USGS
science for a changing world

National Water-Quality Assessment Program

Watershed Regressions for Pesticides (WARP) for Predicting Atrazine Concentration in Corn Belt Streams

Open-File Report 2011–1141

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

Watershed Regressions for Pesticides (WARP) for Predicting Atrazine Concentration in Corn Belt Streams

By Wesley W. Stone and Robert J. Gilliom

National Water-Quality Assessment Program

Open-File Report 2011–1141

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

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

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

U.S. Geological Survey, Reston, Virginia: 2011

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment, visit http://www.usgs.gov or call 1–888–ASK–USGS.

For an overview of USGS information products, including maps, imagery, and publications, visit http://www.usgs.gov/pubprod

To order this and other USGS information products, visit http://store.usgs.gov

Suggested citation:
Stone, W.W., and Gilliom, R.J., 2011, Watershed regressions for pesticides (WARP) for predicting atrazine concentration in Corn Belt streams: U.S. Geological Survey Open-File Report 2011–1141, 18 p.

Foreword

The U.S. Geological Survey (USGS) is committed to providing the Nation with reliable scientific information that helps to enhance and protect the overall quality of life and that facilitates effective management of water, biological, energy, and mineral resources (*http://www.usgs.gov/*). Information on the Nation's water resources is critical to ensuring long-term availability of water that is safe for drinking and recreation and is suitable for industry, irrigation, and fish and wildlife. Population growth and increasing demands for water make the availability of that water, measured in terms of quantity and quality, even more essential to the long-term sustainability of our communities and ecosystems.

The USGS implemented the National Water-Quality Assessment (NAWQA) Program in 1991 to support national, regional, State, and local information needs and decisions related to water-quality management and policy (*http://water.usgs.gov/nawqa*). The NAWQA Program is designed to answer: What is the quality of our Nation's streams and groundwater? How are conditions changing over time? How do natural features and human activities affect the quality of streams and groundwater, 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. From 1991 to 2001, the NAWQA Program completed interdisciplinary assessments and established a baseline understanding of water-quality conditions in 51 of the Nation's river basins and aquifers, referred to as Study Units (*http://water.usgs.gov/nawqa/studies/study_units.html*).

National and regional assessments are ongoing in the second decade (2001–2012) of the NAWQA Program as 42 of the 51 Study Units are selectively reassessed. These assessments extend the findings in the Study Units by determining water-quality status and trends at sites that have been consistently monitored for more than a decade, and filling critical gaps in characterizing the quality of surface water and groundwater. For example, increased emphasis has been placed on assessing the quality of source water and finished water associated with many of the Nation's largest community water systems. During the second decade, NAWQA is addressing five national priority topics that build an understanding of how natural features and human activities affect water quality, and establish links between sources of contaminants, the transport of those contaminants through the hydrologic system, and the potential effects of contaminants on humans and aquatic ecosystems. Included are studies on the fate of agricultural chemicals, effects of urbanization on stream ecosystems, bioaccumulation of mercury in stream ecosystems, effects of nutrient enrichment on aquatic ecosystems, and transport of contaminants to public-supply wells. In addition, national syntheses of information on pesticides, volatile organic compounds (VOCs), nutrients, trace elements, and aquatic ecology are continuing.

The USGS aims to disseminate credible, timely, and relevant science information to address practical and effective water-resource management and strategies that protect and restore water quality. We hope this NAWQA publication will provide you with insights and information to meet your needs and will foster increased citizen awareness and involvement in the protection and restoration of our Nation's waters.

The USGS 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 cost-effective management, regulation, and conservation of our Nation's water resources. The NAWQA Program, therefore, depends on advice and information from other agencies—Federal, State, regional, interstate, Tribal, and local—as well as nongovernmental organizations, industry, academia, and other stakeholder groups. Your assistance and suggestions are greatly appreciated.

William H. Werkheiser
USGS Associate Director for Water

Acknowledgments

Syngenta Crop Protection, Inc., and Waterborne Environmental, Inc., provided watershed boundaries used in development of the explanatory variable data for the AEEMP and AMP sites.

The authors thank Naomi Nakagaki, Jeffrey Martin, Gail Thelin, Michael Wieczorek, and David Wolock of the USGS for their assistance in preparing the data for analysis.

Contents

Figures

Tables

Conversion Factors

Multiply	By	To obtain
	Length	
centimeter (cm)	0.3937	inch (in)
millimeter (mm)	0.03937	inch (in)
meter (m)	3.281	foot (ft)
kilometer (km)	0.6214	mile (mi)
	Area	
square kilometer (km^2)	247.1	acre
square kilometer (km^2)	0.3861	square mile (mi^2)
	Volume	
liter (L)	0.2642	gallon (gal)
	Mass	
gram (g)	0.03527	ounce, avoirdupois (oz)
kilogram (kg)	2.205	pound avoirdupois (lb)
	Hydraulic conductivity	
micrometer per second (µm/s)	0.1417	inch per hour (in/h)

Concentrations of atrazine are given in micrograms per liter (µg/l).

Acronyms

AEEMP	Atrazine Ecological Effects Monitoring Program
AMP	Atrazine Monitoring Program
ARTDRN	percentage of the watershed that is artificially drained
CP	Mallow's Cp
MLR	multiple linear regression
NAWQA	National Water-Quality Assessment Program
NCWQR	National Center for Water Quality Research
PERHOR	percentage of streamflow due to Hortonian overland flow
PI	prediction interval
PMJN	total precipitation during May and June of the year of sampling
R^2	coefficient of multiple determination
RSE	residual standard error
SRL	soil restrictive layer
SRL25	percentage of the watershed with a soil restrictive layer within top 25 centimeters
SSURGO	Soil Survey Geographic database
UI	atrazine use intensity, the annual agricultural use in a watershed
USDA	U.S. Department of Agriculture
USEPA	U.S. Environmental Protection Agency
USGS	U.S. Geological Survey
WA	watershed area
WARP	Watershed Regression for Pesticides
WARP-CB	Watershed Regression for Pesticides for the Corn Belt region

Watershed Regressions for Pesticides (WARP) for Predicting Atrazine Concentration in Corn Belt Streams

By Wesley W. Stone and Robert J. Gilliom

Abstract

Watershed Regressions for Pesticides (WARP) models, previously developed for atrazine at the national scale, can be improved for application to the U.S. Corn Belt region by developing region-specific models that include important watershed characteristics that are influential in predicting atrazine concentration statistics within the Corn Belt. WARP models for the Corn Belt (WARP-CB) were developed for predicting annual maximum moving-average (14-, 21-, 30-, 60-, and 90-day durations) and annual 95th-percentile atrazine concentrations in streams of the Corn Belt region. All streams used in development of WARP-CB models drain watersheds with atrazine use intensity greater than 17 kilograms per square kilometer (kg/km^2). The WARP-CB models accounted for 53 to 62 percent of the variability in the various concentration statistics among the model-development sites. The 95-percent prediction intervals are well within a factor of 10 above and below the predicted concentration statistic. WARP-CB model predictions were within a factor of 5 of the observed concentration statistic for over 90 percent of the model-development sites. The WARP-CB residuals and uncertainty are lower than those of the National WARP model for the same sites. The WARP-CB models provide improved predictions of the probability of exceeding a specified criterion or benchmark for Corn Belt streams draining watersheds with high atrazine use intensities; however, National WARP models should be used for Corn Belt streams where atrazine use intensities are less than 17 kg/km^2 of watershed area.

Introduction

The Watershed Regressions for Pesticides (WARP) models use watershed characteristics as explanatory variables to predict pesticide concentration statistics in streams (Larson and Gilliom, 2001). For watersheds with inadequate direct measurement of pesticide concentrations, WARP model predictions can and have been used in the selection of vulnerable watersheds for more intensive study (U.S. Environmental Protection Agency, 2007). WARP models are based on empirical relations between pesticide concentrations observed at monitoring sites and selected watershed characteristics available nationally. Larson and others (2004) and Stone and Gilliom (2009) developed WARP models specifically for atrazine at the national scale, and these are referred to as “National WARP models” in subsequent discussion. The monitoring sites used to develop the National WARP models are distributed across the United States and cover a wide range of atrazine use intensity and watershed characteristics (Stone and Gilliom, 2009). The most influential variable in the National WARP models is use intensity, expressed as annual mass applied within a watershed divided by watershed area. Atrazine use intensity varies considerably across the United States (Gilliom and others, 2006), and the National WARP model-development sites reflect this wide range in atrazine use intensity. The National WARP models are well suited for guiding and planning of more intensive assessments of vulnerable watersheds at a national level. For some agricultural regions, however, there are opportunities for improving predictions by developing regional models that make use of region-specific data on stream concentrations and watershed characteristics to better represent the conditions in the region—in this case the Corn Belt.

