

**CRWR Online Report 96-6**

**MODELING AGRICHEMICAL TRANSPORT  
IN MIDWEST RIVERS USING  
GEOGRAPHIC INFORMATION SYSTEMS**

**by**

**Pawel Jerzy Mizgalewicz, Ph.D.**

**and**

**David R. Maidment, Ph.D.**

**Principal Investigator**

**December 1996**

**CENTER FOR RESEARCH IN WATER RESOURCES**

**Bureau of Engineering Research .The University of Texas at Austin**

**J. J. Pickle Research Campus. Austin, TX 78712-4497**

**This document is available online via World Wide Web at  
<http://civil.ce.utexas.edu/centers/crwr/reports/online.html>**

**MODELING AGRICHEMICAL TRANSPORT  
IN MIDWEST RIVERS USING  
GEOGRAPHIC INFORMATION SYSTEMS**

**Publication No. \_\_\_\_\_**

**Pawel Jerzy Mizgalewicz, Ph.D.**

**The University of Texas at Austin, 1996**

**Supervisor: David R. Maidment**

A method is proposed for regionalizing watershed scale water quality estimates. Elementary watersheds are delineated using digital elevation data and linked to form a river basin scale watershed network. Elementary watersheds are combined into stream gauge zones for which the only streamflows into and out of a zone are those measured at the zone boundary by stream gauges. Time series of monthly streamflow are obtained by an interpolation procedure in which monthly precipitation over each elementary watershed is converted to streamflow by a runoff coefficient, and then adjusted so that the accumulated streamflow over the gauge zone is equal to the

measured outflow. Concentrations of water quality constituents are found from regression equations in which the mean annual concentration is estimated as a function of watershed, chemical application and climatic characteristics, and a ratio of expected monthly to annual concentration is applied. Parameters of these equations were found for two constituents: nitrate plus nitrite as nitrogen, and atrazine, using data sampled by the US Geological Survey at 151 sites in the Missouri, Upper Mississippi and Ohio River basins. Nitrate plus nitrite concentrations show a fairly uniform seasonal pattern and some dependence on spatial factors; atrazine concentrations show a strong seasonal pattern with high values in May and June, and little dependence on spatial factors. Both constituents appear to increase in concentration with discharge to the 0.3 power approximately. An example application of the method is made to the 32,000 km<sup>2</sup> Iowa-Cedar River basin using elementary watersheds of average area approximately 30 km<sup>2</sup>. In this basin, constituent loading estimates determined using discharge-dependent concentrations appear to be too large when compared with independent loading estimates, which suggests that the sampled water quality database may be somewhat biased towards processes occurring during high runoff rather than baseflow periods. Loading estimates found from discharge-independent concentrations are more reasonable.

# TABLE OF CONTENTS

<b>LIST OF TABLES .....</b>	<b>x</b>
<b>LIST OF FIGURES.....</b>	<b>xiii</b>
<b>LIST OF PROCEDURES .....</b>	<b>xx</b>
<b>1. INTRODUCTION .....</b>	<b>1</b>
1.1 MOTIVATION.....	2
1.2 OBJECTIVES .....	3
1.3 SCOPE OF STUDY.....	4
1.4 PROJECT SUMMARY .....	5
1.5 CONTRIBUTIONS OF STUDY .....	7
<b>2. LITERATURE REVIEW.....</b>	<b>9</b>
2.1 LINKING GIS WITH WATER QUALITY MODELS .....	9
2.2 GIS MODELS OF WATER QUALITY .....	17
2.3 GIS AS A TOOL FOR SPATIAL DATA EXTRACTION.....	19
2.4 COMPARISON OF THE PROPOSED METHOD WITH PREVIOUS STUDIES .....	22
2.4.1 <i>Time domain</i> .....	22
2.4.2 <i>Spatial domain</i> .....	22
2.4.3 <i>Model formulation</i> .....	23
<b>3. DATA AND COMPUTER SOFTWARE DESCRIPTION.....</b>	<b>25</b>
3.1 DATA SOURCES .....	25
3.1.1 <i>Herbicide and nutrient data</i> .....	26

3.1.2 Digital Terrain Representation .....	31
3.1.3 Reach File 1 .....	32
3.1.4 Atrazine and nitrogen fertilizer use .....	34
3.1.5 Hydrologic and climatic data.....	36
3.1.6 Maps of mean annual precipitation and temperature .....	37
3.2 COMPUTER SOFTWARE DESCRIPTION .....	38
3.2.1 GIS software .....	39
3.2.2 Statistical software.....	40
<b>4. METHODOLOGY .....</b>	<b>41</b>
4.1 REPRESENTATIVE AGRICULTURAL CHEMICALS .....	43
4.1.1 Nitrate .....	44
4.1.2 Atrazine .....	48
4.2 SELECTION OF ANALYSIS REGION AND MAP COORDINATE SYSTEM.....	50
4.3 MATHEMATICAL DESCRIPTION .....	54
4.3.1 Overview of transport equations .....	54
4.3.2 GIS and cascade modeling.....	58
4.3.3 Regression equation development .....	61
4.3.4 Agrichemical runoff from the field.....	63
4.3.5 Transport in rivers.....	65
4.3.6 Seasonal variations.....	67
4.3.7 Extracting values of explanatory variables for the regression analysis.....	71
4.3.8 Application of the regression models .....	73
4.4 GIS MODEL DESCRIPTION .....	74
4.4.1 Subdivision of study region into modeling units .....	74
4.4.2 Unit watershed flow system.....	78
4.4.3 Ordering system of the modeling units .....	81
4.4.4 Enhancement of the stream delineation process .....	82

4.5 REDISTRIBUTION OF THE FLOW RECORD OVER UNGAUGED RIVERS .....	85
4.5.1 GIS database of monthly flow rate and the precipitation depth .....	85
4.5.2 Average precipitation depth in modeling units .....	88
4.5.3 Mathematical description .....	90
4.5.4 Example of flow redistribution.....	95
4.6 EXPONENTIAL DECAY MODEL .....	98
4.6.1 Exponential decay model overview.....	99
4.6.2 Travel time approximation.....	101
<b>5. PROCEDURES .....</b>	<b>108</b>
5.1 CONCENTRATION AND FLOW MEASUREMENTS.....	109
5.2 PREPARING DATA FOR MODEL PARAMETER ESTIMATION .....	113
5.2.1 Preparing 500 m (15") DEM for analysis .....	114
5.2.2 Estimation of watershed parameters .....	115
5.2.3 Creating grid of sampling sites .....	120
5.2.4 Adjusting location of sampling sites.....	121
5.2.5 Extracting parameters of the sampled site watersheds.....	124
5.2.6 Agrichemical application.....	126
5.2.7 Annual temperature and annual precipitation depth .....	128
5.3 PREPARING DATA FOR STATISTICAL ANALYSIS.....	130
5.4 AGRICHEMICAL CONCENTRATIONS IN THE IOWA-CEDAR RIVER BASIN .....	132
5.4.1 Creating a map of the flow direction.....	133
5.4.2 Map of the modeling unit outlets.....	139
5.4.3 Watershed connectivity .....	143
5.4.4 Refining modeling units .....	144
5.4.5 Database of monthly precipitation depth and monthly flow rate.....	147
5.4.6 Average precipitation depth in modeling units .....	148
5.4.7 Spatial distribution of flow.....	150

5.4.8 <i>Determining input values for the Iowa-Cedar River model</i> .....	154
5.5 ARCVIEW MODEL OF AGRICHEMICAL TRANSPORT .....	154
5.5.1 <i>Model overview</i> .....	155
5.5.2 <i>Project “Model”</i> .....	158
5.5.3 <i>Project “Results”</i> .....	163
5.5.4 <i>Project “Flwprc”</i> .....	168
5.5.5 <i>Project “Tools”</i> .....	169
<b>6. RESULTS .....</b>	<b>170</b>
6.1 SEASONAL VARIATION OF AGRICHEMICALS IN THE MIDWEST STREAMS .....	171
6.1.1 <i>Seasonal variation of the atrazine concentration</i> .....	172
6.1.2 <i>Seasonal variation of the nitrate plus nitrite as nitrogen concentration</i> .....	177
6.2 AVERAGE ANNUAL AGRICHEMICAL CONCENTRATION IN THE MIDWEST STREAMS	180
6.2.1 <i>Average annual atrazine concentration in the Midwest rivers</i> .....	183
6.2.2 <i>Average annual nitrate plus nitrite as nitrogen concentration</i> .....	188
6.3 ERROR OF MODEL PREDICTIONS .....	192
6.4 COMPARISON OF PREDICTED FLOW WITH OBSERVED ONE.....	197
6.5 AGRICHEMICAL CONCENTRATIONS IN THE CEDAR RIVER BASIN .....	202
<b>7. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS .....</b>	<b>213</b>
<b>APPENDICES .....</b>	<b>223</b>
APPENDIX A C-CODES.....	223
<i>A1 Program newnx.c--reconstructing the flow connectivity between modeling units     after some of units have been removed</i> .....	223
<i>A2 Program fdy4.c--calculating the flow rate in ungauged streams from the     available record</i> .....	226
APPENDIX B AVENUE SCRIPTS .....	232

<i>B1 Changing ArcView projects from the PushButton Bar.</i>	235
<i>B2 Scripts from the project “model”</i>	237
<i>B3 Scripts from the project “results”</i>	263
<i>B4 Scripts from the project “flwprc”</i>	264
<i>B5 Scripts from the project “tools”</i>	279
<b>APPENDIX C ARC/INFO MACROS--AMLs</b>	<b>292</b>
<i>C1 WSHGS.AML</i>	293
<i>C2 NEXTWSH.AML</i>	297
<i>C3 RAININFO.AML</i>	300
<i>C4 RAINMAP.AML</i>	304
<i>C5 RAINM2.AML</i>	307
<i>C6 SELDATA.AML</i>	309
<i>C7 FD4Y.AML</i>	313
<i>C8 SLOPE3.AML</i>	315
<i>C9 DEMALB.AML</i>	318
<b>BIBLIOGRAPHY</b>	<b>319</b>
<b>VITA</b>	<b>339</b>



## LIST OF TABLES

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). .....	26
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). .....	29
Table 4.1 An example of PAT--point attribute table of the USGS gauging station coverage (e.g., item M199001 contains average flow rate in year 1990, month 01). Full table contains monthly flow record for 38 stations, for period from 1940 to 1992 in m <sup>3</sup> /s. ....	88
Table 4.2. Steps of the recorded flow rate redistribution in gauged zone 5465500 for June 1990. ....	97
Table 5.1 Summary statistics of the nitrite plus nitrate and atrazine data sets. ....	113
Table 5.2 Selected statistics of watershed parameters determined from DEM. Length of the flow path used to calculate $E_S$ is in 100 km. ....	125
Table 5.3 Summary statistics of annual application of nitrogen fertilizer and atrazine. ....	127
Table 5.4 Description of tables atra7 and nitr7. ....	130
Table 5.5 Polygon Attribute Table of the cruse coverage. ....	160
Table 5.6 Polygon Attribute Table of the crwsd coverage. ....	160
Table 5.7 Fields of the file model2a.dbf (model specification). ....	161
Table 6.1 Selected coefficients of the regression analysis of the atrazine concentration in the Midwest rivers. Seasonal variation is explained only by the sine-cosine harmonics. Coefficients related to the dummy variables are not shown. ....	173

Table 6.2 Selected coefficients of the regression analysis of the atrazine concentration in the Midwest rivers. Seasonal variation is explained by the sine-cosine harmonics and by the flow rate. Coefficients related to the dummy variables are not shown.....	174
Table 6.3 Seasonal factors of atrazine concentrations in the Midwest rivers estimated by the regression analysis with and without flow rate record. ....	175
Table 6.4 Selected coefficients of the regression analysis of the nitrate concentration in the Midwest rivers. Seasonal variation is explained only by the sine-cosine harmonics.....	177
Table 6.5 Selected coefficients of the regression analysis of the nitrate concentration in the Midwest rivers. Seasonal variations are explained by the sine-cosine harmonics as well as the flow record. Coefficients related to the dummy variables are not shown. ....	178
Table 6.6 Seasonal factors of the nitrate plus nitrite as nitrogen concentrations in the Midwest rivers estimated by the regression analysis with and without flow rate.....	179
Table 6.7 Explanatory variables used to in the analysis of the deseasonalized atrazine and nitrate plus nitrite as nitrogen concentrations in the Midwest rivers. ....	182
Table 6.8 Results of the stepwise regression analysis of average annual atrazine concentration in the Midwest rivers (Data = atrazine concentration with removed seasonal component). ....	184
Table 6.9 Results of the regression analysis of average annual atrazine concentrations in the Midwest rivers (Data = atrazine concentration with removed component explained by the seasonal factor and the flow rate).....	186
Table 6.10 Quartiles of the explanatory variables selected by the regression analysis of the deseasonalized agrichemical concentrations.....	186

Table 6.11 Results of the regression analysis of average annual nitrate plus nitrite as nitrogen concentrations in the Midwest rivers (Data = nitrate concentration with removed seasonal component). .....	188
Table 6.12 Results of the regression analysis of average annual nitrate concentrations in the Midwest rivers (Data = nitrate plus nitrite as nitrogen concentrations with removed component explained by the seasonal factor and the flow rate). .....	190
Table 6.13. Mean Standard Errors for the atrazine concentration models ( $\mu\text{g/L}$ ). .....	192
Table 6.14. Mean Standard Errors for the nitrate plus nitrite as nitrogen concentration models ( $\text{mg/L}$ ). .....	194
Table 6.15 Comparison of predicted and observed flow rate ( $\text{m}^3/\text{s}$ ) for three USGS gauging stations.....	197
Table 6.16 Atrazine concentration, load and the flow rate in modeling units that represent the confluence of the Iowa River with the Cedar River estimated for June 1990. ....	207
Table 6.17 Relation of the atrazine application to the atrazine load in two Iowa rivers: the Cedar River at Palisades and the Old Mans Creek at Iowa City ( Chemical application for year 1989, flow data for 1990). .....	208
Table 6.18 Relation of the nitrogen fertilizer application to the nitrate plus nitrite as nitrogen load in two Iowa rivers: the Cedar River at Palisades and the Old Mans Creek at Iowa City ( Chemical application and flow rate represent year 1990). .....	211

## LIST OF FIGURES

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).....	27
Figure 3.2 Atrazine Concentrations in Old Mans Creek, Iowa (data from USGS Open-File Report 94-396, Scribner et al., 1994).....	28
Figure 3.3 Watersheds sampled during 1989 and 1990 for the USGS reconnaissance study (Battaglin, 1995).....	30
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. ....	32
Figure 3.5 The Cedar River above Waterloo, IA; rivers selected from Reach File 1. ....	33
Figure 3.6 Atrazine use in the Mississippi-Missouri and Ohio River basins in 1989 (from Battaglin and Goolsby, 1995b).....	35
Figure 3.7 Mean annual temperature [°C] in the Mississippi-Missouri and Ohio River basins by NOAA climate division (USGS, 1991).....	37
Figure 3.8 Annual precipitation [mm] in the Mississippi-Missouri and Ohio River basins by NOAA climate division (USGS, 1991).....	38
Figure 4.1 Methodology of the large scale modeling of agrichemical concentrations in the Midwest rivers. ....	42
Figure 4.2 The Mississippi-Missouri and Ohio River Basins. ....	51
Figure 4.3 The Iowa River with tributaries and the USGS gauging stations. ....	52
Figure 4.4 Reactions and transfers in a natural water system (after O'Connor, et al., 1983; Thomann and Mueller, 1987; O'Connor, 1988a, b; University of Mississippi, 1990). ....	57

Figure 4.5 Example of the cascade modeling within the GIS.....	60
Figure 4.6 Nitrate plus nitrite as nitrogen concentrations measured in the Sangamon River at Monticello, Illinois, in May 1990 (Scribner, et al., 1994). .....	71
Figure 4.7 Example of a grid of spatially distributed values of explanatory variables. ....	72
Figure 4.8 Example of watershed outlets: (1) beginning of the river, (2) stream junction, and (3) gauging station. ....	75
Figure 4.9 Comparison of the frequency of modeling unit area subdivided using different sets of watershed outlets. ....	76
Figure 4.10 Division of the Iowa-Cedar River into modeling units using different threshold drainage areas: 25 km <sup>2</sup> , 400 km <sup>2</sup> , and 2500 km <sup>2</sup> . ....	77
Figure 4.11 An example of the flow topology of the unit watersheds; (x,y), x is the unit ID, y is the ID of downstream to x unit. ....	78
Figure 4.12 Flow system in the Iowa-Cedar River basin subdivided into modeling units of different sizes (threshold drainage area: 25 km <sup>2</sup> , 400 km <sup>2</sup> , and 2500 km <sup>2</sup> ). ....	80
Figure 4.13 Comparison of the stream ordering systems: (a) Strahler, (b) Shreve, and (c) utilized in the agrichemical transport model.....	82
Figure 4.14 The Iowa-Cedar River basin: subwatersheds and selected USGS gauging stations (numbers represent station ID). ....	86
Figure 4.15 Weather stations applied to analysis of the hydrologic conditions in the Iowa-Cedar River watershed. ....	87
Figure 4.16 Spatial distribution of monthly precipitation depth [mm/d] in the Iowa- Cedar River watershed in June 1990.....	89
Figure 4.17 Cedar River watershed above Charles City, Iowa: An example of gauging station zones and modeling units. ....	90
Figure 4.18 Location of the gauged zone 5465500, its subdivision into modeling units, and the USGS gauging stations. ....	96

Figure 4.19 Example of time series recorded in gauging stations located along the Cedar River. The data are for the water year 1990. The logarithmic scale is used to show all time series in one picture.....	101
Figure 4.20 Flow path and location of gauging stations which have been used to illustrate the potential method for time of travel estimation.....	102
Figure 4.21 Cross-correlation coefficients: the strength of the linear relationship between the flow rate in the Cedar River recorded near Conesville, IA, and the flow rate recorded at indicated locations. All gauging stations are on one flow path (shown in Figure 4.20). .....	103
Figure 4.22 Analysis of travel time and flow velocity: a) cumulative travel time versus cumulative flow distance; b) travel time vs. flow distance; c) velocity vs. square root of stream slope. ....	104
Figure 4.23 Estimation of the “travel time” for major reaches of the Iowa River and the Cedar River, IA.....	106
Figure 4.24 Comparison of the velocity from the RF1 database with the velocity estimated from a correlation analysis of the flow record in the Iowa River and tributaries .....	107
Figure 5.1 Relationship between atrazine ELISA concentration and concentration by GC/MS estimated from data published in the USGS Open File Report 93-457. ....	110
Figure 5.2 Relationship between atrazine ELISA concentration and concentration by GC/MS estimated by a different authors: (1) Thurman et al. (1992); (2) Goolsby et al. (1993) data gathered in 1990; (3) Goolsby et al. (1993) data gathered in 1991; (4) Moody and Goolsby (1993), and (5) relation estimated in this study utilizing data from USGS open file report 93-457... ..	112
Figure 5.3 Example of the sampling site (from Battaglin 1995) that is located neither on the stream from RF1 nor on the stream delineated from the DEM.....	121

Figure 5.4 Sampled watersheds delineated from the 500 m DEM. ....	124
Figure 5.5 1° quadrangles of DEM utilized to subdivide the Iowa-Cedar River basin into modeling units; Map projection: geographic. ....	134
Figure 5.6 Comparison of the stream network delineated from 100 m DEM (a) with the one delineated from 500 m DEM (b). Both DEMs were adjusted for RF1 stream network.....	138
Figure 5.7 Selected types of the modeling unit outlets (a) type 1, beginning of the stream network; (b) type 2, upstream cell of the stream junction; (c) type 3, USGS gauging station.....	139
Figure 5.8 The Iowa-Cedar River basin subdivided into 1032 unit drainage areas of average area 31.6 km <sup>2</sup> .....	146
Figure 5.9 Runoff (flow in rivers) that occurred in June 1990 in the Iowa-Cedar River watershed. ....	153
Figure 5.10 An example of the Graphical User Interface: a window for the selection of the agricultural application rate in selected counties.....	156
Figure 5.11 Atrazine concentrations along the Cedar River estimated for 1990 (left) and flow rate along the Cedar River in June 1990 (right). ....	157
Figure 5.12 “Bar chart” map of the atrazine concentrations estimated for 1990 (preliminary results)-- an example map created within the ArcView project “Results” (g/m <sup>3</sup> = mg/L).....	164
Figure 5.13 Atrazine concentrations in the Cedar River basin introductory estimated for May 1990 -- a map created within the ArcView project “Results” (g/m <sup>3</sup> = mg/L). ....	165
Figure 5.14 Inconsistency in the assigning IDs to the vectorized grid streams. The GRID function gridline assigned wrong IDs to the Cedar River downstream to the junction with the Iowa River.....	166

Figure 5.15 Flow rate in the Cedar River and tributaries that occurred in June 1990. The map has been created by linking the polygon attribute table of interpolated flow measurements by modeling unit with the arc attribute table of a stream coverage.....	167
Figure 5.16 Monthly flow rate in the lower Cedar River in 1989-1991. Example of visual tools of ArcView project “Flwprc” for hydrologic analysis. ....	168
Figure 6.1 Seasonal factors of atrazine concentrations in the Midwest rivers estimated by the regression analysis with and without flow rate record.....	176
Figure 6.2 Seasonal factors of the nitrate plus nitrite as nitrogen concentrations in the Midwest rivers estimated by the regression analysis with and without flow rate included as an explanatory variable.....	180
Figure 6.3 Influence of the explanatory variables on the average atrazine concentration ( $\mu\text{g/L}$ ). Model without the flow rate (Table 6.7).....	187
Figure 6.4 Influence of the explanatory variables on the average atrazine concentration ( $\mu\text{g/L}$ ). Model with the flow rate (Table 6.8).....	187
Figure 6.5 Influence of the explanatory variables on the average nitrate plus nitrite as nitrogen concentration ( $\text{mg/L}$ ). Model without the flow rate (Table 6.11)..	191
Figure 6.6 Influence of the explanatory variables on the average atrazine concentration ( $\text{mg/L}$ ). Model with the flow rate (Table 6.12). ....	191
Figure 6.7 Difference between measured atrazine concentrations and predicted concentrations in the Midwest rivers: a) model without the flow rate; b) model with flow rate included as an explanatory variable.....	193
Figure 6.8 Difference between measured nitrate concentrations and predicted concentrations in the Midwest rivers: a) model without the flow rate; b) model with flow rate included as an explanatory variable.....	195



Figure 6.9 Observed vs. predicted agrichemical concentrations in the Midwest rivers (model with the flow rate): a) atrazine in $\mu\text{g/L}$ ; b) nitrate plus nitrite as nitrogen in $\text{mg/L}$ . .....	196
Figure 6.10 Location of the USGS gauging stations to verify method of spatial redistribution of the observed flow rate. ....	198
Figure 6.11 Observed and predicted flow rate in selected streams of the Cedar River basin. ....	199
Figure 6.12 Spatial distribution of the surface water balance for the modeling units. Estimated from recorded flow rate, precipitation depth and selected unit features, the Iowa-Cedar River basin, June 1990. ....	201
Figure 6.13 Comparison of the predicted and observed concentrations of nitrate plus nitrite as nitrogen in the Cedar River at Palisades, Iowa and the Old Mans Creek at Iowa City, Iowa, in 1990.....	203
Figure 6.14 Comparison of the predicted atrazine concentrations in the Cedar River at Palisades, Iowa and the Old Mans Creek at Iowa City, Iowa, in 1989 with the observed concentrations in 1990.....	204
Figure 6.15 Profiles of the predicted atrazine concentrations in the Cedar River for June 1989 and the annual atrazine application (based on the results of the ArcView model). ....	205
Figure 6.16 Predicted atrazine concentrations ( $\text{mg}/\text{m}^3 = \mu\text{g}/\text{L}$ ) at the Iowa River and the Cedar River junction (June 1990). ....	206
Figure 6.17 Estimated monthly fractions of the annual atrazine load in the Cedar River at Palisades and the Old Mans Creek near Iowa City, Iowa for year 1990..	209
Figure 6.18 Estimated atrazine load in the rivers of the Iowa-Cedar Basin in June 1990. ....	210
Figure 6.19 Estimated monthly fractions of the annual nitrate load in the Cedar River at Palisades and the Old Mans Creek near Iowa City, Iowa for year 1990..	212

## LIST OF PROCEDURES

Listing 5.1 Creating depressionless DEM, calculating drainage area and delineation of the stream network. ....	115
Listing 5.2 Slope along the flow path. ....	116
Listing 5.3 Average slope of the stream network.....	117
Listing 5.4 Exponentially decaying flow length.....	118
Listing 5.5 Average distance from field to a stream. ....	119
Listing 5.6 Average land slope. ....	119
Listing 5.7 GRID dialog supporting adjustment of sampling sites location.....	123
Listing 5.8 Extracting parameters of the sampled site watersheds.....	125
Listing 5.9 Chemical application in sampled watersheds. ....	127
Listing 5.10 Average climatic parameters.....	128
Listing 5.11 Creating a summary Info table of climatic parameters. ....	129
Listing 5.12 Selecting type 1 watershed outlets. ....	140
Listing 5.13 Selecting type 2 watershed outlets. ....	141
Listing 5.14 Creating point coverage of gauging stations from latitude-longitude coordinates. ....	141
Listing 5.15 Fragment of the file latlonfl.cvs containing station IDs and coordinates....	141
Listing 5.16 Projecting the point coverage of USGS gauging stations from Geographic system into Albers coordinates.....	142
Listing 5.17 Identification of the downstream watershed ID.....	143
Listing 5.18 Eliminating drainage units that are smaller than 1 km <sup>2</sup> .....	144
Listing 5.19 Estimation of the average precipitation depth in modeling units. ....	148

Listing 5.20 Application of the &DO command to repeat action for each month of the selected time period. ....	149
Listing 5.21 Example of the AML that selects precipitation stations that have a complete record in June, 1996 and creates the grid of precipitation depth (adopted from RAINMAP.AML).....	150
Listing 5.22 Example of the Arc/Info macro that prepares data and estimates the flow rate in modeling units (adopted from fd4y.aml, Appendix C7) .....	152
Listing 6.1 The S-plus session for estimation of the model of seasonal atrazine variation. Model does not contain the flow rate component. ....	172
Listing 6.2 The S-plus session for estimation of the model of seasonal atrazine variation. Model contains the flow rate component. ....	173
Listing 6.3 The S-plus session for estimation of the model of seasonal nitrate variation. Model does not contain the flow rate component. ....	177
Listing 6.4 The S-plus specification of the seasonal nitrate model. The model contains the flow rate component. ....	178

# 1. INTRODUCTION

This research is focused on the transport of nutrients and herbicides in surface waters of the Midwestern United States and it examines the extent to which Geographic Information System (GIS) technology can be used to automate and improve the computation of spatio-temporal distribution of selected agricultural chemicals. Up to now, there have been elaborate flow models and pollutant transport models in use at the watershed scale. The application of the most widely used deterministic models such as ANSWERS or SWRRB is limited to agricultural basins of about 200 km<sup>2</sup> and 800 km<sup>2</sup> respectively. An extended version of SWRRB, SWAT model, has been applied to study the hydrology and crop conditions for the conterminous U.S. Deterministic models focus mainly on the detailed description of small-scale physical phenomena and they require huge computer and data resources if implemented on the scale considered here. In the light of the existing projects, the aim of this project is to develop a procedure of estimating the distribution of agricultural chemicals over a region which size is limited only by data availability.

This work proposes a statistical-GIS methodology of determining the amount of selected agrichemicals (nitrate plus nitrite as nitrogen and atrazine) in the rivers of the Upper Mississippi-Missouri River and the Ohio River basins, which have a total drainage area about  $2.4 \cdot 10^6$  km<sup>2</sup> (data used for model parameter calculations were collected in sites scattered over this region). Since the agrichemical transport model requires spatially distributed flow, a technique of spatial redistribution of the discharge recorded in USGS gauging stations, as well as the application of an agrichemical transport model is presented here for the Iowa -Cedar River watershed, Iowa and Minnesota, which has an area of about 32,000 km<sup>2</sup>.

## **1.1 Motivation**

It is commonly known that agricultural activity endangers the quality of surface waters. Farmers apply chemical nutrients to increase soil fertility and use pesticides to control unwanted plants and destructive insects. About 60 percent of pesticides (Gianessi and Puffer, 1990) and nitrogen fertilizers (Kolpin, et al., 1991; US Environmental Protection Agency, 1990) used in the USA are applied to cropland in twelve Midcontinental States (Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin). Four major herbicides: alachlor, atrazine, cyanazine, and metolachlor, account for about 73 percent of the pesticides applied (Gianessi and Puffer, 1990).

Numerous studies have been performed on a small watershed scale. But until recently, few studies have been conducted to evaluate the contamination of rivers in the Upper Mississippi-Missouri and Ohio River basins.

In 1990, the United States Department of Agriculture (USDA) initiated a Management System Evaluation Areas (MSEA) program. The major purpose of this program is to determine the influence of agricultural practices on water quality in the Midwest and to identify management systems that protect water quality. Ten study areas were established to gather data for a better understanding of the factors and processes that control the fate and transport of agricultural chemicals (Hatfield, et al., 1993). Also, the United States Geological Survey (USGS) conducts observational studies to determine the distribution, transport and persistence of selected herbicides, insecticides, and inorganic nutrients in the Mississippi River and its tributaries (Battaglin, et al., 1993).

The Geographic Information System (GIS) offers a unique technology to formulate more objective and consistent methods to synthesize collected data and to assess water quality over large areas (Maidment, 1996). This work explores the applicability of GIS technology to regionalize to large basins the results of small watershed scale water quality studies, and to evaluate the temporal and spatial distribution of the loads and concentrations of selected agricultural chemicals in the surface waters of the Upper Mississippi-Missouri and Ohio River basins.

## **1.2 Objectives**

The general objective of this research is regionalization of watershed scale measurements. This general objective can be divided into the following goals:

- 1) To formulate statistical models capable of representing the spatio-temporal variability of agricultural chemicals in surface waters;
- 2) To evaluate the applicability of the GIS technology for deriving stream and watershed characteristics that influence chemical transport processes;
- 3) To develop a methodology for calculating time series of monthly average flow rate in ungauged streams;
- 4) To build a model for predicting concentrations and loads under different hydrologic conditions and for different agricultural chemical application rates than those historically observed.

### **1.3 Scope of Study**

The following restrictions define the scope of this research:

- 1) The analysis is limited to two selected agrichemicals - the nutrient, nitrate plus nitrite as nitrogen, and the herbicide, atrazine. These two chemicals were chosen because they are representative of nutrients and herbicides, respectively, and because they are present in measurable quantities in many Midwest streams and rivers. Although the research is performed for only two agrichemicals, a similar procedure could be used for other herbicides and nutrients.
- 2) Since the model is constructed using GIS technology, and since the available data are limited, the concentrations and loads of agricultural chemicals in surface waters are described by regression equations, which have limited application for making predictions.
- 3) Observed loads and concentrations as well as chemical application rates are derived from data published in USGS reports. Watershed and stream parameters are derived from a Digital Elevation Model that is available on Internet. Flow rate and the precipitation data are extracted from CD-ROMs (Compact Disk Read Only Memory) published by Hydrosphere Inc.
- 4) Although the regression equations are estimated from data gathered at 151 sampling sites located in eleven Midcontinental States (Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, Ohio, South Dakota, Wisconsin), the complete GIS model is developed only for the Iowa-Cedar River basin, Iowa.

- 5) Model formulation and model parameter estimation are constrained by the available computer resources (SUN Sparc Station IPX and Sun Ultra).

## **1.4 Project Summary**

This research can be divided into the following steps:

- 1) Preparation of measurement data for statistical analysis. This step involves data entry from printed USGS reports, filling in missing values and data correction, and conversion of data into common units.
- 2) Development of the general forms of equations that can explain spatial and temporal variation of the chemical concentration and load in the Upper Mississippi - Missouri River, the Ohio River and their tributaries.
- 3) For each watershed associated with a measurement point, determining the drainage-basin morphometry, climatic characteristics, and chemical application rate. A 15-second (500 m resolution) digital elevation model is utilized in the following steps:
  - delineation of the stream network from the digital elevation model (DEM);
  - location of stations in which the water samples were collected on the delineated river network;
  - estimation of the drainage area associated to each measurement point;
  - calculating parameters that characterize this drainage area;
  - evaluating the watershed average annual precipitation as well as the average annual temperature; and
  - calculating watershed-averaged chemical application rate.



- 4) Preliminary analysis of the spatio-temporal distribution of the constituent concentrations in the Mississippi-Missouri and Ohio River basins. This step includes development of the regression equations that relate measured agricultural chemical concentrations at a given point to such explanatory variables as the annual agricultural chemical application and the flow rate.
- 5) Formulation and application of the methodology for estimation of monthly flow rate in ungauged streams. This methodology is utilized for a selected watershed, i.e., the Iowa-Cedar River basin. The 3 arc-second (100 m resolution) DEM is chosen as a base map for determining the flow direction and for subdivision of the Iowa-Cedar watershed into small drainage units (mostly 20 - 50 km<sup>2</sup>). This step is based on the following:
  - constructing a spatio-temporal database of monthly average flow rate and monthly average precipitation rate for period from 1960 to 1992;
  - dividing the Iowa-Cedar River into small drainage units, converting resulting map from raster into vector format, and building “flow” topology among the drainage units;
  - interpolating or extrapolating recorded flow rate over ungauged drainage units using precipitation, estimated flow rate, an average runoff coefficient, and drainage area as input to an estimating equation.
- 6) Estimation of the spatio-temporal distribution of the agricultural chemical concentration and load. In this step the flow rate identified in step 5 is applied in the regression equation developed in step 4.

## **1.5 Contributions of Study**

This research has the following contributions to the knowledge:

- 1) The development of a method for application of the GIS technology to determine factors that influence the process of mobilization and transport of agricultural chemicals over a very large region;
- 2) The estimation of the general temporal (monthly) pattern of the average atrazine and nitrate concentrations in the Upper Mississippi-Missouri and Ohio River basins;
- 3) The application of the GIS technology to store time series of recorded monthly flow rate and precipitation depth (creation of a spatio-temporal database);
- 4) The development of a spatial analysis routine for estimating monthly time series of expected flow rate in ungauged streams;
- 5) The formulation of a GIS spatial model from which concentrations as well as loads of nitrate plus nitrite and atrazine in streams can be calculated for different hydrologic scenarios.

In addition to the contributions listed above, during this research a set of new GIS tools supporting hydrologic modeling has been developed. The following procedures have been constructed:

- 1) Automated watershed division into hydrologic sub-units;
- 2) Improvement of the major flow paths delineated from a digital elevation model (DEM), and therefore enhancement of the delineated stream network and

watershed boundaries by burning in a mapped stream network to the digital elevation model.

- 3) Calculation of a new stream order system to describe the topologic characteristics of a stream network similar to the Shreve or Strahler ordering methods, that makes flow and transport calculations very efficient; and
- 4) Building in ArcView hydrologic modeling tools that do not exist in ArcView for flowaccumulation--accumulating an entity when traveling downstream, determining drainage area upstream of a given location, and identification of transport paths.

## **2. LITERATURE REVIEW**

This chapter emphasizes research that is focused on GIS applications for agricultural non-point source pollution modeling. Models that include nutrients or pesticides are presented. Soil erosion and sediment transport applications are also mentioned here for completeness of the discussion.

The application of the GIS in modeling non-point source pollution can be grouped into three categories:

- 1) Existing pollution models that are linked with GIS software;
- 2) Pollutant transport is modeled entirely within the GIS;
- 3) GIS is utilized to extract spatial data required for analysis of non-point source pollution;

### ***2.1 Linking GIS with Water Quality Models***

Water quality models linked with GIS programs are the dominant approach in modeling non-point source pollution. The GIS provides the data required for the model, then the model is executed. After calculations, the GIS is used for visual analysis of results. The most popular pollution model that is linked with GIS software is AGNPS (Agricultural Non-Point-Source) developed by Agricultural Research Service (ARS). The model source code is available through the WEB site: <http://www.infolink.morris.mn.us/ars/download.html>.

## ***AGNPS***

AGNPS is an event-based distributed parameter model, that is, it computes flow and pollutant loadings for a single rainfall event. It calculates runoff from agricultural watershed and transport processes of sediment, nitrogen, phosphorous, and COD (Chemical Oxygen Demand). The watershed is represented by square cells of 0.4 - 16 ha (1 - 40 acres). Each cell is characterized by twenty-two parameters that include: SCS curve number, terrain description, channel parameters, soil-loss equation data, fertilization level, soil texture, channel and point source indicators, and an oxygen demand factor. Sediment runoff is estimated from the modified version of USLE (Universal Soil Loss Equation) and its routing is performed for five particle size classes. Calculations of nutrient transport are divided into soluble and sediment-absorbed phases. Due to the large amount of input data required, the application of AGNPS is limited to watersheds not larger than 200 km<sup>2</sup> (Young, et al., 1989; DeVries and Hromadka, 1993; Engel, et al., 1993). However, it has been applied to larger basins, by representing the study area by a grid of cells larger than 16 ha. For example, Morse, et al., (1994) applied AGNPS with 100 ha cells to estimate concentrations in a 1645 km<sup>2</sup> watershed.

At least three interfaces between AGNPS and GRASS (Geographical Resources Analysis Support System) have been constructed: (1) at Michigan State University (He, et al., 1993), (2) by Srinivasan and Engel (Engel, et al., 1993a; Engel, et al., 1993b; Engel 1996; Mitchell, et al., 1993a; Mitchell, et al., 1993b), and (3) by the Soil Conservation Service as a watershed planning tool in the Hydrologic Unit Water Quality Project (HUWQ) (Cronshey, et al., 1993; Geter, et al., 1995; Drungil, et al., 1995).

GRASS is the major public domain GIS that supports a raster data structure with data conversion from vector data. It performs the basic GIS functions of data input, storage, manipulation, analysis, and display (Drungil, et al., 1995). Access to the source code of GRASS provides the flexibility to modify existing GRASS procedures or to add new ones. GRASS has a considerable ability to support hydrologic analysis.

AGNPS has also been linked with Arc/Info GIS. Jankowski and Haddock (1993) coupled AGNPS with PC-Arc/Info, a vector based GIS. The interface was constructed using Arc/Info macro language (SML), Pascal language, and batch programming. Vieux and Needham (1993) studied the AGNPS model sensitivity to grid-cell size. They used Arc/Info to generate AGNPS input files and to display model output. They demonstrated that the variation of channel erosion, sediment yield, and delivery ratio due to the cell size selection may introduce unacceptable errors or erroneous conclusions when analyzing nonpoint pollution using AGNPS.

Morse, et al., (1994) integrated AGNPS, Arc/Info and Oracle, a database system to estimate the nitrogen, phosphorus and COD concentrations for different management scenarios in the Bedford-Ouse catchment, UK. They represented the 1645 km<sup>2</sup> watershed by square cells of 100 ha (1 km \* 1 km). Another AGNPS-Arc/Info integrated system was constructed by Tim and Jolly (1994) to evaluate effectiveness of several alternative management strategies in reducing sediment pollution in a 417 ha watershed located in southern Iowa.

AGNPS has also been linked to other GIS programs, such as: ERDAS (Earth Resources Data Analysis System), a grid cell-based system (Evans and Miller, 1988), Geo/SQL, a vector-based GIS (Yoon, et al., 1993), and IDRISI, a raster based GIS (Klaghofer, et al., 1993). This last interface has been used to evaluate erosion and sediment yields in a lower alpine drainage basin of area of 65 ha (located in Austria). The interface contained EPIC (Erosion/Productivity Impact Calculator, Williams, et al., 1990) a field scale comprehensive model developed to predict the long-term

relationship between erosion and productivity. EPIC's components include weather simulation, hydrology, erosion-sedimentation, nutrient cycling, plant growth, tillage, soil temperature, economics, and plant environment control.

### ***SWRRB***

Cronshey, et al., (1993) describe an interface that combines GRASS and a watershed scale water quality model SWRRB (Simulator for Water Resources in Rural Basins). SWRRB (Arnold, et al., 1990) uses a daily time step for calculations of sediment yield, flow routing, as well as pesticide and nutrient fate and transport studies. Basins are subdivided to account for differences in soils, land use, crops, topography, and weather. The soil profile can be divided vertically into ten layers. Basins of several hundred square miles can be studied, but the number of sub-basins is limited to 10. The hydrology component is based on a daily water balance equation that includes rainfall, runoff, evapotranspiration, percolation, and return flow. Rainfall intensity hyetographs are calculated by using a modified rational method. The Soil Conservation Service curve number technique is used to estimate runoff volume. The evapotranspiration component requires such data as daily solar radiation, mean air temperature, crop cover and snow cover. Daily precipitation as well as air temperatures and solar radiation can be supplied as model input or they can be simulated by a SWRRB weather generator. Sediment yield is computed for each sub-basin by using the Modified Universal Soil Loss Equation. SWRRB is physically based and is intended to be used for situations in which calibration data are not available (DeVries and Hromadka, 1993; Donigian and Huber, 1991).

## ***SWAT-QUAL2E***

SWAT (Soil and Water Assessment Tool) is an extended version of the SWWRB model ( Arnold, et al., 1993). It has been linked with GRASS (Srinivasan and Arnold, 1994; Engel, et al., 1993) and with Arc/Info (Bian, et al., 1996).

The major components of SWAT are similar to SWWRB modules, which include weather, hydrology, erosion, soil temperature, crop growth, nutrients, pesticides, subsurface flow, and agricultural management. The model operates on a daily time step and is capable of simulating 100 or more years. The restriction of only being able to simulate 10 subbasins, in the case of SWRRB, has been removed in SWAT. The watershed can be divided into a practically unlimited number of cells and/or subwatersheds. In addition, each subbasin can be discretized into virtual areas that have a unique soil and land use combination. The hydrologic response is generated in each virtual area. The output of the subbasin is calculated as a weighted average of the virtual area hydrologic responses. The new features of SWAT include: routing of the flow through the basin streams and reservoirs, simulating lateral flow, groundwater flow, stream routing transmission losses, and modeling sediment and chemical transport through ponds, reservoirs, and streams (Mamillapalli, 1996).

The SWAT-GRASS model has been applied for small scale modeling as well as for continental scale hydrologic modeling. For example, Jacobson, et al., (1995) evaluated the water quality impacts of the diverse crops and management practices in a 4.6 km<sup>2</sup> subwatershed of the Herrings Marsh Run Watershed in the North Carolina Coastal Plains. On the other scale extreme, Srinivasan, et al., (1995) applied the SWAT-GRASS interface and such data as a map of soils (STATSGO), map of land use (USGS LULC) and a DEM to estimate the following features for the entire U.S.:

- average annual rainfall;



- average annual total water yield;
- average annual actual evapotranspiration (plant ET was calculated as a function of leaf area, root depth, and irrigation);
- average annual Penman-Montieth potential evapotranspiration; and
- annual grain yield and biomass production.

The U.S. was divided into 78,863 STATSGO polygons for this analysis.

In 1996 the QUAL2E (Enhanced Stream Water Quality Model) water quality component was incorporated into SWAT to simulate instream dynamics. The first-order decay relationships for algae, dissolved oxygen, carbonaceous biochemical oxygen demand, organic nitrogen, ammonium nitrogen, nitrate nitrogen, nitrite nitrogen, organic phosphorus, and soluble phosphorus used in QUAL2E were adopted in SWAT with various adjustments (Ramanarayanan, et al., 1996).

QUAL2E model uses a finite-difference solution of the advective-dispersive mass transport, reaction, and sink/source equation. The stream network is divided into headwaters, reaches, and junctions. The changes in flow conditions are represented as a series of steady-flow water profiles. Such parameters as velocity, cross-sectional area, and water depth that are required for the mass transport calculations are computed from the flow rate. For each river reach, QUAL2E requires specification of as many as 26 physical, chemical, and biological parameters. (DeVries and Hromadka, 1993; Camara and Randal, 1984; Schoellhamer, 1988). Compiling such data at a regional scale would take a very great investment of time and resources.

## ***ANSWERS***

Rewerts and Engel (1991; Engel, 1993; Engel, 1996) integrated the GRASS GIS with ANSWERS (Aerial Nonpoint Source Watershed Environment Response Simulation). ANSWERS (Beasley, et al., 1982) calculates runoff, erosion, sedimentation and phosphorus movement from watersheds. The watershed is divided into a grid of square cells. For each cell the following parameters are defined: slope, aspect, soil porosity, moisture content, field capacity, infiltration capacity, USLE erodibility factor, crop and management factors. A channel is described by width and roughness. Runoff, erosion, sedimentation, and water quality related to sediment associated chemicals (for example, dissolved and sediment-bound ammonium, sediment bound total Kjeldahl nitrogen, and dissolved and sediment-bound phosphorus) are computed for each cell and routed (Beasley and Huggins, 1991; Donigian and Huber, 1991).

Typical cell sizes range from 0.4 to 4 ha with smaller cells providing more accurate simulations. During rainfall events the time step is 60 seconds. If there is no precipitation, the model uses a daily time step (Wolfe, et al., 1995).

De Roo (1993) applied the ANSWERS-GENAMAP GIS-PC-RASTER interface to calculate surface runoff and soil erosion in the Yendacott catchment, UK (147 ha), the Etzenrade catchment (225 ha) and Catsop catchment (46 ha) in The Netherlands. Engel, et al., (1993b) compared the results of GRASS-ANSWERS model, with two other NPS models that were integrated with GIS: AGNPS and SWAT. The simulated results matched observed values reasonably well. Wolfe, et al., (1995) created a user interface that links Arc/Info GIS with ANSWERS. The system has been designed for evaluating the overall effectiveness of selected best management practices at the farm scale.

## ***HSPF***

Al-Abed and Whiteley (1995) linked PC-Arc/Info GIS with the Hydrologic Simulation Program-Fortran (HSPF) to simulate the effects of changes in land use and in resource management strategies on the irrigation water quality in the Grand River, Ontario, with a drainage area of 6,965 km<sup>2</sup>.

HSPF simulates both watershed hydrology and water quality (Johanson, et al., 1980). The rainfall is distributed into interception loss, surface runoff, interflow, and flow into the lower soil zone or groundwater storage. Soil is divided into three moisture zones: an upper soil zone, that influences the rapid runoff, a lower soil zone, and a groundwater storage zone. Some of the water from the groundwater storage becomes stream base flow (DeVries and Hromadka, 1993).

The water quality component simulates silt, clay, and sand sediment transport, including resuspension and settling processes. It can also calculate nutrient and pesticide concentrations. The nutrient processes include DO (dissolved oxygen), BOD (biological oxygen demand), nitrogen, phosphorus, pH, phytoplankton, zooplankton, and benthic algae. HSPF simulates such transfer and reaction processes as hydrolysis, oxidation, biodegradation, volatilization, sorption, and chemical exchange between benthic deposits and the water column. The program user must supply parameters for each of the modeled processes (Donigian and Huber, 1991; DeVries and Hromadka, 1993).

The watershed is divided into segments--parcels of land that are exposed to weather conditions described by one set of meteorological time series. Hydraulic routing requires division of the major streams into modeling segments (Al-Abed and Whiteley, 1995).

## **2.2 GIS models of Water Quality**

Some GIS programs are equipped with a macro language that allows the user to write models within the application. For example, Arc/Info has very powerful macro language, AML (Arc/Info Macro Language). In addition, external procedures written in such programming languages as C/C++ or FORTRAN can be executed by macro, thus the modeling process can be very efficient. This section discuss models of water pollution built using GIS tools.

White and Hofschien (1993) developed a spatial model for assessing nutrient loads in New Jersey rivers using Arc/Info. They used 3 arc-sec digital elevation models (DEM) to partition the study area (15,385 km<sup>2</sup>) into 2,893 drainage basins (polygons) with a network of 10,916 stream segments (arcs). The time of travel was assumed as the basis for calculating predictors of water quality. A simple formula  $v = 0.38 * Q^{0.24}$ , which was estimated for New Jersey, was used to estimate the flow velocity in each reach. A first-order decay reaction was assumed to calculate the non-conservative downstream transport. White and Hofschien attempted to improve the model by representing the decay constant as a function of stream slope, and the nonpoint source yields as a function of subbasin gradient, but the model performance showed no improvement with these refinements. White and Hofschien found that the time of travel, which was calculated from the exponential velocity formula, underestimated by a factor of 0.57 the time of travel of dye-tracer, that is, the dye took approximately twice as long to traverse the stream as the formula suggested. This travel time underestimation was accommodated by assignment of higher values of pollutant decay than those reported in the literature.

Smith, et al., (1993) constructed a GIS model of total phosphorus concentrations in New Jersey streams. The core of this model is a regression equation that relates transformed (natural logarithm) total phosphorus concentration measured at a given point to transformed concentrations resulting from exponentially decayed phosphorous loads in the upstream watershed. In this study, the classical approach of modeling first-order reaction was modified. Instead of using the time of travel and time decay coefficient, the travel distance and a distance decay coefficient for phosphorus were applied in the model respectively. The data from 104 long term sampling stations, collected in the period from 1982 to 1987 were utilized to estimate regression coefficients. The area of the studied region was 15,401 km<sup>2</sup>. The sources of phosphorous were represented by such variables as area of agricultural land, total human population, and total municipal effluent flow.

Zollweg, et al. (1995) constructed a GRASS model of the phosphorus transport for the 25.7 ha Brown Watershed, an upland agricultural watershed in Pennsylvania. The GRASS script language was used to describe the physical processes that originally were modeled by the Soil Moisture-based Runoff Model (SMoRMod). SMoRMod is a distributed spatially variable model. Such parameters as climatic variables, topography, land use, and soils distribution constitute the input. The watershed is divided into rectangular cells. For each cell, the infiltration, soil moisture, groundwater flow, and surface runoff are estimated. The surface runoff is translated through the channel system to the watershed outlet where the storm runoff hydrograph is calculated. The phosphorous module determines the P content of the storm runoff generated over the landscape and transport of this P to the watershed outlet (Zollweg, et al., 1995).

Hession and Shanholtz (1988) incorporated the Universal Soil Loss Equation (USLE) with delivery ratio into the Virginia Geographic Information System

(VirGIS). The model was used to estimate the potential sediment loading in Virginia's Chesapeake Bay drainage area.

### ***2.3 GIS as a tool for spatial data extraction***

The most basic application of the GIS is spatial data manipulation, data extraction for further analysis, and presentation of results in map form. This section discusses work in which the GIS tools have been utilized to support statistical analysis of surface water pollution.

Cressie and Majure (1994) used Arc/Info to determine explanatory variables for a statistical model of the variation in pollutant concentration from dairies in streams of the Upper North Bosque watershed located principally in Erath County, Texas. The Arc/Info GRID and Digital Elevation Models (DEM) were used to determine drainage basins and the lengths along flow paths. Cressie and Majure assumed a spatially constant flow velocity (0.5 m/s), and using simple map algebra, they determined a 3-day flow-time area of influence for each stream measurement site. Seventeen explanatory variables including a number of dairies per acre, a number of animals per acre, lagoons per acre, waste application method, soil hydrologic code, average slope, distance to basin outlet, and precipitation were considered. All variables, except one (seasonal variation), were determined using the GIS. The authors concluded that the GIS was an important tool in observational studies due to its ability to construct explanatory variables at the appropriate scale.

Mueller, et al., (1993) applied logistic regression to relate discrete categories of nitrate concentrations to such explanatory variables as land use in the drainage basins upstream from the sampling sites, percentile of daily streamflow at the time of

sampling, acreage of the basin in corn, acreage in soybeans, density of cattle, and population density.

Logistic regression is used when the independent variable is discrete or categorical rather than continuous. It has the following form:  $\log[p/(1-p)] = a + b_k X$ , where  $p$  is the probability of data value being in one of the possible categories,  $a$  is the intercept,  $b_k$  are  $k$  regression coefficients, and  $X$  is the vector of explanatory variables. The percentiles ( $P_j$ ) are computed using equation  $P_j = A_{(n+1)*j}$ , where:  $A$  is a data set ordered from smallest to largest ( $A_i, i = 1, \dots, n$ ),  $n$  is the sample size of  $A$ , and  $j$  is the fraction of data less than or equal to the percentile value, e.g. for 25th percentile  $j = 0.25$  (Helsel and Hirsh, 1995).

Mueller, et al., (1993) extracted data from GIS databases stored in 1:2,000,000- scale maps of the conterminous United States. The GIS software was used only to areally weight the extracted data and sum it by basin; their model did not include stream transport. Better classification of nitrate concentration was achieved by a model that included the flow percentile, the areal extent of corn and soybean production, the density of cattle, and the density of population, as compared to the model that contained percentile of flow, nitrogen fertilizer application, and population density. In addition, Mueller, et al., found that as the percentile of flow increased, the probability of nitrate concentration being in a higher category also increased. The logistic regression analysis results led these researchers to a conclusion that the level of nitrate contamination in midwestern streams is most strongly related to streamflow and to several characteristics of the upstream basin, including the areal extent of corn and soybean production, the density of cattle, and the population density.

From the observations made on the Mississippi River and four tributaries during a one - year period (from April 1, 1991 to March 31, 1992), Battaglin et al. (1993) estimated a single relationship between the annual use of nitrogen and nitrate transport:  $N_{transport} = -0.2 + 0.1547 N_{use}$  and a linear relationship between annual

atrazine use and the atrazine transport:  $A_{transport} = -12 * + 0.0156 * A_{use}$ , (in metric tons). They used a GIS to estimate the nitrogen and atrazine use within gauged watershed from the county level sales of nitrogen fertilizer and atrazine herbicide. In addition, Battaglin et al. estimated that 321 Mg (ton) of atrazine and 33.7 Mg (ton) of alachlor were discharged from the Mississippi River basin to the Gulf of Mexico in streamflow (from April 1, 1991 to March 31, 1992), while the amounts of these herbicides applied in the basin were approximately equal. This suggests that atrazine is much more persistent agrichemical than is alachlor.

Moody and Goolsby (1993) report the results of a large scale USGS study of herbicide transport in the Lower Mississippi River. Although they did not use a GIS, this work is mentioned here since it is one of the few large scale sampling studies available on the Mississippi River. During May 26-29, 1990, water samples for triazine herbicide analysis were collected every 16 km from Baton Rouge in Louisiana, upriver to the Mississippi-Ohio River confluence (distance of 1900 km). The measurements showed the background level of  $\sim 2.7 \mu\text{g/L}$  of triazine herbicides and an upriver concentration gradient of  $0.2 \mu\text{g/L}$  per 100 km (concentration decreased going downstream). The authors suggest that the longitudinal spatial variability in concentration is a result of cross-channel gradients and the addition of 'slugs' of water from various upriver tributaries. A routing scheme was used to predict the location of water masses. This routing method was tested by using the measurements of the specific conductance. The average flow velocity was  $v = 6 \text{ km/h}$  that gives about 13-day residence time of water in the Mississippi River over the distance sampled in this study. It is interesting that the measurements show about 50% decrease in the load whereas the reported atrazine half-life in water is about 140 days (Thurman, et al., 1992).



## **2.4 Comparison of the proposed method with previous studies**

### **2.4.1 Time domain**

The model presented in this dissertation is designed to represent average monthly values of flow rate, agrichemical concentration, and chemical load in streams of the Midwest. Introduction of a seasonal component to the model fills a gap in existing GIS models of pollutant transport. Most of the hybrid models that have been introduced in Section 2.1, are capable of performing continuous-time simulations (SWAT-GIS), event - related calculations (AGNPS), or daily computations (SWRRB-, ANSWERS-GIS). The models of surface water pollution that are constructed within GIS (discussed in Section 2.2) estimate annual average chemical concentrations for a study period. Even the USGS studies do not evaluate changes in chemical concentration in Midwest surface waters on a monthly basis. Seasonal variations are represented by usually three terms of the year: pre-planting, post-planting, and Fall low flow (Goolsby, et al., 1993a; Goolsby and Battaglin, 1993; Scribner, et al., 1993). The USGS model of agricultural chemical transport in the Midwest rivers relates the annual chemical load with annual agrichemical use (Battaglin, et al., 1993). The statistical model presented here explains seasonal variations of concentration on a monthly basis.

### **2.4.2 Spatial domain**

The model is designed to predict the loads and concentrations in such large basins as the Upper Mississippi River basin ( drainage area 490,000 km<sup>2</sup>), the Ohio

River basin (526,000 km<sup>2</sup>), or the Upper Missouri-Mississippi-Ohio River basin (above Ohio-Mississippi River junction, about 2,400,000 km<sup>2</sup> of drainage area). A more detailed version of the model has been applied for evaluation of nitrate and atrazine concentrations in the Iowa-Cedar River basin, Iowa of 32,000 km<sup>2</sup> - an area larger than the reported limits of application of hybrid models discussed in Section 2.1 and models constructed within GIS, presented in Section 2.2.

Krysanova, et al., (1996), specifies limitations of selected pollution models: for example, AGNPS and ANSWERS are limited to watersheds of about 200 km<sup>2</sup>, SWRRB was developed for agricultural basins as large as 600-800 km<sup>2</sup>, and SWAT is intended to be applied in watersheds up to 25,000 km<sup>2</sup>. Besides the present study the only models known to the author that can be applied for a such large area as the Midwest region are the annual agrichemical load functions presented in Section 2.3, which were estimated by Battaglin (1993).

### **2.4.3 Model formulation**

The proposed model can be classified in the second category discussed previously in Section 2.2, namely models constructed within GIS. Because concentrations and loads estimated by the model are spatially and temporally distributed, the model can be characterized as a distributed system. On the other hand, the agrichemical concentration or load is calculated at a given location by applying average values that characterize the total upstream drainage area in a regression equation. Thus the model can also be considered as a lumped system.

The major differences between existing GIS models and the one presented here result from the spatial extent for which the model has been developed--the Upper Missouri-Mississippi River and the Ohio River basins. It utilizes data available for the

whole US, i.e., digital elevation data and the agrichemical application rates, thus the model contains a limited number of parameters, not including either time-decay nor length-decay coefficients.

