1997

Statistical modeling of agricultural chemical occurrence in midwestern rivers

William A. Battaglin
U.S. Geological Survey

Donald A. Goolsby
U.S. Geological Survey

Journal of Hydrology 196 (1997) 1–25

Statistical modeling of agricultural chemical occurrence in midwestern rivers

William A. Battaglin*, Donald A. Goolsby

US Geological Survey, Water Resources Division, D.F.C., Box 25046, MS 406 Lakewood, CO 80225, USA

Received 2 May 1996; accepted 16 January 1997

Abstract

Agricultural chemicals in surface water may constitute a human health risk or have adverse effects on aquatic life. Recent research on unregulated rivers in the midwestern USA documents that elevated concentrations of herbicides occur for 1–4 months following application in late spring and early summer. In contrast, nitrate concentrations in unregulated rivers are elevated during fall, winter, and spring months. Natural and anthropogenic variables of river drainage basins, such as soil permeability, amount of agricultural chemicals applied, or percentage of land planted in corn, affect agricultural chemical concentration and mass transport in rivers.

Presented is an analysis of selected data on agricultural chemicals collected for three regional studies conducted by the US Geological Survey. Statistical techniques such as multiple linear and logistic regression were used to identify natural and anthropogenic variables of drainage basins that have strong relations to agricultural chemical concentrations and mass transport measured in rivers. A geographic information system (GIS) was used to manage and analyze spatial data. Statistical models were developed that estimated the concentration, annual transport, and annual mean concentration of selected agricultural chemicals in midwestern rivers.

Multiple linear regression models were not very successful (R^2 from 0.162 to 0.517) in explaining the variance in observed agricultural chemical concentrations during post-planting runoff. Logistic regression models were somewhat more successful, correctly matching the observed concentration category in 61–80% of observations. Linear and multiple linear regression models were moderately successful (R^2 from 0.522 to 0.995) in explaining the variance in observed annual transport and annual mean concentration of agricultural chemicals. Explanatory variables that were commonly significant in the regression models include estimates of agricultural chemical use, crop acreage, soil characteristics, and basin topography. © 1997 Elsevier Science B.V.

* Corresponding author.

PII S0022-1694(97)00011-5

1. Introduction

1.1. Background

The US Geological Survey currently (1997) is conducting studies to determine the distribution, transport, and persistence of agricultural chemicals (herbicides, insecticides, other pesticides, and fertilizers) in midwestern USA rivers. Previous studies have documented that the transport of agricultural chemicals in runoff to rivers is seasonal; the largest contributions generally occur during periods of late spring and early summer runoff, and are associated with recent chemical applications during or prior to crop planting (Baker and Richards, 1990; Goolsby et al., 1991; Thurman et al., 1992; Goolsby and Battaglin, 1993, 1995; Battaglin et al., 1993).

These studies determined that herbicides concentrations in some rivers exceed the value of health-based standards for finished drinking water (US Environmental Protection Agency, 1995) at certain times of the year. However, these higher concentrations generally do not persist for the remainder of the year (Goolsby and Battaglin, 1993, 1995), and the health-based standards that are exceeded are for average annual exposure, and not 1- or 10-day exposure.

The presence of agricultural chemicals in rivers of the midwestern United States is of concern to the millions of people who depend on the Ohio, Missouri, and Mississippi Rivers and their tributaries as sources for drinking water. Conventional water-treatment practices do not remove most of these agricultural chemicals (US Environmental Protection Agency, 1989; Baker and Richards, 1990; Adams et al., 1990), and costly alternative treatment methods, such as carbon filtration, reverse osmosis, ozone oxidation, or denitrification, might be required for supplies that do not meet health-based limits.

Natural and anthropogenic variables within river drainage basins, such as soil permeability, the amount of agricultural chemicals applied, or percentage of land planted in various crops, can affect agricultural chemical concentrations and mass transport in rivers. The Mississippi River Basin (MRB) contains some of the most productive cropland in the world. Approximately 80% of the corn and soybeans produced in the United States are grown in the MRB (US Department of Commerce, 1989). Agricultural chemicals are used extensively in the United States to increase production of crops. In the MRB, an estimated 100 000 t of pesticides and 6.3 million t of nitrogen fertilizer are used annually (US Environmental Protection Agency, 1990; Gianessi and Puffer, 1991). Most of this agricultural chemical use is associated with the production of corn, soybeans, sorghum, and wheat. An improved understanding of the spatial and temporal distribution of agricultural chemicals in midwestern rivers is needed so policy makers can make informed decisions about which agricultural chemicals have potential to affect water quality over large areas, and so water managers and suppliers can address problems resulting from the occurrence of agricultural chemicals in drinking-water supplies.

1.2. Objectives

This paper presents a statistical analysis of selected data on agricultural chemicals collected for three studies: (1) the Regional Reconnaissance study (Thurman et al.,

1992; Scribner et al., 1993), (2) the Mississippi River study (Goolsby et al., 1991; Coupe et al., 1995), and (3) the Temporal Variability study (Scribner et al., 1994; Goolsby and Battaglin, 1995). Digital spatial data on agricultural chemical use, land use, climate, and soils, are quantified and managed using a geographic information system (GIS). These data and the water-chemistry data from the three studies are used to (1) identify explanatory variables that are strongly related to the concentration, annual transport, or annual mean concentration of agricultural chemicals in midwestern rivers; (2) estimate annual agricultural chemical transport as a percentage of annual use, or as a function of other explanatory variables; (3) model instantaneous concentration or annual mean concentration of agricultural chemicals as a function of annual use rate, or other explanatory variables; and (4) model the probability of agricultural chemical concentrations being within a particular concentration range.

1.3. Regional Reconnaissance study area

The Regional Reconnaissance study documented agricultural chemical concentrations in streamflow from 147 upper-midwestern basins. Drainage areas of the basins ranged from 260 to 1.8 million km^2. A stratified random-sampling procedure was used to ensure adequate geographic distribution of the sampled basins (Thurman et al., 1992; Scribner et al., 1993). Samples were collected three times in 1989: (1) pre-planting (March–May), (2) post-planting (May–July), and (3) harvest (October–November). Results from 142 basins (Fig. 1(a)) are used in the analysis described in this paper. Sites that are not used either have inadequate daily streamflow records for computing flow percentiles or have drainage areas that are greater than 130 000 km^2.

1.4. Mississippi River and Temporal Variability study areas

The Mississippi River study documented agricultural chemical concentrations in streamflow from eight midwestern basins. The Mississippi River was sampled at three locations, and the Missouri, Ohio, Platte, Illinois, and White Rivers were each sampled at one location. The Temporal Variability study documented agricultural chemical concentrations in streamflow from nine midwestern basins. Two sites, the West Fork of the Big Blue River in Nebraska and the Sangamon River in Illinois, were sampled concurrently with the eight Mississippi River sites and are included in the analysis presented in this paper. The drainage basins of these ten rivers range in size from 1430 to 2.9 million km^2, and are all predominantly within the corn and soybean producing region of the midwestern USA (Fig. 1(b)) (Table 1).

Water samples were collected at the ten sites from April 1991 through March 1992. Except for the Ohio River, all eight sites from the Mississippi River study were sampled every 3–4 days during May through mid-July, once per week during the remainder of the summer and fall, and once every 2 weeks during winter. The Ohio River was sampled once per week, except during winter when samples were collected every 2 weeks (Goolsby et al., 1991). Samples at the eight sites were collected by depth–width integrating methods (Edwards and Glysson, 1988) and composited in stainless steel or glass containers. Automatic samplers were installed at the two sites from the Temporal Variability study.

