributed residual errors) throughout the range of concentration percentiles modeled (fig. 12). The logarithm of concentration was used as the response variable in all subsequent work. In the development of the regression models, the logarithmic, square-root, and fourth-root transformations of agricultural use intensity were considered as explanatory variables as well as the untransformed value. For all other explanatory variables, the logarithmic, square or quadratic, and square-root transformations of each variable were considered as well as the untransformed variable.

Estimates of concentration obtained by retransformation of values predicted by the logarithmic model are of the median concentration expected for a given set of explanatory variables and not the mean concentration. For this study, predicted concentrations were not adjusted for transformation bias (Bradu and Mundlak, 1970; Duan, 1983) because estimates of median concentrations were considered more appropriate for the purposes of this study.

Selection of Explanatory Variables

Initial model development work was done using the 95th percentile atrazine concentrations because only two stations in the model development data set had censored 95th percentile values (values below the method detection limit). Because of this low level of censoring, variable selection and the fitting of initial regression models for the 95th percentile could be done using ordinary least squares (OLS) methods, with the two censored values assigned the value of the detection limit (0.001 µg/L). Final models for all percentiles and for the annual mean were fit using tobit regression because of the presence of censored concentration data.

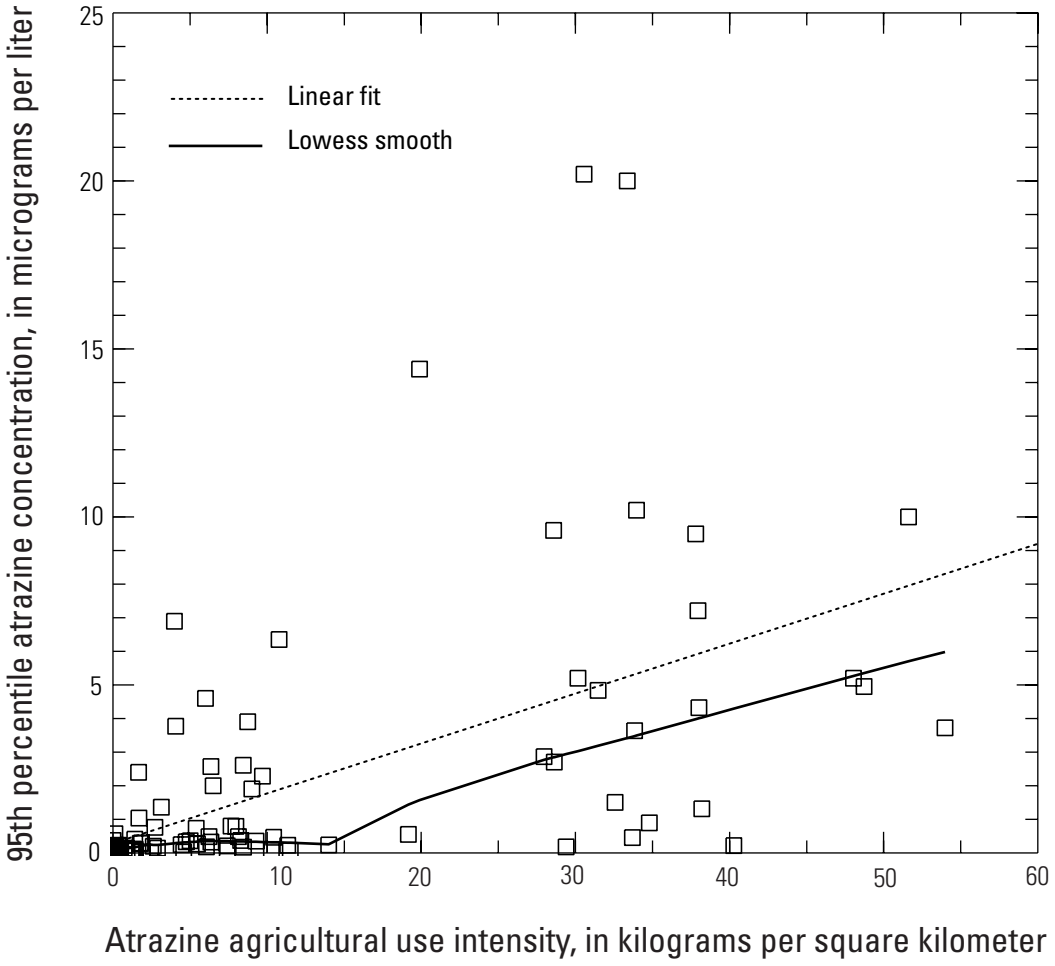


Figure 8. Relation of the 95th percentile atrazine concentration and atrazine agricultural use intensity at selected sampling stations in the United States.

A solid symbol indicates a concentration that is less than a censoring threshold and, therefore, is less than the indicated value.

The regression models were built using a stepwise procedure. The initial model included just the use-intensity variable, as this variable was expected to be the most important in terms of explaining the variance in concentrations among the stations. Thus, the form of the initial model was

$$\log_{10}\ (95\text{th percentile concentration}) = \beta_0 + \beta_1 * \text{use intensity}^{\frac{1}{4}} + \varepsilon' \quad (4)$$

where

β_0 and β_1 are regression coefficients and

ε' is the residual error.

Results for the use-intensity model are shown in figure 13, in which observed 95th percentile concentrations at the model development stations are compared to values predicted by the regression model. This model explains about 59 percent of the variability in the $\log_{10}$ (95th percentile) concentrations at the model development stations ($R^2 = 0.59$).

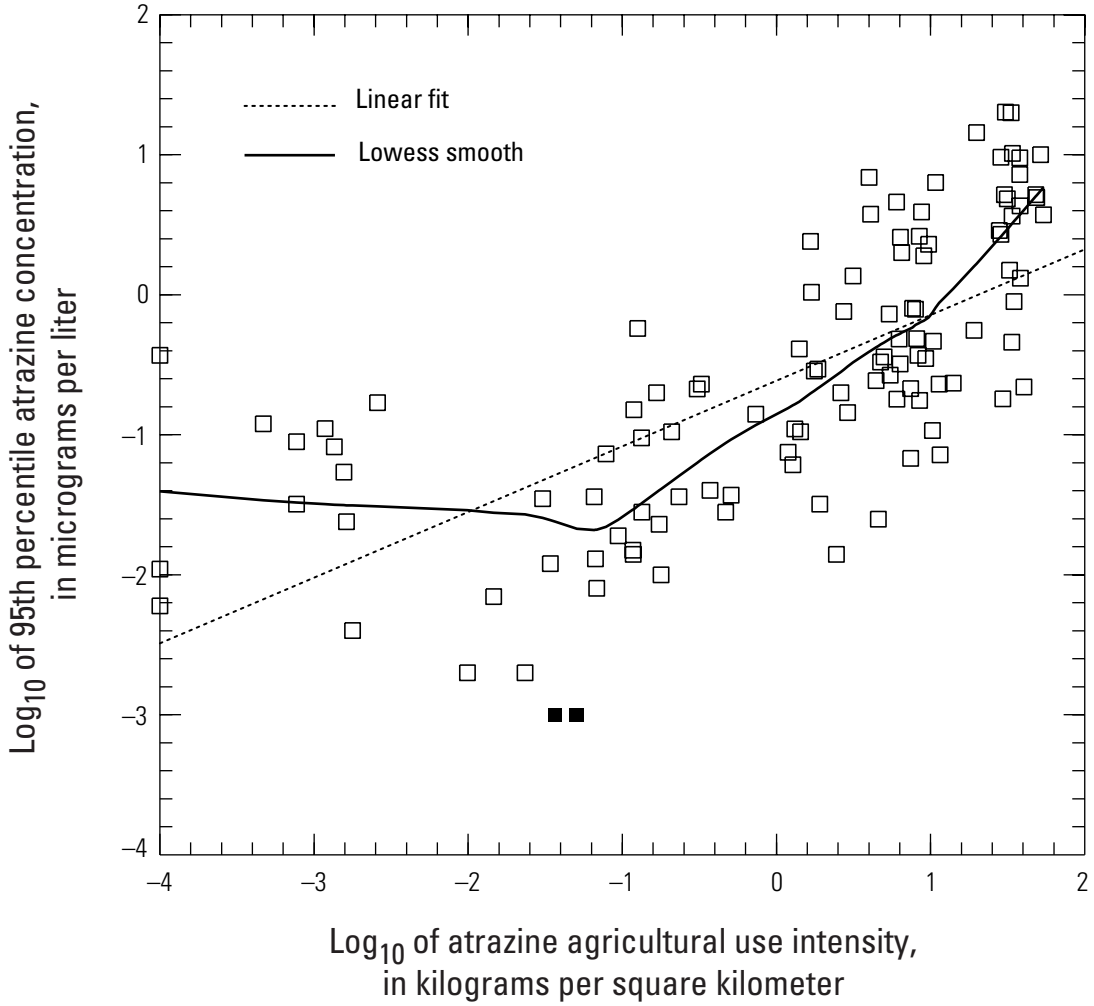


Figure 9. Relation of $\log_{10}$ of the 95th percentile atrazine concentration and $\log_{10}$ of atrazine agricultural use intensity at selected sampling stations in the United States.

A solid symbol indicates a concentration that is less than a censoring threshold and, therefore, is less than the indicated value.

The next step in the stepwise procedure identified the candidate predictor variables (or their transformations) most highly correlated with the residuals from the use-intensity model. The residuals were calculated as

$$\text{residual} = \log_{10}(\text{observed concentration}) - \log_{10}(\text{predicted concentration})$$

Residuals for censored observed values were calculated by subtracting the log of the predicted value from the log of the censoring threshold. A significant correlation between residuals from the use-intensity model and values of another explanatory variable implies that adding the variable to the regression model would explain additional variability. Variables were eliminated from further consideration when there was no evidence of a significant correlation with the residuals from the use-intensity model (correlation coefficient (r) less than 0.03). Additional variables were eliminated from consideration when the relation of the variable (or the transformed variable) with the residuals was not approximately linear or did not have relatively constant variance across the range of values of the variable. This was done subjectively by examining plots of the relations between candidate variables and residuals from the use-intensity model. Examples of the relations between the residuals and several candidate variables are shown in figure 14.

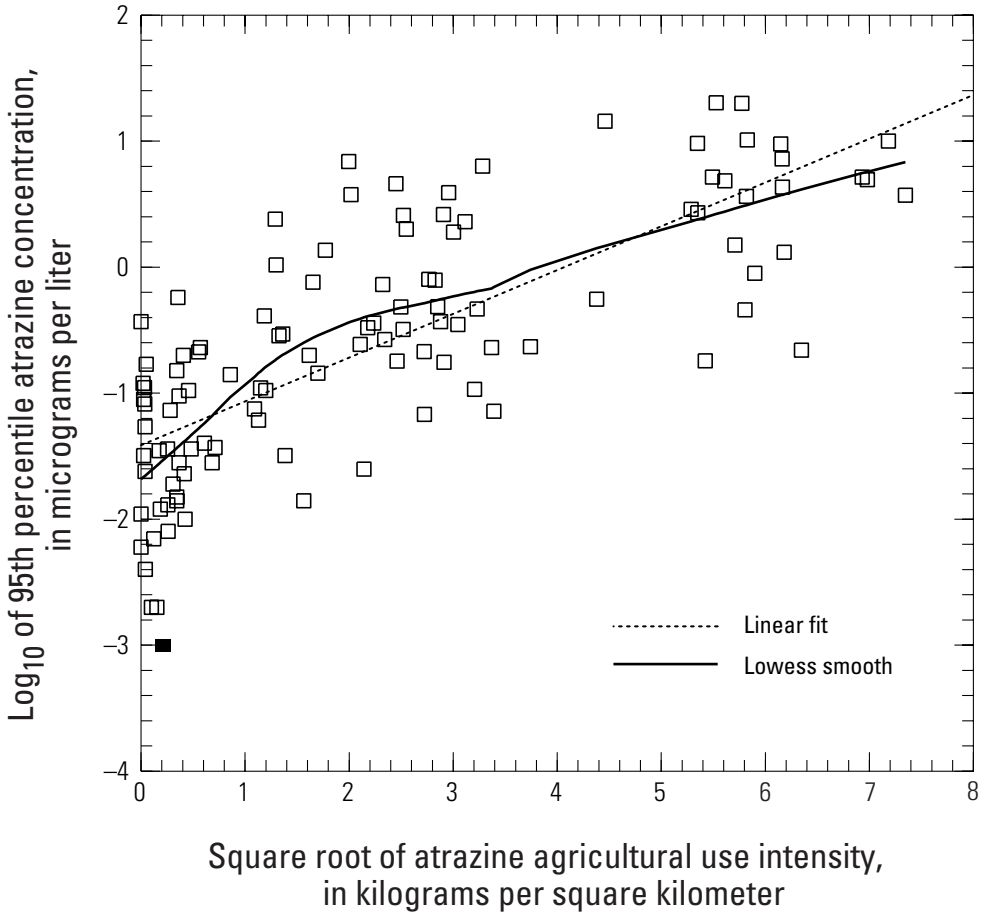


Figure 10. Relation of $\log_{10}$ of the 95th percentile atrazine concentration and the square root of atrazine agricultural use intensity at selected sampling stations in the United States.

A solid symbol indicates a concentration less than a censoring threshold and, therefore, less than the indicated value.

The remaining variables were candidates for addition to the model. A subsampling procedure was used to select which candidate variable to next add to the model. The model development data set of 112 stations was randomly split into two equal parts. Models were fit for the 95th, 75th, and 50th percentiles, and for the annual mean concentration, using the fourth root of use intensity and all of the candidate variables. OLS was used to fit these models, as censoring was relatively low (2 to 14 percent) for all four of these concentration levels. Mallow's Cp, a statistic commonly used in variable selection for regression models, was used to select which of the candidate variables to retain in a two-variable model (along with use intensity) for each of the four concentration levels. Mallow's Cp measures how well the model fits the data, but also incorporates the concept of model parsimony (there is a penalty for added variables). These models, derived from half of the stations, then were used to predict concentrations for the remaining stations (the half not used to derive the models). The process was repeated numerous times, with models being derived from different sets of randomly selected stations. The two-variable model that had the lowest average prediction error for the four concentration levels in the repeated trials was selected for the next stage of model development.

The variable selection procedure was repeated, adding one variable at a time to the model, until no appreciable improvement was seen when a variable was added. This somewhat subjective process balances the desire to have the model explain the maximum amount of the variance while avoiding an overly complex model. Results for three of the models (50th, 75th, and 95th percentile models) derived using this process and the improvement at each step, as shown by decreasing scale and increasing pR^2 values, are given in table 4. Virtually no improvement was seen in going beyond a model with five explanatory variables.

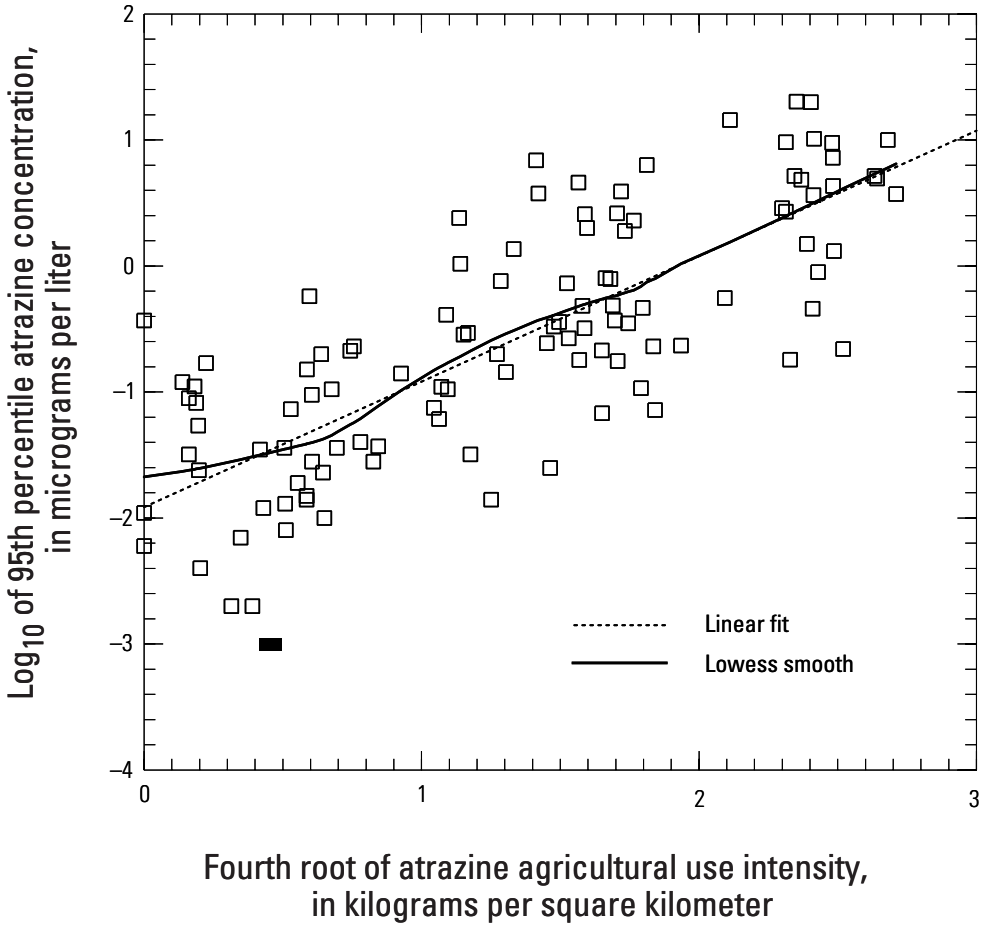


Figure 11. Relation of $\log_{10}$ of the 95th percentile atrazine concentration and the fourth root of atrazine agricultural use intensity at selected sampling stations in the United States.

A solid symbol indicates a concentration less than a censoring threshold and, therefore, less than the indicated value.

At each stage of model development, the explanatory variables were subjectively evaluated for reasonableness (for example, the models predict increasing concentrations with increasing atrazine use) and their overall contribution to explaining the variation in the nine percentiles and mean concentration. Slightly more consideration was given to variables selected for the higher percentiles than for the lower percentiles because the percentage of censored data was lower. The validity of the assumptions of the maximum likelihood methods used to fit the model parameters also was assessed.

The same explanatory variables were used in all of the final atrazine models. Because coefficients were estimated independently for each of the 10 models, no constraint prevents the estimate of a low percentile from exceeding the estimate for a higher percentile. Retaining the same variables in all models reduced how often this occurred. Statistical methods one might use for estimating correlated models, such as seemingly unrelated regressions or simultaneous equations estimation, are not available for regressions involving censored data. In addition, using the same explanatory variables in all models simplifies the data requirements for application of the models.

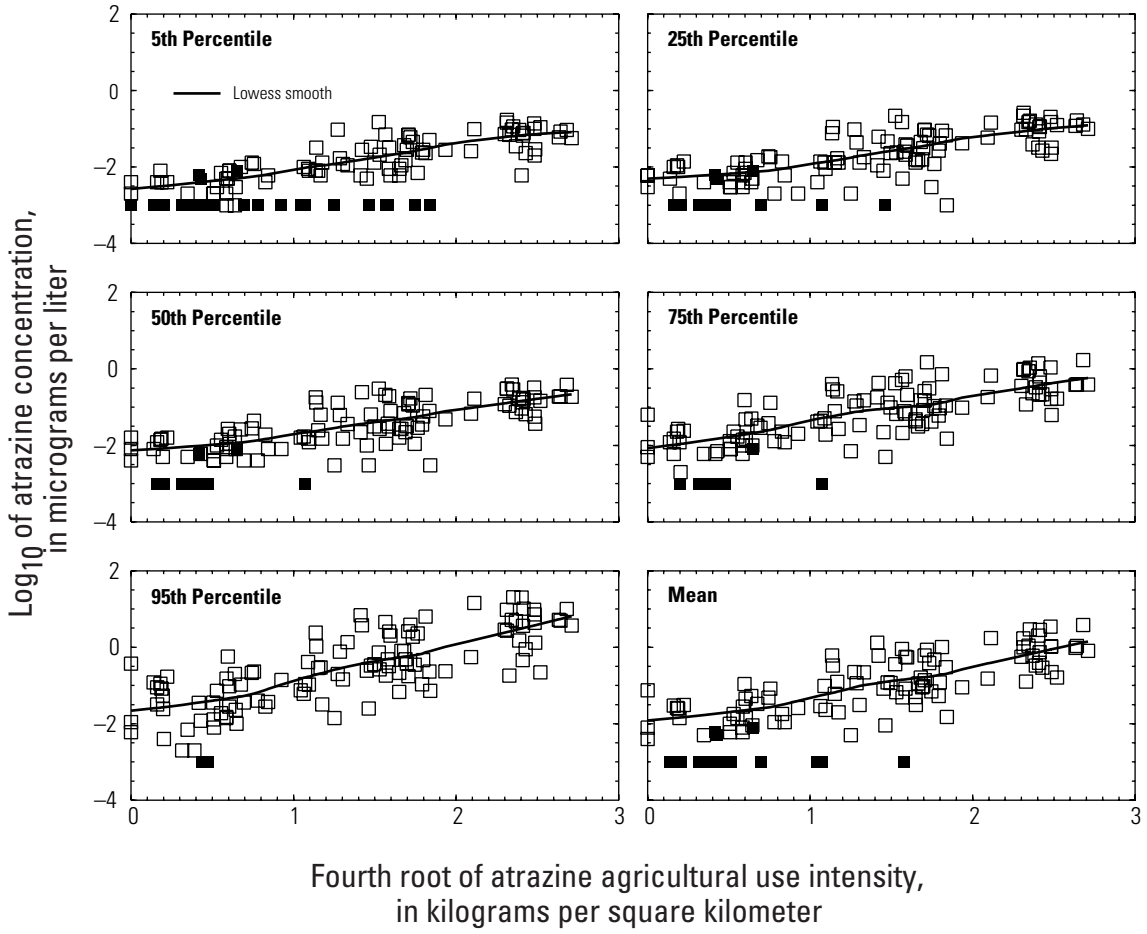


Figure 12. Relation of the $\log_{10}$ of atrazine concentration and the fourth root of atrazine agricultural use intensity at selected sampling stations in the United States.

A solid symbol indicates a concentration less than a censoring threshold and, therefore, less than the indicated value.

Because of the already large number of potential explanatory variables, interactions between variables were not considered as potential explanatory variables. Doing so would have resulted in thousands of additional potential explanatory variables. However, the significance of interactions between variables retained in a final model was evaluated and interaction terms were included when significant.

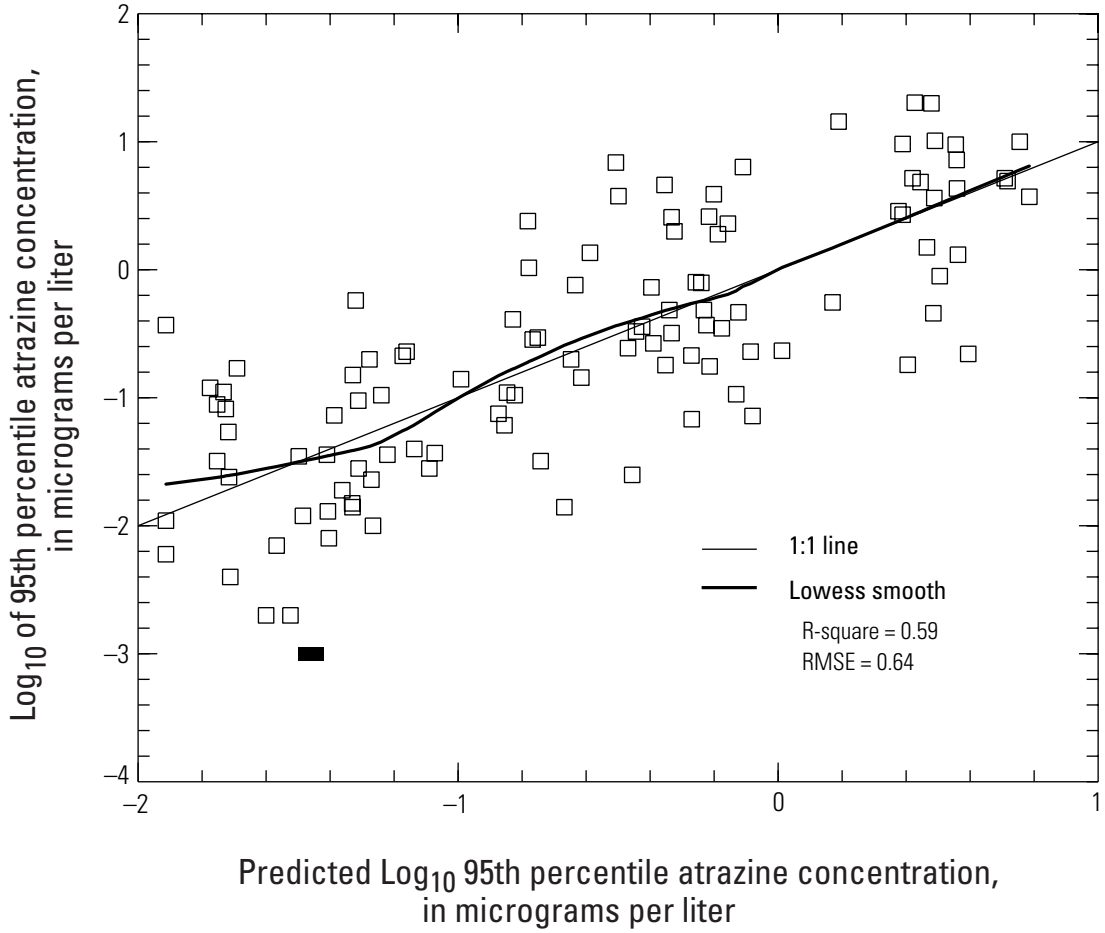


Figure 13. Comparison of 95th percentile atrazine concentrations with predicted 95th percentile atrazine concentrations for the 112 model development stations from a model using only use-intensity as an explanatory variable.