The Corn Belt region of the United States includes all or parts of Iowa, Illinois, Indiana, Minnesota, South Dakota, Nebraska, Kansas, Missouri, Ohio, and Wisconsin (fig. 1). These states were the top 10 states in terms of corn-for-grain acres harvested and accounted for approximately 79 percent of the Nation’s corn-for-grain acreage in the 2007 (U.S. Department of Agriculture, 2007a). Most of the agricultural use of atrazine is associated with corn production, and measured stream concentrations of atrazine closely match the geographic distribution of corn cultivation (Gilliom and others, 2006). Use intensity is the most influential variable in the National WARP models, and atrazine use intensity within the Corn Belt is high and not widely variable (fig. 1). WARP models can be improved for application to the Corn Belt region by developing region-specific models that include important watershed characteristics that are influential in predicting atrazine concentration statistics within the Corn Belt. These new WARP models are called Corn Belt WARP (WARP-CB) and they reduce the uncertainty in predicted atrazine concentrations for high-use areas within the Corn Belt. The primary objectives of this report are to present WARP-CB models for annual maximum moving-average concentrations (14-, 21-, 30-, 60-, and 90-day durations) and annual 95th-percentile concentrations specific to the high-atrazine-use areas of the Corn Belt, evaluate the influence of explanatory variables on the predicted concentration statistics, and evaluate the accuracy and uncertainty associated with the predicted concentration statistics.

Methods

Methods used in this study included selection of stream monitoring sites for the models, calculation of atrazine-concentration statistics, and development and evaluation of regression models to predict atrazine-concentration statistics.

Site Selection

This study was limited to stream monitoring sites with atrazine concentration data suitable for estimating concentration statistics such as annual maximum moving averages (14-, 21-, 30-, 60-, and 90-day durations) and the annual 95th-percentile concentrations. Ideally, a candidate site would have been sampled daily, at least during runoff events throughout the high atrazine use period, in order to accurately characterize short-term concentrations such as the annual maximum 14- and 21-day moving average (U.S. Environmental Protection Agency, 2010b). However, sites with such high sampling frequencies are few. Sites were thus chosen to balance the objective of increasing the accuracy of calculated atrazine concentration statistics with the objective of retaining a large number of sites for model development and evaluation. The site selection criteria were based on the days between sampling dates during the period of annual high atrazine use for the Corn Belt, April 15 through July 30. The maximum moving-average and 95th-percentile concentrations are expected to occur during the high-atrazine-use period for the Corn Belt (Crawford, 2001; Crawford, 2004). Sites where the maximum interval between samples was 10 days or less between April 15 and July 30 were selected for use in model development. This 10-day maximum interval between sample dates ensures that at least two samples will be used in the computation of the 21-day moving-average concentrations during the high-use season. Lerch and others (2011) and U.S. Environmental Protection Agency (2007) both show that maximum moving-average concentrations calculated from data collected at regular intervals as short as every 4 days were biased low when compared to moving-average concentrations calculated from data collected more frequently and targeting runoff events. However, for the purposes of building regression models for the Corn Belt, not enough sites and data are available with such high-frequency targeted sampling as described in Lerch and others (2011) and, therefore, the regression models include sites likely to have low-biased estimates for short duration concentration statistics.

Data from sites sampled by the U.S. Geological Survey National Water-Quality Assessment (NAWQA) program, the National Center for Water Quality Research (NCWQR) of Heidelberg University, Tiffin, Ohio (National Center for Water Quality Research, 2009), the and the U.S. Environmental Protection Agency Atrazine Ecological Effects Monitoring Program (AEEMP) Midwestern Stream Monitoring project (U.S. Environmental Protection Agency, 2010a) were evaluated for site selection. In all, 44 sites met the selection criteria for 1 or more years of sample collection (fig. 1). Of these 44 sites, 2 had only 1 year of data meeting the selection criteria, whereas the other 42 sites all had more than 2 years of data collection that met the criteria. To expand the number of observations available for model development, 2 years of data were used for 42 of the 44 sites with sufficient data. Limiting the number of years used to 2 for each site prevents any one site from having more influence on model development because of having more data in the analysis. When more than 2 years of data met the selection criteria for a site, the 2 years with the highest number of samples were selected for use. The total number of site and year combinations available for model development was 86, including the 2 sites with only 1 year of data. Of the 86 site years selected, 73 of the site and year combinations also had at least 2 samples used to calculate the 14-day moving-average concentrations during the atrazine high-use season. WARP-CB models were not developed for the annual mean or lower annual percentiles because the sites selected for model development did not have sufficient sampling to adequately characterize the atrazine stream concentrations outside of the high-use season. Specifically, because the AEEMP site sampling strategy targeted only the high-use season, use of measured atrazine concentrations to calculate and predict annual means and lower annual percentile concentrations for these sites would produce estimates that may be substantially biased high.

Sites selected for model development include 37 AEEMP sites sampled between 2004 and 2007, 5 NCWQR sites sampled between 1995 and 1998, and 2 USGS NAWQA sites sampled between 1992 and 1998. The AEEMP sampling typically occurred during 4-day fixed intervals; however this fixed-interval sampling was augmented for some site and years with autosamplers that collected stream samples during runoff events (U.S. Environmental Protection Agency, 2007). For the AEEMP site and year selection, years with autosampler-collected data characterizing runoff events were chosen over years with only fixed-interval sampling.

Data used for model evaluation included 11 site and year combinations from the pool of selected AEEMP and NCWQR site years and 10 sites from the U.S. Environmental Protection Agency Atrazine Monitoring Program (AMP) for Community Water Systems (U.S. Environmental Protection Agency, 2011). The model-evaluation site and year combinations chosen from the pool of selected AEEMP and NCWQR site years represent different areas of the Corn Belt and are independent of the remaining 75 site and year combinations that were used to develop the regression models. The AMP sites represent rivers