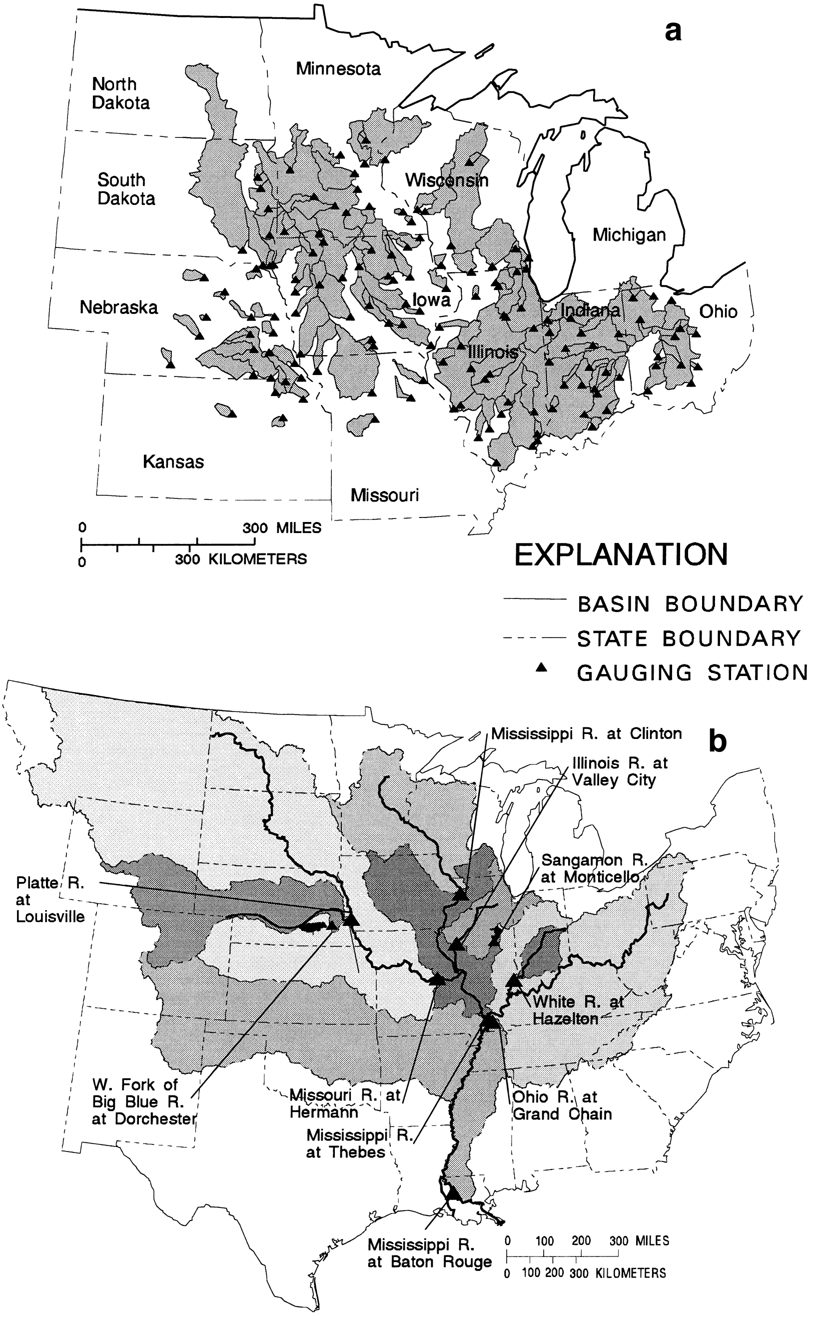

Fig. 1. Water sampling stations and upstream drainage basins for (a) 142 sites in the Regional Reconnaissance study and (b) ten sites in the Mississippi River Basin and Temporal Variability studies.

Table 1
Mississippi River and Temporal Variability study site names, drainage areas, discharge, and selected cropland statistics

Site number	Sampling station name	Drainage basin area (km^2)	Mean daily discharge 4/1/1991–3/31/1992 ($m^3 s^{-1}$)	Harvested corn acreage (% of drainage basin area)	Harvested soybean acreage (% of drainage basin area)	Harvested cropland acreage (% of drainage basin area)
1	Mississippi River at Baton Rouge, LA	2914000	18300	7.0	6.7	25.0
2	Mississippi River at Thebes, IL	1847000	6040	8.6	7.0	27.1
3	Missouri River at Hermann, MO	1357000	1640	4.6	4.1	21.2
4	Ohio River at Grand Chain, IL	526000	6650	7.5	6.8	20.6
5	Mississippi River at Clinton, IA	221700	2030	13.4	7.4	33.1
6	Platte River at Louisville, NE	222200	165	6.9	2.4	18.1
7	Illinois River at Valley City, IL	69000	669	27.7	24.9	56.0
8	White River at Hazelton, IN	29000	178	19.4	16.2	40.1
9	W. Fork Big Blue River, Dorchester, NE	3120	4.8	40.3	8.1	61.8
10	Sangamon River at Monticello, IL	1430	8.8	36.1	40.6	77.8

Samples at these sites were collected as frequently as every few hours during runoff events, at least two or three times a week during the spring runoff period, and about once a week during the remainder of the year.

1.5. Analytical procedures

Upon collection, samples for dissolved nutrients were filtered, preserved with mercuric chloride, and chilled prior to shipment to the laboratory for analysis (Scribner et al., 1993, 1994; Coupe et al., 1995). Nitrate concentrations were determined by automated colorimetric procedures (Fishman and Friedman, 1989). Upon collection, samples for pesticides were filtered though glass fiber filters with a nominal pore size of either 1.0 or 0.7 μm, and chilled prior to shipment to the laboratory. Samples from the Regional Reconnaissance study were analyzed for 11 herbicides and two triazine metabolites by gas chromatography/mass spectrometry (GC/MS) (Thurman et al., 1990; Scribner et al., 1993). Samples from the Mississippi River and Temporal Variability studies were analyzed by one of two methods, both utilizing GC/MS following solid-phase extraction on C_{18} cartridges (Sandstrom et al., 1991, 1994; Coupe et al., 1995).

1.6. GIS data sources

County-level estimates of selected variables used for this study were acquired and managed using GIS. County-level estimates of agricultural land use, crop acreage, and livestock are from the 1987 census of agriculture (US Department of Commerce, 1989). County-level estimates of herbicide use are obtained from Gianessi and Puffer (1991). The herbicide-use estimates are indicative of annual use during the 1987–89 time period, and include agricultural use only; some of these same herbicides are also used for non-selective weed control in industrial areas, fairways, and lawns (Meister and Sine, 1995). County-level estimates of inorganic fertilizer sales (US Environmental Protection Agency, 1990; G. Fletcher, written communication, 1992) are in metric tons of actual nutrient (inorganic nitrogen) for the 1989, 1990, and 1991 fertilizer years (July 1 through June 30). These sales estimates are a surrogate for fertilizer use estimates, which are not uniformly available for the study area. Sales estimates are for agricultural fertilizer only and do not account for residential fertilizer sales or for the use of manure as a natural fertilizer.

Other variables defined from digital and non-digital data sources include: basin area; population density; selected soil variables (hydrologic group, porosity, water-holding capacity and permeability); corn and soybean planting dates; mean annual temperature, precipitation, and runoff; and several hydrologic parameters calculated for a watershed model. Population estimates are from the 1990 census of population and housing (US Department of Commerce, 1990). STATSGO soils data (US Department of Agriculture, 1993) were used to estimate soil hydrologic group, porosity, permeability, and available water-holding capacity. Corn and soybean planting dates were determined from information found in State Agricultural Statistics Service reports and in National Agricultural Statistics reports (US Department of Agriculture, 1984). Climate data are from the US historical climatological network (Karl and Riebsame, 1989; Karl et al., 1990) and from

Table 2
Estimated annual transport in t of atrazine, alachlor, cyanazine, metolachlor, and nitrate; April 1, 1991 through March 31, 1992

Site number	Atrazine	Alachlor	Cyanazine	Metolachlor	Nitrate as N
1	321.0	33.7	127.0	123.0	967000
2	189.0	42.9	113.0	87.2	651000
3	68.4	7.9	31.3	24.7	79900
4	57.0	5.0	13.4	20.2	182000
5	15.1	5.6	8.9	8.1	157000
6	9.3	2.5	5.9	2.5	5640
7	35.8	8.8	19.8	18.9	107000
8	5.7	1.0	1.7	2.1	11500
9	2.8	0.7	0.1	0.7	270
10	0.7	0.2	0.3	0.5	3090

D.M. Wolock (written communication, 1995). Runoff estimates are from a digital version (Rea and Cederstrand, 1994) of a map of average annual runoff in the United States.

The topographic characteristic parameter mean $\ln(a/\tan B)$ for the TOPMODEL (Wolock, 1993) watershed model was calculated for each basin using digital elevation model (DEM) data (a is the upslope area per unit contour length and $\tan B$ is the slope gradient) (D.M. Wolock, written communication, 1995). Mean $\ln(a/\tan B)$ values describe the slope and shape of a basin and are small in basins with steep or short slopes and large in basins with flat or long slopes.

2. Methods of investigation

2.1. Estimating annual transport and annual mean concentration of agricultural chemicals

Daily transport of dissolved atrazine, alachlor, cyanazine, metolachlor, and nitrate as nitrogen was calculated for the ten basins in the Mississippi River and Temporal Variability studies as the product of measured or estimated daily chemical concentrations and daily mean streamflow. Research has shown that less than 1% of the atrazine, cyanazine, and metolachlor transported by the Mississippi River travels in the suspended phase (Pereira and Rostad, 1990; Squillace and Thurman, 1992). Concentrations were estimated by linear interpolation on days when no samples were collected. A concentration of zero was used when herbicide or nitrate concentrations were less than analytical reporting limits (0.05 $\mu g\ l^{-1}$ or less for all chemicals). Estimates of daily transport for the period April 1, 1991, to March 31, 1992 were summed to obtain estimates of annual transport (Table 2). Estimates of annual transport of agricultural chemicals in the Mississippi River and its tributaries for 1991 are similar to those reported by Goolsby et al. (1991) and Dunn (1996), but significantly larger than those reported by Pereira et al. (1989) for 1987.