Results are from the ordinary least squares regression. Filled symbols indicate censored values. RMSE, root mean square error.

Analysis of Model Fit

Diagnostics for censored regression (Escobar and Meeker, 1992) available in the survreg procedure in the S-PLUS program (Insightful Corporation, 1999) were used to assess influential observations and to aid in variable selection. In particular, the deviance residuals described by Escobar and Meeker (1992) were used to assess the appropriateness of transformations used for explanatory variables. Variance inflation factors were used for detecting the presence of multicollinearity among explanatory variables (Neter and others, 1985, p. 391). Box and whisker plots (Tukey, 1977, p. 39) were used to help assess model uncertainty. These plots, also known as boxplots, summarize a group of data by showing a measure of central tendency (the median), the variation (interquartile range), the range (shown by the whiskers), and extreme values (shown by individual points). Boxplots were used for displaying the distribution of model residuals (the logarithm of the observed concentration minus the logarithm of the predicted concentration) and comparing residuals among groups of data (for example, different regions of the country). For the purposes of creating boxplots, the residual for a prediction that had a censored observed concentration was considered zero when the predicted concentration was less than the censoring threshold (because the predicted value was consistent with the observed value). The residual for a censored observation was computed as the difference between the logarithm of one-half the censoring threshold and the logarithm of the predicted concentration when the predicted value was greater than the censoring threshold.

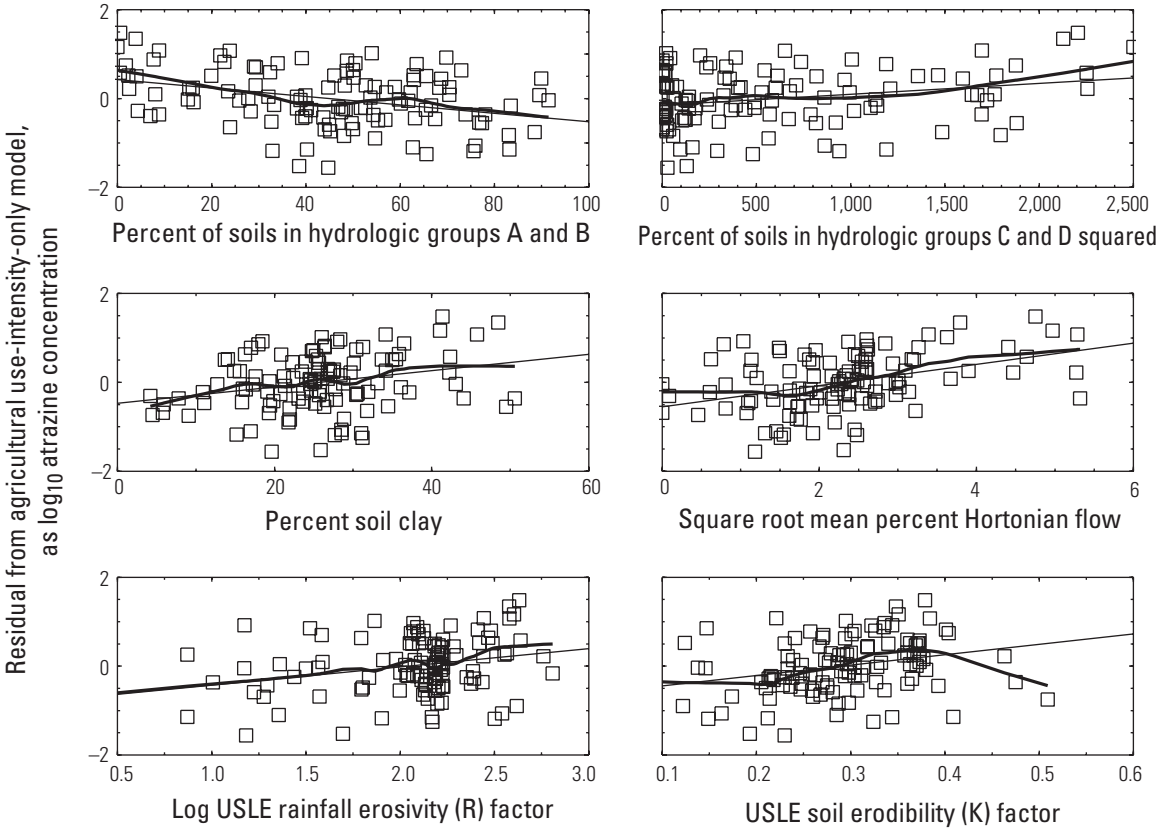


Figure 14. Examples of relations between residuals from the use-intensity-only model and values of candidate explanatory variables for the 112 model development stations.

The thin line is an ordinary least squares regression line. The thick line is a lowess smooth. USLE, Universal Soil Loss Equation.

Estimation of Confidence and Prediction Intervals

Confidence and prediction intervals were approximated using normal theory and the t-distribution. That is, methods for ordinary least square regression were used, which are only approximate when applied to censored data. The standard errors were estimated from the maximum likelihood scale parameter, using the adjustment suggested by Aitkin (1981).

ATRAZINE MODELS

The models selected for all nine percentiles and for the annual mean concentration have the form:

$$\log_{10}(\text{concentration}) = f\left[(\text{use intensity})^{\frac{1}{4}}, \log_{10}(\text{R-factor}), \text{K-factor}, (\text{watershed area})^{\frac{1}{2}}, \text{Dunne overland flow}\right] \quad (5)$$

where

use intensity	is annual agricultural atrazine use in the watershed (kg) / watershed area (km^2),
R-factor	is rainfall erosivity factor from Universal Soil Loss Equation (USLE),
K-factor	is soil erodibility factor from USLE,
watershed area	is area of drainage basin (km^2), and
Dunne overland flow	is percentage of total stream flow derived from surface runoff caused by precipitation on saturated soil.

Table 4. Example results from stepwise model development for the 50th, 75th, and 95th percentile models for atrazine.

[Scale and pseudo-R^2 values are shown for each of the trial models shown on the left. pR^2, pseudo R-squared; fuseint, fourth root of atrazine use intensity in the watershed (see table 5); lrfact, $\log_{10}$(R-factor); kfact, K-factor; sdarea, square root of watershed area; perdun, percent of streamflow due to Dunne overland flow; adry, mean number of consecutive dry days; hgab, sum of percentages of soil hydrologic groups A and B; $hgcd^2$, squared sum of percentages of soil hydrologic groups C and D; appt, mean annual precipitation; latemp, $\log_{10}$(mean annual temperature); stile, square root of percentage of watershed with subsurface tile drainage; awet, mean number of consecutive wet days; orgm, mean percentage of soil organic matter]

Model	50th Percentile		75th Percentile		95th Percentile	
	scale	pR^2	scale	pR^2	scale	pR^2
fuseint	0.485	0.56	0.544	0.56	0.646	0.59
fuseint+lrfact	0.474	0.58	0.519	0.60	0.613	0.63
fuseint+lrfact+kfact	0.457	0.61	0.493	0.64	0.561	0.69
fuseint+lrfact+kfact+sdarea	0.433	0.65	0.459	0.68	0.536	0.72
fuseint+lrfact+kfact+sdarea+perdun	0.413	0.68	0.444	0.70	0.522	0.73
fuseint+lrfact+kfact+sdarea+perdun+adry	0.407	0.69	0.433	0.72	0.516	0.74
fuseint+lrfact+kfact+sdarea+perdun+hgab	0.413	0.68	0.443	0.70	0.515	0.74
fuseint+lrfact+kfact+sdarea+perdun+$hgcd^2$	0.413	0.68	0.442	0.71	0.515	0.74
fuseint+lrfact+kfact+sdarea+perdun+appt	0.412	0.68	0.442	0.71	0.521	0.73
fuseint+lrfact+kfact+sdarea+perdun+latemp	0.411	0.68	0.441	0.71	0.522	0.73
fuseint+lrfact+kfact+sdarea+perdun+stile	0.412	0.68	0.440	0.71	0.509	0.74
fuseint+lrfact+kfact+sdarea+perdun+awet	0.412	0.68	0.443	0.70	0.522	0.73
fuseint+lrfact+kfact+sdarea+perdun+orgm	0.413	0.68	0.443	0.71	0.515	0.74

Statistics for all 10 models are given in table 5. Regression coefficients and statistics are all based on models fit using tobit regression. All five variables were significant ($p < 0.05$) in all models, except for watershed area in the models for the 5th and the 10th percentiles. Values of pR^2 ranged from 0.62 for the 5th percentile to 0.77 for the annual mean concentration, meaning that the models accounted for 62 to 77 percent of the variability in concentrations among the 112 stations used for model development.

Analysis of Significant Explanatory Variables

The primary purpose of the WARP models for atrazine is prediction of concentrations in unmonitored streams. The models were not developed to investigate processes that affect transport of atrazine from agricultural fields to streams. However, the explanatory (predictor) variables in the models should have a reasonable and understandable relation to atrazine concentrations in streams.

Atrazine use intensity in the watershed (as the fourth-root transform) was the most important explanatory variable in all of the regression models. Models using just $(\text{use-intensity})^{1/4}$ as a predictor variable account for 53 to 64 percent of the variability in concentrations among the 112 stations. The coefficient of the transformed use-intensity variable was positive in all models, consistent with the relation shown in figure 11.

Table 5. Statistics and coefficients for 10 atrazine models.

[Pseudo R-square, R-squared value for tobit regression; Scale, tobit regression analogue of the root mean squared error obtained from ordinary least squares regression; fuseint, fourth root of atrazine use intensity in the watershed; lrfact, $\log_{10}$(R-factor); kfact, K-factor; sdarea, square root of watershed area; perdun, percent of streamflow due to Dunne overland flow. <, less than]

Model	Regression coefficients (*p*-value)						Pseudo R-square	Scale	Percent censored observations
	Intercept	fuseint	lrfact	kfact	sdarea	perdun			
5th	– 4.04 (<0.001)	0.70 (<0.001)	0.44 (0.010)	1.55 (0.030)	0.00033 (0.125)	–0.15 (<0.001)	0.62	0.50	22
10th	– 4.20 (<0.001)	0.66 (<0.001)	0.53 (0.001)	1.96 (0.004)	0.00037 (0.080)	– 0.15 (<0.001)	0.63	0.49	20
15th	– 4.17 (<0.001)	0.62 (<0.001)	0.59 (<0.001)	1.95 (0.002)	0.00040 (0.036)	– 0.16 (<0.001)	0.65	0.45	14
25th	– 3.92 (<0.001)	0.60 (<0.001)	0.56 (<0.001)	1.54 (0.011)	0.00041 (0.027)	– 0.13 (<0.001)	0.64	0.44	12
50th	– 4.01 (<0.001)	0.57 (<0.001)	0.66 (<0.001)	1.88 (<0.001)	0.00046 (0.009)	– 0.11 (0.001)	0.68	0.41	9
75th	– 4.15 (<0.001)	0.60 (<0.001)	0.78 (<0.001)	2.38 (<0.001)	0.00064 (<0.001)	– 0.10 (0.008)	0.70	0.44	6
85th	– 4.44 (<0.001)	0.67 (<0.001)	0.87 (<0.001)	3.28 (<0.001)	0.00068 (0.002)	– 0.10 (0.021)	0.71	0.50	5
90th	– 4.47 (<0.001)	0.67 (<0.001)	0.95 (<0.001)	3.46 (<0.001)	0.00072 (<0.001)	– 0.10 (0.011)	0.73	0.50	3
95th	– 4.40 (<0.001)	0.73 (<0.001)	0.96 (<0.001)	3.69 (<0.001)	0.00058 (0.009)	– 0.10 (0.017)	0.73	0.52	2
mean	– 4.67 (<0.001)	0.75 (<0.001)	0.87 (<0.001)	3.28 (<0.001)	0.00047 (0.019)	– 0.11 (0.005)	0.77	0.46	14

The logarithm of the rainfall erosivity factor (R-factor) was a significant predictor in all models and had a positive coefficient. The R-factor is an index that characterizes the energy of storms in a specific area, averaged over a number of years (Brooks and others, 1991, p. 141–143). The R-factor is included in the USLE, which is used to predict soil losses due to erosion. Values of the R-factor for the 112 model development stations are highly correlated with values of precipitation intensity for these stations ($r = 0.94$). The positive coefficient in the regression models implies that more runoff of atrazine occurs in areas of high-energy rain storms.

The soil erodibility factor (K-factor) was also a significant predictor in all models, and was highly significant ($p < 0.001$) in models for the higher percentiles and annual mean. The coefficient was positive in all models (table 5). The K-factor is an index that indicates the susceptibility of soil to erosion (Brooks and others, 1991, p. 143–144). The K-factor for a specific area is dependent on a number of soil characteristics, including sand, silt, and organic matter content, permeability, and soil structure. The K-factor is included in the USLE. Values of the K-factor for the 112 model development stations are positively correlated with values for percentage silt ($r = 0.77$) and negatively correlated with values for percentage sand ($r = -0.70$) and permeability ($r = -0.62$). The positive coefficient for K-factor in the regression models implies that more runoff of atrazine occurs in areas where the soil is susceptible to erosion. This does not necessarily mean that atrazine is being transported on eroded soil particles, as atrazine generally is regarded as being transported in the dissolved phase in runoff water because of its relatively high water solubility (33 mg/L) and low soil sorption coefficient ($K_{oc} = 147$) (U.S. Department of Agriculture, 2001). High K-factor soils have properties that favor runoff of water—low permeability, low sand content, and high silt content. The positive relation between K-factor and atrazine concentrations in the streams may indicate that areas with high K-factor soils have a relatively high potential for the formation of runoff during precipitation, resulting in an increased potential for transport of atrazine.

Watershed area (as the square-root transform) was a significant predictor in all models except the 5th and 10th percentile models, and the coefficient was positive in all models (table 5). Watershed area is already incorporated into the use-intensity variable (use divided by watershed area), so the explanation of its significance in the regression models is not as straightforward as for the other variables. Larson and Gilliom (2001) also determined that watershed area was significant in preliminary WARP models for atrazine and other herbicides, and listed three factors that may contribute to the positive correlation between watershed area and atrazine concentrations. First, in large watersheds, pesticides are contributed from multiple tributaries. The timing of pesticide application and subsequent runoff can vary among the tributaries. The mixing of water from these tributaries can result in elevated concentrations that are sustained for a longer period of time than in the individual tributaries. Second, for some streams with large watersheds, pesticide use is concentrated in a relatively small farmed area of the watershed, with the remainder of the watershed largely untreated. The use-intensity value for these watersheds may be relatively low, despite high use in the farmed areas. It is likely that more water is contributed to the stream through runoff from the farmed areas than from the nonagricultural areas, especially in arid regions where crops are irrigated. For these cases, the watershed area variable may act as a correction factor, accounting for the high concentrations in these streams despite relatively low use-intensity values for the watersheds as a whole. Third, in some of the small streams, where concentrations of pesticides remain elevated for relatively short periods, the highest concentrations may not have been sampled. For these cases, the values computed for the higher concentration percentiles may be biased low, and do not truly reflect the concentrations that would be expected for a given use intensity. These unexpectedly low concentrations would correspond to small watershed areas, further strengthening the positive relation between watershed area and concentration reflected in the coefficients for the watershed area variable in the regression models.

Dunne overland flow was a significant predictor in all models and highly significant ($p < 0.001$) in all models except for the lower percentiles (table 5). The Dunne overland flow values are estimates of the proportion of the total stream flow contributed by overland flow resulting from precipitation falling on saturated soil (that is, precipitation falling on areas where the water table has risen to the land surface). The Dunne value was estimated for each watershed using TOPMODEL, a physically based watershed model (Wolock, 1993). In TOPMODEL, the total contribution of overland flow to streamflow is considered the sum of Dunne overland flow and Hortonian overland flow. Hortonian flow results from precipitation falling at a rate exceeding the infiltration rate of the soil. Hortonian overland flow was included as a potential predictor variable in the regression analysis along with the Dunne value, but Dunne overland flow was a significant predictor much more consistently. The coefficient of the Dunne variable was negative in all models, meaning that concentrations generally were lower in areas where Dunne overland flow was relatively high. Larson and Gilliom (2001) found the same relation in preliminary WARP models. The significance of the Dunne overland flow variable is somewhat unexpected because runoff resulting from Dunne overland flow constitutes a relatively minor proportion of total streamflow in all of the streams included in the model development data set. The maximum and mean values of the estimates of Dunne overland flow for the 112 streams are 7.7 and 2.4 percent of total streamflow, respectively. Larson and Gilliom (2001) hypothesized that Dunne overland flow acts as a source of dilution by relatively atrazine-free water. Areas of a watershed that contribute water to a stream through Dunne overland flow are likely to be lowland and wetland areas relatively close to the stream channel that become saturated at the land surface during wet periods. These areas are less likely to be used for growing crops and, thus, are less likely to have been treated with atrazine. It is also probable that such areas contribute a disproportionately high amount of water to the stream during rainfall, compared to other areas of the watershed, because rainfall on high Dunne areas is readily transformed into surface runoff.

Relation Between Predicted Concentrations and Atrazine Use Intensity

Both the response variable (atrazine concentration) and the use-intensity explanatory variable were transformed for use in the regression models. These transformations resulted in a relatively linear relation between the two variables and improved the behavior of residuals. Despite the use of the transformations, the relation between predicted atrazine concentrations (untransformed) and use-intensity values (untransformed) for a given station is virtually linear. This can be shown by obtaining model-predicted concentrations for a station using a simulated range of use-intensity values while holding the values of the remaining explanatory variables in the models constant (figs. 15–17). Results from application of four of the models to one of the model development stations (Sugar Creek at Milford, Ill.) are shown in figure 15. For each of these plots, 200 separate predictions of atrazine concentration were obtained using randomly generated values for use intensity that ranged from 0 to 80 kg/km^2. Values of the remaining explanatory variables in the models were held constant for all 200 model runs. Predicted concentrations are shown by the solid line in the plots. The diagonal dashed line is a linear reference line obtained from a linear regression of model predictions and use intensity. The two lines nearly coincide in all four plots. The very slight curvature at the low and high ends of the use-intensity range increases at use-intensity values well beyond the range shown, but the relation is essentially linear over the range of atrazine use intensity likely to occur in watersheds where atrazine is used in agricultural applications. The vertical dashed line shows the maximum use intensity among the 112 model development stations. Values of use intensity up to 80 kg/km^2 were used for these plots to demonstrate that the linear relation extends beyond the range of use intensity in the model development data. The observed atrazine concentration for the Sugar Creek station also is shown on each of the plots, plotted at the actual use intensity (use intensity at this station was the highest among the 112 model development stations). Results for the other six models were similar for this station. Similar results also were obtained for other model development stations. Results for the 95th percentile and 50th percentile models for four example stations in different regions of the United States are shown in figures 16 and 17, respectively.

Model Performance

Comparisons between predicted and observed concentrations are made frequently in the following discussion of model performance. Terms used for these comparisons are defined here for clarity. Predicted concentrations within a factor of 10, or "order of magnitude," of the observed values are between one-tenth and 10 times the observed value. More generally, concentrations within a factor of x are between (1/x)*(observed value) and (x)*(observed value). For example, for an observed concentration of 3 µg/L, predicted values between 0.3 and 30 µg/L would be within a factor of 10 of the observed value; predicted values between 0.6 and 15 µg/L would be within a factor of 5.

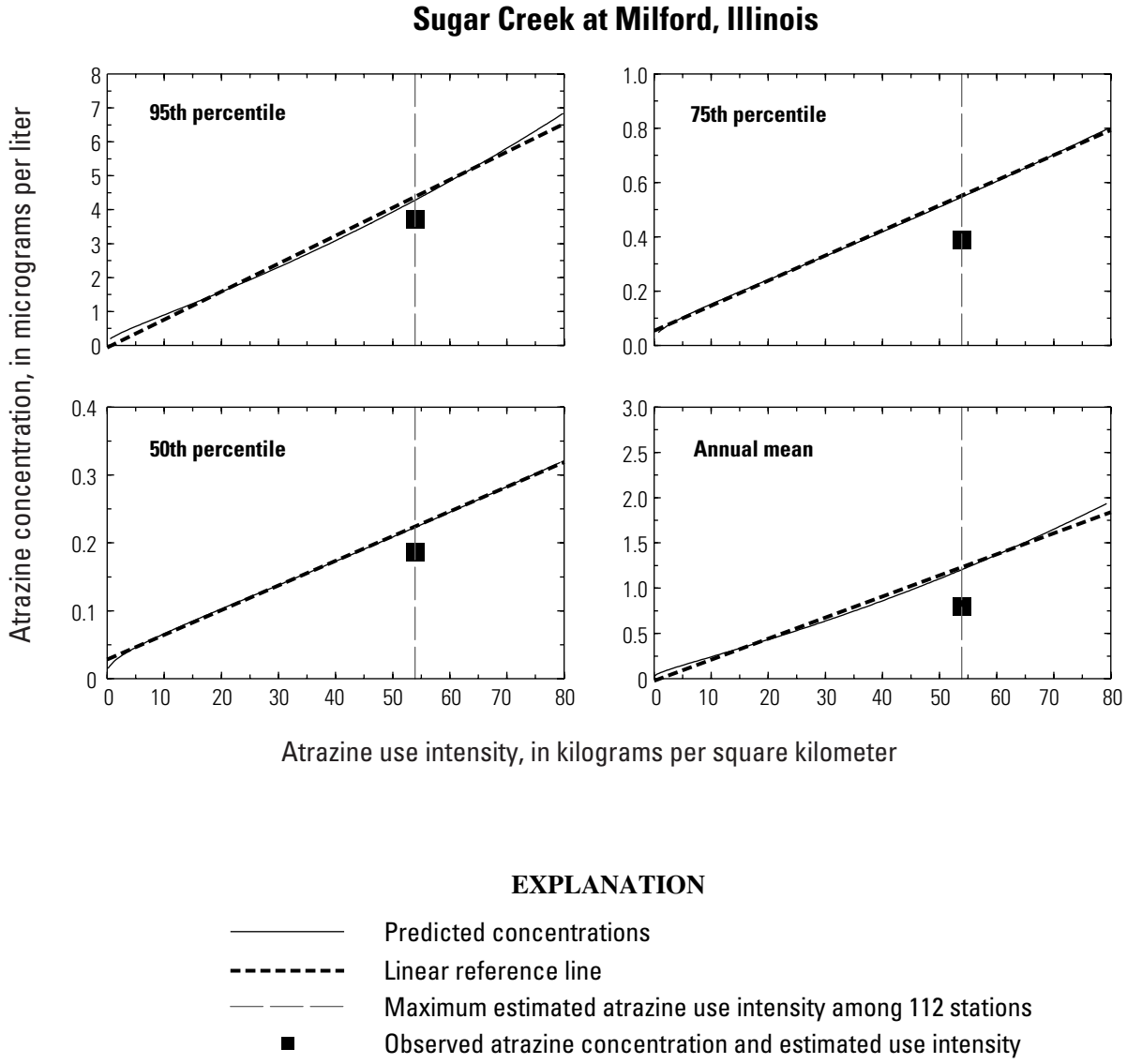


Figure 15. Relation of predicted atrazine concentrations and atrazine use intensity for Sugar Creek at Milford, Ill., for simulated atrazine use-intensity values ranging from 0 to 80 kilograms per square kilometer.

Model Development Stations

The models substantially improve predictive performance over a national average estimate or a model using only use-intensity as a predictor variable (fig. 18). Simply using the mean of the 95th percentile concentrations at the model development stations as an estimate results in errors of up to several orders of magnitude (fig.18*A*). The use-intensity-only model (fig. 18*B*) explains much more of the variance, and most predicted values are within an order of magnitude of the observed value. The full model for the 95th percentile (fig. 18*C*) explains an additional 14 percent of the variability, and 92 percent of predicted values are within a factor of 5 of the observed values at the model development stations. Examples of the fit of the regression models for additional percentiles are shown in figure 19. For all 10 of the atrazine models, 97 percent or more of the predicted values are within a factor of 10, and 91 percent or more within a factor of 5, of the observed values at the model development stations (table 6). Nearly all of the predictions that differ from the observed concentration by more than a factor of 5 are predictions for stations with censored observed concentrations.

