3. DATA AND COMPUTER SOFTWARE DESCRIPTION

3.1 Data sources

The data used in this study are divided into five groups:

1) Measurements of atrazine and nitrate concentrations in surface waters of the Midwest as well as the location of the sampling sites and the stream flow rate at the time chemical measurements were made. The atrazine and nitrate data are described in Section 3.1.1.

2) Digital elevation models (DEM) and digital stream maps (Reach File 1 or RF1). They are used to determine "flow paths" of travel of the chemical from the point of application to the basin outlet. The DEMs are introduced in Section 3.1.2 and the RF1 is discussed in Section 3.1.3.

3) Data that define the source of chemicals in the system, i.e., maps of annual application of agricultural chemicals. The maps of fertilizer and herbicide application are presented in Section 3.1.4.

4) Hydrologic and climatic record. The hydrologic and climatic data serve mainly for the development of the flow model, which spatially and temporally distributes recorded flow rate. The hydrologic data are discussed in Section 3.1.5.

5) Maps of average annual temperature and precipitation. These maps are utilized to characterize the climatic regions in which the sampled watersheds are located. They are introduced in Section 3.1.6.

3.1.1 Herbicide and nutrient data

The data utilized in this study has been extracted from two USGS publications:

1) Open-File Report 94-396, "Concentrations of Selected Herbicides, Two Triazine Metabolites, and Nutrients in Storm Runoff from Nine Stream Basins in the Midwestern United States, 1990-92" (Scribner, et al., 1994);

2) Open-File Report 93-457, "Reconnaissance Data for Selected Herbicides, Two Atrazine Metabolites, and Nitrate in Surface Water of the Midwestern United States" (Scribner, et al., 1993).

Table 3.1 Description of atrazine and nitrate plus nitrite as nitrogen samples published in USGS Open-File Report 94-396, (Scribner, et al., 1994).

Name	Station ID	Drainage Area km^2	Sampled Period	Number of samples		No of sampled days	
				Atrazine (GCMS)	NO_2+NO_3 as Nitrogen	Atrazine (GCMS)	NO_2+NO_3 as Nitrogen
West Fork Big	6880800	3,123	4/03/90-7/27/90	37	176	28	57
Blue R., NE			3/26/91-3/09/92	106	96	64	63
Sangamon R., IL	5572000	1,425	4/04/90-8/23/90	53	255	43	101
			4/12/91-3/25/92	198	168	93	93
Huron R., OH	4199000	961	3/30/90-8/20/90	59	208	53	142
Delaware R., KS	6890100	995	4/04/90-6/29/90	32	154	25	45
Roberts Cr., IA	5412100	267	4/03/90-7/21/90	22	133	20	43
Old Man's Cr. IA	5455100	521	4/16/90-7/02/90	49	171	30	46
Cedar R., IA	5431200	12,261(1)	4/15/90-6/29/90	41	187	33	60
Silver Cr., IL.	5594800	1,202	4/13/90-8/17/90	34	120	29	59
Iroquois R., IL	5526000	5,416	4/04/90-8/18/90	49	177	46	71
			Total:	680	1845	464	780

(1) for analysis a drainage area estimated from DEM has been used (17,409 km^2)

USGS study of agrichemicals in storm runoff. USGS Open-File Report 94-396 contains data that were collected from nine streams in five Midwestern States during late spring-early summer of 1990. Two of the nine streams (the Sangamon River and the West Fork Big Blue River) were sampled from April 1991 to March 1992. Table 3.1 contains detailed information about number of samples and the time interval during which samples were collected. Figure 3.1 shows the location of sampling sites, watersheds that were delineated from the 500 m DEM (Digital Elevation Model), and selected rivers from RF1 (digital map of rivers).