Time-weighted annual mean concentrations of atrazine, alachlor, cyanazine, metolachlor, and nitrate were computed as the average of daily concentration estimates for the period April 1, 1991 to March 31, 1992 (Table 3) (Battaglin and Hay, 1996). Annual

Table 3

Estimated time-weighted annual mean concentrations of atrazine, alachlor, cyanazine, metolachlor in mg l^{-1} and nitrate in mg l^{-1}; April 1, 1991 through March 31, 1992, and standard deviations (in parentheses)

Site number	Atrazine	Alachlor	Cyanazine	Metolachlor	Nitrate as N
1	0.48 (0.71)	0.05 (0.10)	0.20 (0.34)	0.19 (0.28)	1.54 (0.64)
2	0.70 (0.90)	0.16 (0.19)	0.40 (0.63)	0.32 (0.43)	2.98 (1.32)
3	0.86 (1.35)	0.09 (0.17)	0.37 (0.79)	0.33 (0.44)	1.33 (0.72)
4	0.34 (0.47)	0.04 (0.08)	0.08 (0.14)	0.09 (0.17)	0.71 (0.37)
5	0.20 (0.24)	0.07 (0.14)	0.12 (0.23)	0.09 (0.16)	2.29 (0.86)
6	0.85 (1.65)	0.17 (0.52)	0.51 (1.18)	0.20 (0.51)	0.84 (0.58)
7	0.94 (1.70)	0.21 (0.46)	0.50 (0.97)	0.45 (0.74)	4.47 (2.23)
8	0.91 (1.65)	0.14 (0.42)	0.25 (0.58)	0.34 (0.58)	1.67 (1.31)
9	5.98 (15.6)	1.29 (4.56)	0.14 (0.46)	1.08 (2.90)	1.84 (0.93)
10	1.09 (2.28)	0.26 (0.66)	0.30 (1.05)	0.62 (1.33)	7.47 (5.18)

mean concentration estimates are time-weighted, not flow weighted, because time-weighted estimates are likely more representative of an annual mean exposure expected from using rivers or some alluvial aquifers as a source of drinking water. Time-weighted concentration estimates are also relevant to public-water suppliers who must comply with federal regulations (US Environmental Protection Agency, 1991). For herbicides, which tend to occur in higher concentrations during spring runoff when flow is also high, flow-weighted concentration estimates are generally larger than time-weighted estimates. For example, the flow-weighted annual mean metolachlor concentrations for Sites 6 and 9 are 0.51 and 4.12 μg l^{-1}, whereas the time-weighted annual mean concentrations for Sites 6 and 9 are 0.20 and 1.08 μg l^{-1}, respectively.

Insufficient data were collected at the Regional Reconnaissance study sites to estimate either annual transport or annual mean concentrations of agricultural chemicals. Statistical analysis of data collected at these sites was done only on instantaneous concentration values from the post-planting samples.

2.2. GIS data management

GIS was used to manage, manipulate, and display the large quantity of spatial data in this study. Much of the data originated in tabular form, and was converted or entered into GIS thematic data layers (coverages). County-level estimates of agricultural chemical use, agricultural land use, crop acreage, and livestock were constructed by processing tabular data into GIS coverages as described by Battaglin and Goolsby (1995). Graphics showing the relative intensity or distribution of these estimates were created to help visualize their spatial distribution across the United States. Examples of these county-level estimates are given in Fig. 2.

Other data that were converted to or processed as GIS coverages for this study include: population density; soil hydrologic group, porosity, water-holding capacity, and permeability; corn and soybean planting dates; and mean annual temperature, precipitation, and runoff. Population estimates were constructed by processing census geographic unit

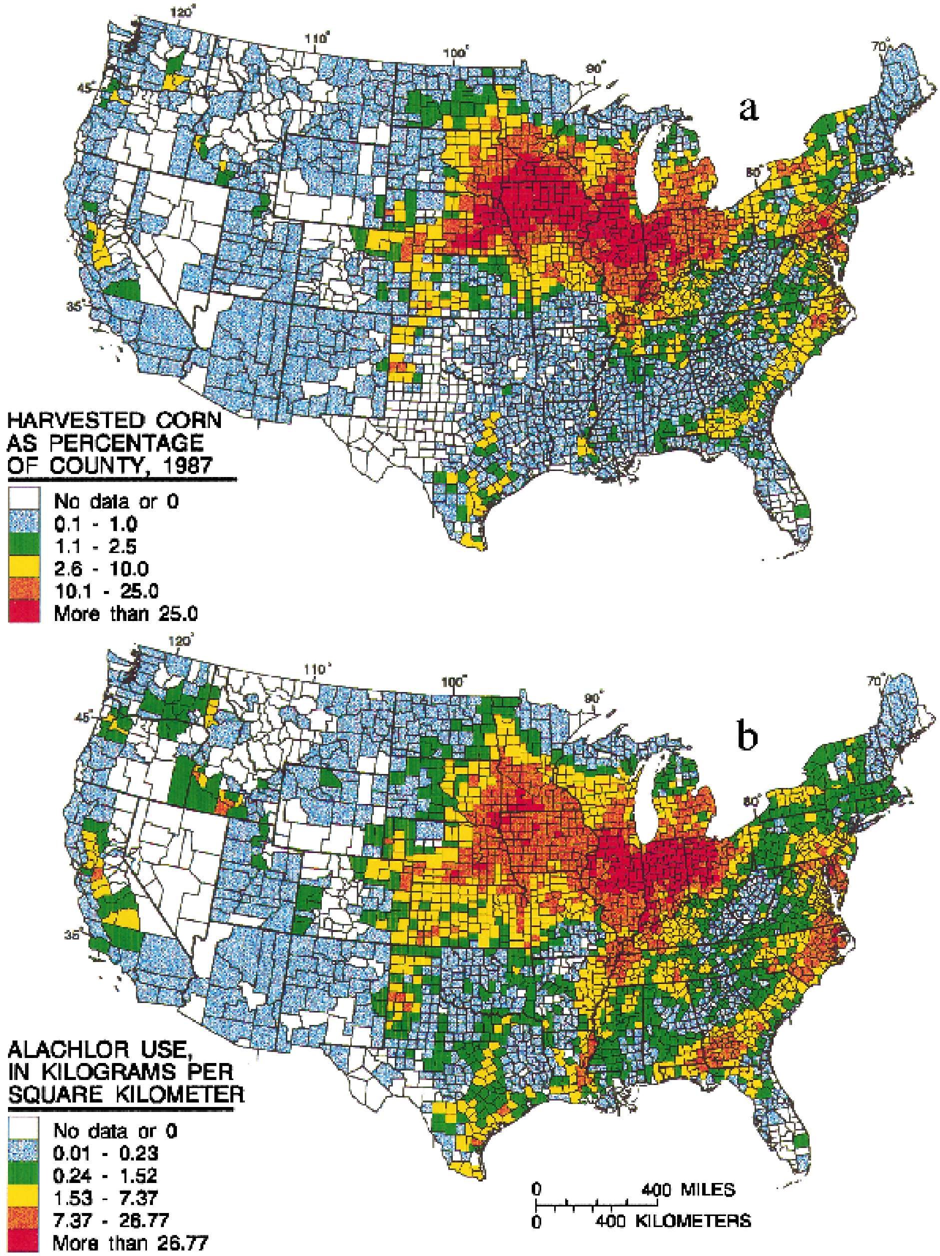

Fig. 2. Estimated (a) harvested corn acreage as a percentage of total county area, 1987 and (b) alachlor use divided by county area, 1989.

(block group) centroids into a GIS point coverage and assigning the population estimates for each block group from the 1990 census of population and housing (US Department of Commerce, 1990) to the appropriate point. STATSGO soils parameters, such as average soil hydrologic group values, for soil association polygons were computed as the area-weighted average value of the soil components that comprise each soil association (US Department of Agriculture, 1993). Dates on which 50% of the corn or soybeans were planted are from State Agricultural Statistics Service reports if available or are the mid-point in the range of usual planting dates from a national agricultural statistics report (US Department of Agriculture, 1984). County-level runoff estimates were computed by interpolating runoff values at county polygon centroids from runoff contours (Rea and Cederstrand, 1994). Some data were not defined using GIS. Mean daily discharge for the period of annual load calculation (FLOW) was computed from daily mean discharge values at streamflow gaging stations. Discharge measured at the time of sampling was converted to percentile of flow (FLWP) based upon the distribution of daily-streamflow values at each site.

Area-weighted transfer and area-weighted sum algorithms were programmed in the GIS and used to generate estimates within drainage basin polygons of the mass of agricultural chemical used, acreages of crops or cropland, and numbers of people or livestock. The algorithms account for cases where the entire county is within a single drainage basin and where only a part of a county is within a drainage basin. In the latter case, values of attributes are weighted by the ratio of area of the county within the basin to total county area. In most cases these raw variable estimates are normalized by dividing by drainage basin area (Mueller et al., 1993).

Statistical models were used to determine if the explanatory variables available could be used to accurately estimate concentration, annual transport, or annual mean concentration of selected agricultural chemicals in midwestern streams. Explanatory and response variables that appear in one or more of the statistical models are listed in Table 4.

2.3. Statistical techniques

Parametric correlations and linear, multiple linear, and logistic regression models were used to identify relations between values of explanatory variables and estimates of the concentration, annual transport, or annual mean concentration of selected agricultural chemicals (response variables). Explanatory variables that were significantly correlated ($P < 0.1$) with the concentration, annual transport, or annual mean concentration of agricultural chemicals were tested in regression models for their ability to predict the response variables. Strong correlation between two explanatory variables was a good indicator that a regression model containing both variables would be undesirably affected by multi-collinearity (Helsel and Hirsch, 1992). Stepwise procedures were used to identify sets of explanatory variables that could best estimate the concentration, annual transport, or annual mean concentration of agricultural chemicals. Explanatory variables with slope parameters having a $P > 0.1$ were considered insignificant, and were not included in the final models. When resulting models had problems with multi-collinearity, sets of explanatory variables were selected manually. Linear and multiple linear regression models were compared based on their R-squared values, root mean square error values, plots of