### **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:

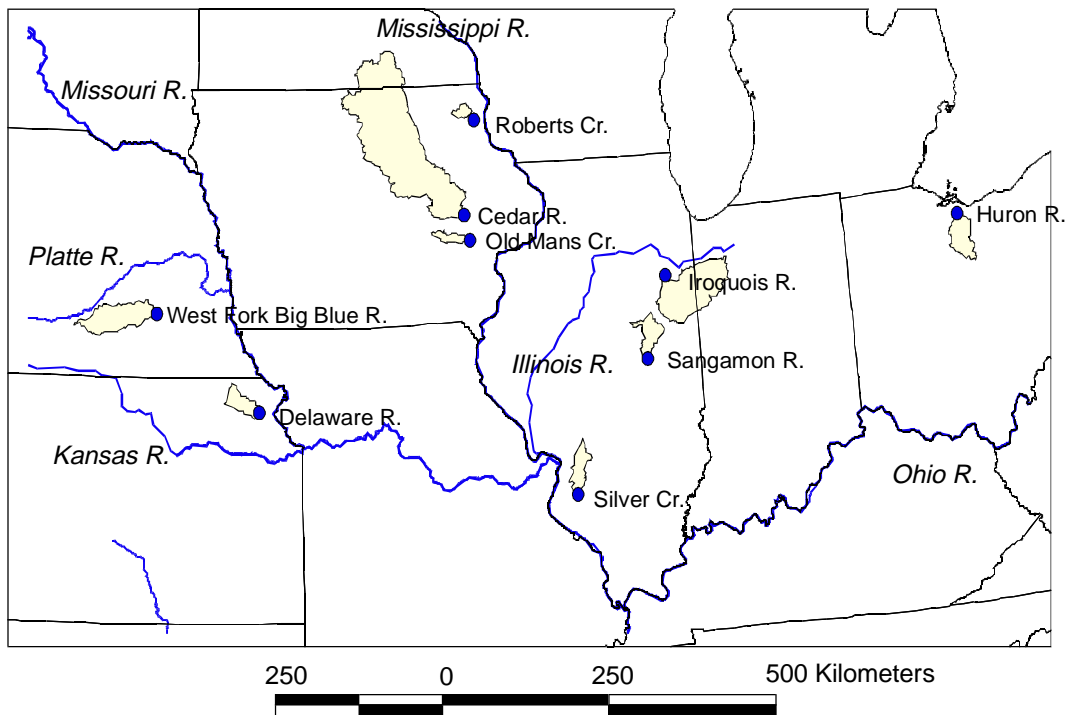
- 1) Open-File Report 94-396, “Concentrations of Selected Herbicides, Two Triazine Metabolites, and Nutrients in Storm Runoff from Nine Stream Basins in the Midwestern United States, 1990-92” (Scribner, et al., 1994);
- 2) Open-File Report 93-457, “Reconnaissance Data for Selected Herbicides, Two Atrazine 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		Number of samples		No of sampled days	
		Area km <sup>2</sup>	Sampled Period	Atrazine (GCMS) as Nitrogen	NO <sub>2</sub> +NO <sub>3</sub> (GCMS) as Nitrogen	Atrazine (GCMS) as Nitrogen	NO <sub>2</sub> +NO <sub>3</sub> (GCMS) as Nitrogen
West Fork Big Blue R., NE	6880800	3,123	4/03/90-7/27/90	37	176	28	57
Sangamon R., IL	5572000	1,425	3/26/91-3/09/92	106	96	64	63
			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 <sup>(1)</sup>	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

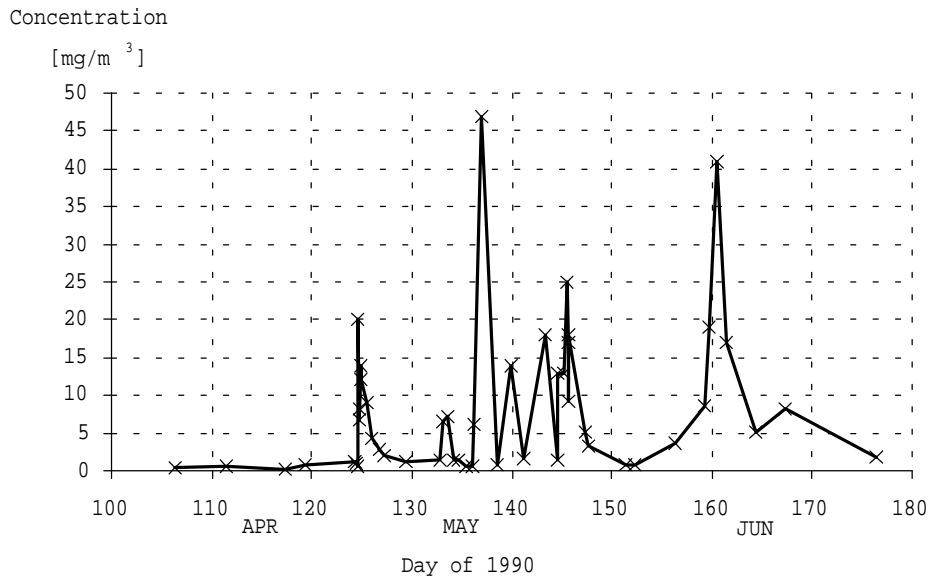
(1) for analysis a drainage area estimated from DEM has been used (17,409 km<sup>2</sup>)

**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.

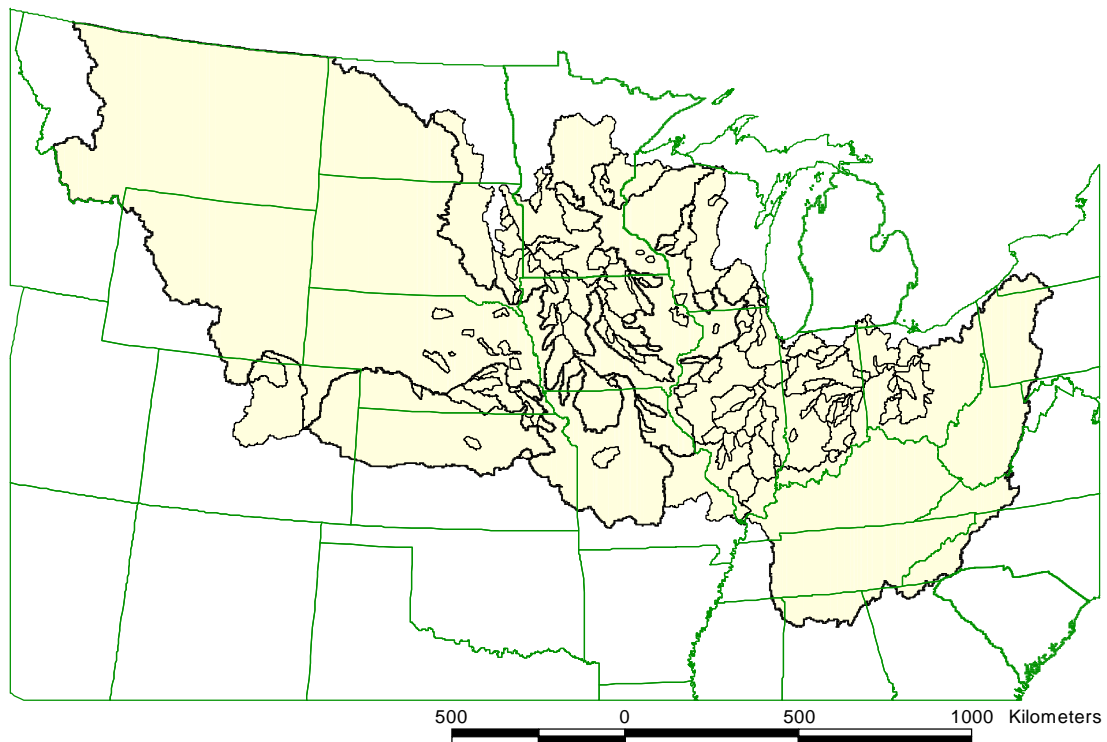
**Table 3.2 Temporal distribution of atrazine and nitrate plus nitrite as nitrogen samples from the Midwestern reconnaissance study; USGS Open-File Report 93-457, (Scribner, et al., 1993).**

Month	Number of samples			
	Atrazine (GC/MS)		Nitrate + Nitrite 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).

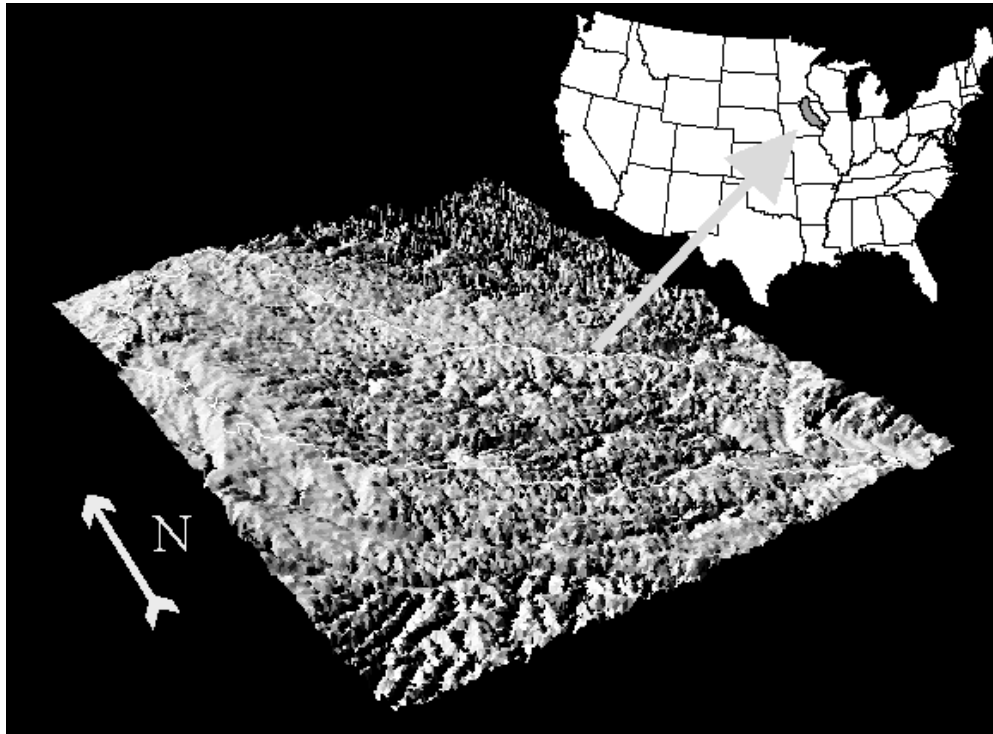
### 3.1.2 Digital Terrain Representation

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](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 m 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

$2.4 \times 10^6 \text{ km}^2$  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.

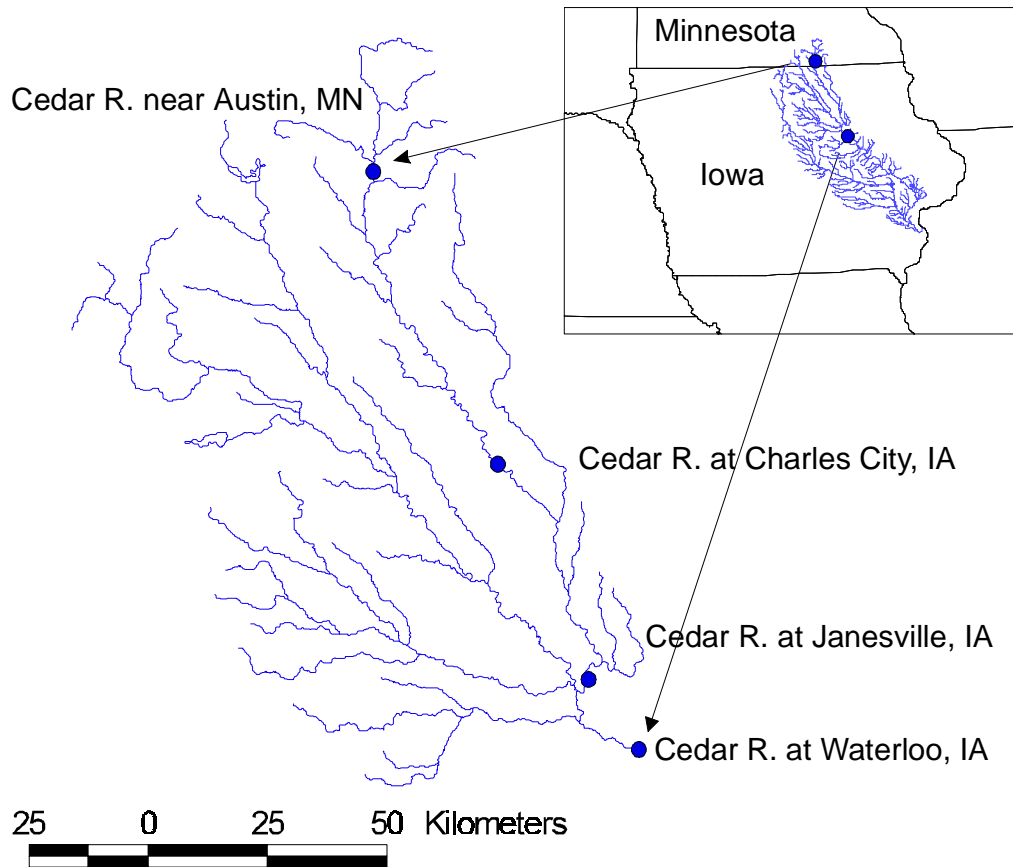


**Figure 3.4** A Digital Elevation Model of the Iowa-Cedar River basin, vertical scale enlarged 500 times. The white line represents the basin boundaries.

### 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-digit 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.

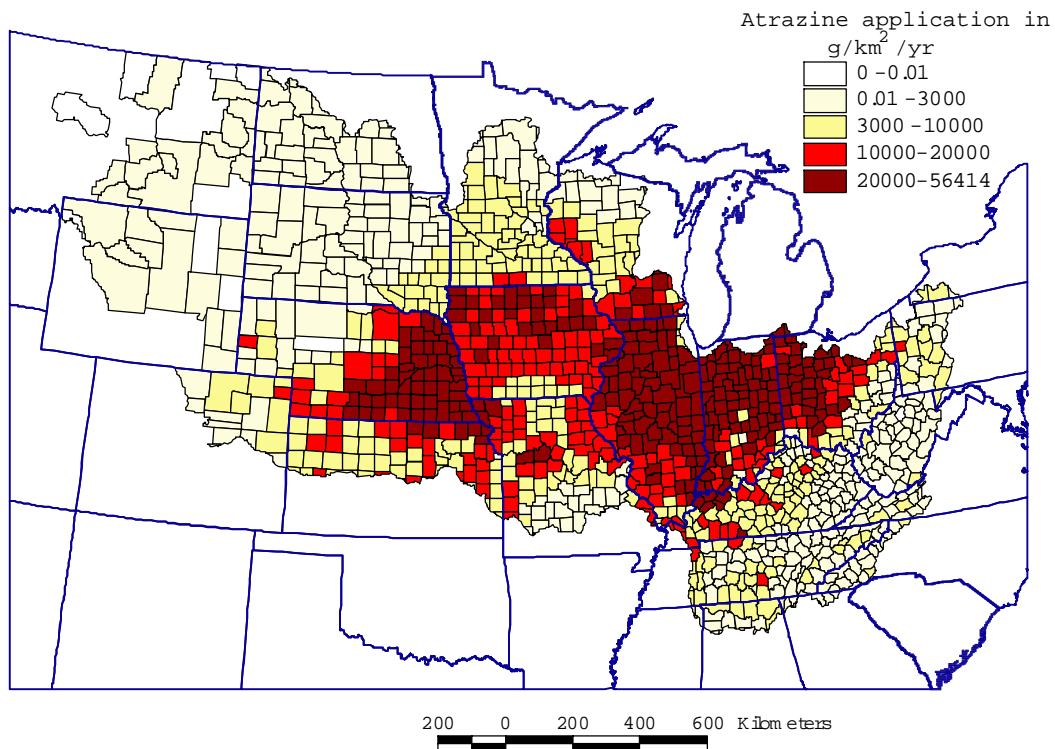
Although the crop acreages used in the assessment were from 1987, the herbicide use estimates generally reflect the 1989 usage amount. Data from such sources as surveys of weed scientists, surveys of farmers, and crop acreage, were utilized to make these estimates. Estimates of the number of pounds of 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).

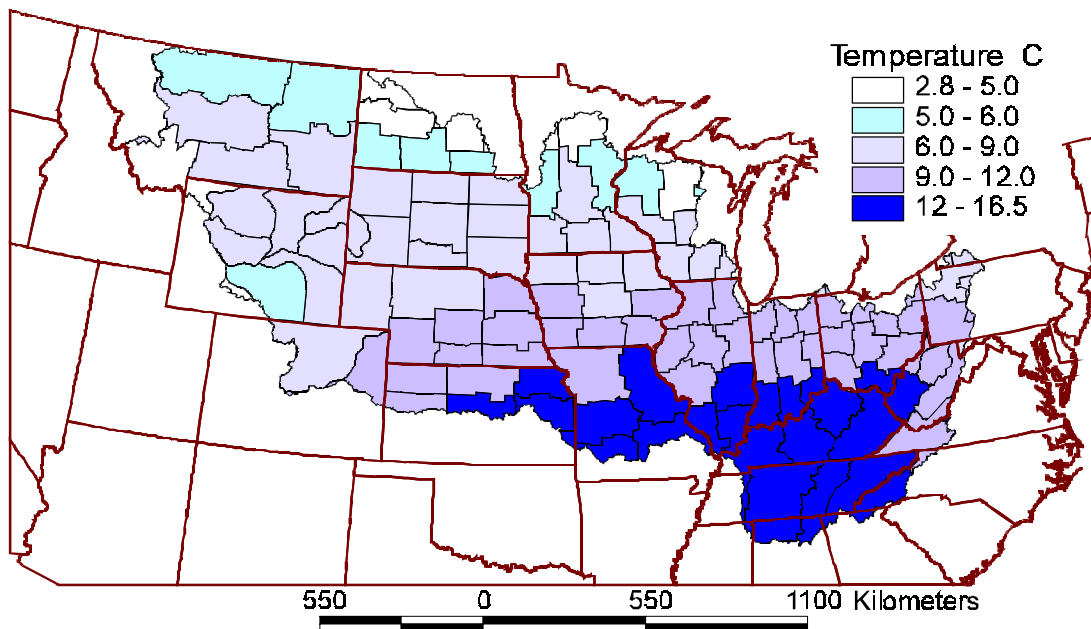
Additional data such as daily rainfall, snowfall, maximum temperature and minimum temperature, which were observed at 17,000 NCDC Stations, are also available on the CD-ROM set published by Hydrosphere Data Products, Inc. (Hydrosphere, 1993b).

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).

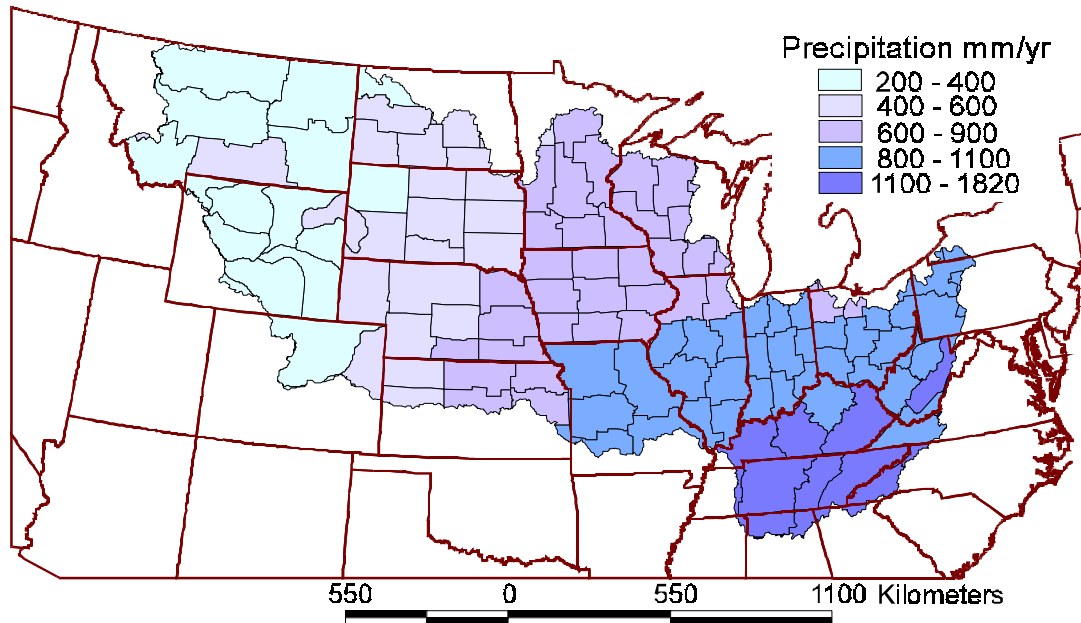
### 3.1.6 Maps of mean annual precipitation and temperature

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](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 [ $^{\circ}\text{C}$ ] 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.

## 4. METHODOLOGY

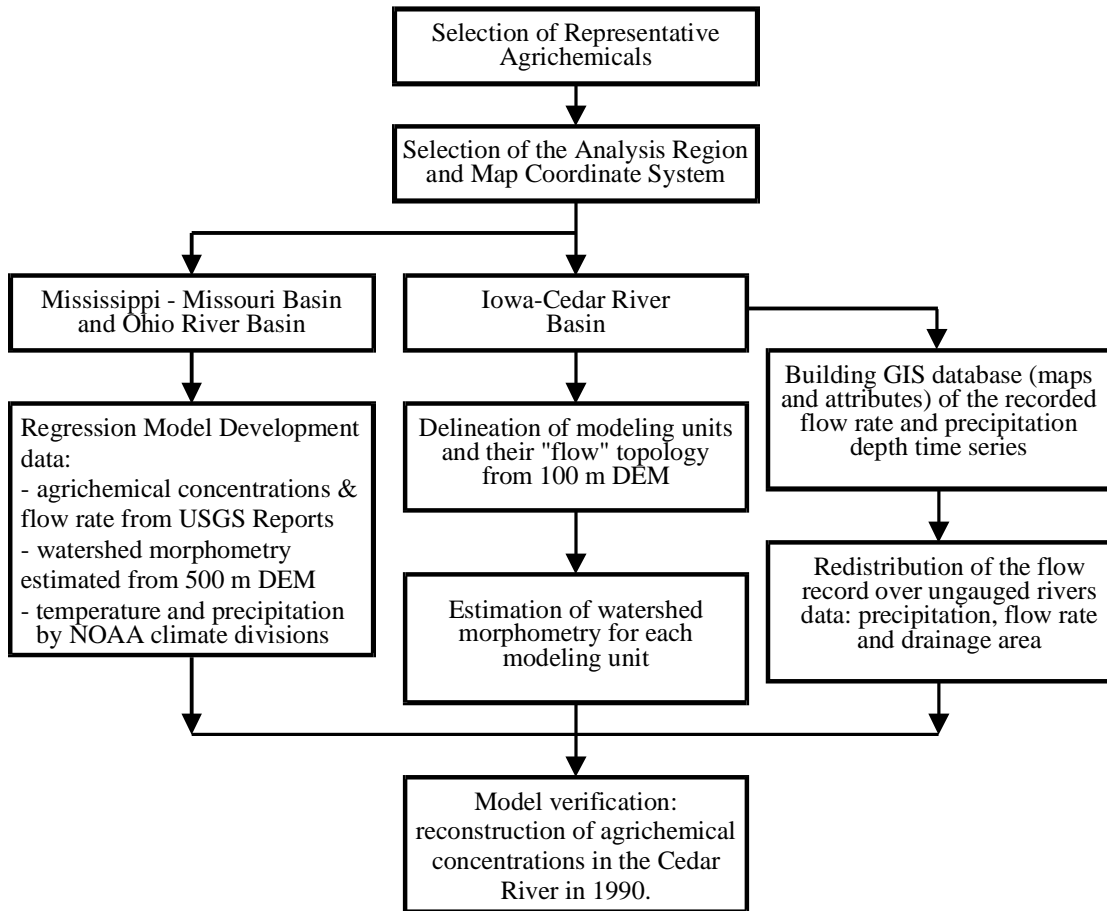
In this section a detailed description of the methodology developed for modeling agrichemical transport is presented. Two agrichemicals are selected: a nutrient, nitrate plus nitrite as nitrogen, and a herbicide, atrazine. These chemicals are introduced in Section 4.1. Section 4.2 describes two geographic analysis regions: one region used for developing a statistical model, and another smaller region used for GIS model development and its verification. The regression equations that relate chemical concentrations to the chemical application on the field, selected watershed descriptors, climatic variables, and the month of a year, are developed utilizing the USGS data collected in more than 150 watersheds scattered over the Mississippi - Missouri Basin above Thebes, Illinois, and the Ohio Basin above Grand Chain, Illinois. A GIS application of these equations is developed and verified for the Iowa-Cedar River watershed located in Iowa and Minnesota.

A detailed mathematical description of the statistical model is presented in Section 4.3. It provides an overview of the transport equations, introduces the explanatory variables which explain the spatial and seasonal (monthly) variations of the nitrate and atrazine concentrations in surface waters, and explains the concept of spatio-temporal “cascade” modeling within GIS. The watershed parameters of the statistical model are estimated from the 500 m DEM.

The GIS model of agrichemical transport in the Iowa-Cedar River basin is presented in Sections 4.4 - 4.6. Section 4.4 describes the methodology of discretizing the watershed, i.e., subdividing the basin into small (about 30 km<sup>2</sup>) hydrologic units and determining the flow connectivity between these units utilizing the 3 arc-sec DEM.

Since the GIS model is designated to estimate the concentrations and loads in all streams, a methodology of redistributing the observed flow rate over ungauged rivers has been developed. Section 4.5 presents details of the monthly flow redistribution technique as well as the application of the GIS capabilities to store and manipulate flow and precipitation time series.

The following flow chart summarizes the methodology developed for the large scale agrichemical transport in the Midwest rivers:



**Figure 4.1 Methodology of the large scale modeling of agrichemical concentrations in the Midwest rivers.**

An alternative and less successful approach to modeling agrichemical transport to the main one developed in this research is briefly introduced in Section 4.6. This model is based on a first order reaction as well as a method of travel time estimation which is presented there.

In this section and in the following sections, the words *river*, *stream*, and *stream/river reach* are used interchangeably. Similarly, *drainage area*, *watershed* and *basin* are considered here as equivalent terms. Sometimes one word *nitrate* is used to represent *nitrate plus nitrite as nitrogen*. *Unit watershed*, *elementary watershed*, and *modeling unit* refer to the smallest drainage area or partial drainage area into which the region under investigation is divided. Each modeling unit is considered as lumped system. Names of maps, computer files, database fields, as well as computer commands are printed in a Courier font.

#### **4.1 Representative agricultural chemicals**

Two constituents are selected for the study of the transport of agricultural chemicals in surface waters: nitrate plus nitrite as nitrogen, and atrazine. The following factors influenced this selection:

- nitrogen and atrazine are representative of nutrients and herbicides, respectively,
- nitrate plus nitrite as nitrogen and atrazine are present in measurable quantities in many Midwest streams, and

- nitrate plus nitrite as nitrogen was the only nutrient measured during both studies performed by the USGS: the reconnaissance study (Scribner, et al., 1993) and the analysis of agrichemical concentrations in storm runoff (Scribner, et al., 1994).

It is assumed that the nitrate plus nitrite concentrations in Midwest streams are mainly derived from chemical fertilizers.

#### 4.1.1 Nitrate

Nitrogen (N) in soils natural waters occurs as organic or inorganic N. The inorganic forms, include ammonium ( $\text{NH}_4^+$ ), nitrite ( $\text{NO}_2^-$ ), nitrate ( $\text{NO}_3^-$ ), nitrous oxide ( $\text{N}_2\text{O}$ ), nitric oxide (NO), and elemental N ( $\text{N}_2$ ). The three most important forms,  $\text{NH}_4^+$ ,  $\text{NO}_2^-$ , and  $\text{NO}_3^-$ , usually represent 2 to 5% of the total soil N. The source of  $\text{NH}_4^+$  is from mineralization of organic N and from fertilizers. During nitrification,  $\text{NH}_4^+$  is converted to  $\text{NO}_2^-$ , (which is toxic to plant roots) by bacteria *Nitrosomonas* ( $2\text{NH}_4^+ + 3\text{O}_2 = 2\text{NO}_2^- + 2\text{H}_2\text{O} + 4\text{H}^+$ ), and then oxidized to  $\text{NO}_3^-$  by *Nitrobacter* ( $2\text{NO}_2^- + \text{O}_2 = 2\text{NO}_3^-$ ). The  $\text{NO}_3^-$  anion is very mobile and subject to leaching losses (Tisdale, et al., 1993).

Nitrate in streams is derived from many anthropogenic and natural resources including chemical fertilizers, animal wastes, domestic sewage, legumes, mineralization of vegetation, soil organic matter, and from the atmosphere through electrical, combustion and industrial processes.  $\text{NO}_3^-$  is a very soluble and mobile anion. It can be transported from agricultural fields in both overland flow and subsurface flow, and by volatilization into the atmosphere. Ammonium is adsorbed by the soil colloids and moves very little until converted to  $\text{NO}_3^-$ . The following are typical concentrations in

streams stated in mg/L as nitrogen, where 1 mg N/L = 1.2159 mg NH<sub>3</sub>/L = 3.2845 mg NO<sub>2</sub>/L = 4.4268 mg NO<sub>3</sub>/L): total nitrogen 0.1-10 mg/L, organic nitrogen 0.1-9 mg/L, ammonia 0.01-10 mg/L, nitrite 0.01-0.5 mg/L, nitrate 0.23 mg/L, and nitrogen gas 0-18.4 mL/L (McCutcheon, et al., 1993).

Standard measurements of nitrogen in surface waters include Kjeldahl nitrogen (ammonia plus organic nitrogen), nitrite plus nitrate, nitrite, and nitrogen in plants. The USGS reports utilized in this research (Scribner, et al, 1993, Scribner, et al., 1994) do not contain data about organic nitrogen in surface waters. Only nitrite plus nitrate concentrations were measured during the USGS reconnaissance study of agrichemicals in Midwest rivers (Scribner, et al., 1993). Nitrate, nitrite, and ammonia were measured in two rivers, the West Fork Big Blue River near Dorchester, Nebraska, and the Sangamon River at Monticello, Illinois, during the analysis of agrichemicals in storm runoff from April, 1991 to March 1990 (Scribner, et al., 1994). In most samples of the Sangamon River the concentrations of ammonia were less than 0.1 mg/L, while the nitrate concentrations were 7 - 14 mg/L during the months from November to June and 0.4-3 mg/L during the months from July to October. The concentrations of nitrate in the West Fork River were smaller than the concentrations in the Sangamon River, rarely exceeding 3 mg/L . The September-October concentrations were below the reporting limit. The ammonia levels varied from about 0.1 to 0.4 mg/L except the summer months in which the ammonia concentrations in most samples were less than 0.05 mg/L.

The proportions of the different forms of the nitrogen in Midwest rivers may be illustrated by the samples collected in five sites (Floyd, Carville, Northwood, Cedar Falls, Gilbertville, and Bertram) along the Cedar River, Iowa, from May 1984 to September 1985 (Squillace and Engberg, 1988). The following median concentrations were reported:



- dissolved nitrite plus nitrate as nitrogen 3.5 - 5.1 mg/L
- dissolved ammonia as nitrogen 0.03 - 0.1 mg/L
- dissolved organic nitrogen 0.25 - 0.85 mg/L

Thus it can be seen that most nitrogen in streams of the Midwest is present as nitrate plus nitrite.

In contrast to overland transport, in which nitrate takes minutes or hours to get to a stream, downward vertical leaching and subsequent underground travel is a long process which takes months or years. The soil system has a strong memory with respect to nitrate production and leaching. Jones and Burt (1993) presented a study in which 64% of annual nitrate concentration in streams was explained by a stepwise regression involving the year of measurement, and each of the previous two years. It may take nitrate years or even decades to appear in rivers as a base flow pollutant.

There are some losses of nitrate due to erosion, but for humid temperate climates, erosion is generally an insignificant process compared with leaching and runoff (OECD, 1986). Other authors indicate that the adsorption has no marked influence on the rate of  $\text{NO}_3^-$  movement (Keeney, 1983; Bailey and Swank 1983). The predominant losses of nitrate in an agricultural field are due to assimilation by row crops and by other terrestrial and aquatic plants.

Although the best time of fertilizer application is at the time of peak N demand of the crop, it is seldom feasible to apply the chemical then. In north central United States most of the N application occurs late summer and fall. It is influenced by the following factors (Tisdale, et al., 1993):

- Temporal and spatial distribution of the rainfall. Because of N mobility in soils, the greater the surplus rainfall, the greater the possibility of loss of N through leaching.

- Temporal and spatial distribution of temperature. Since higher temperatures enhance nitrification, ammonia N applied before planting is more subject to nitrification and leaching.
- Technical factors. In late winter the ground may be too wet for machinery to be operated and spring application is usually too late for small grains to respond in yield to the nitrogen fertilizer applied.

Fall application of nitrogen fertilizers as well as the decrease in temperature during late fall, winter, and early spring, causes high nitrate plus nitrite concentrations in surface waters at those times. These concentrations decrease in late spring and summer when the plant demands for nutrients are high (Goolsby and Battaglin, 1983; Davis and Keller, 1983). In addition, a significant decrease in nitrate concentration in surface water may result from assimilation of nitrate by algae and by instream riparian macrophytes (Heathwaite, 1993; Moore, 1991) as well as nitrate may be converted by the denitrification bacteria and various chemical processes into free nitrogen and nitrogen oxides which escape into the atmosphere. Since these processes are stimulated by high temperatures and low flow rates, the highest loss of nitrate in lakes and rivers occurs during the summer. Lakes and reservoirs act as a "buffer," thus they are less responsive to seasonal changes than are rivers (OECD, 1986).

Denitrification and other processes of biochemical degradation of nitrate can be modeled as a first order reaction with an "overall" decay rate dependent on temperature and carbon content. Some losses may result from infiltration (river water seepage into groundwater).

#### 4.1.2 Atrazine

Atrazine is a herbicide that controls broadleaf weeds in fields of corn and sorghum. It is one of the most widely used herbicides in the United States (Comfort and Roeth, 1996). The EPA has set the drinking water health limit (MCL, Maximum Contamination Level) for atrazine at 3 µg/L (ppb). Conventional water treatment does not remove this herbicide. Recent studies conducted in 29 communities throughout the Midwest, Louisiana, and the Chesapeake Bay detected high concentrations of atrazine in tap water during months from May through July, some of them exceeding EPA MCL of 3 µg/L (EWG, 1996)

Numerous laboratory tests as well as field studies have been performed to determine the behavior of this herbicide in different chemical and physical environments for almost half of the century. Some of the published parameters are as follows: Atrazine (2-chloro-4-ethylamino-6-isopropylamino-s-triazine) is a low solubility herbicide; its water solubility in typical temperature and pH varies from 30 to 35 ppm (mg/L). The volatility of this chemical is very low (vapor pressure =  $0.3 \times 10^{-6}$  mm Hg = 0.00004 Pa). Published values of the octanol extrability coefficient (soil sorption coefficient)  $K_{oc}$  are from 130 to 172. The octanol-water partition coefficient  $K_{ow}$  equals 251 (e.g., Weber, 1972; Hamaker, 1975; Wauchope, 1978; Rao, et al., 1983; Weber, 1988; Plimmer, 1988).

There are two parameters that characterize chemical decay in soil: half-life and chemical persistence. The half-life is the length of time it takes for a sample to reduce to half of its original weight or mass. The chemical persistence is the time for 90% disappearance of chemical from soil. Kruger, et al. (1993) reported half-life of atrazine in soil under unsaturated conditions ranged from 41 d to 231 d, whereas in saturated soil at the 90-120 cm depth, the half-life was 87 d. Similar values were cited by

Goring, et al. (1975). Wauchope (1978) determined that atrazine persistence in soil is 12 months and may vary from 6 to 18 months, which corresponds to a half-life from 55 d to 165 d, depending on climate and soil.

Atrazine decays into many degradation products which can be more persistent and mobile than their parent compound. During a USGS 1991 Mid-continent survey of near-surface aquifers, deethylatrazine, an atrazine metabolite, was the most frequently detected compound followed by atrazine, and then deisopropylatrazine, another atrazine metabolite (Kolpin et al., 1983). The half life of the deisopropylatrazine is much longer in surface water than in soil (Goolsby et al., 1993 pp. 51-62, Comfort and Roeth, 1996). However, Kruger et. al. (1993) found that deisopropylatrazine may be even less persistent under saturated conditions than in saturated soil. Their estimate of the deisopropylatrazine half-life ranged from 32 d to 173 d in the top 30 cm of unsaturated soil, and from 58 d to 173 d in saturated soil at 90 to 120 cm depth.

Weber (1988) presented results of research, which he conducted with J. A. Best, on the effect of pH on the dissipation of atrazine applied to soil. No parent component volatilization was detected; 0.1%-0.2% was found in leachate, plants used two to four percent. Ninety percent of atrazine was retained in the soil layer.

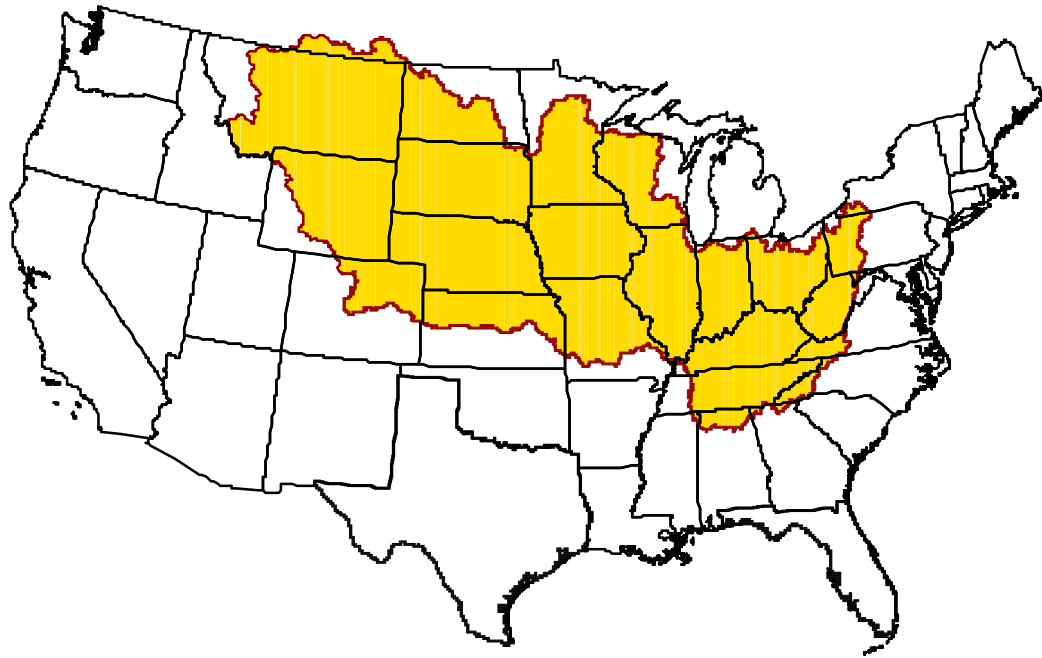
The adsorption and movement of s-triazines in soil depends upon such factors as soil organic matter, clay minerals, pH, temperature, soil moisture, concentration and species of other ions in the system (Weber, 1972; Weber, 1988; Goring, et al., 1975). In the CREAMS model, pesticide in runoff is partitioned between the solution phase and the sediment phase (Knisel, et al., 1983).

Study by the USGS of the occurrence of herbicides in precipitation in the Midwest and Northeastern United States showed that significant amounts of atrazine were lost through volatilization and subsequently returned to the land through precipitation washout. The amount of atrazine in precipitation washout was

approximately equal to one half the loading found in the Mississippi River flowing into the Gulf of Mexico (Goolsby, et al., 1993, pp. 75-86). Plimmer (1988) discussed reports which described the identification of atrazine in fog. These findings were apparently in contradiction to the findings about negligible volatilization published by Weber (1977, 1988). As was pointed out by Kenaga (1975), volatility of a pesticide is applied to heterogeneous surfaces such as natural water, soil, foliage, wood, or glass, is variable because sorption varies. There is a possibility that atrazine enters the atmosphere adsorbed on particulate matter through dust blowing from the land surface. The process of atrazine volatility is not well understood.

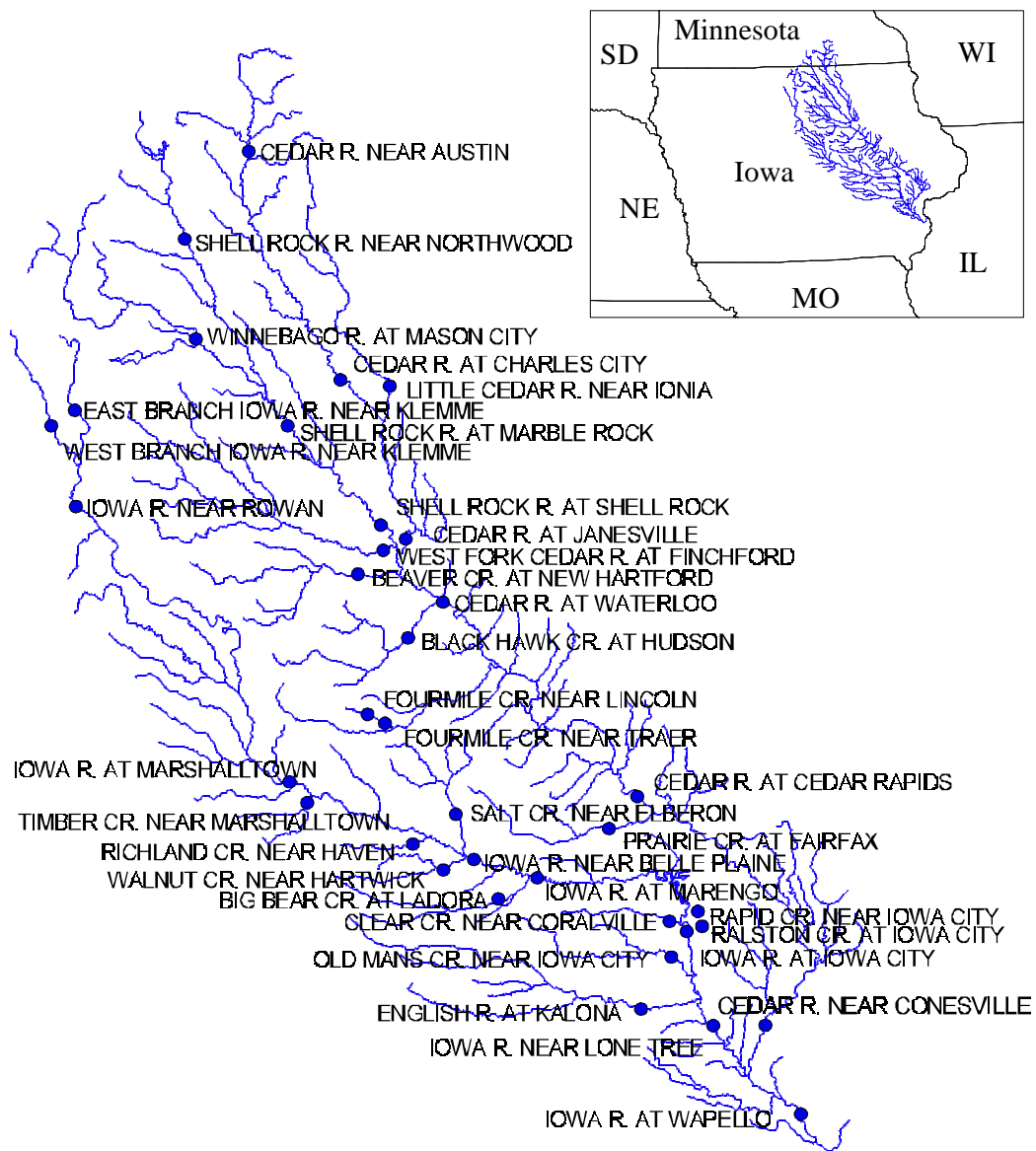
#### ***4.2 Selection of analysis region and map coordinate system***

The Mississippi - Missouri River basin above Thebes, Illinois (drainage area  $1.85 \times 10^6 \text{ km}^2$ ) together with the Ohio River basin above Grand Chain, Illinois (drainage area  $0.53 \times 10^6 \text{ km}^2$ ) constitute the primary region that is used for estimation of statistical model parameters. Its extent is determined by the USGS reconnaissance study of selected herbicides and nitrate in Midwestern United States (Scribner, et al., 1993). This region is one of the most extensive agricultural areas in the country, producing over 80% of all US corn and soybeans (Oberle and Burkart, 1994).



**Figure 4.2 The Mississippi-Missouri and Ohio River Basins.**

Since processing data for an area that covers almost three million square kilometers requires a computer with adequate operational and storage memory, the final model is built, and verified on a selected subregion, i.e., the Iowa-Cedar River watershed in Iowa (area = 32,000 km<sup>2</sup>). There are two sites in the Iowa-Cedar Basin for which measurement of agricultural chemicals are available: Old Man's Creek near Iowa City and Cedar River at Palisades, Iowa. Flow rate is recorded in about 30 USGS gauging stations in the Iowa-Cedar Basin. A diversity of such geographic features as lakes, reservoirs, wetlands, hills, plains, and a wide range of stream sizes makes the Iowa-Cedar Basin a very good representation of the Midwest. Figure 4.3 shows the Iowa and Cedar Rivers, and the location of gauging stations.



**Figure 4.3** The Iowa River with tributaries and the USGS gauging stations.

All maps utilized in this research are represented in Albers Conical Equal Area Projection, generally accepted for large maps of the USA. This projection has the valuable property of equal area representation, combined with a scale error that is practically the minimum attainable in any system covering such a large area in a single sheet.

The following parameters are applied (standard for USA): units = meters, first standard parallel = 29°30'00", second standard parallel = 45°30'00", latitude of projection's origin = 23°00'00", false easting = 0.000 m, false northing = 0.000 m, longitude of central meridian = -96°00'00". Using one common projection for the whole Midwest eliminates the problems associated with merging separately modeling regions into one unit.

The Albers projection is of the conical type, in which the meridians are straight lines meeting in common point beyond the limits of the map, and the parallels are concentric circles, the center of which is at the point of intersection of the meridians. The meridians and the parallels intersect at right angles and the arcs of longitude along any given parallel are of equal length. The spheroid is intersected by a cone at two parallels known as the *standard parallels* for the area to be represented. On the two standard parallels, arcs of longitude are represented by their true lengths, or at an exact scale. Between the standard parallels, the scale along the meridians is too large and beyond them too small (Deetz and Adams, 1969).

The Albers projection is constructed in such a way that the area of the earth's surface between any pair of parallels and meridians is correctly preserved in the flat map representation.



## 4.3 Mathematical description

### 4.3.1 Overview of transport equations

There are two basic mechanisms that are responsible for the transport of dissolved and suspended solutes in surface waters: advection and diffusion/dispersion. These two processes are described by the advection-dispersion equation which is the fundamental equation for majority of the pollutant transport models. Equation (4.1) describes one-dimensional advection-dispersion in a reach with a uniform cross-sectional area.

$$\frac{\partial c}{\partial t} = -\frac{\partial vc}{\partial x} + \frac{\partial}{\partial x} E_x \frac{\partial c}{\partial x} + \sum_{i=1}^n S_i + \sum_{j=1}^m K_j c \quad (4.1)$$

where:

$c$  = concentration of pollutant [ $\text{g}/\text{m}^3$ ]

$t$  = time [s]

$x$  = distance along the river [m]

$-E_x \partial c / \partial x$  = mass flux due to the longitudinal dispersion [ $\text{g}/\text{m}^2\text{s}$ ]

$E_x$  = longitudinal dispersion coefficient [ $\text{m}^2/\text{s}$ ]

$-vc$  = mass flux due to advection [ $\text{g}/\text{m}^2\text{s}$ ]

$S_i$  =  $i$ -th source/sink of the constituent [ $\text{g}/\text{m}^3\text{s}$ ]

$i = 1 \dots n$ ,  $n$  = number of sources/sinks

$K_j$  = decay rate due to the  $j$ -th process [1/s]

$v$  = flow velocity [m/s]

The source/sink term,  $S_i$  and reaction term,  $K_j c$  represent a wide range of features such as lateral flux, transient storage, biotic and abiotic retention, benthic flux, periphyton retention, sediment retention, reaeration, photosynthesis, and nitrification (U. of Mississippi, 1990, James and Elliot, 1993).

Dynamic models solving (Eq. 4.1) have been used mainly for pollution incidents such as spills and runoff discharges. Their applicability for large scale modeling of nitrate/atrazine transport at the scale attempted in this research is limited, mainly because:

- Numerical solution of the advection-dispersion equation requires subdivision of the time domain into relatively short intervals (minutes, hours), but this is less useful when calculating monthly means of chemical transport in extensive stream networks over large areas;
- It is difficult to write a procedure which solves the dispersion-advection equation using a GIS script language, so such a model has to be solved externally to the GIS.

Lagrangian type transport models, such as a Moving Segment Model (MSM), can be efficiently incorporated into GIS. In MSM (James and Elliot, 1993) the stream is subdivided into series of segments. Within each segment the variations of chemical concentration are calculated by summing, for example, hourly changes due to all the processes involved. The process within the block is described by the following equation:

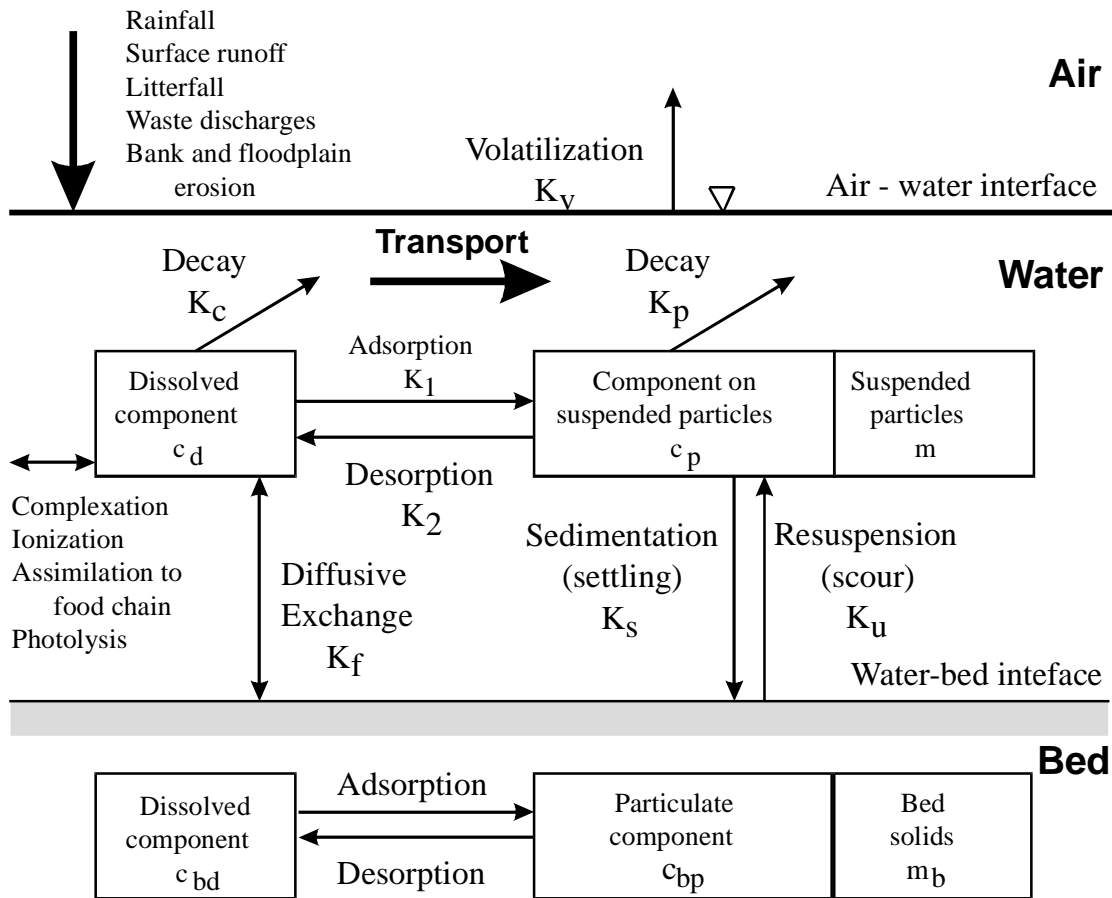
$$\frac{\partial c}{\partial t} = \pm \sum_{i=1}^n S_i + \sum_{j=1}^m K_j c \quad (4.2)$$

where:  $S_i$  is the  $i$ -th source/sink and  $K_j$  is the  $j$ -th reaction coefficient.

The segments (blocks) move downstream with travel time  $\tau$  which may be a function of such parameters as flow rate, cross-sectional area, friction coefficient, slope, and stream curvature. For a detailed representation of the constituent transport in surface water at least the following processes should be represented (O'Connor, et al., 1983; Thomann and Mueller, 1987; O'Connor, 1988a,b; University of Mississippi, 1990):

- Sorption and desorption between dissolved and particulate components in the sediment and water column;
- Settling and resuspension of particles;
- Diffusive exchange between the sediment and water column;
- Loss and gain of the chemical due to the chemical and biochemical reactions such as biodegradation, volatilization, and photolysis;
- Advective and diffusive transport of the chemical in water and as a bed transport;
- Net deposition and loss of chemical to deep sediments.

Figure 4.4 shows these major reaction mechanisms and transfer routes of chemicals and solids in both river water and the river bed.



**Figure 4.4** Reactions and transfers in a natural water system (after O'Connor, et al, 1983; Thomann and Mueller, 1987; O'Connor, 1988a, b; University of Mississippi, 1990).

The processes in surface waters can be described by the mass balance equations for the dissolved and particulate components (Eq. 4.3 and Eq 4.4). For a specific chemical, e.g. nitrate or atrazine, some processes must be included in the transport model whereas some processes may not significant and therefore can be neglected.

$$\frac{dc_s}{dt} = -K_1 c_s + K_2 c_p - K_v c_s + K_f (c_{bs} - c_s) - K_c c_s + K_p c_p - K_q (c_s + c_p) \quad (4.3)$$

$$\frac{dc_p}{dt} = K_1 c_s - K_2 c_p - K_s c_p + K_u (c_{bp} - c_p) - K_c c_{bs} + K_p c_{bp} \quad (4.4)$$

where:

$K_v, K_f, K_s, K_u$  = the bulk transfer coefficients of volatilization, dissolved exchange, settling and scour respectively;

$c_s, c_p$  = dissolved and particulate concentrations in water, respectively;

$c_b, c_{bp}$  = dissolved and particulate concentrations in bed;

$K_c, K_p$  = decay coefficients of dissolved and particulate form;

$K_1 = K_0 r_c m$ ;  $K_0$  = adsorption coefficient,  $r_c$  = adsorptive capacity,

$m$  = concentration of solids;

$K_2$  = desorption coefficient;

$K_q$  = coefficient of dilution due to the groundwater inflow.

### 4.3.2 GIS and cascade modeling

GIS technology gives the opportunity to construct versatile “cascade” models. The idea of cascade modeling incorporated here into GIS has been extracted from the methodology used in forecasting the municipal water use (Maidment and Parzen, 1984; Mizgalewicz, 1991).

The GIS cascade structure is two dimensional. It can be applied both in space and time. In the time domain, a general model describes spatial distribution of annual average amounts of agricultural chemicals in rivers. The annual values are then broken into seasonal or monthly values. For example, the annual average concentrations  $c_y$  for the Midwest can be estimated by the general function:

$$c_y = f(\mathbf{X}) \quad (4.5)$$

where:  $\mathbf{X}$  is a vector of explanatory variables such as annual agrichemical application and watershed morphometry (area, land slope, stream slope, stream length, overland flow length) which do not change with time.

The Equation (4.5) can be further extended by adding a seasonal component, as monthly fractions  $S(m)$ , to break down the annual predictions of concentration  $c_y$  into monthly values  $c(m)$ :

$$c(m) = c_y S(m) = f(\mathbf{X}) S(m) \quad (4.6)$$

Cascade modeling enables further extending of the Eq. (4.6) by adding such elements as year to year trend and an irregular random component.

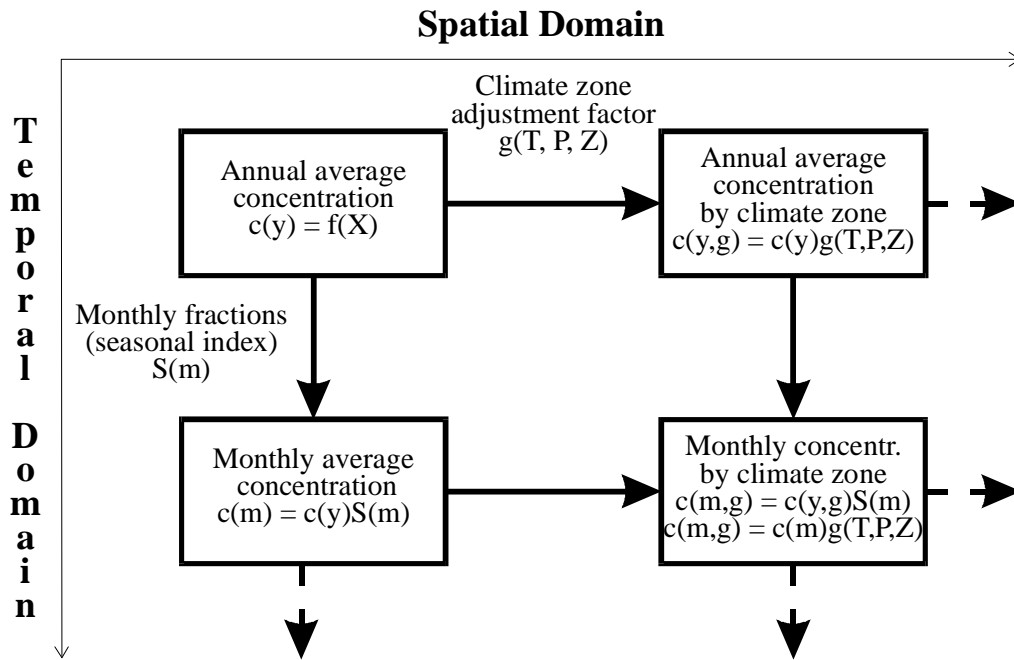
Cascade modeling in the spatial domain implies that the modeling process is subdivided into several levels of resolution, i.e., the Upper Mississippi-Missouri and Ohio River can be subdivided into three basins: the Missouri River Basin above junction with the Mississippi River, the Upper Mississippi River Basin, and the Ohio River Basin. The transport of agricultural pollutants in the Upper Mississippi River is estimated by utilizing results of sub-models that describe transport in individual component basins such as the Des-Moines River, the Skunk River, the Iowa-Cedar River, and Wisconsin River Basins.

Additional spatial subdivision of the Midwest can be performed by introduction of climate zones. Each zone is defined, for instance, by a specific range of temperatures and precipitation depths. Thus, the Eq. (4.6) could have the following form:

$$c(m, g) = f(\mathbf{X}) S(m) g(\mathbf{T}, \mathbf{P}, \mathbf{Z}) \quad (4.7)$$

where:  $c(m, g)$  is the monthly ( $m$ ) agricultural concentration in rivers modified for the climate zones ( $g$ );  $g(T, P, Z)$  is a function of  $T$  - zonal temperature,  $P$  - zonal precipitation depth, and  $Z$  - zone location.

The diagram shown in Figure 4.5 illustrates the example of spatio-temporal cascade modeling within GIS.



**Figure 4.5 Example of the cascade modeling within the GIS.**

Parameters of the mathematical description of the chemical application-runoff process and chemical losses in streams are stored in a GIS attribute table. Moreover, the equations are stored as objects in a database table. Storing equations as a database objects not only permits one to fully implement the cascade modeling technique into the GIS but also it simplifies the structure of the model, allows one to test and compare an unlimited number of equations, and permits changes to the mathematical

description by simple record editing operations. The model prototype is constructed within ArcView, a GIS application (ESRI, 1995).

### **4.3.3 Regression equation development**

There is no complete and consistent data base of parameters required for comprehensive modeling the agrichemical runoff from the field and its transport in surface waters of the Midwest. The runoff and transport parameters have to be estimated from such data as:

- (1) The observed flow rate and the measured agrichemical concentrations in the variety of watersheds scattered over the upper Mississippi-Missouri and Ohio River basins;
- (2) The DEM from which a watershed morphometry can be calculated; and
- (3) Parameters that describe the climatic conditions of the sampled watershed.

Thus, a model based on statistical analysis of observed concentration data is used in this research.

Watershed morphometry, annual agrichemical rate, normal annual temperature, and normal precipitation depth do not depend on the month of the year. These parameters are used here to explain the average annual agrichemical concentration  $c_y$  in Midwest rivers. The following linear relationship is tested by the regression analysis (A detailed discussion of the set of explanatory variables that have potential application in the large scale model of agrichemical runoff from the field and transport in the river network is presented in Section 4.3.4 and Section 4.3.5):



$$c_y = \beta_0 + \beta_1 A_p + \beta_2 S_L + \beta_3 L_L + \beta_4 A + \beta_5 E_S + \beta_6 S_S + \beta_7 Q + \beta_8 T + \beta_9 T_{avg} + \beta_{10} P + \beta_{11} P_{avg} + error \quad (4.8)$$

where:

- $c_y$  = average annual constituent concentration (concentration with removed seasonal component) [mg/L nitrate,  $\mu$ g/L atrazine];
- $A_p$  = annual chemical application rate [kg/km<sup>2</sup>/yr];
- $S_L$  = average slope of the land [dimensionless];
- $L_L$  = average length of the constituent travel path, from the point of application to the stream network within a given watershed to the sampling point [km];
- $A$  = drainage area [km<sup>2</sup>];
- $E_S$  = average of the exponent of negative flow distance in streams
- $$= \frac{1}{n_s} \sum_{i=1}^{n_s} \exp(-L_{ik});$$
- $n_s$  = number of cells that constitute the stream network within a sampled watershed;
- $L_{ik}$  = length of the flow path from the  $i$ -th stream cell to the watershed outlet  $k$  [10<sup>2</sup> km] (since the Missouri River is a very long river, units 10<sup>2</sup> km are applied to enhance the Arc/Info-GRID calculations);
- $S_S$  = average slope of the stream network [dimensionless];
- $Q$  = flow rate at a given stream location [m<sup>3</sup>/s];
- $T$  = normal annual temperature at sampling site [°C]
- $T_{avg}$  = normal annual temperature, average over sampled watershed [°C]
- $P$  = normal annual precipitation depth at sampling site [mm]
- $P_{avg}$  = normal annual precipitation depth, average over sampled watershed [mm]
- $\beta_0, \beta_1, \beta_2 \dots$  = regression coefficients.