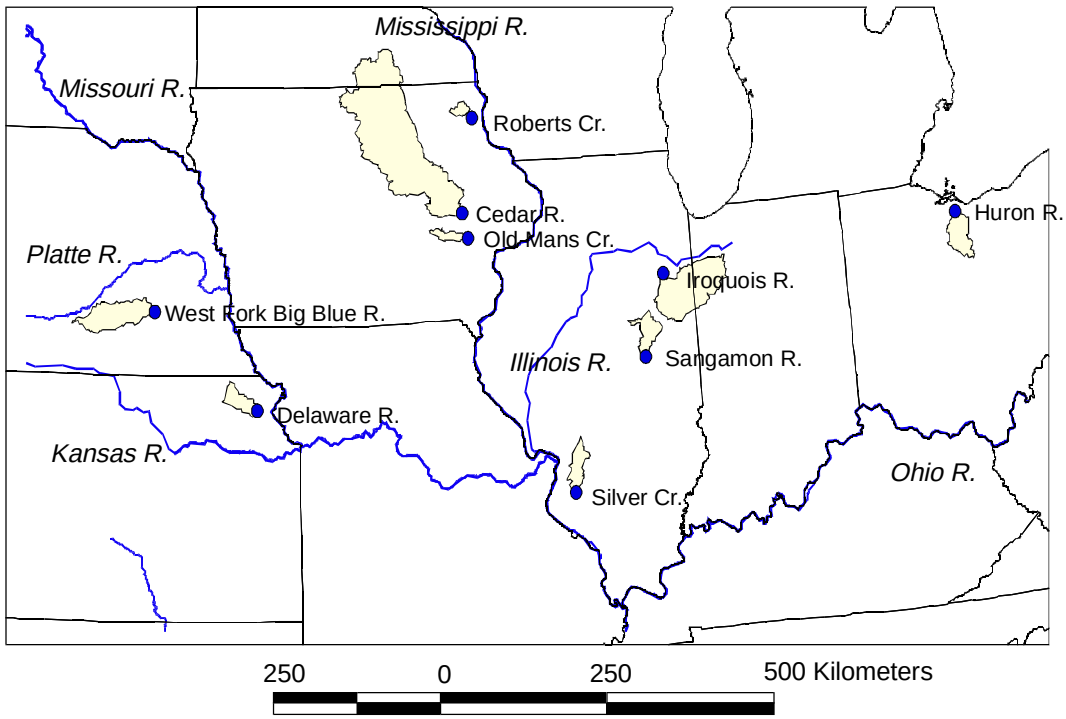

Figure 3.1 Location of sampling sites for study of atrazine and nitrate plus nitrite as nitrogen concentrations in storm runoff; USGS Open-File Report 94-396, (Scribner, et al., 1994).

Concentrations of eleven herbicides, two triazine metabolites, nutrients, streamflow, specific conductance, and pH were measured. The location of each sampling site is described by latitude and longitude. Such estimates as cropland percentage, atrazine use, and nitrogen use are also presented in Open-File Report 94-396. Figure 3.2 presents an example of measured atrazine concentrations in Old Mans Creek in Iowa.

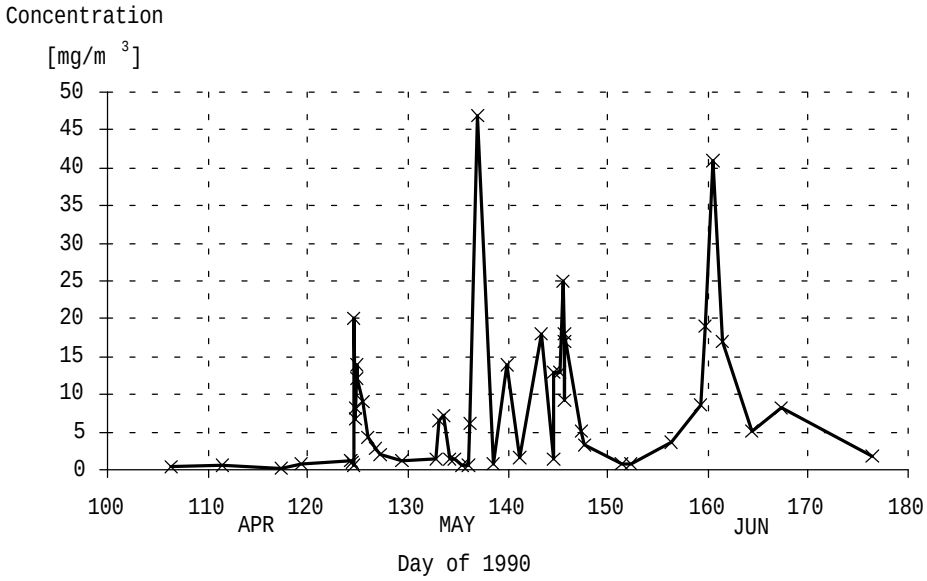

Figure 3.2 Atrazine Concentrations in Old Mans Creek, Iowa (data from USGS Open-File Report 94-396, Scribner et al., 1994).

Reconnaissance study of 147 streams. In Open-File Report 93-457 Scribner, et al., (1993) present data for selected herbicides, two atrazine metabolites, and nitrate

plus nitrite as nitrogen in surface water of the Midwestern United States, 1989-90. One hundred and forty seven Midwestern streams were studied. Two to three samples per site were taken annually--before application of herbicides (March or April), during the first major runoff after application of herbicides (May, June, or July), and in the Fall during a low-flow period when most of the streamflow was derived from the ground water (August - November). Table 3.2 presents the distribution of samples within studied period and Figure 3.3 shows the reconnaissance study drainage areas.