Table 4
Short names and definitions of explanatory and response variables

Short name	Variable type	Definition
ALAC	Response	Annual mean alachlor concentration in mg l^{-1}, 4/1/1991–3/31/1992
ALAT	Response	Annual transport of alachlor in t, 4/1/1991–3/31/1992
ALAR	Explanatory	Alachlor use rate in km km^{-2}, 1987/89
ALAU	Explanatory	Alachlor use in t, 1987/89
ATRC	Response	Annual mean atrazine concentration in mg l^{-1}, 4/1/1991–3/31/1992
ATRT	Response	Annual transport of atrazine in t, 4/1/1991–3/31/1992
ATRR	Explanatory	Atrazine use rate in kg km^{-2}, 1987/89
ATRU	Explanatory	Atrazine use in t, 1987/89
AWC	Explanatory	Available water-holding capacity of the soil in %
CHCK	Explanatory	Number of chickens 3 months or older, 1987
COWD	Explanatory	Density of cattle and calves in number km^{-2}, 1987
CRNH	Explanatory	Corn for grain or seed, harvested, in acres, 1987
CRNP	Explanatory	Harvested corn acreage as % of basin area, 1987
CRPH	Explanatory	Harvested cropland in acres, 1987
CRPP	Explanatory	Harvested cropland acreage as % of basin area, 1987
CYNC	Response	Annual mean cyanazine concentration in mg l^{-1}, 4/1/1991–3/31/1992
CYNT	Response	Annual transport of cyanazine in t, 4/1/1991–3/31/1992
CYNR	Explanatory	Cyanazine use rate in kg km^{-2}, 1987/89
CYNU	Explanatory	Cyanazine use in t, 1987/89
FLOW	Explanatory	Mean daily discharge in m^3 s^{-1}, 4/1/91–3/31/92
FLWP	Explanatory	Flow percentile of streamflow at the time of sample collection
GRNP	Explanatory	Harvested grains (wheat, rye, and oats) acreage as % of basin area, 1987
HYGP	Explanatory	Average soil hydrologic group value (small values represent well drained sandy soils and large values represent poorly drained clay rich soils)
ICRP	Explanatory	Irrigated cropland acreage as % of basin area, 1987
METC	Response	Annual mean metolachlor concentration in mg l^{-1}, 4/1/1991–3/31/1992
METT	Response	Annual transport of metolachlor in t, 4/1/1991–3/31/1992
METR	Explanatory	Metolachlor use rate in kg km^{-2}, 1987/89
METU	Explanatory	Metolachlor use in t, 1987/89
NITC	Response	Annual mean nitrate concentration in mg l^{-1}, 1991
NITT	Response	Annual transport of nitrate in t, 1991
NITR	Explanatory	Nitrogen fertilizer use rate in t km^{-2}, 1991
NU91	Explanatory	Nitrogen fertilizer use in t, 1991
PERM	Explanatory	Soil permeability in m day^{-1}
PIGD	Explanatory	Density of hogs and pigs in number km^{-2}, 1987
PIGS	Explanatory	Number of hogs and pigs, 1987
PLNS	Explanatory	Julian date when 50% of soybean planting was completed, 1989
POLD	Explanatory	Density of poultry (chickens, hens, broilers, and turkeys) in number km^{-2}, 1987
POPD	Explanatory	Density of people in number km^{-2}, 1990
PORO	Explanatory	Average soil porosity in %
PSTP	Explanatory	Pastureland acreage as % of basin area, 1987
Q	Explanatory	Long-term mean discharge in mm day^{-1}
*Q*_PPT	Explanatory	Mean discharge divided by mean precipitation in %
RNOF	Explanatory	Mean annual runoff in cm $year^{-1}$
ROWP	Explanatory	Row crops (corn, soybeans, and sorghum) acreage as % of basin area, 1987
SOYH	Explanatory	Soybeans for beans, harvested, in acres, 1987
SOYP	Explanatory	Harvested soybeans acreage as % of basin area, 1987
SRGP	Explanatory	Harvested sorghum acreage as % of basin area, 1987
TEMP	Explanatory	Long-term annual mean temperature in ° (Fahrenheit)
TOPM	Explanatory	Mean of $\ln(a/\tan B)$ distribution (TOPMODEL) topographic parameter where a is the upslope area, and $\tan B$ is slope

residuals, and variance of inflation (VIF) values. Logistic regression models were compared based on fit statistics such as the Akaike information criterion (AIC) and the Schwartz Criterion (SC) (SAS Institute, 1990), and on their ability to correctly classify agricultural chemical concentrations.

Linear and multiple linear regression (MLR) models were used to investigate relations between explanatory variables and estimates of the concentration, annual transport, or annual mean concentration of selected agricultural chemicals. Linear and multiple linear regression are statistical techniques that use one or more explanatory variables to explain as much of the variation observed in the response variable as possible. Once calibrated, regression models can be used to estimate the response variable from knowledge of the explanatory variable values (Eq. (1)).

$$y = B_0 + B_1 x_1 + \ldots B_k x_k + e \tag{1}$$

where y is the response variable, B_0 is the intercept, B_1 is the slope coefficient for the first explanatory variable x_1, B_k is the slope coefficient for the kth explanatory variable, and e is the remaining unexplained noise (error).

Regression models with between one and six explanatory variables were tested. Though many variables were significantly correlated with concentration, annual transport, or annual mean concentration of agricultural chemicals, the best models generally had five or fewer explanatory variables. Models with more than five explanatory variables often had higher r-squared values, but also had problems with autocorrelation among the explanatory variables, as indicated by VIF values of greater than 10 (Helsel and Hirsch, 1992).

Logistic regression (LGR) is a statistical technique that also uses one or more explanatory variables to predict the probability of a categorical response (Pereira and Itami, 1991; Helsel and Hirsch, 1992; Mueller et al., 1993). The response variable in LGR is the log of the odds ratio $p/(1-p)$, were p is the probability of a data value being in one of the categories. An S-shaped curve is used as a model for the probability (p) of a response. The probability tends to approach zero or one at the extreme values of the explanatory variables, and varies more rapidly near the center of its distribution (Helsel and Hirsch, 1992). The logistic transformation converts a variable constrained between zero and one, into a continuous variable that is linear with respect to the vector of the explanatory variables, X (Eq. (2)).

$$Y = \log(p/(1-p)) = b_0(b_{00}) + b_1 X_1 + \ldots b_k X_k \tag{2}$$

where Y is the response variable, $(p/(1-p))$ is the odds ratio, b_0 is the first intercept and b_{00} is the second intercept, X is the vector of k explanatory variables, b_1 is the slope coefficient for the first explanatory variable X_1, and b_k is the slope coefficient for the kth explanatory variable X_k.

Slope coefficients for the logistic equation are fit to the categorical data using a maximum likelihood method that optimizes the probability that the observed data will be estimated from the set of slope coefficients (Helsel and Hirsch, 1992). The ordinal logistic regression models presented here are comprised of two linear equations. The lines are parallel; only the intercept differs between the two equations. The modeled category was selected as the one with the maximum probability. Model accuracy is defined as the number of correct classifications (modeled category matches observed category)

Table 5
Multiple linear regression models that estimate the concentration of selected agricultural chemicals during 1989 post-planting runoff events in midwestern rivers

Agricultural chemical	Regression models	R-squared	Root mean square error
Atrazine	y = 50.41 + 0.242 ATRR + 0.164 FLWP − 5.945 TOPM + 0.153 PIGD	0.310	13.35
	y = 25.55 + 0.528 CRNP + 0.201 FLWP − 7.324 TOPM + 0.754 TEMP	0.312	13.33
	y = 59.78 + 0.330 CRPP + 0.218 FLWP − 7.647 TOPM	0.259	13.77
Alachlor	y = 53.52 + 0.285 ALAR + 0.074 FLWP − 3.729 TOPM − 0.141 PLNS	0.162	8.66
	y = 20.26 + 0.392 CRNP + 0.078 FLWP − 2.816 TOPM	0.211	8.37
	y = 22.24 + 0.445 CRNP + 0.095 FLWP − 3.168 TOPM − 0.311 ICRP	0.262	8.04
Cyanazine	y = 57.43 + 0.509 CYNR + 0.089 FLWP − 6.374 TOPM + 0.204 SOYP	0.270	8.24
	y = 106.2 + 0.224 ROWP + 0.078 FLWP − 8.053 TOPM − 69.33 PORO	0.274	8.22
	y = 95.31 + 0.403 CRNP + 0.093 FLWP − 7.040 TOPM − 69.75 PORO	0.291	8.12
Metolachlor	y = 31.36 + 0.290 METR + 0.057 FLWP − 2.710 TOPM − 0.056 PLNS	0.305	4.98
	y = 18.22 + 0.253 CRNP + 0.060 FLWP − 3.115 TOPM + 3.091 HYGP	0.239	5.21
	y = 47.30 + 0.147 CRPP + 0.058 FLWP − 3.495 TOPM − 40.623 PORO	0.204	5.33
Nitrate	y = − 15.16+ 1.290 NITR + 0.029 FLWP + 1.151 TOPM − 0.029 PIGD − 0.146 ICRP	0.517	2.51
	y = − 9.91+ 1.007 NITR + 0.017 FLWP + 1.638 HYGP + 2.271 Q + 0.661 PERM	0.485	2.59
	y = − 5.71+ 0.137 ROWP + 0.024 FLWP − 0.029 PIGD + 2.495 Q + 0.532 PERM	0.514	2.51

divided by the number of attempted classifications. Concentration categories for LGR models were picked such that near equal numbers of observations fell within each category.

3. Results and discussion

Linear, multiple linear, and logistic regression were used to develop simple statistical models that may be useful for forecasting the expected concentration of selected agricultural chemicals in midwestern rivers during post-planting runoff events. Linear regression and MLR were also used to develop statistical models that estimate annual transport, or annual mean concentration of agricultural chemicals in larger midwestern rivers. LGR was not tested because of the limited number of sites (ten) with annual transport or annual mean concentration estimates.