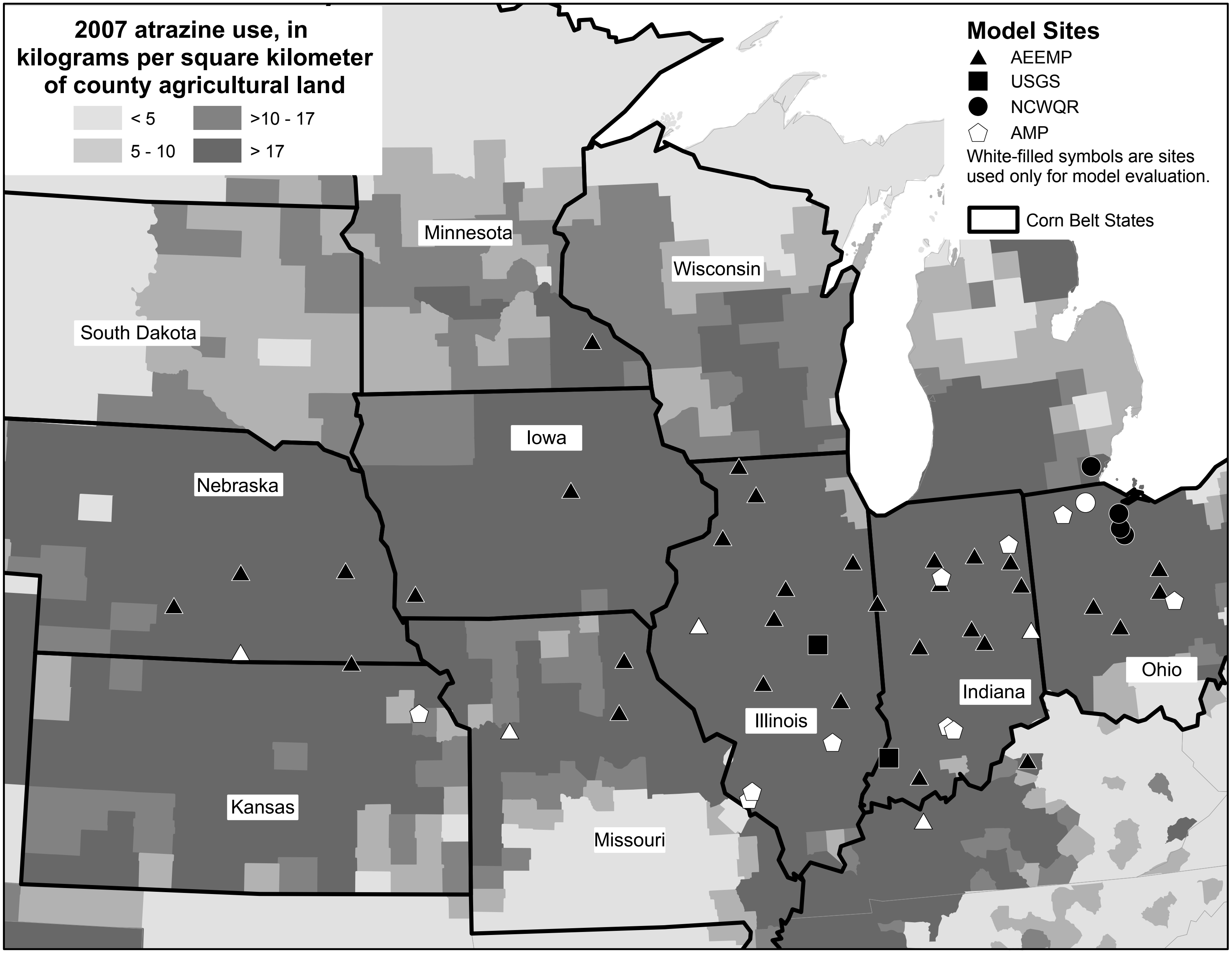

Figure 1. Location of sampling sites used for WARP-CB model development and evaluation. (2007 estimated atrazine use derived from Thelin and Stone [2010]; Abbreviations: AEEMP, U.S. Environmental Protection Agency Atrazine Ecological Effects Monitoring Program; USGS, U.S. Geological Survey; NCWQR, National Center for Water Quality Research; AMP, U.S. Environmental Protection Agency Atrazine Monitoring Program.)

and streams that were sampled weekly from April through July and biweekly the remainder of the year (U.S. Environmental Protection Agency, 2003). The AMP site data used in the model evaluation were collected during 2005 and 2006, and all the selected AMP sites had atrazine use intensity greater than 17 kg/km^2 of watershed area.

Calculation of Concentration Statistics

The annual maximum moving-average and 95th-percentile concentrations of atrazine were calculated by means of the methods described in Stone and others (2008). As described previously, the AEEMP data include results from fixed-interval sampling and samples collected by autosamplers during runoff events. For the AEEMP sites, both the fixed-interval samples and the autosampler data were used in the calculation of concentrations statistics. Hourly atrazine concentrations were estimated for the entire period of record (multiple years) for each site through linear interpolation of actual observations. The hourly concentration estimates were averaged to obtain an estimated daily concentration. The hourly estimates facilitated computations for days with multiple samples but were not used for other purposes. The estimated daily concentrations were used to estimate the annual 95th-percentile concentrations. Moving-average concentrations for the selected durations (14, 21, 30, 60, and 90 days) were computed for each day; for example, the 14-day moving average for a particular day includes the average of that day and the 13 previous days. The estimated moving-average concentrations then were truncated to the 1-year WARP development period for each site and year combination, and the maximum moving-average concentrations were determined for each of the durations. Calculated concentration statistics for the WARP-CB model sites are listed in appendix 1.

Watershed Characteristics Used as Explanatory Variables

The many watershed characteristics evaluated as potential explanatory variables in the WARP-CB model development are listed in table 1. Data and detailed descriptions of the potential explanatory variables used in model development are listed in appendixes 1 and 2. Potential explanatory variables representing pesticide use, land use and population, agricultural management practices, soil properties, physical watershed characteristics, weather and climate characteristics, and hydrological properties were considered.

The above-mentioned watershed characteristics are largely the same used by Stone and Gilliom (2009) for the National WARP models, with the notable exception of an additional variable representing the potential presence of a soil restrictive layer in the watershed. The presence of a shallow clay pan or low-permeability layer in the soils of a watershed has been associated with increased vulnerability for transport of pesticides to streams (Lerch and Blanchard, 2003; U.S. Environmental Protection Agency, 2007). High-potential-runoff soils may have a low-permeability soil layer and may also be associated with hydrologic group C and D soils (U.S. Department of Agriculture, 2007b). Past development of National WARP models included variables intended to characterize high-potential-runoff soils within watersheds, such as the K-factor (mean soil erodibility in the watershed) and the mean percentage of watershed soils classified as hydrologic soil groups C and D. The availability of Soil Survey Geographic (SSURGO) data for the Corn Belt region provides the opportunity to better characterize the extent of a low-permeability soil layer in the development site watersheds. For the purposes of this study, this soil layer of low permeability is called a Soil Restrictive Layer (SRL). The SRL variables were developed from SSURGO soil data tables (U.S. Department of Agriculture, 2010) and were based on three soil depth zones (0–25, 0–50, and 0–100 cm below land surface). For the purposes of this study, a saturated hydraulic conductivity of less than or equal to 1 µm/s was considered to indicate a layer of low permeability. The SSURGO data tables were evaluated to generate a data table of the percent presence for each soil-depth-zone-specific SRL within the SSURGO map units, based on the weighted average of the components of each SSURGO map unit. A 30-m-resolution raster image of the SSURGO map units for the Corn Belt was produced from the original 10-m-resolution raster image of the SSURGO map units. The data tables for the percent presence of the depth-zone-specific SRLs for the SSURGO map units was combined with the 30-m raster image of SSURGO map units to produce a 30-m-resolution raster image of SRL values. Watershed polygons for the WARP-CB sites were converted to 30-m-resolution raster images and overlain with the 30-m-resolution National Land Cover Dataset 2001 to produce a raster image of agricultural land within each watershed. The watershed raster image of agricultural land was overlain with the SRL raster image to determine the mean percentage of the watershed with a depth-zone-specific SRL.

Statistical Analysis

A regression approach was used in the development of the WARP-CB models. Statistical methods used to support the regression approach include transformations of response and explanatory variables, selection of explanatory variables, analysis of model fit, and estimation of prediction intervals.

The National WARP models were developed with maximum likelihood regression methods because of the presence of censored observations in the national model-development data. However, the WARP-CB model-development data do not contain censored observations; therefore, standard multiple linear regression (MLR) methods were used to develop the WARP-CB models.

Table 1. Watershed characteristics considered as explanatory variables for WARP-CB models.

[WARP-CB, Watershed Regression for Pesticides for Corn Belt region; USLE, Universal Soil Loss Equation; cm, centimeter; d, day; km, kilometer; km^2, square kilometer; m, meter; mm, millimeter; yr, year]

Variable	Description
	Pesticide use
UI	Annual atrazine agricultural-use intensity (kg/km^2).
	Land use and population
AG	Percent of basin with agricultural land use.
FOREST	Percent of basin with forest land use.
URBAN	Percent of basin with urban land use.
POPDEN	Mean 2000 population density in watershed (people/km^2).
	Agricultural management practices
ARTDRN	Percent of the watershed that is artificially drained.
CONTILL	Percent of the watershed with corn crop and conservation tillage.
REGTILL	Percent of the watershed with corn crop and regular tillage.
IRRI	Percent of the watershed that is irrigated.
TILE	Percent of the watershed that is drained by subsurface tiles.
	Soil properties
AWC	Mean available water capacity (fraction) in watershed.
CLAY	Mean percent clay in watershed soils.
HGAB	Mean percent of watershed soils classified as hydrologic soil groups A and B.
HGCD	Mean percent of watershed soils classified as hydrologic soil groups C, D, and C/D.
K	Mean soil erodibility of uppermost soil horizon in watershed (K-factor for USLE).
SRL25	Mean percent of agricultural area watershed soils with a soil restrictive layer (saturated hydraulic conductivity $\leq$ 1 µm/s) within top 25 cm of soil layer.
SRL50	Mean percent of agricultural area watershed soils with a soil restrictive layer (saturated hydraulic conductivity $\leq$ 1 µm/s) within top 50 cm of soil layer.
SRL100	Mean percent of agricultural area watershed soils with a soil restrictive layer (saturated hydraulic conductivity $\leq$ 1 µm/s) within top 100 cm of soil layer.
ORGM	Mean percent organic matter in watershed soils.
PERM	Mean soil permeability in the watershed (cm/h).
SAND	Mean percent sand in watershed soils.
SILT	Mean percent silt in watershed soils.
	Physical watershed characteristics
WA	Watershed area (km^2).
ELEV	Mean basin elevation (m).
LATC	Latitude of basin centroid (decimal degrees).
LONC	Longitude of basin centroid (decimal degrees).
SLOPE	Mean percent slope in watershed.
	Weather/climate characteristics
ADRY	Mean annual number of consecutive dry days, 1961–1990.
APPT	Mean annual 1961–1990 precipitation (cm/yr).
APPTI	Mean annual 1961–1990 precipitation intensity (mm/d).
ATEMP	Mean annual 1961–1990 temperature (C).
WET	Mean annual number of consecutive wet days, 1961–1990.
R	Mean annual 1971–2000 rainfall erosivity (R-factor for USLE).
PYEAR	Total precipitation during the year of sampling (mm).
PAPRJUN	Total precipitation during April, May, and June of the year of sampling (mm).