Table 3.2 Temporal distribution of atrazine and nitrate plus nitrite as nitrogen samples from the Midwestern reconnaissance study; USGS Open-File Report 93-457, (Scribner, et al., 1993).

Month	Number of samples			
	Atrazine (GC/MS)		Nitrate + Nirite as Nitrogen	
	1989	1990	1989	1990
MAR	64	38	95	38
APR	25	10	48	10
MAY	69	36	72	37
JUN	51	13	51	13
JUL	11	5	12	5
AUG	1	-	1	-
SEP	1	-	1	-
OCT	114	-	114	-
NOV	30	-	32	-
Total:	366	102	426	103

Atrazine concentrations are reported as two different (but related) components: ELISA atrazine and GC/MS atrazine, since this chemical is usually measured by two methods: ELISA (Enzyme-Linked Immunosorbent Assay) and GC/MS (gas chromatography/mass spectrometry). The ELISA test is easier to perform but is less precise than the GC/MS method. It is affected to varying degrees by the existence of propazine, prometon, simazine, deethylatrazine, cyazine, and deisopropylatrazine (Thurman, et al., 1990).

ELISA and GC/MS results can be related by regression so that results determined by ELISA can be converted to an estimate of what a GC/MS measurement of the same sample would have yielded (Thurman, et al., 1992; Moody and Goolsby, 1993; Goolsby, et al., 1993; Gruessner, et al., 1995). In this study, the atrazine concentrations measured by the GC/MS method are utilized.

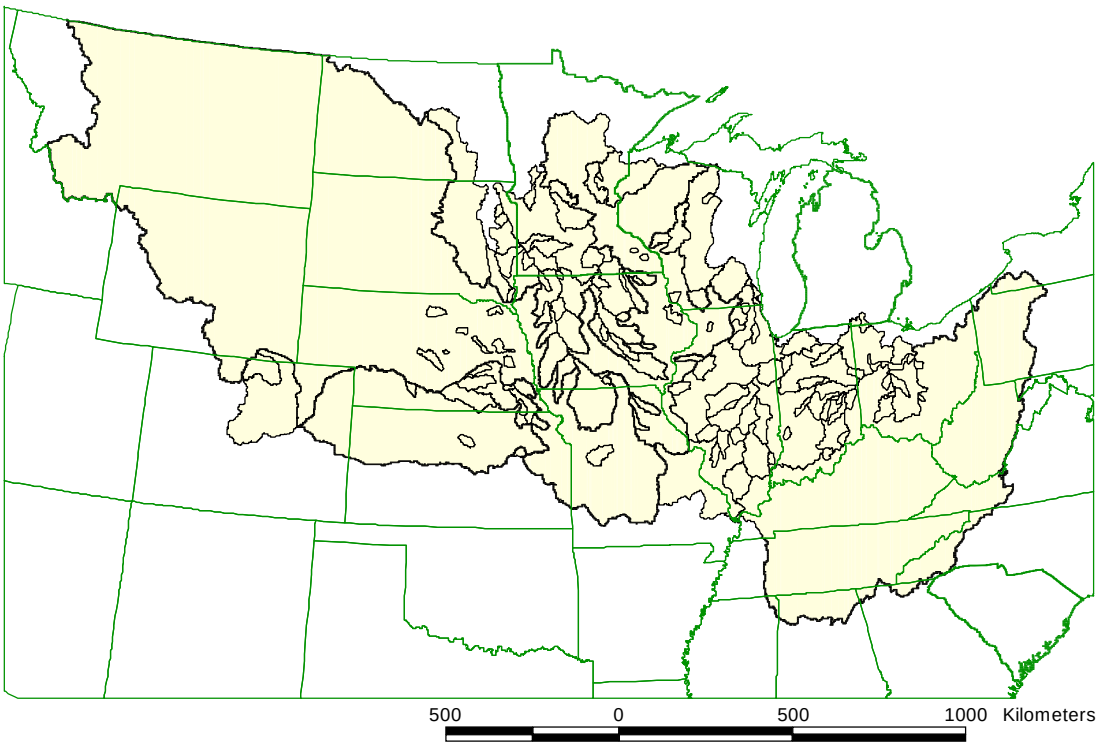

Figure 3.3 Watersheds sampled during 1989 and 1990 for the USGS reconnaissance study (Battaglin, 1995).