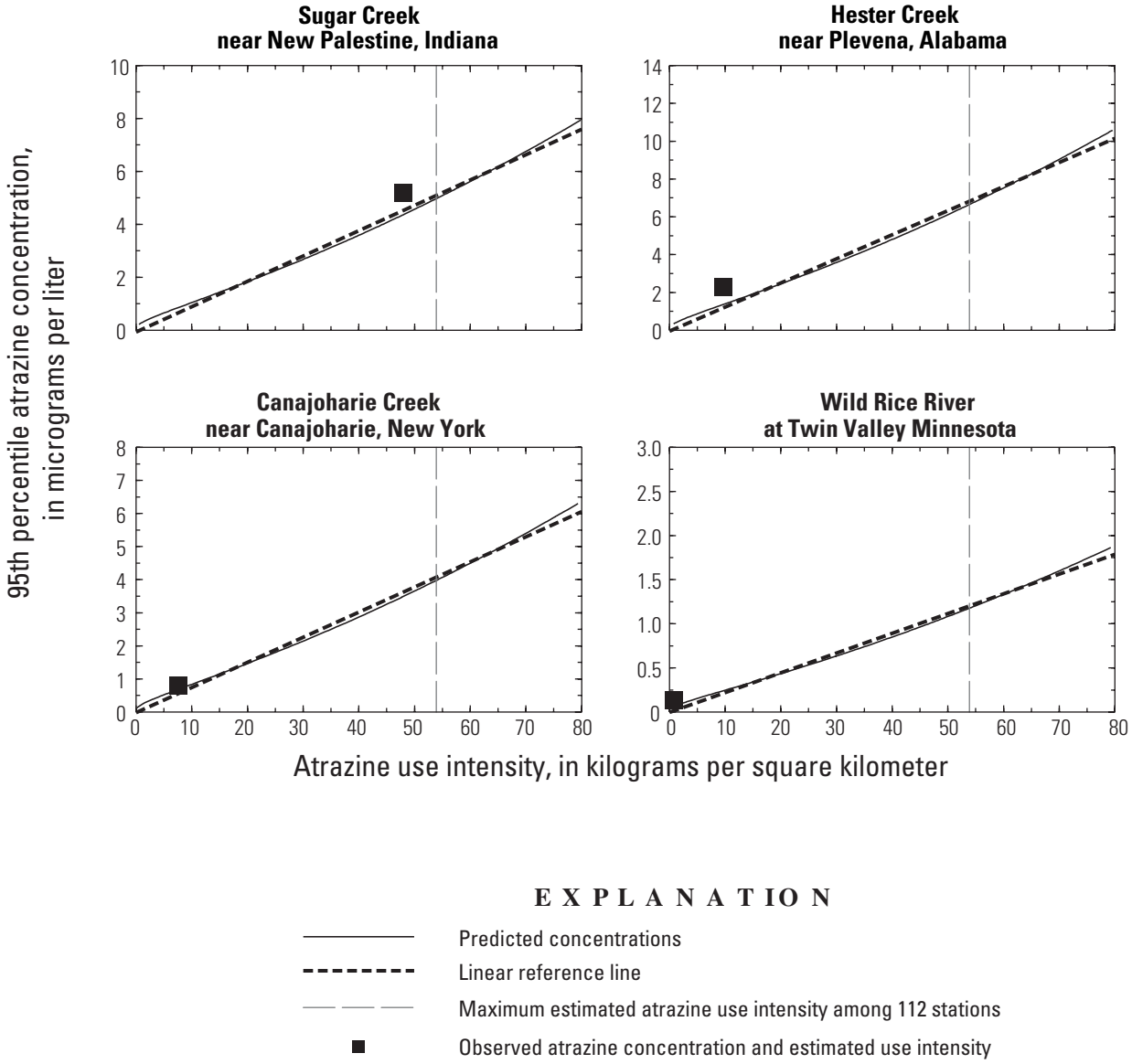


Figure 16. Relation of predicted 95th percentile atrazine concentrations and atrazine use intensity for four stations for simulated atrazine use-intensity values ranging from 0 to 80 kilograms per square kilometer.

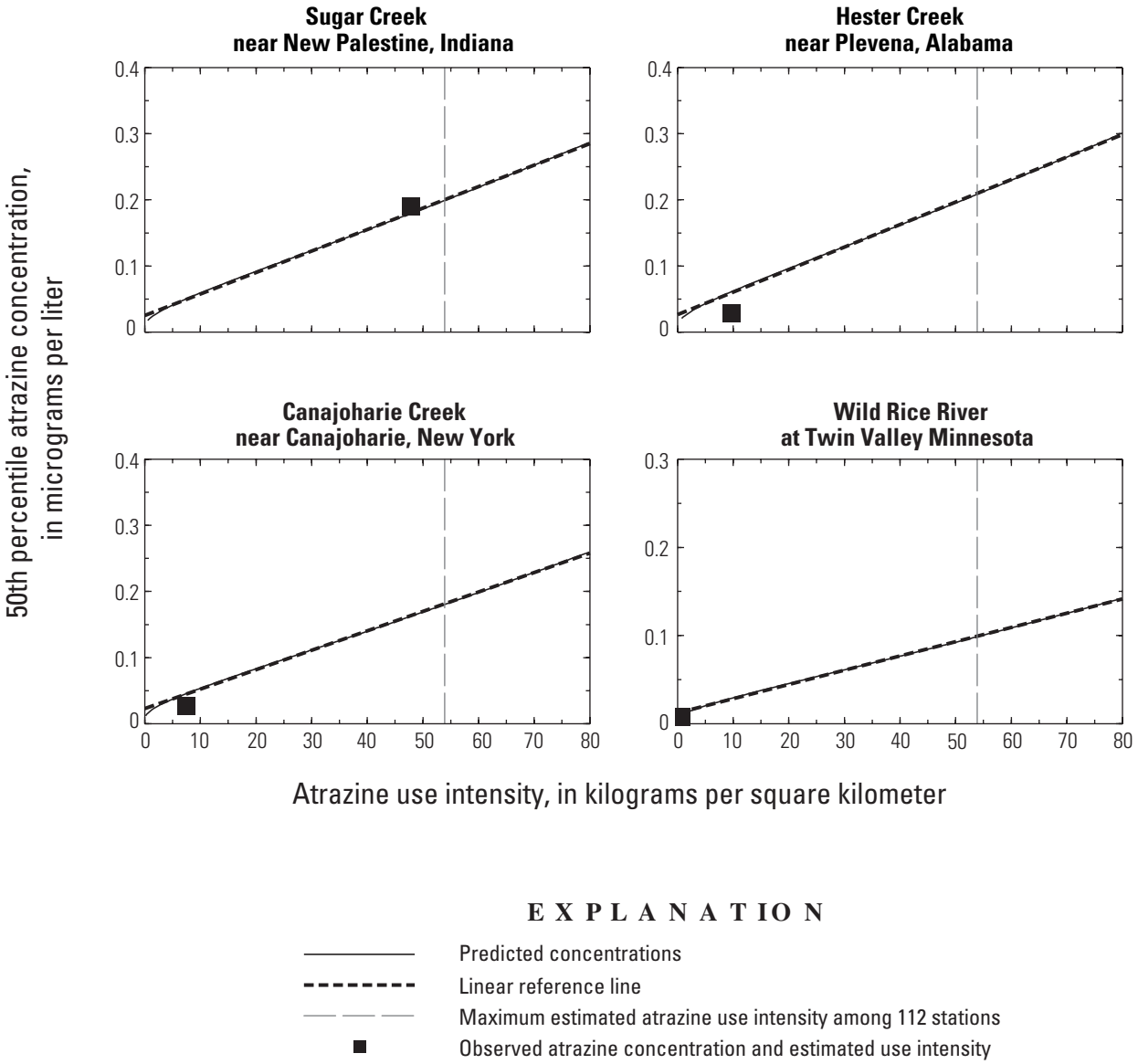


Figure 17. Relation of predicted 50th percentile atrazine concentrations and atrazine use intensity for four stations for simulated atrazine use-intensity values ranging from 0 to 80 kilograms per square kilometer.

Boxplots of the residuals from the models show the performance of the models with respect to geographic region and watershed area for the model development stations. Figures 20 and 21 show residuals grouped by watershed area and geographic region, respectively, for five of the models. Because the residuals are calculated using the base 10 logarithms of observed and predicted values, residuals less than zero indicate overprediction of the concentration and residuals greater than zero indicate underprediction. A residual of zero indicates exact agreement; residuals of –1 and +1 indicate that the prediction for a given station is 10 times and one-tenth of the observed concentration, respectively.

Figure 20 indicates that model predictions are not biased with respect to watershed area. For this plot, watershed areas of the 112 model development stations were divided into quintiles (that is, the first quintile contains the smallest 20 percent of watershed areas, the second quintile contains the second smallest 20 percent, and so on). Residual error was similar for all five groups, which cover over 5 orders of magnitude in watershed area.

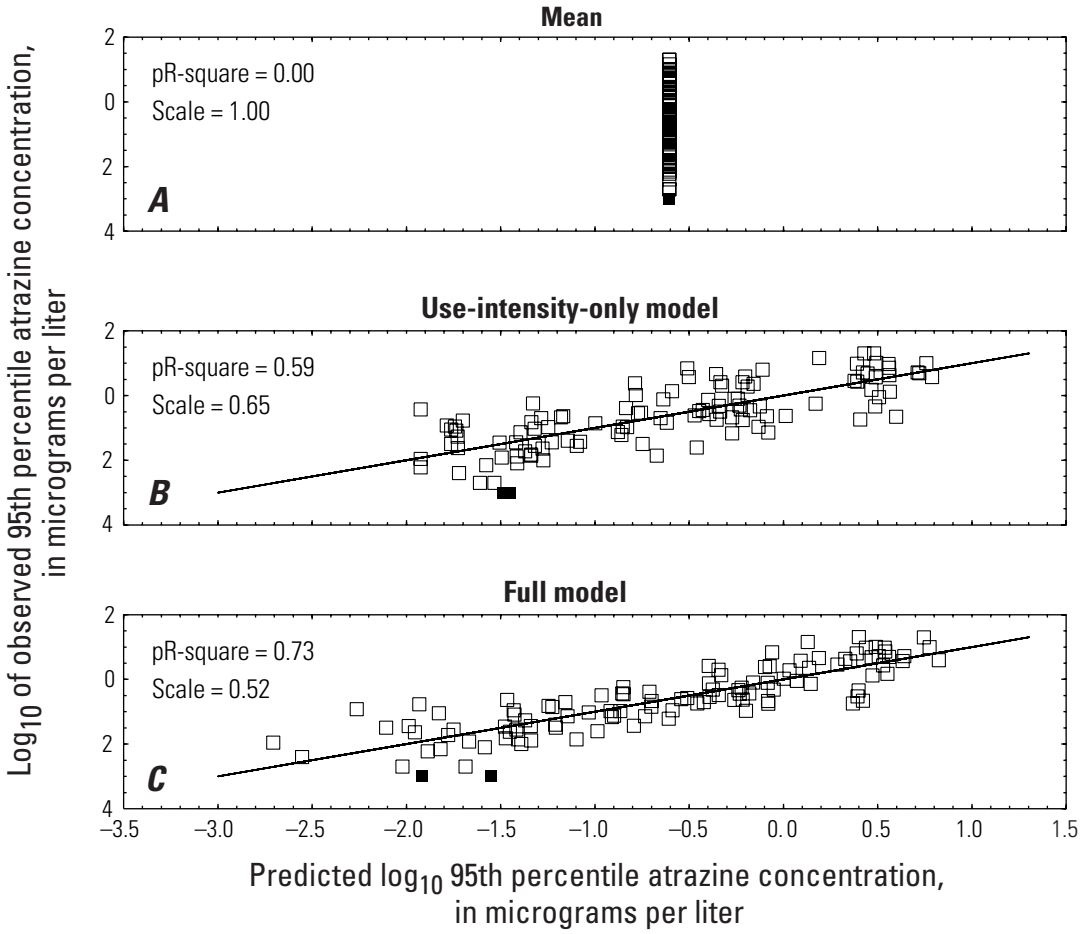


Figure 18. Comparison of observed 95th percentile atrazine concentrations for the 112 model development stations with atrazine concentrations predicted using three different models.

A. Mean. *B*. Use-intensity-only model. *C*. Full model. The solid line is a 1:1 line, indicating exact agreement of the observed and predicted values. Filled symbols indicate censored observed values.

Figure 21 indicates that model predictions of atrazine concentrations have no substantial regional bias. The regional groupings used for this plot are based on the U.S. Department of Agriculture's Farm Resource Regions (U.S. Department of Agriculture, 2000), which categorize agricultural regions of the conterminous United States on the basis of climate, topography, soil types, and dominant agricultural activities. The nine Farm Resource Regions were consolidated into five regions (fig. 22) so that each region would have sufficient data to compute boxplots. The lack of any regional bias partially may be due to similar uses of atrazine in the various regions. Atrazine is primarily used as a pre-emergent herbicide on corn or sorghum throughout the United States.

Model Validation Stations

The 10 models were applied to the 26 stations in the validation data set (fig. 4). Examples of the fit of the models for these validation stations are shown in figures 23 and 24. A summary of the fit of all of the models for the validation stations is shown in table 6. Prediction errors for the validation stations were very similar to errors for the model development sites. Predicted concentrations from the 10 models were within a factor of 10 of the observed concentrations in 96 to 100 percent of cases, and within a factor of 5 of the observed value in 88 to 100 percent of cases. Observed concentrations for these comparisons are from a 1-year period for each of the validation stations. Similar results were obtained when predicted concentrations were compared to all available years of observed concentrations from the validation stations.

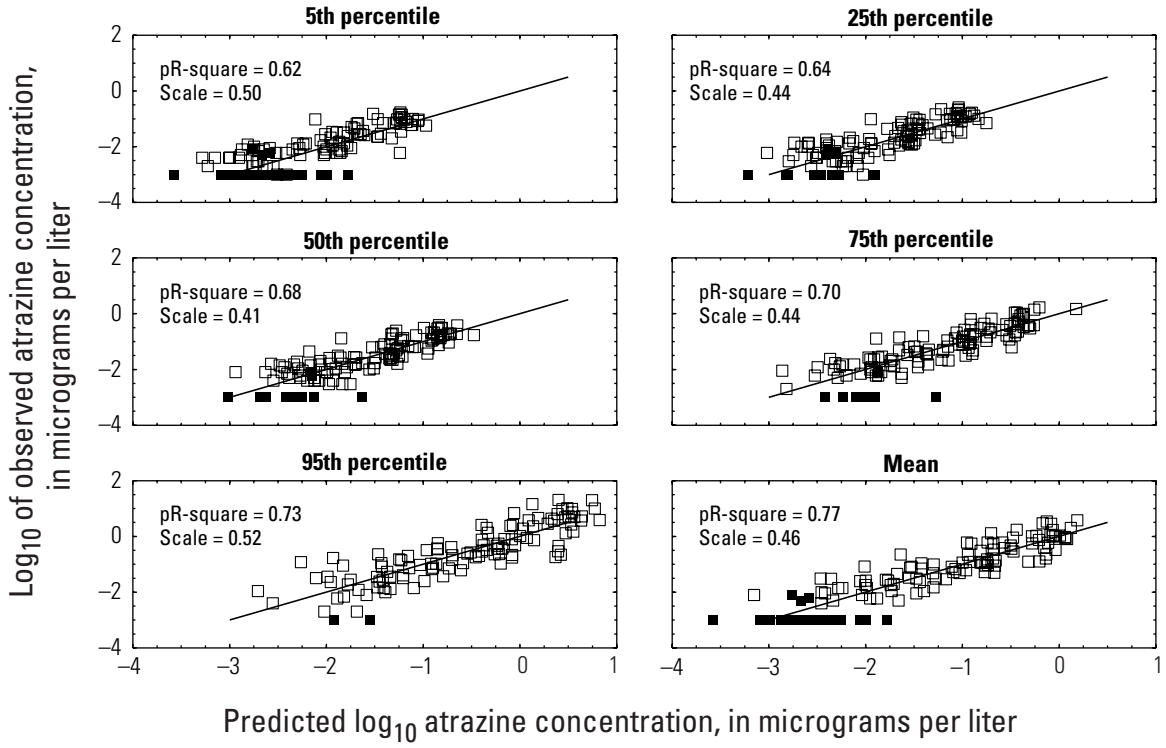


Figure 19. Examples showing the fit of six of the tobit regression models for atrazine.

The solid line is a 1:1 line, indicating exact agreement of the observed and predicted values. Filled symbols indicate censored observed values. pR-square, pseudo R-squared (R-squared value used for tobit regression).

Table 6. Summary of fit of regression models for model development and model validation stations.

[Percentage of predicted values within a factor of 10, 5, and 2 of the observed values are shown for each of the 10 models. Predictions "within a factor of x" of the observed value are between (1/x)*(observed value) and (x)*(observed value). N, number of stations]

Model	Model development stations (N=112)			Validation stations (N=26)		
	Percent of predictions within a factor of			Percent of predictions within a factor of		
	10	5	2	10	5	2
5th	99	95	72	100	100	73
10th	99	96	74	100	100	73
15th	99	95	76	100	100	69
25th	100	96	74	100	100	73
50th	100	96	78	100	92	58
75th	99	96	72	100	92	46
85th	99	91	75	100	96	54
90th	99	93	72	96	88	65
95th	97	92	69	96	88	58
annual mean	99	94	76	100	92	62

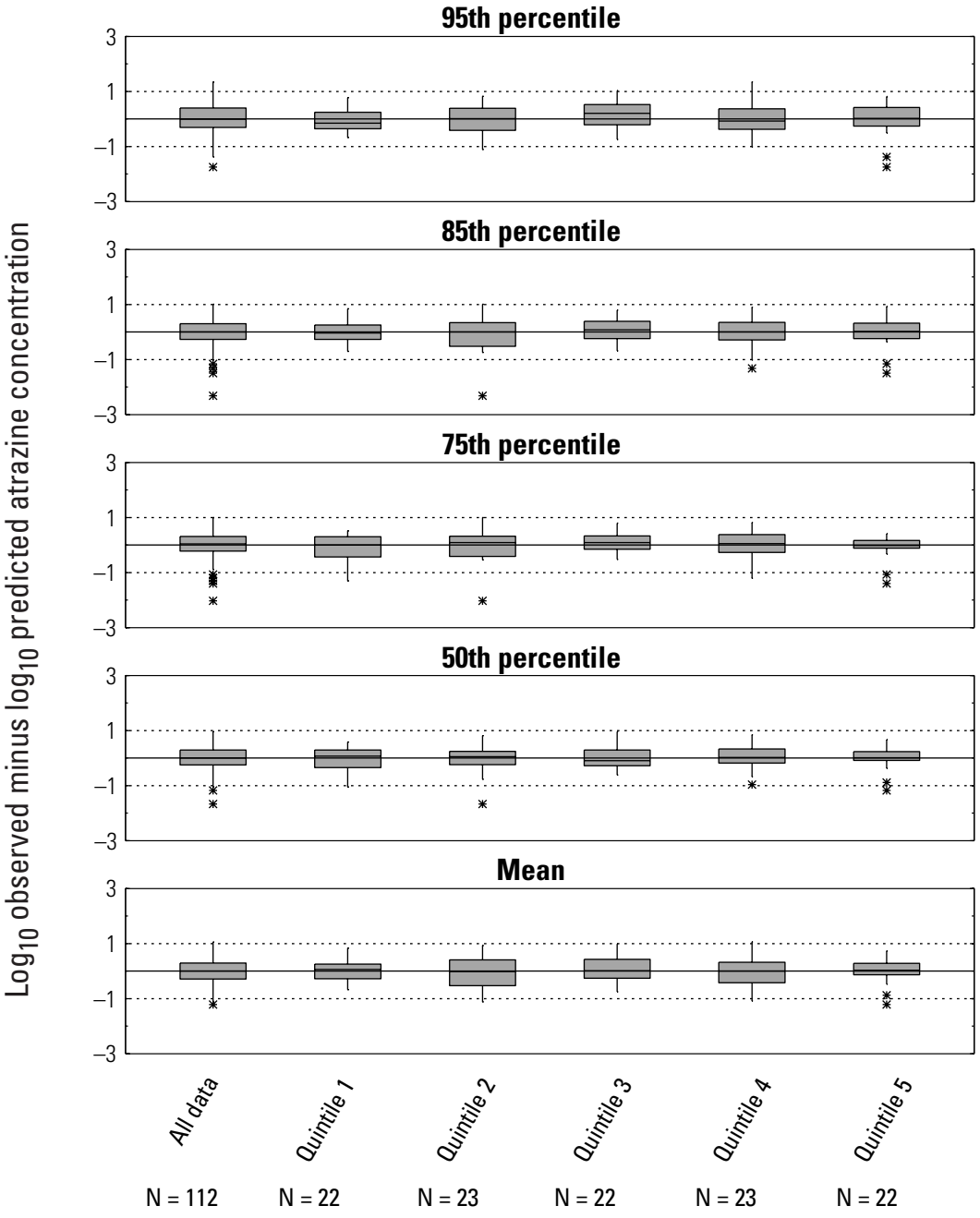


Figure 20. Residual error for five atrazine models for 112 stations grouped into quintiles of watershed area.

First boxplot in each group shows error for all 112 stations. Remaining boxplots show error for stations grouped into five classes based on watershed area quintiles. Residual error is [$\log_{10}$ (observed value) – $\log_{10}$ (predicted value)]. N, number of stations.

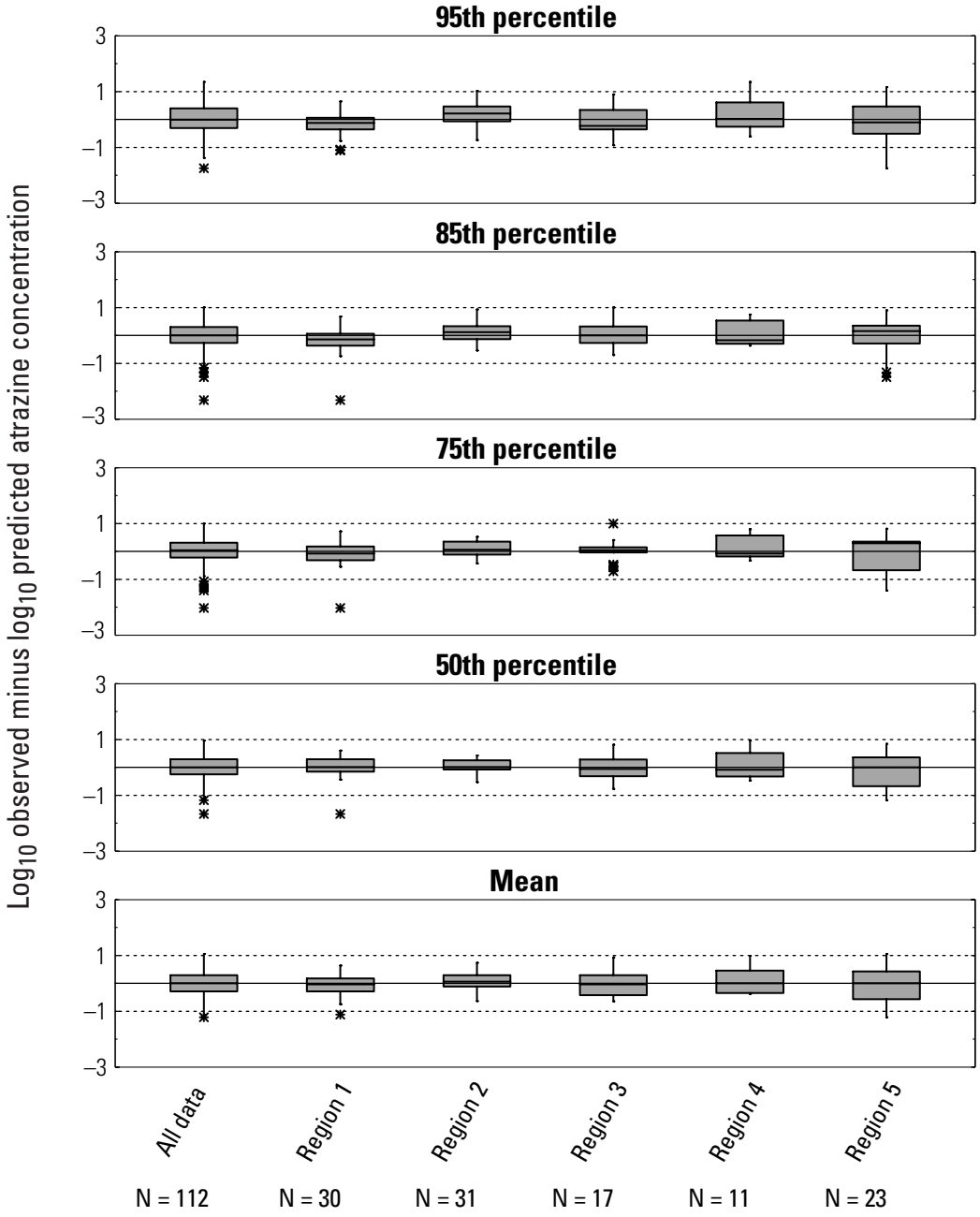


Figure 21. Residual error for five atrazine models for 112 stations grouped by region.

First boxplot in each group shows error for all 112 stations. Remaining boxplots show error for stations grouped by region. Regions from figure 22. Residual error is [$\log_{10}$(observed value) –$\log_{10}$(predicted value)]. N, number of stations.

Despite the relatively good agreement shown in table 6, predictions are biased low for most of the validation stations (figs. 23 and 24). The bias primarily is due to underprediction of low percentiles at ARP stations and of high percentiles at WQL stations. The cause of the bias is not known, but is probably not due to differences in analytical methods used by the different laboratories, as analytical recovery of atrazine was similar for the three methods (J. Fuhrman, Monsanto Company, personal commun., July 7, 2003; R.P. Richards, Heidelberg College Water Quality Laboratory, personal commun., July 7, 2003). The WQL analyzes unfiltered water, whereas the USGS samples are filtered before analysis, but this should have a minimal effect on the measurement of atrazine concentrations, as atrazine is primarily in the dissolved phase in surface waters. Geographic location of the stations could be a factor, as most of the ARP and WQL stations are in the eastern part of the Corn Belt region (fig. 4), whereas the model development stations cover a wide geographic area (fig. 3). Capel and Larson (2001) reported that runoff of atrazine from fields in the eastern Corn Belt is somewhat higher than in other regions of the United States. This could result in a low bias in predictions based on models derived from more widely distributed sampling stations. Predictions for model development stations in Ohio also generally show a low bias, but to a smaller extent than the WQL stations. The low bias in the high percentiles at the WQL stations also may be due to the higher sampling frequency used in their program. At the nine stations sampled by the WQL, the average number of samples collected during the year was 65, compared to an average of 25 samples per year for the model development stations. In addition, the WQL collects extra samples during summer storm events. Thus, the high end of the atrazine concentration distribution is likely more accurately characterized for the WQL-sampled stations than for the model development stations. At the model development stations, especially those on smaller streams, the highest concentrations may not have been sampled and the computed upper percentiles may be biased somewhat low. Application of models derived from these stations to the WQL stations could result in predictions that are biased low. The bias observed for the WQL stations may be due to a combination of the geographic location and the higher sampling frequency.