Table 1. Watershed characteristics considered as explanatory variables for WARP-CB models.—Continued

[WARP-CB, Watershed Regression for Pesticides for Corn Belt region; USLE, Universal Soil Loss Equation; cm, centimeter; d, day; km, kilometer; km^2, square kilometer; m, meter; mm, millimeter; yr, year]

Variable	Description
	Weather/climate characteristics (continued))
PAPRSEP	Total precipitation during April through September of the year of sampling (mm).
PMAY	Total precipitation during May of the year of sampling (mm).
PMJN	Total precipitation during May and June of the year of sampling (mm).
PMJL	Total precipitation during May, June, and July of the year of sampling (mm).
	Hydrologic properties
CONTACT	Mean subsurface contact time (days).
PERDUN	Percent of watershed streamflow contributed by saturation or Dunne overland flow.
PERHOR	Percent of watershed streamflow contributed by infiltration-excess or Horton overland flow.
PET	Mean potential evapotranspiration (cm).
ROFF	Mean annual 1951–1980 runoff (cm/yr).

Transformations of Response and Explanatory Variables

MLR models used for testing hypotheses and estimating confidence or prediction intervals require that the variance of the residuals be constant and that the residuals be independent and normally distributed (Helsel and Hirsch, 2002). Departures from these requirements can result in flawed estimates of model coefficients. A means of addressing possible departures from model assumptions is through transformation of either the response or the explanatory variables or both (Neter and others, 1985). Various transformations were considered to minimize departures from the requirements of the MLR method. The logarithm of concentration was used as the response variable throughout model development. For the explanatory variables, logarithmic, square-root, and third-root transformations, as well as the untransformed value, were considered during development of the regression models.

Because the response variable is a logarithmic transformation, concentrations predicted by the model (after retransformation) are the median concentrations expected for sites that have a given set of explanatory values, rather than the mean concentrations. Predicted concentration statistics were not adjusted for transformation bias because estimates of median values of the statistics were considered more appropriate for the purposes of this study.

Selection of Explanatory Variables

A stepwise approach was used for the initial selection of explanatory variables to include in the regression models. The StepAIC procedure of Venables and Ripley (1999), implemented in S-PLUS, was used for the initial selection of explanatory variables to include in the regression models. The StepAIC procedure, based on Akaike's Information Criterion (Akaike, 1974), balances model goodness of fit with the number of parameters needed to achieve that fit. The procedure attempts to quantify the concept of model parsimony by choosing simpler models over complex ones unless a complex model substantially improves the fit of the model.

Variable selection began with all of the potential explanatory variables and their transformations, too many to include in the StepAIC procedure at one time. Groups of 12 variables were randomly selected and then evaluated with the StepAIC procedure. The random selection and evaluation of variables was repeated 2,000 times. Variables selected greater than 50 percent of the time by the StepAIC procedure for all repetitions were retained in the analysis. When multiple transformations of a variable were retained by the StepAIC selection process, the transformed or untransformed value with the highest frequency of retention was carried forward in the analysis. Less than 50 percent of the potential explanatory variables were retained through this initial StepAIC process; however, this process resulted in the retention of far too many variables than practical for the final models. Final model selection used a combination of StepAIC, Mallow's Cp, multicollinearity evaluation, and scientific judgment. Mallow's Cp, like StepAIC, balances model goodness of fit with the number of parameters needed to achieve that fit. Models and their explanatory variables with the lowest Cp values were retained for further evaluation. Multicollinearity among explanatory variables was measured through use of variance inflation factors (Helsel and Hirsch, 2002). When explanatory variables had high variance inflation factors, they each were evaluated independently in the models. Finally, potential models were subjectively evaluated for reasonableness (for example, the models predict increasing concentration with increasing runoff) and their overall contribution to explaining the variation in atrazine-concentration statistics.

Analysis of Model Fit

Diagnostics used to evaluate the regression models included leverage, studentized residuals, Cook's D, and DFFITS (Helsel and Hirsch, 2002). Measures of goodness of fit, including the residual standard error (RSE) and the coefficient of multiple determination (R-square or R2) were also used in evaluation of the regression models. Box and whisker plots (Tukey, 1977) were used to qualitatively assess model uncertainty and residual errors.

Comparisons between predicted concentration statistics and concentration statistics computed from observations are made frequently in the discussion of model performance. Terms used for these comparisons are defined here for clarity. Concentration statistics generated by the WARP-CB models are referred to as "predicted concentration statistics," and concentration statistics computed from observations are referred to as "observed concentration statistics."

Estimation of Prediction Intervals

Uncertainty in the prediction of a concentration statistic can be expressed in terms of a prediction interval (PI) for a specified confidence level: the confidence level used in this study is 95 percent. Conceptually, each predicted concentration statistic is the median estimate of the particular concentration statistic (for example, percentiles or annual maximum moving averages) for all the stream sites that have the same combination of values for the explanatory variables. The PI is the range of values for a concentration statistic within which 95 percent of the actual concentration-statistic values are expected to occur for all stream sites with the same values of explanatory variables. In addition, the PI can be interpreted as the range within which the actual concentration statistic for an individual site and year is expected to fall 95 percent of the time. Prediction intervals were approximated by using normal theory and the t-distribution; that is, methods for ordinary least squares regression were used (Helsel and Hirsch, 2002).

Atrazine Models

The models for the 95th-percentile and the annual maximum moving-average atrazine-concentration statistics have

$$\log_{10}(\text{concentration}) = f\,[\text{SRL25}, \log_{10}(\text{PMJN}), (\text{PERHOR})^{1/2}, \log_{10}(\text{WA}), \text{ARTDRN}, \text{UI}] \quad (1)$$

where

SRL25 is the percentage of the watershed agricultural land with a soil restrictive layer within top 25 cm of soil surface;

PMJN is the total precipitation (mm) during May and June of the year of sampling;

PERHOR is the estimated percentage of streamflow due to Hortonian overland flow;

WA is the watershed area (km^2);

ARTDRN is the percentage of the watershed that is artificially drained; and

UI is atrazine-use intensity, the annual agricultural use in a watershed (kg) divided by watershed area (km^2).

Coefficients and statistics for all models are given in table 2, and summary statistics for the explanatory variables used in the WARP-CB models are given in table 3. Concentration statistics predicted by these models represent the expected median value of the concentration statistic for all sites with the same values for the explanatory variables.

Table 2. Summary of statistics and coefficients for the atrazine WARP-CB models.

[WARP-CB, Watershed Regression for Pesticides for Corn Belt region; RSE, residual standard error. Model variables are defined in table 1.]

Model	Regression coefficients (p-value)							R-square	RSE
	Intercept	SRL25	$\log_{10}$PMJN	$(\text{PERHOR})^{1/2}$	$\log_{10}$WA	ARTDRN	UI		
14-day	-3.21 (<0.001)	0.015 (<0.001)	1.01 (<0.001)	0.234 (<0.001)	0.244 (0.002)	0.005 (0.003)	0.0055 (0.158)	0.53	0.38
21-day	3.44 (<0.001)	0.015 (<0.001)	1.06 (<0.001)	0.241 (<0.001)	0.259 (<0.001)	0.005 (0.003)	0.0050 (0.173)	0.56	0.37
30-day	-3.45 (<0.001)	0.014 (<0.001)	1.04 (<0.001)	0.244 (<0.001)	0.252 (<0.001)	0.005 (0.001)	0.0045 (0.204)	0.58	0.35
60-day	-3.37 (<0.001)	0.014 (<0.001)	0.95 (<0.001)	0.222 (<0.001)	0.269 (<0.001)	0.005 (0.002)	0.0045 (0.156)	0.60	0.32
90-day	-3.27 (<0.001)	0.014 (<0.001)	0.87 (<0.001)	0.210 (<0.001)	0.269 (<0.001)	0.004 (0.002)	0.0047 (0.122)	0.61	0.30
95th-percentile	-3.25 (<0.001)	0.015 (<0.001)	0.88 (<0.001)	0.245 (<0.001)	0.261 (<0.001)	0.005 (<0.001)	0.0047 (0.134)	0.62	0.31

Table 3. Summary statistics for the WARP-CB explanatory variables.