The statistical models help identify which explanatory variables were strongly related to the concentration, annual transport, or annual mean concentration of agricultural chemicals in midwestern rivers. Interpretations of the effects a particular explanatory variable on a response variable were based on the sign of the *B* or *b* values (Eqs. (1) and (2)). In MLR the *B* values are often called partial regression coefficients (Montgomery and Peck, 1982). Interpretation of the signs of partial regression coefficients is valid unless important regressors have been left out of the model or severe multicollinearity is present (Montgomery and Peck, 1982). The models presented in this paper were screened for these problems and hence can be used to make inferences concerning the relations among explanatory and response variables. While these inferences may have practical value especially when supported by common sense or mechanistic principals they should not be construed to represent or identify cause and effect type relations.

3.1. Agricultural chemical concentrations

Multiple linear regression was used to model the observed concentration of selected agricultural chemicals in 1989 post-planting samples from the Regional Reconnaissance study. These samples were collected during the first major period of runoff occurring after 50% or more of the corn planting and herbicide application in the basin was completed. Herbicide concentrations in midwestern rivers during this time period were likely to be at the highest levels observed during a year. Nitrate concentrations were likely to be elevated, but less likely to represent such a short-term peak concentration (Goolsby and Battaglin, 1993). The best models are listed in Table 5. A chemical use or cropland variable was significant in every model always having a direct relation to agricultural chemical concentration. These terms represent or are surrogate for the source of agricultural chemicals and so a direct relation to concentration was expected. The flow percentile of streamflow at the time of sample collection (FLWP) was a significant variable having a direct relation to concentration in all the models listed. High flows during the post-planting season are in part a result of increased surface runoff. In agricultural areas this runoff is likely to contain elevated concentrations of agricultural chemicals (Goolsby and Battaglin, 1995; Coupe et al., 1995). The TOPMODEL topographic parameter (TOPM) was significant in all herbicide models having an inverse relation to concentration. The implication is that during post-planting runoff events, higher concentrations of herbicides are likely to occur in basins that contain more steep slopes. This relation is consistent with the expected effects of slope on the speed and routing of Hortonian overland flow, and on the magnitude of the contribution of shallow subsurface flow to the storm hydrograph. However, TOPM also was significant in one model having a direct relation to nitrate concentration (Table 5) perhaps suggesting that elevated nitrate concentrations are not directly tied to increased overland runoff. Soybean planting date (PLNS) was significant in two models having an inverse relation to alachlor and metolachlor concentrations. Previous research has shown that herbicides concentration in storm runoff decrease with time after application (Thurman et al., 1992, 1994; Goolsby and Battaglin, 1995). Soil hydrologic group (HYGP) was significant in two models having a direct relation to metolachlor and nitrate concentrations. Larger values of HYGP are associated with poorly drained/clay rich soils which would tend to have higher runoff potential than well drained/sandy soils (Goss,

1992). Prime farmland underlain by poorly drained soils is also likely to be improved by subsurface (tile) or surface drainage. Tile drainage waters often contain elevated concentrations of agricultural chemicals but it is unclear if soil drainage increases or reduces the quantity of agricultural chemicals entering surface water systems (Keeney and Deluca, 1993; Bengtson et al., 1995; Fausey et al., 1995). Porosity (PORO) was significant in three models having an inverse relation to cyanazine and metolachlor concentrations. Soils with low PORO would tend to have higher runoff potential than soils with high PORO. Soil permeability (PERM) was significant in two nitrate models having a direct relation to concentration. This relation is somewhat counter intuitive and suggests that elevated nitrate concentration may not be directly tied to increased overland runoff, but could be associated with increased shallow subsurface flow. All of the MLR models listed in Table 5 have relatively low predictive power and could not be used to estimate expected agricultural chemical concentrations with confidence.

LGR was tested to determine if an improvement could be made over MLR models. LGR requires a discrete response variable; therefore, agricultural chemical concentration values were divided into three categories with roughly equal numbers of observations. For this analysis, these categories can be equated with low, medium, and high agricultural chemical concentrations. The concentration categories and accuracy of two of the best models for atrazine, alachlor, cyanazine, metolachlor, and nitrate are listed in Table 6.

The overall accuracy of the best LGR models, defined as the number of correct classifications (modeled concentration category matches observed category) divided by the number of attempted classifications, averaged about 65% (Table 6). LGR models were most effective at estimating when agricultural chemical concentrations were low (average accuracy of 73%) or high (average accuracy of 68%), but had more difficulty predicting intermediate concentrations (average accuracy of 51%).

As with the MLR models, a chemical use or cropland variable was significant in every model, always having a direct relation to agricultural chemical concentrations. FLWP was a significant variable in every model listed in Table 6, always having the expected direct relation to agricultural chemical concentration. TOPM was significant in all herbicide models, having an inverse relation to concentration. HYGP was significant in four models, having a direct relation to atrazine, alachlor, and metolachlor concentrations. PORO was significant in two models, having an inverse relation to cyanazine and nitrate concentrations. Hog and pig density (PIGD) was significant in three models, having a direct relation with atrazine, alachlor, and cyanazine concentrations. This relation is likely attributable to the frequent juxtaposition of hog production and feed corn production. Available water-holding capacity (AWC) was significant in two models, having a direct relation with alachlor and cyanazine concentrations. AWC is significantly positively correlated with percentage of basin in corn production (CRNP) and likely influences farming practices to some degree. Mean annual runoff (RNOF) was significant in three models, having a direct relation to metolachlor and nitrate concentrations. Greater runoff in agricultural areas may be associated with increased potential for transporting agricultural chemicals. Some variables, for example AWC and population density (POPD), were not significant in MLR models (Table 5), but were significant in some LGR models.

Figs 3 and 4 show a comparison of observed and estimated agricultural chemical classifications for atrazine, alachlor, cyanazine, metolachlor, and nitrate. The figures

Table 6
Comparison of observed agricultural chemical categories to classification from selected logistic regression models

Agricultural chemical	Model	Category	Number of observations	Percentage of correct classifications	Percentage classified low	Percentage classified high
Atrazine post-planting, 1989	1—Y = 7.38(9.55) + 0.11 ATRR + 0.04 FLWP + 2.32 HYGP − 1.97 TOPM + 0.01 PIGD	Less than 2.0 μg l^{-1}	43	72	–	28
		2.0–10.0 μg l^{-1}	39	56	13	31
		More than 10.0 μg l^{-1}	42	76	24	–
		Total	124	69	12	19
	2—Y = − 5.92(−3.67) + 0.11 CRNP + 0.05 FLWP + 2.18 HYGP − 1.84 TOPM + 0.25 TEMP	Less than 2.0 μg l^{-1}	43	77	–	23
		2.0–10.0 μg l^{-1}	39	46	21	33
		More than 10.0 μg l^{-1}	42	69	31	–
		Total	124	65	18	17
Alachlor post-planting, 1989	1—Y = − 1.26(1.06)+ 0.13 ALAR + 0.05 FLWP + 1.73 HYGP − 1.59 TOPM + 32.41 AWC	Less than 0.5 μg l^{-1}	48	75	–	25
		0.5–4.0 μg l^{-1}	43	60	14	26
		More than 4.0 μg l^{-1}	33	55	45	–
		Total	124	65	17	18
	2—Y = − 2.51(0.23)+ 0.04 CRPP + 0.05 FLWP − 0.85 TOPM + 0.05 PIGD + 2.27 Q	Less than 0.5 μg l^{-1}	48	77	–	23
		0.5–4.0 μg l^{-1}	43	65	19	16
		More than 4.0 μg l^{-1}	33	67	33	–
		Total	124	70	15	15
Cyanazine post-planting, 1989	1—Y = 1.95(3.70) + 0.15 CYNR + 0.04 FLWP − 1.07 TOPM + 7.96 Q_PPT + 0.02 PIGD	Less than 0.5 μg l^{-1}	54	80	–	20
		0.5–4.0 μg l^{-1}	34	32	38	30
		More than 4.0 μg l^{-1}	36	67	33	–
		Total	124	63	20	17
	2—Y = 23.93(25.71) + 0.05 CRPP + 0.04 FLWP − 1.49 TOPM − 42.22 PORO − 0.22 GRNP + 31.13 AWC	Less than 0.5 μg l^{-1}	54	72	–	28
		0.5–4.0 μg l^{-1}	34	32	47	21
		More than 4.0 μg l^{-1}	36	75	25	–
		Total	124	62	20	18

Metolachlor post-planting, 1989	1—Y = 8.04(10.66) + 0.16 METR + 0.04 FLWP + 1.61 HYGP − 2.09 TOPM + 0.08 ICRP + 6.61 Q_PPT	Less than 0.5 μg l^{-1}	40	75	–	25
		0.5–3.0 μg l^{-1}	47	64	11	25
		More than 3.0 μg l^{-1}	40	53	47	–
		Total	127	64	19	17
	2—Y = 7.40(9.94) + 0.07 ROWP + 0.04 FLWP + 0.27 RNOF − 1.53 TOPM − 0.52 PERM	Less than 0.5 μg l^{-1}	40	80	–	20
		0.5–3.0 μg l^{-1}	47	51	17	32
		More than 3.0 μg l^{-1}	40	55	45	–
		Total	127	61	21	18
Nitrate post-planting, 1989	1—Y = − 2.18(0.44)+ 0.74 NITR + 0.03 FLWP + 0.005 POPD + 0.14 RNOF − 12.14 PORO	Less than 1.0 mg l^{-1}	41	63	–	37
		1.0–4.0 mg l^{-1}	45	53	27	20
		More than 4.0 mg l^{-1}	40	80	20	–
		Total	126	65	16	19
	2—Y = − 11.58(−8.87) + 0.11 ROWP + 0.02 FLWP + 0.30 RNOF + 0.009 POPD + 0.06 PSTP + 0.03 COWD	Less than 1.0 mg l^{-1}	41	61	–	39
		1.0–4.0 mg l^{-1}	45	47	24	29
		More than 4.0 mg l^{-1}	40	78	22	–
		Total	126	61	16	23