The right hand side of Eq. (4.8) represents two major components of the transport process: runoff from the field and losses in the stream, as well as the climatic conditions that affect the transport processes:

- 1)  $\beta_1 A_p + \beta_2 S_L + \beta_3 L_L + \beta_4 A$  describe changes in mass applied  $U_R$  on a field as it travels from the point of application to a stream;
- 2)  $\beta_5 E_S + \beta_6 S_S + \beta_7 Q$  reflect the losses of agrichemical in rivers; and
- 3)  $\beta_8 T + \beta_9 T_{avg} + \beta_{10} P + \beta_{11} P_{avg}$ , incorporates into the model the effects of the climate zone on the agrichemical transport.

More details on each of the model variables are now presented.

#### **4.3.4 Agrichemical runoff from the field**

***Agrichemical application rate  $A_p$ .*** Since it is a common practice to put more chemical on the field than the amount that can be completely utilized by the vegetation and the chemical - microbiological processes, some of the nutrients and herbicides are transported in runoff from the field. Many studies show that the more chemical is applied on fields, the higher are chemical loads carried by rivers (e.g., Battaglin, et al., 1993).

***Land slope  $S_L$ .*** The higher the slope of the land  $S_L$ , the chemicals are more susceptible to washout, therefore greater losses of the chemical from the field can be expected. This complex process is influenced by gravitational forces. Since vegetation has a smaller chance to uptake nutrient or herbicide when the land slope is higher (due to the shorter residence time), a greater portion of the mass applied on the field should reach

the stream network. Models of erosion and sediment transport include a slope parameter.

***Distance to a stream  $L_L$ .*** The influence of the average distance  $L_L$  between the point of herbicide or nutrient application and the location of the stream network on the amount of mass that enters the stream is not as obvious as watershed slope. (Here, the  $L_L$  is an average distance measured along the flow path between all cells that do not represent a stream network and the cells that constitute the streams, within a given watershed). Intuitively it may be expected that since the longer the average travel path is, the larger the chemical losses are and thus the coefficient  $\beta_3$ , from Eq. (4.8), estimated by the regression analysis should be negative. But, since the predictor variable  $L_L$  represents complex watershed features, the coefficient  $\beta_3$  does not necessarily have to be negative. The average length of the land-flow path depends on the stream density, i.e., on the length of the streams per unit area of watershed. The more streams there are within a given watershed, the shorter the length of the overland flow. The density of the stream network influences the amount of chemical that enters a unit length of the stream. Therefore, for watersheds characterized by a higher value of  $L_L$ , the amount of chemical that enters a stream (mass per unit length) is higher than the amount reaching the stream network which is located in a watershed with smaller  $L_L$ .

***Climate.*** Both precipitation ( $P, P_{avg}$ ) and temperature ( $T, T_{avg}$ ) have an impact on vegetation growth. Moreover, the greater the water surplus (precipitation minus potential evaporation), the greater the possibility of loss of agrichemical through leaching if the crop is not growing vigorously or through washout if the land is not protected by a plant cover. Denitrification depends on the amount of water in soil, whereas nitrification rate is highly correlated with the temperature (Tisdale, et al., 1993).

The influence of climate on the concentrations in surface waters can be illustrated by comparison of studies performed in different climatic regions. For example, the largest number of atrazine detections in Swedish stream waters was in July, August and September (Kreuger and Brink, 1993), whereas in the Midwestern United States the major atrazine runoff occurs in May and June (Scribner, et al., 1994). The Swedish vegetation period is short (6-8 months) and cold (3°-17° C)

Since the region under scrutiny extends from about 37° N to 50° N (latitude) and from 79° W to 114° W (longitude) the spatial distribution of the average temperature and the average precipitation influences not only the spatial distribution of the chemical runoff from the field but it also causes spatially different agrichemical application times. The difference in climate conditions within the studied region could be represented by a spatially distributed adjustment coefficient as well as a time shift introduced into periodic functions that describe seasonal variation of agrichemical concentration in the surface water. The time-shifting of periodic functions is not studied in this research.

#### 4.3.5 Transport in rivers

*Exponent of negative stream-flow distance*  $E_S$ . The explanatory variable  $E_S$  has been introduced to represent agrichemical losses in streams. It is defined by the following formula (excerpted from Eq. 4.8):

$$E_S = \frac{1}{n_s} \sum_{i=1}^{n_s} \exp(-L_{ik}) \quad (4.9)$$

where:

- $n_s$  = number of cells that constitute the stream network within a sampled watershed;
- $L_{ik}$  = length of the flow path from the  $i$ -th stream cell to the watershed outlet  $k$  [ $10^2$  km]

Equation (4.9) is a simplified version of the discrete model of exponential losses of chemical as it travels downstream. The following paragraph explains the development of the explanatory variable  $E_s$ .

The losses of agrichemical in streams can be described by an exponential function of the travel time (the traditional approach) or by an exponential function of the travel distance (Smith. et al., 1993). Although it is possible to estimate the travel time from the observed flow time series, in this research the agrichemical decay has been related to the travel distance, i.e., the amount of chemical that enters the stream in point  $i$  (cell  $i$  in raster representation of river) decays as it travels downstream according to the following equation:

$$R_k = \sum_{i=1}^n R_i \exp(-k_S L_{ik}) \quad (4.10)$$

where:

- $R_k$  = total agrichemical mass in a runoff from  $k$ -th watershed [ $\text{kg}/\text{m}^3$ ];
- $k..$  = index of sampled watershed outlet (cell  $k$ );
- $R_i$  = chemical load that enters the stream in point  $i$  (cell  $i$ ) [ $\text{kg}/\text{m}^3$ ];
- $n$  = number of all cells that constitute the stream network located within  $k$ -th watershed;
- $k_S$  = overall distance decay coefficient [ $1/10^2$  km];
- $L_{ik}$  = length of the flow path from stream cell  $i$  in which the chemical runoff from a

field enters the stream network to the watershed outlet located in cell  $k$  [ $10^2$  km].

Since application of the Eq. (4.10) is very computationally intensive, the following assumptions have been made to construct its simplified form, i.e., equation (4.9):

- a unit mass of chemical enters the stream network ( $\sum R_i = 1$  within sampled watershed);
- the chemical that enters surface water is uniformly distributed over all stream cells, i.e., the amount of chemical that enters the stream at each cell of the stream network within sampled watershed equals  $1/n_s$ ;
- the overall distance decay coefficient is equal one,  $k_S = 1$  (the chemical losses depend only on the chemical travel distance  $L$ )

**Stream slope  $S_s$ .** The stream slope parameter is introduced into the agrichemical transport equation to represent two physical features: flow velocity that is directly related to the chemical travel time, and stream power that affects sedimentation and resuspension of deposits as well as their transport.

#### **4.3.6 Seasonal variations**

The seasonal variations of the nitrate concentration as well as the atrazine concentration in surface waters are modeled by two sets of 12 values, calculated by the following formula:

$$S(m) = \frac{\exp\left(\sum_{k=1}^5 (a_k \sin(2k\pi m / 12) + b_k \cos(2k\pi m / 12))\right)}{\sum_{i=1}^{12} \exp\left(\sum_{k=1}^5 (a_k \sin(2k\pi i / 12) + b_k \cos(2k\pi i / 12))\right)} \quad (4.11)$$

where:

$S(m)$  = the seasonal factor of month  $m$  (average  $S(m) = 1$ )

$\sin(2k\pi m/12)$ , and  $\cos(2k\pi m/12)$  = components of the Fourier series (the cycle corresponding to  $k = 1$  has a 12-month period,  $k = 2, 3, 4$ , and  $5$  are harmonics of period  $12/m$  months;

$m$  = month of a year (1 for January);

$i$  = index of the month ( $i = 1, 2, \dots, 12$ );

$k$  = index of the harmonics;

$a_k$  and  $b_k$  = regression coefficients.

The exponent of the periodic function in Eq. (4.11) is a result of the agrichemical transport model specification in which a linear function that describes the average concentration in a stream is multiplied by a seasonal factor. To estimate the coefficients  $a_k$  and  $b_k$  from the concentration data the logarithmic transformation of concentration data is necessary. Two models have been selected to determine the variation of the constituent concentration in sampled rivers. The first one explains monthly changes by selected harmonics of the Fourier series (Eq. 4.12), whereas the other model, in addition to the sine-cosine components, embodies the flow rate to isolate the flow effect from the systematic seasonal the variations (Eq. 4.13):

$$\ln[c(j, d)] = w_j + \sum_{k=1}^5 (a_k \sin(2k\pi m / 12) + b_k \cos(2k\pi m / 12)) \quad (4.12)$$

and

$$\ln[c(j, d)] = w_j + a_0 \ln[Q(j, d)] + \sum_{k=1}^5 (a_k \sin(2k\pi m / 12) + b_k \sin(2k\pi m / 12)) \quad (4.13)$$

where:

$\ln[\dots]$  = natural logarithm;

$c(j, d)$  = concentration measured at site  $j$  on day  $d$  ( $\mu\text{g/L}$  or  $\text{mg/L}$ );

$Q(j, d)$  = flow rate measured at site  $j$  on day  $d$  ( $\text{m}^3/\text{s}$ );

$w_j$  = an intercept specific for the  $j$ -th sampled watershed, determined by the regression analysis;

$a_0, a_k$  and  $b_k$  = regression coefficients;

$j$  = index of the sampling site;

$d$  = day of sample collection;

$k$  = harmonics number;

$m$  = month of the year.

By combining (Eq. 4.11) with (Eq. 4.12) and (Eq. 4.11) with (Eq. 4.13) the relation between the average annual concentration in a sampled stream can be given by the following equations:

$$c_y(j) = \exp(w_j) \sum_{i=1}^{12} \exp\left(\sum_{k=1}^5 (a_k \sin(2k\pi i / 12) + b_k \sin(2k\pi i / 12))\right) \quad (4.14)$$

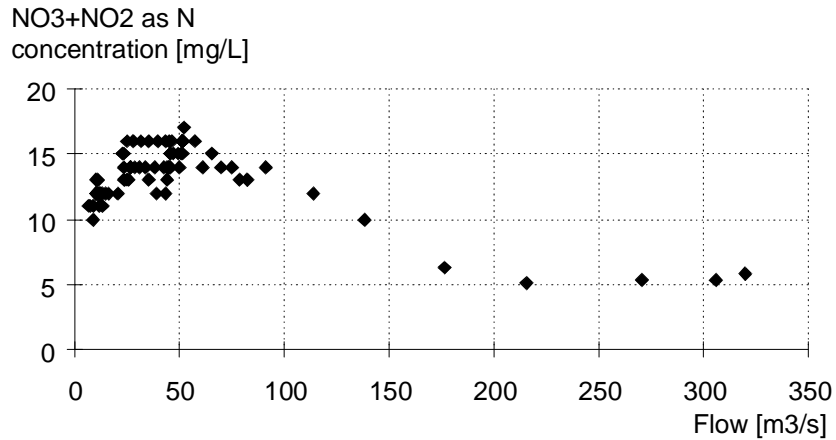
and the model including the flow rate is specified by:

$$c_{yQ}(j) = \exp(w_j) Q^{a_0} \sum_{i=1}^{12} \exp\left(\sum_{k=1}^5 (a_k \sin(2k\pi i / 12) + b_k \sin(2k\pi i / 12))\right) \quad (4.15)$$



The flow rate is introduced into analysis to examine the effects of the flow rate on the concentration level. However, the data utilized in this research may be biased by extreme flow conditions. At least one third of the reconnaissance data (Scribner, et. al., 1993) contain measurements made during the first major runoff after application of herbicides (33% of atrazine samples collected in 1989, and 50% of atrazine samples taken in 1990). Twenty five percent of nitrate plus nitrite as N measurements were collected in May and June of 1989 and 50% of nitrate samples represent concentrations in major runoff events after fertilizer application in 1990. Also, the gathering of samples in nine rivers of Midwest (Scribner, et. al., 1994) was designed to study the concentrations of selected herbicides and nutrients in storm runoff rather than to investigate the seasonal changes of agrichemicals in the Midwest rivers.

The general relation between flow rate and the chemical concentration is as follows: When flow increases, concentration increases. This is a result of the “washout effect”. If the flow increases above a level at which the chemical washout is in balance with the chemical dilution, the concentration does not increase. Further increase in flow may produce further chemical dilution and therefore decrease in concentration. Figure 4.6 shows an example of such a nitrate concentration pattern observed in the Sangamon River at Monticello, Illinois, in May 1990 (Scribner, et al., 1994). As flow increases up to about  $40 \text{ m}^3/\text{s}$ , the nitrate concentration level rises. A further increase of the flow, up to  $80 \text{ m}^3/\text{s}$ , does not produce a change of the concentration level. For flows larger than  $80 \text{ m}^3/\text{s}$  the dilution effect becomes significant and the nitrate concentration decreases when the flow rate increases.



**Figure 4.6 Nitrate plus nitrite as nitrogen concentrations measured in the Sangamon River at Monticello, Illinois, in May 1990 (Scribner, et al., 1994).**

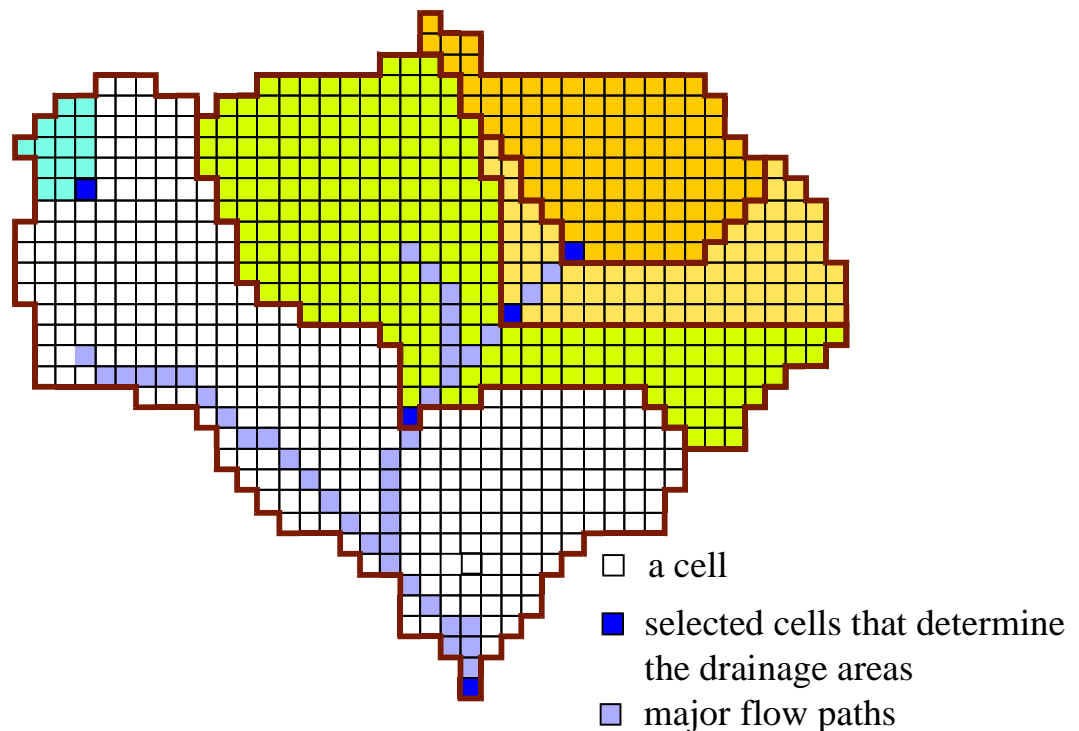
Research performed by Walling and Webb (1984, work cited by Jones and Burt, 1993) shows that in most streams the annual nitrate variations do not have a perfectly symmetrical sinusoidal form; the annual minimum and maximum occur 4-6 weeks later and 2-3 weeks earlier than the timing suggested by a single harmonic. Furthermore, no clear seasonal variations of nitrate concentration exist in catchments where groundwater with a consistently high nitrate concentration mixes with quick flow of varying origin and concentration (Jones and Burt, 1993).

#### **4.3.7 Extracting values of explanatory variables for the regression analysis**

The estimation of values of explanatory variables such as agricultural application rate, total application, average stream slope, average land slope, exponent

of the negative flow length, and average distance from the field to the closest stream is performed in two steps: (1) Grids of spatially distributed parameters are constructed for the Upper Mississippi - Missouri River and the Ohio River basins, and (2) for each cell that represents a sampled watershed outlet the value of the explanatory variable is extracted.

Here, the grid of spatially distributed parameters means a grid which each cell contains average or sum calculated for the total drainage area upstream to the given cell. Figure 4.7 shows selected cells that contain a value that characterize the upstream watershed. This concept is supported by Arc/Info-Grid commands such as `flowaccumulation` and `flowlength`.



**Figure 4.7** Example of a grid of spatially distributed values of explanatory variables.

#### 4.3.8 Application of the regression models

Once their parameters have been determined, the agrichemical transport models (Eq. 4.6) can be used in the following way:

- Estimation of average monthly and average annual concentrations in rivers of the Upper Mississippi-Missouri basin, including the Ohio River basin. The calculations are made using grid-map algebra. The calculations are estimated in all cells ( $9.6 \times 10^6$  cells of size 500 m ) that constitute the basin. The equations that describe chemical concentration in rivers require only a map of distribution of the total agricultural chemical use, the parameters of watersheds that can be easily determined from digital elevation model, and the maps of annual temperature and precipitation depth. Only models that do not utilize the flow rate can be applied. Since the calculation of the concentrations over such a large area is very computationally and computer disk space intensive, it is not presented in this dissertation.
- Calculation of average monthly agrochemical concentrations in the Iowa-Cedar River basin. The map of flow rate is required for obtaining a spatial picture of atrazine and nitrate concentrations in surface waters. The estimation is performed for the watershed subdivided into 1032 subwatersheds of average area  $31.6 \text{ km}^2$  utilizing maps in a vector format. The regression equations are applied to each subwatershed utilizing the data from an attribute table that characterizes the upstream drainage area. A GIS approach to model the concentrations in rivers is discussed in the following Section 4.4. The procedure of the spatial distribution of the historical flow record is presented in Section 4.5.

## **4.4 GIS model description**

### **4.4.1 Subdivision of study region into modeling units**

Elementary watersheds, i.e., modeling units that are considered lumped systems, constitute the smallest units into which a region is subdivided. Each unit is characterized by set of parameters such as area, agricultural chemical application, slope, depth of precipitation, water runoff, average elevation, and an equation that relates the mass of chemical applied and the mass of chemical runoff.

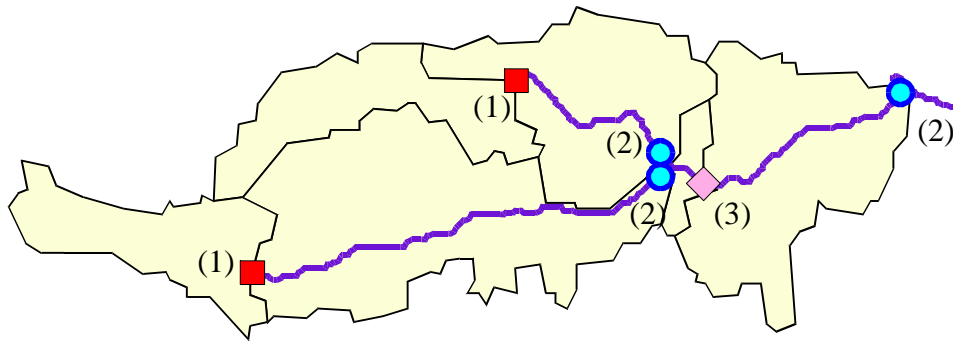
Since the GIS offers very convenient tools for data storage and manipulation, all attributes can be stored and extracted by models that operate on different spatial scales. The order of processing the individual watersheds is the researcher's choice. Moreover, each watershed does not need to be divided into modeling units of the same size. To make the modeling process efficient, such features as density of spatial information and diversity of terrain should influence the assumed size of the elementary unit.

A watershed is explicitly defined by its outlet point. This hydrologic property is utilized here to subdivide the region under investigation into elementary drainage areas. Three types of watershed outlet locations are considered:

- 1) Points in which the drainage area exceeds a threshold value. Streams originate at these points. In this research a threshold value of 25 km<sup>2</sup> drainage area has been assumed;
- 2) Points located immediately upstream of a stream junction; and

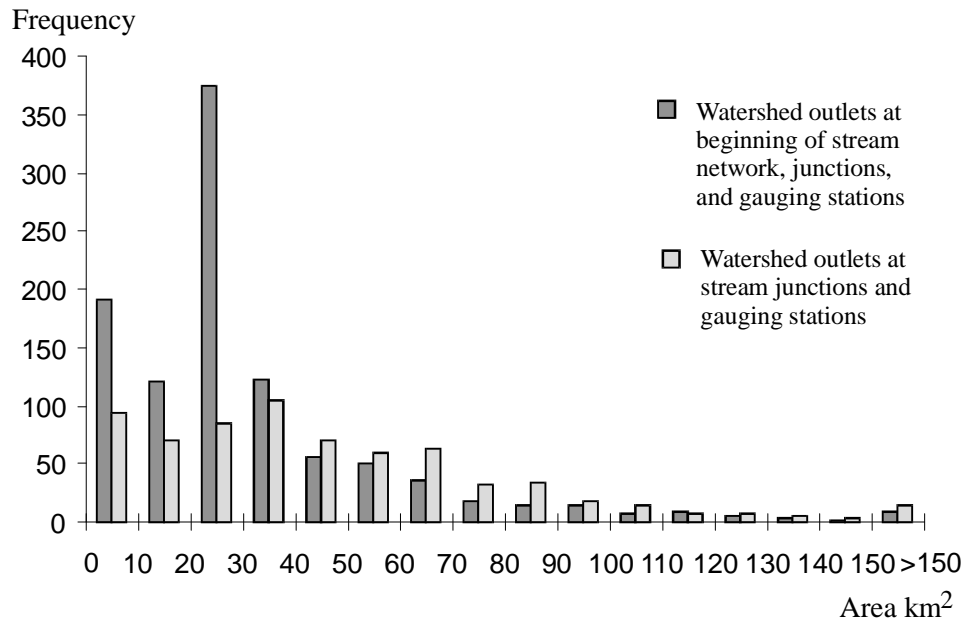
3) Gauging station sites.

Figure 4.8 shows example of selected watershed outlets and corresponding modeling units:



**Figure 4.8 Example of watershed outlets: (1) beginning of the river, (2) stream junction, and (3) gauging station.**

The watershed structure, developed by Maidment (1993) for hydrologic modeling utilizes type (2) and (3) watershed outlets which are positioned at the stream junctions and at the gauging station locations. In this study, this set of watershed outlets has been extended by adding the points in which the stream network, delineated from DEM, begins (type 1 outlet). A test has been performed to determine the influence of additional unit watersheds on the uniformity of region subdivision and the control on the unit area. The Iowa-Cedar River basin has been subdivided using two sets of outlets. Full set: Median unit watershed area = 25.8 km<sup>2</sup> and mean = 31.6 km<sup>2</sup>, reduced set (type 1 outlets not included): median = 37.9 km<sup>2</sup> and mean = 46.7 km<sup>2</sup>. By including the outlets of type 1, the median of unit watershed drainage areas is very close to the threshold drainage area (25 km<sup>2</sup>). Figure 4.9 compares the frequencies of modeling unit area.



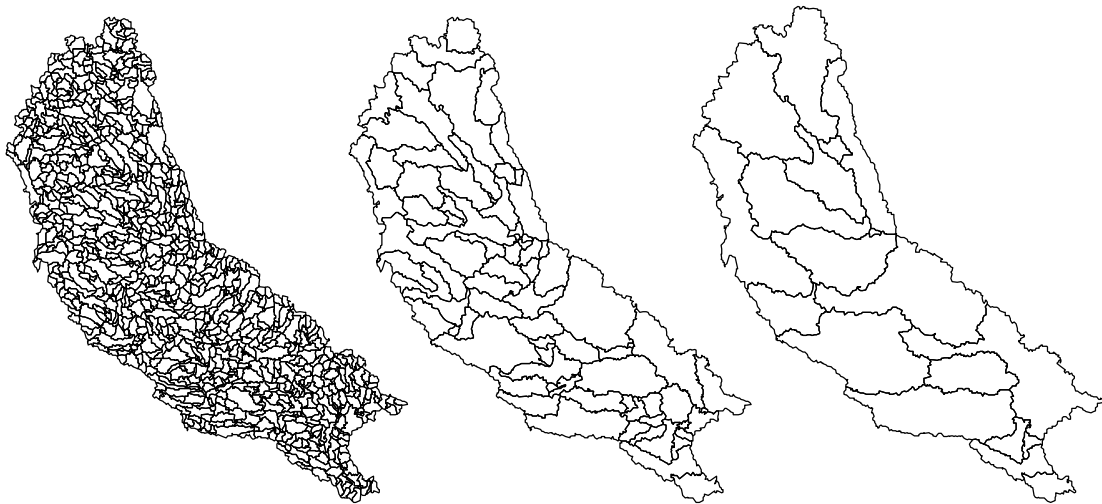
**Figure 4.9 Comparison of the frequency of modeling unit area subdivided using different sets of watershed outlets.**

The following list summarizes the advantages of the addition of type(1) outlets:

- The watershed is subdivided into more uniform (similar area) modeling units and one has more control on the area of units;
- It makes possible the determination of the flow and the constituent load in all nodes of the stream network. Each node has a defined contributing area; and
- The representation of all streams can be standardized. Each reach has input (inflow or load that enters the reach at the upstream end), lateral loss or gain, and output (outflow or load that leaves the reach at its end).

The threshold area of 25 km<sup>2</sup> was selected after many tests with different threshold areas were performed. A smaller threshold area results in very dense stream network and thus large number of areas of size of 1-2 cells. After Arc/Info conversion of these small units from raster format (grid) into vector format many of them disappeared. In addition, smaller than 25 km<sup>2</sup> threshold areas are not justified by the data used in the research. For example, the agricultural application rate is published with a county-size spatial resolution. The stream network delineated from a 3" DEM using 25 km<sup>2</sup> areas, is slightly more dense than the one represented by the RF1 (1:500,000 digital map of rivers).

A larger threshold value than 25 km<sup>2</sup> resulted in very coarse subdivision of studied region and a low density stream network. Figure 4.10 presents the Iowa-Cedar River watershed divided into modeling units of different sizes.

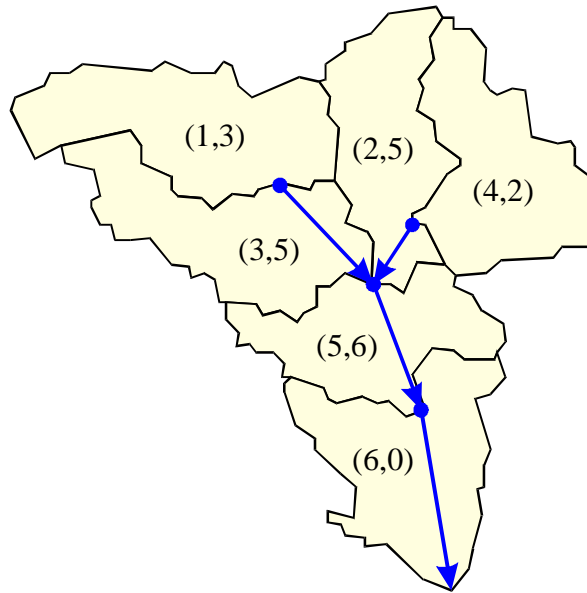


**Figure 4.10** Division of the Iowa-Cedar River into modeling units using different threshold drainage areas: 25 km<sup>2</sup>, 400 km<sup>2</sup>, and 2500 km<sup>2</sup>.



#### 4.4.2 Unit watershed flow system

The network flow system is common function in vector GIS (Maidment, 1993). Here the flow system has been extended: instead of arcs, unit watersheds (polygons) compose the flow system. The flow topology is described by two numbers: the modeling unit ID and the ID of the downstream unit, i.e., the next unit on the flow path. The ID = 0 of the next unit indicates that there are no more downstream units. Figure 4.11 shows an example of the description of the flow topology of the modeling units.



**Figure 4.11** An example of the flow topology of the unit watersheds;  $(x,y)$ ,  $x$  is the unit ID,  $y$  is the ID of downstream to  $x$  unit.

The flow direction indicator is the basic concept in the raster-based hydrologic modeling. For example, Arc/Info Grid denotes the next cell in the flow path by one of eight numbers: 1 represent flow into E (East) neighbor cell, 2 in SE cell, 4 in S cell,

8 in SW cell, 16 in W cell, 32 in NW cell, 64 in N cell, and 128 in NE cell. These numbers are called “flow direction”. Since modeling units are not regular spatial shapes, as the cells are, it is not possible to create an uniform numbering system of the flow direction. The method proposed here describes the flow connectivity by specifying the ID of the next unit on the flow path.

Addition of an item that describes the flow direction into the unit watershed attribute table makes it possible to utilize most of the concepts of hydrologic modeling, so far used in raster GIS, in the vector environment. Such functions as flow accumulation, basin delineation, and flow length can be applied for any shapes and thus for unit drainage areas. The flow system described here may be adopted for all models in which the conditions in a modeling unit do not influence the conditions in an upstream unit such as kinematic wave routing and constituent decay as a pollutant flows downstream.

In this study, the flow connectivity of the unit watersheds is used to calculate cumulative or average parameters of drainage area upstream of a given point. These values are applied to the regression equation to estimate agricultural chemical concentration in the runoff. The following list presents selected examples of application of the unit watershed flow system that have been programmed in the Arc/View script language:

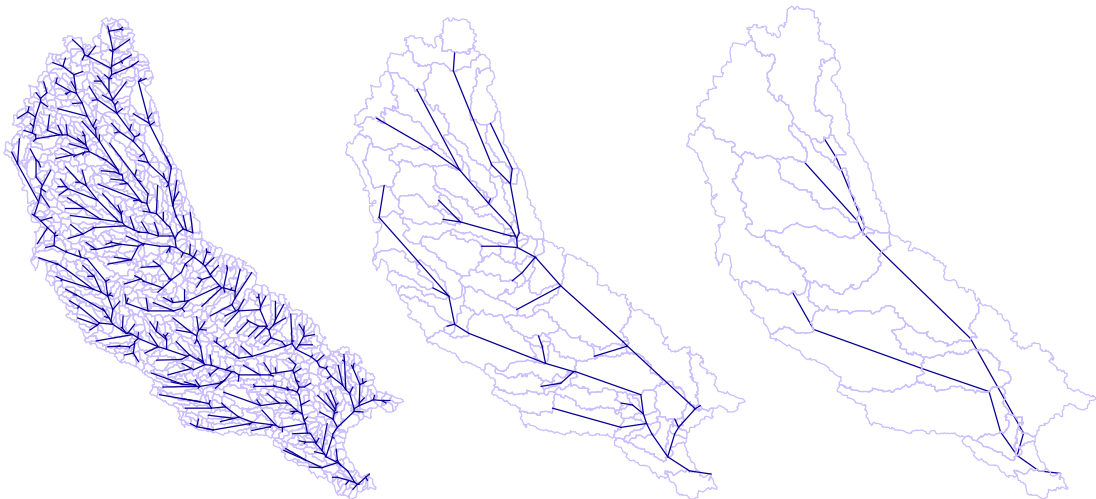
- weighted average of a feature (for example agricultural application ) for the total drainage area upstream of each (or selected) modeling unit (like the Grid function `zonalaverage`);
- accumulated value over total upstream drainage area going along the flow path (vector version of the Grid `flowaccumulation` function); and

- difference between the inputs and the output for each modeling unit, flow difference or “flow-decummulation”.

The unit watershed approach of modeling agricultural transport can be considered as a semantic data model:

*... semantic data modeling --- creating abstractions of geographic data layers which map one to one with their geographic representation but which are simplified in a functional description to the level needed for hydrologic modelling. (Maidment, 1993)*

Figure 4.12 shows examples of the conceptual stream network. The links have been developed by connecting the cells immediately below each modeling unit outlet. Thus, although the system is conceptual, each node is located on the river represented in Grid format.



**Figure 4.12** Flow system in the Iowa-Cedar River basin subdivided into modeling units of different sizes (threshold drainage area: 25 km<sup>2</sup>, 400 km<sup>2</sup>, and 2500 km<sup>2</sup>).

In this system, the outflow from the source watershed (first order watershed) is a point inflow into the stream. The outflow from an intermediate watershed can be represented as a lateral inflow applied on the stream length or can contribute directly to the outflow from the reach, as applied in this research. Thus the conceptual stream network attribute table contains three items that describe the water or agrichemical mass flow conditions:

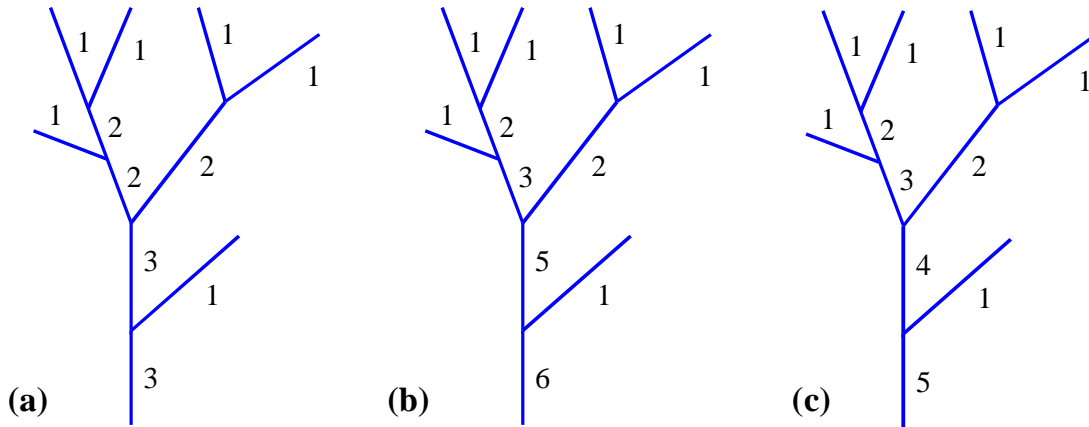
- Inflow (sum of inflows from upstream units);
- Lateral inflow; and
- Outflow.

In addition, the arc attribute table can contain wide range of items that describe the links (length, slope, RF1-ID, agrichemical decay coefficient, travel time) as well as the parameters of the beginning node and the ending node (elevation, coordinates, and length from the watershed outlet). In this research only the conceptual stream network represented by unit watersheds is utilized. Since the unified system of hydrologic unit IDs is developed in this research, the attributes of unit watersheds and the attributes of stream reaches can be linked, i.e., stream parameters such as actual flow length or stream slope can be attributed to modeling units.

#### **4.4.3 Ordering system of the modeling units**

To enhance the calculations performed by the GIS model, the following system of ordering modeling units has been utilized: The most upstream units are assigned an order one. An interior unit has the order equal to the maximum order of the upstream units increased by one. Figure 4.13 compares this ordering method with the Strahler

and the Shreve ordering systems (ESRI, 1992). The proposed ordering system allows one to perform calculations in consecutive manner: initially all the first order units are processed, then the order is increased by one and all units of order two are evaluated. The routing is performed until the unit of the highest order is calculated.



**Figure 4.13 Comparison of the stream ordering systems: (a) Strahler, (b) Shreve, and (c) utilized in the agricultural transport model.**

#### 4.4.4 Enhancement of the stream delineation process

The grid that describes the cell-to-cell flow (flowdirection) is crucial for all hydrologic analysis that is performed in a rasterized environment. Procedures such as stream and watershed boundary delineation, dividing a basin into modeling units, stream slope calculation, length of the flow path estimation, and connection of hydrologic units, are examples of operations that cannot be performed without the map of flow direction. Moreover, the accuracy of all derived information depends on the precision of the flow direction grid. Therefore, an effort has been made to develop a method to improve the map that represents the flow paths. The RF1-digital 1:500,000 map of the US rivers has been selected as a basis for the spatial framework

of the entire flow system. The following explanations support the application of a map of existing rivers to correct the flow system determined from a DEM:

- Since the location of a stream that is delineated from elevation data depends on the cell size, the stream networks determined from the DEMs of different resolutions are not compatible. Thus, the gauging stations linked to one gridded river system will not be in agreement with other gridded river systems derived from DEM grids of different cell sizes.
- The stream system constitutes the best framework for the spatial flow. It took hundreds of thousands of years for a river bed to develop to its current form. Since the river location practically does not change, other information such as position of gauging stations may be related to the location of the stream reach.
- In flat regions, the streams delineated from the DEM tend to be straight lines, whereas, in reality the rivers have a tendency to meander. This causes an overestimation of the stream slope and an underestimation of the river length in purely DEM-derived streams.
- RF1 represents the true river system, whereas, the stream network delineated from the DEM just approximates the same system.

The process of the DEM adjustment is based on the converting the RF1 into grid form and then increasing the elevation of all DEM cells, that do not represent gridded RF1, by an arbitrary value (e.g., 10,000 m). This operation forces the Grid GIS to create a map of flow direction that is compatible with the flow system represented by the vector map of rivers (RF1). The method of enhancing the flow system development has the following disadvantages:

- Since the elevations of the DEM are changed, the modified DEM can not be used for other tasks than the flow direction estimation.
- The stream network has to be represented by a single line. A double line description of rivers can be utilized if the distance between the lines is smaller than the cell size.
- All existing loops in the river network have to be removed or opened to eliminate the ambiguous flow paths.
- All lakes have to be converted into line representations or to polygons.
- The cell width applied for the adjustment process should be smaller than half of the distance between any streams in RF1, to avoid connections of stream networks from different basins that may be created when converting from vector format into raster format.

If the river network does not fulfill the above mentioned requirements, some editing after converting into grid format is necessary.

Incorporating the RF1 into a grid has an additional advantage. By assigning the reach ID from RF1 to the raster river representation, the attribute table of RF1 can be linked with the attribute table of derived grids. Thus, such information as the average flow velocity, average flow rate or stream names that are in the RF1 attribute table can be used for grid models and vice versa, the parameters estimated in grid such as reach slope, drainage area and flow length can be assigned to the streams in vector RF1. The RF1 - Grid link extends the grid-network procedure for hydrologic modeling developed by Maidment (1992).

## **4.5 *Redistribution of the flow record over ungauged rivers***

The flow rate is essential to estimate the agrichemical concentration and load in all rivers of the area under investigation. In this study, the historical flow record is utilized instead of synthetic values. The high or low flow conditions are modeled by selecting a year from the past that had the high or low flow rate recorded. The flow measurements are available only in locations in which the USGS gauging stations are located. Therefore, a procedure that calculates the flow rate in ungauged stations has been developed. This section describes this procedure.

### **4.5.1 GIS database of monthly flow rate and the precipitation depth**

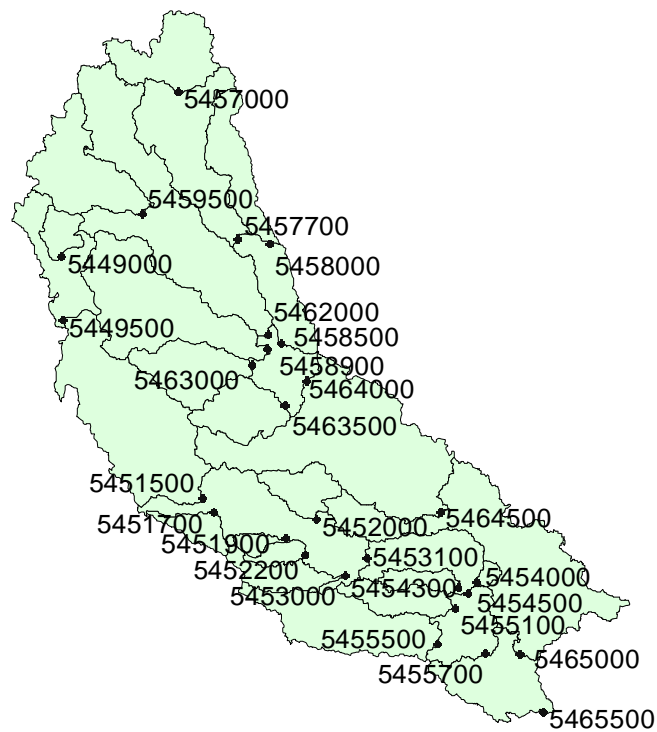
To make the model of transport of agrichemicals capable of reconstructing historical conditions, a database of the recorded average monthly flow rate and a database of the observed average monthly precipitation depth must be constructed. Since the model is developed within the GIS, the data structure of the flow and precipitation time series must be incorporated into the geographic system. The advantages of this approach are as follows:

- Compactness of the system. Spatial and temporal data are stored in one format that is specified by the GIS;
- Efficiency of the system. The spatio-temporal features can be viewed, queried, and processed directly by the procedures built into such GIS software as Arc/Info or ArcView;
- Easy to maintain. Data can be organized by spatial units (e.g., political or hydrologic units) and/or by temporal unit (e.g., year, month or decade);



- Simple to understand. Data are stored in attribute tables, records represent the spatial domain whereas the columns (items) represent the time domain.

The GIS database of monthly flow rate and the database of the average monthly precipitation depth is developed in two steps: first a map of monitoring stations is created and then, for each map, the attribute table with the measurements is built.

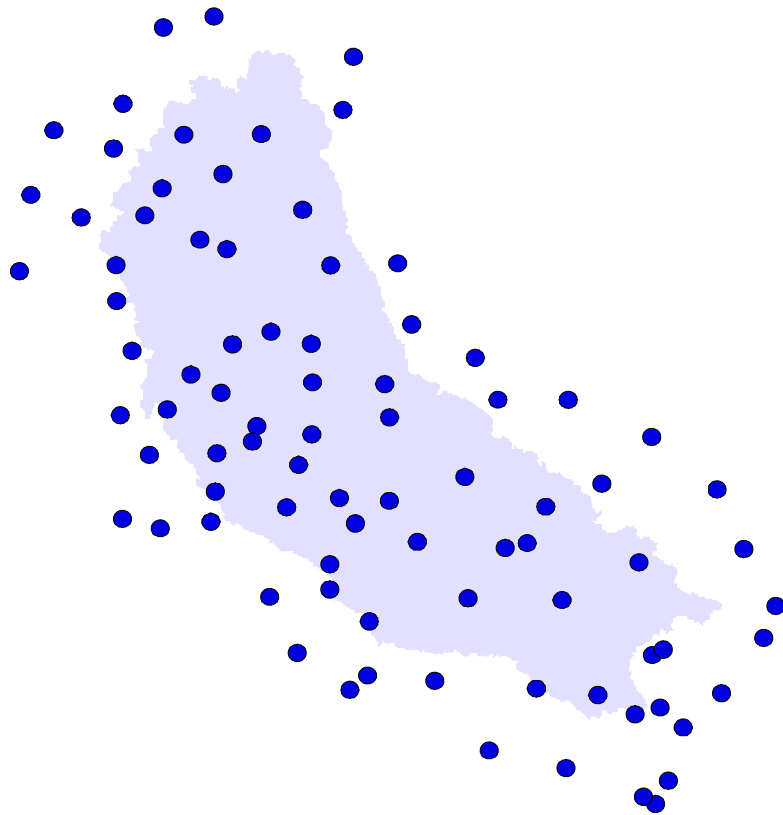


**Figure 4.14 The Iowa-Cedar River basin: subwatersheds and selected USGS gauging stations (numbers represent station ID).**

The gauging station maps are created utilizing the latitude and longitude of gauging sites published by Hydrosphere (1993 a, b; 1994). Although there are 38

USGS stations in the Iowa-Cedar River basin only the 28 stations with complete flow record for period from 1960 to 1992 have been used for analysis. Figure 4.14 shows the map of the USGS gauging stations selected for modeling the spatial distribution of recorded flow rate.

The map of National Weather Service Climate stations contains 86 stations that are located within the Iowa-Cedar River basin and within the 50-km buffer zone outside the basin. Figure 4.15 shows the map of the climate stations. All climate stations are utilized to create precipitation maps.



**Figure 4.15 Weather stations applied to analysis of the hydrologic conditions in the Iowa-Cedar River watershed.**

A fragment of the point attribute table that contains average monthly flow rates is presented in Table 4.1.

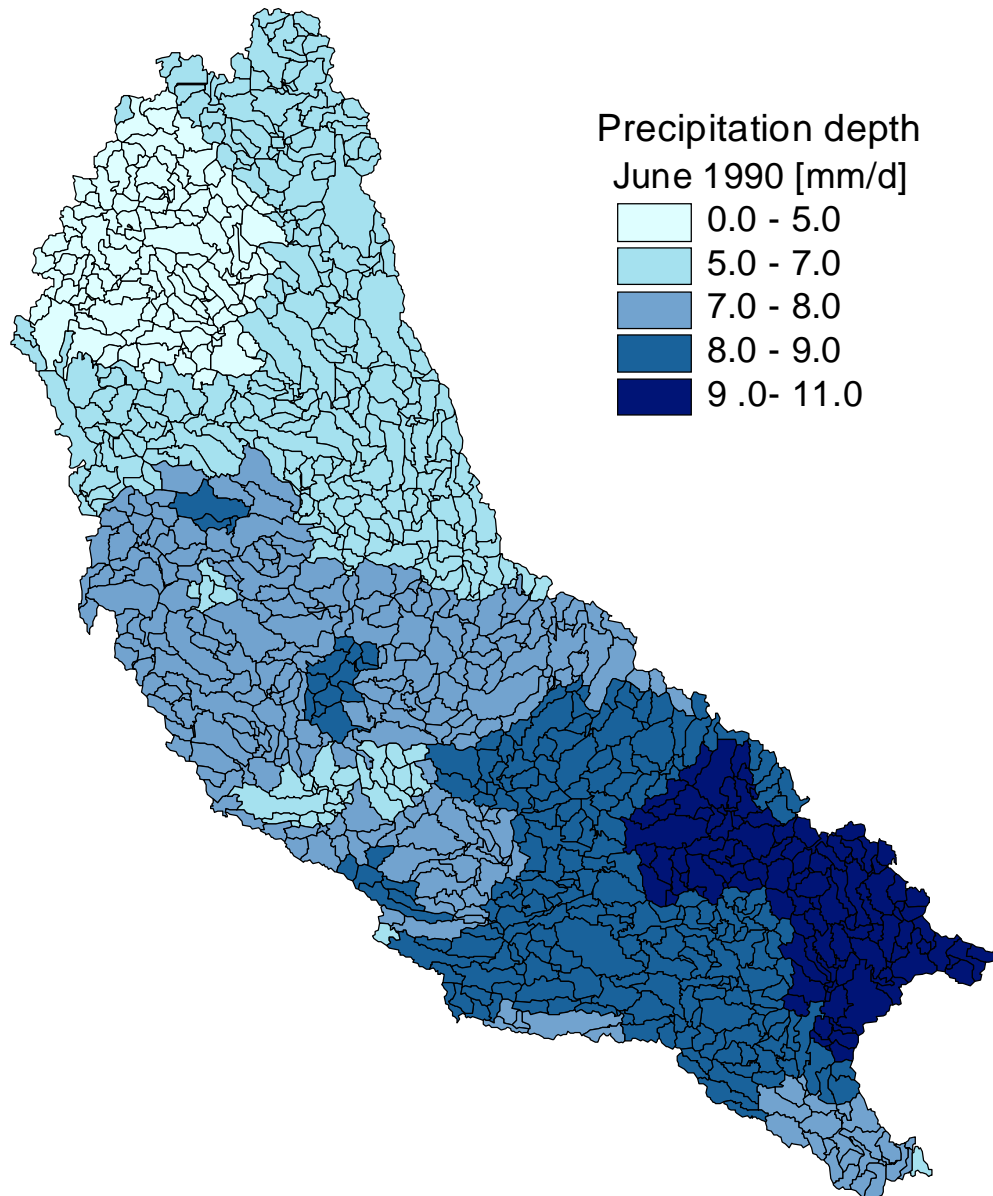
**Table 4.1** An example of PAT--point attribute table of the USGS gauging station coverage (e.g., item M199001 contains average flow rate in year 1990, month 01). Full table contains monthly flow record for 38 stations, for period from 1940 to 1992 in m<sup>3</sup>/s.

GSFLOW	GSFLOW_ID	STATION_ID	M199001	M199002	M199003	M199004	M199005
1	1	5457000	1.339	1.419	13.111	15.036	18.519
2	3	5459500	0.292	0.357	3.228	2.274	6.513
3	4	5457700	2.444	3.596	22.201	23.390	31.630
4	5	5458000	0.165	0.294	5.239	3.143	4.955
5	6	5449000	0.004	0.020	0.697	0.413	0.445
6	9	5449500	0.154	0.204	1.481	1.141	2.917
7	10	5462000	1.444	1.855	15.150	12.601	26.618

#### 4.5.2 Average precipitation depth in modeling units

The process of spatial redistribution of the measured monthly average flow requires the average precipitation depth for each modeling unit. GIS water quality models such as SWAT-GRASS utilize the rainfall depth observed in the closest weather station to the subbasin (Ramanarayan, et al., 1996, Krysanova, et al., 1996). The methods commonly used in hydrology for spatial estimation of rainfall from rain gauges are the Thiessen polygon method (Chow, et al., 1988) and the inverse distance-squared method (Smith, 1993). There are two other functions available in Arc/Info GIS: kriging and trend (fitting a polynomial regression surface). In this research the Arc/Info IDW function (inverse distance-squared method) has been applied. It creates a grid of spatially distributed values extracted from the attribute table of point

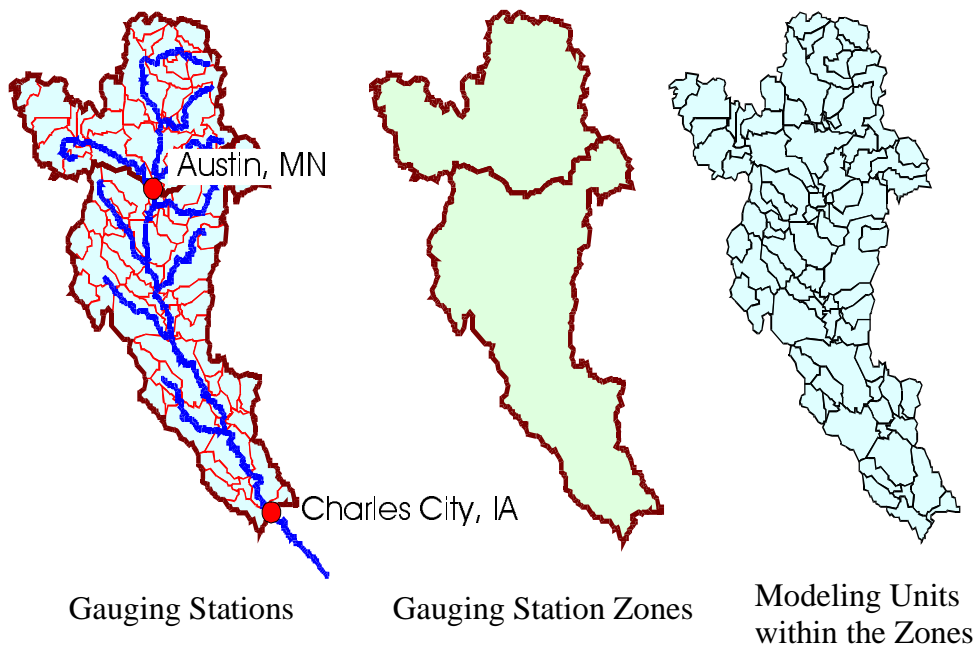
coverage. The calculations are very fast, thus it was feasible to create maps of spatial distribution of monthly average precipitation depth for the Iowa-Cedar River basin for the period from 1950 to 1992. Figure 4.16 presents the map of precipitation depth estimated for June 1990.



**Figure 4.16 Spatial distribution of monthly precipitation depth [mm/d] in the Iowa-Cedar River watershed in June 1990.**

### 4.5.3 Mathematical description

There are about 38 USGS gauging stations in the Iowa-Cedar River basin (some of them are not in service). One or more gauging stations determine a drainage area (or partial drainage area), referred to as a gauging station zone or zone. The gauging station constitutes a point through which a known amount of water flows from one zone to an other zone. The runoff from each modeling unit depends on two factors: 1) the water balance calculated for the unit's respective zone and 2) the distribution of the precipitation depth over the gauging zone. Figure 4.17 illustrates an example of a gauging zone and modeling units.



**Figure 4.17 Cedar River watershed above Charles City, Iowa: An example of gauging station zones and modeling units.**

The process of spatial redistribution of the measured monthly average flow is performed in four steps. Each step is performed separately for a given month. Since the method only redistributes observed values using drainage area and precipitation as a weight, the effect of the snow accumulation and the evaporation on the spatial distribution of the flow is included in the flow record for the gauging zones. Within each zone neither the snow accumulation nor evaporation or groundwater transfer are explicitly taken into consideration but they are considered implicitly by using varying runoff coefficient in each month of each year. The precision of data used for the agricultural transport such as county level chemical application do not justify construction of a very detailed flow model that would require information about spatially distributed temperature, snow depth, land use, aspect, solar radiation etc. The steps of the flow rate interpolation are as follows:

- 1) Estimation of an average runoff coefficient relating monthly precipitation to discharge;
- 2) Approximation of the flow in streams;
- 3) Evaluation of error between estimated flow and gauged flow;
- 4) Correction of estimated flow rate at the gauge and over all modeling units within gauged zone.

***Estimation of average runoff coefficient.*** Using recorded outflow from the first order zones, i.e. watersheds determined by the most upstream gauging stations, an average runoff coefficient  $C_f$  is calculated according to the following equation:

$$C_f(m) = \frac{\sum_j Q_j(m)}{a \sum_j A_j(m) * P_j(m)} \quad (4.16)$$

where:

- $C_f(m)$  = average runoff coefficient for the  $m$ -th month [dimensionless]
- $Q_j(m)$  = average flow recorded in  $j$ -th gauging station during  $m$ -th month [ $\text{m}^3/\text{s}$ ]
- $P_j(m)$  = average monthly precipitation depth [mm/d]
- $A_j(m)$  =  $j$ -th watershed area [ $\text{km}^2$ ]
- $m$  = month
- $j$  = index of the first order watersheds
- $a$  = units conversion factor.

**First approximation of the flow in streams.** The runoff from the modeling units is summed along the flow path, moving downstream from a first order stream toward the outlet of the basin. The mass balance for the  $i$ -th unit is described by the following formula:

$$Q'_i(m) = \sum_k Q'_k(m) + aC_f(m)P_i(m)A_i(m) \quad (4.18)$$

where:

- $Q'_i(m)$  = estimated cumulative flow at the outlet point of the  $i$ -th unit [ $\text{m}^3/\text{s}$ ]
- $Q'_k(m)$  = estimated cumulative flow at the outlet point of the  $k$ -th unit [ $\text{m}^3/\text{s}$ ]
- $k$  = index of units in the immediate upstream vicinity of the  $i$ -th unit
- $C_f(m), P_i(m), A_i(m)$ , = average runoff coefficient, average monthly precipitation depth in unit  $i$ , and  $i$ -th unit area, respectively

To make calculations more efficient, when the flow path crosses the border of the gauging station zone, i.e. at a gauging station, the calculated cumulative flow  $\sum Q'_k(m)$  is substituted by the measured value  $Q_f(m)$ . This substitution of values ensures

that the observed inflow is used to calculate accumulated flow in each zone, and the resulting error at the zone outlet is due only to inaccuracy of water balance estimated within the zone (errors do not propagate from zone to zone). Thus, the water balance in unit downstream to the gauging station is:

$$Q'_i(m) = Q_j(m) + aC_f(m)P_i(m)A_i(m) \quad (4.18)$$

where:

$Q'_i(m), C_f(m), P_i(m), A_i(m)$ , = same as for (Eq. 4.17)

$Q_j(m)$  = observed inflow into modeling unit  $i$ , from zone  $j$ . This value equals to the outflow from zone  $j$  [ $\text{m}^3/\text{s}$ ]

**Error evaluation.** In this step, the difference between estimated and observed flow,  $\Delta Q_j(m)$ , is calculated for each zone  $j$ :

$$\Delta Q_j(m) = Q_j(m) - Q'_i(m) \quad (4.19)$$

where:

$Q_j(m)$  = observed outflow from zone  $j$

$Q'_i(m)$  = estimated cumulative flow at the outlet point of the  $i$ -th unit that is also the outlet point of the  $j$ -th gauging station zone.

**Correction of the cumulative flow.** The correction of the estimated cumulative flow is the crucial process of the flow distribution method. Two weighting coefficients  $k_1$  and  $k_2$  are applied. The coefficient  $k_2$  redistributes error according to the cumulative runoff calculated for each zone separately - the flow in rivers is created only by the estimated runoff  $C_f P(m) A(m)$ . There are no inflows from the upstream zones. Thus the coefficient  $k_2$  redistributes error according to the drainage area and the



precipitation depth within each zone separately. It is calculated for each modeling unit as a proportion of the runoff  $Q_i^x(m)$  from the drainage area upstream of the given,  $i$ -th unit outlet (within the zone  $j$ ) to the total runoff  $Q_j^x(m)$  from the zone  $j$  (at the beginning of the flow path it is very small, at the zone outlet it equals one):

$$k_{2i} = \frac{Q_i^x(m)}{Q_j^x(m)} \quad (4.20)$$

To put more burden of the error correction on the major rivers rather than small streams, the coefficient  $k_l$  has been introduced. There is a higher probability that the estimated error is due to the losses/gains of the river between the two gauging stations than to the losses/gains in small streams located far from measurement points. In another words, the further from the gauged reach a stream is located, the higher is the uncertainty of the flow is and therefore, the more appropriate it is to apply the average basin runoff coefficient to estimate the stream flow rate.

The coefficient  $k_l$  is calculated using total cumulative flow, i.e., the approximated flow in the streams of the whole basin as a proportion of the estimated cumulative flow  $Q'_i(m)$  at the outlet point of the  $i$ -th modeling unit to the cumulative flow  $Q'_j(m)$  at the outlet of zone  $j$  in which the  $i$ -th unit is located:

$$k_{1i} = \frac{Q'_i(m)}{Q'_j(m)} \quad (4.21)$$

The adjustment of estimated flow (Eq. 4.18 and Eq. 4.19) can be represented by the following formula:

$$Q''_i(m) = Q'_i(m) + k_{1i}k_{2i}\Delta Q_j(m) \quad (4.22)$$

where:

$Q''_i(m)$  = corrected cumulative flow at the outlet point of the  $i$ -th modeling unit;

$Q'_i(m)$  = estimated cumulative flow at the outlet point of the  $i$ -th modeling unit;  
 $\Delta Q_j(m)$  = error, difference between observed and estimated cumulative flow at the outlet of zone  $j$ ;  
 $k_1, k_2$  = weighting coefficients.

#### 4.5.4 Example of flow redistribution

This section shows an example of the redistribution of the recorded flow rate in the gauged zone 6565500 in which the Iowa River joins the Cedar River for June 1990. The location of this zone is shown in Figure 4.18.

The average runoff coefficient for the June 1990 is calculated using recorded outflow and redistributed measured precipitation depth in the first order gauged zones of the Iowa-Cedar basin from Eq. (4.17):

$$C_f(m) = \frac{345.4}{891.6} = 0.387$$

where  $m$  indicates June 1990,  $345.4 \text{ m}^3/\text{s}$  is the cumulative discharge from the first order basins, and  $891.6 \text{ m}^3/\text{s}$  is the product of monthly precipitation and drainage area over all first order basins.

The outflow from the zone, station 5465500 (the Iowa River at Wapello), equals to  $827.621 \text{ m}^3/\text{s}$ . The measured inflows at stations 5455700 (the Iowa River near Lone Tree) and 5465000 the Cedar River near Conesville) are  $285.039 \text{ m}^3/\text{s}$  and  $491.895 \text{ m}^3/\text{s}$  respectively.



**Figure 4.18** Location of the gauged zone 5465500, its subdivision into modeling units, and the USGS gauging stations.

The estimated outflow from the zone is equal to the sum of inflows from upstream zones plus the runoff from zone due to the precipitation ( $40.034 \text{ m}^3/\text{s}$ ):

$$\begin{aligned}
 Q'_{5465500}(m) &= Q_{5455700}(m) + Q_{5465000}(m) + aC_f(m)P_{5465500}(m)A_{5465500}(m) \\
 &= 285.039 + 491.895 + 40.034 = 816.968 \text{ m}^3 / \text{s}
 \end{aligned}$$

The error of estimated flow is calculated according to the equation (4.19):

$$\Delta Q_{5465500}(m) = 827.621 - 816.968 = 10.653 \quad m^3 / s$$

The correction coefficients  $k_{1i}$  and  $k_{2i}$  are calculated according to the equations (4.20) and (4.21) respectively.  $k_{1i}$  is calculated by dividing column  $Q_i'(m)$  by  $Q_j'(m) = 816.968$   $m^3/s$  and  $k_{2i}$  is calculated by dividing column  $Q_i^x(m)$  by  $Q_j^x(m) = 40.03$   $m^3/s$ .