3.1.2 Digital Terrain Representation

Elevation data form the spatial framework for modeling basic hydrologic processes. In this research, a Digital Elevation Model is used not only to determine the paths of travel of agricultural chemicals, but also to divide the study region into small watersheds or modeling units, to determine the "flow" topology between the modeling units, to delineate the stream network, and to determine the characteristics of the drainage area. Two versions of DEM are utilized: 100 m (derived from 3-second) DEM is applied for the Iowa-Cedar River watershed and 500 m (15 second) DEM is used for the Upper Mississippi-Missouri and Ohio River basin.

The Digital Elevation Model (DEM) consists of a 2-D array of ground positions at regularly spaced intervals. One-Degree DEM files (3x3 arc-second data spacing) are available throughout the US and therefore they have been used in this research. The majority of 1-Degree DEMs were produced by the Defense Mapping Agency (DMA) either from cartographic (maps 1:24,000 scale through 1:250,000 scale) or from photographic sources. Some of these DEMs were created by regriding 7.5-minute and 30-minute DEMs. The compressed (and uncompressed) DEM files are available via the Internet at ftp://edcftp.cr.usgs.gov/pub/data/DEM/250/. The maps needed can also be selected from the map of the USA that is displayed on the screen: http://edcwww.cr.usgs.gov/glis/hyper/guide/1_dgr_demfig/index1m.html

The 500m DEM was created by resampling 1-Degree DEM files. It has been released by the USGS on a CD-ROM (Rea and Cederstrand, 1995). In this research, the 500 km DEM has been applied to create a map of the drainage area of each sampling site from which measurements have been utilized for model development. Although for some regions the quality of determined flow direction was poor and some editing had to be done, the 500 m DEM appeared to have an optimal resolution. At the cost of precision, the Mississippi-Missouri-Ohio River basin of area about

$2.4*10^6$ km^2 could be treated as a one processing unit using 500m or 15” data, which was impossible for a 3” DEM. Since the size of this basin grid is about 90 Mb (500 m DEM, integer values), the complexities of the analysis are restricted by the available computer power and memory.

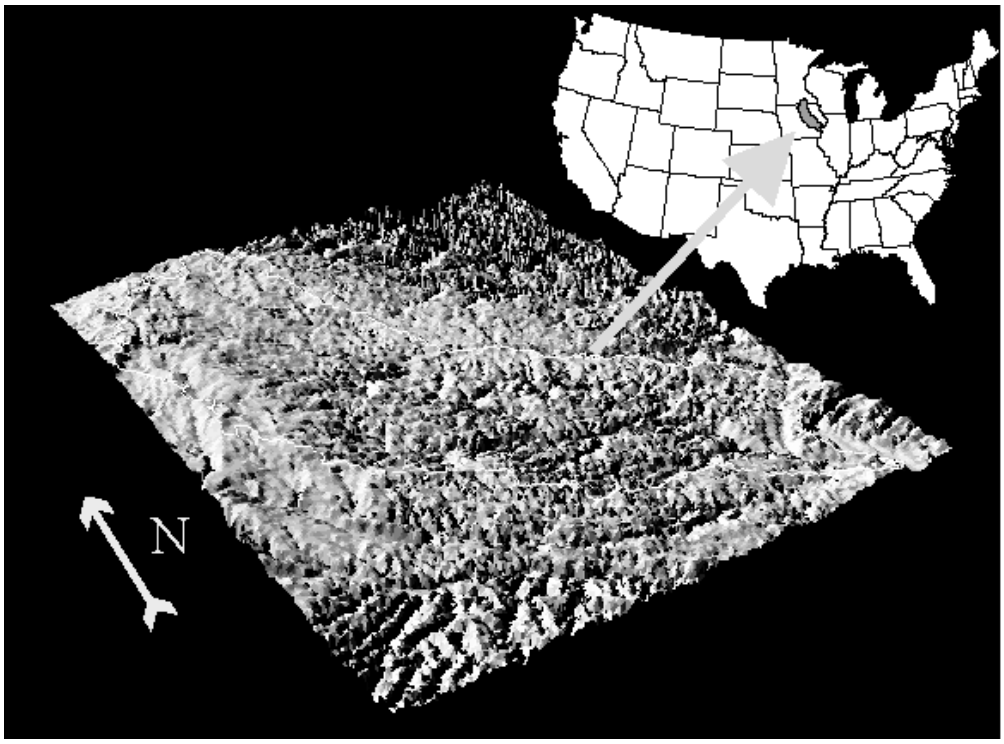

Figure 3.4 A Digital Elevation Model of the Iowa-Cedar River basin, vertical scale enlarged 500 times. The white line represents the basin boundaries.