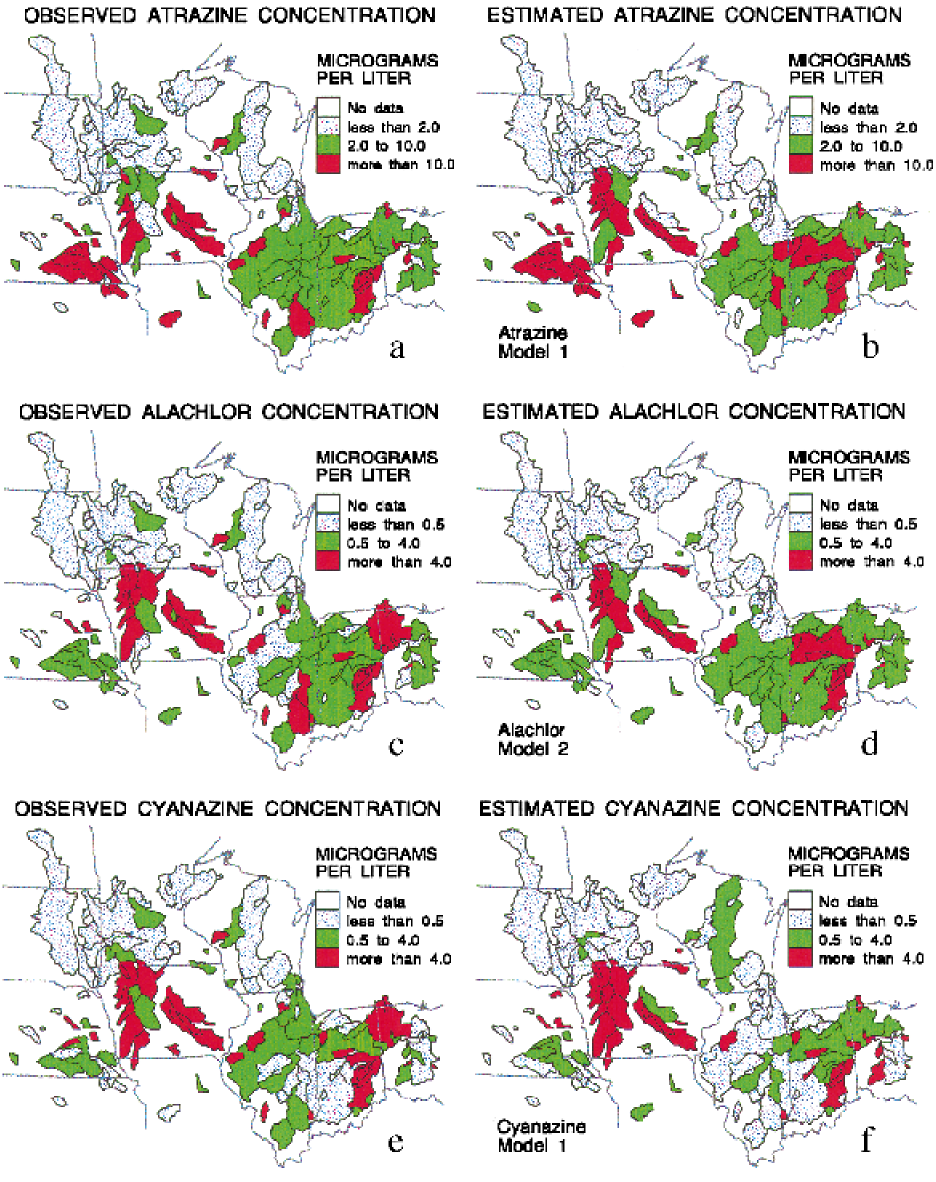

Fig. 3. Observed and estimated classifications of the 1989 post-planting concentration of (a + b) atrazine, (c + d) alachlor, and (e + f) cyanazine.

indicate that regional patterns of observed concentrations were well represented by LGR models. Atrazine LGR models were most effective at matching observed concentrations for basins in Kansas (KS), Nebraska (NE), Wisconsin (WI), Minnesota (MN), and Ohio (OH); and less effective in Illinois (IL), Indiana (IN), and Iowa (IA). Alachlor LGR models were most effective at matching observed concentrations for basins in KS, NE,

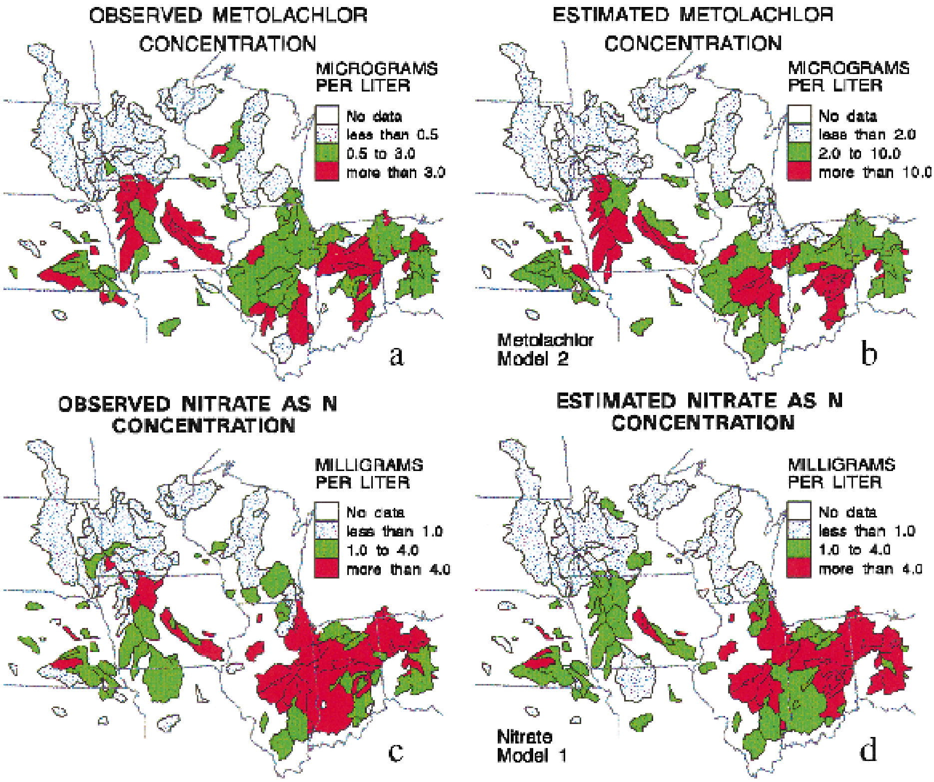

Fig. 4. Observed and estimated classifications of the 1989 post-planting concentration of (a + b) metolachlor, (c + d) nitrate as N.

and WI; and less effective in IL, IN, OH, MN, and IA. Cyanazine LGR models were most effective at matching observed concentrations for basins in KS, MN, IA, and IN; and less effective in IL, NE, WI, and OH. Metolachlor LGR models were most effective at matching observed concentrations for basins in KS, MN, IA, and OH; and less effective in IL, NE, WI, and IN. Nitrate LGR models were most effective at matching observed concentrations for basins in KS, NE, and IL; and less effective in WI, MN, IA, OH, and IN.

3.2. Annual transport and mean concentrations of agricultural chemicals

Linear and multiple linear regression were used to model observed annual transport or annual mean concentration of agricultural chemicals in the ten basins of the Mississippi River and Temporal Variability studies (Fig. 1). The best models of annual transport are given in Table 7, and the best models of annual mean concentration are given in Table 8. The slope coefficient for single variable linear regression models with agricultural chemical use as the explanatory variable and agricultural chemical transport as the response variable can be used to estimate chemical transport as a percentage of use. For the ten basins, significant models ($P < 0.05$) were determined when estimates of annual

Table 7
Regression models of the annual transport of selected agricultural chemicals in ten midwestern rivers

Agricultural chemical	Regression models	*R*-squared	Root mean square error
Atrazine	ATRT = − 8.07 + 0.0153 ATRU	0.988	12.15
	ATRT = − 7.45 + 6.27 × 10^{-6} CRNH	0.968	19.81
	ATRT = − 6.04 + 4.57 × 10^{-6} CRNH + 0.0055 FLOW	0.988	12.81
Alachlor	ALAT = 0.55 + 0.0022 ALAU	0.818	6.76
	ALAT = − 0.78+ 0.0046 ALAU − 0.0026 FLOW	0.947	3.91
	ALAT = 0.17 + 8.55 × 10^{-7} CRNH	0.886	5.35
Cyanazine	CYNT = − 2.45+ 0.0166 CYNU	0.970	8.66
	CYNT = − 3.33+ 2.85 × 10^{-6} CRNH	0.978	7.39
	CYNT = − 2.86+ 0.020 CYNU − 0.0019 FLOW	0.981	7.39
Metolachlor	METT = − 1.16 + 0.0079 METU	0.994	3.42
	METT = − 0.42 + 2.88 × 10^{-6} SOYH	0.990	4.52
	METT = − 0.63 + 3.38 × 10^{-6} SOYH − 0.0014 FLOW	0.994	3.58
Nitrate	NITT = − 14031+ 0.1532 NU91	0.943	82778
	NITT = − 9029+ 0.0222 SOYH	0.972	58424
	NITT = − 13761+ 0.0159 PIGS + 0.0039 CHCK	0.970	64511

transport were regressed with estimates of annual use for atrazine, alachlor, cyanazine, metolachlor, and nitrate as nitrogen (Table 7) (Battaglin et al., 1993). The slope coefficients for these models indicate that, on an annual basis, estimated masses equivalent to about 1.5% of the atrazine, 0.4% of the alachlor, 1.7% of the cyanazine, 0.8% of the metolachlor, and 15% of the nitrogen applied within the ten basins studied were transported out of those basins in river water. However, previous results indicate that there is significant basin-to-basin variability in the estimates of transport as a percentage of use (Battaglin and Goolsby, 1994). Results presented here are for April 1, 1991 to March 31, 1992; the effects of climatic conditions and other variables that might change from year to year were not addressed in this paper. Estimates of the transport of atrazine to the Gulf of Mexico from other years show significant variability (Goolsby et al., 1993), indicating that estimates of annual atrazine transport as a percentage of use would also show significant variability.