**Table 4.2. Steps of the recorded flow rate redistribution in gauged zone 5465500 for June 1990.**

Unit ID	Next unit ID	Order	Area km <sup>2</sup>	P mm/d	$C_f P_i A_i$ m <sup>3</sup> /s	$Q_i^x(m)$ m <sup>3</sup> /s	$Q_i'(m)$ m <sup>3</sup> /s	$k_{2i}$ -	$k_{1i}$ -	$Q_i''(m)$ m <sup>3</sup> /s
100644	644	1	29.29	9.1	1.198	1.198	1.198	0.0299	0.0015	1.198
644	675	2	69.53	8.9	2.773	3.970	3.970	0.0992	0.0049	3.975
100666	666	1	25.96	8.8	1.027	1.027	1.027	0.0256	0.0013	1.027
663	671	58	28.22	8.5	1.076	1.076	286.115	0.0269	0.3502	286.216
657	665	74	11.59	9.1	0.475	0.475	492.370	0.0119	0.6027	492.446
665	680	75	64.89	8.5	2.469	4.066	495.961	0.1016	0.6071	496.618
100664	664	1	25.03	9.5	1.065	1.065	1.065	0.0266	0.0013	1.065
100668	668	1	30.68	8.3	1.139	1.139	1.139	0.0285	0.0014	1.139
668	667	2	19.56	8.3	0.727	1.866	1.866	0.0466	0.0023	1.867
666	671	2	10.85	8.4	0.409	1.436	1.436	0.0359	0.0018	1.437
667	671	3	36.08	8.2	1.333	4.192	4.192	0.1047	0.0051	4.197
664	665	2	1.39	9.2	0.057	1.122	1.122	0.0280	0.0014	1.123
100673	673	1	25.08	8.3	0.933	0.933	0.933	0.0233	0.0011	0.933
100676	676	1	26.31	8.3	0.977	0.977	0.977	0.0244	0.0012	0.978
671	675	59	20.17	8.2	0.738	7.442	292.481	0.1859	0.3580	293.190
670	667	2	1.41	8.2	0.052	0.993	0.993	0.0248	0.0012	0.993
673	681	2	64.83	8.2	2.379	3.312	3.312	0.0827	0.0041	3.315
100672	672	1	25.09	9.0	1.008	1.008	1.008	0.0252	0.0012	1.008
100670	670	1	25.66	8.2	0.941	0.941	0.941	0.0235	0.0012	0.941
100678	679	1	29.79	8.1	1.087	1.087	1.087	0.0271	0.0013	1.087
100674	674	1	30.8	8.0	1.110	1.110	1.110	0.0277	0.0014	1.110
675	677	60	6.4	7.9	0.227	11.639	296.679	0.2907	0.3631	297.803
674	677	2	42.37	7.8	1.478	2.587	2.587	0.0646	0.0032	2.590
672	679	2	47.07	8.4	1.764	2.772	2.772	0.0692	0.0034	2.775
676	681	2	59.99	8.2	2.216	3.194	3.194	0.0798	0.0039	3.197
681	685	3	73.53	7.8	2.565	9.071	9.071	0.2266	0.0111	9.097
677	680	61	1.37	7.7	0.047	14.274	299.313	0.3566	0.3664	300.705
680	682	76	27.33	7.7	0.949	19.289	796.223	0.4818	0.9746	801.225
5465500	684	78	62.05	7.8	2.173	40.034	816.968	1.0000	1.0000	827.621
100683	683	1	25.13	8.0	0.904	0.904	0.904	0.0226	0.0011	0.904
679	5465500	3	11.98	7.9	0.427	4.285	4.285	0.1070	0.0052	4.291
685	682	4	57.49	7.8	2.001	14.168	14.168	0.3539	0.0173	14.234
682	5465500	77	3.36	7.9	0.118	33.576	810.510	0.8387	0.9921	819.373
683	685	2	63.18	7.7	2.193	3.096	3.096	0.0773	0.0038	3.100

For example, consider modeling unit number 663 which is located on the Iowa River just downstream of the Lone Tree gauge (station 5455700). The January 1990 precipitation on this model unit was 8.5 mm/d on average, thus with a drainage area of 28.22 km<sup>2</sup> and a runoff coefficient  $C_f$  of 0.3874, the runoff from this modeling unit is  $28.22 * 8.5 * 0.3874 = 92.926 \text{ km}^2\text{mm/d} = 92.926/86.400 \text{ m}^3/\text{s} = 1.076 \text{ m}^3/\text{s}$ .

The outflow from unit 663 is a sum of inflow and estimated outflow:

$$Q_i'(m) = 285.039 + 1.076 = 286.115 \text{ m}^3/\text{s}.$$

The correction coefficients are:

$$k_1 = 286.115/816.968 = 0.3502, \quad \text{and} \quad k_2 = 1.076/40.034 = 0.0269,$$

thus, the adjusted flow in the Iowa River at the outlet point of the unit 663 is:

$$Q_i''(m) = 286.115 + 0.3502 * 0.0269 * 10.653 = 286.216 \text{ m}^3/\text{s}.$$

#### **4.6 Exponential decay model**

This section discusses a model which estimates loads in rivers assuming the chemical losses in rivers are governed by a first order reaction, i.e., the agricultural mass exponentially decays as it travels from one modeling unit to the next downstream unit. The preliminary model has been developed to calculate the concentrations in rivers utilizing the results of a CEEPS (Comprehensive Environmental Economic Policy Evaluations System) metamodel developed by the Iowa State University's Center for Agricultural and Rural Development (Bouzaher and Monale, 1993).

The CEEPS applies two models of chemical runoff from the field: PRZM (Mullins, et al., 1993) for pesticides and EPIC for nutrients. These results specify the proportion of agrichemical mass applied that leaves the field as runoff. Unfortunately, the author of this dissertation was unable to obtain the spatial distribution of these chemical application-loss relations over the Iowa-Cedar river to estimate the rate of agrichemical losses in rivers necessary to develop and test the complete GIS model with CEEPS data as input. Since it is not in the scope of the research to develop such a CEEPS model, only its outline for the completeness of discussion, is included below.

#### 4.6.1 Exponential decay model overview

The spatial model frame is based on the watershed divided into modeling units described in previous sections. The mass that enters the stream in a unit watershed is calculated using export factors:

$$M_i = U_i Z_i \quad (4.23)$$

where:

$M_i$  = agrichemical mass that enters surface water in the  $i$ -th modeling unit;

$U_i$  = total agrichemical application in the  $i$ -th unit; and

$Z_i$  = export factor that depends, e.g., on agricultural management practices.

This factor also includes all chemical losses in  $i$ -th unit.

As the agrichemical mass load travels along the flow path it decays. The decay process can be related to travel time or to travel distance. It is assumed here that the time decay coefficient or distance decay coefficient represent such processes as decay,

sorption, desorption, volatilization, exchange between sediment and water column, settling, scour, and dilution. The equations (4.24 and 4.25) describe this process utilizing a travel time and a time decay coefficient:

$$M_{in,i} = \sum_j M_{out,j} \quad (4.24)$$

and

$$M_{out,i} = M_{in,i} \exp(-k_{T,i} t_i) + M_i \quad (4.25)$$

where:

- $M_{in,i}$  = agrichemical mass that enters  $i$ -th modeling unit;
- $M_{out,j}$  = output mass from the  $j$ -th unit;
- $j$  = indicator of units that are immediate upstream of the unit  $i$ ;
- $M_{out,i}$  = agrichemical mass that leaves  $i$ -th modeling unit;
- $M_i$  = agrichemical mass in runoff from  $i$ -th unit;
- $k_{T,i}$  = time loss coefficient, applied to the mass that enters  $i$ -th unit;
- $t_i$  = time that takes the constituent to travel through  $i$ -th unit.

To describe the agrichemical losses by an exponential function of river reach length the  $k_{T,i} t_i$  component of the equation (4.25) must be substituted by a distance loss coefficient  $k_{L,i}$  and a travel distance  $L_i$  (Eq. 4.26).

$$M_{out,i} = M_{in,i} \exp(-k_{L,i} L_i) + M_i \quad (4.26)$$

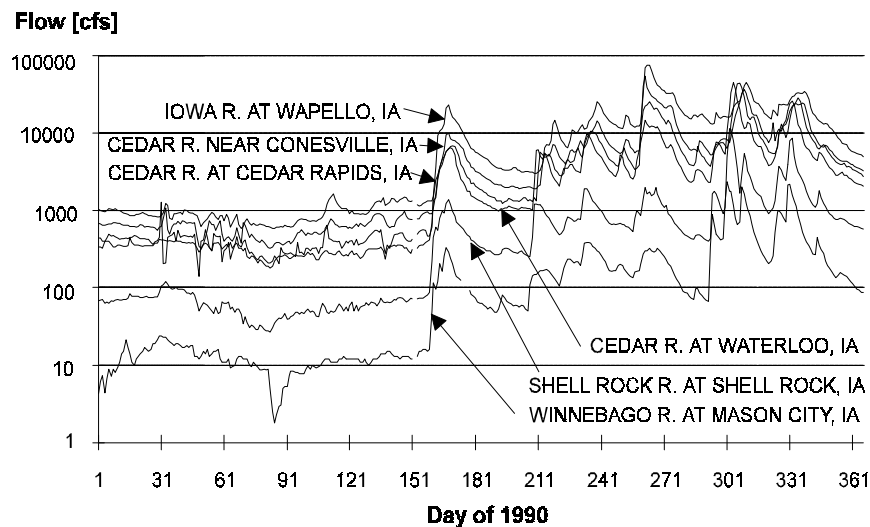
where:

- $k_{L,i}$  = distance loss coefficient, applied to the mass that enters  $i$ -th unit;
- $L_i$  = travel distance through  $i$ -th unit.

Since the database of the average monthly precipitation depth, the monthly water runoff from the field as well as the average monthly flow rate in river for all modeling units it is feasible to apply concentrations (traditional approach:  $c_1 = c_0^{-k_T t}$ ) instead of loads.

#### 4.6.2 Travel time approximation

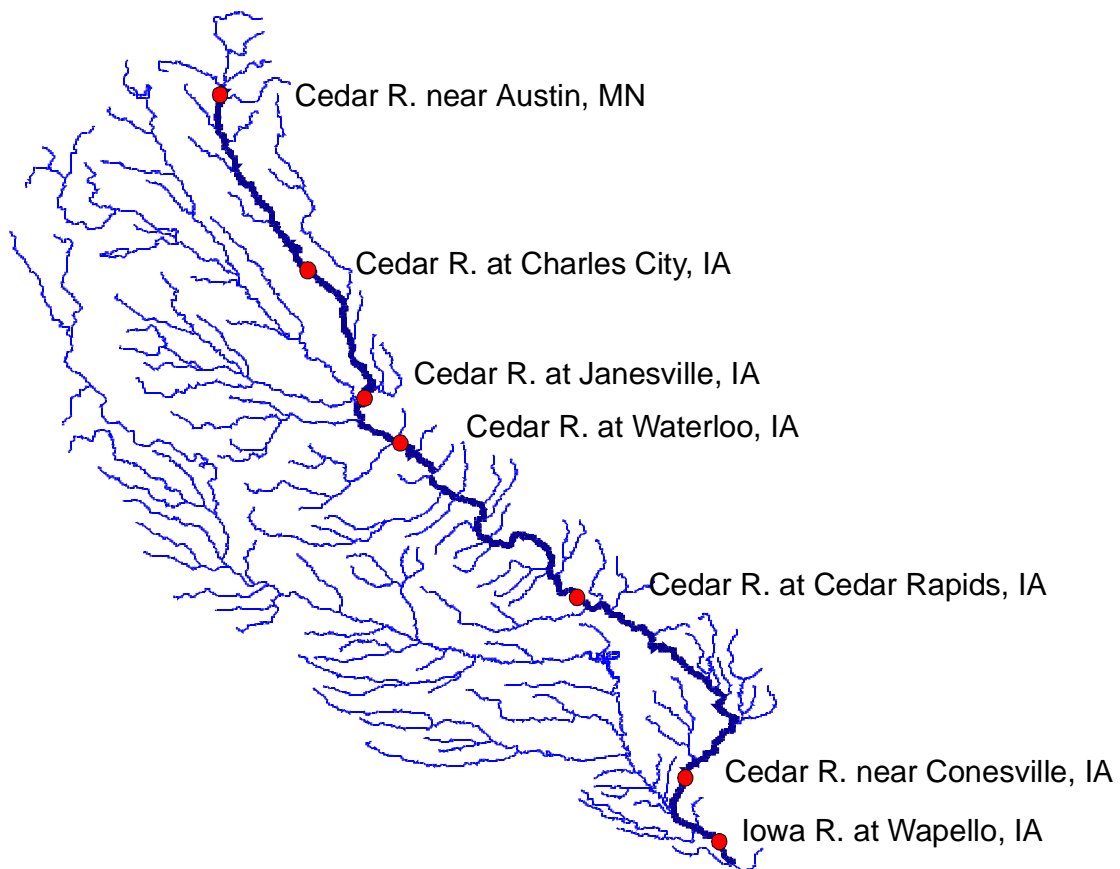
Since no information about the river cross sections is available as yet, a special procedure has been developed to determine the distribution of time of travel and flow velocity. The visual analysis of daily flow record along the flow path has revealed that there is a time shift between flow time series. Figure 4.19 presents an example of the flow rate time series along the flow path from the gauging station "Winnebago River at Mason City, Iowa" to the gauging station "Iowa River at Wapello, Iowa".



**Figure 4.19** Example of time series recorded in gauging stations located along the Cedar River. The data are for the water year 1990. The logarithmic scale is used to show all time series in one picture.



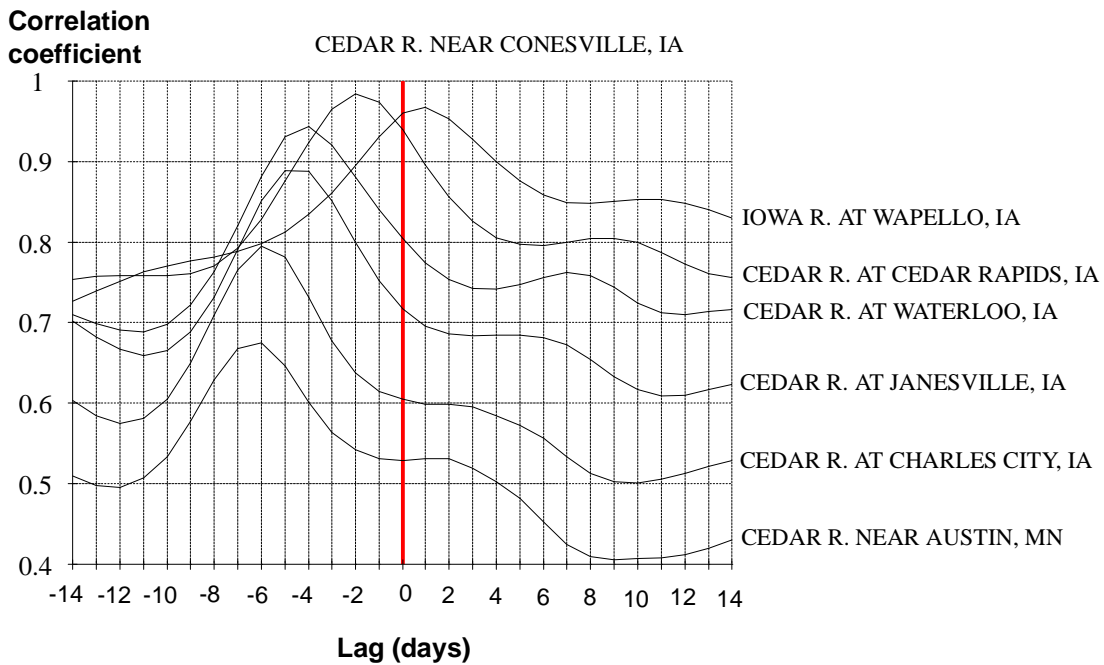
Lagged cross-correlation coefficients can be applied to test the linear relationship between flow time series recorded in the gauging stations located on the same flow path. Six gauging stations along the Cedar River shown in Figure 4.20 have been selected for preliminary analysis.



**Figure 4.20 Flow path and location of gauging stations which have been used to illustrate the potential method for time of travel estimation.**

The cross-correlation coefficients between the time series recorded at the gauging station "Cedar River near Conesville, IA" and the time series recorded in remaining gauging stations have been calculated. Figure 4.21 shows the estimated

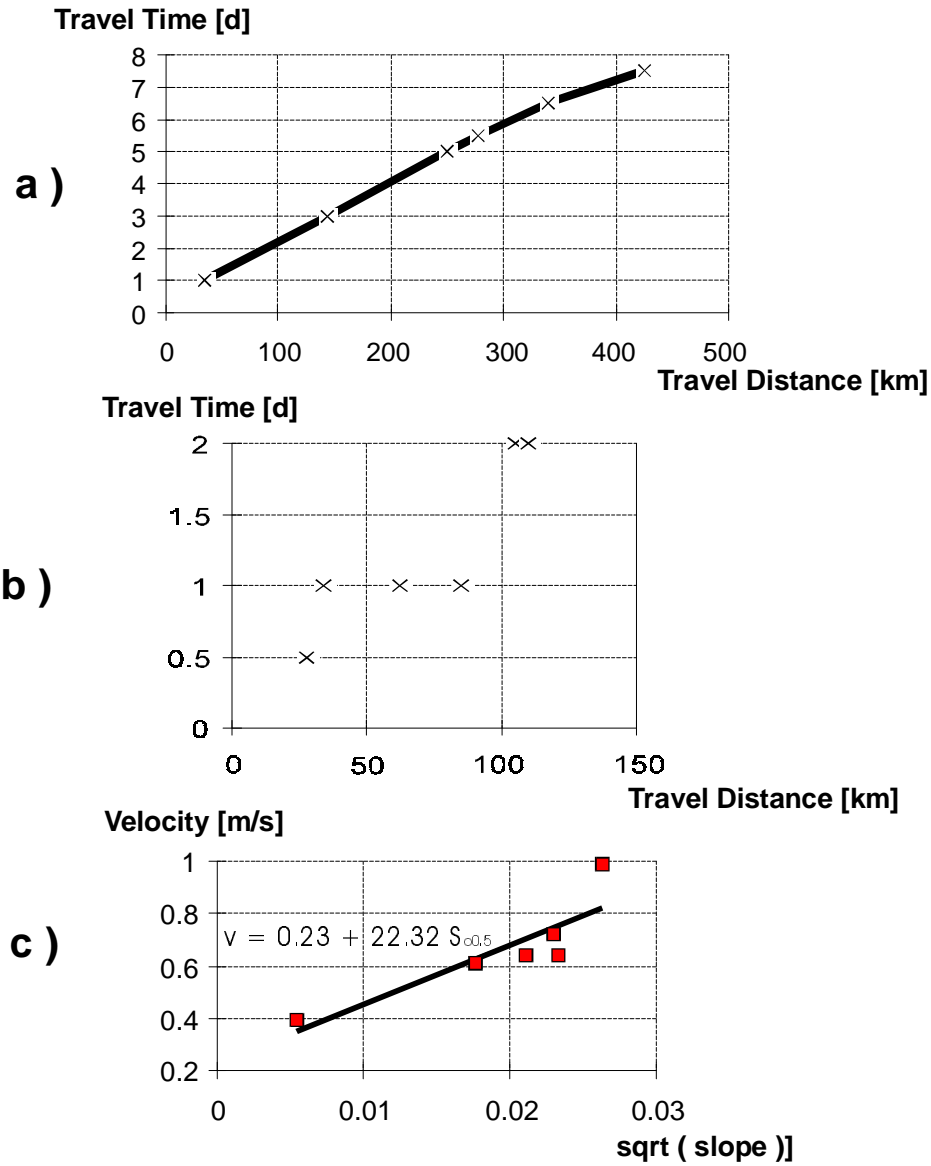
coefficients. For example, the cross-correlation between the flow recorded in gauging station "Cedar River near Conesville" and the flow recorded in gauging station "Cedar River at Cedar Rapids" has the maximum value for a two-day lag. This suggests that it takes about two days for the water to flow from Cedar Rapids to Conesville.



**Figure 4.21 Cross-correlation coefficients: the strength of the linear relationship between the flow rate in the Cedar River recorded near Conesville, IA, and the flow rate recorded at indicated locations. All gauging stations are on one flow path (shown in Figure 4.20).**

The projection of the travel time over the ungauged regions is much more difficult than the projection of the flow velocity, except the case when a simple relation between the travel time and the travel distance is applied. Figures 4.22a and 4.22b show such a relationship which indicates that the travel time can be substituted by the

travel distance. Figure 4.22c presents the flow velocity as a function of the square root of the stream slope.



**Figure 4.22 Analysis of travel time and flow velocity:**  
**a) cumulative travel time versus cumulative flow distance;**  
**b) travel time vs. flow distance;**  
**c) velocity vs. square root of stream slope.**

The travel time shown in Figure 4.22 represents the lag time for which the cross-correlation coefficient has the maximum value. The slope for each stream reach has been calculated from the digital elevation model (DEM). The flow length has been calculated by the Arc/Info function `flowlength`.

The same exercise has been repeated for the Iowa River. Two river sections have been excluded from the analysis: one between Rowan and Marshalltown (maximum correlation between no-lagged series) and other between Marengo and Iowa City (the Lake Coralville causes that the maximum correlation coefficient to be estimated for nine-day lagged series). The following relationship between “travel time”, the stream slope, and the stream length has been estimated for both the Iowa River and the Cedar River:

$$t = \frac{L}{0.3 + 20.0S_o^{0.5}} \quad (4.27)$$

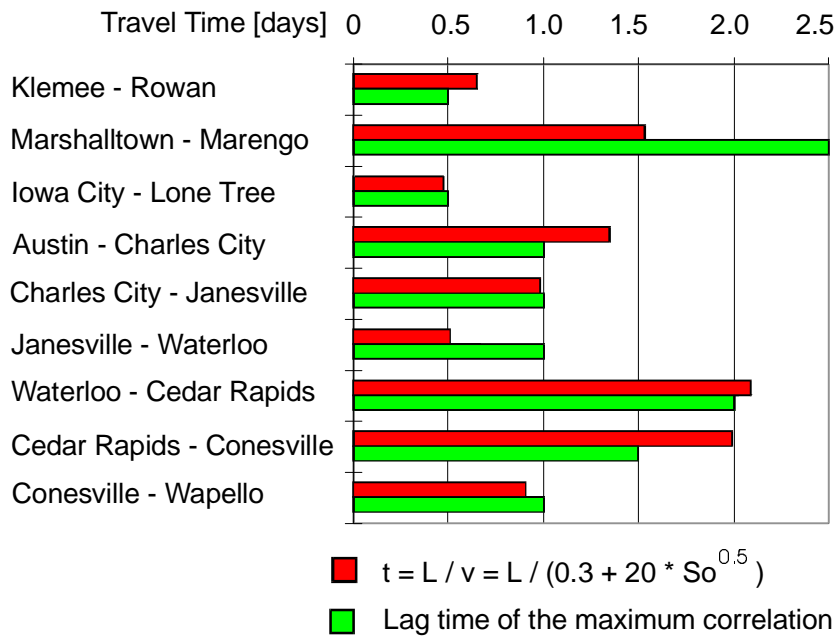
where:

$t$  = travel time [d]

$S_o$  = stream slope

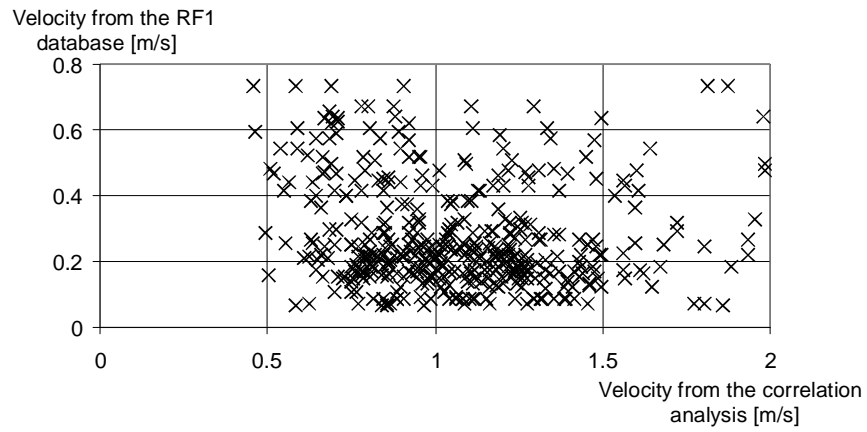
$L$  = stream length [km]

The Equation (4.27) gives a good approximation of the “travel time” in the Iowa River and the Cedar River. Figure 4.23 shows that the error of prediction for all stream reaches is below 0.5 day (the precision of determining a lag time of the maximum correlation is 0.5 day.)



**Figure 4.23 Estimation of the “travel time” for major reaches of the Iowa River and the Cedar River, IA.**

Figure 4.24 compares the flow velocity extracted from the RF1 database with the flow velocity estimated by the correlation analysis. There is no visible relation between these two data sets. The flow velocity estimated from the correlation of flow record is about three times higher than the one published with RF1. According to the RF1 velocities, it takes constituent almost a month to travel from Austin, MN to Wapello, IA. Further analysis is needed to clarify the reasons for this discrepancy.



**Figure 4.24 Comparison of the velocity from the RF1 database with the velocity estimated from a correlation analysis of the flow record in the Iowa River and tributaries**

## 5. PROCEDURES

This Section describes the procedures utilized to construct the model of agrichemical transport in the Midwest rivers. Section 5.1 shows the steps of preparing published concentrations and flow rate for statistical analysis. Section 5.2 presents Arc/Info procedures that have been prepared to estimate parameters that characterize watersheds within the Upper Mississippi-Missouri River and the Ohio River basins. These parameters include average slope of the rivers, average land slope, distance from the field to the stream along the flow path, exponentially decayed stream length, agrichemical application, average annual temperature, and average annual precipitation depth. Section 5.3 explains the process of converting the database of observations and the sampled watershed characteristic into Splus data objects for further statistical analysis. To make the presentation clearer and more consistent, the detailed description of the statistical analysis is discussed together with presentation of the results in Chapter 6.

Sections 5.4 and 5.5 explain the models of the agrichemical transport in the rivers of the Iowa-Cedar River basin. These models apply results of the statistical analysis of the data that describe the atrazine and nitrate concentrations in the Midwest rivers, the flow rate, watershed morphometry and the climatic parameters. Section 5.4 discusses the method of subdividing the Iowa-Cedar River basin into hydrologic units utilizing a digital elevation model and a digital map of rivers. This section describes also the GIS database of the flow rate and the precipitation depth and the process of the spatial redistribution of the recorded flow rate. The prototype of the agrichemical transport model programmed within the ArcView GIS is presented in Section 5.5.

## 5.1 Concentration and flow measurements

The data printed in the USGS Open File Report 94-396 (Scribner, et al., 1994) have been manually entered into a computer. The following items have been extracted from the USGS Report: sampling-site name, date of collection (month/day/year), collection time (24-hour), stream-flow ( $\text{ft}^3/\text{s}$ ), nitrate plus nitrite as nitrogen ( $\text{mg}/\text{L}$ ), atrazine ELISA ( $\mu\text{g}/\text{L}$ ), and atrazine GC/MS ( $\mu\text{g}/\text{L}$ ). If more than one sample was taken during one day, the average daily flow-weighted concentration has been calculated as  $\bar{c} = \sum_i c_i Q_i / \sum_i Q_i$  where:  $c$  = concentration,  $Q$  = flow rate.

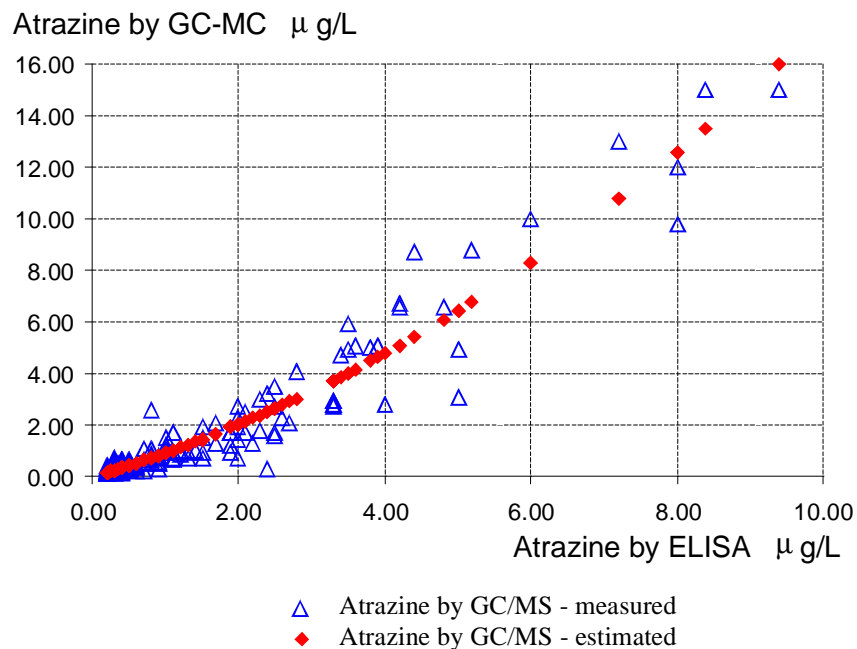
The following values have been entered into a computer from the USGS reconnaissance study (USGS Open File Report 93-457, Scribner, et al., 1993): State, site name, date of collection (month/day/year), stream-flow ( $\text{ft}^3/\text{s}$ ), nitrate plus nitrite as nitrogen ( $\text{mg}/\text{L}$ ), atrazine ELISA analyzed in Iowa laboratory ( $\mu\text{g}/\text{L}$ ), atrazine ELISA analyzed in Kansas laboratory ( $\mu\text{g}/\text{L}$ ), and atrazine GC/MS ( $\mu\text{g}/\text{L}$ ). About 43 missing values of GC/MS atrazine concentration have been determined from ELISA tests using following regression equation that relates ELISA atrazine and GC/MS atrazine (the equation has been estimated utilizing data published in USGS Open File Report 93-457):

$$\text{atrazine} = -0.0101 + 0.82547 \text{ELISA} + 0.09316 \text{ELISA}^2 \quad (5.1)$$

$R^2 = 0.92$ , 175 observations, Standard Error 0.74  $\mu\text{g}/\text{L}$ .

Figure 5.1 shows observed GC/MS atrazine vs. predicted atrazine (Eq. 5.1).





**Figure 5.1 Relationship between atrazine ELISA concentration and concentration by GC/MS estimated from data published in the USGS Open File Report 93-457.**

The estimation of the CG/MS atrazine concentrations from the ELISA tests is commonly used in practice. Figure 5.2 compares the estimated in this research relationship with the functions developed by different authors.

Thurman et al. (1992) estimated the following quadratic function (Fig. 5.1, line 1):

$$\text{atrazine} = 0.21 + 2 \text{ELISA} - 0.73 \text{ELISA}^2 + 0.15 \text{ELISA}^3 \quad (5.2)$$

$R^2 = 0.93$ ,  $n = 127$ . It is evident from Figure 5.2 that there is an error in the published equation.

Goolsby et al. (1993) fitted two regression equations for GC/MS values (Fig. 5.2, line 2 and line 3, respectively), as a function of ELISA values, for samples taken in 1990:

$$\text{atrazine} = -0.01 + 0.70 \text{ ELISA} \quad (5.3)$$

$R^2 = 0.94$ ; st. error of estimate = 0.09  $\mu\text{g/L}$ , and for samples taken in 1991:

$$\text{atrazine} = 0.80 \text{ ELISA} \quad (5.4)$$

$R^2 = 0.78$ ; st. error of estimate = 0.10  $\mu\text{g/L}$

In another study, Moody and Goolsby (1993) estimated the following relationship ( $R^2 = 0.94$ , Fig. 5.2, line 4):

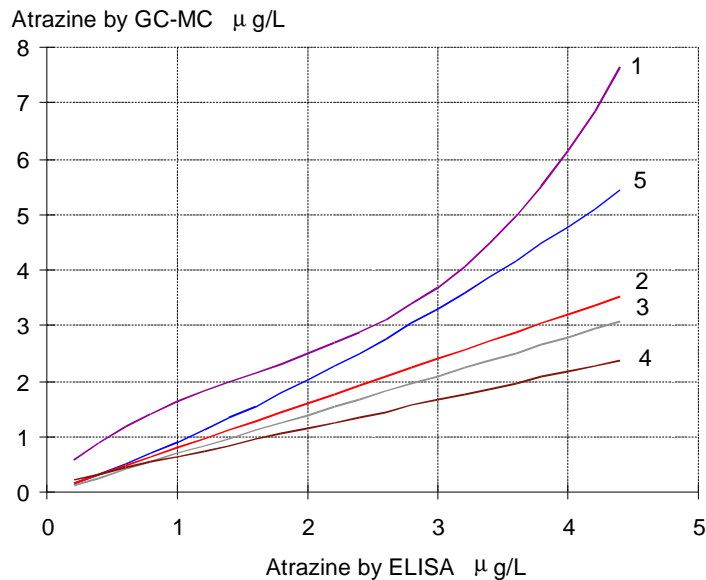
$$\text{atrazine} = 0.13 + 0.51 \text{ ELISA} \quad (5.5)$$

Gruessner et al. (1995) described the relation between ELISA atrazine and GC/MS atrazine by the equation (5.6):

$$\text{ELISA} = 0.029 + 1.16 \text{ atrazine} \quad (5.6)$$

sample size = 217,  $r = 0.96$ .

This survey of equations shows that slightly different relationships between CG/MS atrazine and ELISA atrazine are estimated from different data sets.



**Figure 5.2 Relationship between atrazine ELISA concentration and concentration by GC/MS estimated by a different authors:**  
**(1) Thurman et al. (1992);**  
**(2) Goolsby et al. (1993) data gathered in 1990;**  
**(3) Goolsby et al. (1993) data gathered in 1991;**  
**(4) Moody and Goolsby (1993), and**  
**(5) relation estimated in this study utilizing data from USGS open file report 93-457.**

The streamflow record has been converted from cubic feet per second (ft<sup>3</sup>/s) into cubic meters per second (m<sup>3</sup>/s). The two chemical data sets, one from the USGS Open File Report 94-396, and the other from USGS Open File Report 93-457, have been merged and the column containing the site identification number (USGS Station ID) has been added. A simple substitution method (Helsel and Hirsh, 1995) has been applied to values below the reporting limit. For atrazine, 0.025 µg/L replaced a less-than 0.05 µg/L value and for nitrite-nitrate as nitrogen, 0.05mg/L was substituted for a less than 0.1 mg/L value. The summary statistics of the atrazine, nitrite/nitrate-nitrogen concentrations and loads, and the corresponding flow rates are presented in Table 5.1.

**Table 5.1 Summary statistics of the nitrite plus nitrate and atrazine data sets.**

<i>Statistics</i>	<i>Nitrite/Nitrate data set</i>			<i>Atrazine data set</i>		
	<i>Flow</i> <i>m<sup>3</sup>/s</i>	<i>Load</i> <i>g/s</i>	<i>Concentr.</i> <i>mg/L</i>	<i>Flow</i> <i>m<sup>3</sup>/s</i>	<i>Load</i> <i>mg/s</i>	<i>Concentr.</i> <i>µg/L</i>
Mean	82.22	249.8907	5.22	101.15	296.62	5.35
Standard Error	19.20	16.95	0.14	26.82	39.04	0.37
Median	8.85	28.92	2.91	9.72	11.77	1.00
Standard Dev.	694.55	613.18	5.21	819.46	1192.48	11.33
Skewness	22.43	5.43	1.23	19.09	14.06	4.92
Minimum	0.0028	0.0002	0.05	0.0028	0.0011	0.03
Maximum	20416.56	7603.12	26.00	20416.56	27124.85	116.00

## **5.2 Preparing data for model parameter estimation**

To estimate sampled watershed characteristics, a system of distributed watershed parameters has been developed. This system is composed of a set of grids that describe spatial distribution of such parameters as average agrichemical application rate, average river slope, drainage area, land slope, distance from a field to the river, average annual temperature, and average annual precipitation depth for the Upper Mississippi-Missouri and the Ohio Basins. Each grid cell contains a value that describes a feature of the drainage area upstream of that cell. The 500 m resolution grid, based on the 15 second DEM, is used to accomplish this task. The following factors justify this approach:

- 1) Watershed parameters, that explain the measured loads and concentrations of selected agrichemicals in Midwest rivers, are used to estimate loads and concentrations in unsampled streams;
- 2) Estimating parameters for a region that has area almost  $3 \times 10^6$  km<sup>2</sup>, and which is represented by a grid of about 3800 rows and 5800 columns is a

very computer intensive process, so a 100 m cell size would be too intensive;

- 3) A grid-based distributed system of parameters does not require a prior subdivision of region into watersheds nor specification of a watershed outlet.

### 5.2.1 Preparing 500 m (15'') DEM for analysis

The 500 m DEM released by Rea and Cederstrand (1995) has been used in this study. Since it was published using the Albers Equal Area coordinate system with standard projection parameters, no projection was needed.

To ensure that the whole region contributes to the outflow, the DEM depressions must be removed. This is performed by the GRID procedure `fill` (Listing 5.1, line 1) which produces a “filled” DEM, `mwfill`, and a grid that contains the flow direction, `mwfdr`. GRID determines one of the eight directions of the flow from the steepest descent: value 1 represents flow in E (east) direction, 2 in SE direction, 4 - S, 8-SW 16 - W, 32 - NW, 64 - N, and 128 in NE direction. No adjustment of the elevations, i.e., “burning in” RF1 streams has been applied to the 500 m DEM due to the size of the files.

The drainage area upstream of each cell is calculated by summing the number of contributing cells and is calculated by the procedure `flowaccumulation(mwfdr)`, and multiplying by the area of single cell that is equal to a certain value (in this case,  $500\text{ m} * 500\text{ m} = 0.25\text{ km}^2$ : Listing 5.1, line 2). The stream network is delineated by assuming that a stream is created from the outflow from area at least of  $50\text{ km}^2$ , threshold number of cells (line 3 of Listing 5.1).

The grid function `con ( condition , true , false )` is a conditional statement. The expression `condition` is evaluated for each cell in the analysis window. If it is TRUE, the expression `true` identifies the value to be used to calculate the output cell values. If `condition` is FALSE the output cell is evaluated according to the expression `false`, or, if this expression is not stated, the output cell value is set to NODATA.

**Listing 5.1 Creating depressionless DEM, calculating drainage area and delineation of the stream network.**

```
1: fill mwdem mwfil sink # mwfdr
2: mwarea = 0.25 * flowaccumulation ( mwfdr )
3: mwstr = con ( mwarea > 50, 1 )
```

Before further calculations are made, the stream network delineated from the DEM is compared with the RF1 map to visually verify the flow system. One major inaccuracy was found: the Wisconsin River, a Mississippi River tributary, flowed North to the Great Lakes. To correct this error the elevation of some cells of the DEM were changed to force the flow in proper direction. Correction of the flow system can also be made by making changes directly to the flow direction grid. The process shown in Listing 5.1 was repeated until satisfactory map of the flow system was constructed.

### **5.2.2 Estimation of watershed parameters**

This section presents procedures for estimating the grids which each cell contains the following parameters of the upstream drainage area: average slope of the stream network, average exponent of the stream network length, average distance from field to river, and average slope of the watershed.

**Slope along the flow path.** To estimate the average land slope and the average stream slope, a grid that contains values of the cell slope, measured along the flow path, has to be constructed. Listing 5.2 shows an example of the AML that calculates the slope according to the relation  $(H_{i+1} - H_i)/\Delta_L$ , where  $H_{i+1}$  is the elevation stored in the cell, e.g., `mwfil(1,0)` pointed by the flow direction grid `mwfdr`,  $H_i$  is the elevation of the processed cell (`mwfil`), and  $\Delta_L$  is the distance; cell width (variable `%grd$dx%`) or cell diagonal (variable `%diag%`), depending on the flow direction value. The full listing of the procedure that calculates the slope along the flow path, is presented in Appendix C8.

**Listing 5.2 Slope along the flow path.**

```

1: &describe mwfil
2: &sv diag = 1.414213562 * %grd$dx%
3: DOCELL
4:   if (mwfdr == 1)
5:     slope3 = ( mwfil - mwfil(1,0) ) / %grd$dx%
6:   else
7:     if (mwfdr == 2)
8:       slope3 = ( mwfil - mwfil(1,1) ) / %diag%
9:     else
10:
11:
12: . . .
13:
14:   if (mwfdr == 128)
15:     slope3 = ( mwfil - mwfil(1,-1) ) / %diag%
16:   else
17:     slope3 = 0
18: END

```

**Average slope of the stream network.** First, the cells that compose the stream network are selected from the grid which contains a description of the direction of flow, `mwfdr`, - the grid `strfdr` is created. The number of cells in the upstream stream network is then calculated, using the grid command `flowaccumulation`.

The same command is used to calculate the sum of the slopes. A grid of average river slope upstream of each cell is obtained by dividing the grid of sum of slopes, `slpfacfac`, by the grid with the number of cells in the upstream network, `strfac`. To save disk space, the resulting grid, `strslp`, is multiplied by 1,000,000 and its values are converted into integer representation. Listing 5.3 shows the GRID dialog which performs this procedure.

**Listing 5.3 Average slope of the stream network.**

```

1: strfdr = con ( mwstr > 0 , mwfdr )
2: strfac = flowaccumulation ( strdir )
3: slpsfac = flowaccumulation ( strdir, slope3 )
4: strslp = slpsfac / strfac
5: slps6i = int ( strslp * 1000000 )

```

**Average exponent of the negative stream network length.** According to the information presented in Section 4.3.5 the methodology has been simplified by assuming that the mass that enters the surface waters upstream of each cell is equal to one and is uniformly distributed along the river. The reduced formula describing variable  $E_S$  (Eq. 4.9) has the following form:

$$E_{Sk} = \frac{1}{n_k} \sum_{i=1}^n \exp(-L_{ik}) \quad (5.7)$$

where:

- $E_{Sk}$  = weighted average of the exponent of negative flow distance in rivers upstream to the  $k$ -th cell (decay),
- $n_k$  = number of cells that constitute the upstream network (`strfac`),
- $L_{ik}$  = length of the flow path from the  $i$ -th stream cell to the  $k$ -th cell [ $10^2$  km].

Listing 5.4 shows the GRID dialog for computing parameter  $E_{Sk}$ . The command `flowlength`, that determines the downslope distance along the flow path, from



each cell to the outlet on the edge of the grid, `mwflg`. The exponent of the negative flow distance, expressed in 100 km, is then calculated. The cells that comprise the stream network are selected and the cumulative value of the exponent of the negative flow distance is determined.

Since function `flowlength` calculates the distance from each cell to the outlet on the edge of the grid, not to the watershed outlet, the grid, `strefac`, must be divided by the grid, `mwexpflg`, according to relation:

$$\sum_{i=1}^n \exp(-L_{ik}) = \frac{\sum_{i=1}^n \exp(-L_{io})}{\exp(-L_{ko})} \quad (5.8)$$

where:

$L_{ik}$  = length of the flow path from the  $i$ -th stream cell to the  $k$ -th cell,

$L_{io}$  = length of the flow path from the  $i$ -th stream cell to the outlet on the edge of the grid (cell  $o$ ),

$L_{ko}$  = length of the flow path from the watershed outlet (cell  $k$ ) to the outlet on the edge of the grid (cell  $o$ ).

The division by `strfac` in line 5, Listing 5.4, is introduced to account for the assumption of a unit mass entering river network. The purpose of the operation stated in the last line of Listing 5.4 is to reduce the size of the computer file; integer grids occupy considerably less disk space than the real-number grids (30%-90% less).

**Listing 5.4 Exponentially decaying flow length.**

```

1: mwflg = flowlength ( mwfdr )
2: mwexpflg = exp ( -0.00001 * mwflg )
3: mwexpstr = con ( mwstr > 0 , mwexpflg )
4: strefac = flowaccumulation ( mwfdr , mwexpstr )
5: decay = ( strefac / mwexpflg ) / strfac
6: decs6i = int ( decay * 1000000 )

```

**Average length of the flow path from field to the stream.** To exclude the cells that represent a stream network from calculations, a grid containing one in the cells that do not belong to the river network, and zero for the stream cells is created in line 1, Listing 5.5. The downslope distance, along the flow path, from each cell that does not represent a stream to all cells that stand for river is calculated in line 2, Listing 5.5. The average value is determined by summation of all distances upstream (line 3, Listing 5.5) and by dividing the resulting grid, `lndflfac`, by the total number of cells within the given drainage area that are not streams, `lndfac` (line 5, Listing 5.5).

**Listing 5.5 Average distance from field to a stream.**

```

1: lnd1str0 = con ( isnull(mwstr) , 1 , 0 )
2: lndflg = flowlength ( mwfdr , lnd1str0 )
3: lndflfac = flowaccumulation ( mwfdr , lndflg )
4: lndfac = flowaccumulation ( mwfdr , lnd1str0 )
5: alndflg = lndflfac / lndfac
6: alndflgi = int ( alndflg )

```

**Average land slope.** Listing 5.6, line 1 shows the command used to create a grid, `xlndslp`, containing the value of the slope in the cells that do not belong to the stream network, and zero for the stream cells. The average slope, `lndslp`, is calculated by first, summing all of the upstream values of slope, (line 2, Listing 5.6) and by dividing this cumulative slope grid, `slplfac`, by the total number of cells within the given drainage area that do not represent streams, `lndfac`, (line 3, Listing 5.6). The last operation converts the slope grid into an integer format.

**Listing 5.6 Average land slope.**

```

1: xlndslp = con ( isnull ( mwstr ) , slope3 , 0 )
2: slplfac flowaccumulation ( mwfdr , xlndslp )
3: lndslp = slplfac / lndfac
4: slpl6i = int ( lndslp * 1000000 )

```

### 5.2.3 Creating grid of sampling sites

A grid that represents the USGS sampling sites (watershed outlets) is necessary to extract the parameters of the reconnaissance watersheds from the grids of distributed parameters. This section describes the process used to create such a grid from available data.

The Arc/Info map `sta_recon` obtained from Battaglin (1995) of 147 sampling sites that are described in USGS Open File Report 93-457 (Scribner, et al., 1993) was converted into raster format by the following GRID operation:

```
sites1 = pointgrid ( sta_recon, station, #, #, 500 )
```

Unfortunately, four of the nine stations from USGS Open-File report 94-396 (Scribner, et al., 1994) are not represented in the `sta_recon` coverage. Therefore, a coverage of these four sampling sites was constructed from published latitude-longitude coordinates (`generate`). The attribute table was created (`build`) and the coordinates of the resulting coverage were changed from geographic projection into Albers projection (`project`). Finally, this coverage was converted into grid `sites2` and merged with grid `sites1`:

```
site_all = con ( isnull(sites1) , sites2, sites1)
```

The site identification numbers, given in attribute table of the `sta_recon` coverage (Battaglin, 1995) of the three sites, have been changed to site identifiers that are published in USGS Open-File Report 93-457 (Scribner, et al., 1993), as follows:

The Mississippi River at Winfield, ID = 1 has been changed to ID = 05587450  
(the Mississippi River at Grafton);

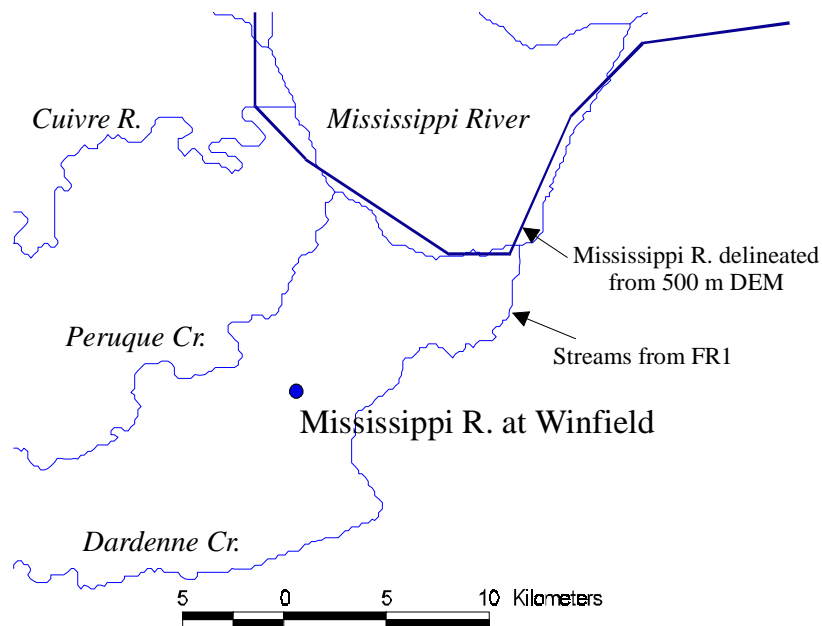
The Ohio River near Olmstad, ID = 2 has been assigned ID = 03612500  
(the Ohio River near Grand Chain);

The Wabash River near Griffin ID = 3 has been replaced by ID = 03378500  
(the Wabash River near New Harmony).

The drainage area of the original station and the one which replaced it is identical for all three sites.

#### 5.2.4 Adjusting location of sampling sites

The location of most of the sampling sites is not completely consistent with the stream network delineated from the 500 m DEM. Figure 5.3 presents an example of a sampling site (Sta-recon map from Battaglin, 1995) that is located neither on the RF1 stream, nor on the stream network that has been delineated from the DEM.



**Figure 5.3** Example of the sampling site (from Battaglin 1995) that is located neither on the stream from RF1 nor on the stream delineated from the DEM.

The cells that represent both the sampling site and the watershed outlet must be located on the proper flow path. The following procedure has been developed to adjust the position of gauging stations:

- constructing a grid of sampling sites that contains published values of drainage area (`areap`). This step has been performed within ArcView. The new item, `dareakm2`, has been added to the value attribute table (VAT) of the grid `site_all`. This VAT is linked to the point attribute table (PAT) of the coverage `sta_recon` that contains in item `darea` (the USGS estimates of drainage area in  $\text{mi}^2$ ). The values of `dareakm2` have been calculated in ArcView using the following formula in “field calculate” entry box:  $[\text{dareakm2}] = [\text{darea}] * 2.59$ , where 2.5900 is the square mile-to-square kilometer conversion factor. The grid, `areap`, is created from the grid, `site_all`, by a simple GRID assignment statement (line 1, Listing 5.7);
- assigning the drainage area value to each cell in the circular neighborhood of cells in grid `areap`. A five cell radius is assumed (line 2, Listing 5.7);
- calculating the absolute value of the relative difference between the published drainage area and the distributed values that have been estimated from the DEM in each sampling site cell within its neighborhood (line 3, Listing 5.7);
- creating grid of zones (`zone5`) by assigning a unique number to every continuous area with the same value (line 4, Listing 5.7);
- within each zone, finding a cell that has the smallest difference between the drainage area determined from DEM and the area given by USGS (smallest value in each zone of grid, `error5`) (line 5, Listing 5.7);

- assigning the site identification number to the cells that have minimum error (lines 6 and 7 of Listing 5.7);
- selecting cells with a relative error of drainage area smaller than 20% (line 8, Listing 5.7)

**Listing 5.7 GRID dialog supporting adjustment of sampling sites location.**

```

1: areap = site_all.dareakm2
2: usgsarea = focalmax (areap, CIRCLE, 5 )
3: error5 = int ( abs ( ( 100000 * (usgsarea - mwarea) ) /
   usgsarea ) )
4: zones5 = regiongroup ( usgsarea, #, FOUR)
5: er_min5 = zonalmin (zones5 , error5 )
6: id_zone = focalmax ( site_all, CIRCLE, 4)
7: siteid5 = con ( ( er_min5 == error5 ) , id_zone )
8: siteer20 = con ( er_min5 < 20000 , siteid5 )

```

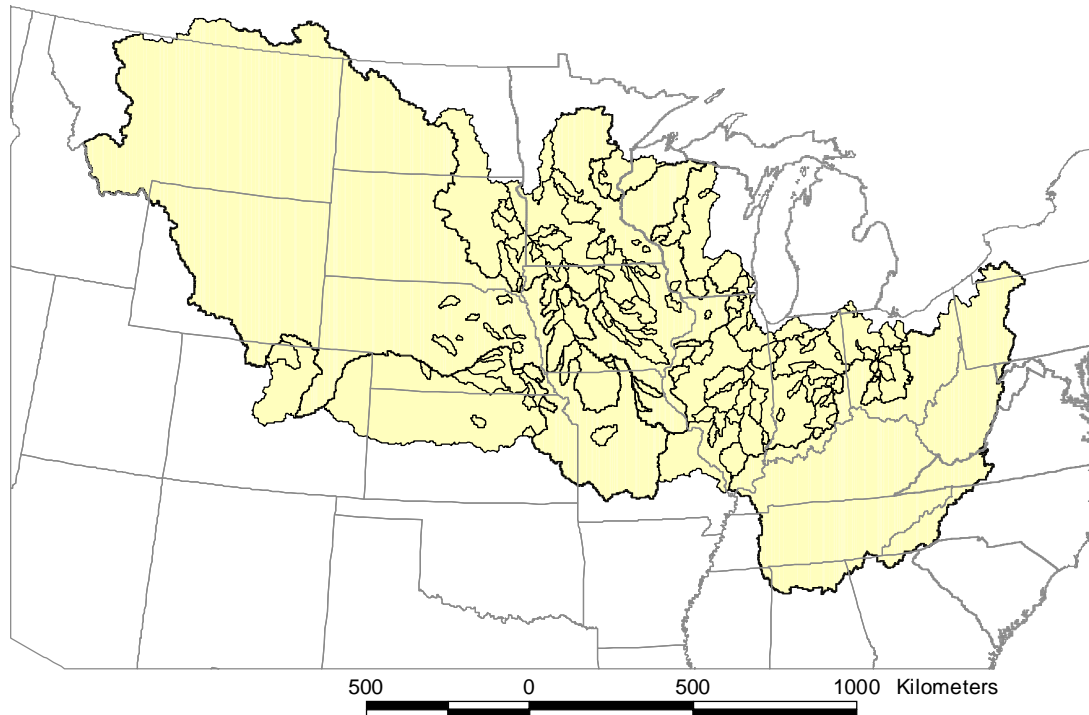
Wherever possible, the locations of watershed outlets that have a difference in drainage area larger than 20% have been adjusted manually using the grid editor included in ArcTools processor. The results have been stored as mwout grid. About 13% of stations had an error in drainage area larger than 20%. The discrepancies are due to the relatively low resolution of the DEM used in this analysis and, therefore, small precision of the watershed boundaries. These discrepancies are also due to errors in the estimation of drainage area by the USGS. For example, Kalkhoff and Kuzniar (1991) state that the drainage area of sample site 05412100, the Roberts Creek above St. Olaf, Iowa, is equal to 70 mi<sup>2</sup>, whereas Scribner et al. (1993) published a number that is 47% larger (103 mi<sup>2</sup>). The grid mwout has been used to delineate watersheds from the 500 DEM:

```

mwvsh = watershed ( mwfdr, mwout )

```

Figure 5.4 shows the sampled watersheds delineated from the 500 m DEM.



**Figure 5.4** Sampled watersheds delineated from the 500 m DEM.

### **5.2.5** Extracting parameters of the sampled site watersheds

Listing 5.8 shows the process of extracting watershed parameters from the grids of distributed values and associating them with the site identification number. As a result a VAT is created that contains the following items: `mwout` (site ID), `slps6i2` (stream slope), `decs6i2` (decayed flow length), `flg10i2` (flow length from field to stream), and `slp16i2` (land slope).

**Listing 5.8 Extracting parameters of the sampled site watersheds.**

In GRID:

```

1: slps6i2 = con ( mwout > 0 , slps6i )
2: decs6i2 = con ( mwout > 0 , decs6i )
3: flgl0i2 = con ( mwout > 0 , alndflgi )
4: slpl6i2 = con ( mwout > 0 , slpl6i )
5: comslps = combine ( mwout , slps6i2 )
6: comdecs = combine ( mwout , decs6i2 )
7: comflgl = combine ( mwout , flgl0i2 )
8: comslpl = combine ( mwout , slpl6i2 )

```

In Arc/Info:

```

1: copy comslpl comtot
2: joinitem comtot.vat comflgl.vat comtot.vat mwout mwout
3: joinitem comtot.vat comslps.vat comtot.vat mwout mwout
4: joinitem comtot.vat comdecs.vat comtot.vat mwout mwout

```

Since the parameters were multiplied by 100000 to convert them to integers, a new Info table that contains the recalculated floating point values has been constructed using Arc/Info and ArcView tools. The items and their units in the new table are as follows: *station\_id*, *area* (km<sup>2</sup>), *decstr* (e<sup>100km</sup>), *slpstr* (dimensionless), *slplnd* (dimensionless), and *alflgkm* (km). The summary of statistics for a selected parameters of sampled watersheds are presented in Table 5.2. The average land surface slope is approximately one percent while the average stream slope is about six times flatter, or 0.17%.

**Table 5.2 Selected statistics of watershed parameters determined from DEM. Length of the flow path used to calculate  $E_S$  is in 100 km.**

<i>Statistics</i>	<i>Area</i> <i>km<sup>2</sup></i>	<i>Decayed</i> <i>stream length</i> <i><math>E_S</math></i>	<i>Stream</i> <i>Slope</i> <i><math>S_S</math></i>	<i>Land</i> <i>Slope</i> <i><math>S_L</math></i>	<i>Land</i> <i>length</i> <i><math>L_L</math> [km]</i>
Mean	36493	0.59124	0.00168	0.01091	3.34
Standard Error	18360	0.01877	0.00007	0.00057	0.04
Median	1685	0.63221	0.00146	0.00871	3.23
Standard Deviation	224117	0.22912	0.00091	0.00700	0.50
Minimum	173	0.00353	0.00039	0.00130	2.45
Maximum	2335354	0.91928	0.00702	0.04486	5.51



### 5.2.6 Agrichemical application

The mass of agricultural chemical that was applied within each watershed in which the concentrations and flow rate were measured, has been calculated using Arc/Info-GRID. This procedure is presented in Listing 5.9. A detailed description of the process is given below:

- In the first four lines of Listing 5.9 the polygon coverages NIT89, NIT90, NIT91, and HERBICIDE1, are converted into grid maps that have cell size equal to 500 m, and grid values equal to the values from items NTOT89.USE, NTOT90.USE, NTOT91.USE, and H1080.USE, respectively. Lines 5-8 transform the application units into grams per square kilometer ( $\text{g}/\text{km}^2$ ) and modify the format of numbers from double precision to integer to reduce the size of the files. The total application of the nitrogen-fertilizer in all cells upstream of a given cell is calculated in line 9, 10, and 11 for 1989, 1990, and 1991, respectively. The total application of atrazine is determined in line 12.
- The cells that contain the amount of chemical application in the watersheds under investigation are selected in line 13 and they are converted into integer format in line 14 in order to construct a VAT (line 15). The total chemical application map is combined with the grid that contains the site ID number to create a table (VAT) that contains two important items: `mwout`--the watershed ID and `nit89g`--the value of chemical use. Lines 13-16 process nitrogen-fertilizer use in 1989. By a similar procedure, the use of the remaining agrichemicals is estimated for each sampled watershed.

**Listing 5.9 Chemical application in sampled watersheds.**

```

1: xn89 = polygrid(nit89 , ntot89.use, #, #, 500 )
2: xn90 = polygrid(nit90 , ntot90.use, #, #, 500 )
3: xn91 = polygrid(nit91 , ntot91.use, #, #, 500 )
4: xa89 = polygrid(herbicide1 , h1980.use, #, #, 500 )
5: n89gkm2 = int ( xn89 *350262.198 )
6: n90gkm2 = int ( xn90 *350262.198 )
7: n91gkm2 = int ( xn91 *350262.198 )
8: a89gkm2 = int ( xa89 *175.1312741 )
9: n89g = flowaccumulation ( mwfdr, n89gkm2 * 0.25)
10:n90g = flowaccumulation ( mwfdr, n90gkm2 * 0.25)
11:n91g = flowaccumulation ( mwfdr, n91gkm2 * 0.25)
12:a89g = flowaccumulation ( mwfdr, a89gkm2 * 0.25)

13:na89g = con (mwout > 0, n89g )
14:ni89g = int ( na89g )
15:buildvat ni89g
16:comn89 = combine (mwout, ni89g)

```

Table 5.3 shows a summary of the statistics of estimated agricultural chemical use in the sampled watersheds. The application rate has been determined by dividing the total agrichemical application by the drainage area.