3.1.3 Reach File 1

Reach File 1 (RF1) is a representation of streams in the conterminous United States at a scale of approximately 1:500,000. The original file was prepared by the US Environmental Protection Agency (USEPA). In 1994 it was translated from a

mainframe computer into an Arc/INFO coverage. RF1 in Arc/Info export format is available via Internet: http://nsdi.usgs.gov/nsdi/wais/water/rf1.HTML.

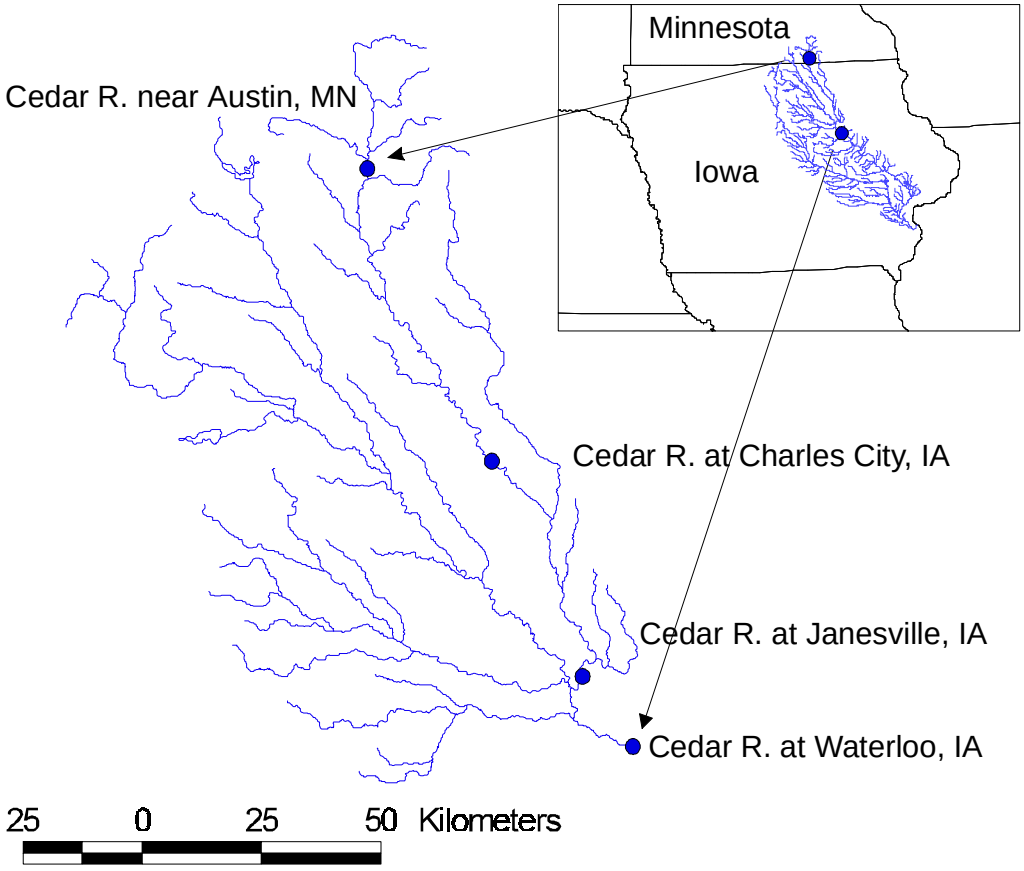

Figure 3.5 The Cedar River above Waterloo, IA; rivers selected from Reach File 1.

Because of the size of the entire RF1 data set (its Arc/INFO Export file occupies 57 MB), it also is available in 18 separate files, each covering a 2-digital hydrologic unit code (Water Resources Region).

In this research, RF1 is applied to adjust the DEM in order to ensure that the streams delineated from the DEM are compatible with the streams from RF1. Figure 3.5 shows an example of RF1 for the Cedar River above Waterloo, Iowa.

3.1.4 Atrazine and nitrogen fertilizer use

In 1995, the USGS published maps of herbicide and nitrogen-fertilizer use (Battaglin and Goolsby, 1995a, b). Five coverages summarizing the use of 96 herbicides were constructed from tabular estimates of herbicide use by county and by crop published in Gianessi and Puffer (1991). Atrazine use in the Mississippi-Missouri and Ohio River basins in 1989 extracted from the USGS coverage is presented in Figure 3.6.