[WARP-CB, Watershed Regressions for Pesticides for Corn Belt region; mm, milimeters; km^2, square kilometers; kg, kilograms; Model variables are defined in Table 1.]

Statistic	SRL25 (percent)	PMJN (mm)	PERHOR (percent)	WA (km^2)	ARTDRN (percent)	UI (kg/km^2)
Minimum	0.00	64.5	0.20	23	0.00	17.27
25th percentile	0.00	137.5	6.22	53	5.42	28.69
Median	0.09	191.5	8.00	64	30.93	38.21
Mean	6.37	197.4	8.63	1050	32.86	39.84
75th percentile	7.00	248.0	9.89	112	59.69	49.84
Maximum	92.04	348.7	28.7	29290	81.60	73.66

Analysis of Significant Explanatory Variables

SRL25 was highly significant ($p<0.001$) and the most important explanatory variable in the regression models (table 2). Models using only SRL25 as an explanatory variable account for 19 to 21 percent of the variability in concentration statistics among the development sites. SRL25 is an indication of the percentage of the watershed that may have a soil restrictive layer (saturated hydraulic conductivity less than or equal to 1 μm/s) within the top 25 cm of the soil column. Lerch and Blanchard (2003) found that the presence of a layer impeding the downward movement of water to groundwater was associated with higher vulnerability of a watershed for pesticide transport to streams. The coefficient for SRL25 is positive in all the models (table 2), indicating that watersheds with more area underlain by soils with a low-permeability layer have an increased potential for transport of atrazine from application areas to streams.

PMJN was highly significant ($p<0.001$) and the second most important explanatory variable in the regression models (table 2). Models using only PMJN as an explanatory variable account for 14 to 17 percent of the variability in concentration statistics among the development sites. PMJN is the total precipitation during May and June of the year of sampling and a significant explanatory variable in the National WARP models. For watersheds within the Corn Belt, the months April through June include the highest atrazine application period. High precipitation during May and June creates substantial surface runoff and increases the potential for recently applied atrazine loss to streams, which is consistent with the positive coefficient observed for PMJN in all models.

PERHOR was highly significant ($p<0.001$) and the third most important explanatory variable in the regression models (table 2). Models using only PERHOR as an explanatory variable account for 12 to 15 percent of the variability in concentration statistics among the development sites. PERHOR is the estimated percent Hortonian overland flow contribution to the stream. Hortonian overland flow is also called infiltration-excess overland flow and occurs when precipitation rates exceed infiltration rates. An area with high precipitation intensity and low soil permeability has a higher percentage of streamflow from Hortonian overland flow than an area with low precipitation intensity and high soil permeability. The positive coefficient of this variable in the WARP-CB models implies that areas with higher percentage of streamflow from Hortonian overland flow also have higher atrazine stream concentrations than those with lower contributions to streamfow from Hortonian overland flow.

WA (watershed area) was significant ($p<0.01$) and positive for all the WARP-CB models (table 2). Watershed area is also a significant explanatory variable (positive coefficient) in the National WARP models. Larson and Gilliom (2001) listed several factors that may contribute to the positive relation between watershed area and atrazine concentration in streams. For the WARP-CB model-development sites, two of the factors listed by Larson and Gilliom (2001) appear most relevant. First, the contribution of water from multiple small streams to a large river in a region dominated by agricultural land use can result in elevated concentrations that are sustained for a longer time than in the individual streams because the timing of pesticide application and subsequent runoff can vary among the individual streams. Second, in some small streams, the highest concentrations may not have been sampled because concentrations of pesticides may remain elevated for relatively short periods. The computed concentration statistics for these small streams may thus be biased low, further strengthening the positive relation between watershed area and concentration reflected in the coefficients for the watershed-area variable in the regression models. Lerch and others (2011) found that the moving-average concentration statistics calculated from AEEMP monitoring data for Goodwater Creek (MO-02) at an interval of every 4 days were biased low in comparison to more intensive samples collected during their study at the same site.

ARTDRN was significant ($p<0.01$) and positive in all models (table 2). ARTDRN is the percentage of the watershed that has artificial surface and subsurface drainage. ARTDRN, as used in the WARP-CB models, is defined by conservation practice categories 606, 607, and 608 in the 1992 Natural Resources Inventory (U.S. Department of Agriculture, 1995). Artificial drainage routes excess water to streams that may otherwise slowly percolate to groundwater or pond and eventually evaporate. Artificial drainage used in depressional areas or areas of low topographical relief without natural surface

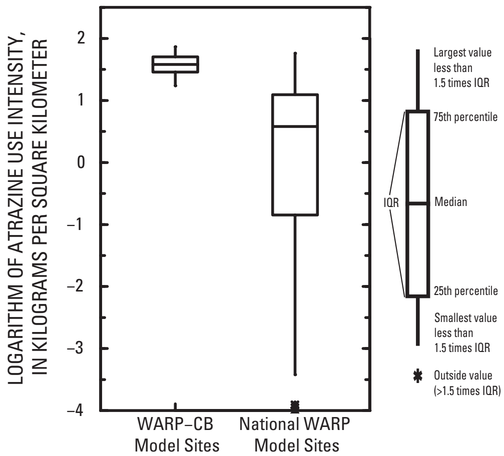

Figure 2. Atrazine use intensities for the WARP-CB and National WARP model-development sites.

runoff pathways will increase pesticide loss to streams because water that may naturally infiltrate to groundwater or evaporate will be transported to streams by the increased surface or subsurface drainage (Baker and others, 2007).

UI was not significant in any of the WARP-CB models (table 2). UI is the atrazine-use intensity in the watershed. Larson and others (2004) and Stone and Gilliom (2009) found that atrazine-use intensity in the watershed, as the fourth-root transformation, was the most important explanatory variable in the National WARP models. However, this is not true for WARP-CB models (table 2). The WARP-CB model-development sites represent a much higher and narrower distribution of atrazine-use intensity than the sites used in development of the National WARP models (fig. 2). Given the small variation in use intensity between the WARP-CB model-development sites, atrazine use intensity was not expected to be a significant explanatory variable for atrazine concentration statistics for the WARP-CB models. However, atrazine-use intensity is an important indicator for proper application of the WARP-CB and the National WARP models. Sites where atrazine use intensities are lower than those used in the development of the WARP-CB (atrazine use intensity less than 17 kg/km^2 watershed area) should be modeled with the National WARP models and not the WARP-CB models.

Model Performance

The performance of WARP-CB models was evaluated by assessment of (1) goodness of fit and model uncertainty, (2) performance of the model with the model-development and evaluation sites, and (3) comparison to the National WARP models applied to WARP-CB development sites.

Regression-model results for the annual maximum moving-average (14-, 21-, 30-, 60-, 90-day durations) and annual 95th-percentile concentration WARP-CB models are shown in figure 3. Values of R2 ranged from 0.53 for the 14-day moving-average model to 0.62 for the 95th-percentile model, meaning that the models accounted for 53 to 62 percent of the variability in the concentration statistics among the 75 site and year combinations used for model development. All of the predicted concentration statistics are within a factor of 10 of the observed concentration statistics at the development and evaluation sites. WARP-CB model predictions for development sites were within a factor of 5 of the observed concentration statistic for more than 90 percent of the site and year combinations (table 4). Regression-model results for the evaluation sites in relation to the model-development sites also are shown in figure 3. The WARP-CB models do not show substantial bias for either the development or the evaluation sites.

The PIs for the WARP-CB models are shown in figure 4 (model-development site years were arranged and numbered by order of increasing predicted 95th percentile concentration statistic). Concentration statistics are expressed as logarithms, resulting in symmetrical intervals for the plotted PIs; the high and low bounds of the intervals are the same distance from the predicted value. However, expressing the concentration statistics as logarithms obscures the fact that the intervals are skewed—the upper part of the PI interval covers a wider range of values than the lower part. Comparison of the PIs among the predicted concentration statistics shows that the PIs are largest for the annual maximum 14-day moving-average model and smallest for the annual maximum 90-day moving-average model (fig. 4). The PIs for all models are within a factor of 10 of the observed concentration statistics. The PI is a function of the fit of the model and the amount of variability explained by the model. The annual maximum 90-day

Table 4. Summary of performance of the WARP-CB and National models with the 75 WARP-CB model-development site years.