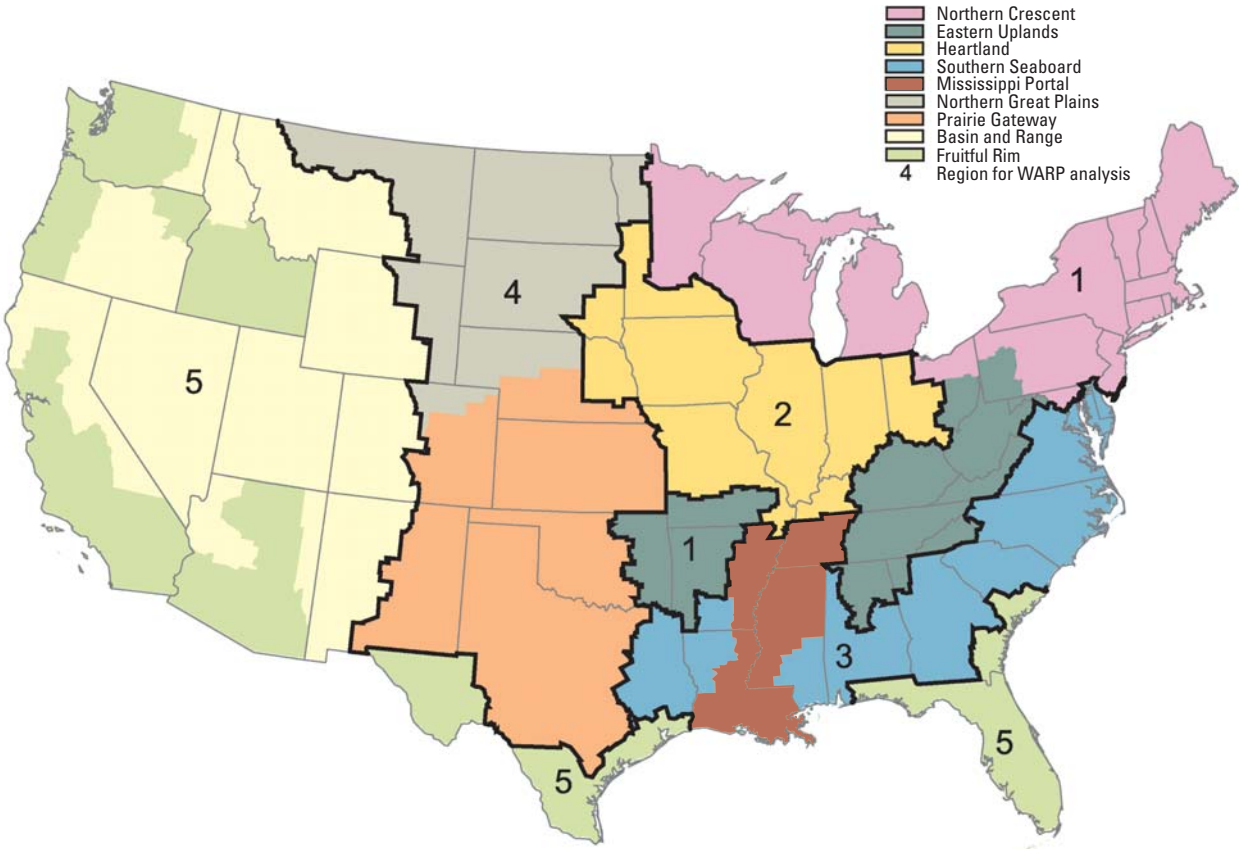


Figure 22. U.S. Department of Agriculture Farm Resource Regions and regions used for WARP model evaluation.

WARP, Watershed Regressions for Pesticides.

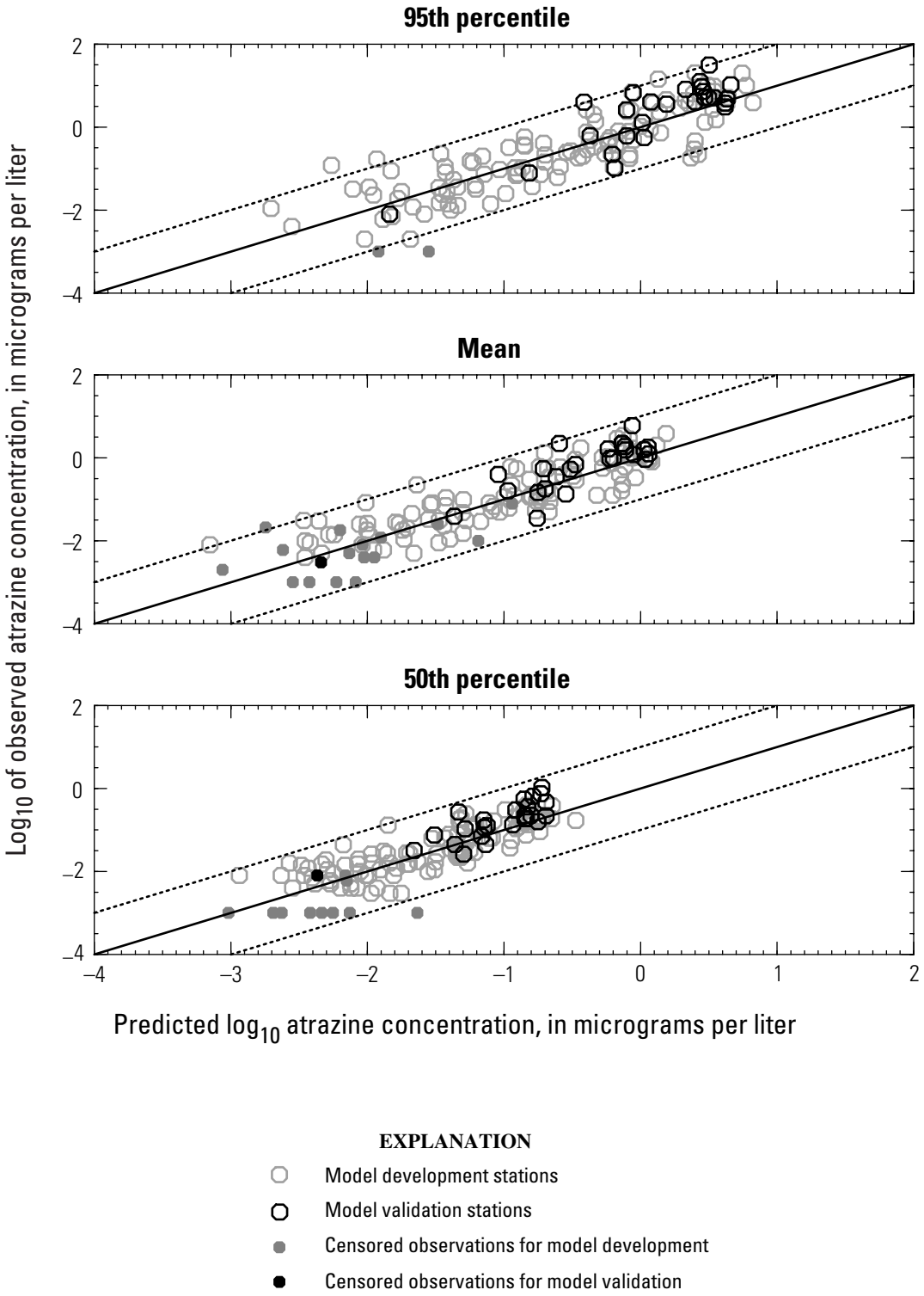


Figure 23. Comparison of observed and predicted atrazine 95th percentile, annual mean, and 50th percentile concentrations for the 26 model validation stations.

Results for the model development stations are shown for comparison. The solid line is a 1:1 line, indicating exact agreement of the observed and predicted values. Dashed lines correspond to predictions of one-tenth (upper line) and 10 times (lower line) the observed value. Filled symbols indicate censored observed values.

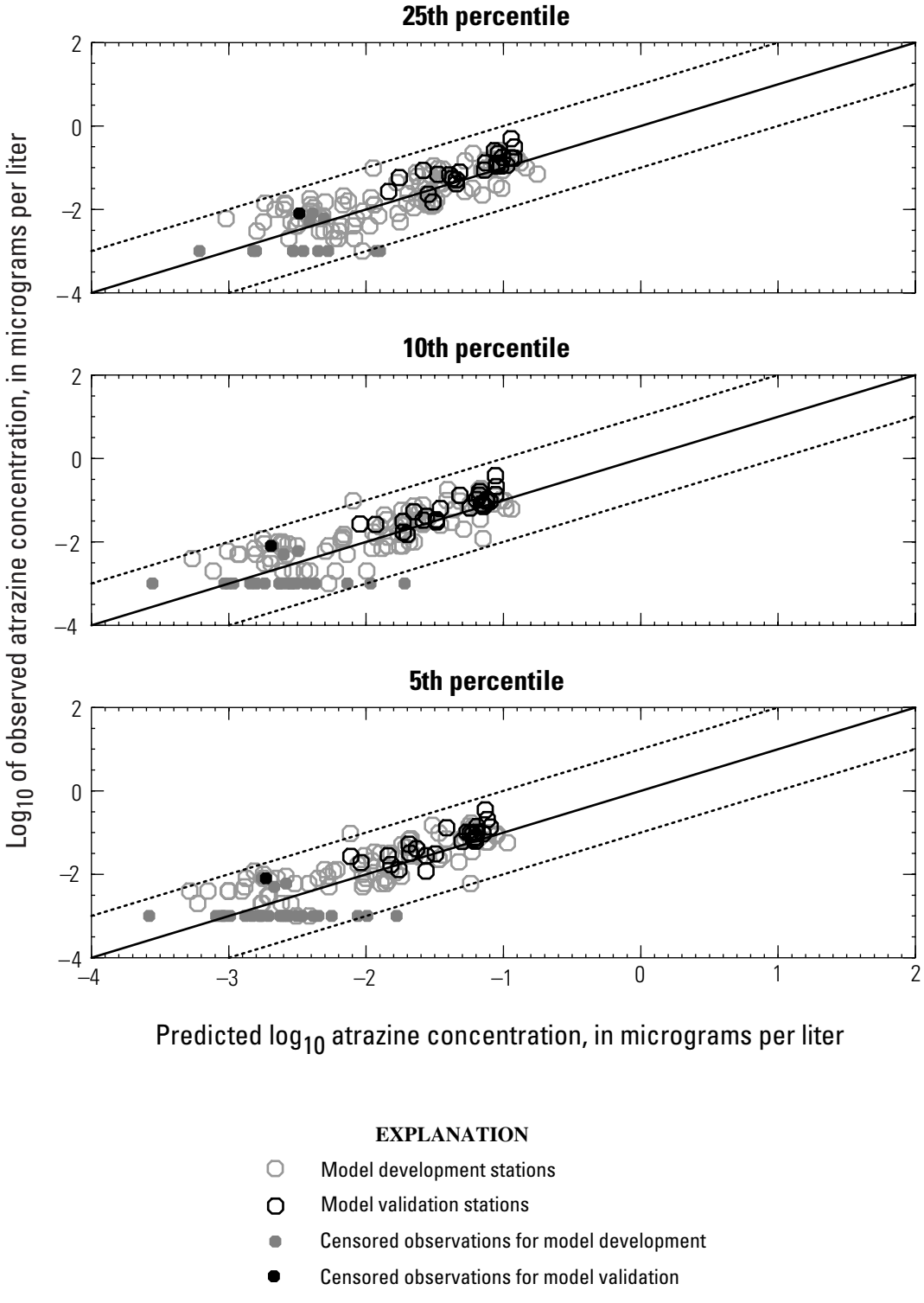


Figure 24. Comparison of observed and predicted atrazine 25th, 10th, and 5th percentile concentrations for the 26 model validation stations.

Results for the model development stations are shown for comparison. The solid line is a 1:1 line, indicating exact agreement of the observed and predicted values. Dashed lines correspond to predictions of one-tenth (upper line) and 10 times (lower line) the observed value. Filled symbols indicate censored observed values.

Predictions for Lakes and Reservoirs

The models also were applied to 16 stations on lakes and reservoirs where untreated water was sampled as part of the ARP program. All of these stations are located in the midwestern Corn Belt region, and the sampling frequency was similar to the sampling frequency of the ARP stream stations in the validation data set (14 to 15 samples per year). These lake and reservoir stations are not regarded as model validation stations, as the models were developed using data from flowing water systems. The models were applied to these stations to illustrate how the models perform when applied to lakes and reservoirs. A bias is to be expected in the predictions for lakes and reservoirs, as the temporal pattern of atrazine concentrations in lakes and reservoirs is somewhat different than the pattern in flowing water systems. In general, peak concentrations are lower in lakes and reservoirs, but the period of elevated concentration lasts longer (Larson and others, 1997; Battaglin and Goolsby, 1998). This results in higher values at the low end of the concentration distribution and lower values at the high end of the concentration distribution for lakes and reservoirs compared to streams with similar atrazine use-intensity in the watershed. Because the models were developed using concentration data from streams, model predictions for the low end of the concentration distribution for lakes and reservoirs are expected to be biased low. Concentrations at the high end of the distribution for lakes and reservoirs, on the other hand, should be overpredicted using models developed from stream data. Results are shown in figure 25, in which residuals from all 10 models are plotted as boxplots. Residuals for the 26 validation stations are shown for comparison.

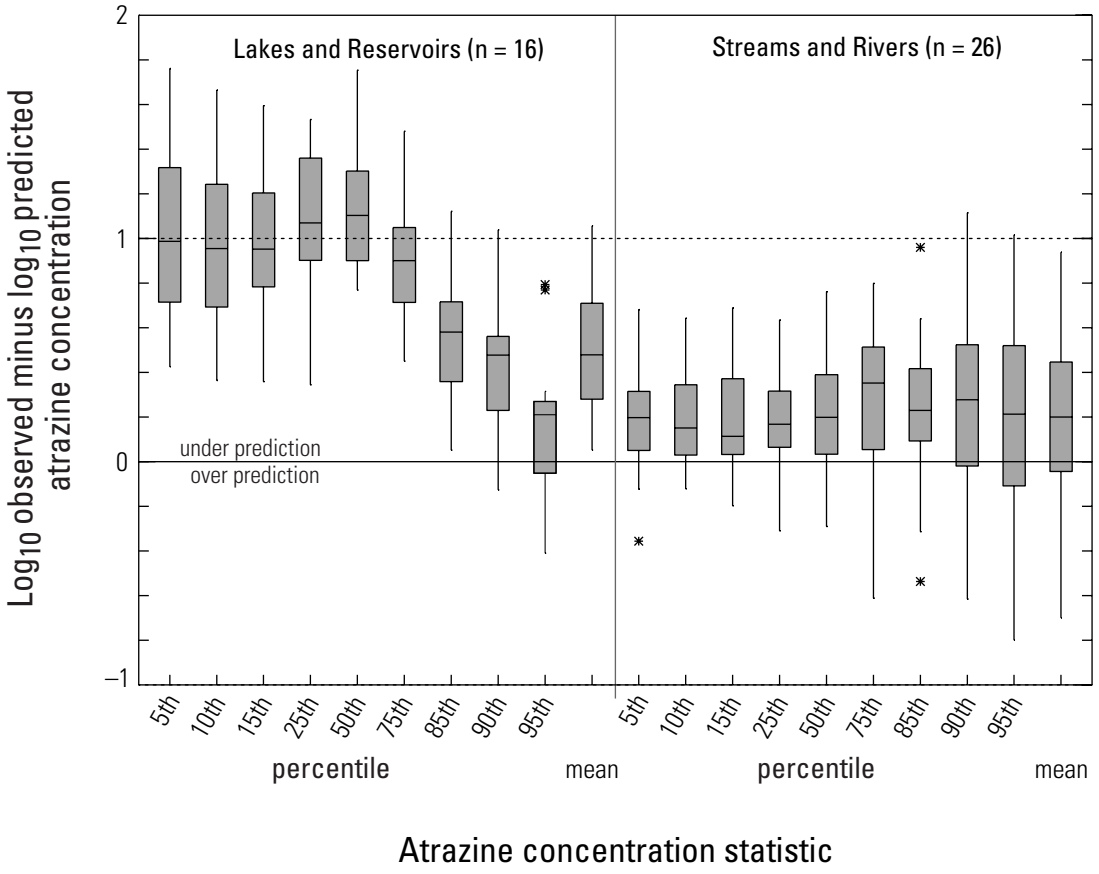


Figure 25. Prediction error for validation stations on lakes and reservoirs and for validation stations on streams and rivers.

Positive residuals imply underprediction (predicted concentrations lower than observed values); negative residuals imply overprediction.

As expected, predicted concentrations for the lower percentiles at many of the lake and reservoir stations were biased low, with predicted concentrations 1/5 to 1/20 of the observed concentrations at most stations. The high bias (overprediction) expected for the upper percentiles generally did not occur for these 16 stations; the upper percentiles also were underpredicted. This could be related to the sampling frequency used for the model development stations, which is somewhat low for capturing peak concentrations of atrazine in streams, especially small streams. If peak concentrations were not sampled in some model development streams, the upper percentiles computed for these streams would be biased low. For lakes and reservoirs, which generally have a longer period of elevated atrazine concentrations, peak concentrations are more likely to be sampled and values computed for the upper percentiles are probably more accurate. This would result in a low bias for upper percentile predictions in lakes and reservoirs. Regression models developed using monitoring data from lakes and reservoirs may be needed to more accurately predict concentrations of atrazine in lakes and reservoirs.

Uncertainty in Model Predictions

Uncertainty in a predicted concentration can be expressed in terms of a confidence interval (CI) and a prediction interval (PI). Each predicted value corresponds to a particular combination of values for the explanatory variables, corresponding to a particular sampling station in this study. If data were collected at a number of stations that have identical watershed characteristics or over a number of years at the same station, so that many values of the modeled concentration measures were available, the mean of these values would be expected to fall within the CI computed for the model prediction, with a specified level of confidence. The confidence level used for the CI in this study is 95 percent, implying a 95 percent probability that the "true" mean value of repeated measurements at the same station falls within the CI. The prediction interval, on the other hand, represents the likelihood of a particular value for an individual site and year (rather than a mean value) falling within a specified interval of the predicted value. The confidence level used for the PI in this study is also 95 percent.

Examples of CIs and PIs are shown in figure 26. In these plots, the model development stations are arranged in order of increasing predicted 95th percentiles (26*A*) and predicted 50th percentiles (26*B*). The upper and lower bounds of the CI and PI for each predicted value are shown as lines above and below the line for the predicted values. Concentrations are expressed as logarithms in these plots, resulting in symmetric intervals for both the CI and PI (the high and low bounds of the intervals are the same distance from the predicted value). The CI is smaller than the PI for a given predicted value, reflecting the greater certainty in a mean value compared to an individual value. The width of the CIs is more variable than the width of the PIs among the different stations. The width of the PIs is nearly uniform throughout the range of predicted values for the 112 stations. Variations in the width of intervals for predictions from a given model reflect differences in the values of the explanatory variables for individual stations. Stations with extreme values in one or more of the explanatory variables have wider intervals, and greater uncertainty in their predicted concentrations, than stations with explanatory variables that fall closer to the center of the explanatory data. Intervals for the 50th percentile predictions are somewhat smaller than intervals for the 95th percentile predictions, reflecting the better fit (lower scale) of the 50th percentile model compared with the 95th percentile model.

The logarithmic plots in figure 26 clearly show the relative widths of the CIs and PIs over the entire range of predicted values for a given model. However, expressing the concentrations as logarithms obscures the fact that the intervals are actually skewed—the upper part of both intervals covers a wider range of concentrations than the lower part. This is illustrated in figure 27, in which the same data are plotted without the logarithmic transformation of concentrations. In these plots, the upper part of the intervals is considerably wider than the lower part, especially for the higher predicted concentrations. In the case of the highest predicted value for the 95th percentile (6.7 µg/L for the Mississippi River at St. Francisville, La.), the CI extends from 1.26 to 35.1 µg/L, and the PI extends from 0.34 to 129.6 µg/L.

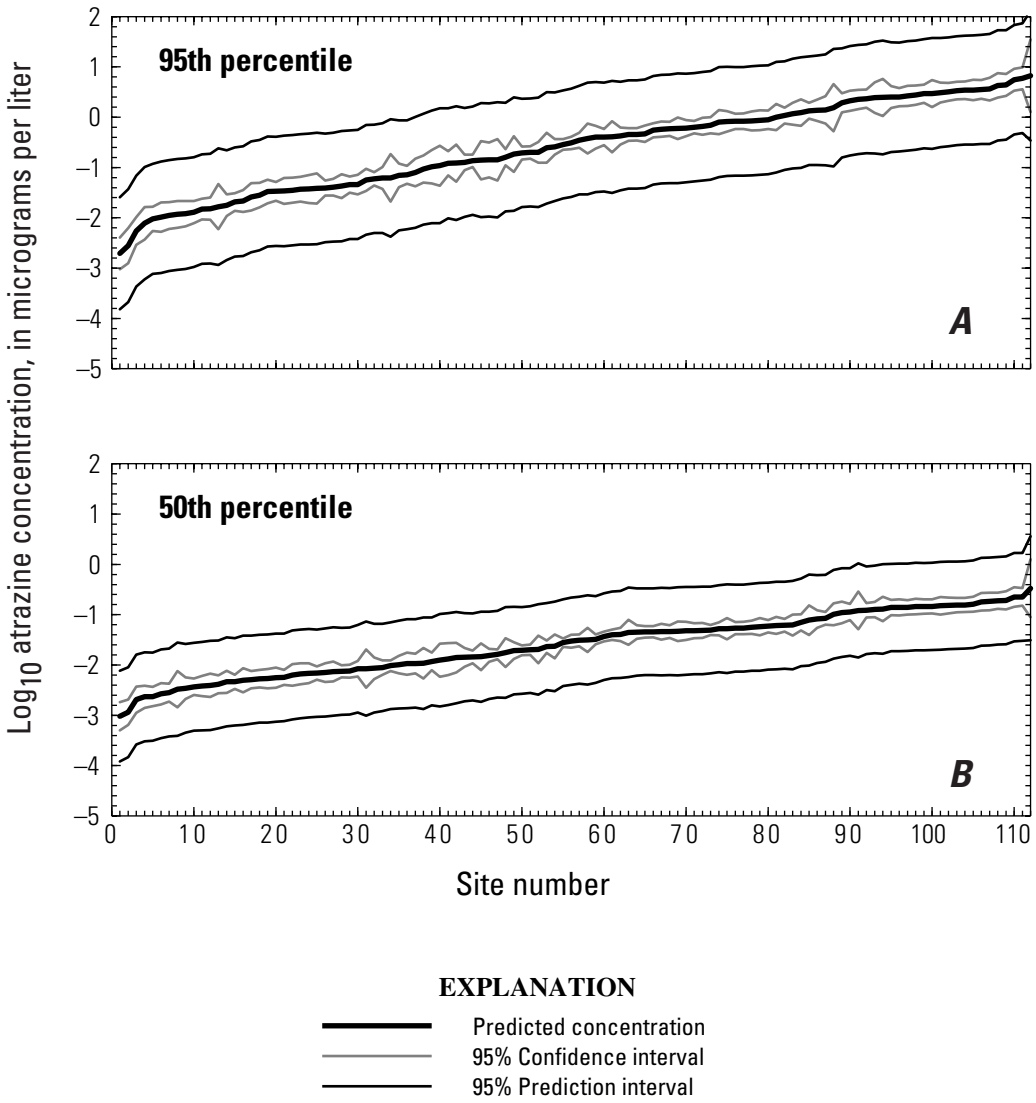


Figure 26. Examples of confidence intervals and prediction intervals for predicted atrazine concentrations (log-transformed) for the 112 model development stations.

A. 95th percentile. *B.* 50th percentile. Stations are arranged in order of increasing predicted concentrations.

The levels of uncertainty among the 10 models are compared in figure 28. The size of the CIs and PIs are represented as the ratio of the upper bound of the interval to the predicted value (this is the same as the ratio of the predicted value to the lower bound of the interval). For example, for the Mississippi River case mentioned in the previous paragraph, the ratio for the CI is 5.24 (35.1/6.7) and the ratio for the PI is 19.3 (129.6/6.7). Expressing the size of the intervals in this way allows comparisons with predicted values similar to comparisons made for predicted and observed values (table 6). The 95 percent CIs for most predictions extend to less than a factor of 2 above and below the predicted value for all 10 models (fig. 28*A*). The extreme values (shown as asterisks) are for the same four stations for all 10 models. In each case, one or more of the explanatory variables for the station has a relatively extreme value compared with the rest of the stations. The highest ratios, shown by the asterisks well above the others, are for the Mississippi River at St. Francisville, La, which has a watershed area far greater than the rest of the stations. The width of 95 percent PIs varies among the 10 models (28*B*). PIs for the 15th, 25th, 50th, and 75th percentile predictions, and for predictions of the annual mean, extend to a factor of 7 to 10 above and below the predicted value for most stations. PIs for the remaining models extend to a factor of 10 to 13 above and below the predicted values for most stations. Thus, the width of prediction intervals generally is similar to or somewhat larger than the largest differences seen between predicted and observed atrazine concentrations, which are mostly less than a factor of 10. Examples of the CI and PI for specific stations, and their relation to predicted concentrations and multiple years of observed concentrations, are shown in the section on year-to-year variability.