Although the crop acreages used in the assessment were from 1987, the herbicide use estimates generally reflect the 1989 usage amount. Data from such sources as surveys of weed scientists, surveys of farmers, and crop acreage, were utilized to make these estimates. Estimates of the number of pounds of atrazine used per square mile in 1989 are contained the Arc/Info coverage HERBICIDE1 (attribute H1980.USE), obtained from National Spatial Data Infrastructure (NSDI), National Geospatial Data Clearinghouse:
http://nsdi.usgs.gov/nsdi/wais/water/herbicide1.HTML.

Battaglin and Goolsby (1995a) constructed seven coverages summarizing annual nitrogen sales in U. S. counties, for the fertilizer years 1985-1991. The fertilizer year starts July 1 of the previous year and ends June 30, e. g. fertilizer year 1990 starts 07/01/89 and ends 06/30/90). The maps were prepared from estimates reported by the U. S. Environmental Agency (years 1985-89) and from Jerald Fletcher in cooperation with the National Fertilizer and Environmental Research center, Tennessee Valley

Authority (years 1990 and 1991). The sales of fertilizer do not account for the use of manure.

Coverages of nitrogen fertilizer use estimates are available in Arc/Info export format from National Spatial Data Infrastructure (NSDI), National Geospatial Data Clearinghouse: http://nsdi.usgs.gov/nsdi/wais/water/nit89.HTML, .../nit90.HTML, and .../nit91.HTML). Total nitrogen-fertilizer use in tons per square mile from July, 1989 to June, 1990 is stored in attribute NTOT90 of coverage NIT90; nitrogen-fertilizer use from July, 1990 to June, 1991 is in attribute NTOT91 of coverage NIT91.

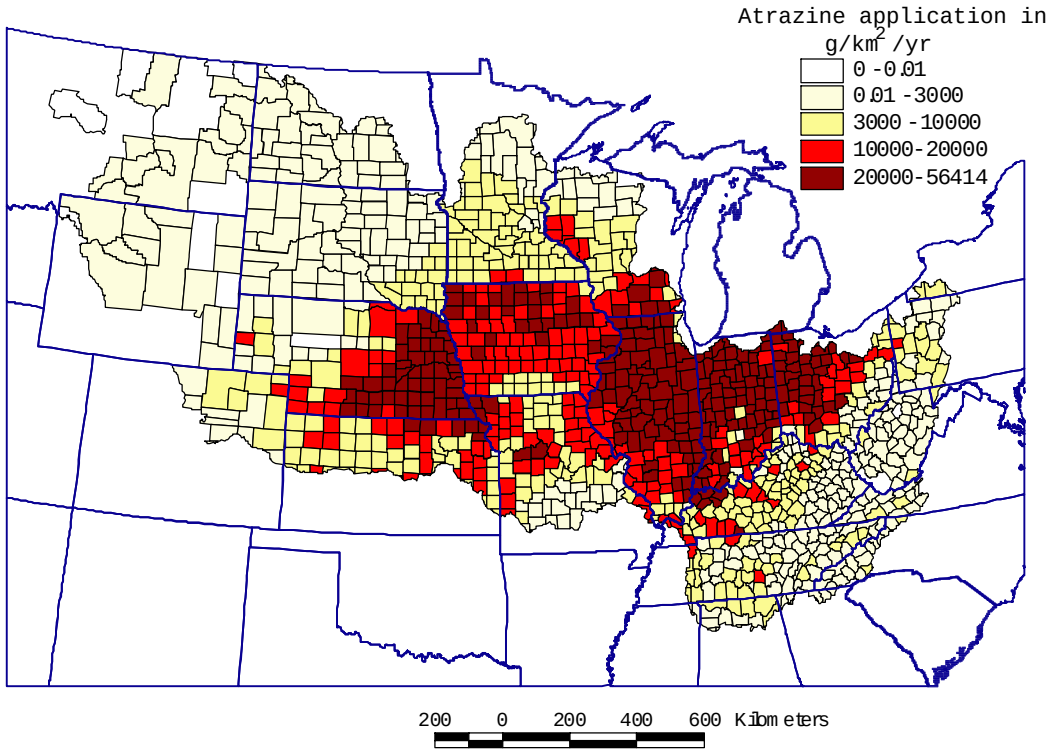

Figure 3.6 Atrazine use in the Mississippi-Missouri and Ohio River basins in 1989 (from Battaglin and Goolsby, 1995b)

3.1.5 Hydrologic and climatic data

The gauging station locations, drainage areas, and the daily flow data for 31,000 USGS gauging stations are available from the CD-ROM set published by Hydrosphere Data Products, INC. (Hydrosphere, 1993a). These data originate from the USGS WATSTORE system (Daily and Peak Values Files For Stream Flows).