Model	WARP-CB models			National WARP models		
	Percent of predictions within a factor of			Percent of predictions within a factor of		
	10	5	2	10	5	2
14-day	100	93	57	--	--	--
21-day	100	95	60	92	83	45
30-day	100	96	64	--	--	--
60-day	100	97	72	94	88	39
90-day	100	100	75	96	88	43
95th-percentile	100	97	71	96	89	51

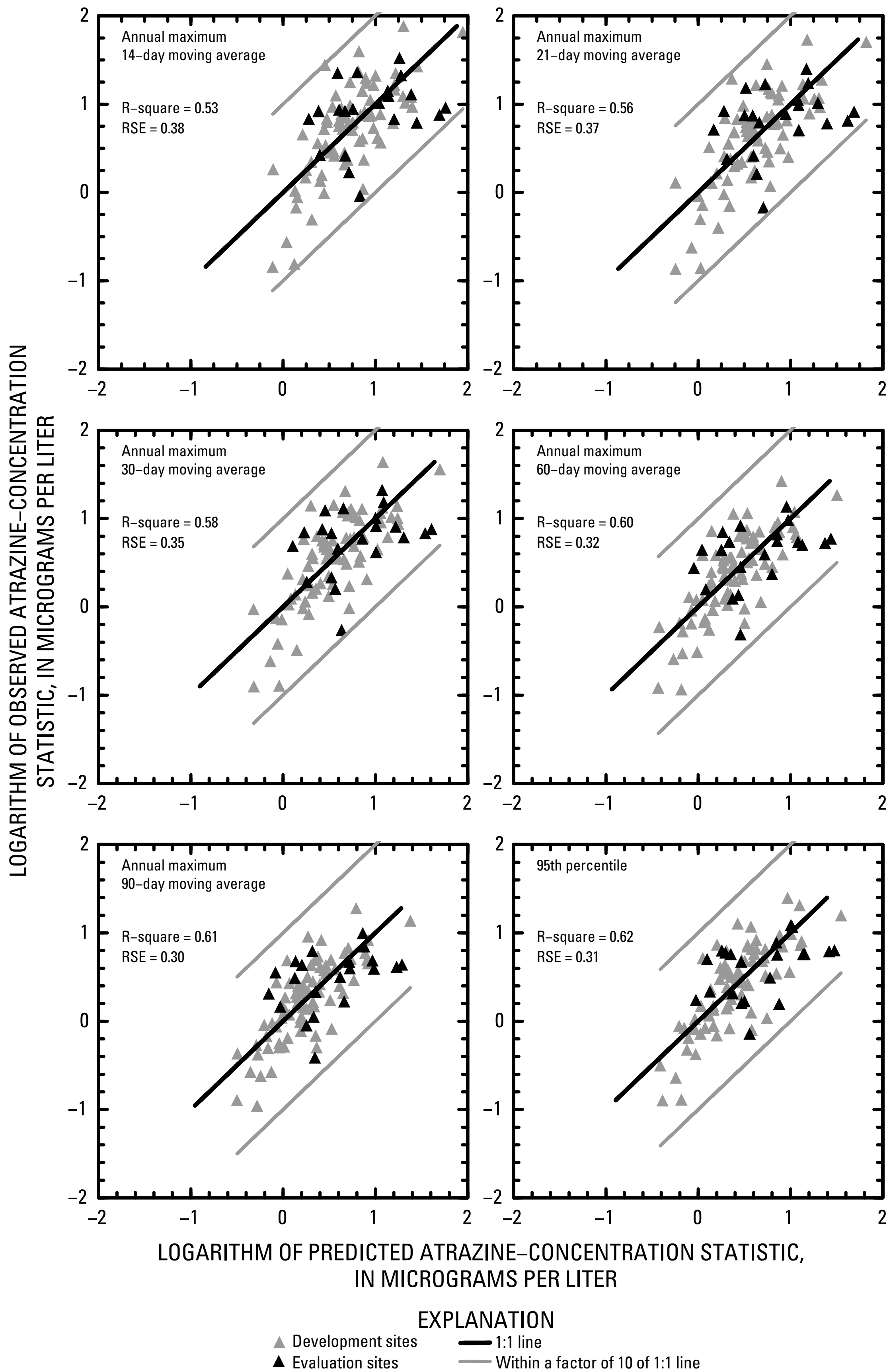

Figure 3. Atrazine concentration statistics from concentrations observed at model-development and evaluation sites in relation to values of the same statistics predicted by the WARP-CB models.

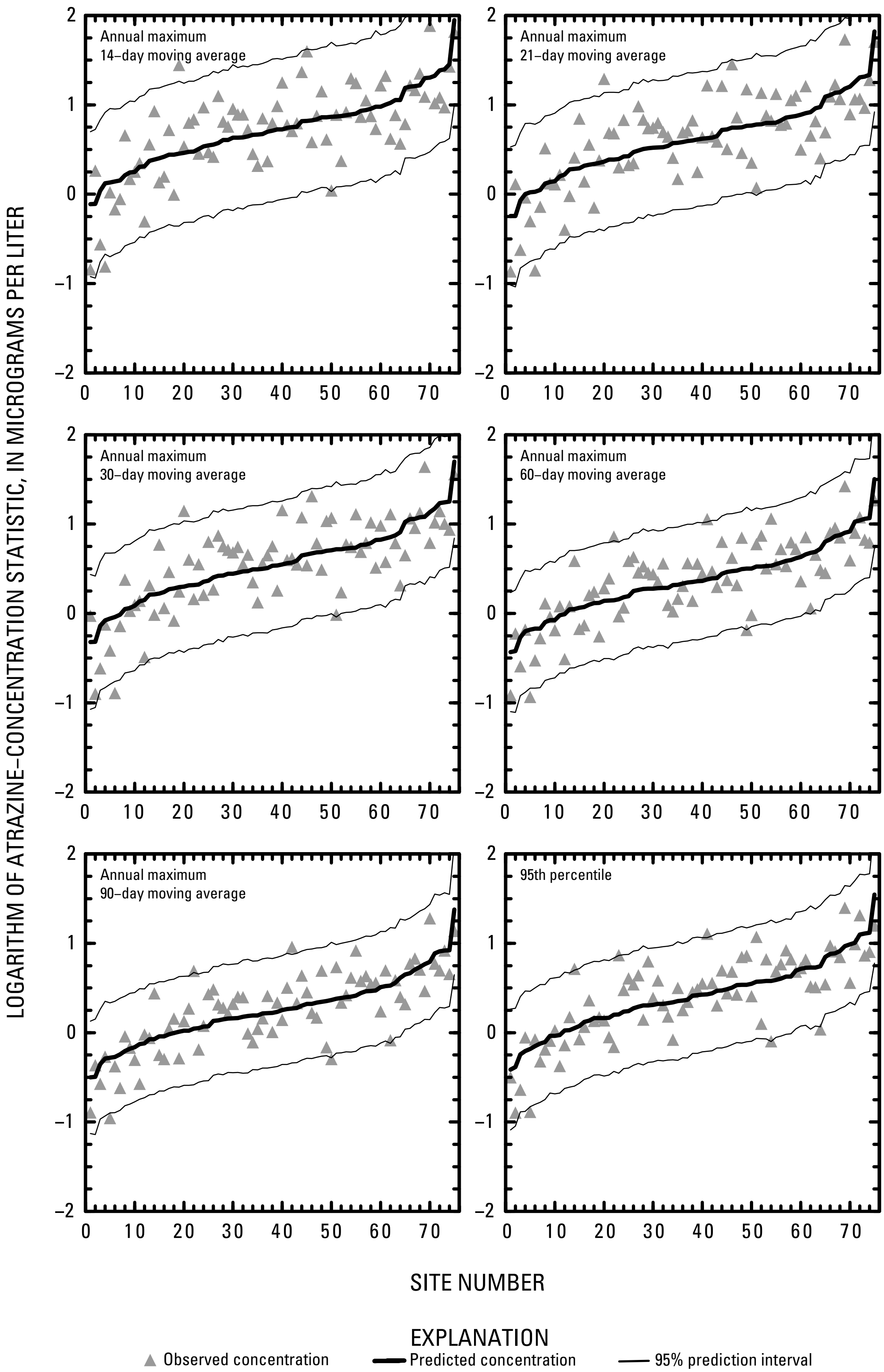

Figure 4. Prediction intervals for atrazine concentration statistics predicted by the WARP-CB models. (Model-development site years were arranged and numbered by order of increasing predicted 95th percentile concentration statistic.)

moving-average model has a better model fit and explains more variability (RSE = 0.30; R2 = 0.61) than the annual maximum 14-day moving-average model does (RSE = 0.38; R2 = 0.53).