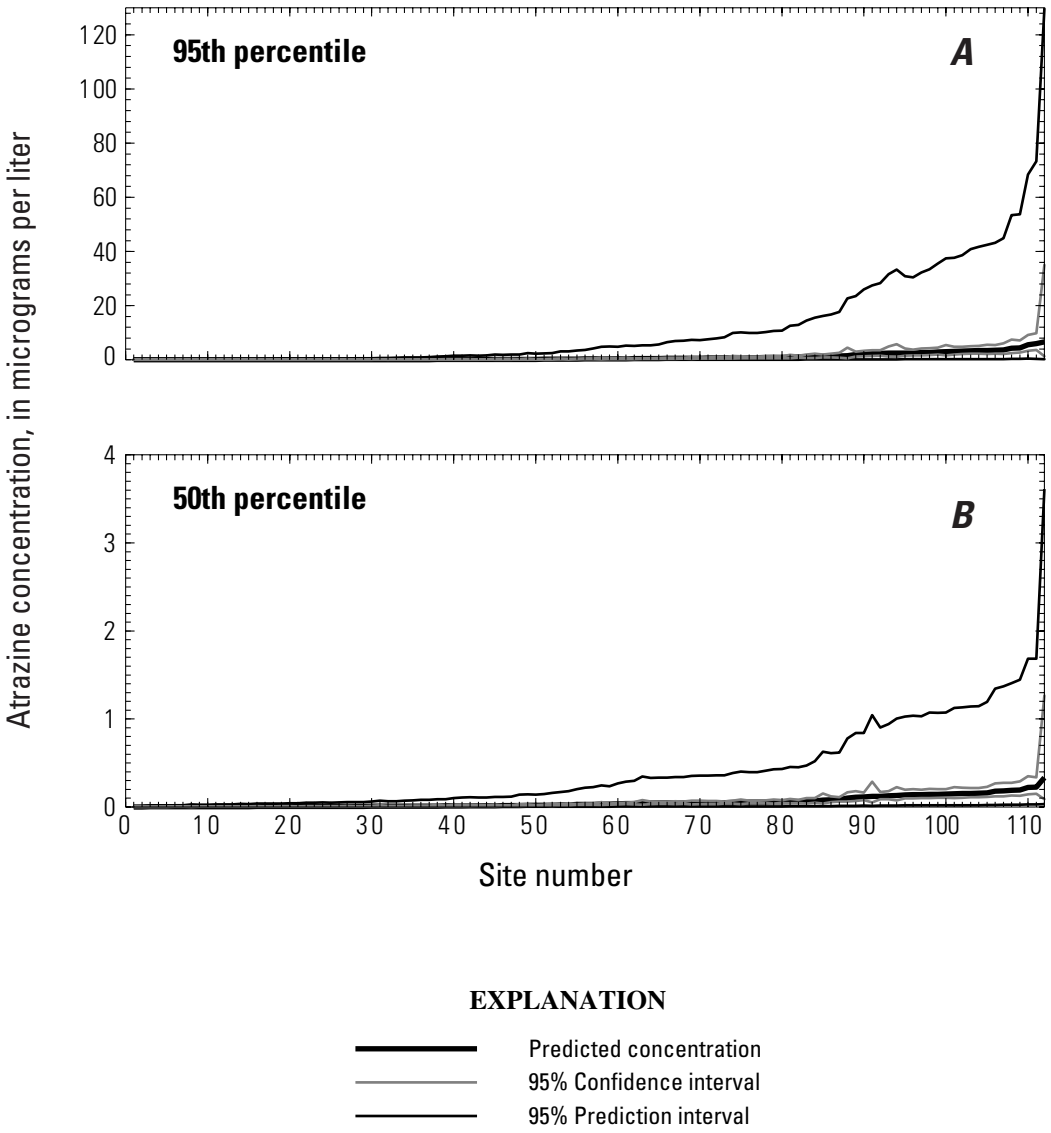


Figure 27. Examples of confidence intervals and prediction intervals for predicted atrazine concentrations (not log-transformed) for the 112 model development stations.

A. 95th percentile. *B*. 50th percentile. Stations are arranged in order of increasing predicted concentrations.

Prediction of Annual Concentration Distributions

Model Development Stations

The nine predicted percentiles for a given stream can be combined to provide an estimate of the frequency distribution of concentrations in the stream for a 1-year period. This estimate then can be compared to the distribution obtained by using the percentiles computed from the measured concentrations. Examples of estimated and observed concentration distributions are shown in figures 29, 30, and 31 for rivers and streams that have very large, medium to large, and small watersheds, respectively. All of these rivers and streams were used for model development. The estimated distributions reasonably reproduce the observed distributions in nearly all of these examples. For all of the examples, all estimated percentiles are within a factor of 10 of the observed percentiles.

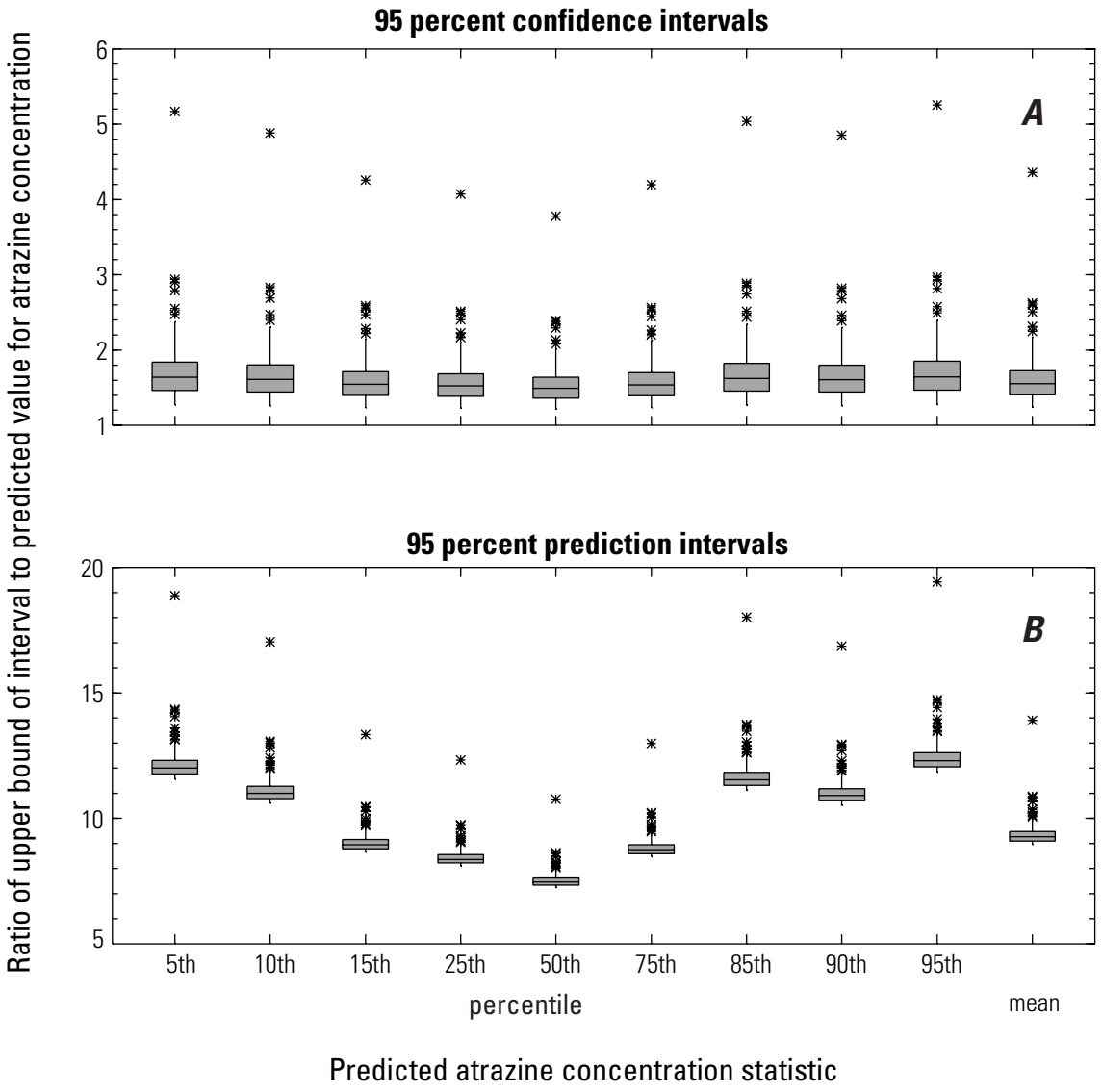


Figure 28. Ratios of the upper bound of 95 percent confidence and prediction intervals to predicted atrazine concentrations from 10 models.

A. Confidence intervals. *B*. Prediction intervals. Each boxplot shows the distribution of ratios among the 112 model development stations.

Model Validation Stations

Predicted and observed distributions for four stream stations from the validation data set (fig. 32) show that the predicted concentrations generally are biased low, consistent with results shown in figure 25. The general shape of the distributions is reasonably reproduced by the model predictions, however, and most predicted concentrations are well within a factor of 10 of the observed concentrations. In figure 33, predicted and observed distributions are shown for four lakes and reservoirs. Underprediction of the lower percentiles is much more evident for these stations, and the estimated distributions do not match the observed distributions nearly as well as for the stream stations. This is consistent with the earlier discussion of biased predictions for lakes and reservoirs, and reflects the fact that the models were developed using data from flowing-water systems. Estimates of the higher percentiles for the lakes and reservoirs were more accurate, with most estimates for the 95th, 90th, and 85th percentiles within a factor of 10 of the observed values.

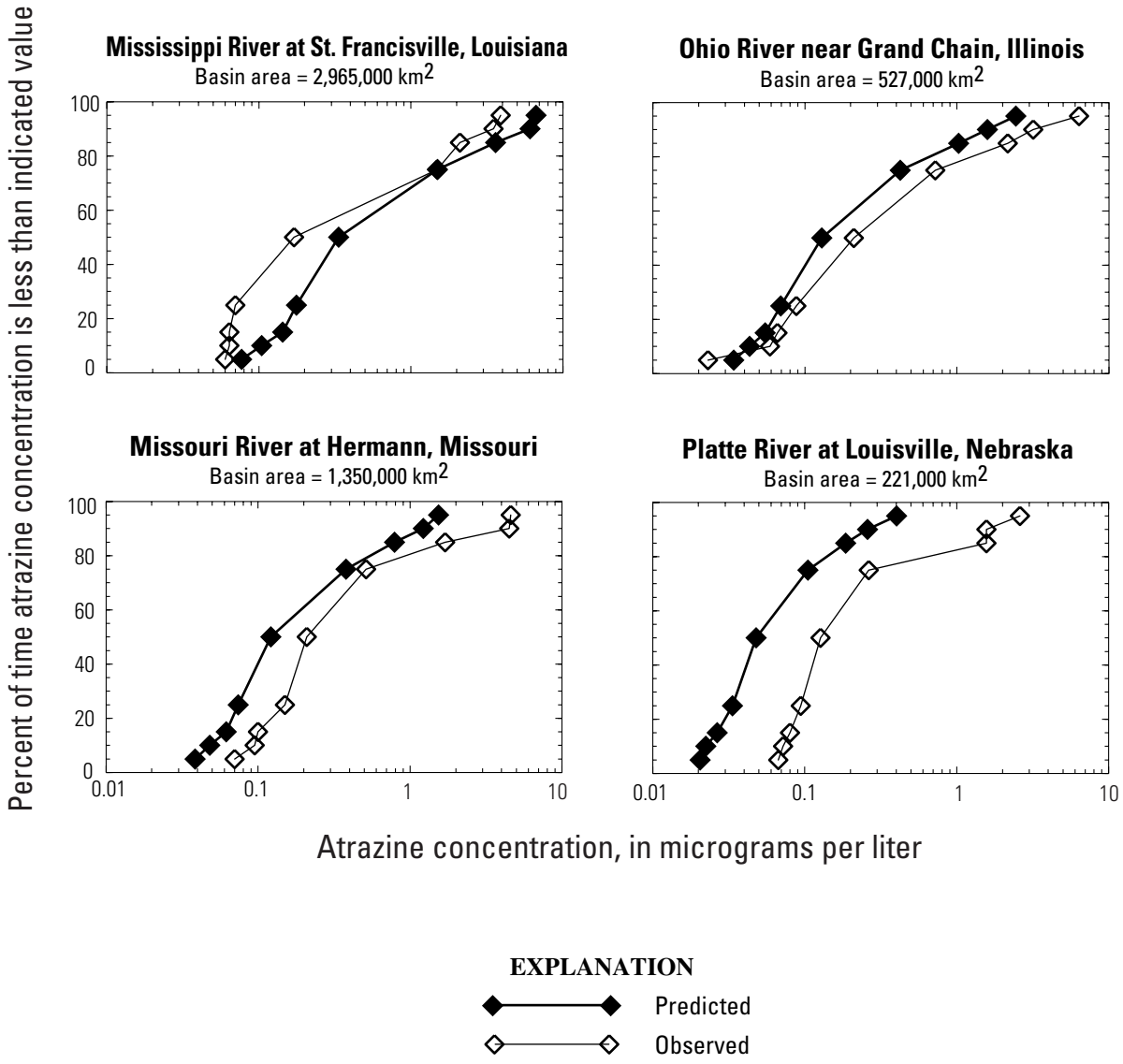


Figure 29. Examples of predicted and observed annual atrazine concentration distributions for model development rivers with very large watersheds.

Year-to-Year Variability

Predicted concentrations (and the predicted annual concentration distributions) obtained from the regression models for a given station do not correspond to specific years. The only explanatory variable in the atrazine models that has a temporal component is atrazine use intensity, with different estimates used for stations where samples were collected before 1995 and for stations where samples were collected during 1995 and later years. The other explanatory variables in the regression models are either long-term average values (R-factor and Dunne overland flow) or essentially constant with respect to time (watershed area and K-factor) for a given watershed. Thus, predictions for a particular station are the same for all years before1995, and for 1995 and later years. This implies that the accuracy of the predictions for a station depends, in part, on how the distribution of concentrations at that station varies from year to year.

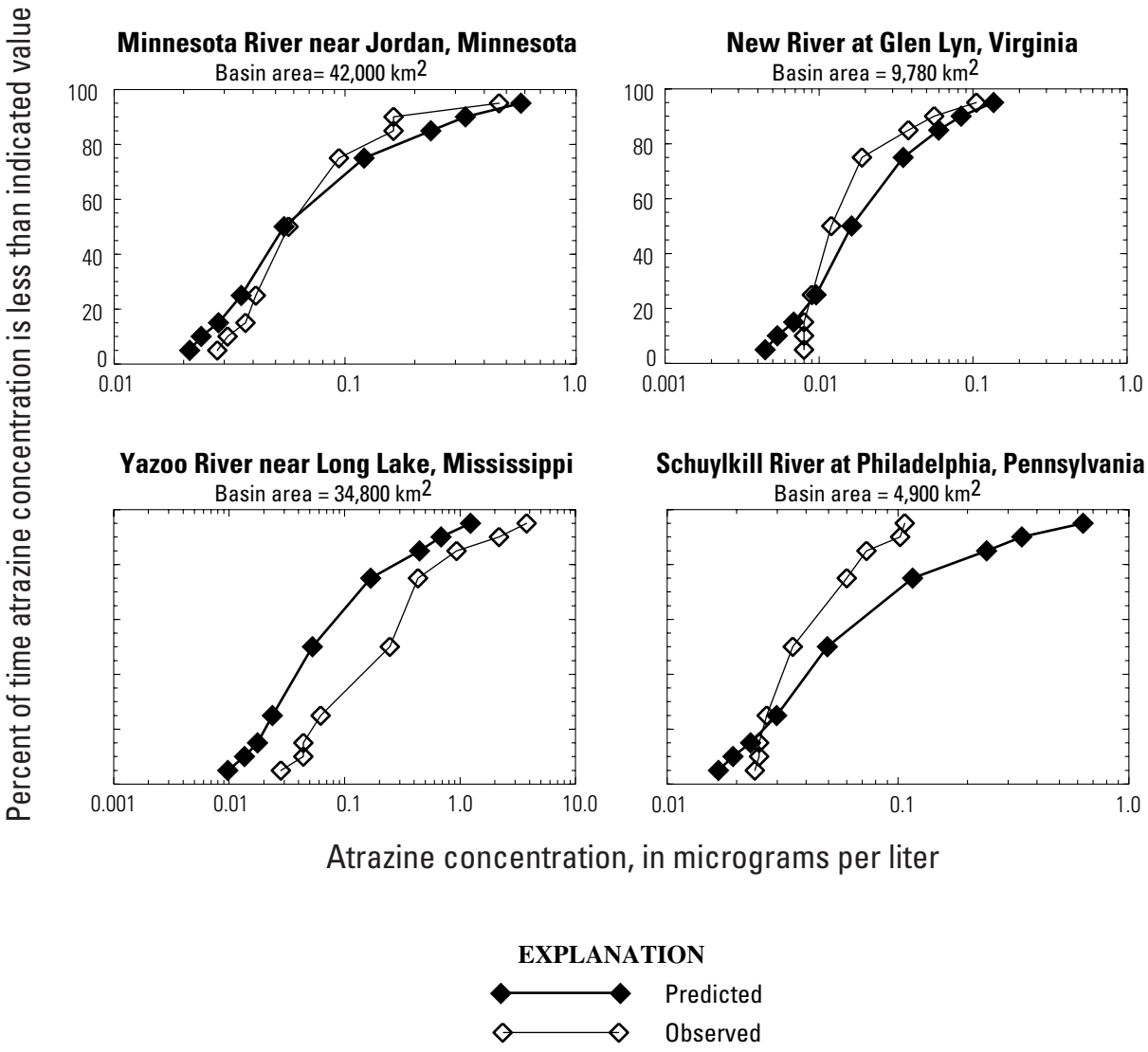


Figure 30. Examples of predicted and observed annual atrazine concentration distributions for model development rivers with medium to large watersheds.

km 2, square kilometer.

Computed percentiles and annual mean concentrations are available for multiple years at some of the model development and model validation stations, and these data can be used to evaluate year-to-year variability. A total of 54 stations on streams (31 model development stations and 23 validation stations) had multiple years of data: 19 stations had 2 years, 11 stations had 3 years, 10 stations had 4 years, 4 stations had 5 years, 6 stations had 6 years, 3 stations had 7 years, and 1 station had 8 years. The ratio of the maximum value to the minimum value of a concentration statistic for a station can be used as an indication of the year-to-year variability at the station. The boxplots in figure 34 show the distribution of ratios among the 54 stations for each of the concentration statistics. The ratios are expressed as base 10 logarithms in this plot, so that a value of 1.0 indicates that the maximum value for a station was 10 times the minimum value for the station. Ratios greater than 10 occurred in only 7 of the 540 total cases. In most cases, the ratios were well below 10, with a mean ratio of less than 5 for all of the concentration statistics. Ratios greater than 10 for the 5th, 10th, and 25th percentiles (four stations) all involved a censored value for the minimum value, and the maximum value also was very low. The one ratio greater than 10 for the 95th percentile was for the Missouri River at Omaha, one of the NASQAN stations used for model validation. Four years of data were available for this station, with 95th percentile values of 4.0, 0.18, 2.2, and 1.08 µg/L, resulting in a ratio of 22.2 (4.0/0.18).

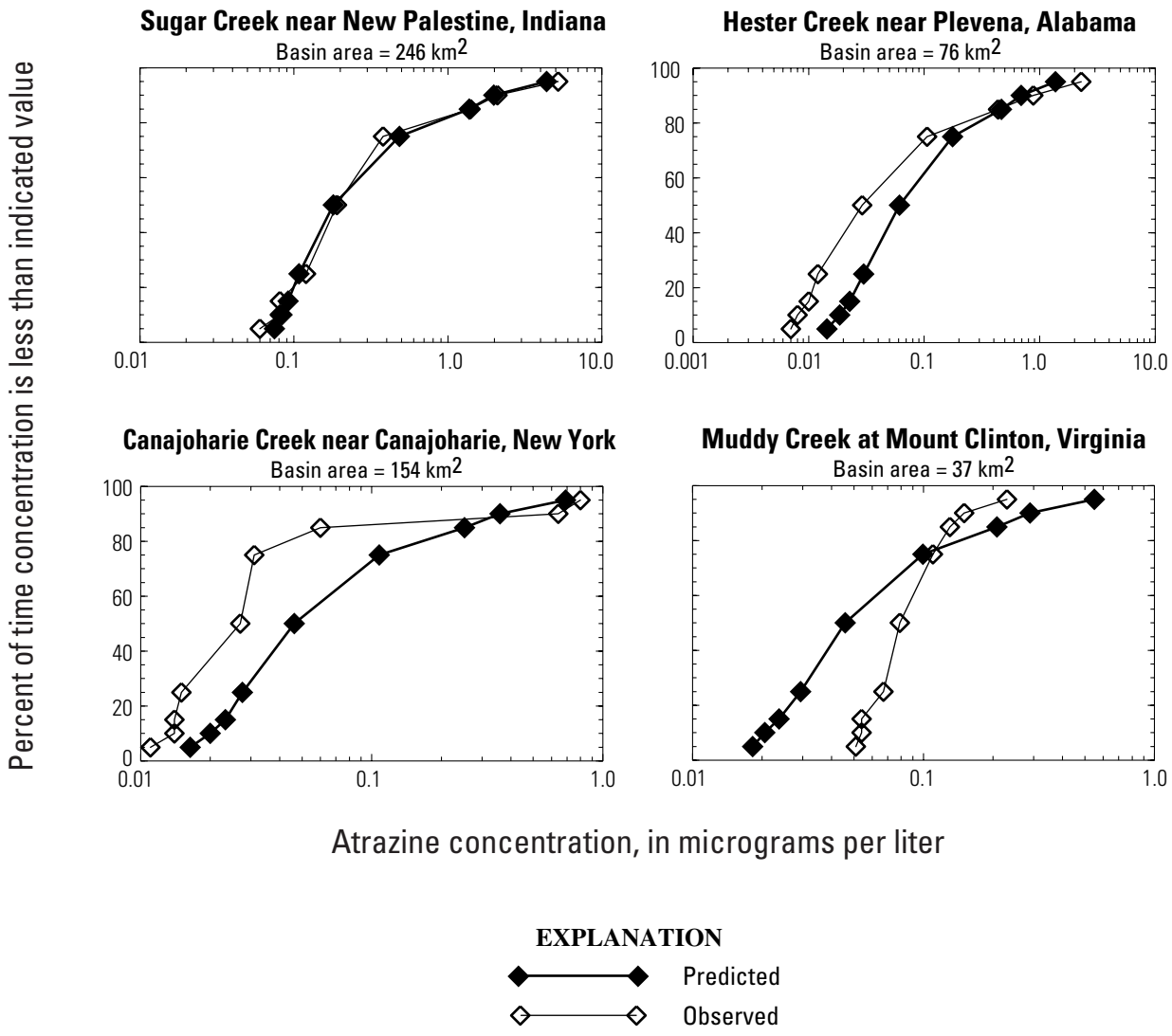


Figure 31. Examples of predicted and observed annual atrazine concentration distributions for model development streams with small watersheds.

km ^{2}, square kilometer.

Comparisons of predicted atrazine concentration distributions and multiyear observed concentration distributions are shown in figures 35 and 36. Figure 35 shows the model-estimated distribution and the observed distributions for 1992 through 1999 for the White River in Indiana. This river was included in the model development data set. The range of observed values for each percentile all fall within an order of magnitude. The estimated percentiles are within an order of magnitude of the observed values, so that the predicted distribution reasonably matches the observed distribution for each of the 8 years. Prediction intervals computed for each of the estimated percentiles also are shown. The prediction intervals would be expected to contain 95 percent of observed values obtained from repeated years of sampling at the station. The observed values for the White River fall within the prediction intervals for all 8 years. Figure 36 shows similar results for one of the validation stations, the Sandusky River in Ohio, sampled by the Heidelberg College WQL. Observed concentration distributions are shown for 1990 through1991 and 1994 through 1998. The model-estimated distribution is shown along with the prediction interval for each predicted value. Although there is a negative bias (underprediction) for some of the estimated percentiles, nearly all of the observed percentiles fall within the prediction intervals of the estimated values.