Additional data such as daily rainfall, snowfall, maximum temperature and minimum temperature, which were observed at 17,000 NCDC Stations, are also available on the CD-ROM set published by Hydrosphere Data Products, Inc. (Hydrosphere, 1993b).

For the purpose of this research the following data have been exported from Hydrosphere CD-ROMS in ASCII (text) format:

- Monthly flow rate measured in all USGS gauging stations in the Iowa Cedar River basin, for the years up to 1992;
- Monthly precipitation depth measured in all NCDC gauging stations in the Iowa Cedar River basin and within a 50 km buffer outside the basin for the years up to 1992; and
- Description of the gauging sites (to extract such information as drainage area and station location-- latitude and longitude).

3.1.6 Maps of mean annual precipitation and temperature

Arc/Info coverages of the mean annual precipitation and temperature in U.S. have been published by the USGS in 1991 (USGS, 1991). The mean values as well as the standard deviations were calculated for the period from 1951 to 1980 and stored in the attribute table of the map of National Oceanic and Atmospheric Administration (NOAA) Climate Divisions (each state is divided into nine climate zones). The map originator is National Climatic Data Center. The Arc/Info coverage is available via Internet at http://nsdi.usgs.gov/nsdi/wais/water/climate_div.HTML. Figure 3.7 and Figure 3.8 present the mean annual temperature and the mean annual precipitation in the Upper Mississippi-Missouri and Ohio River basins, respectively.

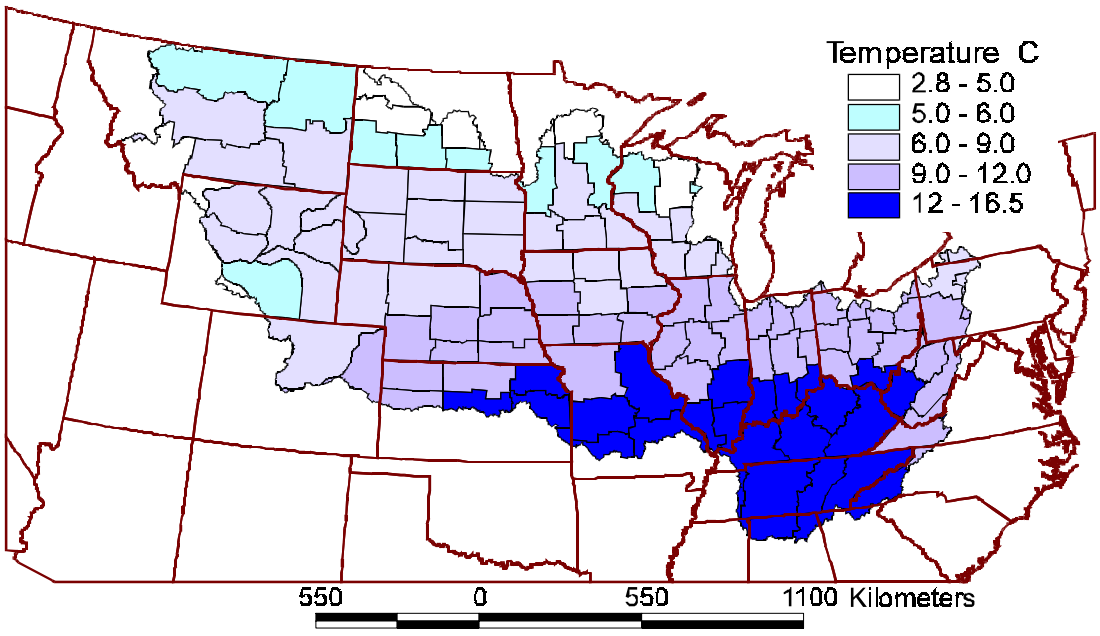

Figure 3.7 Mean annual temperature [°C] in the Mississippi-Missouri and Ohio River basins by NOAA climate division (USGS, 1991).