Residual errors for the WARP-CB development sites from application of the WARP-CB and National WARP models are shown in figure 5. The National WARP models have a much larger range of residual errors for the WARP-CB model-development sites in comparison to the WARP-CB models. In addition, the National WARP models tend to overestimate atrazine concentration statistics for the WARP-CB model-development sites in comparison to WARP-CB models. Atrazine use intensity is the most important explanatory variable in the National WARP models, the tendency of the National WARP models to overestimate atrazine concentration statistics for the WARP-CB development sites may be influenced by the high use intensities represented by these sites (fig. 2). Table 4 shows the performance of the WARP-CB and National WARP models for the 75 WARP-CB model-development site and year combinations. The percentages of model predictions within a factor of 10, 5, and 2 of the observed concentration statistic are higher for the WARP-CB models than the National WARP models.

The levels of uncertainty for the annual maximum moving-average (21-, 60-, and 90-day durations) and 95th-percentile WARP-CB and National WARP models are compared in figure 6. PI size is represented as the ratio of the upper boundary of the interval to the predicted concentration statistic. This is the same as the ratio of the predicted concentration statistic to the lower boundary of the interval. The extreme values (shown as asterisks) represent sites where one or more explanatory variables are relatively extreme in value when compared to values for the rest of the sites. The sites with relatively extreme values in one or more explanatory variables have wider PIs, meaning greater uncertainty in their predicted concentration statistic, than sites with explanatory variable values that fall closer to the center of the explanatory data. For all atrazine concentration statistics, the PIs for National WARP models are approximately twice those for the WARP-CB models, indicating that the National WARP models have a higher level of uncertainty than the WARP-CB models.

Model Limitations

Application of the regression models for predicting atrazine-concentration statistics, and the WARP methodology in general, are subject to several important limitations.

- The WARP-CB models are based on sites with high atrazine use intensities (fig. 2). For stream sites where atrazine use intensities are lower than the least value used in the WARP-CB model development (17 kg/km^2 watershed area), the National WARP models should be used, with WARP-CB used only to provide an estimated upper limit for comparison to the National WARP result.

- The sampling frequencies of the model-development sites may not be sufficient to reliably characterize the highest moving-average concentrations during a year. Thus, application of the models to predict the annual maximum moving-average concentrations for relatively short durations such as the 14- and 21-day moving average is expected to generally underpredict the actual annual maximum moving-average concentrations for these durations.

- The regression models are designed for prediction of atrazine-concentration statistics for streams within the Corn Belt region of the United States. Although the sites used for model development represent a wide variety of environmental settings and a large range of watershed areas, it is likely that other watersheds within the Corn Belt have one or more characteristics outside the ranges of the watershed parameters used to develop the models (table 3). Application of the models to streams draining such watersheds would result in increased uncertainty in predicted concentrations.

- The models were developed by using concentration data from streams. Application of the models to lakes, reservoirs, or streams heavily influenced by reservoirs will result in biased predictions because of the influence of water storage on the temporal distribution of concentrations.

- The atrazine-use data used in the models are estimates for applications to agricultural land only. Substantial nonagricultural use of atrazine in a watershed could result in underprediction of atrazine concentrations in a stream, if such use cannot be estimated.

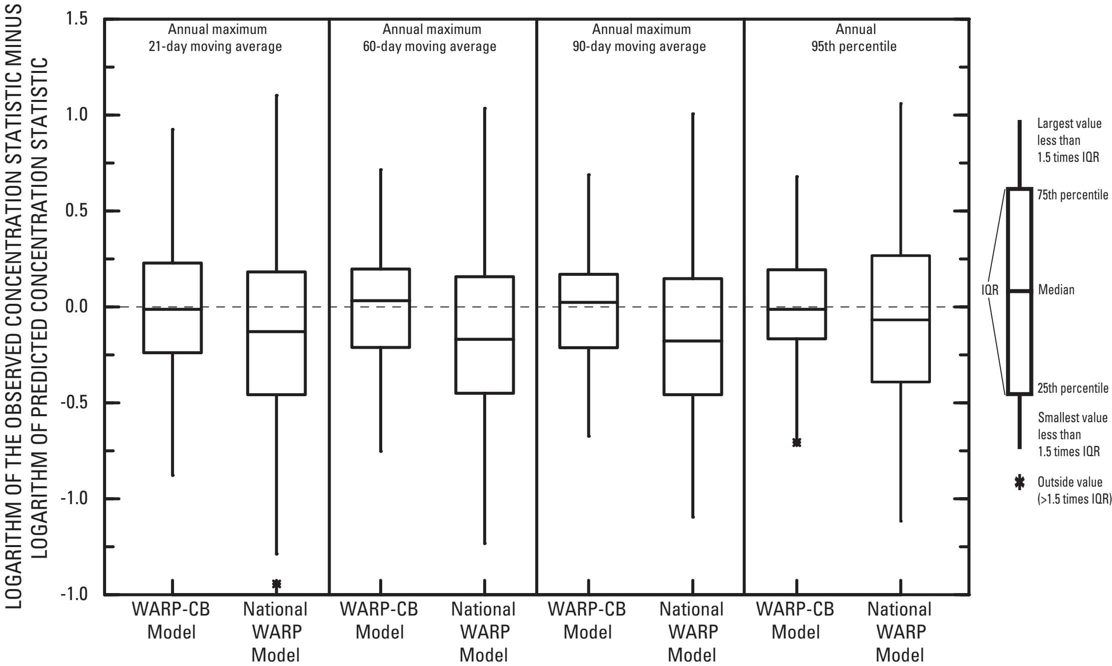

Figure 5. Residual errors for atrazine concentration statistics predicted by WARP-CB and National WARP models. (Residual error is $\log_{10}$(observed concentration statistic) – $\log_{10}$(predicted concentration statistic).)

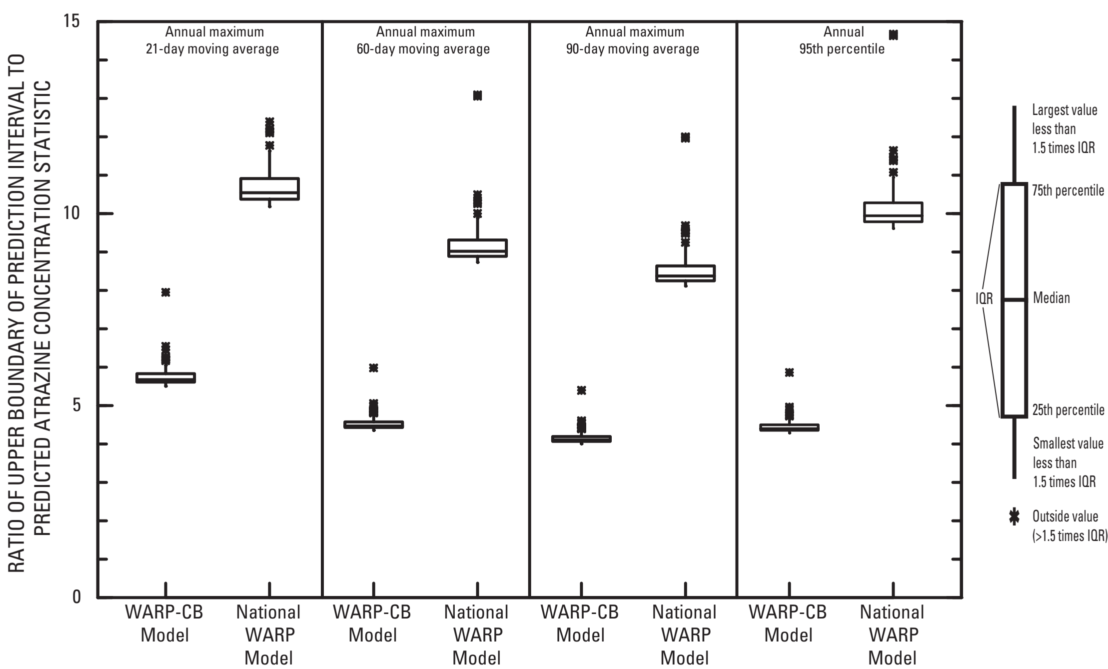

Figure 6. Potential prediction errors among concentration statistics predicted by use of the atrazine WARP-CB models. (Potential prediction error is represented by the ratio of the upper boundary of the 95-percent prediction interval to the predicted atrazine concentration statistic. Each boxplot shows the ratios among the 75 model-development sites and years.)

Summary and Conclusions

Regression models (WARP-CB) were developed for predicting annual maximum moving-average (14-, 21-, 30-, 60-, and 90-day durations) and annual 95th-percentile atrazine concentrations in streams of the Corn Belt region of the United States, using data on watershed characteristics. Predictions of atrazine-concentration statistics generated by these models can be used to characterize the concentrations of atrazine for comparison to specific water-quality benchmarks for evaluation of potential concerns regarding human health and aquatic life.