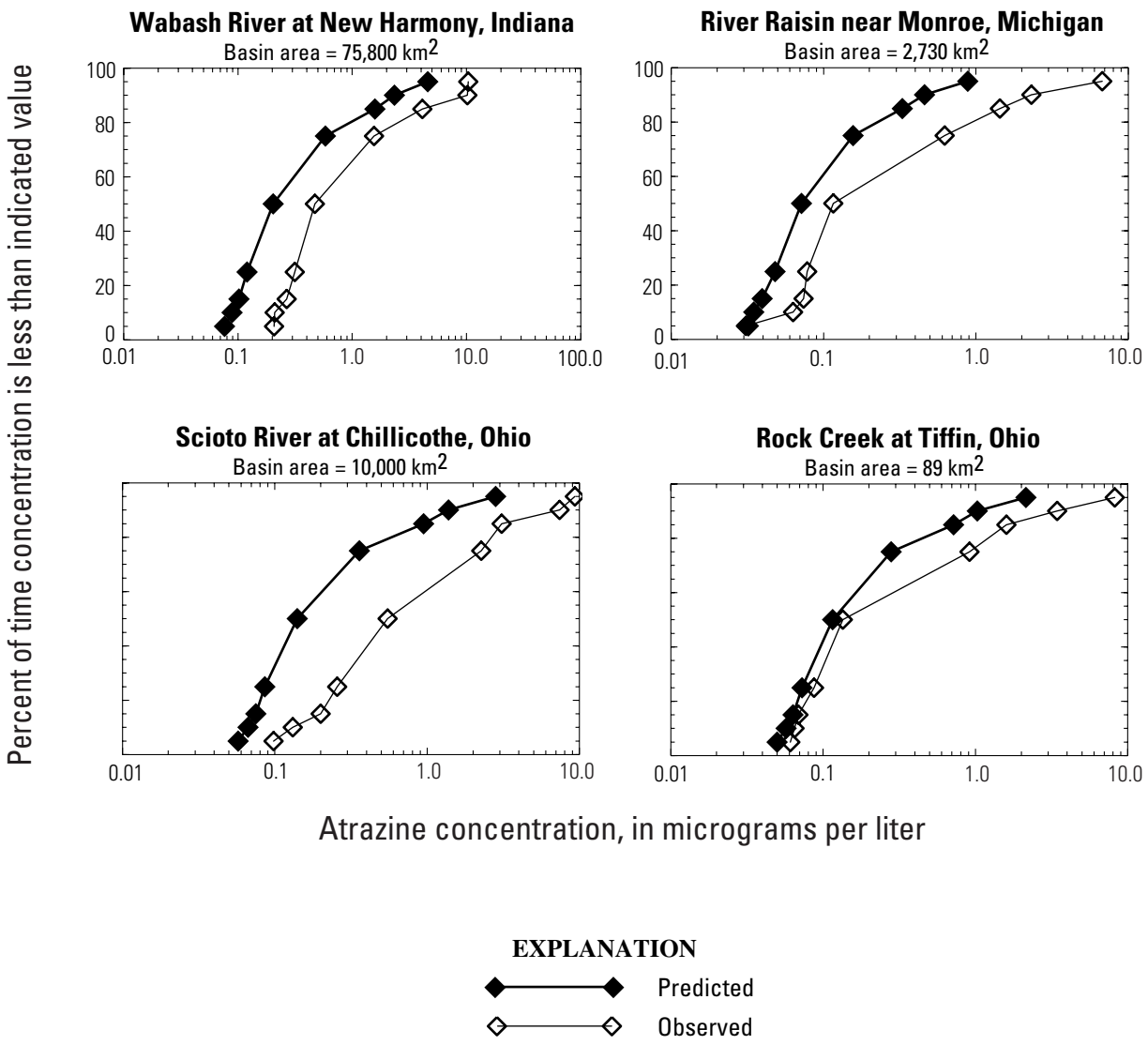


Figure 32. Examples of predicted and observed annual atrazine concentration distributions for four rivers included in the model validation data set.

km ^{2}, square kilometer.

Overall, results shown in figures 34 through 36 indicate that values of the selected percentiles and the annual mean concentration generally varied by less than a factor of 10 from year to year at most of these stream stations during the period of sampling. This variability does not exceed the uncertainty in the model-estimated concentrations, implying that the models adequately predict concentrations for these stations during most years of the study period. The patterns of year-to-year variability described here are based on only a few years of data at most stations. In addition, most of the stations with more than 2 years of data were sampled primarily during the 1995 to 1999 period. The variability in concentration statistics derived from these data may not reflect the true long-term variability in atrazine concentrations at these stations, especially the variability in concentrations at the higher and lower ends of the frequency distribution. Longer-term monitoring at a larger number of stations would be needed to substantiate the results obtained for the stations in this study.

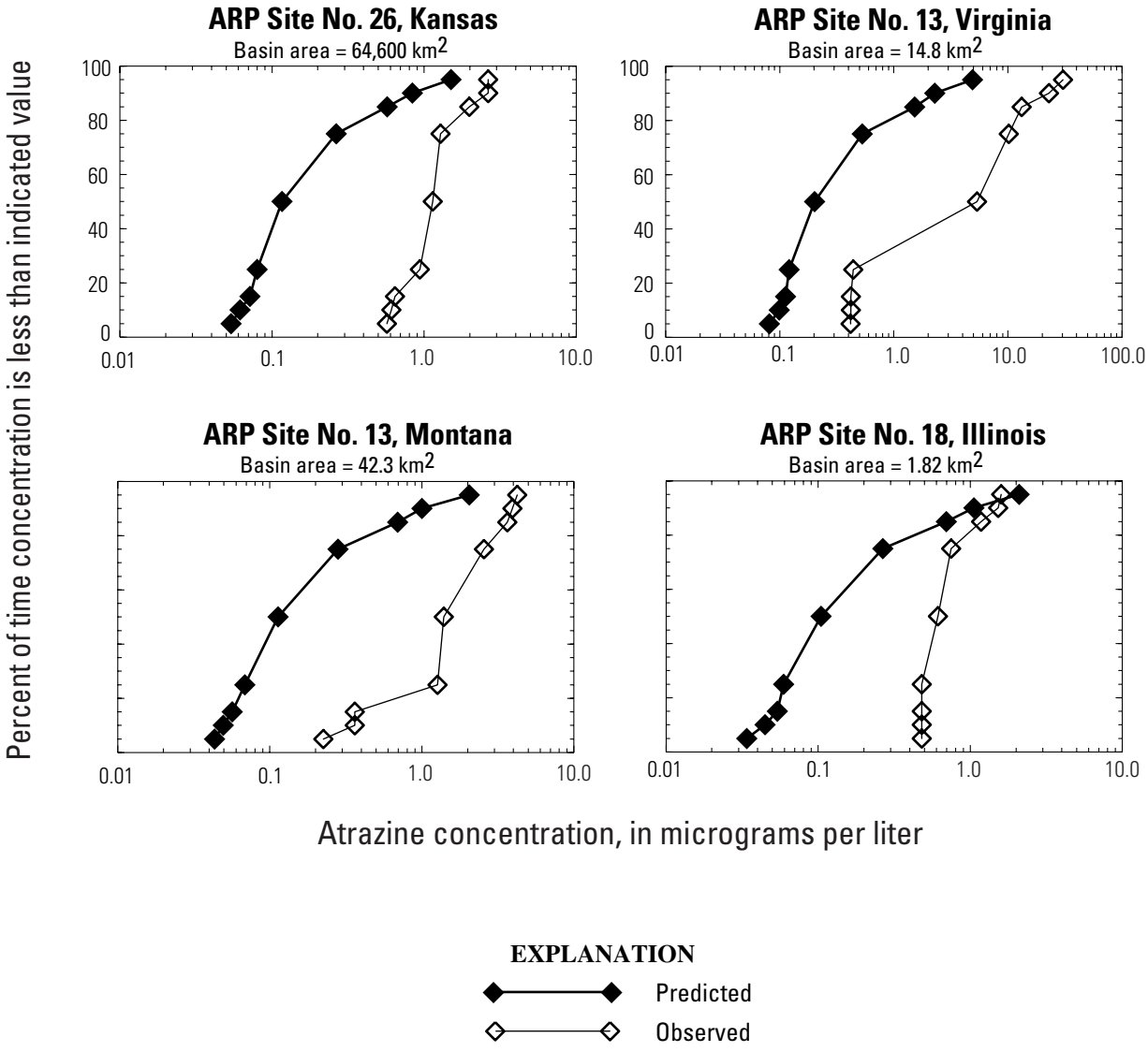


Figure 33. Examples of predicted and observed annual atrazine concentration distributions for four stations located on lakes or reservoirs. km ^{2}, square kilometer. ARP, Acetochlor Registration Partnership.

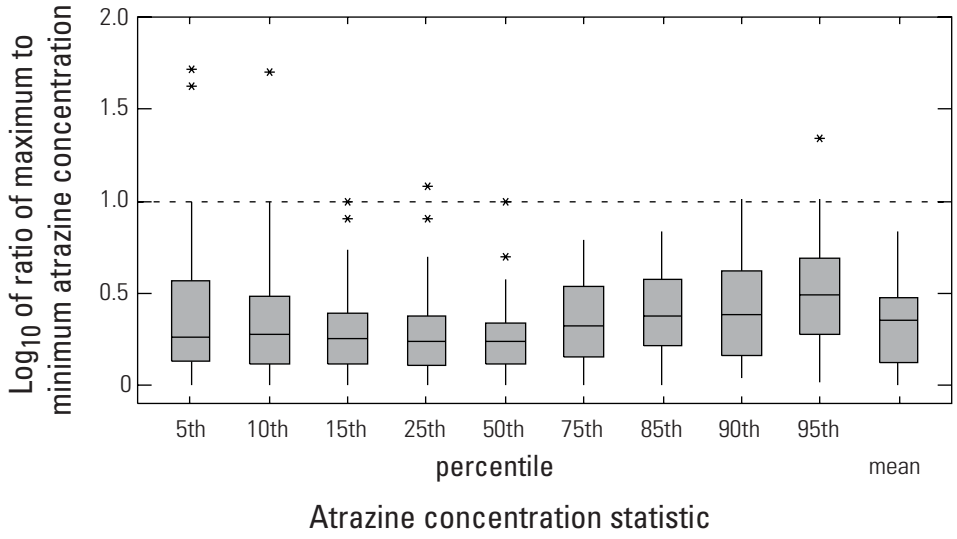


Figure 34. The ratio of the maximum and minimum observed values for nine atrazine concentration percentiles and the annual mean atrazine concentration for 54 stations with 2 or more years of computed statistics.

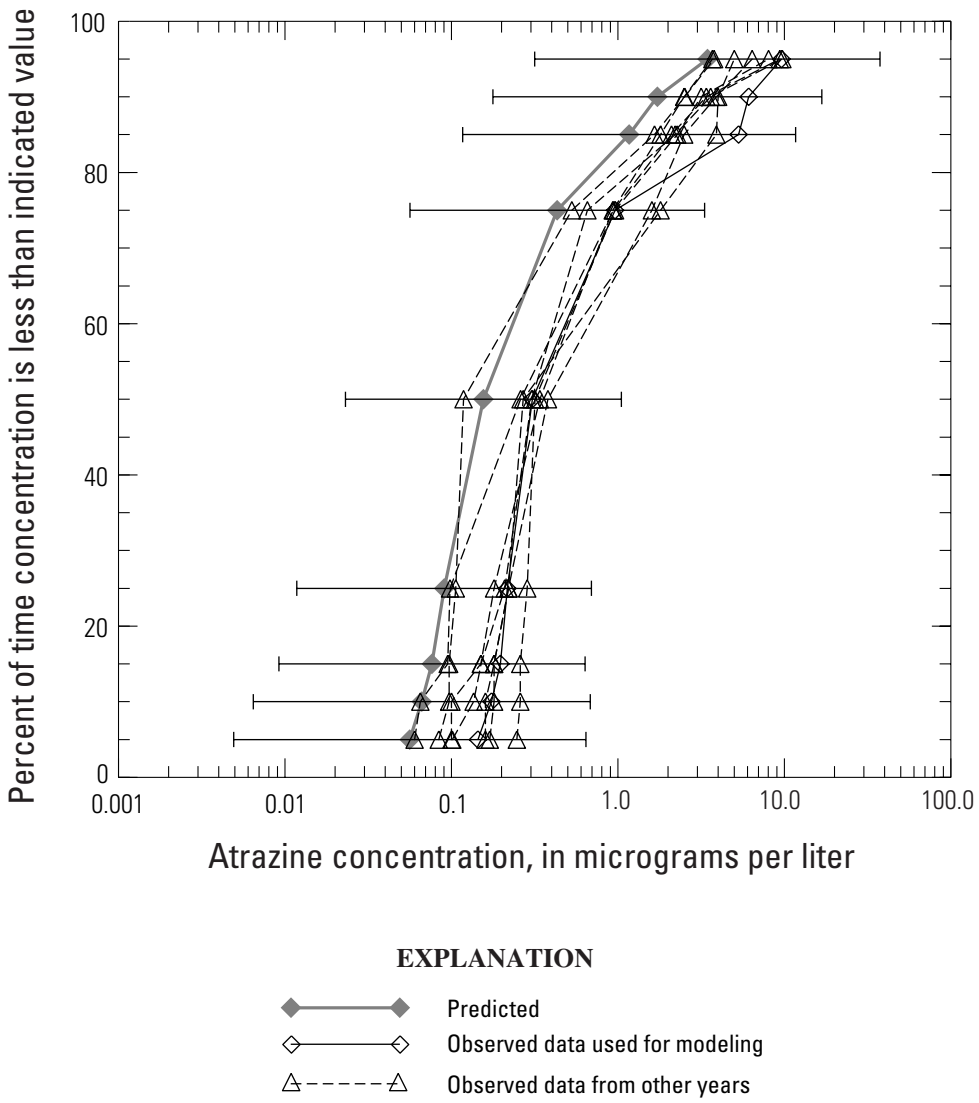


Figure 35. Observed annual atrazine concentration distributions for 8 years (1992–99) and predicted atrazine concentration distribution for the White River in Indiana.

Prediction intervals (95 percent) for each predicted concentration are shown as horizontal bars.

The WARP method does not directly provide estimates of daily concentrations. However, the WARP method may be applied to periods of less than a full year. Many pesticides, including atrazine, exhibit distinct seasonal patterns in rivers and streams. For atrazine, the seasonal pattern during a 1-year period can be divided into high and low seasons (fig. 37). The high season corresponds to the period following application to cropland and subsequent runoff to streams. The low season covers the rest of the year when concentrations decline to low or undetectable levels in most streams. Separate sets of regression models can be developed for each of these two seasons using the WARP method. Estimated concentration percentiles from the two sets of models then can be used to obtain estimates of the distribution of concentrations in a given stream during the two periods. Separate concentration distributions for high and low seasons provide a more realistic picture of the seasonal nature of atrazine occurrence in streams than annual distributions. The assignment of high and low seasons, while relatively straightforward for atrazine, may be more difficult for some other pesticides. Atrazine is applied to cropland once per year in most agricultural applications, and has little nonagricultural use in most parts of the United States. Many other pesticides have a number of different applications, and may be applied at various times during the year in some areas. Some insecticides, for example, may be applied in response to specific outbreaks of insects and at different times in different years.

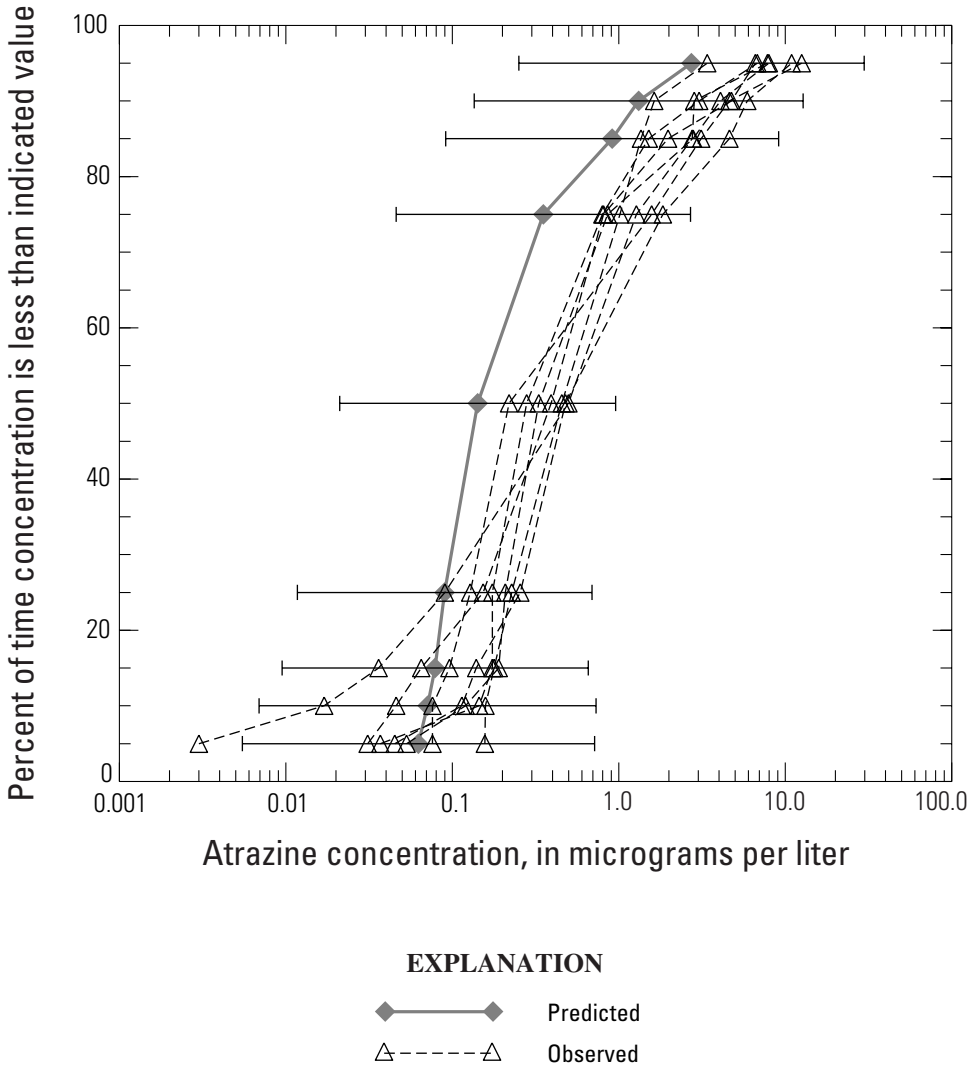


Figure 36. Observed annual atrazine concentration distributions for 7 years (1990–91, 1994–98) and predicted atrazine concentration distribution for the Sandusky River in Ohio.

Prediction intervals (95 percent) for each predicted concentration are shown as horizontal bars.

SEASONAL ATRAZINE MODELS

One limitation of regression models derived using the WARP method is the lack of a temporal component. The concentrations predicted for a water body are not directly related to a specific date. Model predictions of concentration percentiles provide only an estimate of the time during the year that the concentration of a pesticide in a stream remains below the predicted value. Assessment of the risk associated with pesticide occurrence in surface waters usually requires knowledge of both the magnitude and the timing of the occurrence. Currently, the USEPA is addressing requirements of the Food Quality Protection Act (FQPA) passed by the U.S. Congress in 1996 (U.S. Environmental Protection Agency, 2002b). The FQPA requires that concentrations of pesticides in drinking water be included in risk assessments used in setting tolerances (maximum legal limits) for pesticide residues in food. Ideally, daily values of pesticide concentrations would be available for use in these risk assessments for all water bodies used as sources of drinking water.

To evaluate the utility of seasonal regression models for atrazine, models for the high and low seasons were developed using the same data set used to develop the annual models. The high and low seasons were defined for individual stations based on the likelihood of atrazine being applied and runoff occurring during the period in a given region. The high season used for the seasonal models is the same as the high runoff period defined for each station (tables 1 and 2). The low season used for the seasonal models corresponds to the combined low and medium runoff period defined for each station. For example, for stations in Indiana (pesticide runoff group F in table 1), the high period extended from May through July and the low period from August through April. These classifications were subjectively determined for each region after consultation with NAWQA personnel familiar with the area.

Concentration percentiles and mean concentrations were computed for each period using the same methods used for the annual percentiles and mean concentrations. It should be noted, however, that the concentration percentiles computed for the high and low seasons represent different amounts of time than the percentiles computed for the entire year. The time represented by the seasonal concentration percentiles is determined by the length of time included in the high and low seasons for a given station. For example, for stations in Indiana (pesticide runoff group F in table 1) the high season includes the months of May, June, and July, or 92 days. Thus, the 95th percentile concentration computed for a station in Indiana represents the concentration equaled or exceeded on approximately 5 days during the high season [92 days − (0.95 × 92 days) = 4.6 days]. For stations in areas with more extended planting seasons, the same percentile represents a longer period. For example, for a station in Arizona (runoff group C in table 1), the high season covers the period March through October, or 245 days. The 95th percentile concentration computed for a station in this group represents the concentration equaled or exceeded on approximately 12 days during the high season. Thus, the high-season 95th percentile for a station in runoff group F represents an estimate of the 99th percentile concentration for the entire year [(365−4.6)/365 = 0.987]. For stations in group C, the high-season 95th percentile represents an estimate of the 97th percentile concentration for the entire year [(365−12)/365 = 0.967].

Seasonal regression models were developed using the same stepwise procedure used for the annual models. Although a number of different combinations of the explanatory variables were evaluated using this procedure, the variables used for the annual models were as good or better than any other group of variables considered, in terms of explaining the variance in the observed concentrations during both the high and low seasons. Thus, the high- and low-season models include the same variables as the annual models: (use intensity)$^{1/4}$, log (R-factor), K-factor, Dunne overland flow, and (watershed area)$^{1/2}$. The coefficients for the variables differ from the coefficients in the annual models, as the high-season regression equations were derived using concentration data for the high season only and the low-season equations were derived using concentration data from the low season only.

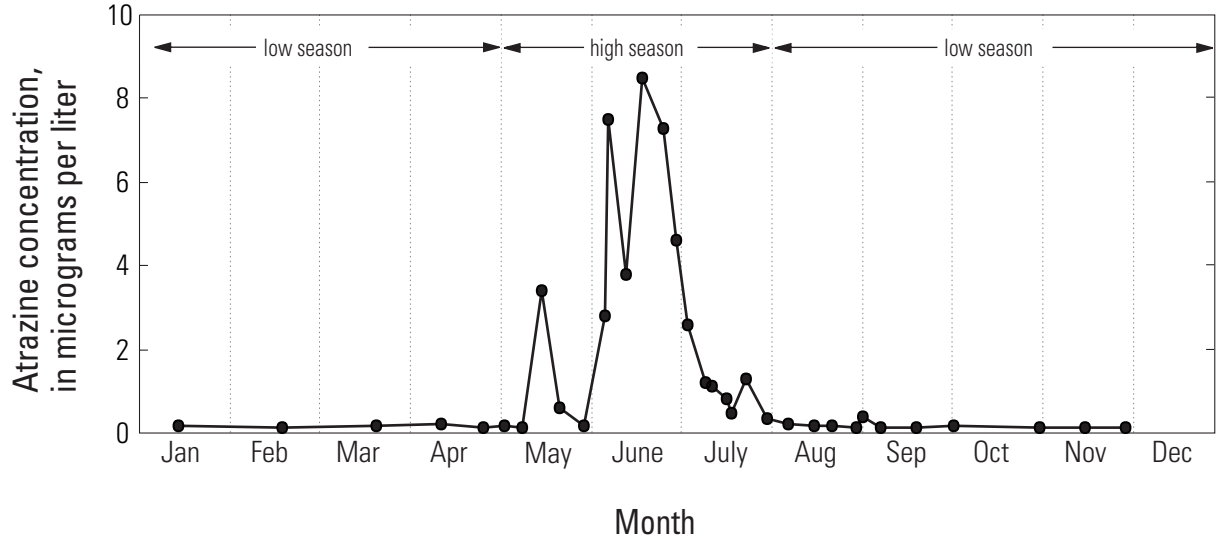


Figure 37. Example of the assignment of samples to seasons for use in seasonal atrazine models.

Statistics for the seasonal models are given in table 7. The high-season models explained 67 to 76 percent of the variability among the 112 stations (pR^2 values of 0.67–0.76). The variability explained by the high-season models for the mean and the higher percentiles was similar to that of the corresponding annual models (table 5). The high-season models for the lower percentiles explained slightly more of the variability than the corresponding annual models. The low-season models explained from 58 to 69 percent of the variability among the stations (table 7). The variability explained by the low-season models for the lower percentiles and the mean was similar to that of the corresponding annual models; the low-season models explained somewhat less of the variability for the higher percentiles.

As with the annual models, boxplots of the residual errors from the seasonal models can be used to examine the magnitude of the residual error of the models and the performance of the models with respect to watershed area and geographic region for the model development stations (figs. 38–41). These plots are analogous to figures 20 and 21, and show residual errors grouped by watershed area (figs. 38 and 39) and geographic region (figs. 40 and 41) for five of the high-season and five of the low-season models, respectively. A residual of zero indicates exact agreement; residuals of –1 and +1 indicate that the prediction for a given station is 10 times and one-tenth of the observed concentration, respectively. The regional groupings for figures 40 and 41 are the same ones used for figure 21, and are based on the U.S. Department of Agriculture Farm Resource Regions (U.S. Department of Agriculture, 2000). The nine Farm Resource Regions were consolidated into five groups (fig. 22) so that each group would have sufficient data to compute boxplots. Figures 38 through 41 indicate no substantial bias with respect to watershed area or region for the seasonal-model predictions. Prediction errors for the seasonal models are similar to errors for the annual models, with nearly all predicted concentrations within a factor of 10 of the observed concentration for the model development stations.

Uncertainty in the predictions from the seasonal models is evaluated in figures 42 and 43, using the same method that was used to evaluate uncertainty in the annual model predictions (fig. 28). Confidence intervals (CIs) for predictions from both the high- and low-season models extend less than a factor of two above and below the predicted concentration for most stations. The largest CIs extend to concentrations four to seven times the predicted concentrations. Prediction intervals (PIs) for predictions from both seasonal models extend to 10 to 18 times the predicted concentrations for most stations, with the largest PIs extending to more than 30 times the predicted concentrations.