**Table 5.3 Summary statistics of annual application of nitrogen fertilizer and atrazine.**

<i>Statistics</i>	<i>Nitrogen fertilizer</i>		<i>Atrazine</i>	
	<i>Application</i> <i>kg/yr</i>	<i>Appl. rate</i> <i>kg/km<sup>2</sup>/yr</i>	<i>Application</i> <i>kg/yr</i>	<i>Appl. rate</i> <i>kg/km<sup>2</sup>/yr</i>
Mean	43705590	5878.1	219644	26.66
Standard Error	7170678	51.7	40274	0.42
Median	10298910	6034.0	48695	25.32
Standard Deviation	259336782	1871.2	1230168	12.74
Minimum	245134	481.6	466	1.00
Maximum	4.79E+09	9481.1	19294461	50.55

### 5.2.7 Annual temperature and annual precipitation depth

Listing 5.10 shows the procedure of creating grids of annual precipitation depth. In line 1 grid `prinyr` (annual precipitation depth in inches) is created from Arc/Info coverage `clim_div`, utilizing data stored in its polygon attribute table item `pre.mean`. Line 2 converts vector map of average temperature (item `tmp.mean`) into grid representation `tmf`. In lines 2 and 3 the units are changed from inches to millimeters for precipitation, and from degrees Fahrenheit to degrees Celsius for temperature. Since the Missouri Basin extends behind the US borders, and the `clim_div` coverage contains data only for the US, temperature of 5 °C (line 4) and precipitation depth of 333 mm (line 3) has been assumed for the Canadian part of the Missouri Basin. A grid of annual precipitation depth in watershed upstream of each cell (`prmmavg`) is calculated by summing all precipitation-cells over drainage area (line 7, grid `mwprfac`) and then by dividing the result by the number of cells that constitute upstream drainage area (line 8). The distributed system of the watershed-average temperature `tmcavg` is calculated in lines 9 and 10.

#### Listing 5.10 Average climatic parameters.

```
1: prinyr = polygrid (clim_div, pre.mean)
2: tmf = polygrid (clim_div, tmp.mean )
3: prmm0 = 25.4 * prinyr
4: tmc0 = 5 * (tmf - 32 ) / 9
5: prmm = con (isnull (prmm0), 333 , prmm0 )
6: tmc = con (isnull (tmc0 ), 5 , tmc0 )
7: mwprfac = flowaccumulation(mwfdrr , prmm)
8: prmmavg = mwprfac / mwfac
9: mwtcfac = flowaccumulation(mwfdrr , tmc)
10: tmcavg = mwtcfac / mwfac
```

The process of extracting temperature and precipitation data for the sampling sites is presented in Listing 5.11. In lines 1-4 the climate grids are converted into

integer format. Before conversion, all grids are multiplied by 1000 to preserve the decimal component (three decimal digits). Lines 5-8 shows constructing grids that contain values only in cells that represent sampled watershed outlets. The VATs that relate sampling site ID stored in grid `mwout`, and the estimated parameter (command `combine`) are created in lines 9-12. The summary Info table `clim.dat` is constructed in lines 13-15. This table contains the following items:

- `mwout` = sampled watershed ID;
- `tca103s` = average temperature over sampled watershed;
- `tc103s` = annual temperature at sampling site;
- `pmma103s` = annual precipitation depth, average for sampled watershed;
- `pmm103s` = annual precipitation depth at sampling site.

The temperature stored in the `clim.dat` is in  $10^{-3}$  °C and the precipitation depth is in  $10^{-3}$  mm.

**Listing 5.11 Creating a summary Info table of climatic parameters.**

```

1: pmm1000 = int ( prmm * 1000 )
2: pmma1000 = int ( prmmavg * 1000 )
3: tc1000 = int ( tmc * 1000 )
4: tca1000 = int ( tmcavg * 1000 )
5: pmm103s = con ( mwout > 0 , pmm1000 )
6: pmma103s = con ( mwout > 0 , pmma1000 )
7: tc103s = con ( mwout > 0 , tc1000 )
8: tca103s = con ( mwout > 0 , tca1000 )
9: com1 = combine (mwout , pmm103s)
10: com2 = combine (mwout , pmma103s)
11: com3 = combine (mwout , tc103s)
12: com4 = combine (mwout , tca103s)
13: arc joinitem com1.vat com2.vat clim.dat mwout mwout
14: arc joinitem clim.dat com3.vat clim.dat mwout mwout
15: arc joinitem clim.dat com4.vat clim.dat mwout mwout

```

### 5.3 Preparing data for statistical analysis

The Info tables that contain selected watershed features and annual atrazine and nitrogen-fertilizer application have been merged with the table that contains measured flow and concentration in 153 sampling sites on the Midwest rivers. Two tables have been created: `atra7` and `nitr7`. The table, `atra7`, contains the following data:

**Table 5.4 Description of tables `atra7` and `nitr7`.**

atra7		nitr7	
state	state name abbreviation	state	state name abbreviation
name	name of the sampling site;	name	name of the sampling site;
id	station id;	id	station id;
area	drainage area in km <sup>2</sup> ;	area	drainage area in km <sup>2</sup> ;
use	total mass of atrazine use in kg/yr;	use	total mass of N-fertilizer use in kg/yr;
appl	atrazine application rate in kg/km <sup>2</sup> /yr;	appl	nitrogen-fertilizer application rate in kg/km <sup>2</sup> /yr;
date	month/day/year of sample collection;	date	month/day/year of sample collection;
day	day in which sample was collected, 1...365;	day	day in which sample was collected, 1...365;
month	month in which sample was collected, 1-12;	month	month in which sample was collected, 1-12;
flowm3s	flow rate in m <sup>3</sup> /s;	flowm3s	flow rate in m <sup>3</sup> /s;
loadmgs	atrazine load in g/s;	loadgs	nitrate load in g/s;
loadgd	atrazine load in g/d;	loadkgd	nitrate load in kg/d;
concmgm3	concentration in µg/L (mg/m <sup>3</sup> );	concmgm3	concentration in mg/L (g/m <sup>3</sup> );
decstr	exponent of negative flow length;	decstr	exponent of negative flow length;
declnd	exponent of negative distance from field to stream network;	declnd	exponent of negative distance from field to stream network;
slpstr	slope of the streams;	slpstr	slope of the streams;
slplnd	average slope of the land;	slplnd	average slope of the land;
alflgkm	average length from field to stream network in km;	alflgkm	average length from field to stream network in km;
prmm	precipitation at sampling site mm/yr	prmm	precipitation at sampling site mm/yr
prmmavg	precipitation depth in sampled watershed mm/yr	prmmavg	precipitation depth in sampled watershed mm/yr
tmc	annual temperate at sampling site °C	tmc	annual temperate at sampling site °C
tmcavg	average temperature in sampled watershed °C	tmcavg	average temperature in sampled watershed °C

The statistical program Splus operates only on objects. All data and results of statistical analysis are stored as objects. A table can be stored as an *array* object or as a *frame* object. The `atra7` and `nitr7` data sets have been imported into Splus data object *frames*:

```
atra7 _ read.table("atra7" , header=T)
nitr7 _ read.table("nitr7" , header=T)
```

Stepwise regression has been used to estimate model parameters. The following commands show an Splus session for selecting significant components of the model by the stepwise regression:

```
alnc.lm _ lm(log(concmgm3) ~ 1, data = atra7)
acla _ step(alnc.lm, ~ log(appl) + log(flowm3s)
+         + log(decstr) + slpstr
+         + log(alflgkm) + log(slp1nd)
+         + month + sin(pi*month/12) + cos(pi*month/12)
+         + sin(2*pi*month/12) + cos(2*pi*month/12)
+         + sin(4*pi*month/12) + cos(4*pi*month/12)
+         + sin(6*pi*month/12) + cos(6*pi*month/12)
+         + sin(8*pi*month/12) + cos(8*pi*month/12)
+         + sin(10*pi*month/12) + cos(10*pi*month/12)
+         + sin(12*pi*month/12) + cos(12*pi*month/12) )
```

The results of regression analysis are displayed by the Splus command `summary()`

```
summary (acla)
```

Different models have been tested by adding or removing model components. The following line shows removing variable `log(appl)` from the model `acla`:

```
acla2 _ update(acla, . ~ . - log(appl) )
```

Two diagnostic plots have been used to visually validate the assumption that the residuals are normally distributed:

1) plot of residuals versus fitted values

```
plot(fitted(acla), studres(acla) )
```

2) a normal quantile plot of residuals

```
qqnorm(resid(ac1a))  
qqline(resid(ac1a))
```

The data can be transformed within the specification of the regression model, as well as the new columns in the data frames can be derived from the original data set.

Below are two examples of creating new values from existing ones:

```
nitr7[, "srslpstr"] _ sqrt ( nitr4[, "slpstr" ] )  
nitr7[, "etstr" ] _ nitr6[, "decstr" ] ^ ( 1 / nitr7[, "srslpstr" ] )
```

The results of the statistical analysis of data and the agrichemical transport models selection are discussed in Chapter 6.

## **5.4 *Agrichemical concentrations in the Iowa-Cedar River basin***

The procedure of modeling of agricultural chemical concentration and load in the Iowa River and its tributaries includes:

- The Iowa-Cedar River basin is subdivided into modeling units, using the 100 m DEM. This step includes such operations as downloading 1° DEMs via Internet, converting the RF1 river map from vector format to raster format, determining a map of the flow direction, creating a map of the USGS gauging stations, constructing a map of modeling unit outlets, and determining modeling unit boundaries;
- The map of modeling units is converted from grid into vector format;

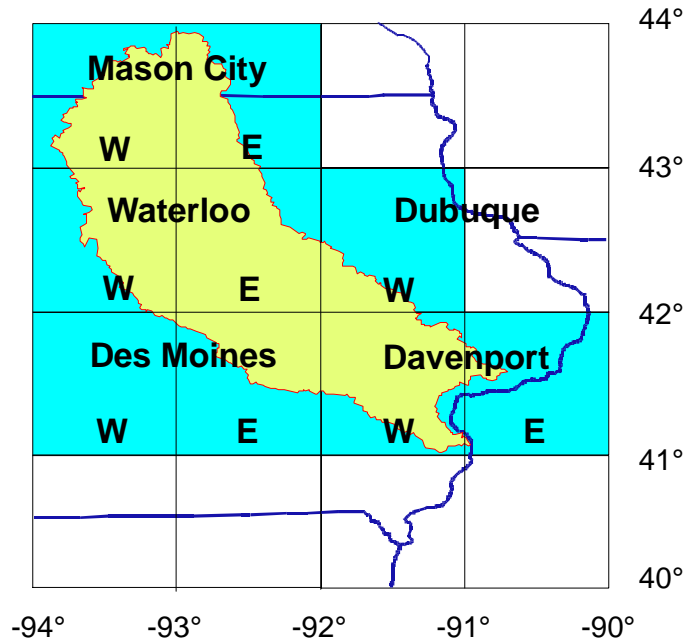
- The parameters of the total drainage area that is defined by the outlet point of the modeling unit, such as average distance from the field to the stream, and watershed slope, are estimated from the 100 m DEM, and inserted into modeling unit polygon attribute table;
- The mean monthly runoff at each outlet of the modeling unit is calculated, utilizing the measured runoff in 28 USGS gauging stations and observed precipitation depth from 86 NCDC stations;
- The link between map of counties that contain data about the agrichemical application and the map of unit watersheds is established;
- The equation that describes the concentration of agrichemicals in streams is incorporated into the GIS database;
- For each spatial unit and each month of the year, a regression equation is applied to calculate chemical concentration at the elementary watershed outlet;
- The spatial distribution of chemical loads is estimated by multiplying the concentration by the mean monthly flow rate at each outlet.

#### **5.4.1 Creating a map of the flow direction**

**Downloading DEM via Internet.** Nine, one-degree quadrangles (mason city-w.gz, mason city-e.gz, waterloo-w.gz, waterloo-e.gz, dubuque-w.gz, des moines-w.gz, des moines-e.gz, davenport-w.gz, and davenport-e.gz) of DEMs that contain digital elevation data for the region enclosed by meridians: 90° and 94° West (longitude) and by parallels 41° and 44° North (latitude), have been downloaded from the USGS National Geospatial Data Clearinghouse via Internet



([http://edcwww.cr.usgs.gov/glis/hyper/guide/1\\_dgr\\_dem](http://edcwww.cr.usgs.gov/glis/hyper/guide/1_dgr_dem)). Figure 5.5 shows the location and the names of the downloaded quadrangles.



**Figure 5.5** 1° quadrangles of DEM utilized to subdivide the Iowa-Cedar River basin into modeling units; Map projection: geographic.

The files have been uncompressed (UNIX command `gunzip`), and modified as explained below. The DEM files do not contain record delimiters, thus they should be added by the following UNIX command:

```
dd if=inputfilename of=outputfilename ibs=4096 cbs=1024 conv=unblock
```

No direct conversion of the USGS DEM's into Arc/Info GRID format is supported. The Arc command `demlattice` has been used to convert a DEM in USGS or TAME format into a lattice. The following example shows the conversion of the DEM `waterloo_e.dem` into Arc lattice `us4292lat`:

```
Arc: demlattice waterloo_e.dem us4292lat USGS
```

After all nine DEMs were converted into lattices, they were merged to construct a single map. For example, the following grid command has been used to merge four lattices:

```
Grid: crlat1 = merge(us4394lat, us4393lat, us4294lat, us4293lat)
```

In Arc/Info, the concepts of lattice and grid are similar, with one important exception. A lattice is a regularly-spaced sample of points representing a surface. Elevation values in a lattice correspond to discrete points. In a grid, the values apply to the entire cell. Although both lattice and grid apply the same data structure, the fact that a lattice represents data at points (has no area) and grid cell has an area, causes differences in how various operators interpret the data contained in a lattice and within a grid. The further analysis described here, in which the lattices were treated as grids, was not affected by these differences.

Since imported maps are in geographic coordinate system (latitude/longitude) they have been projected into Albers system (`project`), assuming the cell size of output grid equal to 100 m. Appendix C9 contains a macro that converts a grid into Albers Equal Area coordinates. The following parameters have been used:

```
Units          METERS
Datum          NAD83
Spheroid       GRS1980
1st standard parallel = 29 30 0.000
2nd standard parallel = 45 30 0.000
central meridian   = -96 0 0.000
latitude of projection's origin = 23 0 0.000
```

**Converting river map from vector format into grid representation.** The following GRID commands have been applied to specify the cell size and the map extent of the rasterized RF1:

```
setcell crdem  
setwindow crdem
```

The portion of the RF1 map that represents rivers in the Iowa-Cedar River basin has been converted into grid format by applying the `linegrid` command:

```
crrf1 = linegrid ( rf1, rf1_id )
```

The item, `rf1-id`, from the AAT (arc attribute table) has been used to assign values to the cells that represent rivers. This item has been selected since it allows one to relate the grid VAT table with the RF1.FLOW Info table, through one of the following identification fields: CUSEG (hydrologic cataloging unit code and reach segment number) or RR (Reach File ID --an unique identifier for each river reach). The table, RF1.FLOW, contains useful information for hydrologic modeling. For example, it comprises such items as: LFVEL--low flow velocity, MFVEL--mean flow velocity, MNFLO--USGS mean annual flow, and SVTEN--USGS 7-10 Year Flow (Lanfear, 1994).

The raster version of RF1 was edited using Arc/Info interactive editing environment, `arctools`. The lakes were replaced by a single line or, if the region of lake was enclosed, by the series of cells that fill up the enclosed region. The resulting data set was saved under the original grid name, `crrf1`.

### **Creating a map of the flow direction and delineation of the stream network.**

To ensure that the flow paths delineated from the DEM are in line with the RF1 river network, the terrain map has been adjusted. Elevations of all cells that do not represent the real stream network have been raised by 10000 m:

```
crdem2 = con ( isnull ( crrf1 ), crdem + 10000 , crdem )
```

The value 10,000 was selected arbitrarily. During this research it was found that this increment should be larger than the maximum elevation of the DEM within the analysis region for two reasons:

1) To ensure that the cells that constitute the stream network are lower than the cells that do not represent the stream. For example, increasing the elevation of all cells that do not represent the stream network by 5 m in the DEM which values vary from 0 to 50 m does not ensure the stream-cells to be below the neighboring land-cells (except for smoothed DEMs). Thus an uncertainty about the compatibility between RF1 rivers and the delineated rivers is introduced;

2) To easily select all cells that have been adjusted. Selection can be made by a single conditional grid expression, for example, all adjusted cells in grid `crdem2` can be set to 1 by the following grid assignment:

```
output = con ( crdem2 > 9000 , 1 ).
```

The depressions have been removed to ensure that the whole region contributes to the runoff, and the flow direction grid `crfdr` has been built:

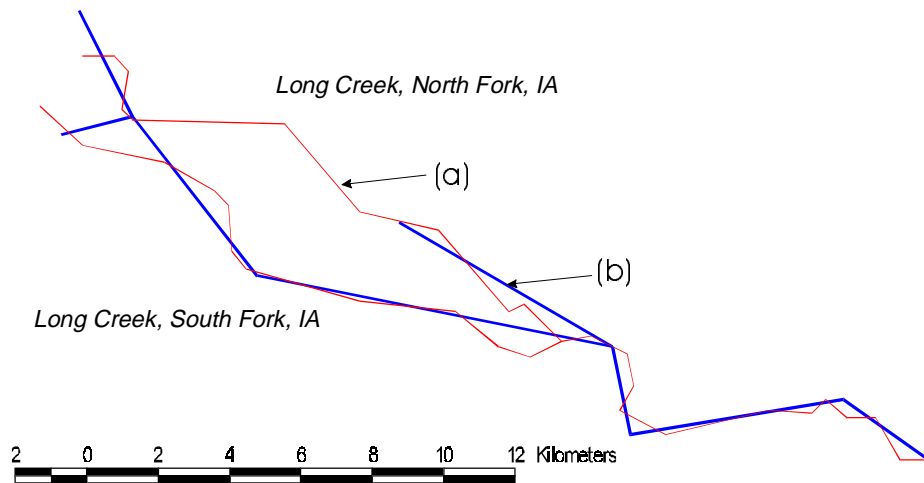
```
fill crdem2 crfill12 crfdr
```

Rivers have been delineated under the assumption that the runoff from a drainage area of 25 km<sup>2</sup> produces a stream. To all cells that have an accumulated number of cells greater than 2500 (= 25 km<sup>2</sup> with 100 m DEM cells) “flowing” into them, the value one has been assigned:

```
crfac = flowaccumulation ( crfdr )  
crstr25 = con ( crfac > 2500 , 1, 0 )
```

Conversion of the RF1 file into a 500 m by 500 m raster format revealed that the cell size of 500 m is too large to represent 1:500,000 river network, without

ambiguity. Some of RF1 reaches are closer to each other than one kilometer, and if converted into 500 m grid they are artificially connected. When applied in the procedure described above, i.e., if the RF1 streams are burned into the 500 m DEM, these connected rivers cause the GRID command `flowdirection` to determine a false direction of flow (due to the difference in elevation of watersheds). As a result, the water flows from one watershed into another through connected streams that belong to different basins. Figure 5.6 shows the low precision of the streams delineated from 500 m DEM adjusted for RF1 streams. The same streams but delineated from a 100 m DEM adjusted for RF1 almost perfectly represent the RF1 rivers--this justifies the application of the 100 m DEM in this study.



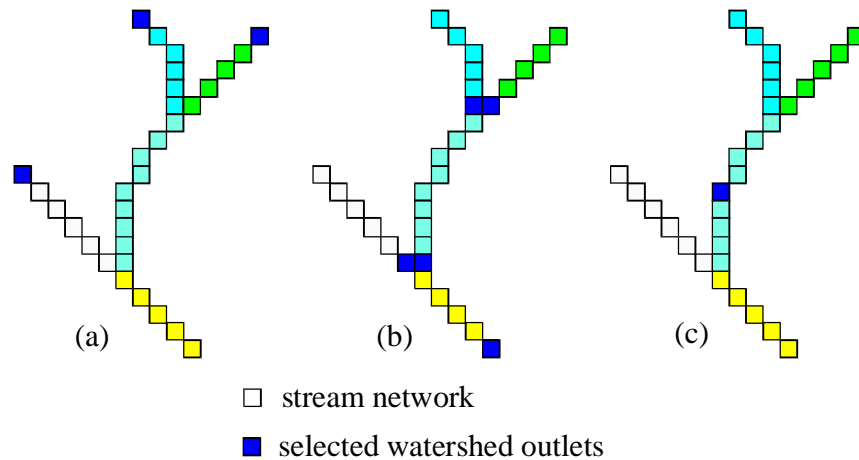
**Figure 5.6** Comparison of the stream network delineated from 100 m DEM (a) with the one delineated from 500 m DEM (b). Both DEMs were adjusted for RF1 stream network.

## 5.4.2 Map of the modeling unit outlets

A watershed is explicitly defined by its outlet point. Therefore, to subdivide the basin into elementary watersheds, a set of points -- elementary watershed outlets-- must be specified. Three types of watershed outlets are utilized to subdivide the Iowa-Cedar River basin into units:

- The most upstream cell of the first order stream;
- The most downstream cell of each reach; and
- The cells that represent the USGS stations.

Figure 5.7 shows examples of watershed outlets



**Figure 5.7 Selected types of the modeling unit outlets**

**(a) type 1, beginning of the stream network;**

**(b) type 2, upstream cell of the stream junction;**

**(c) type 3, USGS gauging station**

**Beginnings of the delineated stream network (type 1 outlets).** As was mentioned in Section 4, the type 1 outlets have been introduced to have more control on the average area of the modeling unit and to determine the flow (and chemical load) conditions at the beginning of the river system. Listing 5.12 shows the Grid dialog that selects appropriate cells. In line 1, the raster stream network `crstr25` is divided into “grid zones” `crls1`. In Arc/Info GIS a grid zone is developed from the cells that have the same value. Command `streamlink` assigns an unique value to each stream reach. In line 2, `zonalmin` function finds the minimum value in flowaccumulation grid `crfac` delineated by the zones of `crls1` and assigns it to all cells of the reach. Function `streamorder` (line 3) assigns to each reach a numeric order (Strahler or Schreve). In line 4, the most upstream cell in the first order stream is found by selecting the cell that has the same value in both grids, `crlmn` and `crfac`. The value of the selected cells is equal to the value from the grid `crls1` increased by 100,000, to indicate the type 1 outlet and thus the type of watershed (source watershed). This number also indicates (after subtracting 100,000) the ID of the downstream unit (intermediate watershed) making the numbering system of the hydrologic structure consistent and efficient.

**Listing 5.12 Selecting type 1 watershed outlets.**

```
1: crls1 = streamlink ( crstr25 , crfdr )
2: crlmn = zonalmin (crls1 , crfac )
3: crlso = streamorder ( crstr25 , crfdr )
4: crlpu = con(crlso == 1 AND crlmn == crfac, crls1 + 100000)
```

**River junctions (type 2 outlets).** The type 2 outlets are determined by selecting the most downstream cell in each stream reach. The selection is made by finding in each “grid zone,” defined by grid `crsls`, a cell that has the maximum value in

flowaccumulation grid `crfac` (Listing 5.13). The last operation shown in Listing 5.13 merges the grid of type 1 outlets with the grid of type 2 outlets.

**Listing 5.13 Selecting type 2 watershed outlets.**

```
1: crlmx = zonalmax (crlsl , crfac )
2: crlpd = con(crlmx == crfac, crlsl)
3: crlud = con ( isnull( crlpu), crlpd, crlpd)
```

**USGS gauging stations (type 3 outlets)** The description of the USGS gauging stations published by Hydrosphere (Hydrosphere, 1994) contains latitude and longitude. These coordinates have been used to produce a point coverage of the USGS stations (Listing 5.14). Listing 5.15 shows a part of the file, `latlonfl.csv`, that contains point coordinates.

**Listing 5.14 Creating point coverage of gauging stations from latitude-longitude coordinates.**

```
Arc: generate xgsflow
Generate: input latlonfl.csv
Generate: points
Generate: q
Arc: build xgsflow point
```

**Listing 5.15 Fragment of the file `latlonfl.csv` containing station IDs and coordinates.**

```
5448500,-337370,154670
5449000,-337051,154831
5449500,-337056,153936
5451700,-334285,151225
. . .
```

To make the point coverage compatible with other maps used in this study, it has been projected into Albers system of coordinates as follows (Listing 5.16):



**Listing 5.16 Projecting the point coverage of USGS gauging stations from Geographic system into Albers coordinates.**

```
Arc: project cover xgsflow gsflow
Input
Projection geographic
units ds
Parameters
output
Projection      ALBERS
Zunits          NO
Units           METERS
Spheroid        GRS1980
Xshift          0.0000000000
Yshift          0.0000000000
Parameters
  29 30 0.000 /* 1st standard parallel
  45 30 0.000 /* 2nd standard parallel
-96  0 0.000 /* central meridian
  23  0 0.000 /* latitude of projection's origin
0.00000 /* false easting (meters)
0.00000 /* false northing (meters)
end
```

The grid representation of the USGS stations that is compatible with the delineated stream network has been developed using Arc/Info grid editor tools (arctools). The following line combines the grid with the USGS stations `crgs` with the grid of watershed outlets `crlud`:

```
crou = con ( isnull( crlud), crgs, crlud)
```

Since the gauging station cells may overlap with other cells representing watershed outlets units the following statement can be used to establish the priority of which cell needs to be selected in the final map of the modeling units (grids priority from the highest to the lowest: `crlpu`, `crgs`, `crlpd`):

```
crou = con( isnull(crlpu) , con( isnull(crgs), crlpd, crgs), crlpu)
```

The map of modeling units is created by the following Grid statement:

```
crwsh = watershed (crfdr , crou)
```

An Arc/Info script has been developed to create the following grids: stream network, watersheds, and watershed outlets. The complete listing of this program, `wshgs .aml`, is presented in Appendix C1. A first order watershed, and its outlet cell, have an ID number in the range  $100000 < ID < 200000$ . This number is the ID of the first order stream increased by 100000. A watershed that drains through the point at which the gauging station is located assumes the USGS gauging station identification number, 5448500 for example.

### 5.4.3 Watershed connectivity

To determine the average value of the model parameters for the drainage area upstream of a given point, the sum of values for all upstream modeling units must be calculated. All these units can be identified if the position of each unit on the flow path is known. The Arc/Info script `nextwsh .aml` creates an Info file with two items that store the modeling unit ID number and the ID number of the next unit on the flow path (downstream unit). The listing of the macro `nextwsh .aml` is shown in Appendix C2. The major part of this AML assigns to the cell that corresponds to the watershed outlet `crout` a value from the cell, located in watershed grid `crwsh`, that is pointed by the flow direction `crfdr`. In line 1, Listing 5.17, a value zero is assigned to all cells that have NODATA (`xcrwsh`). This step is necessary to ensure that the next watershed to the last unit on the flow path is indicated by the ID equal zero (no downstream units). Lines 2-9 create a grid `crnxt` similar to the grid of watershed outlets `crout` but with ID of the downstream unit.

#### Listing 5.17 Identification of the downstream watershed ID.

```
1: xcrwsh = con ( isnull ( crwsh ), 0, crwsh )
2: crnxt = con ( crout > 0, con ( crfdr == 1, xcrwsh(1,0), ~
3:           con ( crfdr == 2, xcrwsh(1,1), ~
```

```

4:      con (crfdr == 4, xcrwsh(0,1), ~
5:      con (crfdr == 8, xcrwsh(-1,1), ~
6:      con (crfdr == 16, xcrwsh(-1,0), ~
7:      con (crfdr == 32, xcrwsh(-1,-1), ~
8:      con (crfdr == 64, xcrwsh(0,-1), ~
9:      con (crfdr == 128, xcrwsh(1,-1), -1)))))))))

```

#### 5.4.4 Refining modeling units

Two problems occurred when the grid of modeling units was converted into vector format:

- 1) Small zones were not converted into visible polygons, although the records in polygon attribute tables were created. This makes it impossible to select records in the PAT by pointing to polygons on the screen;
- 2) A simple grid zone had more than one component area and the components were connected by the cell corner. It was converted into more than one polygon, i.e., one modeling unit was described by two records in the PAT. For this unit, the value that represents an area in the VAT, after conversion into a polygon coverage, was listed in the PAT twice. Thus, the modeling unit had two times larger area in vector map than the same unit represented in the grid.

To solve the problems mentioned above, all zones of smaller area than one km<sup>2</sup> (1 km<sup>2</sup> = 100 cells 100 m \* 100 m) were incorporated into neighboring units. The GRID nibble command replaces areas in a grid corresponding to a mask with the values of the nearest neighbors (Listing 5.18):

#### Listing 5.18 Eliminating drainage units that are smaller than 1 km<sup>2</sup>

```

aaa = crwsh.count
aaa2 = con (aaa > 100 , 1 )
crwsh2 = nibble(crwsh, aaa2, DATAONLY )

```

The `nibble` command has also been used to remove the upstream portion of the unit composed of two parts connected only by the cell corner.

The removal of small drainage areas introduced discontinuity of the flow system. A program was written in C language to determine the new relation between modeling units. The C code of the program `newnx.c` is listed in Appendix A1. This procedure requires two text (ASCII) files. In the first file, each line contains two numbers separated by a comma: `unit_id` and the downstream `unit_id`, for the original set of modeling units. The second file contains only the list of `unit_id` --identification numbers of units that remained from the original set after some of them have been removed. Program `newnx` creates an ASCII file in which it stores updated information about system connectivity. In each line, the following numbers are written, separated by a comma: `unit_id`, downstream `unit_id`, and numeric order of the unit in the flow system.

By assigning a numeric order to each unit, the further calculations are more efficient. The method of stream numbering is as follows: all exterior units, i.e., the most upstream units are assigned an order 1. The order of each interior unit is calculated as the maximum order of the upstream units increased by one.

The advantage of the program written in C over the procedure written in Arc/Info macro language or in ArcView script language Avenue, is that the C program, after minor changes, can be used by both Arc/Info and ArcView, whereas program written in AML can be executed only from Arc/Info, and a script written in Avenue language can only be run from ArcView. The additional advantage of writing more complex procedures in such languages as FORTRAN or C is that the time of execution is very fast. The time gain is on the order 1000, i.e., the computational hours can be reduced to seconds.

The following assignment creates the final map of the modeling units in vector format (weed tolerance = 180 m):

```
crwsd = gridpoly ( crwsh3 , 180 )
```

The map of subdivision of the Iowa -Cedar River basin into 1032 modeling units is presented in Figure 5.8.



**Figure 5.8** The Iowa-Cedar River basin subdivided into 1032 unit drainage areas of average area 31.6 km<sup>2</sup>.

#### **5.4.5 Database of monthly precipitation depth and monthly flow rate**

The time series of the average monthly flow rate and the average precipitation depth for 38 USGS gauging stations and 86 climate stations, for the time period from January 1940 to September of 1992, have been extracted from the Hydrosphere CD-ROMs (Hydrosphere 1993 a, b) and stored in a worksheet format.

The units of measurements have been changed:

- cubic feet per second (cfs) into cubic meter per second ( $\text{m}^3/\text{s}$ ) (flow) and
- inch per day into centimeter per day (cm/d) (precipitation).

For each parameter, two sets of ASCII comma delimited files have been created. One set contains identification numbers of observation stations and respective longitude and latitude, expressed in decimal seconds. The other set of files comprise the station ID number and the time series of flow or precipitation.

The point coverage of 86 precipitation stations has been created from the ASCII files and projected into Albers Equal Area system of coordinates. This procedure is identical to the one described in a previous section for generation of point coverage of USGS stations.

The ASCII files that contain flow and precipitation measurements have been converted into Arc/Info INFO files and linked with the appropriate maps.

#### 5.4.6 Average precipitation depth in modeling units

The method of spatial redistribution of measured flow rate utilizes the precipitation depth as a weighting factor. Therefore, the monthly average precipitation depth must be calculated for each modeling unit. The example of the GRID dialog which calculates a raster map of the average precipitation depth, in modeling each modeling unit for May 1990 is shown in Listing 5.19.

The grid of distributed precipitation depth has been created directly from the climate station map. The inverse distance weighting procedure `idw` has been applied. To make the calculations shorter and to save the disk space, 100 times larger cells (width = 1000 m, area = 1 km<sup>2</sup>) than those used for the stream and modeling unit delineation have been assumed (`setcell 1000`). The average values for each modeling unit have been determined by the `zonalmean` function. To create an attribute table, the cell values must be represented as integers. In line 7 of listing 5.19, the grid, `xxx1`, that contains zonally averaged precipitation depth is first multiplied by 10000 to preserve the decimal part and then converted into integer format. The command `combine` has been used to create a table in which the modeling unit identification number (stored in grid `crwsd`) is related to the estimated average precipitation depth (stored in grid `xxx2`).

**Listing 5.19 Estimation of the average precipitation depth in modeling units.**

```
1: setcell 1000
2: setwindow MAXOF
3: xxx = idw ( gsrain , m199005 )
4: setcell crwsd
5: setwindow crwsd
6: xxx1 = zonalmean ( crwsd , xxx )
7: xxx2 = int ( 10000 * xxx1 )
8: xxxcom = combine ( crwsd , xxx2 )
```

Since the procedure described above must be executed for each month, it has been included into AML's DO feature that repeats calculations for each month of the specified time period, i.e., from month %fm% of the year %fy% to month %tm% of the year %ty%. An example of the &DO block is presented in Listing 5.20

**Listing 5.20 Application of the &DO command to repeat action for each month of the selected time period.**

```

1: &DO yr = %ty% &to %fy% &by -1
2:   &DO mt = 12 &to 1 &by -1
3:     &s ab = [calc %yr% = %fy%] AND [calc %mt% lt %fm%]
4:     &s bb = [calc %yr% = %ty%] AND [calc %mt% gt %tm%]
5:     &s cc = %ab% OR %bb%
6:     &IF NOT %cc% &THEN
7:       &do
8:         . . . /* here, block to be repeated
9:       &end
10:    &end
11:&end

```

Appendix C3 contains the code of the Arc/Info macro RAININFO.AML. This program calculates average values for specified zones and then writes them into an INFO file. The user must specify the period for which the estimates have to be made, the point coverage that holds the time series, and the grid that divides the region into zones. The output of the macro is the Arc/Info INFO file that contains for each zone the ID number and the series of computed zonal averages.

In the approach presented above, only stations that have complete time series, i.e., stations that have no missing values, have been applied. The macro SELDATA.AML that is included in Appendix C6 selects stations that have a complete record in the specified time interval.

To include all available measurements of precipitation depth, another approach has been tested. The stations that have data for the processing month have been used, unlike the method discussed above in which only the stations that have data for all



months of the specified time period have been utilized. For each month, station selection is performed (Arc/Info command `reselect`) and then the grids of spatial distribution of the precipitation depth are calculated (IDW procedure). This task is performed by the macro `RAINMAP.AML` (attached in Appendix C4). The results were slightly different, but since the zonal averages were calculated later, the differences were insignificant. Listing 5.21 shows an example procedure that selects all stations from the database (coverage `prcmap`) that have a complete record in June, 1966 (point coverage `xxx`), and then it creates the grid of precipitation (`pm196606`) by the inverse distance-squared method:

**Listing 5.21 Example of the AML that selects precipitation stations that have a complete record in June, 1996 and creates the grid of precipitation depth (adopted from `RAINMAP . AML`).**

```
1: arc reselect prcmap xxx point
2: res m196606 ge 0
3: ~
4: n
5: n
6: arc build xxx point
7: pm196606 = int ( 1000 * idw ( xxx, m196606 ) )
8: kill xxx all
```

#### **5.4.7 Spatial distribution of flow**

Three approaches have been made to incorporate the procedure described in Section 4.5 into GIS. Coding has been made in:

- 1) Avenue, the script language of ArcView;
- 2) Arc/Info macro language, AML;
- 3) C programming language.

For the first two approaches, the calculations of the flow rate for a single month took hours, whereas the program written in C language redistributed the flow record in a few seconds. Therefore a method that uses the Arc/Info Tables to prepare the data and then utilizes the C program to calculate the flow rate in modeling units has been developed. It is explained here for a single month: March 1990. The full Arc/Info macro--`fd4y.aml`--is presented in Appendix C7. The C program--`fd4y.c`--is listed in Appendix A2. The `fd4y` program requires two input files and it stores the results in one output file. In lines 2 and 3 of Listing 5.22, the Arc/Info processor, Tables, creates the first input file `xxxgsin` in which the following items of the point attribute table `gsfl28.pat` are stored:

- `station_id` = ID number of the USGS gauging station;
- `station_nx` = ID number of the downstream USGS station;
- `m199003` = item that contains flow record for March 1990

The second input table, `xxxunin`, is prepared in lines 4 and 5. The following items are unloaded from the polygon attribute table `unprec.pat`:

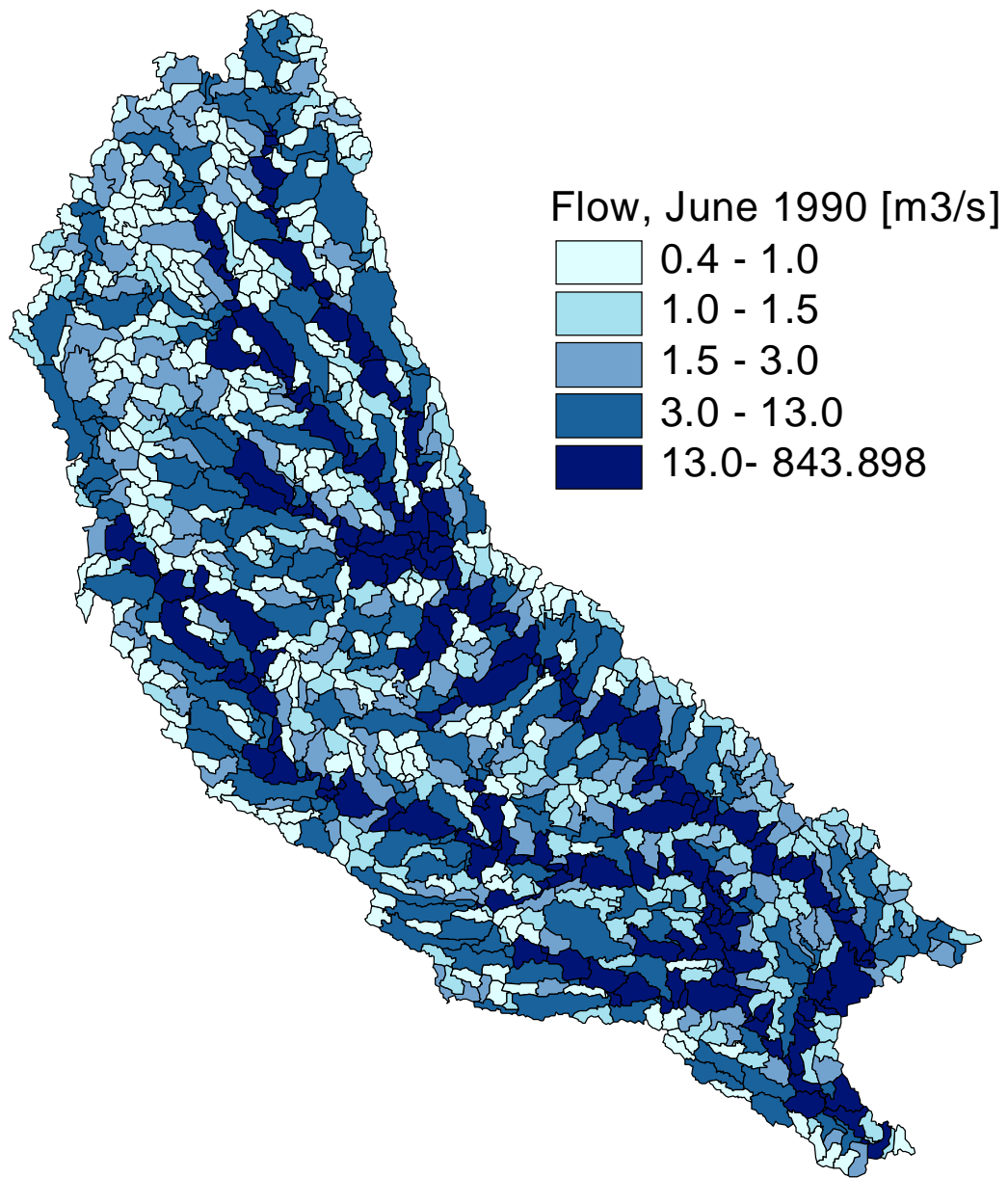
- `unit_id` = modeling unit ID number;
- `unit_nx` = ID number of the downstream unit
- `gswsh` = ID of the USGS gauging station that is located downstream from the given unit;
- `area_km2` = area of the unit;
- `order` = modeling unit order in the flow system; and
- `pm199003` = item that contains average precipitation depth.

The program `fd4y` is executed in line 6. Since UNIX is a multitasking system, the AML must wait until `fd4y` finishes calculations. Lines 7 to 10 contain the loop in which AML checks if `fd4y` program created output file `xxxunout`. In lines 11 to 14, the temporary INFO file `m199003.dat` is created in which the results of the

fd4y calculations are stored (line 15). Then, the estimated flow rate is attached to the Arc/Info database file unflow.pat (line 17) and the temporary file m199003.dat is deleted (line 18). In the complete version of this procedure (fd4y.aml), the process described here is repeated for all months of the specified time period. Figure 5.9 presents an example of estimated runoff that occurred in June 1990.

**Listing 5.22 Example of the Arc/Info macro that prepares data and estimates the flow rate in modeling units (adopted from fd4y.aml, Appendix C7)**

```
1: tables
2: select gsfl28.pat
3: unload xxxgsin station_id station_nx m199003 DELIMITED INIT
4: select unprec.pat
5: unload xxxunin unit_id unit_nx gswsh area_km2 order pm199003
  DELIMITED INIT
6: &sys fdy4 xxxgsin xxxunin xxxunout
7: &s i = 0
8: &do &until [exists xxxunout -file]
9:   &s i = %i% + 1
10:&end
11:define m199003.dat
12:unit_id,4,8,B
13:qm199003,4,12,F,4
14:~
15:add from xxxunout
16:select m199003.dat
17:&sys arc joinitem unflow.pat m199003.dat unflow.pat unit_id
  order
18:kill m199003.dat
```



**Figure 5.9** Runoff (flow in rivers) that occurred in June 1990 in the Iowa-Cedar River watershed.

#### **5.4.8 Determining input values for the Iowa-Cedar River model**

The flow rate at the outlet of each modeling unit has been estimated by the procedure described in Section 5.4.7.

The distributed system of watershed parameters and normal weather parameters have been created for the Iowa-Cedar River using 100 m grid according to the methodology discussed in Section 5.2.2. The concentration models determined by the regression analysis relate agrichemical concentration to parameters of the upstream drainage area. Thus it is important to get parameters for the entire upstream area at each unit watershed outlet rather than data just for the modeling unit belonging to that point. The characteristics of the drainage area that is upstream of each modeling unit outlet, have been extracted from the distributed system of watershed parameters according to the methodology discussed in Section 5.2.5.





#### **5.5 *ArcView model of agrichemical transport***

The computer version of the agrichemical transport model is completely constructed within the GIS software ArcView. A modular structure has been utilized. Such a structure not only makes the model simple and easy to understand, but also it allows improvement of the model or addition of new blocks without changing the model's core components. The program execution is supported by a customized

graphical user interface (GUI). Tools that enhance the graphical representation of the results are developed. The ArcView model contains procedures for representing hydrologic processes.

### 5.5.1 Model overview

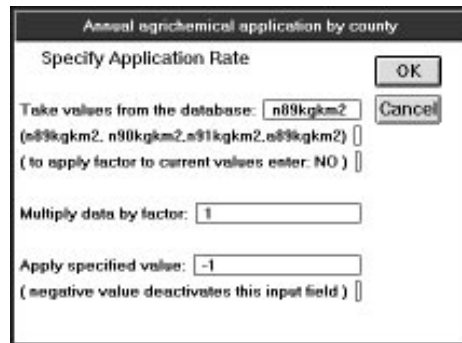
The ArcView application is divided into four projects (a project is a collection of associated documents: views, tables, layouts, charts, and scripts). Each project is designed to perform a different task. Four buttons that allow user to switch between projects have been placed in the PushButton Bar (the avenue scripts assigned to each button are listed in Appendix B1). These buttons have the following functions:

-  Runs the “model” project. This project is designed to prepare entry data for the agrichemical transport model as well as to calculate concentrations and loads.
-  Runs the “results” project which is designed to display the results as maps of agrichemical concentration and load.
-  Executes project “flwprc” which provides visualization of the flow record for a selected modeling unit.
-  Project “tools” contains the fundamental tools for hydrologic modeling.

The project “model” is the main component of the ArcView application. Projects “results”, “flwprc”, and “tools” are supporting modules. The scripts associated with the project “model” are listed in Appendix B2. The scripts used by the

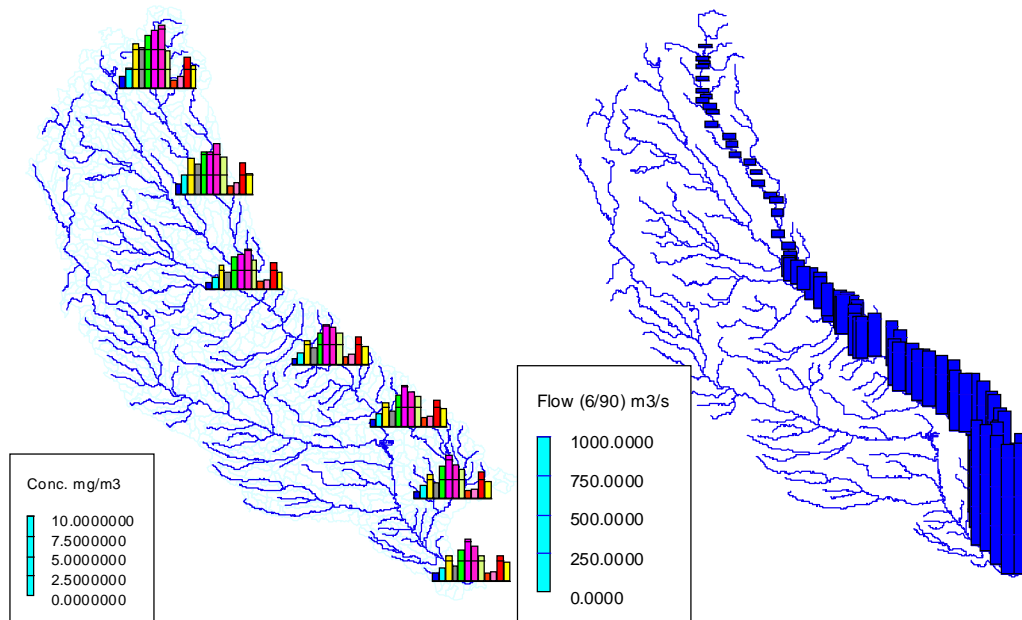
projects “results”, “flwprc”, and “tools” are presented in Appendices B3, B4, and B5, respectively.

The agrichemical transport model is navigated through a set of buttons that, when pressed, execute particular model components. The data entry and the data selection is maintained by a set of windows. Figure 5.10 presents an example of a window that supports selection of the agrichemical application rate to be used for concentration estimation. This window contains three input fields. In the first field, the user can enter the name of agrichemical application database that will be utilized for predicting the chemical concentrations and loads in the Iowa-Cedar River basin. The second field allows user to multiply the application rate in the counties which are selected by ArcView tools. This feature is introduced to study the response of the concentration in rivers after the application rate has been increased or decreased by a specified percentage. The last entry field permits one to specify an application rate for a selected county or all counties without the necessity of editing the attribute tables directly. This field is deactivated if a negative value is entered.



**Figure 5.10** An example of the Graphical User Interface: a window for the selection of the agrichemical application rate in selected counties.

The presentation tools support drawing bar charts at the center of selected units. They are developed to enhance the presentation of the data and the model results on a map. Figure 5.11 shows examples of the application of the presentation tools.



**Figure 5.11 Atrazine concentrations along the Cedar River estimated for 1990 (left) and flow rate along the Cedar River in June 1990 (right).**

The hydrologic modeling is supported by scripts that perform the following tasks:

- Selection of the units downstream of a selected point;
- Selection of all units upstream of a selected unit;
- Estimation of the watershed weighted average of an entity such as chemical application rate;
- Accumulation of values when moving downstream (flow accumulation);



- Calculation of the unit mass balance (reverse operation to flow accumulation); and
- Determination of the order of the unit (all exterior units/streams are assigned an order of one, an interior unit has the order equal to the maximum order of the upstream units increased by one).

### 5.5.2 Project “Model”

When opened, the project “Model” displays the following maps:

- 1) Agrichemical application rate by county (window: Application Rate, theme: Use);
- 2) The Iowa-Cedar River basin divided into modeling units (window: Modeling Units, theme: Units).

Five buttons are designed to select different activities:





Edit Application. This button displays dialog boxes that allow the user to specify the agrichemical application rate. If the view-window “Application Rate” is active, the application in either all, or selected counties can be specified. Additionally, the application rate of nitrogen fertilizers estimated for 1989, 1990, and 1991, and application rate of atrazine estimated for 1989 can be selected. A multiplier that increases/decreases application rate by a given percentage also can be used to simulate the different policy scenarios. If the view-window “Modeling units” is active, the application rate in each modeling unit can be modified.





Edit Cumulative Flow Rate. This button shows dialog boxes for selecting the flow record (by selecting a year) that will be used for calculations and for

adjusting the selected values. For example, all data can be multiplied by a factor that represents extreme conditions. The adjustment may be performed for all months of the specified year or for each individual month.

 Select Year. This button allows the user to specify a year for a model that has a trend component. There is no trend in the current model.

 Select Month. The dialog box displayed by this button specifies if the calculations will be performed for all months or only for selected months of a year.

 Run. Executes the script that performs the calculations.

 First Order Reaction. Calculates concentrations and loads assuming exponential decrease of the agricultural constituent along the flow path. The current model does not contain a map of spatially distributed decay rates that allows one to estimate concentrations by this method.

The following coverages and data files constitute the core database of the model:

- 1) Coverage `cruse`--a map of 47 counties that are within the Iowa-Cedar River basin. Table 5.5 presents the items of `cruse`'s PAT.

**Table 5.5 Polygon Attribute Table of the cruse coverage.**

Field (Item)	Description
Fips	county FIPS code
St	state
Cntyname	county name
N89kgkm2, N90kgkm2, N91kgkm2	estimated average nitrogen fertilizer application rate for the years 1989, 1990, and 1991, respectively [kg/km <sup>2</sup> ]
A89kgkm2	estimated average atrazine application rate for 1989 [kg/km <sup>2</sup> ]
Temp	field reserved for storage of results of partial calculations
Use	application rate selected for the estimation of concentrations and loads

2) Coverage crwsd--map of 1032 modeling units. Table 5.6 lists the items of the attribute table.

**Table 5.6 Polygon Attribute Table of the crwsd coverage.**

Field (Item)	Description
Unit_id	ID of the modeling unit
Gswsh	ID of the USGS gauging station downstream of the modeling unit
Unit_nx	ID of the next unit on the flow path (downstream unit ID)
Order	a number that specifies location of the unit on the flow path
Areakm2	area of the unit [km <sup>2</sup> ]
Careakm2	drainage area upstream of the unit outlet [km <sup>2</sup> ]
Alndslp	average land slope of unit drainage area
Alndlgkm	average length on the flow path from the land to the stream network [km]
Chemuse	Agrichemical application rate in the unit [kg/yr/km <sup>2</sup> ]
Cchemuse	average agrichemical application rate over the unit drainage area [kg/yr/km <sup>2</sup> ]
Travtime	Agrichemical travel time across the unit [d] (no data available)
Losscoef	Overall loss coefficient in the unit [1/d] (no data available)
Expofact	Export factor, fraction of the applied agricultural chemical in the unit that enters the surface water (no data available)
Tcavg	Annual average temperature for upstream drainage area [°C]
Pmmavg	Annual precipitation depth, average over upstream drainage area [mm]
Flow01...Flow12	monthly flow rate that flows through the unit outlet [m <sup>3</sup> /s]
Conc01...Conc12	concentration at the unit outlet (atrazine [mg/m <sup>3</sup> ], nitrate plus nitrite as nitrogen [g/m <sup>3</sup> ])
Load01...Load12	chemical load (flow * concentration, atrazine [mg/s], nitrate plus nitrite as nitrogen [g/s])

3) File Model2a.dbf -- model specification. Table 5.7 shows the fields of this file.

**Table 5.7 Fields of the file model2a.dbf (model specification).**

Field	Description
Year	a year that is used to estimate the trend coefficient;
Ftrend	mathematical description of the trend. Current version of the model assumes no trend, i.e. the trend is described by the following equation: $1.00+(0.0000*Year)$
Trendcf	trend coefficient which is calculated according to the information stored in items Year and Ftrend
Si01...Si12	seasonal index, 12 values that represent the monthly variations of the agrichemical concentration around the annual average.
Model	model name, e.g. Nitrate01, Atrazine01, Atrazine02
Sel	contains 1 if the model is selected for calculations, 0 otherwise
Equation	equation for estimation of agrichemical concentrations. Four models are available, 1) atrazine model without the flow rate [ $\mu\text{g/L}$ ]: $(-1.5575+(0.026*U)+(0.7998*LL)+(0.4559*TA)-(0.0048*PA))*TR*SI$ 2) atrazine model with the flow rate [ $\mu\text{g/L}$ ]: $(-0.8142+(0.0133*U)+(0.3346*LL)+(38.8804*LS)+(0.2732*TA)-(0.0029*PA))*(Q^{0.2899})*TR*SI$ 4) nitrate plus nitrite as nitrogen model without the flow rate [ $\text{mg/L}$ ]: $(-7.42451+(0.001062*U)-(1.033063*TA)+(0.019339*PA))*TR*SI$ 5) nitrate plus nitrite as nitrogen model with the flow rate [ $\text{mg/L}$ ]: $(-7.57848+(0.00064886*U)+(0.776683*LL)+(173.6409*LS)-(0.520245*TA)+(0.0088545*PA))*(Q^{0.3432})*TR*SI$  Three multiplicative models are included only for testing purposes: x1) nitrate plus nitrite as nitrogen [ $\text{g/m}^3$ ]: $0.000143208*(U^{1.3814})*(A^{(-0.5556)})*(Q^{0.4240})*(LL^{0.9522})*TR*SI$ x2) for atrazine [ $\text{g/m}^3$ ]: $0.0001265546*(U^{0.8323})*(SL^{0.3591})*(Q^{0.0940})*(LL^{0.9208})*TR*SI$ x3) for atrazine, same as (x2) but different output units [ $\text{mg/m}^3$ ]: $0.1265546*(U^{0.8323})*(SL^{0.3591})*(Q^{0.0940})*(LL^{0.9208})*TR*SI$ where: U = application rate, A = drainage area, Q = flow rate, LL = average overland flow length, SL = land slope, TA = average temperature, PA = average precipitation, TR = trend coefficient, and SI = seasonal index.
Comments	comments and model description

- 4) File `Linkuse.dbf`--description of the link between counties and modeling units (1517 records). Three items describe the link: `Fips` (county FIPS code), `Unit_id` (ID number of the modeling unit or part of the modeling unit that is located in the county described by the FIPS code), and `Area_km2` (area of the unit or part of the unit that is located in the county specified by the FIPS code [ $\text{km}^2$ ]).
- 5) Coverage `Unflow` -- map of 1032 modeling units. The PAT file contains item `Unit_id`, and items `Qm196001...Qm199209`, that store the cumulative flow rate estimated for the period from January 1960 to September 1992.
- 6) Coverage `Unprec` -- map of 1033 modeling units. The PAT file contains item `Unit_id`, and items `Pm195001...Pm199306`, that store the average precipitation depth estimated for the period from January 1950 to June 1993.




The ArcView script `equat6` (Appendix B2) calculates concentrations and loads. This script first selects the record that contains the value of item `[Sel]` equal to 1. Then `equat6` extracts the value from the item `[Year]`, retrieves the equation that describes the trend from the item `[Ftrend]`, calculates the trend coefficient, and stores it in the item `[Trendcf]`.

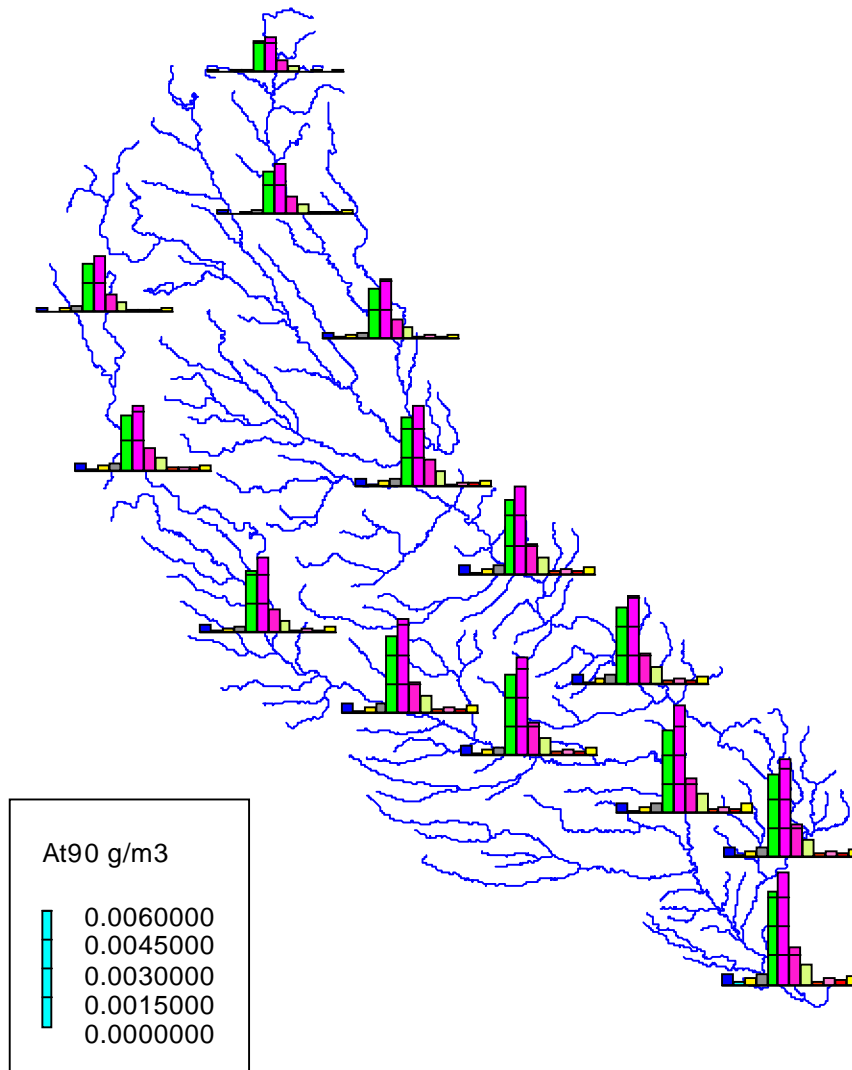
In the next step, the script extracts the equation from the item `[Equation]` and replaces the symbolic names `U`, `A`, `LS`, `LL`, and `Q` with the corresponding item names `[Cchemuse]`, `[Careakm2]`, `[Alndslp]`, `[Alndlgkm]`, and one of the item `[Qm01]...[Qm12]`. The symbolic name `TR` is replaced by the value stored in the item `[Trendcf]` and the name `SI` is replaced by the value from one of the items `[Si01]...[Si12]`. The concentrations are calculated for each month of the year by the Avenue request `calculate` and the results are stored in the items

[Conc01]..[Conc12]. The loads are estimated by multiplication of the flows and concentrations. The products are stored in the items [Load01]...[Load12].

Calculations of the agrichemical concentration in surface waters, according to the first order process, are performed by the script `decay1`. This Avenue program has been written for future extentions of the agrichemical transport model, for example, to utilize the results of the CEEPES (Comprehensive Environmental Economic Policy Evaluation System) modeling program developed by Iowa State University's Center for Agricultural and Rural Development, CARD (Bouzaher and Monale, 1993).