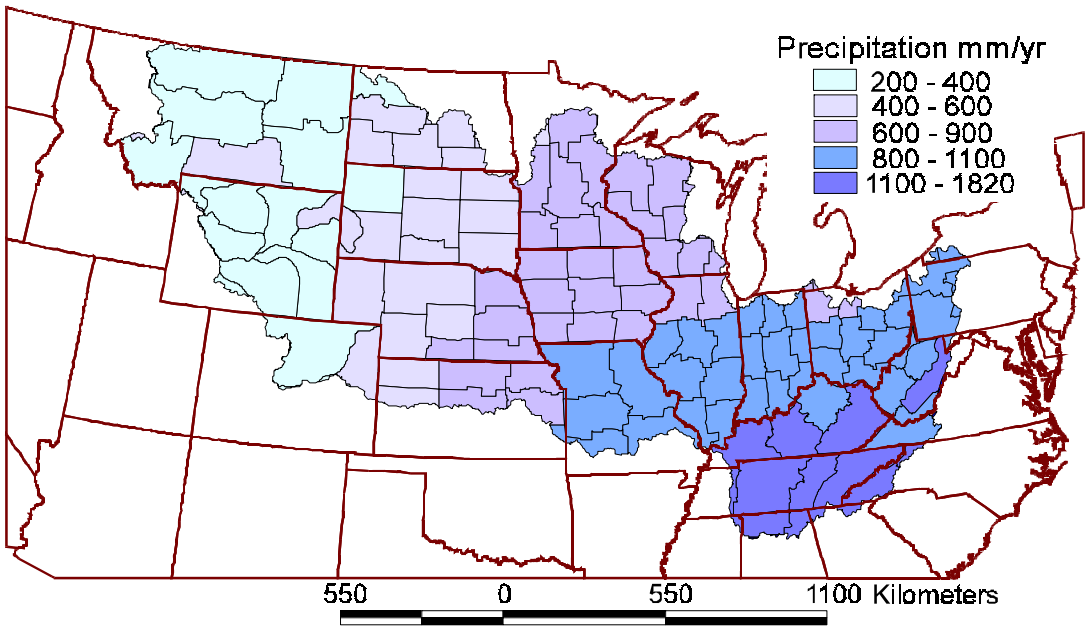

Figure 3.8 Annual precipitation [mm] in the Mississippi-Missouri and Ohio River basins by NOAA climate division (USGS, 1991).

3.2 Computer software description

Section 3.2.1 presents the Geographic Information System software used in this research. Section 3.2.2 introduces statistical program, S-Plus, used to develop a regression model. The research was performed on SUN SPARC station IPX, SUN ULTRA (operation system UNIX) and 486DX2-66 (operation system DOS).

3.2.1 GIS software

Arc/Info and ArcView constitute the GIS software used in this research. (Arc/Info and ArcView are registered trademarks of the Environmental Systems Research Institute, Inc., Redlands, California). Arc/Info is a spatial analysis system which represents spatial data in separate layers and it provides operators for manipulating these data. It contains three basic spatial primitives for vector data: Points, Lines, and Polygons. This software supports also three derived data structures: Grids (a rectangular mesh of points), Triangulated Irregular Networks (a TIN is a set of points connected by triangles), and Networks (a set of connected arcs with assigned flow properties). Each spatial primitive can have an associated record in a database, an Info file. The fields of this record contain user specified descriptive attributes, such as area, length, category, name, etc. The one to one correspondence between the spatial features (point, arc or polygon) and data records (Info) is the basis of the hybrid Arc/Info data model.

Besides the Arc/Info core system, two Arc/Info processors are extensively used in this research: GRID and TABLES. GRID manipulates maps in raster format or grids. TABLES is used to handle the data stored in the following attribute tables:

- point attribute table (PAT) associated with a point coverage;
- arc attribute table (AAT), an Info table of an arc coverage;
- polygon attribute table (PAT), an Info table of a polygon coverage; and,
- value attribute table (VAT), an Info table attached to a grid.

The Arc/Info macro language (AML) enables the automation of complex or repeated tasks.

ArcView is a GIS software completely operated from a graphical user interface. Although it can perform only simple spatial operations on the maps in vector

format, it has very powerful and convenient tools to manage attribute tables. Avenue, an object oriented script language within ArcView, allows the user to build complex GIS applications. ArcView is available for workstations and for personal computers. ArcView version 2.1 was used in this research (ESRI, 1995).

3.2.2 Statistical software

The data analysis and model parameter estimation were performed by the application of two computer programs: Microsoft Excel spreadsheet version 5.0a (Microsoft and Excel are registered trademarks of Microsoft Corporation) and the statistical software, S-Plus version 3.2 (S-PLUS is a registered trademark of MathSoft, Inc.) MS Excel was used mainly as a preprocessor for data editing, simple calculations such as unit conversion, and data verification. All the statistical analysis which is presented in this project was performed using S-Plus.

S-Plus is an interpreted language that evaluates expressions. The results of the expression is an object. All input data must be one of the S-Plus objects. Two types of objects have been utilized in this research:

- vector, a set of elements in a special order, and
- data frame, that represents data in a two dimensional table.

There are two more S-Plus objects that are designed to store tables, array and matrix, but the data frame, unlike the matrix, can have values of different modes in different columns (MathSoft, 1993; Venables and Ripley, 1994). A detailed description of the application of S-Plus in this study is presented in Section 6.