Examples of concentration distributions estimated from the seasonal and annual models are shown in figure 44. The example plots show results for a very large river (Ohio River), two large rivers (Minnesota and New Rivers), and one small stream (Sugar Creek), all of which were included in the model development data set. Similar results were obtained for the other model development stations. Taken together, the two distributions derived from the seasonal models provide more information than the single distribution derived from the annual models, and better reflect the seasonal nature of atrazine occurrence in streams. For example, the estimated annual distribution for Sugar Creek implies that concentrations range from about 0.1 to 4 µg/L during about 90 percent of the year, and that the median (50th percentile) concentration for the year is approximately 0.15 µg/L. The combined estimates from the seasonal models imply that concentrations during the low season (August through April) cover approximately 0.08 to 0.3 µg/L, a very small range, and during the high season (May through July) approximately 0.2 to 10 µg/L, a much wider range. The median concentration during the low season is approximately 0.1 µg/L, very close to the annual median concentration, whereas the median concentration during the high season is much higher, approximately 1 µg/L.

Table 7. Statistics and coefficients for seasonal models for atrazine.

[Pseudo R-square, R-squared value for tobit regression; Scale, tobit regression analogue of the root mean squared error obtained from ordinary least squares regression; fuseint, fourth root of atrazine use intensity in the watershed; lrfact, $\log_{10}$(R-factor); kfact, K-factor; sdarea, square root of watershed area; perdun, percent of streamflow due to Dunne overland flow. <, less than]

Model	Regression coefficients (*p*-value)						Pseudo R-square	Scale	Percent censored observations
	Intercept	fuseint	lrfact	kfact	sdarea	perdun			
HIGH-SEASON MODELS									
5th	–4.80 (<0.001)	0.82 (<0.001)	0.75 (<0.001)	2.61 (0.002)	0.00085 (<0.001)	–0.15 (0.002)	0.67	0.60	17
10th	–4.63 (<0.001)	0.79 (<0.001)	0.70 (<0.001)	2.76 (<0.001)	0.00080 (<0.001)	–0.14 (0.003)	0.69	0.55	14
15th	–462 (<0.001)	0.76 (<0.001)	0.74 (<0.001)	2.96 (<0.001)	0.00085 (<0.001)	–0.14 (0.001)	0.71	0.52	12
25th	–4.59 (<0.001)	0.73 (<0.001)	0.75 (<0.001)	3.33 (<0.001)	0.00084 (<0.001)	–0.13 (0.004)	0.68	0.55	10
50th	–4.60 (<0.001)	0.80 (<0.001)	0.86 (<0.001)	3.19 (<0.001)	0.00081 (<0.001)	–0.14 (0.002)	0.71	0.55	9
75th	–4.62 (<0.001)	0.90 (<0.001)	0.94 (<0.001)	3.27 (<0.001)	0.00072 (0.004)	–0.14 (0.004)	0.73	0.58	7
85th	–4.43 (<0.001)	094 (<0.001)	0.88 (<0.001)	3.28 (<0.001)	0.00055 (0.034)	–0.12 (0.016)	0.71	0.61	5
90th	–4.25 (<0.001)	0.94 (<0.001)	0.86 (<0.001)	3.22 (<0.001)	0.00045 (0.077)	–0.12 (0.015)	0.72	0.60	4
95th	–4.14 (<0.001)	1.00 (<0.001)	0.88 (<0.001)	2.99 (<0.001)	0.00025 (0.031)	–0.13 (0.009)	0.75	0.58	2
mean	–4.98 (<0.001)	0.96 (<0.001)	0.96 (<0.001)	3.68 (<0.001)	0.00058 (0.017)	–0.13 (0.009)	0.76	0.57	14
LOW-SEASON MODELS									
5th	–5.13 (<0.001)	077 (<0.001)	0.78 (<0.001)	2.45 (0.005)	0.00049 (0.046)	–0.20 (<0.001)	0.66	0.56	34
10th	–4.78 (<0.001)	0.74 (<0.001)	0.57 (0.005)	2.76 (0.001)	0.00049 (0.046)	–0.16 (0.002)	0.62	0.57	31
15th	–4.62 (<0.001)	0.7 (<0.001)	0.62 (<0.001)	2.62 (<0.001)	0.0005 1 (0.014)	–0.18 (<0.001)	0.69	0.48	26
25th	–4.6 (<0.001)	0.63 (<0.001)	0.66 (<0.001)	2.85 (<0.001)	0.00053 (0.018)	–0.17 (<0.001)	0.63	0.52	23
50th	–4.66 (<0.001)	0.54 (<0.001)	0.76 (<0.001)	3.07 (<0.001)	0.00058 (0.005)	–0.12 (0.004)	0.64	0.49	16
75th	–4.33 (<0.001)	0.49 (<0.001)	0.76 (<0.001)	3.05 (<0.001)	0.00051 (0.011)	–0.14 (<0.001)	0.63	0.47	9
85th	–4.59 (<0.001)	0.42 (<0.001)	0.94 (<0.001)	3.63 (<0.001)	0.00063 (0.003)	–0.15 (<0.001)	0.61	0.50	7
90th	–4.43 (<0.001)	0.42 (<0.001)	0.92 (<0.001)	3.67 (<0.001)	0.00051 (0.017)	–0.16 (<0.001)	0.6	0.50	5
95th	–4.25 (<0.001)	0.4 (<0.001)	0.96 (<0.001)	3.64 (<0.001)	0.00047 (0.036)	–0.19 (<0.001)	0.58	0.52	3
mean	–5.37 (<0.001)	0.54 (<0.001)	1.00 (<0.001)	4.42 (<0.001)	0.00050 (0.019	–0.19 (<0.001)	0.66	0.49	31

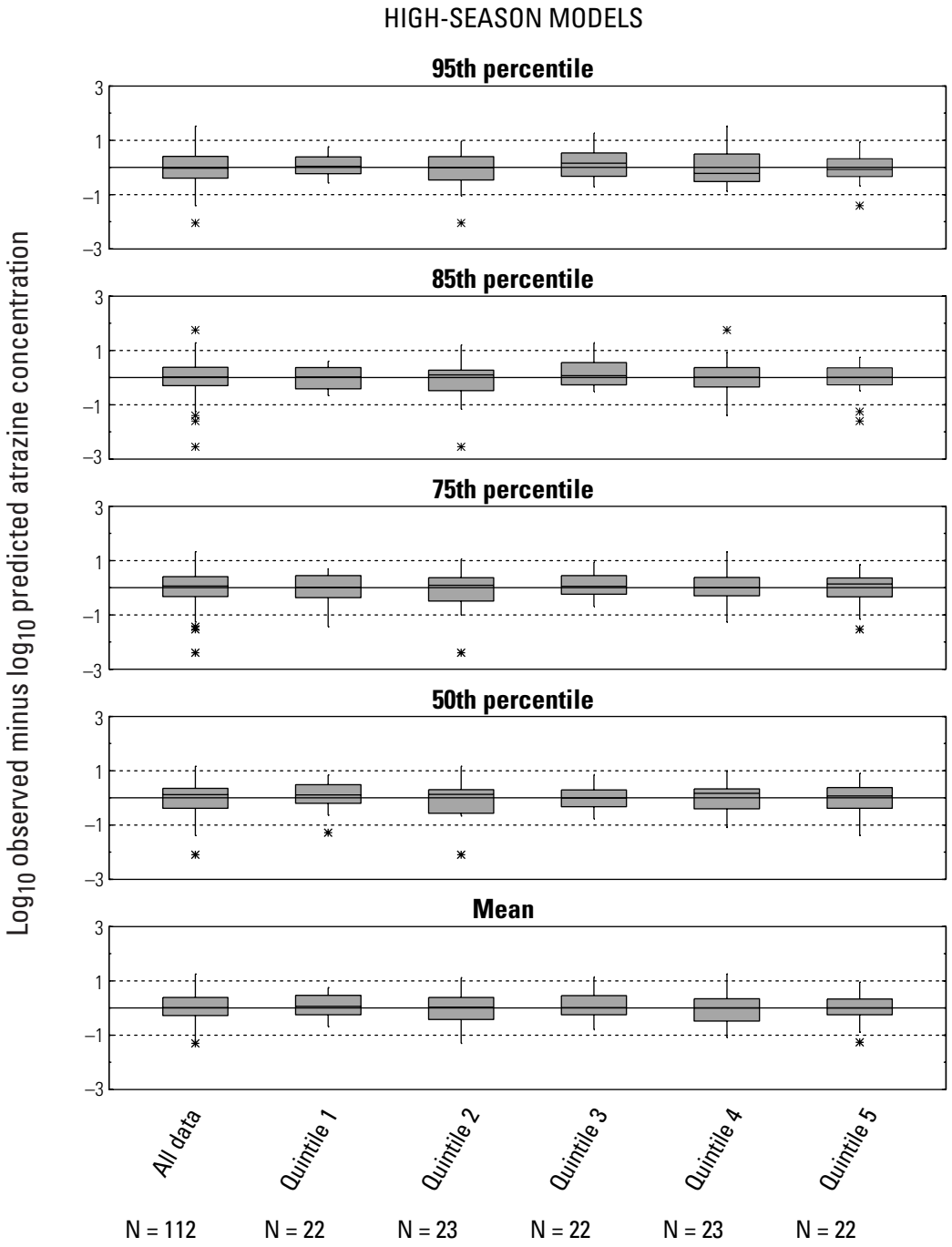


Figure 38. Residual error for five high-season atrazine models for 112 stations grouped into quintiles of watershed areas.

First boxplot in each group shows error for all 112 stations. Remaining boxplots show error for stations grouped into five classes based on watershed area quintiles. Residual error is [$\log_{10}$(observed value) — $\log_{10}$(predicted value)]. N, number of stations.

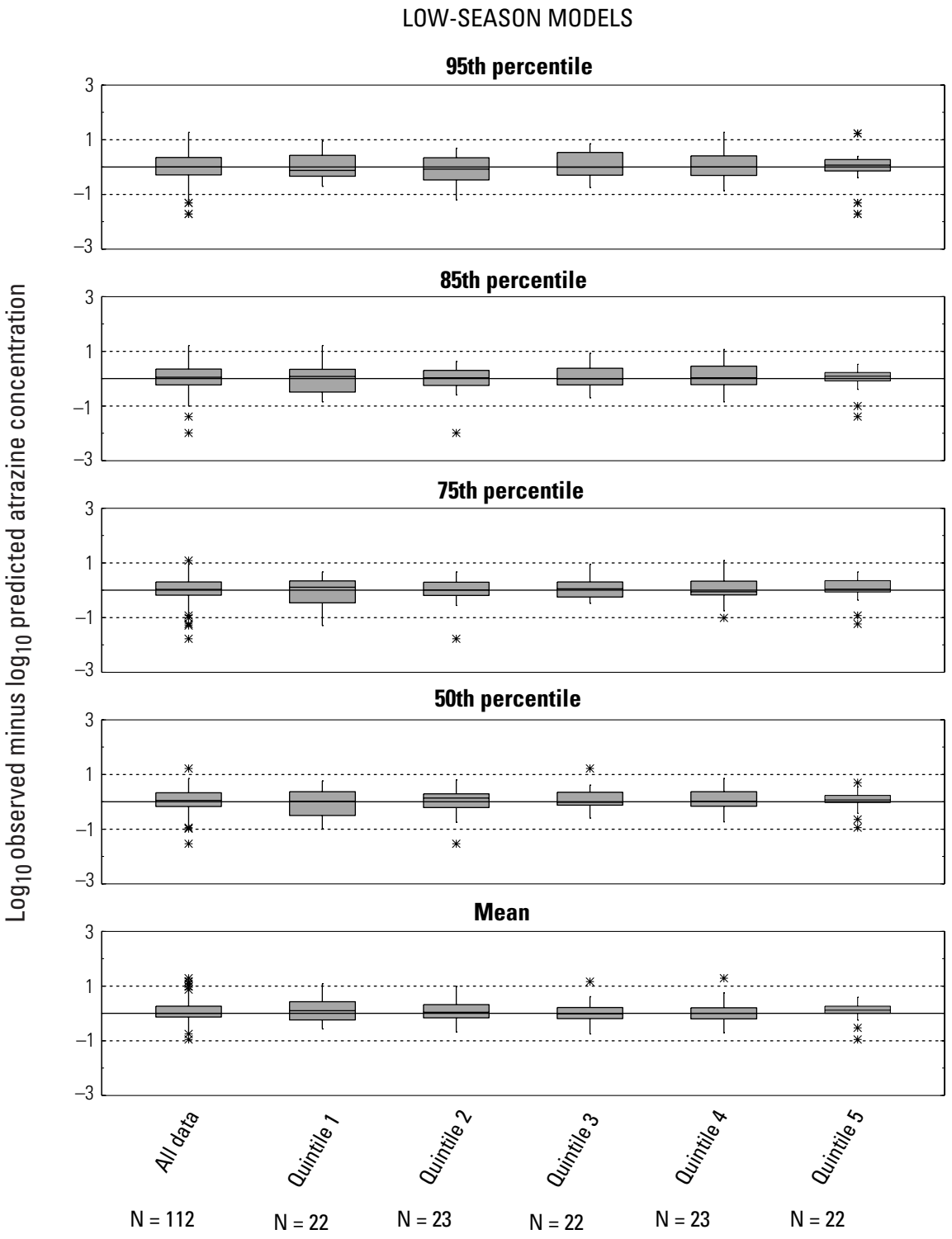


Figure 39. Residual error for five low-season atrazine models for 112 stations grouped into quintiles of watershed areas.

First boxplot in each group shows error for all 112 stations. Remaining boxplots show error for stations grouped into five classes based on watershed area quintiles. Residual error is [$\log_{10}$(observed value) — $\log_{10}$(predicted value)]. N, number of stations.

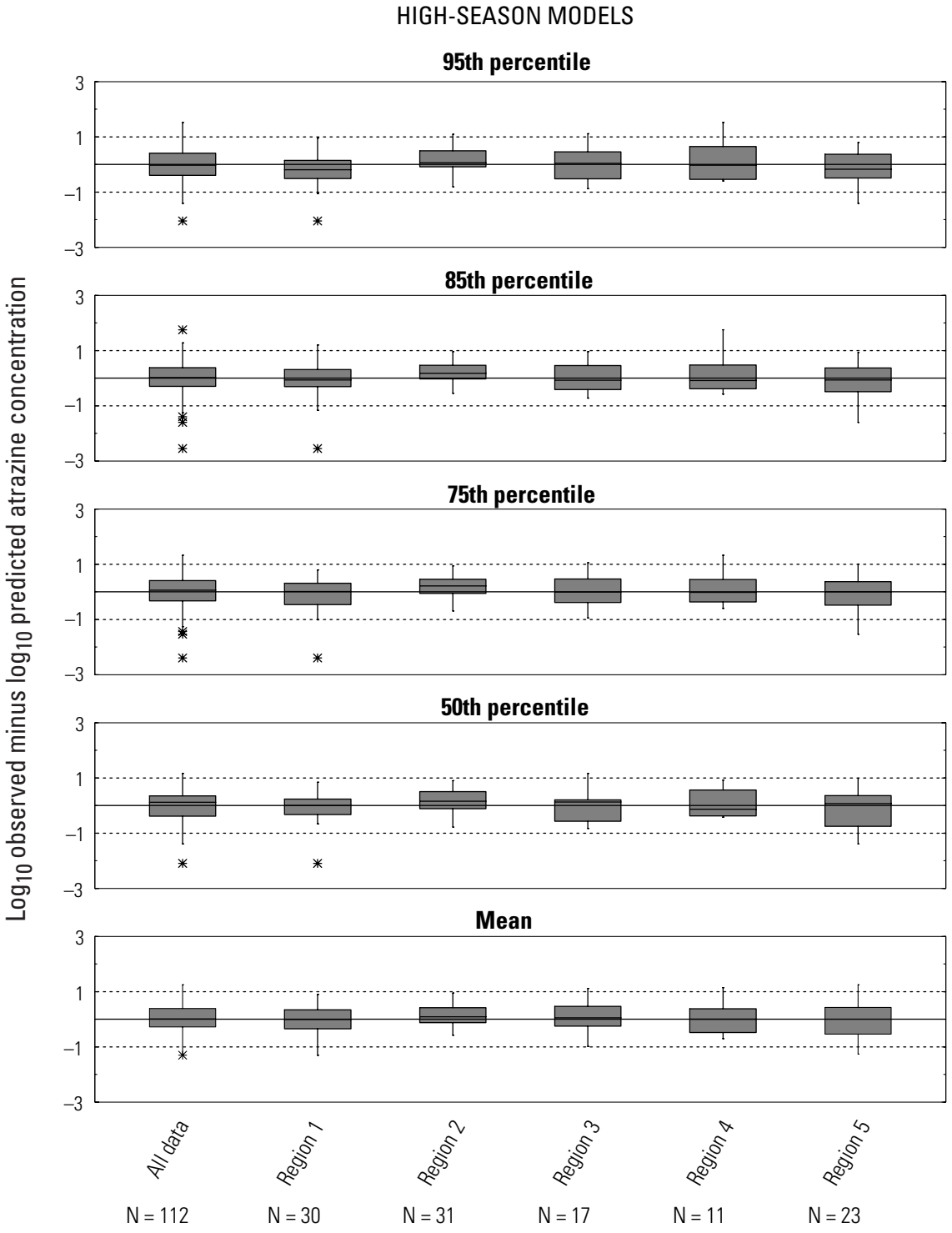


Figure 40. Residual error for five high-season atrazine models for 112 stations grouped by region.

First boxplot in each group shows error for all 112 stations. Remaining boxplots show error for stations grouped by region. Regions from figure 22. Residual error is [$\log_{10}$ (observed value) — $\log_{10}$ (predicted value)]. N, number of stations.

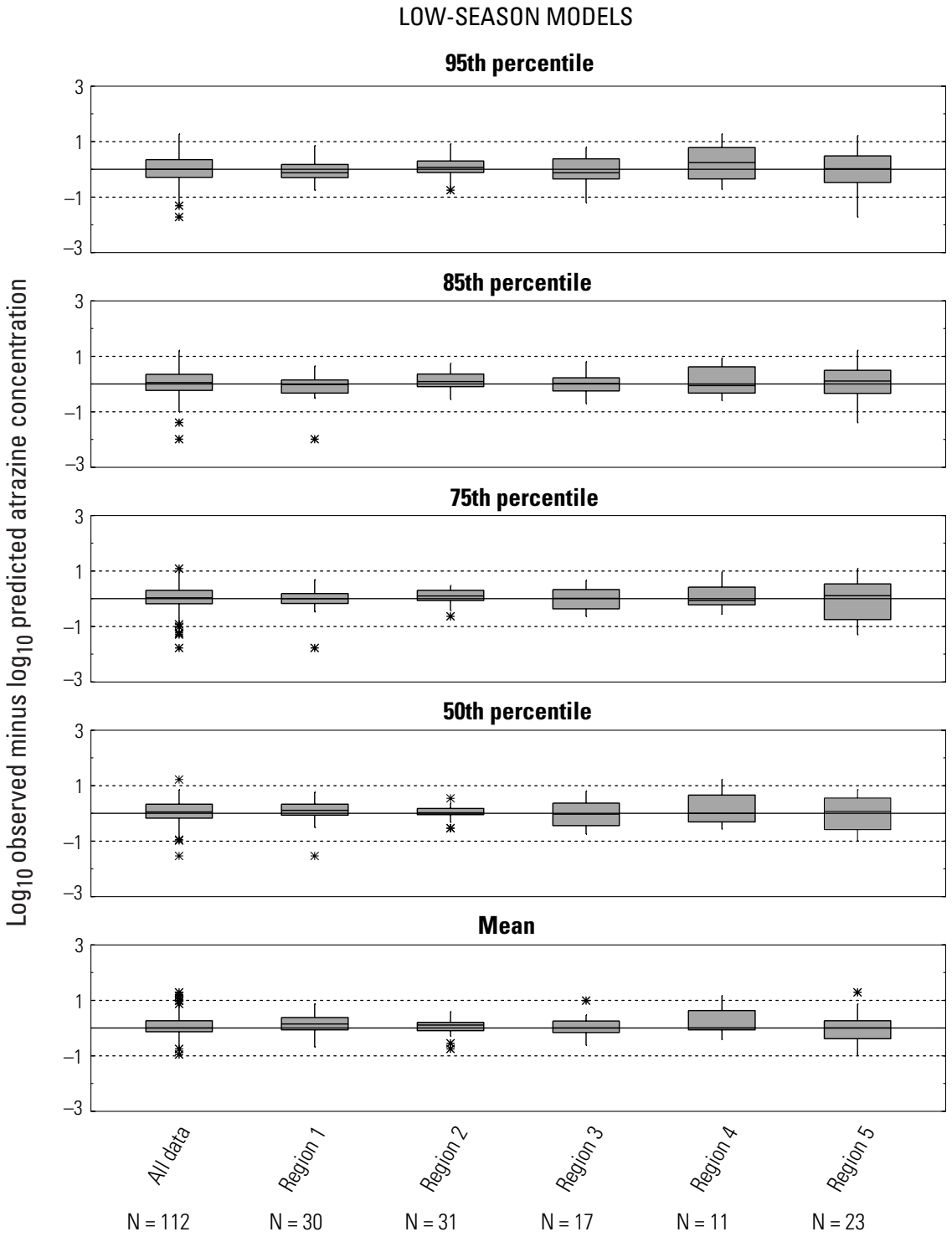


Figure 41. Residual error for five low-season atrazine models for 112 stations grouped by region.

First boxplot in each group shows error for all 112 stations. Remaining boxplots show error for stations grouped by region. Regions from figure 22. Residual error is [$\log_{10}$(observed value) − $\log_{10}$(predicted value)]. N, number of stations.

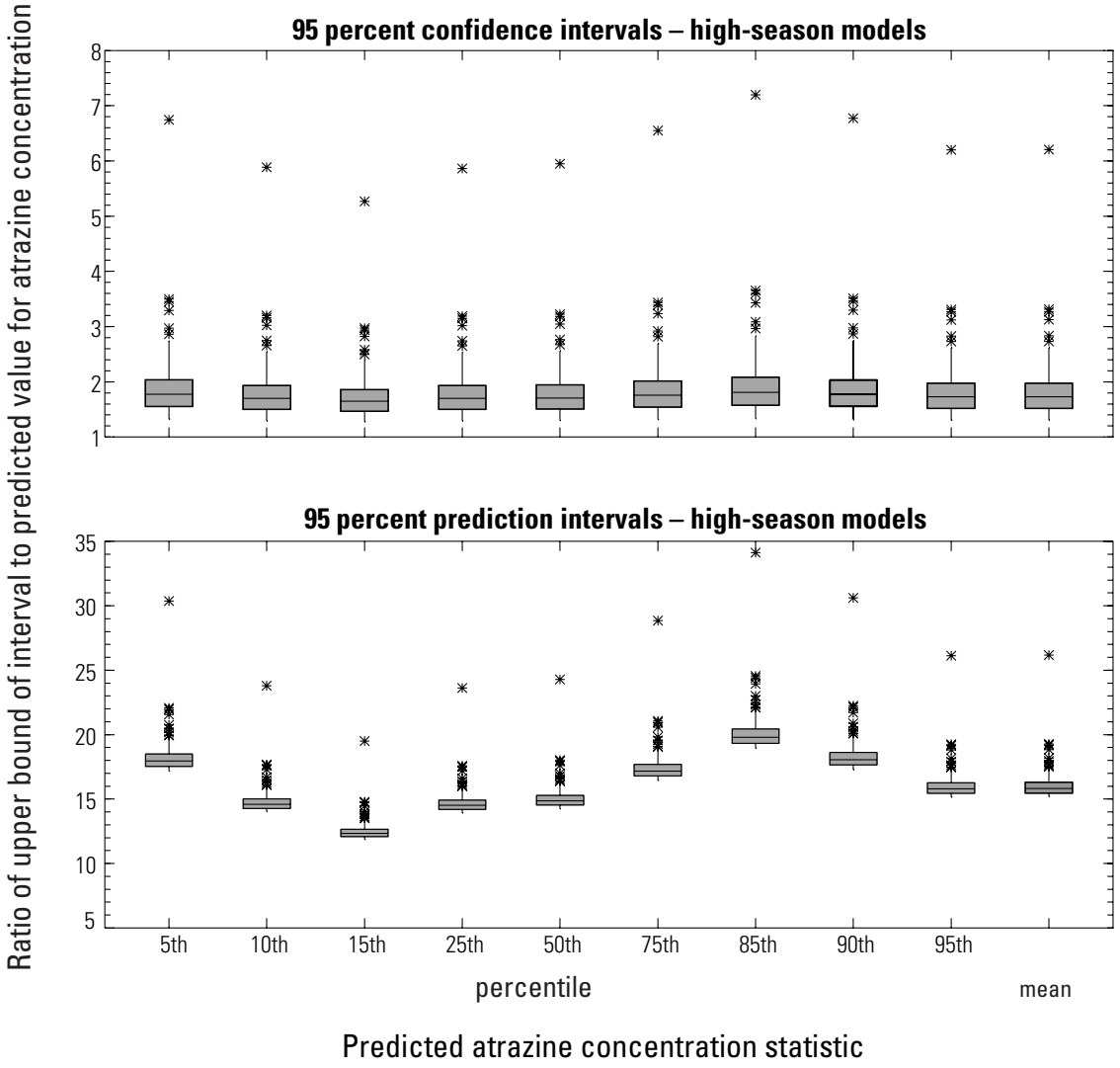


Figure 42. Ratios of the upper bound of 95 percent confidence and prediction intervals to predicted atrazine concentrations from 10 high-season models.

Each boxplot shows the distribution of ratios among the 112 model development stations.