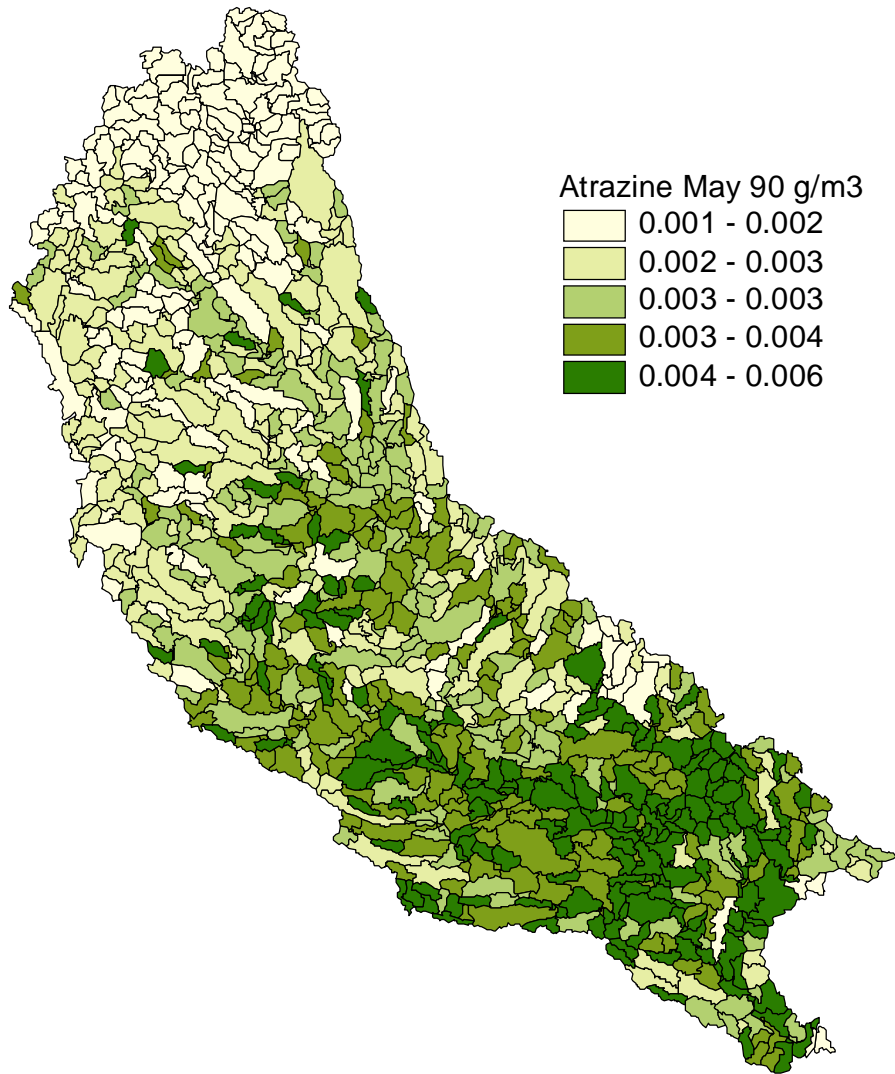
### **5.5.3 Project "Results"**

This project reserves the space for preparing the maps of estimated agrichemical concentration and load in surface water. Three scripts, executed from the button bar, , , and , draw the bar charts of the monthly average chemical concentration, chemical load and flow rate, respectively. Figure 5.13 presents the preliminary agrichemical model results of the estimation atrazine concentrations in 1990 in the form of bar charts located along the Iowa River and the Cedar River.



**Figure 5.12 “Bar chart” map of the atrazine concentrations estimated for 1990 (preliminary results)-- an example map created within the ArcView project “Results” (g/m3 = mg/L).**

The results of the calculations can be presented on the map of modeling units. Figure 5.13 shows an example of atrazine concentrations in streams for May 1990.

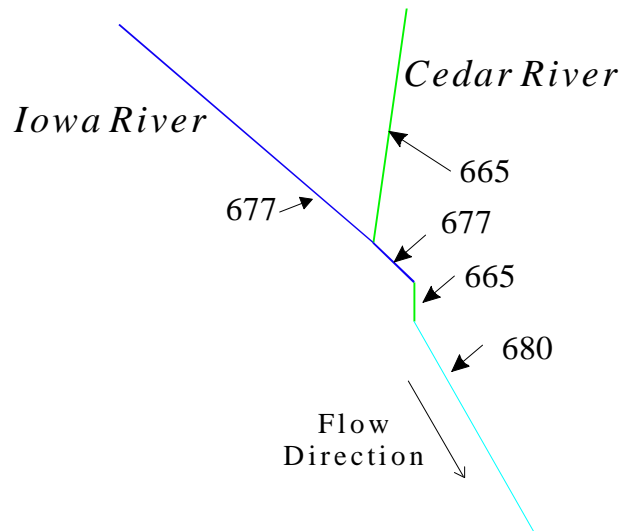


**Figure 5.13 Atrazine concentrations in the Cedar River basin introductory estimated for May 1990 -- a map created within the ArcView project "Results" (g/m<sup>3</sup> = mg/L).**

The ArcView model of the agrichemical transport represents the flow system as a network of connected modeling units. Since the Arc/Info-Grid procedures that

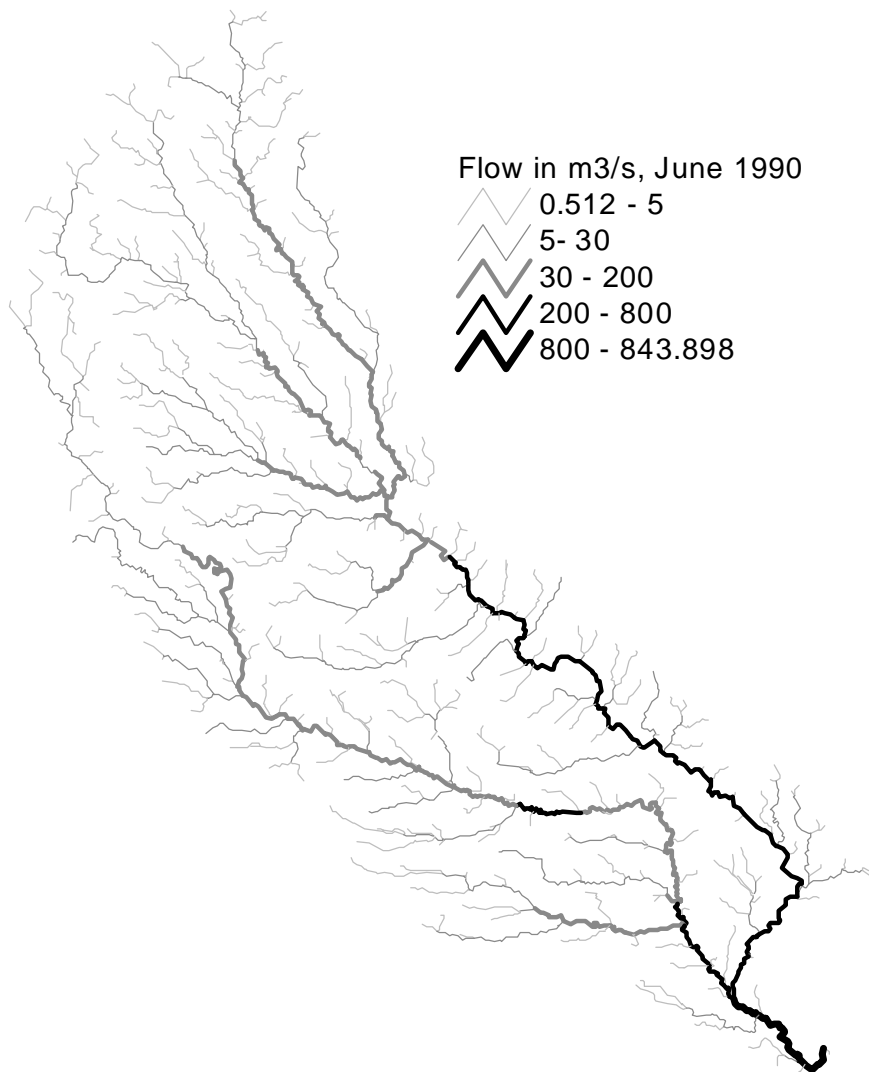


convert grid of rivers into vector format do not preserve proper numbering of the stream reaches, i.e., some reaches are split and they are identified by two or more IDs (Fig. 5.14), the modeling unit approach has been used to simplify the model.






**Figure 5.14 Inconsistency in the assigning IDs to the vectorized grid streams. The GRID function `gridline` assigned wrong IDs to the Cedar River downstream to the junction with the Iowa River.**

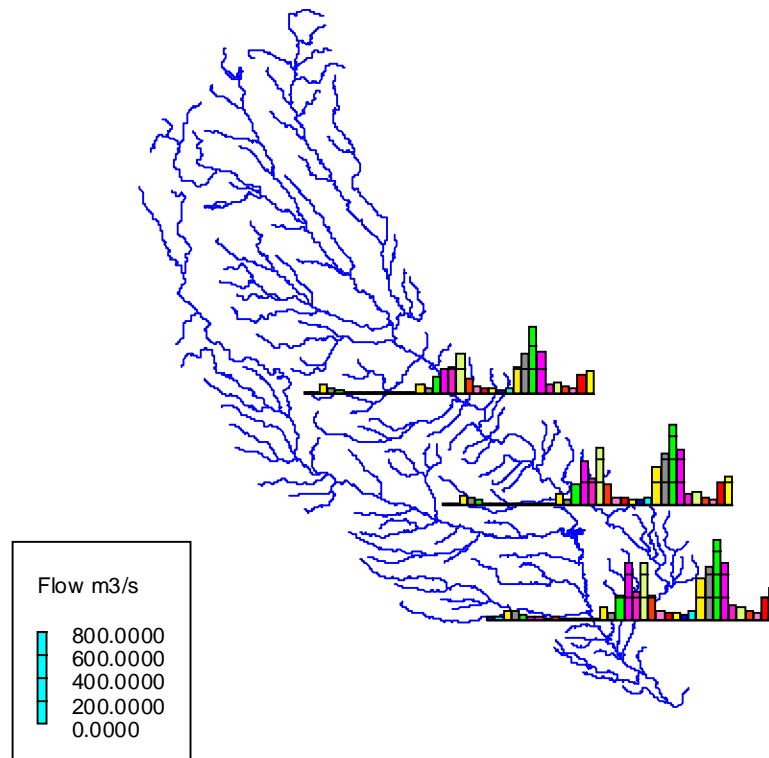
The vector representation of the stream network created by both, `streamline` and `gridline` commands can not be used for the modeling without extensive editing of the stream coverage. However, such a map of rivers can be used for data presentation purposes. Figure 5.15 shows flow rate in the streams of the Iowa-Cedar Basin created by linking the polygon attribute table of modeling unit coverage with the arc attribute table of the streams map.



**Figure 5.15** Flow rate in the Cedar River and tributaries that occurred in June 1990. The map has been created by linking the polygon attribute table of interpolated flow measurements by modeling unit with the arc attribute table of a stream coverage.

#### 5.5.4 Project “Flwprc”


The bar charts that represent the monthly flow rate and the average precipitation depth for a selected time period can be drawn at the center of selected modeling units (  draws charts of the flow rate and  draws charts of the precipitation depth). The script associated with the button  displays all recorded flow rate (monthly values from January 1960 to December 1992) for a specified modeling unit. Since charts of monthly flow rate for multiple year periods can be drawn, the wet or dry years can be easily identified. Figure 5.16 presents the monthly flow for years from 1989 to 1991.





**Figure 5.16 Monthly flow rate in the lower Cedar River in 1989-1991. Example of visual tools of ArcView project “Flwprc” for hydrologic analysis.**


### 5.5.5 Project “Tools”

The following tools have been created to support the model of agrichemicals in surface waters (Avenue scripts are included in Appendix B5):

 - Determines the unit/stream order in the flow system (Avenue script `order6`);

 - Calculates weighted average for each modeling unit total drainage area (script `upavg2`);

 - Accumulates values going along the flow path (script `cumul2`); and

 - Calculates the difference between the unit inputs and the unit output (script `decom2`).

## 6. RESULTS

According to the methodology described in Section 4 and the procedures discussed in Section 5, the models of the agrichemical concentrations in the Midwest rivers are developed in two steps. First, the seasonal components are estimated, one for the nitrate plus nitrite as nitrogen concentration and the other for the atrazine concentration. The seasonal changes of the concentrations are discussed in Section 6.1. The second step of the model development estimates a regression equation that explains the average annual concentration in the sampled rivers. The process of representing the deseasonalized agrichemical concentration by such explanatory variables as chemical application, watershed morphometry, and climatologic parameters is presented in Section 6.2. Section 6.3 discusses the differences between the modeled concentrations and the measured ones. The statistical models of agrichemical concentrations have been developed utilizing data that characterize the Upper Mississippi-Missouri and the Ohio River basins.

The precision of the spatial redistribution of the monthly flow record is tested in Section 6.4. The flow measured in the USGS stations distributed over the Iowa-Cedar River watersheds is utilized in that portion of the study. Section 6.5 describes the application of the statistical models that are incorporated into the GIS - ArcView to predict atrazine and nitrate plus nitrite concentrations as well as the chemical loads in the Iowa River, the Cedar River and in their tributaries.

For the clarity of presentation, some information from Section 4 and Section 5 is repeated here.

## 6.1 Seasonal variation of agrichemicals in the Midwest streams

All seasonal models assume each station to have an annual average concentration represented by a constant around which the amount of agrichemical oscillates seasonally. By utilizing dummy variables, created by the S-plus function `factor()`, the regression equation describes the mean annual concentration by constants that are different for each site and by a set of sine-cosine functions describing monthly variation that are common for all sampling sites. Equation 6.1 (presented in Section 4.3.6 as Eq. 4.12) represents the regression model utilized to estimate the seasonal variations of the atrazine concentration as well as the nitrate plus nitrite as nitrogen concentration. The model that includes the flow rate as an explanatory variable is described by the Equation 6.2 (Eq. 4.13).

$$\ln[c(j, d)] = w_j + \sum_{k=1}^5 (a_k \sin(2k\pi m / 12) + b_k \cos(2k\pi m / 12)) \quad (6.1)$$

and

$$\ln[c(j, d)] = w_j + a_0 \ln[Q(j, d)] + \sum_{k=1}^5 (a_k \sin(2k\pi m / 12) + b_k \cos(2k\pi m / 12)) \quad (6.2)$$

where:

$\ln[\dots]$  = natural logarithm;

$c(j, d)$  = concentration measured at site  $j$  on day  $d$  ( $\mu\text{g/l}$  or  $\text{mg/l}$ );

$Q(j, d)$  = flow rate measured at site  $j$  on day  $d$  ( $\text{m}^3/\text{s}$ );

$w_j$  = intercept specific for the  $j$ -th sampled watershed;

$a_0, a_k$  and  $b_k$  = coefficients;

$j$  = index of the sampling site;

$d$  = day of sample collection;

$k$  = harmonics number;

$m$  = month of the year.

### 6.1.1 Seasonal variation of the atrazine concentration

The following listing shows an example of the S-plus regression model specification for seasonal variation of the atrazine concentrations in surface waters of the Mississippi-Missouri River and the Ohio River basins.

**Listing 6.1 The S-plus session for estimation of the model of seasonal atrazine variation. Model does not contain the flow rate component.**

```
1: a4si _ lm ( log(Concmgm3) ~ factor(Id)
2: + + sin(2*pi*Month/12) + cos(2*pi*Month/12)
3: + + sin(4*pi*Month/12) + cos(4*pi*Month/12)
4: + + sin(6*pi*Month/12) + cos(6*pi*Month/12)
5: + + sin(8*pi*Month/12) + cos(8*pi*Month/12)
6: + + sin(10*pi*Month/12) + cos(10*pi*Month/12) , data = atra7)
```

In Listing 6.1 the results of the least square calculations are stored in an S-plus object that is named by the user `a4si`. The natural logarithm of the atrazine concentration `Concmgm3` is the dependent variable. Column `Concmgm3` in the data frame `atra7` contains concentrations in  $\mu\text{g/L}$ . The independent variables are trigonometric functions and the factorized variable `Id`. `Id` is the name of a column in the data frame `atra7` that contains sampling site identification numbers. Function `factor` tells the least square procedure `lm` to treat the variable as a factor that has  $p = 151$  levels (the number of distinct `Id` values). The  $p-1$  columns are added to the

model matrix, and then the procedure `lm` estimates a value (intercept) for each `Id` category.

Table 6.1 shows selected results of the regression analysis of the atrazine concentration in the Midwest rivers. Since only the seasonal component is important in the analysis, Table 6.1 does not contain the 151 intercept terms.

**Table 6.1 Selected coefficients of the regression analysis of the atrazine concentration in the Midwest rivers. Seasonal variation is explained only by the sine-cosine harmonics. Coefficients related to the dummy variables are not shown.**

Explanatory variable	Value	Std. Error	t value	Pr(> t )
<code>sin((2 * pi * Month)/12)</code>	-0.1174	0.0975	-1.2040	0.2289
<code>cos((2 * pi * Month)/12)</code>	-1.5085	0.1256	-12.0105	0.0000
<code>sin((4 * pi * Month)/12)</code>	-0.1655	0.1085	-1.5261	0.1274
<code>cos((4 * pi * Month)/12)</code>	0.7244	0.1191	6.0829	0.0000
<code>sin((6 * pi * Month)/12)</code>	0.0397	0.1018	0.3903	0.6964
<code>cos((6 * pi * Month)/12)</code>	-0.2813	0.1128	-2.4933	0.0129
<code>sin((8 * pi * Month)/12)</code>	-0.3014	0.0871	-3.4594	0.0006
<code>cos((8 * pi * Month)/12)</code>	-0.0256	0.1153	-0.2218	0.8245
<code>sin((10 * pi * Month)/12)</code>	0.1521	0.1003	1.5167	0.1297
<code>cos((10 * pi * Month)/12)</code>	0.0700	0.0795	0.8809	0.3787

Residual standard error: 1.021 on 772 degrees of freedom  
Multiple R-Squared: 0.7555  
F-statistic: 14.91 on 160 and 772 degrees of freedom

Table 6.2 presents the regression coefficients of the seasonal model that contains the flow rate. A similar S-plus dialog has been used to the one presented in Listing 6.1, except the independent variable `log(Flowm3s)` -- the natural logarithm of the flow rate in  $\text{m}^3/\text{s}$  -- has been added to the model specification. Listing 6.2 shows this dialog.

**Listing 6.2 The S-plus session for estimation of the model of seasonal atrazine variation. Model contains the flow rate component.**

```
1: a4siq _ lm ( log(Concmgm3) ~ factor(Id) + log(Flowm3s)
2: + + sin(2*pi*Month/12) + cos(2*pi*Month/12)
```



```

3: + + sin(4*pi*Month/12) + cos(4*pi*Month/12)
4: + + sin(6*pi*Month/12) + cos(6*pi*Month/12)
5: + + sin(8*pi*Month/12) + cos(8*pi*Month/12)
6: + + sin(10*pi*Month/12) + cos(10*pi*Month/12) , data = atra7)

```

**Table 6.2 Selected coefficients of the regression analysis of the atrazine concentration in the Midwest rivers. Seasonal variation is explained by the sine-cosine harmonics and by the flow rate. Coefficients related to the dummy variables are not shown.**

Explanatory variable	Value	Std. Error	t value	Pr(> t )
log(Flowm3s)	0.2899	0.0297	9.7696	0.0000
sin((2 * pi * Month)/12)	-0.4237	0.0972	-4.3589	0.0000
cos((2 * pi * Month)/12)	-1.3266	0.1200	-11.0542	0.0000
sin((4 * pi * Month)/12)	-0.1536	0.1024	-1.5002	0.1340
cos((4 * pi * Month)/12)	0.5409	0.1140	4.7457	0.0000
sin((6 * pi * Month)/12)	0.0401	0.0960	0.4174	0.6765
cos((6 * pi * Month)/12)	-0.2170	0.1067	-2.0337	0.0423
sin((8 * pi * Month)/12)	-0.3110	0.0823	-3.7814	0.0002
cos((8 * pi * Month)/12)	0.0200	0.1089	0.1832	0.8547
sin((10 * pi * Month)/12)	0.1359	0.0947	1.4352	0.1516
cos((10 * pi * Month)/12)	0.0970	0.0751	1.2920	0.1967

Residual standard error: 0.9638 on 771 degrees of freedom  
Multiple R-Squared: 0.7825  
F-statistic: 17.23 on 161 and 771 degrees of freedom

Regardless of the statistical significance, all harmonics have been utilized to determine the monthly seasonal factors of atrazine concentration. The seasonal factors  $S(m)$  have been calculated by the following formula (Eq. 4.11) that ensures that the average of 12 seasonal factors equals one:

$$S(m) = \frac{\exp\left(\sum_{k=1}^5 (a_k \sin(2k\pi m / 12) + b_k \cos(2k\pi m / 12))\right)}{\sum_{i=1}^{12} \exp\left(\sum_{k=1}^5 (a_k \sin(2k\pi i / 12) + b_k \cos(2k\pi i / 12))\right)} \quad (6.3)$$

where:

$S(m)$  = the seasonal factor of month  $m$ ;

$i$  = index of the month ( $i = 1, 2, \dots, 12$ );

$k$  = index of the harmonics;

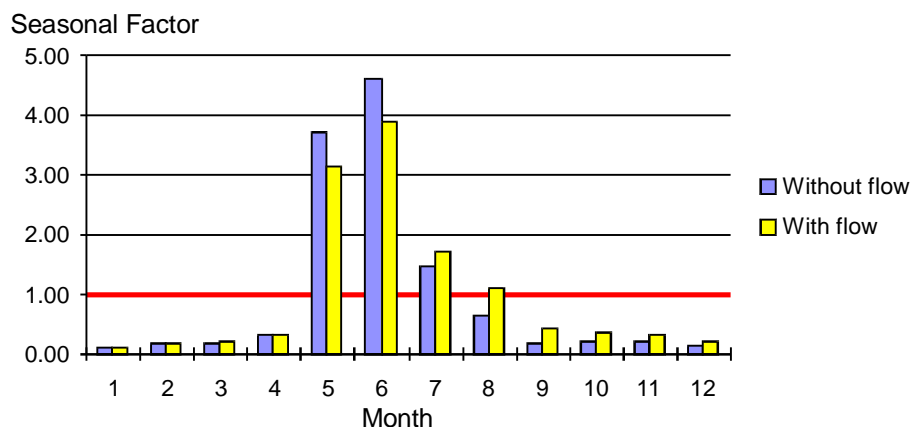
$a_k$  and  $b_k$  = regression coefficients from Table 6.1 or Table 6.2.

Table 6.3 shows non-normalized seasonal factors (described by the numerator of right hand side of Equation 6.3) as well as the normalized seasonal factors  $S(m)$  in the way that the average value over a year is equal to one (Eq. 6.3).

**Table 6.3 Seasonal factors of atrazine concentrations in the Midwest rivers estimated by the regression analysis with and without flow rate record.**

Month	Not-normalized		Normalized	
	without flow	with flow	without flow	with flow
1	0.26	0.23	0.11	0.12
2	0.41	0.36	0.17	0.19
3	0.47	0.43	0.19	0.22
4	0.77	0.60	0.32	0.31
5	9.06	6.00	3.72	3.13
6	11.23	7.44	4.61	3.88
7	3.60	3.30	1.48	1.72
8	1.55	2.09	0.64	1.09
9	0.47	0.82	0.20	0.43
10	0.51	0.72	0.21	0.37
11	0.52	0.63	0.22	0.33
12	0.36	0.41	0.15	0.21
Average	2.43	1.92	1.00	1.00

Seasonal factors of the atrazine concentration are also presented in Figure 6.1. Some trace amount of atrazine exists in the rivers throughout the year. The major transport occurs after chemical application on the field, in May and in June. The average monthly concentration in late Spring and early Summer is about 20 times higher than the average concentration in the months from September to April.



**Figure 6.1** Seasonal factors of atrazine concentrations in the Midwest rivers estimated by the regression analysis with and without flow rate record.

The seasonal factors estimated without utilizing the flow rate have slightly higher amplitude than the ones that have been calculated by the regression model with the flow rate. This indicates, that the seasonally varying flow rate is correlated with the atrazine concentration. The flow coefficient listed in Table 6.2 describes a positive relationship between the atrazine concentration and the flow rate:  $c = \dots Q^{0.2899}$ . It represents not only the relationship between flow and concentration at an individual site, but also reflects the “spatial” relationship of concentrations in rivers of different sizes.

### 6.1.2 Seasonal variation of the nitrate plus nitrite as nitrogen concentration

Listing 6.3 presents the S-plus dialog which has been applied to estimate the seasonal variations of the nitrate plus nitrite as nitrogen. The nitrate concentrations (mg/L) are stored in the in the column `Concgm3` of the S-plus data frame `nitr7`. The results of the least square procedure `lm` are written to the S-plus object `n4si`.

**Listing 6.3 The S-plus session for estimation of the model of seasonal nitrate variation. Model does not contain the flow rate component.**

```

1: n4si _ lm ( log(Concgm3) ~ factor(Id)
2: + + sin(2*pi*Month/12) + cos(2*pi*Month/12)
3: + + sin(4*pi*Month/12) + cos(4*pi*Month/12)
4: + + sin(6*pi*Month/12) + cos(6*pi*Month/12)
5: + + sin(8*pi*Month/12) + cos(8*pi*Month/12)
6: + + sin(10*pi*Month/12) + cos(10*pi*Month/12) , data = nitr7)

```

Table 6.4 shows selected results of the regression analysis of seasonal nitrate concentration changes in the Midwest rivers. The coefficients related to the `factor(Id)` (150 coefficients plus one intercept) are not shown.

**Table 6.4 Selected coefficients of the regression analysis of the nitrate concentration in the Midwest rivers. Seasonal variation is explained only by the sine-cosine harmonics.**

Explanatory variable	Value	Std. Error	t value	Pr(> t )
$\sin((2 * \pi * \text{Month})/12)$	0.8789	0.0871	10.0874	0.0000
$\cos((2 * \pi * \text{Month})/12)$	0.0892	0.1095	0.8147	0.4154
$\sin((4 * \pi * \text{Month})/12)$	0.1397	0.0964	1.4495	0.1475
$\cos((4 * \pi * \text{Month})/12)$	0.6929	0.1040	6.6653	0.0000
$\sin((6 * \pi * \text{Month})/12)$	-0.3028	0.0873	-3.4706	0.0005
$\cos((6 * \pi * \text{Month})/12)$	-0.0526	0.0974	-0.5407	0.5889
$\sin((8 * \pi * \text{Month})/12)$	-0.0318	0.0755	-0.4215	0.6735
$\cos((8 * \pi * \text{Month})/12)$	-0.2373	0.1010	-2.3494	0.0190
$\sin((10 * \pi * \text{Month})/12)$	0.2794	0.0888	3.1445	0.0017
$\cos((10 * \pi * \text{Month})/12)$	-0.0878	0.0628	-1.3986	0.1622
Residual standard error: 0.948 on 1147 degrees of freedom				
Multiple R-Squared: 0.6803				
F-statistic: 15.26 on 160 and 1147 degrees of freedom				

Nitrate plus nitrite as nitrogen model specification is displayed in Listing 6.4. The natural logarithm of the flow rate  $\log(\text{Flowm3s})$ , has been added to the model determined in Listing 6.3. The results are stored in the object `n4siq`. Estimated model parameters, excluding coefficients related to the `factor(Id)`, are displayed in Table 6.5.

**Listing 6.4 The S-plus specification of the seasonal nitrate model. The model contains the flow rate component.**

```

1: n4siq _ lm ( log(Concgm3) ~ factor(Id) + log(Flowm3s)
2: + + sin(2*pi*Month/12) + cos(2*pi*Month/12)
3: + + sin(4*pi*Month/12) + cos(4*pi*Month/12)
4: + + sin(6*pi*Month/12) + cos(6*pi*Month/12)
5: + + sin(8*pi*Month/12) + cos(8*pi*Month/12)
6: + + sin(10*pi*Month/12) + cos(10*pi*Month/12) , data = nitr7)

```

**Table 6.5 Selected coefficients of the regression analysis of the nitrate concentration in the Midwest rivers. Seasonal variations are explained by the sine-cosine harmonics as well as the flow record. Coefficients related to the dummy variables are not shown.**

Explanatory variable	Value	Std. Error	t value	Pr(> t )
log(Flowm3s)	0.3432	0.0213	16.1054	0.0000
sin((2 * pi * Month)/12)	0.5789	0.0809	7.1570	0.0000
cos((2 * pi * Month)/12)	0.3487	0.1003	3.4776	0.0005
sin((4 * pi * Month)/12)	0.1035	0.0871	1.1882	0.2350
cos((4 * pi * Month)/12)	0.5066	0.0946	5.3546	0.0000
sin((6 * pi * Month)/12)	-0.2964	0.0788	-3.7596	0.0002
cos((6 * pi * Month)/12)	-0.0136	0.0880	-0.1550	0.8769
sin((8 * pi * Month)/12)	-0.0290	0.0682	-0.4255	0.6706
cos((8 * pi * Month)/12)	-0.1488	0.0914	-1.6275	0.1039
sin((10* pi * Month)/12)	0.2397	0.0803	2.9851	0.0029
cos((10* pi * Month)/12)	-0.0326	0.0568	-0.5733	0.5666

Residual standard error: 0.8565 on 1146 degrees of freedom  
Multiple R-Squared: 0.7393  
F-statistic: 20.19 on 161 and 1146 degrees of freedom

All sine-cosine terms have been used to calculate the seasonal factors. Likewise for the atrazine model (Eq. 6.1), the exponent of the sum of harmonics, non-normalized seasonal factors, has been normalized to make the average of the seasonal

factors equal one. Table 6.6 compares the seasonal factors for two models, one without flow rate and the other with the flow rate.

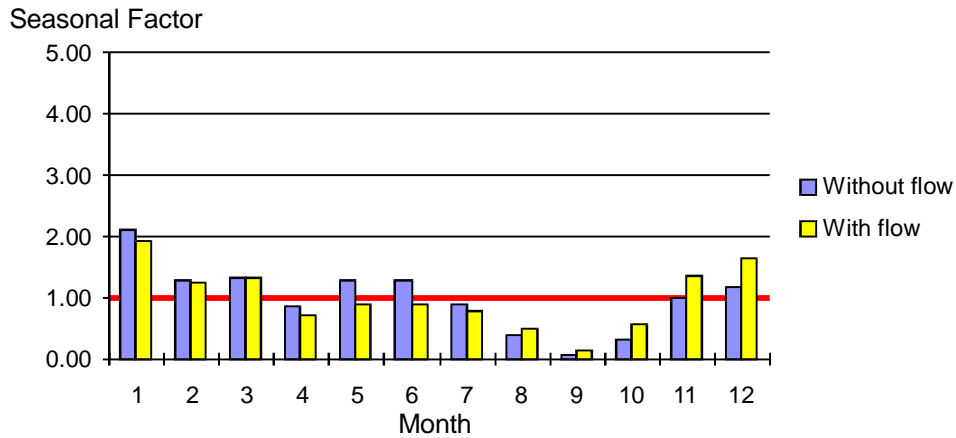
**Table 6.6 Seasonal factors of the nitrate plus nitrite as nitrogen concentrations in the Midwest rivers estimated by the regression analysis with and without flow rate.**

Month	Not-normalized		Normalized	
	without flow	with flow	without flow	with flow
1	2.69	2.31	2.10	1.94
2	1.64	1.49	1.28	1.26
3	1.70	1.58	1.33	1.33
4	1.09	0.84	0.85	0.71
5	1.64	1.05	1.28	0.88
6	1.66	1.06	1.30	0.89
7	1.14	0.95	0.89	0.80
8	0.52	0.59	0.41	0.50
9	0.09	0.17	0.07	0.14
10	0.43	0.66	0.34	0.56
11	1.28	1.62	1.00	1.36
12	1.50	1.94	1.17	1.63
Average	1.28	1.19	1.00	1.00

The seasonal variations of the nitrate plus nitrite as nitrogen in the Midwest rivers are also visualized in Figure 6.2. The variations of nitrate exhibit a different pattern than do the variations of atrazine. The amplitude of nitrate oscillations is much smaller than the one for atrazine. For the nitrate, the range of seasonal factors varies between 0.1 and 2.1, whereas the range of atrazine factors is from 0.1 to 4.6.

For the months from February to June, nitrate concentrations fluctuate very little around the annual average value. From June the concentration level decreases reaching the minimum in September (about 10% of the annual average). Then nitrate concentration increases to a maximum value, 200% of annual average, in January. The

variations of the nitrate plus nitrite concentration evaluated from the reduced model generally agree with the published patterns (for example, for Great Britain: Jones and Burt, 1993, for Slovakia: Mendel and Repa, 1994), except for April, when the concentrations are lower than the ones estimated for May, June and July. The high May - July concentrations can be explained by the late Spring (May) application of fertilizers.



**Figure 6.2** Seasonal factors of the nitrate plus nitrite as nitrogen concentrations in the Midwest rivers estimated by the regression analysis with and without flow rate included as an explanatory variable.

## **6.2 Average annual agrichemical concentration in the Midwest streams**

One hundred fifty one average concentration levels, each for one sampling site, have been assumed to calculate seasonal changes of the agrichemicals. This section

presents the results of the regression analysis of deseasonalized concentration data. The seasonal component has been removed from the concentration measurements  $c(j,d)$  by dividing the concentrations by the respective seasonal factor  $S(m)$ :

$$c_{ns}(j,d) = \frac{c(j,d)}{S(m)} \quad (6.4)$$

where:

$c_{ns}(j,d)$  = deseasonalized agrichemical concentration;

$c(j,d)$  = observed concentration at site  $j$  on day  $d$ ;

$j$  = sample site indicator;

$m$  = month of the year when the sample was collected;

$S(m)$  = seasonal factor;

A second set of the deseasonalized concentration measurements has been calculated by removing from the data a component that is explained by both the flow rate  $Q$  and the seasonal factor  $S_q(m)$ . The process of preparing agrichemical concentration observations for further analysis is described by the Eq. 6.5.

$$c_{nsq}(j,d) = \frac{c(j,d)}{S_q(m)Q(j,d)^a} \quad (6.5)$$

where:

$c_{nsq}(j,d)$  = concentration with removed seasonal and flow components;

$Q(j,d)$  = observed flow rate at  $j$ -th site, on day  $d$ ;

$S_q(m)$  = seasonal factor estimated from a model that includes flow rate;

$a$  = a coefficient estimated by the regression during seasonal changes of concentration analysis;

$d, m, j, c(j,d)$  = same as in Eq (6.4)



The deseasonalized observations have been utilized to determine the model that relates the average annual concentration level at a given location to the annual agrichemical application rate, parameters that describe the watershed upstream to that location, and to selected climatologic variables. A linear form was assumed to model the average annual atrazine and nitrate concentrations (Eq. 6.6):

$$c_{ns}(j, d) = a + bX(j) \quad (6.6)$$

where:

$c_{ns}(j, d)$  = observed, deseasonalized concentration at site  $j$  on day  $d$  (for the seasonal model with the flow rate this dependent variable is  $c_{nsq}(j, d)$ );

$a$  = intercept;

$b$  = vector of regression coefficients;

$X(j)$  = vector of explanatory variables.

**Table 6.7 Explanatory variables used to in the analysis of the deseasonalized atrazine and nitrate plus nitrite as nitrogen concentrations in the Midwest rivers.**

Variable description	Symbol	S-plus var.	Units
Agrichemical total application	$U$	Use	kg/yr
Agrichemical application rate	$A_p$	Appl	kg/km <sup>2</sup> /yr
“Decayed” stream network length	$E_S$	Decstr	e <sup>(-100km)</sup>
Average slope of the streams	$S_S$	Slpstr	-
Average travel distance from the field to the closest stream	$L_L$	Alflgkm	km
Land slope	$L_S$	Slplnd	-
Drainage area	$A$	Area	km <sup>2</sup>
Average annual temperature at sampled site	$T$	Tc	°C
Average annual temperature over sampled watershed	$T_{avg}$	Tcavg	°C
Annual precipitation depth at sampled site	$P$	Pmm	mm
Annual precipitation depth over sampled watershed	$P_{avg}$	Pmmavg	mm

Table 6.7 shows the explanatory variables used to develop these models. The notation of the predictor variables shown in column “Symbol” is compatible with the notation that has been introduced in Section 4, whereas the column “S-plus var.” contains names of variables used in S-plus sessions.

The following Section 6.2.1 discusses selected models of the deseasonalized (average annual) atrazine concentrations whereas Section 6.2.2 presents models of the deseasonalized nitrate concentrations.

### 6.2.1 Average annual atrazine concentration in the Midwest rivers

The following S-plus stepwise variable selection procedure has been used to determine the variables and their coefficients to explain the average annual atrazine concentration in rivers studied:

```
stepcs1 _ step ( csnq.lm, ~ Area + Appl + Flowm3s + Decstr + Slpstr  
+ Slplnd + Alflgkm + Tc + Tcavg + Pmm + Pmmavg )
```

where: `stepcs1` is the object in which the results are stored and `step` is the S-plus stepwise regression procedure. The object `csnq.lm` contains results of the least squares analysis of the simplest model (the simplest model is composed only of an intercept). It is created by the following command:

```
csnq.lm _ lm ( csnq ~ 1, data = atra8)
```

Selected explanatory variables, the regression coefficients as well as their significance are presented in Table 6.8.

**Table 6.8 Results of the stepwise regression analysis of average annual atrazine concentration in the Midwest rivers (Data = atrazine concentration with removed seasonal component).**

Variable	Value	Std. Error	t value	Pr(> t )
(Intercept)	-1.5575	1.2660	-1.2303	0.2189
Appl	0.0260	0.0120	2.1785	0.0296
Alflgkm	0.7998	0.2651	3.0171	0.0026
Tcavg	0.4559	0.1166	3.9093	0.0001
Pmmavg	-0.0048	0.0014	-3.3401	0.0009

Residual standard error: 4.002 on 928 deg. of freedom  
Multiple R-Squared: 0.04599  
F-statistic: 11.18 on 4 and 928 degrees of freedom

The model has a very low  $R^2$  which raises questions about the application of the regression equation to explain the average concentration. The low  $R^2$  is partially a result of applying daily deseasonalized concentrations to estimate the annual average concentration. Even after the seasonal component was removed, the deseasonalized daily concentration varies significantly, for example, the deseasonalized concentrations in the Sangamon River at Monticello, Illinois, vary from 0.2  $\mu\text{g/L}$  to 19  $\mu\text{g/L}$  (mean =1.8, standard deviation = 2.7), or those in the West Fork Big Blue River near Dorchester, Nebraska, vary from 0.08  $\mu\text{g/L}$  to 8.7  $\mu\text{g/L}$  (mean =3.2 , st. dev. = 2.0),

The other reason for the low variance explained is that only 5 stations have data available for period longer than 3 months. The majority of Midwest rivers (94%) were sampled on average three times per year, a number too small to construct a statistically sound spatial model of the average annual concentration. It must be noted, that although the reconnaissance samples were collected by depth integrating techniques at three to five locations across each stream (Thurman, et al., 1992, work cited by Scribner, et al., 1993) they represent the conditions of the stream only at the time the sample was taken. The herbicide concentration in a river after application

during runoff can change significantly in a short period of time. For example, the atrazine concentration in the Old Mans Creek, Iowa, increased during one day 05/16/1996, from 0.57  $\mu\text{g/L}$  ( $Q = 256$  cfs, time = 00:15) to 6.2  $\mu\text{g/L}$  ( $Q = 307$  cfs, time = 6:15), and then to 47  $\mu\text{g/L}$  ( $Q = 304$  cfs, time = 22:15) (Scribner, et al., 1993). Thus, characterizing the average annual river conditions by three samples can not be supported by a good summary statistics.

Despite the low  $R^2$  the model is further analyzed. The coefficients have expected signs. Increase in atrazine application causes an increase of the atrazine concentration in rivers. The longer the average distance of the overland flow, the less dense river network is, and the higher are the concentrations in rivers. This is a result of atrazine accumulation when it travels from a field to the surface water.

The positive coefficient for temperature indicates that in “warmer” regions, where more agricultural activity is performed, the rivers are more polluted. Thus, in such regions higher river pollution may be expected than in colder watersheds.

The negative relationship between atrazine concentration and the annual precipitation depth suggests, that considering average annual conditions, the rainfall “dilutes” polluted water. Regions that have smaller annual precipitation tend to have higher atrazine concentrations. On the other hand, if a single event is considered, the rainfalls that occur in a short time after atrazine application cause a positive relationship between the river flow rate and the atrazine concentrations, which is indicated by the regression analysis of seasonal variations discussed in Section 6.1.1.

A similar analysis to the one presented above has been performed for the atrazine concentrations with the seasonal cycle removed as well as the flow related components. The atrazine application rate, despite its statistical insignificance, was forced into the equation selected by the stepwise regression analysis. The coefficients of the final model are listed in Table 6.9.

**Table 6.9 Results of the regression analysis of average annual atrazine concentrations in the Midwest rivers (Data = atrazine concentration with removed component explained by the seasonal factor and the flow rate).**

Variable	Value	Std. Error	t value	Pr(> t )
(Intercept)	-0.8142	0.9148	-0.8901	0.3737
Appl	0.0133	0.0087	1.5353	0.1251
Slplnd	38.8804	16.8116	2.3127	0.0210
Alflgkm	0.3346	0.1751	1.9109	0.0563
Tcavg	0.2732	0.0736	3.7131	0.0002
Pmmavg	-0.0029	0.0009	-3.1999	0.0014

Residual standard error: 2.516 on 927 deg. of freedom  
Multiple R-Squared: 0.03118  
F-statistic: 5.966 on 5 and 927 degrees of freedom

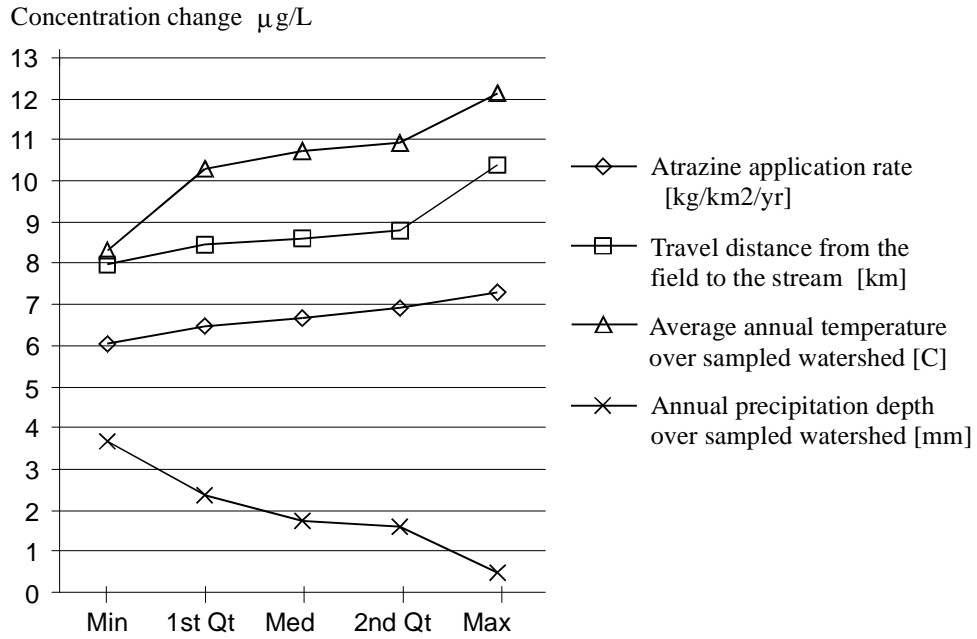
The regression included the average slope of the watershed into the model. This suggests, that the watershed slope has influence on the concentration. Indeed, it is easier to mobilize and transport agrichemical in steep slope-watersheds than in flat-watersheds.

**Table 6.10 Quartiles of the explanatory variables selected by the regression analysis of the deseasonalized agrichemical concentrations.**

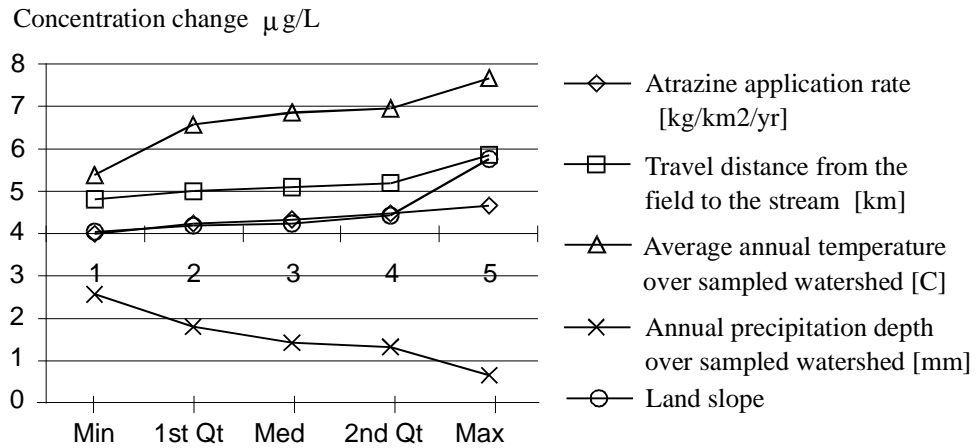
Statistics	Area	Appl	Appl	Slplnd	Alflgkm	Tcavg	Pmmavg
Statistics		atrazine	nitrogen				
	km <sup>2</sup>	kg/km <sup>2</sup> /yr	kg/km <sup>2</sup> /yr	-	km	°C	mm
Minimum	173	1.00	482	0.001	2.4	5.0	487
First quartile	961	18.08	4544	0.005	3.1	9.4	759
Median	1425	25.32	6034	0.006	3.3	10.4	886
Third quartile	3521	34.18	7372	0.011	3.5	10.7	921
Maximum	2335354	50.55	9481	0.045	5.5	13.5	1148
Average	19109	26.66	5878	0.009	3.4	10.0	848

Figure 6.3 and Figure 6.4 show the influence of the selected explanatory variables on the average concentration represented by the model described in Table 6.7 and Table 6.8, respectively. The change in concentration is related to the change in each variable assuming the minimum, first quartile (25th percentile), median (50th

percentile), third quartile (75th percentile), and maximum value. The values of the quartiles are listed in Table 6.10.



**Figure 6.3 Influence of the explanatory variables on the average atrazine concentration ( $\mu\text{g/L}$ ). Model without the flow rate (Table 6.7).**



**Figure 6.4 Influence of the explanatory variables on the average atrazine concentration ( $\mu\text{g/L}$ ). Model with the flow rate (Table 6.8).**

## 6.2.2 Average annual nitrate plus nitrite as nitrogen concentration

Parallel to the analysis of the atrazine concentrations, an analysis of the deseasonalized nitrate plus nitrite as nitrogen concentration has been conducted. The investigation was initiated by the model selected by the following S-plus stepwise regression procedure:

```
stepcs1 _ step ( csnq.lm, ~ Area + Appl + Flowm3s + Decstr + Slpstr
+               + Slplnd + Alflgkm + Tc + Tcavg + Pmm + Pmmavg )
```

where `scnq.lm` is an object that contains results of the simplest regression model

```
csnq.lm _ lm(ncsnq ~ 1 , data = n8v1).
```

The final model `atcs1` of the average annual nitrate concentration is presented in Table 6.11.

**Table 6.11 Results of the regression analysis of average annual nitrate plus nitrite as nitrogen concentrations in the Midwest rivers (Data = nitrate concentration with removed seasonal component).**

Variable	Value	Std. Error	t value	Pr(> t )
(Intercept)	-7.424541	0.8887	-8.3541	0.0000
Appl	0.001062	0.0001	17.7787	0.0000
Tcavg	-1.033063	0.0855	-12.0820	0.0000
Pmmavg	0.019339	0.0012	16.3178	0.0000

Residual standard error: 3.862 on 1304 deg. of freedom  
Multiple R-Squared: 0.3025  
F-statistic: 188.5 on 3 and 1304 degrees of freedom

The nitrate model has much better statistics than the analogous atrazine model. The coefficients for the climate variables have opposite signs to the respective coefficients in the atrazine equation. This proves that nitrate transport proceeds quite differently than atrazine transport does. The estimated inverse relationship between

nitrate concentration and the temperature shows that higher temperature enhances microbial activity and the vegetation uptake, which affects the nitrate concentration not only seasonally but also spatially (due to the climatic differences).

Tisdale, et al., (1993) pointed out that because of nitrogen mobility in soils, the greater the surplus rainfall, the greater the possibility of loss of nitrogen through leaching. This association is supported here by the estimated positive relationship between annual precipitation and the deseasonalized nitrate plus nitrite as nitrogen concentration in the Midwest rivers.

Regions with higher rainfall have greater surface runoff and greater leaching through the soil. Both, the surface flow and the groundwater transport atrazine and nitrate. Since atrazine decays, and since the groundwater transport takes months or years, the atrazine concentration in the surface waters is mainly related to surface runoff events that occur after atrazine application on the field in late spring and early summer. Thus for a single event that occurs after atrazine application, the expected relationship between precipitation and concentration is positive (except for extremely large rainfalls). But for the long period of time this relationship becomes an opposite one, since a large fraction of the river flow is from groundwater (e.g., groundwater constitutes 80% of the flow in the Cedar River, Iowa) and a high portion of the annual precipitation occurs before atrazine use. Thus, for regions with higher annual precipitation depth the lower average annual atrazine concentration can be expected if other explanatory variables are constant.

Nitrate is very a persistent chemical. It enters the river not only with the surface runoff but also it is transported by the subsurface flows that supply the river with nitrate all year around. The long-term average of the annual precipitation depth is an indicator of the magnitude of agrichemical transport by the leaching and groundwater flow.



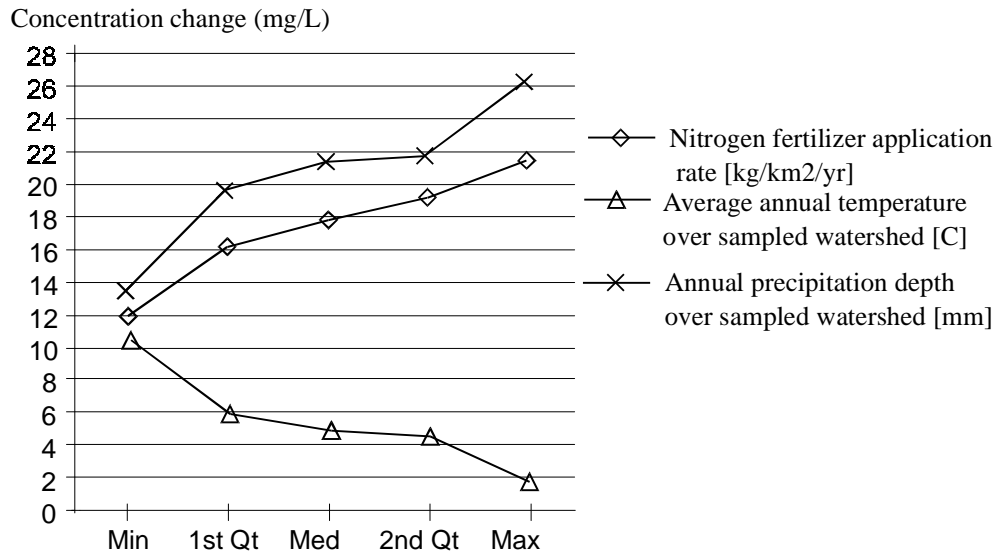
Table 6.12 shows selected variables and coefficients estimated for the model of the nitrite and nitrate as nitrogen concentration without the seasonal part and without the component explained by the measured flow rate.

**Table 6.12 Results of the regression analysis of average annual nitrate concentrations in the Midwest rivers (Data = nitrate plus nitrite as nitrogen concentrations with removed component explained by the seasonal factor and the flow rate).**

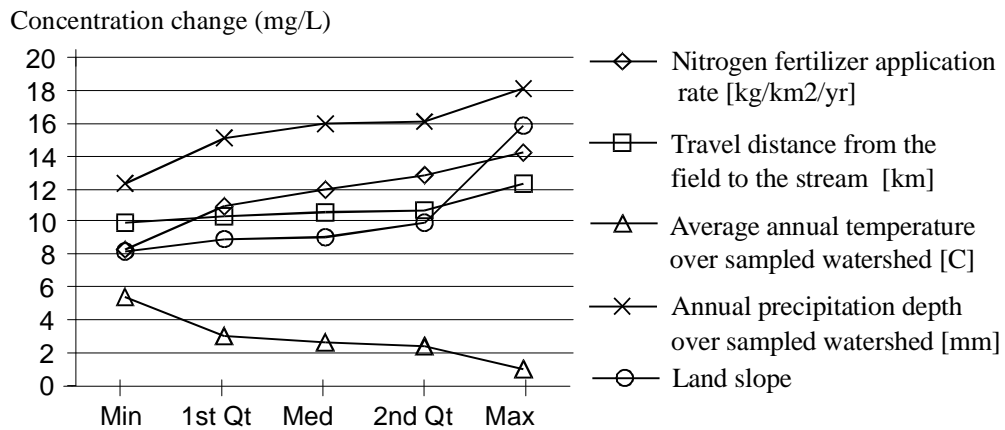
Variable	Value	Std. Error	t value	Pr(> t )
(Intercept)	-7.57848	1.0244	-7.3977	0.0000
Appl	0.00064886	0.0001	11.5591	0.0000
Tcavg	-0.520245	0.0643	-8.0874	0.0000
Pmmavg	0.0088545	0.0009	9.7573	0.0000
Slplnd	173.6409	18.3246	9.4758	0.0000
Alflgkm	0.776683	0.1674	4.6390	0.0000

Residual standard error: 2.879 on 1302 degrees of freedom  
Multiple R-Squared: 0.1636  
F-statistic: 50.95 on 5 and 1302 degrees of freedom

Figure 6.5 and Figure 6.6 show the range of the influence of the selected explanatory variables on the average concentration represented by the model described in Table 6.11 and Table 6.12, respectively. The change in concentration is related to the change in each variable assuming the minimum, first quartile (25th percentile), median (50th percentile), third quartile (75th percentile), and maximum value. The values of the quartiles are listed in Table 6.10 (Section 6.2.1). The slope of lines between first and third quartiles (Fig. 6.11 and Fig.6.12) indicates that the nitrate fertilizer application rate has relatively high influence on the nitrate concentration in the Midwest rivers.



**Figure 6.5** Influence of the explanatory variables on the average nitrate plus nitrite as nitrogen concentration (mg/L). Model without the flow rate (Table 6.11).



**Figure 6.6** Influence of the explanatory variables on the average atrazine concentration (mg/L). Model with the flow rate (Table 6.12).

### 6.3 Error of model predictions

The mean squared error (*MSE*) of the estimate for the agrichemical concentrations about the model is calculated from the following equation:

$$MSE = \sqrt{\frac{\sum (c - \hat{c})^2}{n}} \quad (6.7)$$

where:

$c$  = measured agrichemical concentration;

$\hat{c}$  = modeled concentration;

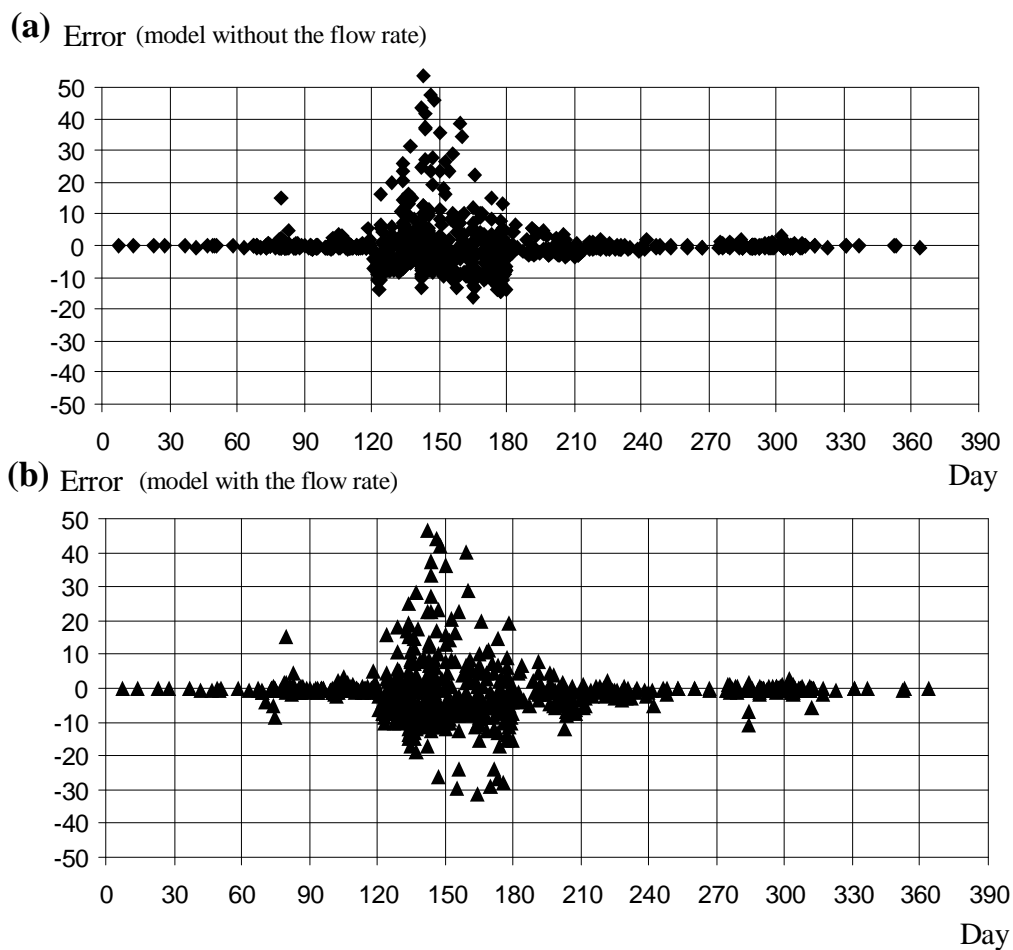
$n$  = sample size.

Figure 6.7 shows the difference between measured atrazine concentrations and the predicted ones by two models: one developed without utilizing flow rate as an independent variable, and the other one calculated utilizing recorded flow rate. Since the atrazine concentrations are high in May and the June, the prediction errors are also much higher in these month than the errors in the other months of the year. Thus two standard errors have been calculated for each model. One MSE for May and June and the other for July - April. Table 6.13 summarizes the results:

**Table 6.13. Mean Standard Errors for the atrazine concentration models ( $\mu\text{g/L}$ ).**

Model	Mean Squared Error	
	May, June	July-April
Model 1(no flow)	14.23	1.34
Model 2 (with flow)	16.55	1.92
Sample size	442	491

The relatively high error for the period from late summer to early spring is due to the high variability of measured concentrations in June (10% of sample) that often are larger than 3  $\mu\text{g/L}$ , and some as large as 16  $\mu\text{g/L}$  observations in March. For example, the concentration measured in the Auglaize River near Fort Jennings, Ohio on 03/14/89 was 15  $\mu\text{g/L}$  and on 03/21/90 was 16  $\mu\text{g/L}$ . (Scribner, et al., 1993).



**Figure 6.7** Difference between measured atrazine concentrations and predicted concentrations in the Midwest rivers:  
a) model without the flow rate;  
b) model with flow rate included as an explanatory variable.

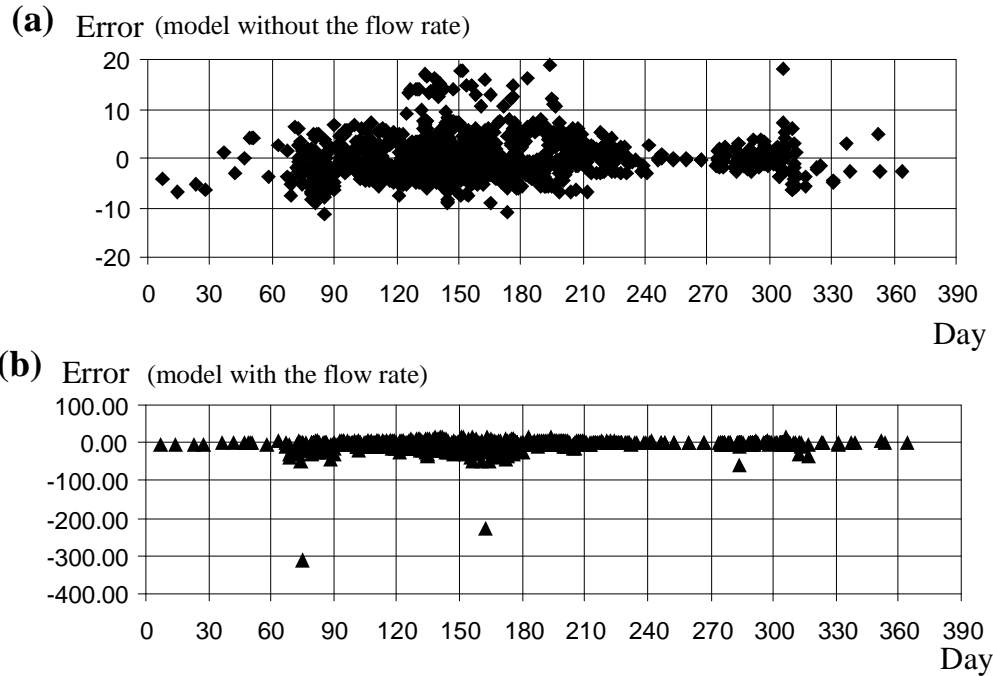
Figure 6.8 presents the difference between measured nitrate concentrations and the predicted ones. The differences do not exhibit clear seasonal variations as was the case for the atrazine models. Therefore, just one mean squared error has been calculated for each model. The errors are listed in Table 6.14.