Both the annual transport and mean annual concentration of the selected agricultural chemicals were fairly well modeled by the simple regression models. However, no significant models were discovered for annual mean cyanazine concentration. The limited sample size (10) of this data set precludes broad extrapolation of the results presented here to the rivers of the midwest in general. A chemical use, cropland, or livestock variable was significant in every model, having a direct relation to agricultural chemical transport and mean annual concentration. In many cases, models that used cropland estimates as surrogates for the harder to come by agricultural chemical use estimates worked nearly as well as models using the chemical use estimates. FLOW was significant in one transport MLR models for each herbicide, but was not significant in the nitrate transport MLR models. It was expected that FLOW would be directly related to chemical transport if the solute source was distributed and the solute behaved conservatively in the river. Atrazine appeared to fit this expectation, but FLOW was inversely related to the annual transport of alachlor, cyanazine, and metolachlor (Table 7) suggesting that instream transformations

Table 8
Regression models of the mean annual concentration of selected agricultural chemicals in ten midwestern rivers

Agricultural chemical	Regression models	*R*-squared	Root mean square error
Atrazine	ATRC = − 0.28+ 0.0829 ATRR	0.615	1.113
	ATRC = 0.15 + 0.0263 ATRR + 0.0990 ICRP	0.992	0.172
	ATRC = 0.14 + 0.0265 CRNP + 0.1051 ICRP	0.987	0.222
	ATRC = 0.08 + 0.0298 CRNP + 0.5362 SRGP	0.986	0.227
Alachlor	ALAC = − 0.10+ 0.0205 CRNP	0.522	0.274
	ALAC = − 0.02+ 0.0076 CRNP + 0.0221 ICRP	0.995	0.031
	ALAC = − 0.03+ 0.0071 SOYP + 0.0273 ICRP	0.993	0.036
	ALAC = − 0.01+ 0.0068 ALAR + 0.0249 ICRP	0.992	0.039
Metolachlor	METC = 0.09 + 0.0141 METR + 0.0919 SRGP	0.943	0.080
	METC = 0.04 + 0.0193 CRNP	0.735	0.162
	METC = 0.08 + 0.0133 CRNP + 0.0553 SRGP	0.916	0.097
Nitrate	NITC = 0.52 + 0.1611 SOYP	0.863	0.811
	NITC = − 0.60+ 0.0817 CRPH	0.658	1.280
	NITC = 1.88 + 0.1538 SOYP − 0.0078 POLD	0.910	0.701

or other processes may significantly affect the concentrations of these three herbicides in rivers. However, relatively high VIF values for the explanatory variables in these relations (4.52, 7.97, 4.87, and 7.55, respectively) suggest that the sign of the FLOW coefficient could be misleading. Irrigated cropland as a percentage of basin area (ICRP) was significant and directly correlated with atrazine and alachlor concentrations in several models (Table 8). Drainage from irrigated cropland can transport herbicides to both surface water and groundwater (Board on Agriculture, National Research Council, 1993). Relations with ICRP were strongly influenced by Site 9 (Table 1) which has a large percentage of the basin as irrigated cropland (44%) relative to the other nine sites (all less than 8%) and elevated atrazine and alachlor concentrations (Table 3). Harvested sorghum as a percentage of basin area (SRGP) was significant and directly correlated with atrazine and metolachlor concentrations. Both atrazine and metolachlor are used to control weeds in sorghum crops. Relations with SRGP were strongly influenced by Site 9 (Table 1) which has a large percentage of the basin as sorghum (8.8%) relative to the other nine sites (all less than 1.1%), and elevated atrazine and metolachlor concentrations (Table 3). Livestock variables representing the number of hogs and pigs (PIGS) and number of chickens (CHCK) were significant and directly correlated with nitrate transport in one model (Table 7). These terms are surrogates for among other things manure production. However, the density of poultry within a basin (POLD) was significant and inversely related to nitrate concentration, perhaps because of correlation with the other explanatory variable (SOYP) (Table 8).

4. Conclusions

Results presented in this paper demonstrate that various regression techniques can be used to model the concentration, annual transport, and annual mean concentration of

selected agricultural chemicals in midwestern rivers. The regression models developed for this paper identify land use, chemical use, soil, and climatic variables in upstream drainage basins that affect agricultural chemical concentrations in rivers. These models can contribute to the understanding of agricultural chemical concentrations in midwestern rivers and ultimately may prove helpful to researchers, water suppliers, and water managers attempting to forecast when and where water supplies may not meet health-based drinking-water standards.

The success of the regression techniques used to model the concentration, annual transport, and annual mean concentration of selected agricultural chemicals in midwestern rivers was variable. Three and four variable MLR models were not very successful (R^2 from 0.162 to 0.312) in explaining the variance in observed herbicide concentrations during 1989 post-planting runoff in midwestern rivers. MLR models of nitrate concentration were somewhat better (R^2 from 0.485 to 0.517), but utilized more (five) explanatory variables. Five and six variable LGR models were moderately successful at identifying expected categories of agricultural chemical concentration, correctly matching the observed concentration category in 61–70% of observations. One and two variable regression models were successful (R^2 from 0.818 to 0.994) in explaining the variance in observed annual transport of agricultural chemicals. One and two variable regression models were nearly as successful (R^2 from 0.522 to 0.995) in explaining the variance in observed mean annual concentrations of agricultural chemicals. However, the limited number of sites (ten) with estimates of the annual transport and annual mean concentration of agricultural chemicals precludes broad extrapolation of these results.

All of the regression models presented in this paper support the hypothesis that the dominant process controlling agricultural chemical concentration and mass transport in midwestern rivers was the use of those agricultural chemicals within the associated drainage basins. Hence, farming best management practices targeted at reducing the use of agricultural chemicals and their lost to surface water or groundwater may prove successful in lowering the post-planting runoff concentration, annual transport, and annual mean concentration of those chemicals in midwestern rivers. Other processes also affected agricultural chemical concentration and mass transport in midwestern rivers. Flow conditions at the time of sample collection had a strong affect on post-planting agricultural chemical concentrations with higher concentration observed during high streamflow. Basin slope often had a significant effect on post-planting herbicide concentrations, higher concentrations occurred more often in basins that contained more steep slopes. Soil characteristics had a significant effect on post-planting agricultural chemical concentrations, with higher concentrations observed in rivers that drained basins with soils that had low porosity, high permeability, high water holding capacity, or poor drainage.

The results from this paper support the importance of the continued collection and publication of water quality data and information and statistics on agricultural land use and agricultural practices. Although some of the explanatory variables that were significant in the regression models presented in this paper are relatively static (for example, soil hydrologic group), many of the most important variables (agricultural chemical use, crop acreages, and streamflow) vary from year to year. For these models to be effective at forecasting concentrations or transport amounts, accurate and timely estimates of agricultural chemical use, cropland, and flow variables are required.

Acknowledgements

This research was conducted as part of the US Geological Survey's Toxic Substances Hydrology Program with support from the National Water Quality Assessment Program, the National Research Program and numerous USGS employees at local offices. I express my gratitude for critical and editorial reviews from D. Kolpin, J. Stamer, J. Flager, P. Van Cappellen, D. Helsel, D. Cooper, and one anonymous reviewer.

References

Adams, C.D., Randtke, S.J., Thurman, E.M., Hulsey, R.A., 1990. Atrazine and its degradation products in soil and ground water and the effectiveness of water-treatment processes for their removal. In: Proc. of the 40th Annual University of Kansas Environmental Engineering Conference.

Baker, D.B., Richards, R.P., 1990. Transport of soluble pesticides through drainage networks in large agricultural river basins. In: Kurtz, D.A. (Ed.), Long Range Transport of Pesticides. Lewis Publishers, Inc., Chelsea, MI.

Battaglin, W.A., Goolsby, D.A., 1994. Relations between herbicide use and annual transport of herbicides in midwestern rivers, 1991–1992. In: Weigmann, D.L. (Ed.), New Directions in Pesticide Research, Development, Management, and Policy, Proceedings of the Fourth Conference on Pesticides, 1–3 November 1993. Virginia Water Resources Research Center, Blacksburg, VA, pp. 711–723.

Battaglin, W.A., Goolsby, D.A., 1995. Spatial Data in Geographic Information System Format on Agricultural Chemical Use, Land Use, and Cropping Practices in the United States. US Geol. Surv. Water-Resour. Invest. Rep. 94-4176.

Battaglin, W.A., Hay, L.E., 1996. Effects of sampling strategies on estimates of annual mean herbicide concentrations in midwestern rivers. Environ. Sci. Tech. 30 (3), 889–896.