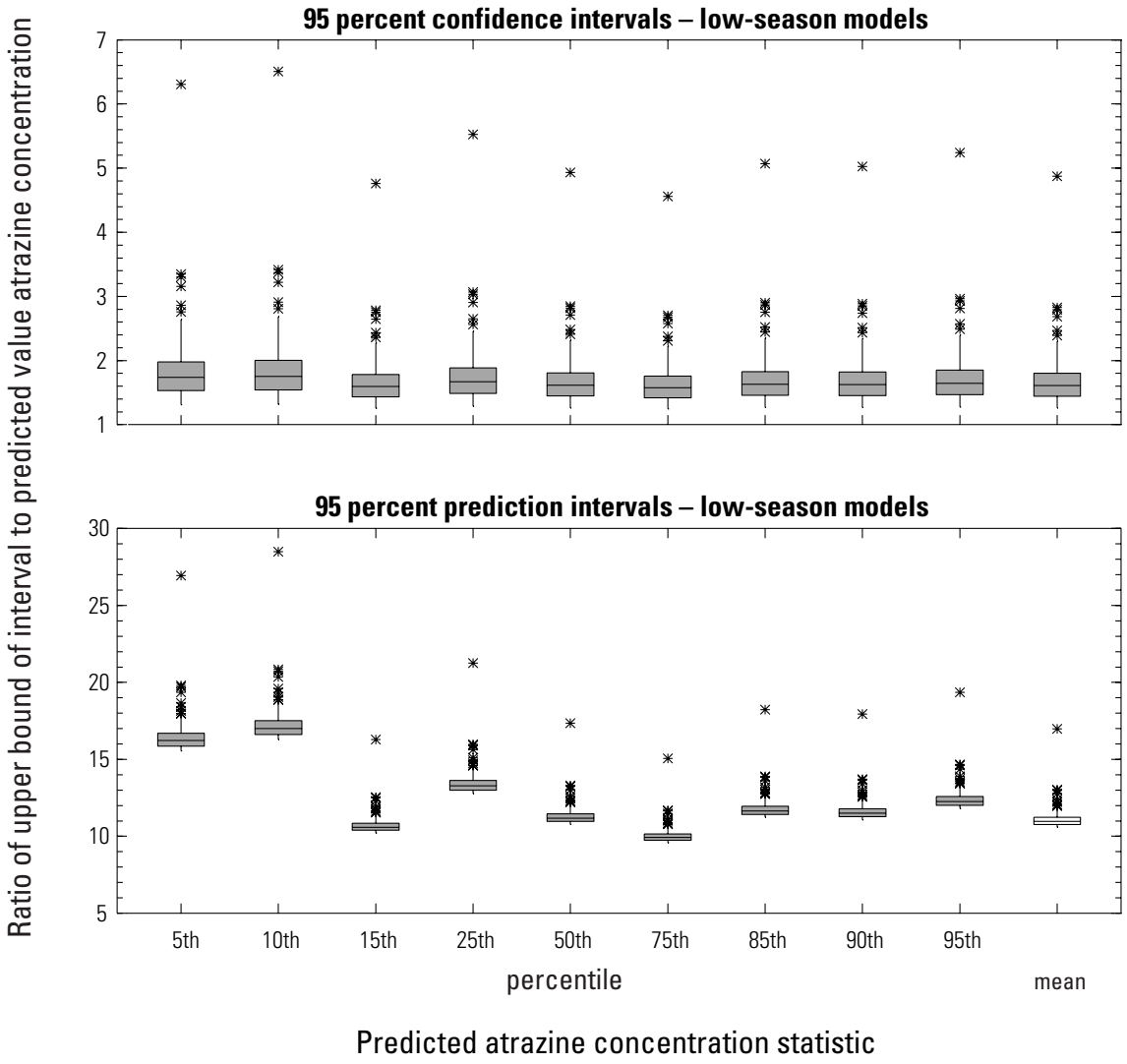


Figure 43. Ratios of the upper bound of 95 percent confidence and prediction intervals to predicted atrazine concentrations from 10 low-season models.

Each boxplot shows the distribution of ratios among the 112 model development stations.

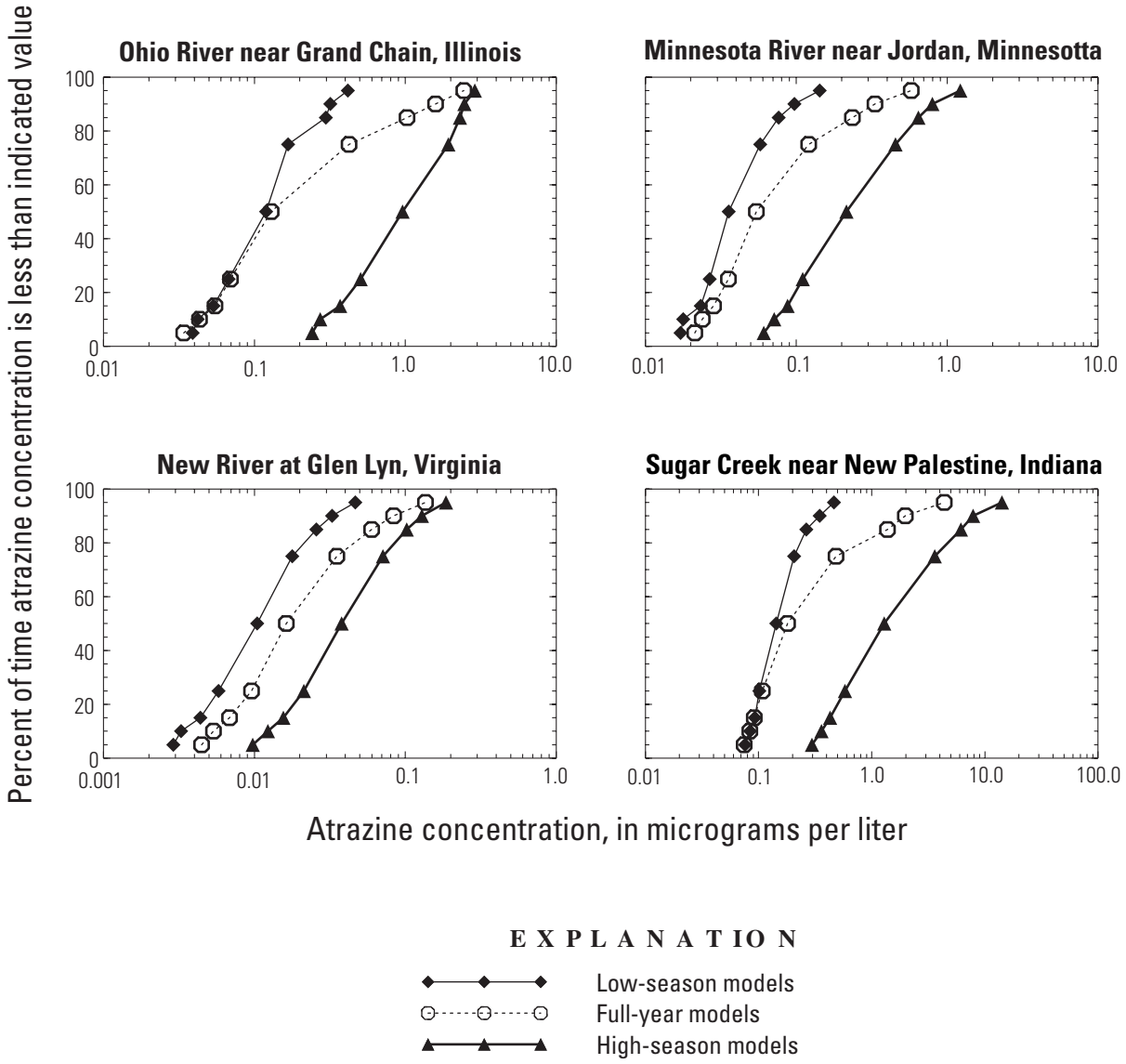


Figure 44. Examples of atrazine concentration distributions estimated from seasonal and annual models for a very large river (Ohio River), two large rivers (Minnesota and New Rivers), and a small stream (Sugar Creek).

MODEL LIMITATIONS

Use of the regression models for atrazine, and the WARP methodology in general, are subject to several limitations:

(1) The regression models were derived for estimation of atrazine concentrations in rivers and streams of the conterminous United States. While the 112 stations used for model development represent a wide variety of environmental settings and a large range of watershed area, it is likely that some watersheds in the conterminous United States or elsewhere have one or more characteristics outside the ranges of the watershed parameters used to develop the models. Application of the models to streams draining these watersheds would result in increased uncertainty in predicted concentrations and potentially biased results.

(2) The models were developed using data from flowing water systems. Application of the models to lakes or reservoirs may result in biased predictions, as shown by the results of application of the models to a small number of stations on lakes and reservoirs (fig. 25).

(3) The atrazine use data used for the models are estimates for agricultural applications to cropland only. Nonagricultural use of atrazine in a watershed, if significant, could result in inaccurate (low) predictions of atrazine concentrations in a stream. In addition, the use estimates are based on periodic summaries of data on land use, agricultural crops, and pesticide use, with a gap of 3 to 5 years between updates. Significant changes in farming practices or pesticide use in a watershed between updates of the data used for the use estimates could result in decreased accuracy of model predictions of atrazine concentrations. This would be more likely for streams draining small agricultural watersheds and watersheds undergoing changes in land use (for example, urban areas) than for large rivers.

(4) More generally, development of regression models for additional pesticides using the WARP methodology may not be straightforward in some cases. In several ways, atrazine may be the pesticide for which the WARP method is most likely to yield the best results. The widespread, primarily agricultural use and similar application methods used for atrazine across the United States result in a fairly predictable relation between atrazine use and concentrations in surface waters compared to many other pesticides. In addition, atrazine has been detected often in surface waters across the United States, and a large amount of atrazine concentration data is available for model development. The WARP methodology has been used to develop models for several other herbicides, all of which also have widespread, primarily agricultural use and frequent detections in surface waters (Larson and Gilliom, 2001). Extension of the WARP method to additional pesticides will require reliable use data and sufficient monitoring data for computing reliable concentration percentiles and annual mean concentrations. For pesticides with substantial nonagricultural use, a surrogate measure of use (for example, population density for pesticides used in urban areas) may be needed because data on nonagricultural use of pesticides are not currently available on a national scale.

SUMMARY AND CONCLUSIONS

Regression models were developed for estimation of annual distributions of atrazine concentrations in rivers and streams, using nationally available data on watershed characteristics and atrazine use. Separate models were derived for nine specific percentiles of the annual distribution of atrazine concentrations and for the annual mean concentration. Estimated concentrations from the models can be combined to provide an estimate of the annual distribution of atrazine concentrations in unmonitored streams. The models account for 62 to 77 percent of the variability in concentration percentiles among 112 streams used for model development. Atrazine use intensity in the watershed was the most important explanatory variable, accounting for 53 to 64 percent of the variability among the 112 streams. Four additional watershed characteristics were included in the models, accounting for an additional 8 to 16 percent of the variability. Uncertainty in predicted concentrations was expressed in terms of prediction intervals. For the annual models, 95 percent prediction intervals extend to a factor of 7 to 13 above and below the predicted concentration in most cases.

Results for 26 model validation stations on rivers and streams show that concentrations were predicted within a factor of 10 of the observed concentrations in nearly all cases. Concentration distributions for the validation stations were reasonably reproduced by the models. A relatively small number of stations were available for model validation, and many of the validation stations were located in the Corn Belt region of the midwestern United States. A larger number of validation stations, representing a wider range of environmental and agricultural settings, is needed for a more complete evaluation of the WARP atrazine models.

Predicted concentrations for 16 stations on lakes and reservoirs were biased low (underpredicted), especially for the lower percentiles. This bias was expected, and reflects known differences in the temporal patterns of atrazine concentrations in lakes and reservoirs and temporal patterns in flowing water systems. Regression models derived using concentration data from lakes and reservoirs may be needed to adequately model atrazine concentration distributions in these water bodies.

Year-to-year variability in annual atrazine concentration distributions and annual mean concentrations among the stations used for model development and validation was relatively small. At stream stations with multiyear data (54 stations), the variability in concentration percentiles and annual means for different years was usually less than an order of magnitude. Prediction intervals computed for model-estimated concentrations included observed concentrations in the additional years in most cases, implying that model predictions, along with the associated estimates of uncertainty, can reasonably be expected to represent concentrations in a stream for several years. Longer-term monitoring data from a larger number of stations would be needed to substantiate these results.

Seasonal models for atrazine concentration distributions in streams were developed to add a temporal component to the WARP models. Separate sets of models were developed for the high season (the application period and the period of potential runoff to streams) and the low season (the remainder of the year). The same explanatory variables included in the annual models were significant predictors in both sets of seasonal models. The high-season models accounted for 67 to 76 percent of the variability in high-season concentrations among the 112 model development stations. The low-season models accounted for 58 to 69 percent of the low-season variability. For most seasonal-model predictions, 95 percent prediction intervals extend to a factor of 10 to 18 above and below the predicted concentration. Comparison of estimated annual atrazine concentration distributions with estimated seasonal distributions shows that additional information can be obtained by developing seasonal models covering less than a full year, and that concentration distributions derived from the seasonal models better reflect the seasonal nature of atrazine occurrence in streams.

REFERENCES CITED

Acetochlor Registration Partnership, 2002, Acetochlor Registration Partnership (ARP) internet web site, accessed online May 1, 2002, at http://www.arpinfo.com.

Aitken, M., 1981, A note on the regression analysis of censored data: Technometrics, v. 23, no. 2, p. 161–163.

Amemiya, T., 1984, Tobit models—A survey: Journal of Econometrics, v. 24, p. 3–61.

Battaglin, W.A., and Goolsby, D.A., 1998, Regression models of herbicide concentrations in outflow from reservoirs in the midwestern USA, 1992–1993: Journal of the American Water Resources Association, v. 34, no. 6, p. 1369–1390.

Bradu, D., and Mundlak, Y., 1970, Estimation in lognormal linear models: Journal of the American Statistical Association, v. 65, no. 329, p. 198–211.

Brooks, K.N., Ffolliott, P.F., Gregersen, H.M., and Thames, J.L., 1991, Hydrology and the management of watersheds: Ames, Iowa, Iowa State University Press, 392 p.

Buckley, J., and James, I., 1979, Linear regression with censored data: Biometrika, v. 66, no. 3, p. 429–436.

Capel, P.D., and Larson, S.J., 2001, Effect of scale on the behavior of atrazine in surface waters: Environmental Science and Technology, v. 35, no. 4, p. 648–657.

Center for International Earth Science Information Network, 1996, Archive of census related products: Palisades, New York, Center for International Earth Science Information Network, Socioeconomic Data and Applications Center [also available at http://sedac.ciesin.org/plue/cenguide.html].

Chatterjee, S., and McLeish, D.L., 1986, Fitting linear regression models to censored data by least squares and maximum likelihood methods: Communications in Statistics—Theory and Methods, v. 15, no. 11, p. 3227–3243.

Coupe, R.H., and Goolsby, D.A., 1999, Monitoring the water quality of the nation's large rivers—Mississippi River Basin NASQAN Program: U.S. Geological Survey Fact Sheet FS-055-99, 4 p. [also available at http://water.usgs.gov/nasqan/missfact/msfact.html].

Daly, Christopher, Taylor, George, and Gibson,Wayne, 1997, The PRISM approach to mapping precipitation and temperature: in Proceedings of the Conference on Applied Climatology, Reno, Nevada, 1969, American Meteorological Society, p. 208–209.

Duan, N., 1983, Smearing estimate—A nonparametric retransformation method: Journal of the American Statistical Association, v. 78, no. 383, p. 605–610.

Duncan, G.M., 1986, A semi-parametric censored regression estimator: Journal of Econometrics, v. 32, p. 5–34.

Escobar, L.A., and Meeker, W.Q., Jr., 1992, Assessing influence in regression analysis with censored data: Biometrics, v. 48, p. 507–528.
Gebert, W.A., Graczyk, D.J., and Krug, W.R., 1987, Average annual runoff in the United States: U.S. Geological Survey Hydrologic Atlas 710.
Gianessi, L.P., and Anderson, J.E., 1995, Pesticide use in U.S. crop production: Washington, D.C., National Center for Food and Agricultural Policy, unpaginated.
Gianessi, L.P., and Marcelli, M.B., 2000, Pesticide use in U.S. crop production—1997, National Summary Report: Washington, D.C., National Center for Food and Agricultural Policy, unpaginated [also available at http://www.ncfap.org/ncfap/nationalsummary1997.pdf.
Gilliom, R.J., and Helsel, D.R., 1986. Estimation of distributional parameters for censored trace level water quality data 1—Estimation techniques: Water Resources Research, v. 22, no. 2, p. 135–146.
Greene, W.H., 2000, Econometric analysis (4th ed.): Upper Saddle River, N.J., Prentice-Hall, 1004 p.
Haas, C.N., and Jacangelo, J.G., 1993, Development of regression models with below-detection data: Journal of Environmental Engineering, v. 119, no. 2, p. 214–230.
Helsel, D.R., and Gilliom, R.J., 1986, Estimation of distributional parameters for censored trace level water quality data 2—Verification and applications: Water Resources Research, v. 22, no. 2, p. 147–155.
Hirsch, R.M., Alley, W.M., and Wilber, W.G., 1988, Concepts for a National Water-Quality Assessment Program: U.S. Geological Survey Circular 1021, 42 p.
Hooper, R.P., Goolsby, D.A., Rickert, D.A., and McKenzie, S.W., 1997, NASQAN—A program to monitor the water quality of the nations's rivers: U.S. Geological Survey Fact Sheet FS-055-97, 4 p.
Horowitz, J.L., 1986, A distribution-free least squares estimator for censored linear regression models: Journal of Econometrics, v. 32, p. 59–84.
Hughes, P.Y., Mason, E.H., Karl, T.R., and Brower, W.A., 1992, United States historical climatology network daily temperature and precipitation data: Oak Ridge, Tenn., Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Report ORNL/CDIAC-50, NPD-042, 40 p.
Insightful Corporation, 1999, S-PLUS 2000 user's guide: Seattle, Wash., Insightful Corporation, Data Analysis Products Division, 558 p.
Judge, G.G., Griffiths, W.E., Carter-Hill, R., Lütkepohl, H., and Lee, T.C., 1985, The theory and practice of econometrics (2nd. ed.): New York, John Wiley and Sons, 1019 p.
Laitila, T., 1993, A pseudo-R-square measure for limited and qualitative dependent variable models: Journal of Econometrics, v. 56, p. 341–356.
Larson, S.J., and Gilliom, R.J., 2001, Regression models for estimating herbicide concentrations in U.S. streams from watershed characteristics: Journal of the American Water Resources Association, v. 37, no. 5, p. 1349–1368.
Larson, S.J., Capel, P.D., and Majewski, M.S., 1997, Pesticides in surface waters—Distribution, trends, and governing factors: Ann Arbor Press, Chelsea, Mich., 373 p. [available from CRC Press].
Lindley, C.E., Stewart, J.T., and Sandstrom, M.W., 1996, Determination of low concentrations of acetochlor in water by automated solid phase extraction and gas chromatography with mass-selective detection: Journal of AOAC International, v. 79, no. 4, p. 962–966.
Maddala, G.S., 1983, Limited-dependent and qualitative variables in econometrics: Cambridge, United Kingdom, Cambridge University Press, 401 p.
Martin, J.D., 1999, Quality of pesticide data for environmental water samples collected for the National Water-Quality Assessment Program, 1992–96, and examples of the use of quality-control information in water-quality assessments: U.S. Geological Survey unpublished data, accessed October 1, 2002, at http://ca.water.usgs.gov/pnsp/rep/qcsummary/, updated 10/27/1999.
Martin, J.D., Gilliom, R.J., and Schertz, T.J., 1999, Summary and evaluation of pesticides in field blanks collected for the National Water-Quality Assessment Program: 1992–1995: U.S. Geological Survey Open-File Report 98-412, 102 p. [also available at http://ca.water.usgs.gov/pnsp/rep/ofr98412.pdf].
Miller, R.G., 1976, Least squares regression with censored data: Biometrika, v. 63, p. 449–464.
Miller, R., and Halpern, J., 1982, Regression with censored data: Biometrika, v. 69, no. 3, p. 521–531.

Mueller, D.K., Martin, J.D., and Lopes, T.J., 1997, Quality-control design for surface water sampling in the National Water-Quality Assessment Program: U.S. Geological Survey Open File Report 97-223, 17 p. [also available at http://water.usgs.gov/nawqa/protocols/OFR97-223/ofr97-223.pdf].

Myers, L., Riggs, M., Lashley, J., Whitmore, R., and Moya, J., 2000, Options for development of parametric probability distributions for exposure factors: U.S. Environmental Protection Agency Report EPA/600/R-00/058.

Natural Resources Conservation Service, 1994, State soil geographic (STATSG0) data base—Data use information: U.S. Department of Agriculture, National Soil Survey Center, Miscellaneous Publication Number 1492, 110 p.

Natural Resources Conservation Service, 2000, 1997 National resources inventory: U.S. Department of Agriculture, Natural Resources Conservation Service.

Neter, J., Wasserman, W., and Kutner, M.H., 1985, Applied linear statistical models (2nd. ed.): Homewood, Illinois, Irwin, 1127 p.

Owensby, J.R., and Ezell, D.S., 1992, Climatography of the United States No. 81—Monthly station normals of temperature, precipitation, and heating and cooling degree days, 1961–90: U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Climatic Data Center.

Powell, J.L., 1984, Least absolute deviations estimation for the censored regression model: Journal of Econometrics, v. 25, p. 303–325.

Richards, R.P., and Baker, D.B., 1993, Pesticide concentration patterns in agricultural drainage networks in the Lake Erie Basin: Environmental Toxicology and Chemistry, v. 12, p. 13–26.

Richards, R.P., Baker, D.B., Kramer, J.W., and Ewing, D.E., 1996, Annual loads of herbicides in Lake Erie tributaries in Ohio and Michigan: Journal of Great Lakes Research, v. 22, p. 414–428.

SAS Institute, 1999, SAS procedures guide, version 8, v 2: Cary, N.C., SAS Institute, p. 1317–1454.

Schneider, H., and Weissfeld, L., 1986, Estimation in linear models with censored data: Biometrika, v. 73, no. 3, p. 741–745.

Shelton, L.R., 1994, Field guide for collecting and processing streamwater samples for the National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 94-455, 42 p.

Thelin, G.P., and Gianessi, L.P., 2000, Method for estimating pesticide use for county areas of the conterminous United States: U.S. Geological Survey Open-File Report 00-250, 62 p. [also available at http://water.wr.usgs.gov/pnsp/rep/ofr00250/ofr00250.pdf].

Therneau, T.M., 1999, A package of survival routines for S: Rochester, MN, Mayo Clinic Department of Health Sciences Research, 83 p. [also available at http://www.mayo.edu/hsr/people/therneau.html]

Tobin, J., 1958, Estimation of relationships for limited dependent variables: Econometrics, v. 26, p. 24–36.

Tukey, J.W., 1977, Exploratory data analysis: Reading, Mass., Addison-Wesley, 506 p.

U.S. Department of Agriculture, 1999a, Pesticide data program—Annual summary calendar year 1999: U.S. Department of Agriculture, Agricultural Marketing Service, 112 p.

———1999b, 1997 Census of agriculture, volume 1, Geographic Area Series: U.S. Department of Agriculture.

———2000, Farm resource regions: U.S. Department of Agriculture, Economic Research Service, Agricultural Information Bulletin Number 760 [also available at http://www.ers.usda.gov/publications/aib760/aib-760.pdf].

———2001, Pesticide properties database: U. S. Department of Agriculture, Agricultural Research Service, accessed May 28, 2002, at wizard.arsusda.gov/acsl/ppdb.html (updated October 1, 2001).

U.S. Department of Commerce, 1995, 1992 Census of agriculture; Geographic area series 1A; U.S. summary state data file, and cross-tab data file: U.S. Department of Commerce, Bureau of the Census, Data User Services Division, CD-AG-92-1C [CD-ROM, dBase format].

U.S. Environmental Protection Agency, 2000, Guidance for data quality assessment—Practical methods for data analysis: U.S. Environmental Protection Agency, Office of Environmental Information Report EPA/600/R-96/084, 68 p.

———2002a, Acetochlor web page: U.S. Environmental Protection Agency, Office of Pesticide Programs, accessed May 1, 2002, at http://www.epa.gov/oppefed1/aceto/.

———2002b, The Food Quality Protection Act (FQPA) of 1996: U.S. Environmental Protection Agency, Office of Pesticide Programs, accessed October 7, 2002, at http://www.epa.gov/oppfead1/fqpa/ (updated April 30, 2002).

U.S. Geological Survey, 1997 to present, National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1–A9, 2 v., variously paged [also available at http://water.usgs.gov/owq/FieldManual/. Chapters originally were published from 1997–1999; updates and revisions are ongoing and are summarized at http://water.usgs.gov/owq/FieldManual/mastererrata.html].

———1999, The quality of the nation's waters—Nutrients and pesticides: U.S. Geological Survey Circular 1225, 82 p.

———2000, U.S. GeoData Digital Elevation Models: U.S. Geological Survey Fact Sheet FS-040-00, 2 p. [also available at http://mac.usgs.gov/mac/isb/pubs/factsheets/fs04000.pdf].

———2002, The USGS national stream quality accounting network: U.S. Geological Survey, accessed May 12, 2002, at http://water.usgs.gov/nasqan/.

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.

Wolock, D.W., 1993, Simulating the variable-source-area concept of streamflow generation with the watershed model TOPMODEL: U.S. Geological Survey Water- Resources Investigations Report 93-4124, 33 p.

Zaugg, S.D., Sandstrom, M.W., Smith, S.G., and Fehlberg, K.M., 1995, Methods of analysis by the U.S. Geological Survey National Water Quality Laboratory—determination of pesticides in water by C-18 solid-phase extraction and capillary-column gas chromatography/ mass spectrometry with selected-ion monitoring: U.S. Geological Survey Open-File Report 95-181, 49 p.