**Table 6.14. Mean Standard Errors for the nitrate plus nitrite as nitrogen concentration models (mg/L).**

Model	Mean Squared Error
Model 1(no flow)	4.04
Model 2 (with flow)	13.43
Sample size	1308

Both Figure 6.8 and Table 6.14 indicate that the model that uses the flow rate for predictions overestimates the concentrations for high discharges. It is clearly visible for rivers with the extremely high flow events. For example, the flow of 20,500 m<sup>3</sup>/s (721,000 cfs) was recorded in the Ohio River near Grand Chain, Illinois, on 3/16/1989 and almost 9,000 m<sup>3</sup>/s (309,000 cfs) on 6/11/89. The model predicted an unrealistic concentration of 310 mg/L and 200 mg/L respectively, whereas the observed levels were less than the reporting limit of 0.1 mg/L.

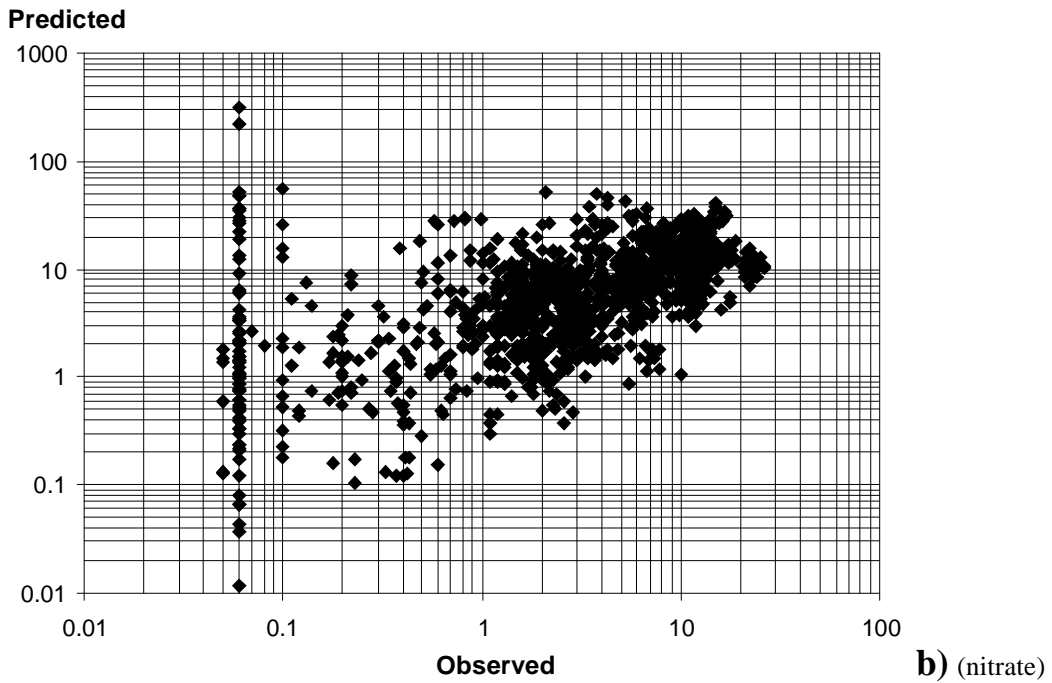
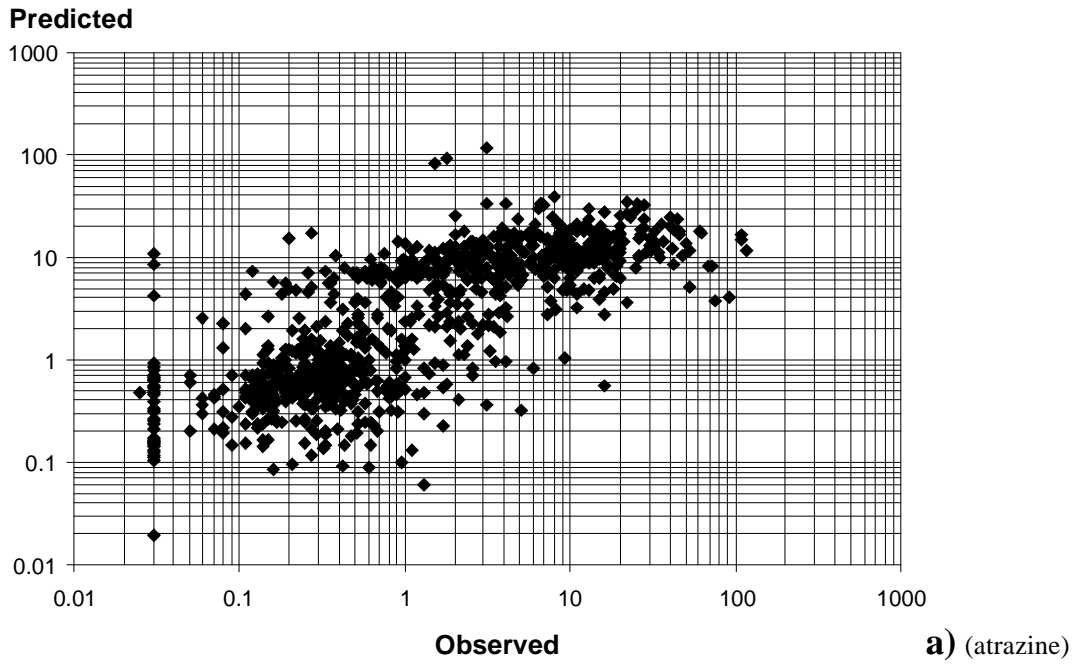
It must be noted that the standard errors are estimated using daily observations and daily predictions. The models are not intended to calculate the agrichemical in the surface waters on a daily basis but they are designed to estimate average monthly conditions. Thus such a very high flow rates as the one recorded for the Ohio River near Grand Chain on 3/16/1989 can not be used as a representation of the mean monthly flow. The errors discussed in this section serve only as a rough model verification and as an indicator of how the daily concentrations scatter around the mean monthly prediction.



**Figure 6.8** Difference between measured nitrate concentrations and predicted concentrations in the Midwest rivers:

- a) model without the flow rate;
- b) model with flow rate included as an explanatory variable.

The observed atrazine concentrations and the observed nitrate concentrations are compared to the predicted values by the model with the flow rate in Figure 6.9a and Figure 6.9b respectively. The plot of observed values vs. predicted concentrations by the model without the flow rate exhibits a similar pattern to the one shown in Figure 6.9.



**Figure 6.9** Observed vs. predicted agrichemical concentrations in the Midwest rivers (model with the flow rate): a) atrazine in  $\mu\text{g/L}$ ; b) nitrate plus nitrite as nitrogen in  $\text{mg/L}$ .

## 6.4 Comparison of predicted flow with observed one

Verification of the method of spatial redistribution of the recorded flow rate is performed within the Iowa-Cedar River basin. Three USGS gauging stations: (a) Shell Rock R. near Northwood, IA (5459000) (b) Fourmile Cr. near Traer, IA (5464137); and (c) Prairie Cr. at Fairfax, IA (5464640) has not been utilized in redistribution process. Figure 6.10 shows location of these gauging stations as well as the location of selected USGS stations whose record have been used in the flow rate estimations. Table 6.15 compares the predicted flow with the observed values.

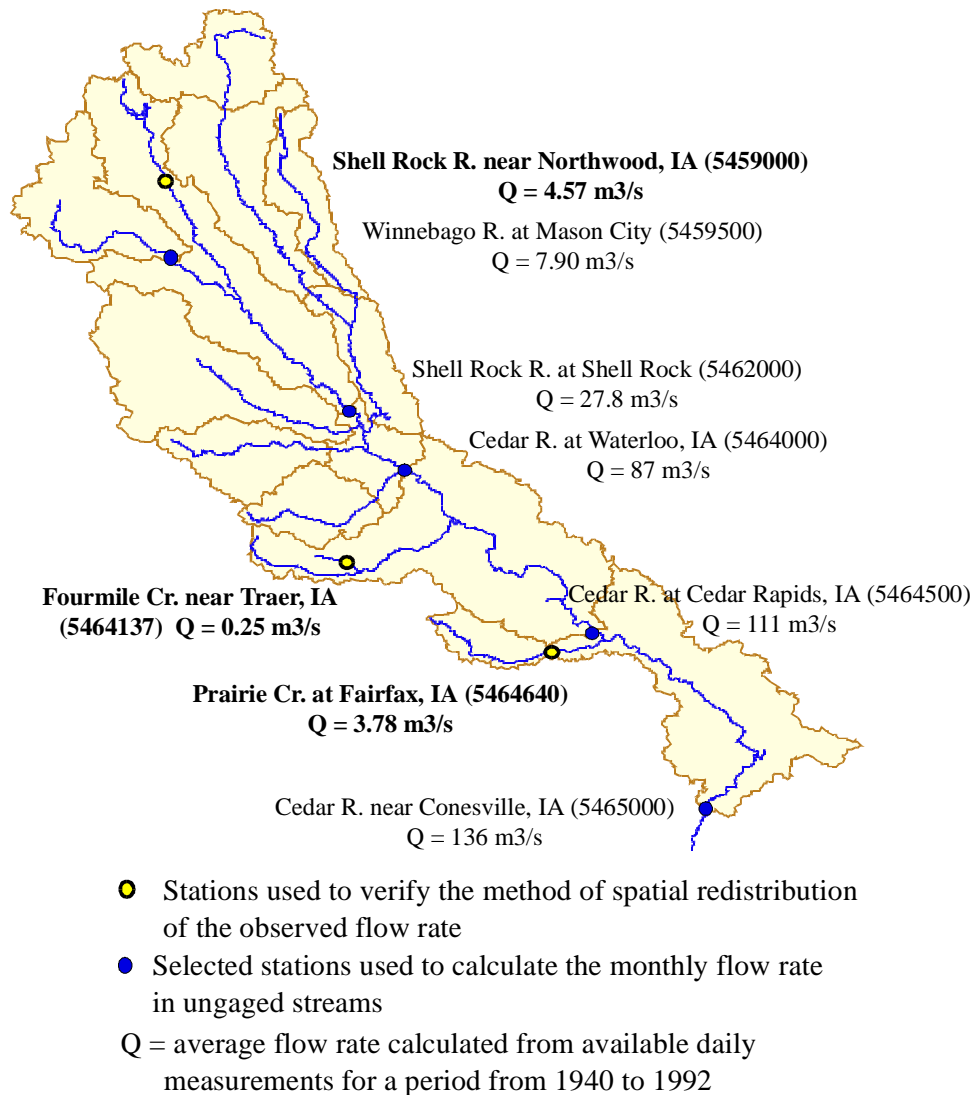
**Table 6.15 Comparison of predicted and observed flow rate (m<sup>3</sup>/s) for three USGS gauging stations.**

	Sample 1	Sample 2	Sample 3
USGS station ID	5459000	5464137	5464640
Modeling unit ID	71	388	486
Time period	1960/01-1986/09	1962/10-1980/12	1966/10-1982/9
Sample size	321.00	209.00	192.00
Mean predicted	5.37	0.48	3.66
Mean measured	5.33	0.33	3.78
Std. dev. of predicted	5.23	0.54	3.86
Std. dev. of measured	6.24	0.42	4.45
Mean difference	-0.05	-0.15	0.12
Std. dev. of difference	3.48	0.29	1.88

Although, the normality of the error has not been verified, the following normal deviates of difference between observed and estimated flow have been estimated (as an approximation of the valid statistical indicator), Sample1:  $z_1 = 0.24$ , Sample2:  $z_2 = -7.5$ , and Sample 3:  $z_3 = 0.89$ . The error is significant for Sample2. This comparison shows that the inaccuracy in predicted flow is larger for units located



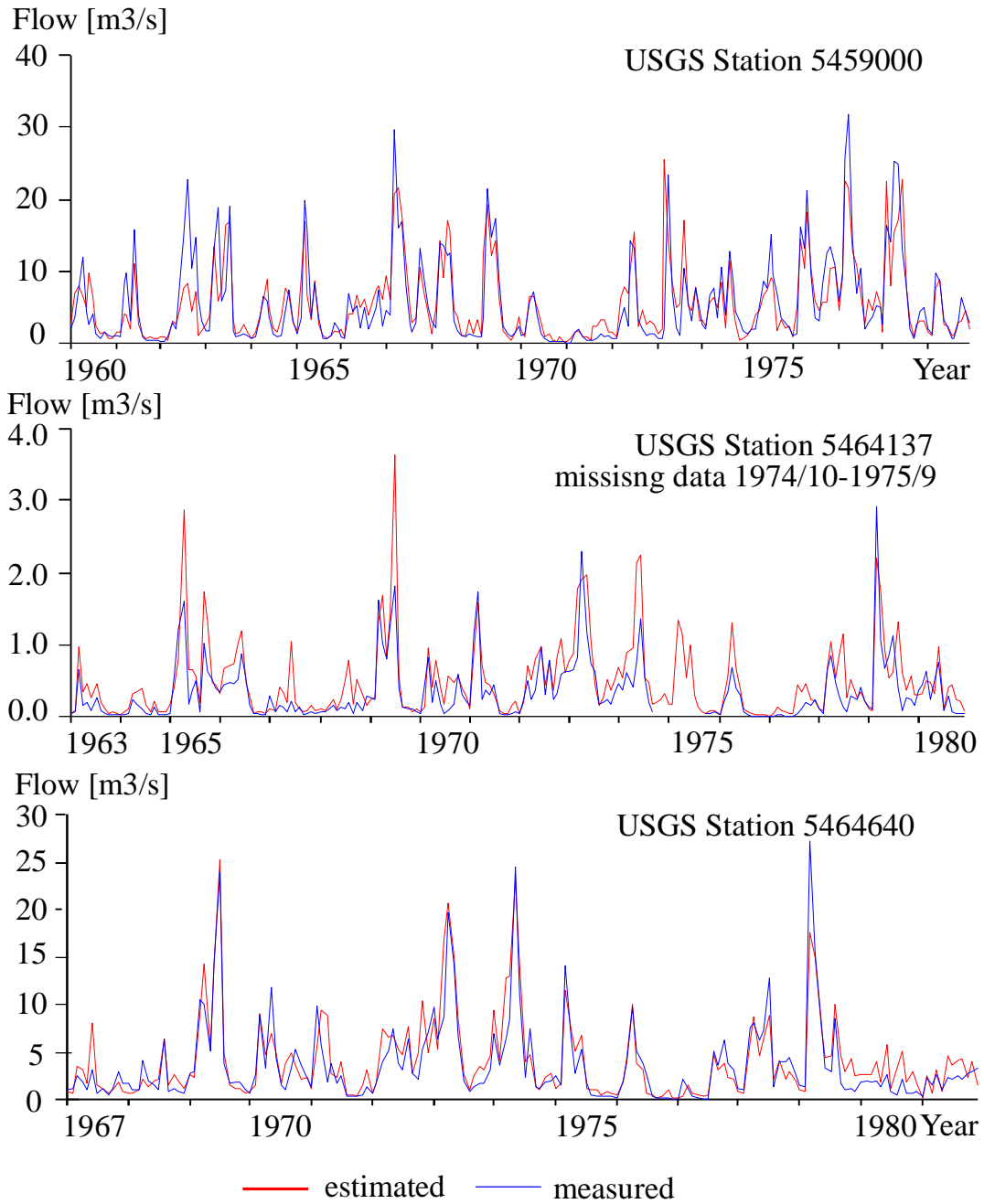
farther from gauging stations that the error for units that are closer to a measurement point.



**Figure 6.10 Location of the USGS gauging stations to verify method of spatial redistribution of the observed flow rate.**

Figure 6.11 presents the time series of the observed and predicted flow rate in three selected streams of the Cedar River basin. Although the flow redistribution

method is a very simple one, estimated monthly flow rates in ungauged streams are very close to the true values.

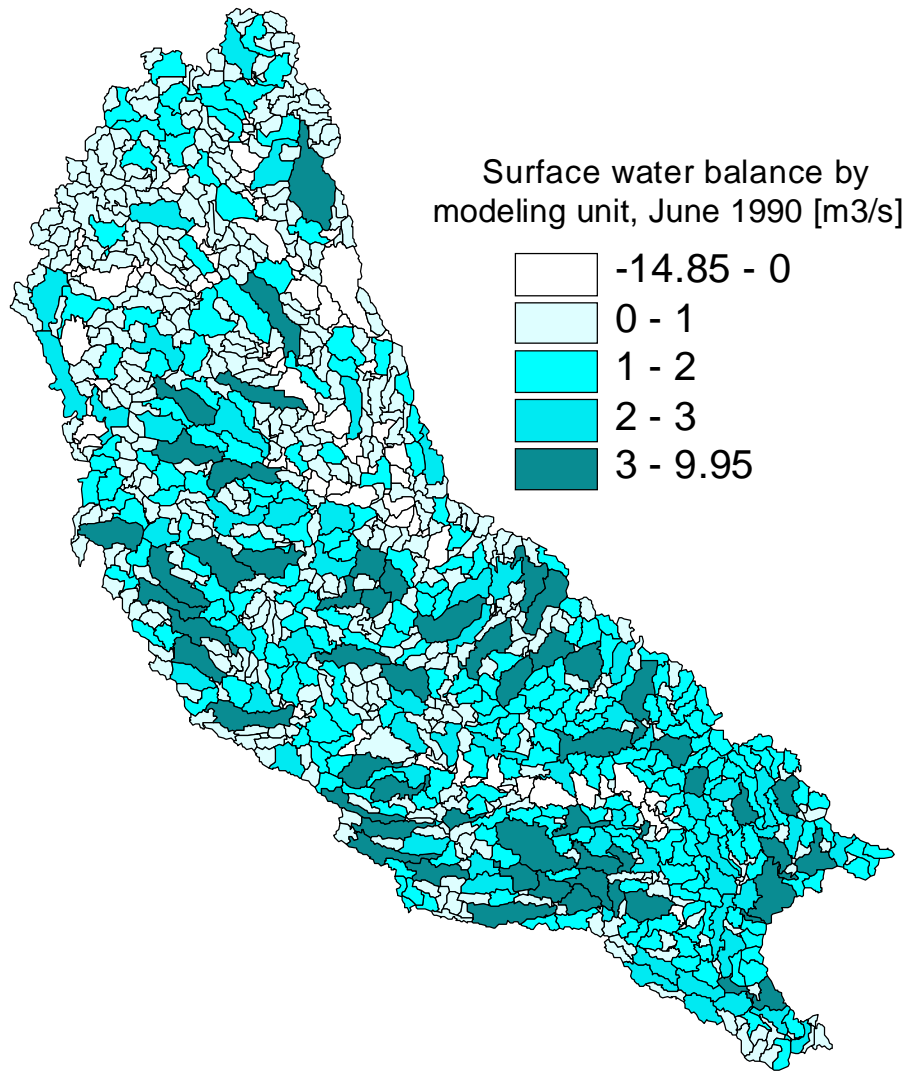


**Figure 6.11** Observed and predicted flow rate in selected streams of the Cedar River basin.

The method of spatial redistribution of the measured flow has a potential application in estimation of water losses and identification of modeling units that do not contribute to the flow in rivers. In addition, the information about the difference between outflow and inflow for each individual modeling watershed can be used to model the transport of agrichemicals between a field and the stream network. Negative balance, i.e. inflow into an unit is greater than the outflow from the unit, means water losses, occurs mainly in regions where lakes exist. It also may indicate groundwater recharge zones. Since the unit with a negative surface water balance do not contribute to the flow in the river, it is unlikely that such a unit significantly contributes to the river pollution.

On the other hand, the high positive surface water balance indicates a big contribution of the unit surface and subsurface flow to the river flow, and thus a large agrichemical contribution may be expected.

Figure 6.12 shows the estimated surface water balance for each modeling unit of the Iowa-Cedar River basin in June 1990. The ArcView script `decom2` has been applied to calculate the difference between the cumulative flow at the unit outlet and the sum of the cumulative inflows which enter the unit. The spatially distributed flow in rivers in the Iowa - Cedar River, in June 1990 (Figure 5.7), has been utilized.



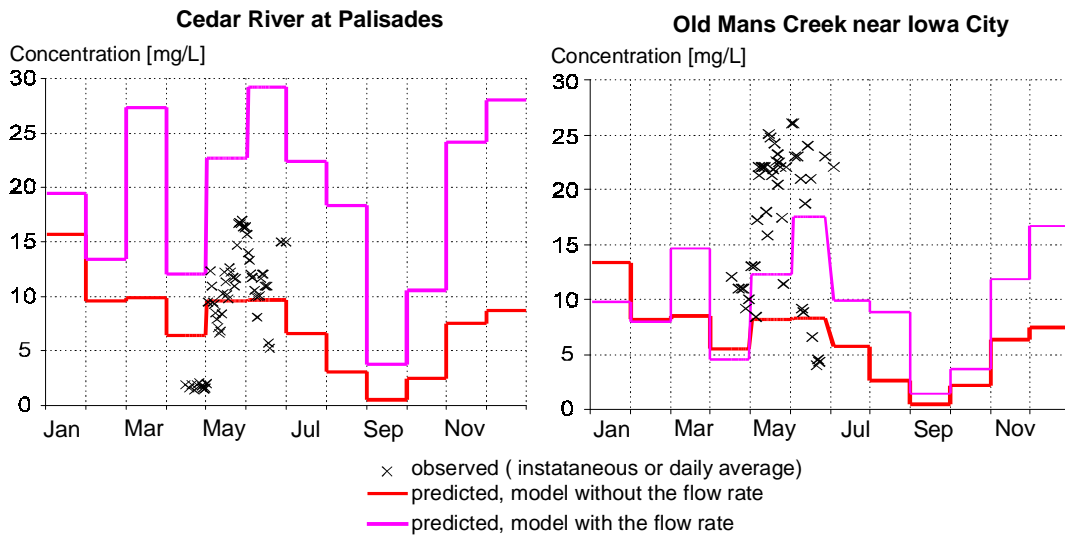
**Figure 6.12** Spatial distribution of the surface water balance for the modeling units. Estimated from recorded flow rate, precipitation depth and selected unit features, the Iowa-Cedar River basin, June 1990.

## **6.5 Agrichemical concentrations in the Cedar River basin**

This section presents the application of the methodology developed in this dissertation to estimate the agrichemical concentrations in the Iowa-Cedar River basin. The regression equations that describe the atrazine and nitrate concentration in the midwest rivers have been derived from the data gathered in more than 150 sampling points scattered over the Upper Mississippi-Missouri River and the Ohio River basins. Therefore the models reflect the regional average conditions in the rivers, and for a specific watershed, the model results may be considerably different from the available measurements.

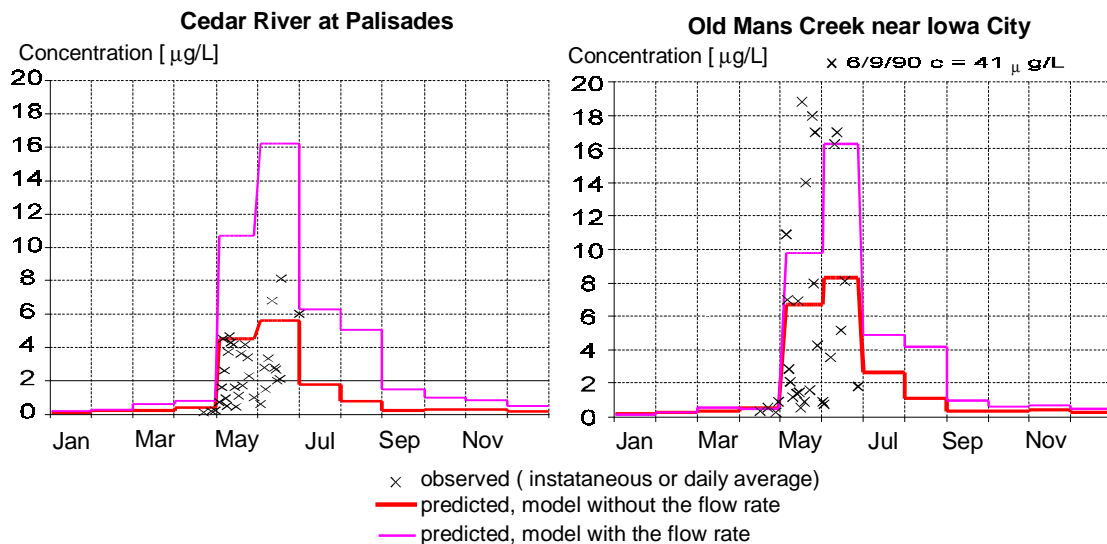
The monthly flow rate for the year 1990, as well as the annual nitrogen fertilizer use in 1990 have been extracted from the GIS database for the Iowa-Cedar Basins. After the Arc/View model performed calculations, the results for two units that represent Old Mans Creek at Iowa City (unit\_ID = 5455100) and the Cedar River at Palisades (unit\_ID = 475) have been extracted for further analysis. Figure 6.13 compares calculated concentrations of nitrate plus nitrite as nitrogen in the selected locations with the values published by the USGS (Scribner, et al., 1994). The predictions are much lower than the observations, except for the Cedar River in April and the Old Mans Creek in June, when the estimated nitrate concentrations are close to the measured ones. The plot of the measurements shows a significant increase of agrichemical at the end of April-beginning of May: from about 2 mg/L (April) to as high as 13 mg/L (May) in the Cedar River at Palisades, and from about 10 mg/L (April) to as high as 25 mg/L (May) in the Old Mans Creek near Iowa City. Most likely the jump in the concentration level was caused by the late Spring application of nitrogen fertilizers..

Modeling nitrogen in the surface waters is a very difficult task. Even such advanced model as GRASS-SWAT-QUAL2E has difficulties to make predictions that agree with the measurements. “Researchers agree that modeling nitrogen is one of the most challenging tasks even at field scale.” (Ramanarayanan, et al., 1996)



**Figure 6.13 Comparison of the predicted and observed concentrations of nitrate plus nitrite as nitrogen in the Cedar River at Palisades, Iowa and the Old Mans Creek at Iowa City, Iowa, in 1990.**

An exercise similar to the one for nitrate plus nitrogen has been performed for the atrazine. Since no herbicide usage has been available for 1990, the values estimated for the 1989 have been applied in the calculations. Figure 6.14 shows the predicted values of the atrazine concentration in the Cedar River at Palisades and in Old Mans Creek near Iowa City. The estimated values are compared with the measured concentrations in 1990 (Scribner, et al., 1994). Unlike the nitrate, the predictions of the atrazine are within the range of the observed values.



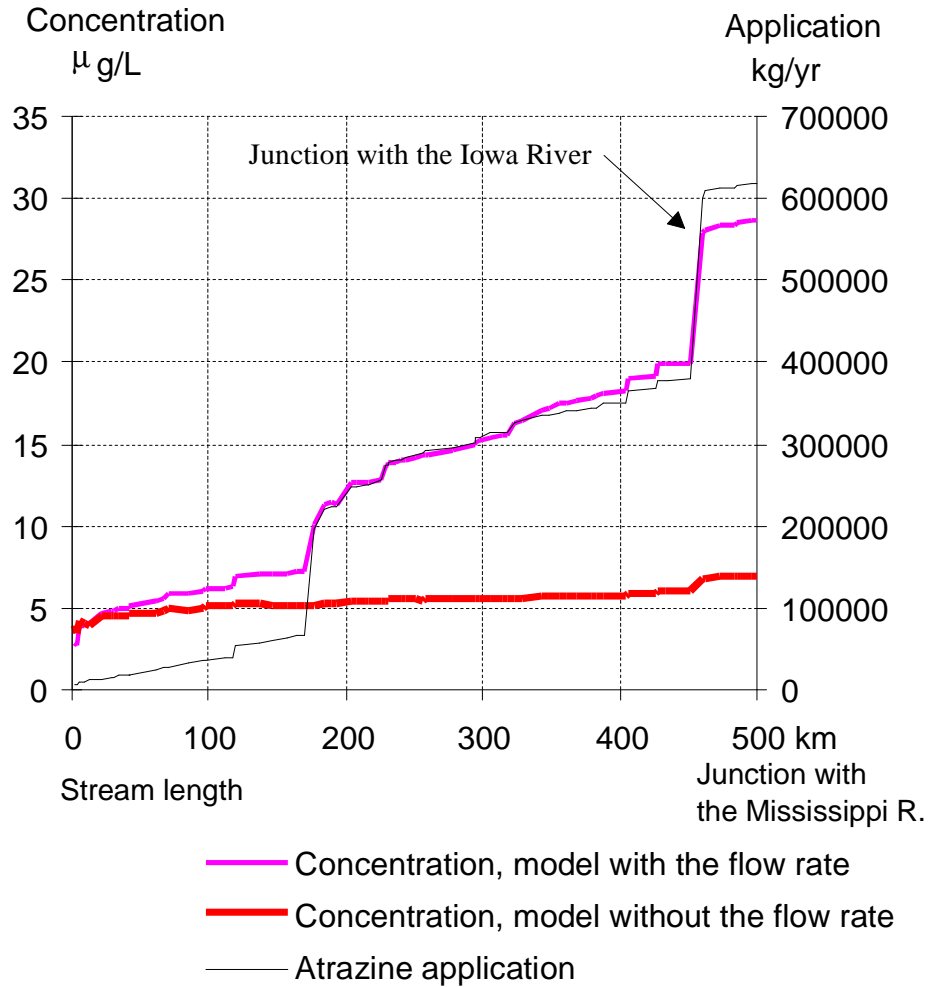
**Figure 6.14 Comparison of the predicted atrazine concentrations in the Cedar River at Palisades, Iowa and the Old Mans Creek at Iowa City, Iowa, in 1989 with the observed concentrations in 1990.**

The ArcView model of the agricultural transport in surface waters has very versatile tools to create profiles along a selected flow path. Figure 6.15 presents such profiles of the atrazine application and the atrazine concentration in the Cedar River estimated for June 1989. For selected unit watersheds that represent the Cedar River, three items have been extracted from the attribute table:

- calculated concentration for June 1989;
- annual atrazine application; and
- length of the flow path.

Moving downstream, the average increase of the herbicide mass applied to the field is about 1.25 t/km of the Cedar River (if major tributaries such as the Iowa River and the Rock Shell River are excluded from calculations, the increase of atrazine application is

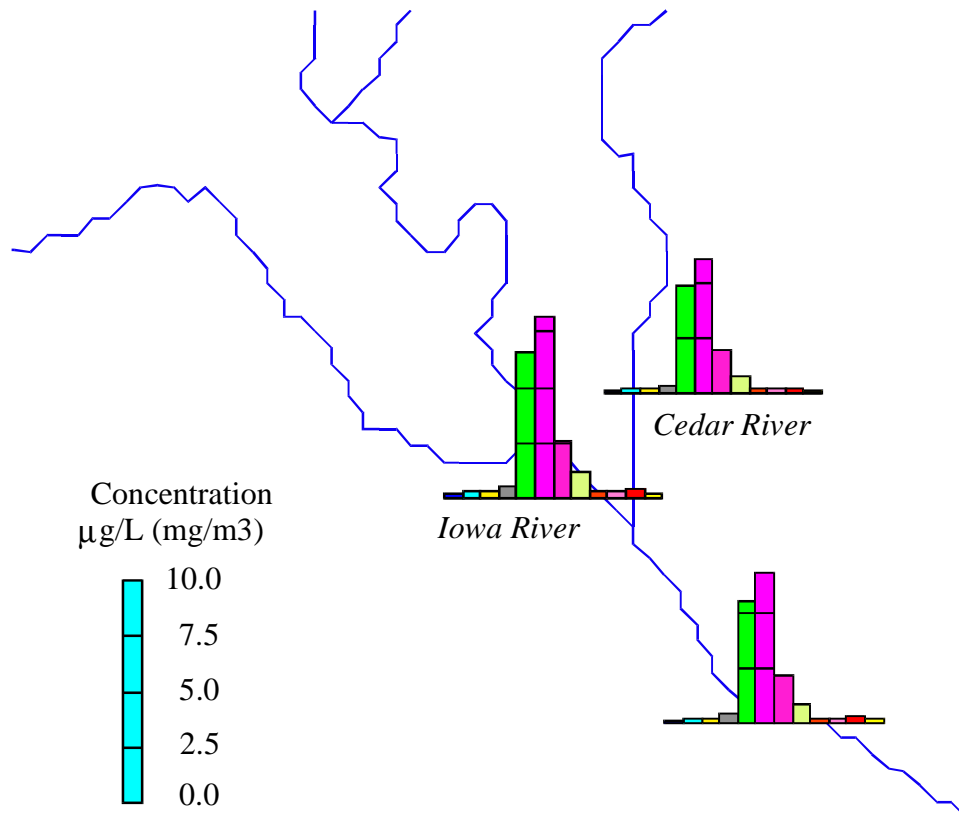
400 kg/km). The concentration in the upper portion of the river under study increases going downstream, with the rate of 1.0  $\mu\text{g/L}/100\text{ km}$ . In the downstream portion of the river, the rate decreases to 0.3  $\mu\text{g/L}/100\text{ km}$ .



**Figure 6.15 Profiles of the predicted atrazine concentrations in the Cedar River for June 1989 and the annual atrazine application (based on the results of the ArcView model).**



The model estimates the concentrations based on the agrichemical application in the upstream drainage area, and the selected parameters of the watershed. To check if the mass balance at river junctions is reasonable, i.e., if the amount of the atrazine in the Iowa River is high enough to create in the Cedar River an increase in concentration of about 0.8  $\mu\text{g/L}$  (Figure 6.15), the bar charts of the concentration for units close to the junction have been drawn utilizing the model tools (Figure 6.16). Indeed, the concentration in the Iowa River exceeds 8  $\mu\text{g/L}$ , a value that is high enough to produce concentration of 6.8  $\mu\text{g/L}$  after water from Iowa River mixes with the Cedar River.



**Figure 6.16 Predicted atrazine concentrations ( $\text{mg/m}^3 = \mu\text{g/L}$ ) at the Iowa River and the Cedar River junction (June 1990).**

Table 6.16 shows estimated values of atrazine concentration, load and the flow rate in the modeling units that represent the confluence of the Iowa River with the Cedar River.

**Table 6.16 Atrazine concentration, load and the flow rate in modeling units that represent the confluence of the Iowa River with the Cedar River estimated for June 1990.**

Parameter	Units	Unit 677 (Iowa River)	Unit 665 (Cedar River)	Unit 680 (Iowa R + Cedar R)
Concentration	µg/L	8.2	6.0	6.8
Flow rate	m <sup>3</sup> /s	301	496	800
Load	kg/month	6,375	7,764	14,143
Drainage Area	km <sup>2</sup>	11,500	20,146	31,637
Application	t/yr	222.4	379.0	601.4
Applic. rate	kg/km <sup>2</sup> /yr	19.335	18.812	19.000

Further analysis has been performed to verify if the models (with and without flow rate) can be used for making predictions of total annual chemical transport in the midwest rivers. The annual load has been calculated for two Iowa rivers, the Cedar river at Palisades and the Old Mans Creek at Iowa City. The atrazine application rate in 1989 was extracted from the USGS maps (Battaglin and Goolsby, 1995 a, b). The 1990 monthly flow rate was used to predict concentrations and loads. The calculations have been performed by the ArcView agrichemical transport model. The selected results have been extracted from the polygon attribute table of the map of modeling units and they are presented in Table 6.17.

**Table 6.17 Relation of the atrazine application to the atrazine load in two Iowa rivers: the Cedar River at Palisades and the Old Mans Creek at Iowa City ( Chemical application for year 1989, flow data for 1990).**

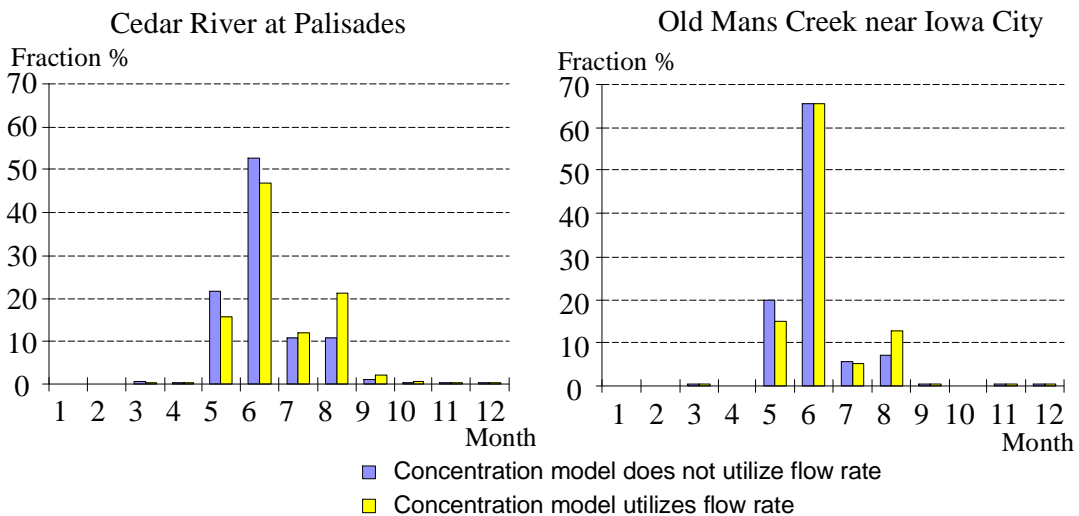
Sampling site	Application kg/yr	Model with flow rate	Estimated load kg/yr	Load as a fraction of application %
Cedar River at Palisades	325416	no	9639	3.0
		yes	31456	9.7
Old Mans Creek at Iowa City	9440	no	845	8.9
		yes	1660	17.6

The model that applies flow rate to evaluate atrazine concentration, predicts high atrazine loads: the predicted annual load for the Old Mans Creek is as high as 18% of total atrazine application and in the Cedar River it is about 10 % of herbicide use.

The model that does not utilize flow rate to calculate chemical concentrations gives smaller load estimates, 3% and 9% for the Old Mans Creek and the Cedar River respectively. These results are very close to the published agrichemical runoff from the field: Squillace and Engberg (1988) estimated that 1.5% - 4% of atrazine applied, depending on the assumed rate of chemical application, was transported by the Cedar River, Iowa in 1985. Two - three percent applied atrazine was carried by the Wye River, Maryland when substantial runoff occurred within two weeks of application (Wu, et al., 1983, cited by Squillace and Engberg, 1988). Based on the extensive review of literature on pesticides losses in runoff waters Wauchope (1978) stated that losses as high as 5 % can be expected for pesticides formulated as wettable powders (atrazine is such a pesticide), and, in addition, losses may be three times higher if a large runoff occurs about 2 weeks of application. The estimated transport of atrazine in the Mississippi River and its four tributaries (from 04/1991 to 03/1992) varied from 0.58% of use for Platte River at Louisville, Nebraska, to 1.83% in Illinois River at Valley City, Illinois (Battaglin, et al., 1993).

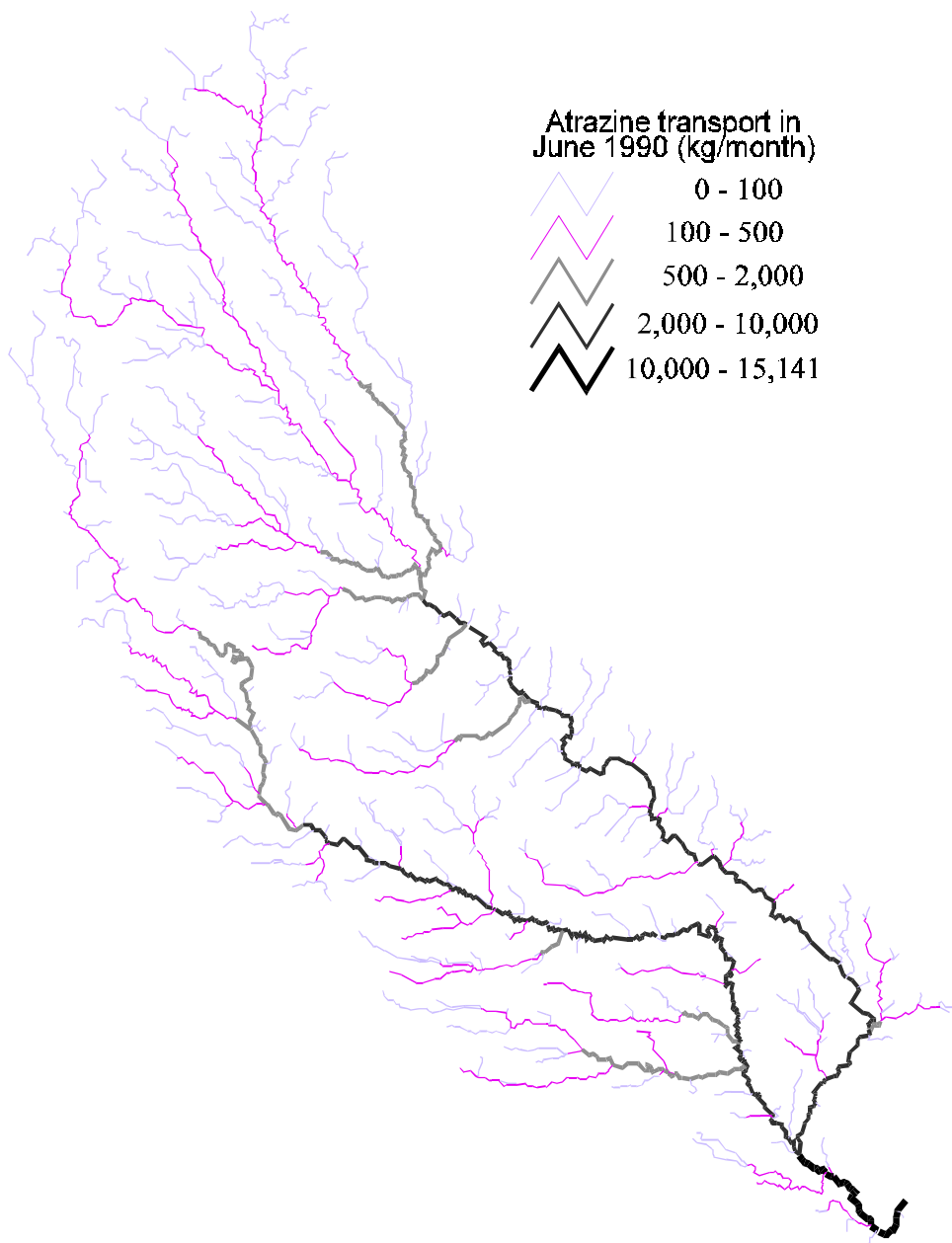
Figure 6.17 shows the monthly distribution of the loads, represented as a fraction of the annual load, in the sites under investigation. These bar charts indicate that the seasonal variation of atrazine loads is very realistic, about 70% of total atrazine load in the Cedar River at Palisades and about 80% of total load in the Old Mans Creek have occurred in May and June (average for 1990).

Squillace and Engberg (1988) estimated that 70% of the annual chemical load in the Cedar River, calculated for period from May 1984 through November 1985, occurred in June 1984. Thus, the models give realistic temporal distribution of the atrazine loads in the rivers of the Iowa-Cedar Basin.



**Figure 6.17 Estimated monthly fractions of the annual atrazine load in the Cedar River at Palisades and the Old Mans Creek near Iowa City, Iowa for year 1990.**

The spatially distributed atrazine loads in the rivers of the Iowa-Cedar Basin in June 1990 are presented in Figure 6.18. The concentration were evaluated by the “without-flow” model.



**Figure 6.18** Estimated atrazine load in the rivers of the Iowa-Cedar Basin in June 1990.

The nitrate plus nitrite as nitrogen loads were calculated for the same rivers as were the atrazine loads. The 1990 nitrogen fertilizer application has been extracted from the USGS maps of the annual nitrogen fertilizers sales in US counties (Battaglin and Golsby, 1995a). Table 6.18 presents selected results of the calculations made by the ArcView agricultural transport models.

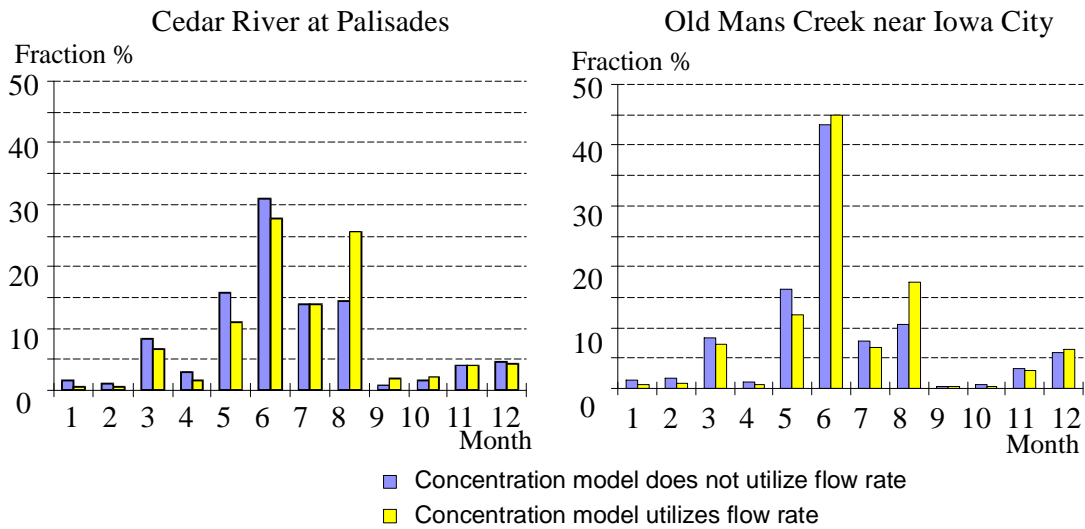
**Table 6.18 Relation of the nitrogen fertilizer application to the nitrate plus nitrite as nitrogen load in two Iowa rivers: the Cedar River at Palisades and the Old Mans Creek at Iowa City ( Chemical application and flow rate represent year 1990).**

Sampling site	Application t/yr	Model with flow rate	Estimated load t/yr	Load as a fraction of N application %
Cedar River at Palisades	123651	no	28234	22.8
		yes	95221	77.0
Old Mans Creek at Iowa City	3165	no	1267	40.0
		yes	2589	81.8

The nitrogen loads constitute large portion of the nitrogen fertilizers application. The model that utilizes the flow rate to estimate the concentrations forecasts relatively large loads. However, as shown in Figure 6.13, the model “without flow rate” predicts concentrations that are in agreement with the observed values (Cedar River at Palisades) or are lower than the observed ones (Old Mans Creek near Iowa City). Thus the estimated fractions of total chemical use are realistic. For comparison, the average transport of nitrate in the Mississippi River and its four tributaries (from 04/1991 to 03/1992) was about 15.5% of use (Battaglin, et al., 1993). It must be noted that the nitrate plus nitrite as nitrogen in Midwest rivers is not only a result of the nitrogen fertilizer application but it is also related to other sources of the nitrogen such as discharge from municipal treatment plants and manure nitrogen

inputs. For example, Coote, et al., 1978 explained 92% of the nitrate variability in the Canadian Great Lakes Basin rivers by row crops and manure nitrogen inputs.

The monthly fractions of the annual nitrate loads are shown in Figure 6.19



**Figure 6.19** Estimated monthly fractions of the annual nitrate load in the Cedar River at Palisades and the Old Mans Creek near Iowa City, Iowa for year 1990.

Although the seasonal pattern of the nitrate plus nitrite as nitrogen concentration in the midwest rivers differs from the atrazine concentration variations, the seasonal pattern of the nitrate load shown in Figure 6.19 is similar to the atrazine one. The high transport in Summer months is due to high flow rate that is recorded in the Iowa-Cedar Rivers, especially in June, July, and August.

## **7. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS**

The primary result of this work is the development of a GIS (Geographical Information System) methodology for the large scale modeling of agricultural chemical concentrations in surface waters. Two chemicals were selected for model development: a nutrient, nitrate plus nitrite as nitrogen, and a herbicide, atrazine. Statistical models which allow one to estimate concentrations in a region as large as the upper Mississippi River, the Ohio River, and their tributaries have been constructed. Two sets of models are presented in this dissertation: one for nitrate plus nitrite concentration predictions and another one for calculation of the atrazine concentrations. Each set contains a function which utilizes the flow rate, among other explanatory variables, to estimate chemical concentrations and another function which allows one to calculate concentrations without using the flow rate. All models are of a cascade type in which mean annual concentrations are derived from local watershed characteristics and the result is multiplied by a monthly factor to account for seasonal variations. Spatial variability within the region under investigation is represented indirectly by climatological parameters: mean annual temperature and mean annual precipitation depth.

The methodology developed in this research can be applied using raster GIS as well as vector GIS, however, since raster GIS requires huge computer resources to do calculations cell by cell, the final agrichemical model has been built using ArcView, a vector GIS where the cells are grouped into modeling zones.



The raster GIS, Arc/Info GRID, proved to be very convenient environment to determine a wide range of watershed characteristics using digital elevation models, and maps of agrichemical application, normal temperature and precipitation. This research revealed that it is possible to create a grid of a  $2.4 \times 10^6$  km<sup>2</sup> basin, in which each cell (here 500 m by 500 m) contains a characteristic of the drainage area upstream of this cell, such as:

- watershed morphometry (watershed area, longest stream length, average exponent of the negative stream length, average stream slope, average distance of overland flow, average land slope);
- average agrichemical application rate; and
- normal temperature and normal precipitation depth.

The resulting stack of grids of distributed watershed parameters can be used for data extraction from cells that represent sampling sites, and it can also be used after the relationship between constituent concentration and the watershed parameters has been established, to calculate agrichemical concentration in each cell for which the relationship is applicable.

GIS tools can be applied to a digital elevation model (DEM) to divide the region under investigation into hydrologic units whose size depends on such features as the diversity of terrain, density of spatial information, size of the basin, and computer resources. A method was developed in this research to divide a large basin into hydrologic modeling units and to create a system of flow connectivity between these units which allows one to build efficient hydrologic models using a concept of a set of modeling units within a stream gauge zone. This method was extensively tested for the Iowa-Cedar River watershed, Iowa, using 100 m grid resolution. It includes the following five steps:

- correction of the DEM to make both the delineated stream network and the delineated watershed boundaries compatible with the digital map of the rivers RF1 (River Reach File 1).
- stream network delineation using threshold value of 25 km<sup>2</sup> drainage area;
- development of a map of the modeling unit outlets (points in which the drainage area exceeds a threshold value, points located immediately upstream of a stream junction, and gauging station sites);
- delineating modeling units and converting them from grid representation into vector form;
- determining the flow topology between the units.

The method of adjusting the 100 m DEM by “burning in” streams improves the stream and watershed delineation. The major streams delineated from the adjusted DEM are compatible with the RF1 (River Reach File) stream reach network. This feature allows one to exchange stream attributes between RF1 and the DEM-derived watershed characteristics. “Burning in” streams was found to be the simplest, and quickest method of making a realistic digital stream network. Moreover, this method makes less changes to the original DEM compared to other methods (e.g. smoothing, spline surface fitting), which may be important for such tasks as calculation of the flood volume. A shortcoming of this method is that it can produce parallel streams, although in the Iowa River basin only two, insignificant parallel streams were noticed. Since some of the RF1 streams are close enough to be connected after they are converted into a 500 m grid, the maximum cell size for adjusting DEM using RF1 should be smaller than 500 m.

Three types of modeling unit outlets have been utilized here: cells that represent the beginning points of the stream network, cells immediately upstream of the stream junctions, and cells that represent USGS (United States Geological Survey) gauging stations. Including stream starting points in the set of modeling unit outlets makes possible the determination of the drainage area, discharge and constituent load in each node of the stream network. It also gives one more control over the average area of unit watershed. For example, in the Iowa-Cedar River basin, the average area of the modeling units determined utilizing all types of watershed outlets was 31.6 km<sup>2</sup> (threshold area for stream delineation = 25 km<sup>2</sup>) whereas excluding stream starting points from the set of watershed outlets points produced units of average area 46.7 km<sup>2</sup>.

The threshold area of 25 km<sup>2</sup> or 2500 cells for stream delineation was found to be an optimal area for dividing the Iowa-Cedar River basin into modeling units. A smaller threshold area resulted in very dense stream network and a large number of very small unit watersheds (represented by 1-2 cells) and was not justified by the spatial resolution of data used in this research. A larger threshold value than 25 km<sup>2</sup> resulted in very coarse subdivision of the region studied and a low density stream network. The stream network delineated from a 100 m DEM using 25 km<sup>2</sup> limit, was slightly more dense than the one represented by the 1:500,000 digital map of rivers, Reach File 1.

The flow topology of irregular spatial shapes such as unit watersheds can be efficiently described by assigning to each modeling unit the identification number of the downstream unit.

The GIS can be used to store and manipulate spatially distributed time series. In this research, a point coverage of the 86 National Climatic Data Center weather stations that are located within the Iowa-River basin and within the 50-km buffer zone

outside the basin was created. The attribute table of this coverage stored the monthly values of the precipitation depth for years from 1960 to 1992. These data were used to calculate the average monthly precipitation depth in 1032 modeling units by the inverse squared distance weighting procedure. A similar GIS database of the monthly flow rate was created for the 38 USGS gauging stations located in the drainage area. The attribute table of the stations point coverage contains flow time series from 1960 to 1992 (384 columns or items).

A methodology that redistributes the observed flow record over all modeling units according to the spatial distribution of precipitation and the drainage area was developed. A coverage of modeling units with an attribute table containing 32 years of monthly flow rate was created. Tests performed for three stream gauges not included in the flow redistribution process showed that the method of observed discharge interpolation/extrapolation gives very good estimates of the monthly flow in ungauged rivers. Although this method was developed for a vector representation of the watershed, it could also be applied to redistribute the flow rate using a grid representation of the study region. The procedure of the discharge calculations in ungauged rivers was programmed in both the ArcView script language Avenue and in the C language.

The approach of storing of the equations that describe the agrichemical concentration as well as the seasonal factors in a database file, instead of including them in the Avenue code, is a very efficient way of model specification. The equations can be easily updated or changed without the necessity of making changes in the computer program. The ArcView script extracts both the equations and their parameters from the database files and then calculates agrichemical concentrations and loads in the Iowa-Cedar River basin.

The credibility of the statistical model predictions is influenced by the data used for parameter estimation. The estimates of agrichemical monthly variability may be influenced by the fact that 90% of the atrazine samples were taken in the months March-July, and October. Similarly, the winter months are not well represented in the nitrate plus nitrite as nitrogen data. Only 1.2% of all nitrate samples were taken in the months from December to February. Since the majority of data represent the conditions of rivers during major runoff events after herbicide and nitrogen application, the predictions of the average monthly concentration level may be overestimated.

The regional model of the seasonal atrazine variations in Midwest rivers shows that the major transport occurs after chemical application on the field, in May and June. The average monthly concentrations estimated without considering the influence of flow rate on concentration level for months May, June and July are 3.7, 4.6 and 1.5 times higher than the annual average respectively. The atrazine concentrations in most of the remaining months of the year are less than 30% of annual average concentration.

If flow rate is included in the study of monthly concentration variation, the seasonal factors for the May, June, July, and August are respectively 3.1, 3.8, 1.7, and 1.1 times higher than the annual average. The monthly fractions for months from September to April vary between 0.2 and 0.4, except January in which the concentration is the smallest (10% of annual average).

Nitrate plus nitrite as nitrogen exhibits a different seasonal pattern compared to that for atrazine. The highest concentrations can be expected in January (twice the average annual level), and the lowest concentrations in months from August to November: 10%-50% of average. The monthly factors estimated utilizing the flow rate are slightly higher than the factors estimated without the flow rate in months from

April to July, and they are lower than the “no-flow” factors from August to December. Both atrazine and nitrate plus nitrite as nitrogen appear to increase in concentration with discharge to the 0.3 power approximately.

The average annual atrazine concentration in the Midwest rivers is described by linear functions of chemical application rate, average overland flow length, normal temperature and normal precipitation depth. In addition, the model that was developed utilizing the flow rate contains the average land slope. The spatial regression models for atrazine have very low  $R^2$  (less than 0.05, F statistic  $> 6$ ). There are two major reasons for the low variance explained by the linear models:

- 1) Highly variable daily concentration values were used to estimate the regression equations;
- 2) Ninety four percent of the Midwest rivers whose data were used, were sampled on average three times a year, a number too small to construct a statistically sound spatial model of the average annual concentrations.

It is possible that better statistics could be obtained if instead of daily, the monthly average atrazine concentrations for all months of a year were used for the regression analysis.

The average annual nitrate plus nitrite concentration in the Midwest rivers is explained by the nitrogen fertilizer application rate, normal temperature, and normal precipitation (for the model without the flow rate,  $R^2 = 0.30$ ,  $F = 188$ ), and by fertilizer application rate, land slope, average overland flow length, normal temperature, and normal precipitation (for the model with the flow rate,  $R^2 = 0.16$ ,  $F = 51$ ).

The atrazine model that uses the flow rate in its calculation of atrazine concentrations overestimates the loads. For example, for the flow conditions of 1990,

the atrazine model predicted annual load in Old Mans Creek, Iowa as 18% of the herbicide application, and for the Cedar River at Palisades, Iowa as 10 % of the application. The analogous nitrate model overestimated the concentrations measured in 1990 in the Cedar River at Palisades but it predicted reasonably well the concentrations in Old Mans Creek.

The atrazine model that does not utilize the flow rate to calculate the concentration predicted atrazine transport as 3% and 9% of the application for the Cedar River at Palisades and Old Mans Creek respectively. These numbers are similar to the values published in literature. The concentrations predicted by the nitrate model that does not utilize the flow rate, were close to the observations made in the Cedar River in May and June 1990, but they were in the lower range of the measurements in Old Mans Creek near Iowa City.

Since the coefficients of the agrichemical transport models have been determined from the data collected in streams scattered over the Midwest, the reasonable extent of the region for concentration predictions is the Upper Missouri - Mississippi - Ohio River basin (above the Ohio-Mississippi River junction).

The major findings of this research can be grouped into four categories listed below:

- **Modeling on different spatial scales**

The GIS technology was successfully applied to perform analysis on following spatial scales:

- data region (area 2,400,000 km<sup>2</sup>);
- simulation region (area 32,000 km<sup>2</sup>);
- gauged zone (average area 1,100 km<sup>2</sup>); and
- modeling unit (average area 30 km<sup>2</sup>).

- **Terrain - based modeling**

- “burning in” streams in DEM method significantly improves the stream network and the watershed boundaries delineated from a digital elevation model;
- automatic division of the simulation region into consistent system of hydrologic modeling units grouped into stream gauge zones is a useful tool supporting hydrologic modeling process;
- a distributed system of drainage area parameters allows evaluation of any watershed morphometry descriptor in all points of the region under investigation with the resolution of the DEM (here 500 m for data region, and 100 m for simulation region).

- **Spatial interpolation of streamflow**

The method used for calculating monthly discharge in ungauged streams predicted accurately the discharges in ungauged streams as shown by predicted and observed flow comparisons at three test gauges.

- **Regional modeling of agrichemicals in surface waters**

- Spatio-temporal cascade modeling can be applied within GIS to construct models agrichemical transport;
- Atrazine showed a strong seasonal pattern with high values in May and June;
- Nitrate plus nitrite as nitrogen exhibited a weak seasonal pattern gradually changing from lowest value in September to the highest value in January.
- Both constituents appeared to increase in concentration with discharge to the 0.3 power approximately;



- Regression equations of average annual atrazine concentrations had a very low explanatory power ( $R^2 < 0.05$ ), whereas equations that describe average annual nitrate concentrations were satisfactory ( $R^2 > 0.16$ );
- The negative relationship between atrazine concentration and the normal precipitation depth and the positive relationship between nitrate concentration and the precipitation suggests that the atrazine transport occurs mainly with an overland flow whereas nitrate can be also transported by infiltration and ground water flow.
- Constituent loading estimates determined using discharge-dependent concentrations appeared to be too large when compared with independent loading estimates;
- Water quality data used in this research may be biased towards processes occurring during high runoff rather than baseflow periods.