Battaglin, W.A., Goolsby, D.A., Coupe, R.H., 1993. Annual use and transport of agricultural chemicals in the Mississippi River, 1991–92. In: Goolsby, D.A., Boyer, L.L., Mallard, G.E. (Eds.), Selected Papers on Agricultural Chemicals in Water Resources of the Midcontinental United States. US Geol. Surv. Open-File Rep. 93-418.

Bengtson, R.L., Carter, C.E., Fouss, J.L., Southwick, L.M., Willis, G.H., 1995. Agricultural drainage and water quality in Mississippi delta. J. Irrig. Drain. Eng., ASCE 121 (4), 292–295.

Board on Agriculture, National Research Council, 1993. Soil and Water Quality—An Agenda for Agriculture. National Academy Press, Washington, DC, 516 pp.

Coupe, R.H., Goolsby, D.A., Iverson, U.L., Markovchick, D.J., Zaugg, S.D., 1995. Pesticide, Nutrient, Water-Discharge and Physical Property Data for the Mississippi River and Some of its Tributaries, April 1991–September 1992. US Geol. Surv. Open-File Rep. 93-657.

Dunn, D.D., 1996. Trends in Nutrient Inflows to the Gulf of Mexico from Streams Draining the Conterminous United States, 1972–93. US Geol. Surv. Water-Resour. Invest. Rep. 96-4113.

Edwards, T.K., Glysson, G.D., 1988. Field Methods for Measurement of Fluvial Sediment. US Geol. Surv. Open-File Rep. 86-53.

Fausey, N.R., Brown, L.C., Belcher, H.W., Kanwar, R.S., 1995. Drainage and water quality in Great Lakes and cornbelt states. J. Irrig. Drain. Eng., ASCE 121 (4), 283–288.

Fishman, M.J., Friedman, L.C., 1989. Methods for the Determination of Inorganic Substances in Water and Fluvial Sediments. US Geol. Surv. Techniques of Water Resour. Invest., Book 5.

Gianessi, L.P., Puffer, C., 1991. Herbicide Use in the United States. Resources for the Future, Washington, DC.

Goolsby, D.A., Battaglin, W.A., 1993. Occurrence, distribution, and transport of agricultural chemicals in surface waters of the midwestern United States. In: Goolsby, D.A., Boyer, L.L., Mallard, G.E. (Eds.), Selected Papers on Agricultural Chemicals in Water Resources of the Midcontinental United States. US Geol. Surv. Open-File Rep. 93-418, pp. 1–24.

Goolsby, D.A., Battaglin, W.A., 1995. Occurrence and distribution of pesticides in rivers of the midwestern United States. In: Leng, M.L., Leovey, E.M.K., Zubkoff, P.L. (Eds.), Agrochemical Environmental Fate: State of the Art. CRC Press Inc., Boca Raton, FL, pp. 159–173.

Goolsby, D.A., Coupe, R.H., Markovchick, D.J., 1991. Distribution of Selected Herbicides and Nitrate in the Mississippi River and its Major Tributaries, April Through June 1991. US Geol. Surv. Water-Resour. Invest. Rep. 91-4163.

Goolsby, D.A., Battaglin, W.A., Thurman, E.M., 1993. Occurrence and Transport of Agricultural Chemicals in the Mississippi River During July and August 1993. US Geol. Surv. Circ. 1120-C.

Goss, D.W., 1992. Screening procedure for soils and pesticides for potential water quality impacts. Weed Technol. 6, 701–708.

Helsel, D.R., Hirsch, R.M., 1992. Statistical Methods in Water Resources. Elsevier, New York.

Karl, T.R., Riebsame, W.E., 1989. The impact of decadal fluctuations in mean precipitation and temperature on runoff: a sensitivity study over the United States. Clim. Change 15 (3), 423–447.

Karl, T.R., Williams, C.N., Quinlan, F.T., Boden, T.A., 1990. United States Historical Climatology Network (HCN) Serial Temperature and Precipitation Data. Oak Ridge National Laboratory, Environmental Sciences Division Publication No. 3404.

Keeney, D.R., Deluca, T.H., 1993. Des Moines river nitrate in relation to watershed agricultural practices: 1945 versus 1980s. J. Environ. Qual. 22, 267–272.

Meister, R.T., Sine, C. (Eds.), 1995. Farm Chemicals Handbook ’95. Meister Publishing Company, Willoughby, OH.

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

Mueller, D.K., Ruddy, B.C., Battaglin, W.A., 1993. Relation of nitrate concentrations in surface water to land use in the upper-midwestern United States, 1989–90. In: Goolsby, D.A., Boyer, L.L., Mallard, G.E. (Eds.), Selected Papers on Agricultural Chemicals in Water Resources of the Midcontinental United States. US Geol. Surv. Open-File Rep. 93-418, pp. 41–50.

Pereira, J.M., Itami, R.M., 1991. GIS-based habitat modeling using logistic multiple regression: a study of the Mt. Graham red squirrel. Photogramm. Eng. Remote Sensing 57 (11), 1475–1486.

Pereira, W.E., Rostad, C.E., 1990. Occurrence, distribution, and transport of herbicides and their degradation products in the lower Mississippi River and its tributaries. Environ. Sci. Technol. 24 (9), 1400–1406.

Pereira, W.E., Rostad, C.E., Leiker, T.J., 1989. Preliminary assessment of the fate and transport of synthetic organic agrochemicals in the lower Mississippi River and its tributaries. In: Mallard, G.E., Ragone, S.E. (Eds.), US Geol. Surv. Water-Resour. Invest. Rep. 88-4220, pp. 453–464.

Rea, A., Cederstrand, J.R., 1994. GCIP Reference Data Sets (GREDS). US Geol. Surv. Open-File Rep. 94-388 (CD-ROM).

Sandstrom, M.W., Wydoski, D.S., Schroeder, M.P., Zamboni, J.L., Foreman, W.T., 1991. Methods of Analysis by the National Water Quality Laboratory: Determination of Organonitrogen Herbicides in Water by Solid Phase Extraction and Capillary Column Gas Chromatography/Mass Spectrometry with Selected Ion Monitoring. US Geol. Surv. Open-File Rep. 91-519.

SAS Institute, 1990, SAS/STAT User’s Guide, Version 6, 4th ed. SAS Institute, Cary, NC.

Scribner, E.A., Goolsby, D.A., Thurman, E.M., Meyer, M.T., Mills, M.S., Pomes, M.L., 1993. Reconnaissance Data for Selected Herbicides, Two Atrazine Metabolites, and Nitrate in Surface Water of the Midwestern United States, 1989–90. US Geol. Surv. Open-File Rep. 93-457.

Scribner, E.A., Thurman, E.M., Goolsby, D.A., Meyer, M.T., Pomes, M.L., 1994. Concentrations of Selected Herbicides, Two Triazine Metabolites, and Nutrients in Storm Runoff From Nine Stream Basins in the Midwestern United States, 1990–92. US Geol. Surv. Open-File Rep. 94-396.

Squillace, P.J., Thurman, E.M., 1992. Herbicide transport in rivers: importance of hydrology and geochemistry in nonpoint-source contamination. Environ. Sci. Technol. 26 (3), 538–545.

Thurman, E.M., Meyer, M.T., Pomes, M.L., Perry, C.E., Schwab, A.P., 1990. Enzyme-linked immunosorbent assay compared with GC/MS for the determination of herbicides in water. Anal. Chem. 62, 2043–2048.

Thurman, E.M., Goolsby, D.A., Meyer, M.T., Mills, M.S., Pomes, M.L., Kolpin, D.W., 1992. A reconnaissance study of herbicides and their metabolites in surface water of the midwestern United States using immunoassay and gas chromatography/mass spectrometry. Environ. Sci. Technol. 26 (12), 2440–2447.

Thurman, E.M., Meyer, M.T., Mills, M.S., Zimmerman, L.R., Perry, C.A., Goolsby, D.A., 1994. Formation and transport of deethylatrazine and deisopropylatrazine in surface water. Environ. Sci. Technol. 28 (13), 2267–2277.

US Department of Agriculture, 1984. Usual Planting and Harvesting Dates for U.S. Field Crops. US Department of Agriculture, Statistical Reporting Service, Agriculture Handbook Number 628.

US Department of Agriculture, 1993. State Soil Geographic Data Base (STATSGO)—Data Users Guide. US Department of Agriculture, Soil Conservation Service, Misc. Publication Number 1492.

US Department of Commerce, 1989. Census of Agriculture, 1987—Final County File. US Department of Commerce, Bureau of the Census, Washington, DC (machine-readable data file).

US Department of Commerce, 1990. Census of Population and Housing, 1990. US Department of Commerce, Bureau of the Census, Data Users Division, Washington, DC.

US Environmental Protection Agency, 1989. Drinking Water Health Advisory: Pesticides. Office of Drinking Water Health Advisories, Lewis Publishers, Inc., Chelsea, Michigan.

US Environmental Protection Agency, 1990. County-Level Fertilizer Sales Data. US Environmental Protection Agency, Office of Policy, Planning, and Evaluation PM-221.

US Environmental Protection Agency, 1991. Organic chemicals other than total trihalomethanes, sampling and analytical requirements (section 141.24 or part 141, national primary drinking-water regulations). US Code of Federal Regulations, Title 40, 100-149, revised 1/30/1991, pp. 3583–3587.

US Environmental Protection Agency, 1995. Drinking Water Regulations and Health Advisories. US Environmental Protection Agency, Office of Water, Washington, DC.

Wolock, D.M., 1993. Simulating the Variable-Source-Area Concept of Streamflow Generation with the Watershed Model TOPMODEL. US Geol. Surv. Water-Resour. Invest. Rep. 93-4124.