All sites used in development of the WARP-CB models drain watersheds with high atrazine use intensity. The models accounted for 53 to 62 percent of the variability in concentration statistics among the 75 sites and years used for model development. For all the models, 95-percent prediction intervals are well within a factor of 10 above and below the predicted concentration statistic. Over 90-percent of the WARP-CB model predictions for the model-development sites are within a factor of 5 of the observed concentrations statistic.

The WARP-CB model-development site predictions show smaller overall residual error and reduced uncertainty when compared to the National WARP model predictions for the same sites. Prediction intervals for National WARP models are approximately twice those for the WARP-CB models, indicating that the National WARP models have a higher level of uncertainty than the WARP-CB models. This lower level of uncertainty associated with the WARP-CB models will improve the predicted probabilities of exceeding a specified criterion or benchmark during application of the models. However, the WARP-CB models are based on sites with high atrazine use intensities. Even within the Corn Belt region, the National WARP models should be used instead of the WARP-CB models for sites where atrazine use intensities are lower than approximately 17 kg/km^2, although WARP-CB can be used to provide estimates of upper limits for comparison to the National WARP models.

The National WARP and WARP-CB models are tools for predicting atrazine concentrations in unmonitored or inadequately monitored streams. The performance of the models for the development and evaluation sites supports the application of the WARP-CB models for predicting annual atrazine concentration statistics in streams draining high atrazine use areas of the Corn Belt. The WARP-CB models also provide a framework to interpret the predictions in terms of uncertainty.

References Cited

Akaike, Hirotugu, 1974, A new look at the statistical model identification: IEEE Transactions on Automatic Control, v. AC-19, no. 6, p. 716–723, doi:10.1109/TAC.1974.1100705.

Baker, J.L., Helmers, M.J., and Laflen, J.M., 2007, Water management practices, rain-fed cropland, in Schnepf, Max, and Cox, Craig, eds., Environmental benefits of conservation on Cropland—The status of our knowledge: Ankeny, Iowa, Soil and Water Conservation Society, p. 89–130.

Crawford, C.G., 2001, Factors affecting pesticide occurrence and transport in a large midwestern river basin: Journal of the American Water Resources Association (JAWRA), v. 37, no. 1, p. 1–15, doi: 10.1111/j.1752-1688.2001.tb05470.x.

Crawford, C.G., 2004, Sampling strategies for estimating acute and chronic exposures of pesticides in streams: Journal of the American Water Resources Association (JAWRA), v. 40, no. 2, p. 485–502, doi:10.1111/j.1752-1688.2004.tb01045.x.

Gilliom, R.J.; Barbash, J.E.; Crawford, C.G.; Hamilton, P.A.; Martin, J.D.; Nakagaki, Naomi; Nowell, L.H.; Scott, J.C.; Stackelberg, P.E.; Thelin, G.P.; and Wolock, D.M., 2006, The quality of our Nation's waters—Pesticides in the Nation's streams and ground water, 1992–2001: U.S. Geological Survey Circular 1291, 172 p. (Also available at *http://pubs.usgs.gov/circ/2005/1291/*.)

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. A3, 522 p. (Also available at *http://pubs.usgs.gov/twri/twri4a3/*.)

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-Resources Investigations Report 03–4047, 68 p. (Also available at *http://pubs.usgs.gov/wri/wri034047/*.)

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 (JAWRA), v. 37, no. 5, p. 1349–1367, doi:10.1111/j.1752-1688.2001.tb03644.x.

Lerch, R.N., and Blanchard, P.E., 2003, Watershed vulnerability to herbicide transport in northern Missouri and southern Iowa streams: Environmental Science & Technology, v. 37, no. 24, p. 5518–5527, doi:10.1021/es030431s.

Lerch, R.N., Sadler, E.J., Sudduth, K.A., Baffaut, C., and Kitchen, N.R., 2011, Herbicide transport in Goodwater Creek Experimental Watershed—I. Long-term research on atrazine: Journal of the American Water Resources Association (JAWRA), v. 47, no. 2, p. 209–223, doi:10.1111/j.1752-1688.2010.00503.x.

National Center for Water Quality Research, 2009, Tributary data download, accessed August 2009 at *http://www.heidelberg.edu/academiclife/distinctive/ncwqr/data.*

Neter, John, Wasserman, William, and Kutner, M.H., 1985. Appliedlinear statistical models (2d ed.): Homewood, Ill., R.D. Irwin, 1,127 p.

Stone, W.W., Gilliom, R.J., and Crawford, C.G., 2008, Watershed Regressions for Pesticides (WARP) for predicting annual maximum and maximum moving-average concentrations of atrazine in streams: U.S. Geological Survey Open-File Report 2008–1186, 19 p. (Also available at *http://pubs.usgs.gov/of/2008/1186/*.)

Stone, W.W., and Gilliom, R.J., 2009, Update of Watershed Regressions for Pesticides (WARP) for predicting atrazine concentration in streams: U.S. Geological Survey Open-File Report 2009–1122, 22 p. (Also available at *http://pubs.usgs.gov/of/2009/1122/*.)

Thelin, G.P., and Stone, W.W., 2010, Method for estimating annual atrazine use for counties in the conterminous United States, 1992–2007: U.S. Geological Survey Scientific Investigations Report 2010–5034, 129 p. (Also available at *http://pubs.usgs.gov/sir/2010/5034/*.)

Tukey, J.W., 1977. Exploratory data analysis: Reading, Mass., Addison-Wesley, 688 p.

U.S. Department of Agriculture, 1995, 1992 national resources inventory: Washington, D.C., Natural Resources Conservation Service, and Ames, Iowa, Statistical Laboratory, Iowa State University.

U.S. Department of Agriculture, 2007a, Census of agriculture, table 26, Field crops, accessed January 2011 at *http://www.agcensus.usda.gov/Publications/2007/Full_Report/Volume_1,_Chapter_2_US_State_Level/st99_2_026_026.pdf*.

U.S. Department of Agriculture, 2007b, Hydrologic soil groups: Natural Resources Conservation Service, National Engineering Handbook, chap. 7, Part 630, Hydrology, 210-VI-NEH, accessed January 2011 at *http://directives.sc.egov.usda.gov/OpenNonWebContent.aspx?content=17757.wba*.

U.S. Department of Agriculture, 2010, Soil Survey Geographic (SSURGO) Database: Natural Resources Conservation Service, accessed January 2011 at *http://soildatamart.nrcs.usda.gov*.

U.S. Environmental Protection Agency, 2003, Interim reregistration eligibility decision for atrazine: Washington, D.C., Office of Prevention, Pesticides, and Toxic Substances, Office of Pesticide Programs, Case No. 0062, 304 p. (Also available at *http://www.epa.gov/pesticides/reregistration/REDs/atrazine_ired.pdf*.)

U.S. Environmental Protection Agency, 2007, Preliminary interpretation of the ecological significance of atrazine stream-water concentrations using a statistically-designed monitoring program: Washington, D.C., Office of Prevention, Pesticides, and Toxic Substances, Office of Pesticide Programs, Environmental Fate and Effects Division, 221 p., accessed January 2011 at *http://www.epa.gov/scipoly/sap/meetings/2007/december/whitepaper_sap.pdf*.

U.S. Environmental Protection Agency, 2010a, Atrazine Ecological Effects Monitoring Program—Midwestern Stream Monitoring project data: Accessed January 2011 at *http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2003-0367-0178*.

U.S. Environmental Protection Agency, 2010b, Meeting minutes of the FIFRA Science Advisory Panel meeting held April 26–29, 2010, on the re-evaluation of the human health effects of atrazine—Review of experimental animal and in vitro studies and drinking water monitoring frequency: SAP Minutes No. 2010–04, accessed January 2011 at *http://www.epa.gov/scipoly/sap/meetings/2010/april/042610minutes.pdf*.

U.S. Environmental Protection Agency, 2011, Atrazine Updates: Monitoring in Community Water Systems. Atrazine Monitoring Program Drinking Water data: Accessed January 2011 at *http://www.epa.gov/opp00001/reregistration/atrazine/atrazine_update.htm*.

Venables, W.N., and Ripley, B.D., 1999. Modern applied statistics with S-PLUS (3d ed.): New York, Springer-Verlag, 501 p.

Appendixes

Appendixes are separate online documents, accessible at *http://pubs.usgs.gov/ofr/2011/1141/*.

1. Values of response and explanatory variables used in the WARP-CB model development.
2. Detailed description of the explanatory variables used in the development of the WARP-CB models.

Stone and Gilliom—**Watershed Regressions for Pesticides (WARP) for Predicting Atrazine Concentrations in Corn Belt Streams**—Open-File Report 2011–1141