### **The recommendations for extending and modifying the work**

- expand the analysis for baseflow periods. This requires an additional set of concentration measurements which together with the data used in this research will represent average monthly conditions rather than runoff events.
- instead of using daily measurements, use average monthly concentrations and discharges to determine concentration regression equations. The average monthly concentration values need to be estimated before the spatio-temporal atrazine and nitrate pattern is estimated by regression.
- introduce an intervention function. Such a function will incorporate into the model the time of agrichemical application on the field.
- since the mean annual concentration models were unsatisfactory, additional analysis of average annual concentration is required perhaps provided by deterministic simulation models.

# Appendices

## Appendix A C-codes

This Appendix contains the following programs written in C language:

- 1) `newnx.c` - reconstructs the flow connectivity between modeling units after some of units have been removed;
- 2) `fdy4.c` -- estimates discharge in ungauged streams.

### A1 Program `newnx.c` -- reconstructing the flow connectivity between modeling units after some of units have been removed

```
/* newnx.c
* from new1.c, 9/21/95
* Pawel Mizgalewicz, CRWR UT at Austin */
/*****
* newnx.c -- reconstructs the flow connectivity between units
* after some of them have been removed. Requires two ASCII, comma
* delimited files:
*   file 1) full set of units. Each line should contain:
*           unit_id, next unit_id
*   file 2) reduced set of units. Each line should contain the
*           unit_id number
* Output file: unit_id, id of downstream unit, unit order
*
* USAGE: newnx infile1 infile2 outfile
*****/
#include <stdio.h>
main (argc, argv)
    int argc;
    char *argv[];
    /* arg 1) input_file (oid, onx => old ID, old NEXT)
```

```

        2) input_file (nid => new ID)
        3) output_file (nid, nmx, nor => ID, new NEXT, order) */
{
char loop, found;
int i, j, k, onn, nnn, nni, nor[1200], nix[1200];
long id, nx, ix,
    oid[1200], onx[1200], nid[1200], nmx[1200];
FILE *fgsin, *funin, *ftable2;
if (argc < 3 )
{
    printf ("wrong number of arguments = %d \n", argc);
    exit(1);
}
fgsin = fopen (argv[1], "r");
i = 0;
while (fscanf(fgsin, "%li,%li", &id,&nx) != EOF)
{
    oid[i] = id;
    onx[i] = nx;
    ++i;
}
onn = i-1;
fclose(fgsin);
funin = fopen (argv[2], "r");
i = 0;
while (fscanf(funin, "%li", &id) != EOF)
{
    nid[i] = id;
    nmx[i] = 0;
    ++i;
}
nnn = i-1;
nni = i;
fclose(funin);
/* To each new unit assign old_next          */
for (i = 0; i <= nnn; ++i)
    for (k = 0; k <= onn; ++k )
        if ( nid[i] ==  oid[k] )
            {
                nmx[i] = onx[k];
                break;
            }
/* check if all next are valid (have new_id specified)*/
for (i = 0; i <= nnn; ++i)
{
    found = 0;
    ix = 0;
    while (!found)
    {
        ++ix;
        for (j = 0; j <= nnn; ++j)
            if ( ( nmx[i] == nid[j] ) || ( nmx[i] == 0 ) )
                {
                    found = 1;
                    break;
                }
    }
}
}

```

```

    }
    if (found) continue;
    /* if not found, go to old table and read next_id */
    for (j = 0; j <= onn; ++j)
    if ( nmx[i] == oid[j] )
    {
        nmx[i] = onx[j];
        break;
    }
    if(ix > 1000000) printf("error %i\n", ix);
}
}
/* ===== WATERSHED ORDER ===== */
/* Set initial value of order array */
for (i = 0; i <= nnn; ++i)
{
    nor[i] = 1;
    nix[i] = nni;
}
/* Select first order items */
for (i = 0; i <= nnn; ++i)
{
    j = 0; loop = 1;
    while ( loop )
    {
        if (nmx[i] == nid[j])
        {
            nor[j] = 0;
            nix[i] = j;
            loop = 0;
        }
        ++j;
        if (j == nni) loop = 0;
    }
}
/* order of the remaining streams */
for (i=0; i<=nnn; ++i)
    if(nor[i] == 1)
    {
        j = nix[i];
        k = 2;
        while ((nor[j] < k) && (j != nni))
        {
            nor[j] = k;
            j = nix[j];
            k = k + 1;
        }
    }
}
/* Write results to output file */
ftable2 = fopen (argv[3], "w");
for (j = 0; j <= nnn; ++j )
    fprintf(ftable2, "%li,%li,%i\n", nid[j], nmx[j], nor[j] );
fclose(ftable2);
return 0;
}

```

**A2                    Program fdy4.c--calculating the flow rate in ungauged streams  
from the available record**

```

/*****
/* fdy4.c  -- calculates flow rate in all modeling units from
*           the available record. The average precipitation
*           depth is used as a weight.
* arguments:
*           1) file name that contains USGS flow record
*           2) file name that contains modeling unit data
*           3) name of the output file.
* input: two ASCII files, each record contains the following
*         values, coma delimited:
*         file one specified by the first argument:
*         gsid = USGS station identification number,
*         gsnx = ID number of the downstream USGS station,
*         gsqo = flow rate;
*         file two specified by the second argument:
*         unid = modeling unit identification number,
*         unnx = ID number of the downstream modeling unit,
*         ungsid = ID number of the USGS station that
*                 determines the zone in which the unit
*                 is located (ID of the gauging station that
*                 is downstream of the modeling unit),
*         unar = area of the modeling unit,
*         unor = modeling unit order in the flow system,
*         unpr = average precipitation depth;
* output: ASCII file specified by the third argument:
*         unid = modeling unit identification number,
*         unqc = flow rate estimated in modeling unit
*              (this flow is cumulative, i.e., it
*              is a runoff from the drainage area
*              determined by the modeling unit outlet. */
/*****

#include <stdio.h>
#define size 512

main (argc, argv)
    int argc;
    char *argv[];
    /* arg 1) input_file (gs) must have: id, next, and flow.
       2) input (modeling units--un)
          must have: id, next, gsid, area, order, precip.
       3) output_file: unid, qc (cumulative)
       used in ArcView version:
       4) name_id (copied from first column)
       5) name_out */

```

```

{
char loop, notfound;
int i, j, k, l, n, unormx, d;
int unor[2000], unixx[2000], ungsix[2000], gsused[100], or,
    gsni, gsnn, unni, unnn, unormxi;
long gsid[100], gsnx[100], id, nx, idgs,
    unid[1200], unnx[1200], ungsid[1200];
float gsqo[100], gsqi[100], gsvl[100], gscf[100],
    unar[1200], unqo[1200], unpr[1200], unqc[1200], unqct[1200],
    ar, qo, pr, rncf, sum1, sum2, xxx, x2;
/* gsor[100], gsar[100], gspr[100] deleted, gsvl[100] created */
char buf[size];
FILE *fgsin, *funin, *ftable2;
if (argc < 3 )
{
    printf ("wrong number of arguments = %d \n", argc);
    exit(1);
}
fgsin = fopen (argv[1], "r");
/* fgets(buf, size, fgsin); */
i = 0;
while (fscanf(fgsin, "%li,%li,%f", &id,&nx,&qo) != EOF)
{
    gsid[i] = id;
    gsnx[i] = nx;
    gsqo[i] = qo;
    ++i;
}
gsnn = i-1;
gsni = i;
fclose(fgsin);
/* look for records that has Q = 0. All records that has Q=0 will
have instead of next-id, the id they should have as a new unit.
*/
for ( i = 0; i <= gsnn; ++i )
{
    if ( gsqo[i] > 0.0 ) continue;
    for ( j = 0; j <= gsnn; ++j )
        if ( gsid[i] == gsnx[j]) gsnx[j] = gsnx[i];
}

/* =====*/

funin = fopen (argv[2], "r");
fgets(buf, size, funin); /* tables unload first, dummy record */
i = 0; /* that have negative area !!!! */
while (fscanf(funin, "%li,%li,%li,%f,%i,%f",
    &id,&nx,&idgs,&ar,&or,&pr) != EOF)
{
    unid[i] = id;
    unnx[i] = nx;
    ungsid[i] = idgs;
    unar[i] = ar;
    unor[i] = or;
    unpr[i] = pr;
}

```

```

    ++i;
  }
  unnn = i-1;
  unni = i;
  fclose(funin);
  /* update ungsid[i] */
  for ( i = 0; i <= gsnn; ++i )
  {
    if ( gsqo[i] > 0.0 ) continue;
    for ( j = 0; j <= unnn; ++j )
      if ( gsid[i] == ungsid[j]) ungsid[j] = gsnx[i];
  }
  /* rebuild GS arrays, i.e. remove records GSQO <= 0 */
  i = 0;
  for ( j = 0; j <= gsnn; ++j )
  {
    if ( gsqo[j] <= 0.0 ) continue;
    gsid[i] = gsid[j];
    gsnx[i] = gsnx[j];
    gsqo[i] = gsqo[j];
    gsvl[i] = 0.0;
    gsqi[i] = 0.0;      /* initialization of the inflow table */
    gsused[i] = 0;
    ++i;
  }
  gsnn = i-1;
  gsni = i;

  /* Set UNGSIX index, i.e. index that relates UNID with the GSID
  = record in UN table with the record in GS table,
  if ungsid[] = 0, i.e.,. no gauging station assigned to this
  record
  index = gsnn + 1 = gsnx      */
  for ( i = 0; i <= unnn; ++i )
  {
    if ( ungsid[i] == 0 )
    {
      ungsix[i] = gsni;
      continue;
    }
    for ( k = 0; k <= gsnn; ++k )
    {
      if ( ungsid[i] == gsid[k] )
        ungsix[i] = k;
    }
  }
  for ( i = 0; i <= gsnn; ++i )
  {
    for ( j = 0; j <= unnn; ++j )
      if ( ungsid[j] == gsid[i] )
      {
        gsvl[i] = gsvl[i] + ( unpr[j] * unar[j] );
      }
  }

```

```

/* Calculate runoff coefficient, use first order GS watersheds only
*/
sum1 = 0.0;
sum2 = 0.0;
xxx = 1.0 / 8640.0 ; /* area in km2, prec in 0.001 cm/d, -> m3/s
*/
notfound = 1;
for (i = 0; i <= gsnn; ++i)
{
    notfound = 1;
    for (j = 0; j <= gsnn; ++j )
    {
        if (gsnx[j] == gsid[i] )
        {
            notfound = 0;
            break;
        }
    }
    if ( notfound == 1 )
    {
        ++n;
        sum1 = sum1 + gsqo[i];
        sum2 = sum2 + ( gsvl[i] * xxx);
    }
}
/* printf("s1 s2 %f,%f\n",sum1,sum2); */
rncf = sum1 / sum2;

/* 1) Calculate runoff from all UN watersheds */
/* 2) assign initial values to both, unqc[] and unqct[] */
/* 3) fill array unixx[], index of the next/downstream unit */
/* is set in "find maximum order" module */

for (i = 0; i <= unnn; ++i)
{
    unqo[i] = rncf * unpr[i] * unar[i] * xxx ;
    unqc[i] = unqo[i];
    unqct[i] = unqo[i];
    for ( j = 0; j <= unnn; ++j)
    {
        if (unid[j] == unnx[i])
        {
            unixx[i] = j;
            break;
        }
    }
}

/* Find maximum UN order */
unormx = 0;
for (i = 0; i <= unnn; ++i)
{
    if ( unormx < unor[i] )
    {
        unormx = unor[i];
    }
}

```



```

        unormxi = i;
    }
}
unixx[unormxi] = unni;
/* Calculate zonal cumulative UN flow (zonal = within GS zone) */
for (k = 1; k < unormx; ++k)
{
    for (i = 0; i <= unnn; ++i)
    {
        if ( unor[i] != k )
            continue;
        j = unixx[i];
        if ( ungsid[j] != ungsid[i] )
            continue;
        unqc[j] = unqc[j] + unqc[i];
    }
}
for (i = 0; i <= unnn; ++i)
{
    unqct[i] = unqc[i];
}
/* Calculate sum of GS inflow (inflows ? */
for (i = 0; i <= gsnn; ++i)
{
    for (j = 0; j <= gsnn; ++j)
    {
        if ( gsnx[i] == gsid[j] )
        {
            gsqi[j] = gsqi[j] + gsqi[i];
        }
    }
}
/* 1) Estimate correction factors GSCF */
/* 2) Assign correction factor for the UN watersheds which are
/* outside GS zones, i.e., most downstream UN watersheds
/* (next wsh for the last unit has GSID = 0) */
for (i = 0; i <= gsnn; ++i)
{
    x2 = rncf * gsvl[i] * xxx;
    gscf[i] = ( gsqi[i] / ( gsqi[i] + x2 ) - 1.0 ) /x2 ;
    if ( gsnx[i] == 0 )
    {
        gscf[gsni] = gscf[i];
        gsid[gsni] = 0;
    }
}

/* Calculate cumulative flow (total= observed inflow + calculated
cumulative flow */
for (i = 0; i <= unnn; ++i )
{
    k = i;
    j = unixx[k];
    l = ungsix[k];

```

```

    if ( ungsid[k] == ungsid[j] ) continue;
    if ( gsused[l] == 1 ) continue;
    xxx = gsqo[l];
    gsused[l] = 1;
    while ( j < unni )
    {
        unqct[j] = unqct[j] + xxx;
        k = j;
        j = unixx[k];
        if ( ungsid[j] != ungsid[k] ) break;
    }
}
/* Final distribution of flow */
for ( i = 0; i <= unnn; ++i)
{
    j = ungsix[i];
    unqc[i] = unqct[i] * (1. + ( gscf[j] *unqc[i] ) );
}

/* Write results to output file */
ftable2 = fopen (argv[3], "w");
/* arc view version:
fprintf ( ftable2, "\"%s\", \"%s\"\n", argv[4], argv[5]);
*/
for ( j = 0; j <= unnn; ++j )
    fprintf(ftable2, "%li,%f\n", unid[j], unqc[j] );
fclose(ftable2);
return 0;
}

```

## **Appendix B Avenue scripts**

This Appendix contains the following ArcView-Avenue scripts

1) Scripts that allow user to switch between different modules of the agrichemical transport model;

gotoapr	Opens a project that is specified in the first parameter
gtflwprc	Opens the project "flwprc.apr", (maps of the flow rate and precipitation depth), property: Click
gtflwu	Opens the project "flwprc.apr", property: Update
gtmodel	Opens the project "model.apr" (data preparation and model execution), property: Click
gtmodu	Opens the project "model.apr", property: Update
gtresult	Opens the project "results.apr" (results of last model execution), property: Click
gtresu	Opens the project "results.apr", property: Update
gttools	Opens the project "tools.apr" (selected tools for hydrologic maps preparation), property: Click
gttoou	Opens the project "tools.apr", property: Update

2) Scripts from the project "model";

edflow3	Displays the dialog box to select and modify the cumulative flow rate. Changes are made for the selected (or all) polygons active on the "Modeling Units" view. Assigned to the button EQc, property: Click.
edflow3u	Assigned to the button EQc (flow rate selection and modification), property: Update.

eduse2	Runs scripts "eduse2a" and "eduse2a" that support editing the agrichemical application rate. Assigned to button EAP. Property: Click.
eduse2a	Displays the dialog box to edit agrichemical application rate. Changes will be made to selected (or all) county polygons displayed on an active view (executed from the script "eduse2")
eduse2m	Displays the dialog box to edit agrichemical application rate. Changes will be made to selected (or all) modeling unit polygons displayed on active view (executed from the script "eduse2")
eduse2u	Assigned to button EAP (edit application rate). Property: update.
equat6	Calculates concentrations and loads for all or selected months of the year. Assigned to the button RUN, property: Click.
Selmodel	Displays the dialog box to select a model: equation: $c = f(\text{Flow, Area, Time, Use, X})$ . Assigned to the button SMO, property: Click.
Stime1	Displays the dialog box to select a year ( time variable for a model). Assigned to the button SYr, property: Click.
upwavq1	Calculates weighted average for upstream units. Incorporated into scripts eduse.
selup2	Selects upstream units (active theme must have the following fields: unit_id, unit_nx, and order). Assigned to a button in category: Tools, property: Apply.
Seldown2	Selects downstream units (active theme must have the following fields: unit_id, unit_nx, and order). Assigned to a button in category: Tools, property: Apply.
Selup2u	Update event for buttons that select upstream and downstream units ("selup2", "seldown2"). Assigned to a button in category: Tools, property: Update.
decay1	Calculates the concentrations and loads of a chemical that exponentially “decays” as it is carried by water. Assigned to the button FOR, property: Click.

### 3) Scripts from the project “model”;

Cchart1	Draw charts of the concentration (12 months) at the center of selected features, property Click (runs script "schart1").
Lchart	Draw charts of the load (12 months) at the center of selected features, property Click (runs script "schart1").

Qchar1 Draw charts of the flow rate (12 months) at the center of selected features, property Click (runs script "schart1").

#### 4) Scripts from the project "flwprc";

pmchar1 Draws charts of the precipitation depth at the center of selected features, for selected months, property: Click, (executes script "schart1")

pmchar1u Draws charts of the precipitation depth, property: Update

qmchar1 Draws charts of the flow rate at the center of selected features, for selected months, property: Click, (runs script "schart1").

qmchar1u Draws charts of the flow rate, property: Update.

schart1 Draws charts, is executed from such scripts as pmchar1 and qmchar1 (Adopted from ESRII examples supplied with Avenue).

movie2 Displays charts in a sentence, category: tools, property: apply (script qmchar1u is used in the property: Update).

aliaset Sets the alias names for fields, that describe a chart axis.

#### 5) Scripts from the project "tools";

order6 Determines the order of a stream/watershed in a flow system

upavg2 calculates weighted average over upstream units

cumul2 Accumulates values, going downstream (flow accumulation)

decom2 calculates the difference between the inflows and the outflow, the reverse process to the flow accumulation

## **B1 Changing ArcView projects from the PushButton Bar.**

### **gotoapr Opens a project that is specified in the first parameter**

```
'gotoapr.ave
gotoname = self.get(0)
theProject = av.GetProject
prname = theproject.getname
if (nil <> theProject) then
  if (theProject.IsModified) then
    res = MsgBox.SaveChanges("Do you want to save changes to " +
theProject.GetName + "?", "ArcView", true)
    if (nil = res) then exit end
    if (res) then
      av.Run("Project.Save", nil)
      if (theProject.IsModified) then exit end
    end
  end
end
xxx = theproject.getfilename.asstring.substitute(prname, gotoname)
theFName = xxx.asfilename
if (nil <> theProject) then
  theProject.Close
end
av.ClearGlobals
av.PurgeObjects
Project.Open(theFName)
```

### **gtflwprc Opens the project "flwprc.apr", (maps of the flow rate and precipitation depth), property: Click**

```
xxx = av.run("gotoapr", {"flwprc.apr"})
```

### **gtflwu Opens the project "flwprc.apr", property: Update**

```
prname = av.GetProject.getname
if(prname = "flwprc.apr") then
  self.setenabled(false)
else
  self.setenabled(true)
end
```

**gtmodel**      **Opens the project “model.apr” (data preparation and model execution), property: Click**

```
xxx = av.run("gotoapr", {"model.apr"})
```

**gtmodu**      **Opens the project “model.apr”, property: Update**

```
prname = av.GetProject.getname  
if(prname = "model.apr") then  
  self.setenabled(false)  
else  
  self.setenabled(true)  
end
```

**gtresult**      **Opens the project “results.apr” (results of last model execution), property: Click**

```
xxx = av.run("gotoapr", {"results.apr"})
```

**gtresu**      **Opens the project “results.apr”, property: Update**

```
prname = av.GetProject.getname  
if(prname = "results.apr") then  
  self.setenabled(false)  
else  
  self.setenabled(true)  
end
```

**gttools**      **Opens the project “tools.apr” (selected tools for hydrologic maps preparation), property: Click**

```
xxx = av.run("gotoapr", {"tools.apr"})
```

**gttoou**      **Opens the project “tools.apr”, property: Update**

```
prname = av.GetProject.getname  
if(prname = "tools.apr") then  
  self.setenabled(false)  
else  
  self.setenabled(true)  
end
```

## B2 Scripts from the project “model”

**edflow3**      **Displays the dialog box to select and modify the cumulative flow rate. Changes are made for the selected (or all) polygons active on the "Modeling Units" view. Assigned to the button EQc, property: Click.**

```
'edflow3.ave
'=====
'data table name:    Cumulative flow
'            fields:    "Unit_id", m199001 ...
'model table name:  feature table of "units" ("Crwsd")
'            fields:    "Unit_id", "Qm01 .. Qm12"
'=====
theView = av.getproject.finddoc("Modeling Units")
FTmodl = theview.findtheme("Units").Getftab
FTflow = av.getproject.finddoc("Cumulative flow").getVtab
'FTflow = fldoc.findtheme("unflow").getftab
if ((FTmodl = nil) or (FTflow=nil)) then
    msgbox.error("Can't get ftab of units or cumulative flow",
        " Units or flow unaccessible")
    exit
end
'check if all fields can be found:
'current version of edflow1 and v2 assumes that all fields exist
sp = "xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx"
ln1 = "Assume flow conditions of:"
ln3 = "( to apply factor to current values enter: NO )"
ln4 = "Multiply flow rate for all months by:"
ln5 = "            Apply multiplication factor to each month:"
ln6 = "Multiply flow rate for January by....."
ln7 = "Multiply flow rate for February by....."
ln8 = "Multiply flow rate for March by....."
ln9 = "Multiply flow rate for April by....."
ln10 = "Multiply flow rate for May by....."
ln11 = "Multiply flow rate for June  by....."
ln12 = "Multiply flow rate for July  by....."
ln13 = "Multiply flow rate for August by....."
ln14 = "Multiply flow rate for September by."
ln15 = "Multiply flow rate for October by....."
ln16 = "Multiply flow rate for November by.."
ln17 = "Multiply flow rate for December by.."
labls = {ln1,ln3,ln4,sp,ln5,ln6,ln7,ln8,ln9,ln10,ln11,ln12,
         ln13,ln14,ln15,ln16,ln17}
defts = {"1990"," ","1.00"," ","",
         ", "1.00", "1.00", "1.00", "1.00", "1.00",
         "1.00", "1.00", "1.00", "1.00", "1.00", "1.00", "1.00", "1.00" }
```



```

tytul = "    Select flow rate"
while (true)
  xlist = msgBox.Multiinput(tytul,
    "Monthly average flow rate m3/s", labls, defts )
  if (xlist.count = 0) then
    exit
  end
  for each i in 0..16
    if ((i=0) or (i=1) or (i=3) or (i=4))then
      continue
    end
    if (xlist.get(i).isnumber.not) then
      msgbox.error(xlist.get(i)+"is not a number",
        "Error in line "+i.asstring)
      continue
    end
  end
  if (xlist.get(0) = "no" )then
    break
  end
  yr = xlist.get(0).asnumber
  if ((yr < 1960) or (yr>1991)) then
    msgbox.error(xlist.get(0)+"out of range (1960-1991)",
      "Error in year field ")
    continue
  end
  break
end
if (xlist.get(0) = "no" ) then
  if(Ftmdl.iseditable.not) then
    Ftmdl.seteditable(true)
  end
  for each m in 1..12
    av.setstatus (m * 100 /12)
    av.showmsg("Processing month "+m.asstring++" ... ")
    if (m <10) then
      nameto = "Flow"+"0"+m.asstring
    else
      nameto = "Flow"+m.asstring
    end
    im = m + 4
    expr = "["+nameto+"] *"+xlist.get(2)+"*"+xlist.get(im)
    Ftmdl.calculate(expr,Ftmdl.findfield(nameto))
  end
  if(Ftmdl.iseditable) then
    Ftmdl.seteditable(false)
  end
  exit
end
' Find the table and join fields for the tables
field1 = Ftmdl.FindField("Unit_id")
field2 = Ftflow.FindField("Unit_id")
' Now perform the join....
Ftmdl.Join( field1, Ftflow, field2)
if(Ftmdl.iseditable.not) then

```

```

    Ftmodl.seteditable(true)
end
for each m in 1..12
  if (m <10) then
    namefrom = "Qm"+xlist.get(0)+"0"+m.asstring
    nameto = "Flow"+"0"+m.asstring
  else
    namefrom = "Qm"+xlist.get(0)+m.asstring
    nameto = "Flow"+m.asstring
  end
  im = m + 4
  expr = "["+namefrom+"] *"+xlist.get(2)+"*"+xlist.get(im)
  Ftmodl.calculate(expr,Ftmodl.findfield(nameto))
end
if(Ftmodl.iseditable) then
  Ftmodl.seteditable(false)
end
FTmodl.unjoinall

```

**edflow3u** Assigned to the button EQc (flow rate selection and modification),  
property: Update.

```

' edflow3u.ave : update event for "edit/select flow rate"
' Is enabled when "Modeling Units" view is active.
' =====
theView = av.GetActiveDoc
if (theView.GetEditableTheme <> NIL) then
  SELF.SetEnabled(FALSE)
  exit
end
t = theView.GetName
if ( t = "Modeling Units" ) then
  SELF.SetEnabled(TRUE)
  exit
end
SELF.SetEnabled(FALSE)

```

**eduse2** Runs scripts "eduse2a" and "eduse2a" that support editing the  
agricultural application rate. Assigned to button EAP. Property:  
Click.

```

'eduse2.ave
' =====
t = av.GetActiveDoc.GetName
if ( t = "Modeling Units") then
  av.run("eduse2m","")
  exit
end
if (t = "Application Rate") then

```

```

    av.run("eduse2a","")
    exit
end
msgbox.Info("Modeling Units or Application Rate must be active",
            "Raquired View is not Active")

```

**eduse2a      Displays the dialog box to edit agrichemical application rate.  
Changes will be made to selected (or all) county polygons displayed  
on an active view (executed from the script "eduse2")**

```

'eduse2a.ave (updated eduselb.ave)
'edtuse.update script checked the existence of themes.
'=====
'link table name:   "County link unit file" extracted from
'                   feature table of "Useunit"
'       fields:    "Fips", "Unit_id", "area_km2"
'model table name: feature table of Units (alias of "Crwsd")
'       fields:    "Unit_id", "Chemuse"
'county table name: feature table of Use feature from
'                   "Application Rate" view, (alias of "cruse")
'       fields:    "Fips", "use"
'=====

theView = av.getproject.finddoc("Modeling Units")
FTmodl = theview.findtheme("Units").Getftab
FTlink = av.getproject.finddoc("linkuse.dbf").getvtab
FTcnty = av.getproject.finddoc("Application
Rate").findtheme("use").GetFtab
if ( FTlink.haserror ) then
    msgbox.error ( " Can't find link table",
                  "theme "++FTmodl.asstring)
    exit
end
'check if all fields can be found:
FDumdl = FTmodl.findfield("Unit_id")
if ( FDumdl = nil ) then
    msgbox.error ( " Can't find Unit_id field",
                  "theme "++FTmodl.asstring)
    exit
end
FDfcty = FTcnty.findfield("Fips")
if ( FDfcty = nil ) then
    msgbox.error ( " Can't find Fips field",
                  "theme "++FTcnty.asstring)
    exit
end
FDucty = FTcnty.findfield("use")
if ( FDucty = nil ) then
    msgbox.error ( " Can't find Use field",
                  "theme "++FTcnty.asstring)
    exit
end

```

```

FDtcty = FTcnty.findfield("temp")
  if ( FDucty = nil ) then
    msgbox.error ( " Can't find Temp field",
                  "theme "++FTcnty.asstring)
    exit
  end
FDclnk = FTlink.findfield("Fips")
  if ( FDclnk = nil ) then
    msgbox.error ( " Can't find Fips field",
                  "theme "++FTlink.asstring)
    exit
  end
FDulnk = FTlink.findfield("Unit_id")
  if ( FDulnk = nil ) then
    msgbox.error ( " Can't find Unit_id field",
                  "theme "++FTlink.asstring)
    exit
  end
FDalnk = FTlink.findfield("area_km2")
  if ( FDalnk = nil ) then
    msgbox.error ( " Can't find Area_km2 field",
                  "theme "++FTlink.asstring)
    exit
  end
'FDumdl = FTmodl.findfield("Unit_id")
FDrmdl = FTmodl.findfield("ChemUse")
  if ( FDrmdl = nil ) then
    msgbox.error ( " Can't find ChemUse field",
                  "theme "++FTmodl.asstring)
    exit
  end
end
'=====
'This part is added for upstream weighetd average procedure
theFtab = FTmodl
FDunit = FDumdl
FDnext = theFtab.FindField("unit_nx")
  if ( FDnext = nil ) then
    msgbox.error ( " Can't find unit_nx field",
                  "theme "++FTmodl.asstring)
    exit
  end
end
FDorder = theFtab.FindField("order")
  if ( FDorder = nil ) then
    msgbox.error ( " Can't find order field",
                  "theme "++FTmodl.asstring)
    exit
  end
end
FDitem = FDrmdl
FDadd = theFtab.FindField("Cchemuse")
  if ( FDadd = nil ) then
    msgbox.error ( " Can't find Cchemuse field",
                  "theme "++FTmodl.asstring)
    exit
  end
end
FDarea = theFtab.FindField("areakm2")

```

```

if ( FDarea = nil ) then
  msgbox.error ( " Can't find areakm2 field",
                "theme "++FTmodl.asstring)
  exit
end
parlst = {theFTab, FDunit, FDnext, FDorder, FDadd, FDitem, FDarea}
' av.run("upwavg1",parlst)
sp = "xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx"
ln1 = "Take values from the database:"
ln2 = "(n89kgkm2, n90kgkm2,n91kgkm2,a89kgkm2)"
ln3 = "( to apply factor to current values enter: NO )"
ln4 = "Multiply data by factor:"
ln5 = "Apply specified value:"
ln6 = "( negative value deactivates this input field )"
ln7 = "( units: nitrogen [kg/km2], atrazine [kg/km2])"
labls = {sp, ln1, ln2, ln3,sp ,ln4,sp,ln5,ln6,ln7}
multf = 1.000
userv = -1.0
defts = {" ", "n89kgkm2", " ", " ", " ", " ", multf.asstring, " ",
userv.asstring, " ", " "}
tytul = " Specify Application Rate"
while (true)
  xlist = msgBox.Multiinput(tytul,
  "Annual agrichemical application by county", labls, defts )
  if (xlist.count = 0) then
    exit
  end
  if (xlist.get(7).isnumber.not) then
    msgbox.error(xlist.get(7)++"is not a number", "Error in field
#3")
    continue
  end
  aaa = xlist.get(7).asnumber
  if (aaa >= 0) then
    'edit "application by county" (table in LTpar list) and exit
    'set cruse table editable
    if(FTcnty.iseditable.not) then
      FTcnty.seteditable(true)
    end
    FTcnty.calculate(xlist.get(7),FDucty)
    if(FTcnty.iseditable) then
      FTcnty.seteditable(false)
    end
    'exit
    break
  end
  'check multiplication factor
  if (xlist.get(5).isnumber.not) then
    msgbox.error(xlist.get(5)++"is not a number", "Error in field #2")
    continue
  end

  ' now check the historical record field (AV is not case sensitwiwe !
  ' NO = no:
  if(xlist.get(1)="no") then

```

```

    'set cruse table editable
    if(FTcnty.iseditable.not) then
        FTcnty.seteditable(true)
    end
    expres = xlist.get(5)+"*[use]"
    FTcnty.calculate(expres,FDtcty)
    FTcnty.calculate("[temp]",FDucty)
    if(FTcnty.iseditable) then
        FTcnty.seteditable(false)
    end
    'exit
    break
end
'last possibility that user wants historical record !
hrecord = FTcnty.findfield(xlist.get(1))
if (hrecord = nil) then
    msgbox.error(xlist.get(1)+" is neither NO nor field name"++ln2,
        "Error in field #1")
    continue
end
'set cruse table editable
if(FTcnty.iseditable.not) then
    FTcnty.seteditable(true)
end
expres = xlist.get(5)+"*["+xlist.get(1)+"]"
FTcnty.calculate(expres,FDucty)
if(FTcnty.iseditable) then
    FTcnty.seteditable(false)
end
'exit
break
end 'of while error = false
'make dictionaries:
dlink = dictionary.make(FTlink.getnumrecords)
'dc (dict.) relates fips and edited use (field "use")
dc = dictionary.make(FTcnty.getnumrecords)
'dm (dict.) relates unit_id and cumulative (use * area)
dm = dictionary.make(FTmodl.getnumrecords)
'da (dict.) relates unit id and cumulative area
da = dictionary.make(FTmodl.getnumrecords)
'assign initial values to cumulative (use * area) and (area)
' the average use over modelling unit = cum(use*area)/cum(area)
for each rec in FTmodl
    dm.add(FTmodl.returnvalue(FDumdl,rec),0)
    da.add(FTmodl.returnvalue(FDumdl,rec),0)
end
'fill up the dc dictionary: Fips ==> use (cruse counties coverage)
for each rec in FTcnty
    dc.add(FTcnty.returnvalue(FDfcty,rec),FTcnty.returnvalue(FDucty,rec))
end
'calculate cumulative values
for each rec in FTlink
    unitid = FTlink.returnvalue(FDulnk,rec)
    fips = FTlink.returnvalue(FDclnk,rec)
    area = FTlink.returnvalue(FDalnk,rec)

```

```

    applic = dc.get(fips)
    oldarea = da.get(unitid)
    oldappl = dm.get(unitid)
    da.set(unitid, oldarea + area )
    dm.set(unitid, oldappl + (area * applic))
end
'Write results into table "crwsd", item "Chemuse"
'set crwsd (model parameters) table editable
if(FTmodl.iseditable.not) then
    FTmodl.seteditable(true)
end
'Write results to "crwsd" table
for each rec in FTmodl
    unitid = FTmodl.returnvalue(FDumdl,rec)
    use = dm.get(unitid)/da.get(unitid)
    av.ShowMsg("Writing to chemuse: "++rec.asString++use.asstring)
    FTmodl.SetValue(FDrmdl, rec, use)
end
av.run("upwavg1",parlst)
'stop edit session (and refresh) "crwsd" table
if(FTmodl.iseditable) then
    FTmodl.seteditable(false)
end
end

```

**eduse2m     Displays the dialog box to edit agrichemical application rate.  
Changes will be made to selected (or all) modeling unit polygons  
displayed on active view (executed from the script "eduse2")**

```

'eduse2m.ave (old eduselc.ave)
'edtuse.update script checked the existence of themes.
'=====
'model table name:  feature table of "Units"
'      fields:  "Unit_id", "Chemuse"
'=====
theView = av.getproject.finddoc("Modeling Units")
FTmodl = theview.findtheme("Units").Getftab
if(theview.findtheme("Units").isactive.not) then
    msgbox.warning("Theme >Units< is not active", "")
end
'check if all fields can be found:
FDrmdl = FTmodl.findfield("ChemUse")
if ( FDrmdl = nil ) then
    msgbox.error ( " Can't find ChemUse field",
                  "theme "++FTmodl.asstring)
    exit
end
'=====
'This part is added for upstream weighthd average procedure
theFTab = FTmodl
FDunit = theFTab.FindField("unit_id")
if ( FDunit = nil ) then
    msgbox.error ( " Can't find unit_id field",

```

```

        "theme "++FTmodl.asstring)
    exit
end
FDnext = theFtab.FindField("unit_nx")
if ( FDnext = nil ) then
    msgbox.error ( " Can't find unit_nx field",
        "theme "++FTmodl.asstring)
    exit
end
FDorder = theFtab.FindField("order")
if ( FDorder = nil ) then
    msgbox.error ( " Can't find order field",
        "theme "++FTmodl.asstring)
    exit
end
FDitem = FDrmdl
FDadd = theFtab.FindField("Cchemuse")
if ( FDadd = nil ) then
    msgbox.error ( " Can't find Cchemuse field",
        "theme "++FTmodl.asstring)
    exit
end
FDarea = theFtab.FindField("areakm2")
if ( FDarea = nil ) then
    msgbox.error ( " Can't find areakm2 field",
        "theme "++FTmodl.asstring)
    exit
end
parlst = {theFtab, FDunit, FDnext, FDorder, FDadd, FDitem, FDarea}
' av.run("upwavg1",parlst)
'=====
sp = "xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx"
ln4 = "Multiply data by factor:"
ln5 = "Apply specified value:"
ln6 = "( negative value deactivates this input field )"
ln7 = "( units: nitrogen [kg/km2], atrazine [kg/km2])"
labls = {sp, ln4,sp,ln5,ln6,ln7}
multf = 1.000
userv = -1.0
defts = {" ", multf.asstring," ", userv.asstring," "," "}
tytul = "    Specify Application Rate"
while (true)
    xlist = msgBox.Multiinput(tytul,
        "Annual agrichemical application by modeling unit",
        labls, defts )
    if (xlist.count = 0) then
        exit
    end
    if (xlist.get(3).isnumber.not) then
        msgbox.error(xlist.get(3)+"is not a number", "Error in field
#1")
        continue
    end
    aaa = xlist.get(3).asnumber
    if (aaa >= 0) then

```



```

    'put the user value into "ChemUse" field (selected or all recds)
    'set crwsd table editable
    if(FTmdl.iseditable.not) then
        FTmdl.seteditable(true)
    end
    FTmdl.calculate(xlist.get(3),FDrmdl)
    av.run("upwavg1",parlst)
    if(FTmdl.iseditable) then
        FTmdl.seteditable(false)
    end
    break
end
'check multiplication factor
if (xlist.get(1).isnumber.not) then
    msgbox.error(xlist.get(1)+"is not a number", "Error in field
#1")
    continue
end
'set crwsd table editable
if(FTmdl.iseditable.not) then
    FTmdl.seteditable(true)
end
expres = xlist.get(1)+"*[ChemUse]"
FTmdl.calculate(expres,FDrmdl)
av.run("upwavg1",parlst)
if(FTmdl.iseditable) then
    FTmdl.seteditable(false)
end
break
end
end

```

**eduse2u Assigned to button EAP (edit application rate). Property: update.**

```

' eduse2u.ave : update event for "edit application"
' Is enabled when either "Application Rate" view or "Modeling Units"
' view is active.
' =====
theView = av.GetActiveDoc
if (theView.GetEditableTheme <> NIL) then
    SELF.SetEnabled(FALSE)
    exit
end
t = theView.GetName
if ( ( t = "Modeling Units") or ( t = "Application Rate") ) then
    SELF.SetEnabled(TRUE)
    exit
end
SELF.SetEnabled(FALSE)

```

**equat6        Calculates concentrations and loads for all or selected months of the year. Assigned to the button RUN, property: Click.**

```
'equat6 modified equat5, time component changed
'   into seasonal index, separate trend,
'   equation includes land-length and land-slope
'   model stored in model2.dbf
'equat4 calculates concentration and load in
'feature attribute table of "crwsd"
'model equations and parameters are stored in modell.dbf
'same as equat3 plus recalculation for selected months
'=====
theView = av.getproject.finddoc("Modeling Units")
FTmodl = theview.findtheme("Units").Getftab
if (FTmodl = nil) then
    msgbox.error("Can't find feature table", "Units unaccessible")
    exit
end
tabl = av.getproject.finddoc("model2.dbf")
vtabl = tabl.getvtab
if (vtabl = nil) then
    msgbox.error("Can't find model2.dbf", "model2.dbf unaccessible")
    exit
end
i = -1
fsel = vtabl.findfield("sel")
for each rec in vtabl
    i = i + 1
    if (vtabl.returnvalue(fsel,rec) = 1 ) then
        idmod = i
        break
    end
end
fmodel = vtabl.findfield("Model")
xtxt = "Model:
"+vtabl.returnvalue(vtabl.findfield("Model"),idmod)++"?"
if(msgbox.miniYesNo(xtxt, True).not) then
    exit
end
'===== months selection =====
yyy = "all OR selected months 0=NO, 1=YES"
mylist = {"1","0","0","0","0","0","0","0","0","0","0","0","0","0"}
lbs = {"All months . . . . . ",
"January . . . . . ",
"February . . . . . ",
"March . . . . . ",
"April . . . . . ",
"May . . . . . ",
"June . . . . . ",
"July . . . . . ",
"August . . . . . ",
"September . . . . . ",
"October . . . . . ",
"November . . . . . "}
```

```

    "December . . . . ."}
xloop = true
while(xloop)
newlist = msgbox.multiinput(yyy, "Recalculate",lbs, mylist)
if(newlist.count = 0) then
    msgbox.info("Operation Canceled", "Process: RUN")
    exit
end
xloop = false
ind = 0
for each x in newlist
    ind = ind + 1
    if(x.isnumber.not) then
        msgbox.error("error in line"++ind.asstring+"::"+x,"")
        '
        msgbox.info(x.asstring, "xxx")
        xloop = true
        mylist = newlist
        break
    end
end
end
if(newlist.get(0).asnumber > 0) then
    for each ind in 1..12
        newlist.set(ind, 1)
    end
else
    for each ind in 1..12
        newlist.set(ind, newlist.get(ind).asnumber)
    end
end
'==== end of months selection =====
feq = vtab1.findfield("equation")
exp101 = vtab1.returnvalue(feq,idmod)
exp102 = exp101.substitute("U", "[Cchemuse]")
exp103 = exp102.substitute("A", "[Careakm2]")
exp104 = exp103.substitute("LS", "[Alndslp]")
exp105 = exp104.substitute("LL", "[Alndlgkm]")
' calculate trend coefficient
fyear = vtab1.findfield("year")
yr = vtab1.returnvalue(fyear,idmod)
xyr = yr.asstring
trfunc = vtab1.findfield("Ftrend")
trcoeff = vtab1.findfield("Trendcf")
exp400 = vtab1.returnvalue(trfunc,idmod)
exp401 = exp400.substitute("Year", xyr)
if(Ftmodl.iseditable.not) then
    Ftmodl.seteditable(true)
end
if(vtab1.iseditable.not) then
    vtab1.seteditable(true)
end
xxx0 = vtab1.calculate(exp401,trcoeff)
xtr = vtab1.returnvalue(trcoeff,idmod).asstring
exp110 = exp105.substitute("TR", xtr)
' clear fields:

```

```

'   xxx0 = vtbl.calculate("0",trcoeff)
av.showstopbutton
for each mt in 1..12
  xstop = av.setstatus (mt * 100 /12)
  av.showmsg("Processing month "++mt.asstring++" ... ")
  if(newlist.get(mt) = 0) then
    continue
  end
  if(xstop.not) then
    xyes =msgbox.Mininyesno("Do you want to stop ?", False)
    if(xyes) then
      if(Ftmodl.iseditable) then
        Ftmodl.seteditable(false)
      end
      if(vtbl.iseditable) then
        vtbl.seteditable(false)
      end
      av.clearstatus
      exit
    else
      av.ClearStatus
      av.showstopbutton
    end
  end
  end
  xmt = mt.asstring
  if (mt <10) then
    Qname = "[Flow"+"0"+xmt+]"
    Cname = "Conc"+"0"+xmt
    Lname = "Load"+"0"+xmt
    Sname = "Si"+"0"+xmt
  else
    Qname = "[Flow"+xmt+]"
    Cname = "Conc"+xmt
    Lname = "Load"+xmt
    Sname = "Si"+xmt
  end
  end
  'clear fields,
'   xxx0 = Ftmodl.calculate("0",Ftmodl.findfield(Cname))
'   xxx0 = Ftmodl.calculate("0",Ftmodl.findfield(Lname))
xsi = vtbl.returnvalue(vtbl.findfield(Sname),idmod).asstring
expl11 = expl10.substitute("SI", xsi)
expl12 = expl11.substitute("Q", Qname)
xxx3 = Ftmodl.calculate(expl12,Ftmodl.findfield(Cname))
exp200 = "["+Cname+"] *"+Qname
xxx4 = Ftmodl.calculate(exp200,Ftmodl.findfield(Lname))
end
if(Ftmodl.iseditable) then
  Ftmodl.seteditable(false)
end
if(vtbl.iseditable) then
  vtbl.seteditable(false)
end
av.clearstatus
av.clearmsg

```

**Selmodel**     **Displays the dialog box to select a model: equation:  $c = f(\text{Flow, Area, Time, Use, X})$ . Assigned to the button SMO, property: Click.**

```
' selmodel (puts 1 into field "sel" of "modell.dbf"
' if selected, 0 otherwise.
' =====
tabl = av.getproject.finddoc("modell.dbf")
vtabl = tabl.getvtab
if (vtabl = nil) then
  msgbox.error("Can't find modell.dbf", "modell.dbf inaccessible")
  exit
end
fsel = vtabl.findfield("sel")
fmodel = vtabl.findfield("Model")
modlist = list.make
for each rec in vtabl
  modlist.add(vtabl.returnvalue(fmodel,rec))
end
xselect = msgbox.choiceasstring(modlist,"Select model:",
  "Model selection")
if(xselect <> nil) then
  if(vtabl.iseditable.not) then
    vtabl.seteditable(true)
  end
  for each rec in vtabl
    if(xselect = vtabl.returnvalue(fmodel,rec)) then
      vtabl.setvaluenumber(fsel,rec,1)
    else
      vtabl.setvaluenumber(fsel,rec,0)
    end
  end
  if(vtabl.iseditable) then
    vtabl.seteditable(false)
  end
end
else
  msgbox.info("No selection made", "Nothing changed !")
end
```

**Stime1**     **Displays the dialog box to select a year ( time variable for a model). Assigned to the button SYr, property: Click.**

```
'stime1
' writes year into the field "year" of "modell.dbf"
' =====
tabl = av.getproject.finddoc("modell.dbf")
vtabl = tabl.getvtab
if (vtabl = nil) then
  msgbox.error("Can't find modell.dbf", "modell.dbf inaccessible")
```

```

    exit
end
fyear = vtab1.findfield("Year")
if (fyear = nil) then
    msgbox.error("Can't find field: Year",
        "Field does not exist ?")
    exit
end
yrdef = vtab1.returnvalue(fyear,0).asstring
'prompt for year:
while(true)
    tx = "Specify a value for model's time variable (enter a year)"
    year = msgbox.input(tx,"Setting Model's Time Variable", yrdef)
    if(year = nil) then
        exit
    end
    if (year.isnumber.not) then
        msgbox.error(year++"is not a number",
            "Wrong entry ")
        continue
    end
    yr = year.asnumber
    if ((yr < 1900) or (yr>2999)) then
        msgbox.error(year++"out of range (1900-2999)",
            "Error in year field ")
        continue
    end
    break
end
if(vtab1.iseditable.not) then
    vtab1.seteditable(true)
end
vtab1.calculate(year,fyear)
if(vtab1.iseditable) then
    vtab1.seteditable(false)
end
end

```

**upwavq1    Calculates weighted average for upstream units. Incorporated into scripts eduse.**

```

' upwavg1.ave calculates weighted average for upstream units
' This is run from eduse... script
theFTab = self.get(0)
FDunit = self.get(1)
FDnext = self.get(2)
FDorder = self.get(3)
FDadd = self.get(4)
FDitem = self.get(5)
FDarea = self.get(6)
ltunit = list.make
ltnext = list.make
ltorder = list.make

```

```

ltitem = list.make
ltadd = list.make
unnx = theFTab.GetNumRecords
unnn = unnx - 1
dcarea = dictionary.make(unnx)
dcmass = dictionary.make(unnx)
for each rec in theFTab
    Nunit = theFTab.ReturnValueNumber(FDunit, rec)
    ltunit.Add(Nunit)
    ltnext.Add(theFTab.ReturnValueNumber(FDnext, rec))
    ltorder.Add(theFTab.ReturnValueNumber(FDorder, rec))
    Narea = theFTab.ReturnValueNumber(FDarea, rec)
    Nmass = theFTab.ReturnValueNumber(FDitem, rec) * Narea
    dcarea.add(Nunit, Narea )
    dcmass.add(Nunit, Nmass )
end
' find maximum order
unmaxord = 1
for each i in 0..unnn
    av.ShowMsg("Finding maximum order ..."++i.asString)
    istatus = i * 100 / unnn
    av.SetStatus(istatus)
    if (ltorder.get(i) > unmaxord ) then
        unmaxord = ltorder.get(i)
    end
end
'calculate cumulative values
unmaxord1 = unmaxord - 1
for each k in 1..unmaxord1
    av.ShowMsg("Calculating cumulative values ..."++k.asString)
    istatus = k * 100 / unmaxord
    av.SetStatus(istatus)
    for each i in 0..unnn
        if ( ltorder.get(i) <> k ) then
            continue
        end
        Nunit = ltunit.get(i)
        Nnext = ltnext.get(i)
        Nxmass = dcmass.get(Nunit) + dcmass.get(Nnext)
        Nxarea = dcarea.get(Nunit) + dcarea.get(Nnext)
        dcmass.set(Nnext, Nxmass)
        dcarea.set(Nnext, Nxarea)
    end
end
av.clearMsg
'write results to Ftab
for each rec in theFTab
    Nunit = theFTab.ReturnValueNumber(FDunit, rec)
    Xavg = dcmass.get(Nunit) / dcarea.get(Nunit)
    TheFTab.SetValue( FDadd, rec, Xavg )
end
theFTab.Refresh
if(TheFTab.isEditable) then
    theFTab.SetEditable(false)
end

```

**selup2**      **Selects upstream units (active theme must have the following fields: unit\_id, unit\_nx, and order). Assigned to a button in category: Tools, property: Apply.**

```
'three fields are required:
nm_id = "unit_id"
nm_nx = "unit_nx"
nm_or = "order"
theView = av.GetActiveDoc
theThemes = theView.GetActiveThemes
p = theView.GetDisplay.ReturnUserPoint
shiftk = System.IsShiftKeyDown
if (ShiftK) then
    op = #VTAB_SELTYPE_XOR
else
    op = #VTAB_SELTYPE_NEW
end
t = thethemes.get(0)
if (t.CanSelect.not) then
    exit
end
t.SelectByPoint(p, op)
recs = t.FindByPoint(p)
if(recs.count = 0) then
    exit
end
rec = recs.get(0)
v = t.getFTab
fid = v.findfield(nm_id)
fnx = v.findfield(nm_nx)
ford = v.findfield(nm_or)
if ((fid=nil)or(fnx=nil)or(ford=nil)) then
    msgbox.error ( " Can't find necessary fields",
                  " Upstream selection ")
    exit
end
xid = v.returnvalue(fid, rec)
xnx = v.returnvalue(fnx, rec)
rno = rec
rord = v.returnvalue(ford,rno)
if(rord = 1) then
    exit
end
bmap = v.getselection
stk = stack.make
av.showmsg("Selecting upstream elements ...")
while (true)
    for each rcd in v
        'don't look for upstream if first order
        if(rord = 1 ) then
            break
```



```

    end
    xnx = v.returnvalue(fnx,rcd)
    if(xid = xnx) then
        stk.push(rcd.asstring)
    end
end
rno = stk.pop.asnumber
if(shiftk) then
    if(bmap.get(rno)) then
        bmap.clear(rno)
    else
        bmap.set(rno)
    end
else
    bmap.set(rno)
end
t.blinkrecord(rno)
rord = v.returnvalue(ford,rno)
xid = v.returnvalue(fid,rno)
if((stk.depth = 0 )and (rord = 1)) then
    break
end
end
av.Clearmsg

```

**Seldown2** Selects downstream units (active theme must have the following fields: unit\_id, unit\_nx, and order). Assigned to a button in category: Tools, property: Apply.

```

'seldown2
'three fields are required:
nm_id = "unit_id"
nm_nx = "unit_nx"
nm_or = "order"
theView = av.GetActiveDoc
theThemes = theView.GetActiveThemes
p = theView.GetDisplay.ReturnUserPoint
shiftk = System.IsShiftKeyDown
if (ShiftK) then
    op = #VTAB_SELTYPE_XOR
else
    op = #VTAB_SELTYPE_NEW
end
t = thethemes.get(0)
if (t.CanSelect.not) then
    exit
end
t.SelectByPoint(p, op)
recs = t.FindByPoint(p)
if(recs.count = 0) then
    exit
end
end

```

```

rec = recs.get(0)
v = t.getFTab
fid = v.findfield(nm_id)
fnx = v.findfield(nm_nx)
ford = v.findfield(nm_or)
if ((fid=nil)or(fnx=nil)or(ford=nil)) then
  msgbox.error ( " Can't find necessary fields",
                " Downstream selection ")
  exit
end
xid = v.returnvalue(fid, rec)
xnx = v.returnvalue(fnx, rec)
rno = rec
rord = v.returnvalue(ford,rno)
if(xnx = 0) then
  exit
end
bmap = v.getselection
av.showmsg("Selecting downstream elements ...")
while (true)
  for each rcd in v
    xid = v.returnvalue(fid,rcd)
    if(xid = xnx) then
      bmap.set(rcd)
      if(shiftk) then
        if(bmap.get(rcd)) then
          bmap.clear(rcd)
        else
          bmap.set(rcd)
        end
      else
        bmap.set(rcd)
      end
    end
    xnx = v.returnvalue(fnx,rcd)
    if (xnx = 0) then
      exit
    end
  end
end
msgbox.error("I can't find the most downstream reach",
            "Last reach not found")

```

**Selup2u**     **Update event for buttons that select upstream and downstream units ("selup2", "seldown2"). Assigned to a button in category: Tools, property: Update.**

```

'selup2.update
' Is enabled when prpoer fields are found in active
'three fields are required:
nm_id = "unit_id"
nm_nx = "unit_nx"

```

```

nm_or = "order"
theView = av.GetActiveDoc
if (theView.GetEditableTheme <> NIL) then
    SELF.SetEnabled(FALSE)
    exit
end
for each t in theView.GetThemes
    if (t.IsVisible.not) then
        continue
    end
    if (t.IsActive.not) then
        continue
    end
    fid = t.getftab.findfield(nm_id)
    fnx = t.getftab.findfield(nm_nx)
    ford = t.getftab.findfield(nm_or)
    if ((fid<>nil)and(fnx<>nil)and(ford<>nil)) then
        SELF.SetEnabled(TRUE)
        exit
    end
end
break
end
SELF.SetEnabled(FALSE)

```

**decay1**      **Calculates the concentrations and loads of a chemical that exponentially “decays” as it is carried by water. Assigned to the button FOR, property: Click.**

```

' decay1.ave Calculates loads and concentrations
'           in surface waters. Chemical losses
'           are governed by the first order reaction.
'unit conversion factor
' Assumptions:
' application rate is in kg/km2/yr
' area is in km2
' then, the mass runoff is in kg/year
' to make load units g/s
' the conversion factor equal to 0.00003170979
' 1000 (g/kg) / ( 365 days * 86400 seconds )
' must be applied:
conversf = 0.00003170979
' flow is in m3/s, thus concentration is in g/m3
' or mg/L
' The usercoef represents seasonal index, confidence
' limits, extreme conditions, application timing etc.
' It is a multiplier of mass runoff from the field.
usercoef = 1

' This part is for an Attribute table !!!
' note: dbf table contains a dummy (first) record
' that may influence calculations (unlikely)
' but will produce avenue errors (dividing by 0)

```

```

theView = av.getproject.finddoc("Modeling Units")
theFtab = theview.findtheme("Units").Getftab
TBunit = theFtab
' Ask user for needed fields:
TXunit = "unit_id"
TXnext = "unit_nx"
TXorder = "order"
TXarea = "Areakm2"
TXuse = "Chemuse"
TXtrat = "Travtime"
TXloss = "Losscoef"
TXexpt = "Expofac"
TXflow = "Flow"
FDunit = theFtab.FindField(TXunit)
FDnext = theFtab.FindField(TXnext)
FDorder = theFtab.FindField(TXorder)
FDarea = theFtab.FindField(TXarea)
FDuse = theFtab.FindField(TXuse)
FDtrat = theFtab.FindField(TXtrat)
FDloss = theFtab.FindField(TXloss)
FDexpt = theFtab.FindField(TXexpt)
FDflow = theFtab.FindField(TXflow)
' Make a list of VTunit fields
LSFDunit = theFtab.GetFields
' If the default fields can not be found, ask user for
' the required fields:
if(FDunit=nil) then
  ' Select current_unit id field (from entity)
  FDunit = MsgBox.List(LSFDunit, "Select a field that contains"
    ++NL++"the watershed/stream ID", "Unit ID")
end
' If not selected, quit
IF (FDunit = nil) then
  exit
end
if(FDnext=nil) then
  ' Select downstream_unit id field
  FDnext = MsgBox.List(LSFDunit, "Select a field that contains"
    ++NL++" the downstream unit ID", "Downstream unit ID")
end
' If not selected, quit
IF ( FDnext = nil) then
  exit
end
if(FDorder=nil) then
  ' Select order field
  FDorder = MsgBox.List(LSFDunit, "Select field that contains"
    ++NL++" the watershed/stream order", "Order")
end
' If not selected, quit
IF ( FDorder = nil) then
  xxx = msgbox.info("Create a field of stream order",
    "Missing fields")
  exit
end

```

```

if(FDarea=nil) then
  ' Select field that contains area of the modeling unit
  FDarea = MsgBox.List(LSFDunit, "Select a field that contains"
  ++NL++"modeling unit area (km2)", "Area")
end
' If not selected, quit
IF ( FDarea = nil) then
  exit
end
if(FDuse=nil) then
  ' Select field that contains the application rate within unit
  FDuse = MsgBox.List(LSFDunit, "Select a field that contains"
  ++NL++"application rate (kg/km2/yr)", "Application Rate")
end
' If not selected, quit
IF ( FDuse = nil) then
  exit
end
if(FDexpt=nil) then
  ' Select field that contains the export factors
  FDexpt = MsgBox.List(LSFDunit, "Select a field that contains"
  ++NL++"the export factors", "Export factor")
end
' If not selected, quit
IF ( FDexpt = nil) then
  exit
end
if(FDtrat=nil) then
  ' Select field that contains travel time through modeling unit
  FDtrat = MsgBox.List(LSFDunit, "Select a field that contains"
  ++NL++"unit travel time (day)", "Travel Time")
end
' If not selected, quit
IF ( FDtrat = nil) then
  exit
end
if(FDloss=nil) then
  ' Select field that contains the loss coefficient
  FDloss = MsgBox.List(LSFDunit, "Select a field that contains"
  ++NL++"the loss coefficient (1/d)", "Loss coefficient")
end
' If not selected, quit
IF ( FDloss = nil) then
  exit
end
if(FDflow=nil) then
  ' Select field that contains the accumulated flow rate
  FDflow = MsgBox.List(LSFDunit, "Select a field that contains"
  ++NL++"the total flow rate", "Flow rate")
end
' If not selected, quit
IF ( FDflow = nil) then
  exit
end
' Prompt for a name (name of the field, that will be created)

```

```

TXconc = msgbox.input("Enter the name of concentration field ",
"Concentration", "conc_gm3")
FDconc = theFtab.FindField(TXconc)
xx = false
if (FDconc = nil) then
  xx = msgbox.YesNo( TXconc++"field is not in the table "
  ++TBunit.asString+nl+"Do you want to create this field ?",
  "Field not Found !", true)
  if (xx.not) then
    exit
  end
  fenum = FDflow.gettype
  fprec = 6
  fwidth = 12
  xxx = msgbox.input("Please specify a width of the conc. field",
  "Width of the new field" , fwidth.asstring)
  fwidth = xxx.asnumber
  IF ( fwidth = nil) then
    exit
  end
  xxx2 = msgbox.input("Please specify number of decimal places:",
  "Precision of the new field" , fprec.asstring)
  fprec = xxx2.asnumber
  IF ( fprec = nil) then
    exit
  end
  'check if the theFtab can be edited if not, set it editable
  if(theFtab.isEditable.not) then
    theFtab.SetEditable(true)
  end
  ' Create a new field
  if(theFtab.CanAddFields) then
    FDconc = field.Make(TXconc,fenum,fwidth,fprec)
    theFtab.AddFields({FDconc})
  else
    msgBox.warning("Can not add a field to"++theFtab.asString,
    "Can not edit table")
    exit
  end
else
  if(theFtab.isEditable.not) then
    theFtab.SetEditable(true)
  end
end
' Prompt for a name (name of the field, that will be created)
TXload = msgbox.input("Enter the name of load field ",
"Load", "load_gm3")
FDload = theFtab.FindField(TXload)
xx = false
if (FDload = nil) then
  xx = msgbox.YesNo( TXload++"field is not in the table "
  ++TBunit.asString+nl+"Do you want to create this field ?",
  "Field not Found !", true)
  if (xx.not) then
    exit
  end
end

```

```

end
fenum = FDflow.gettype
fprecl = 6
fwidthl = 16
xxx3 = msgbox.input("Please specify a width of the load field",
  "Width of the new field" , fwidthl.asstring)
  fwidthl = xxx3.asnumber
IF ( fwidthl = nil) then
  exit
end
xxx4 = msgbox.input("Please specify number of decimal places:",
  "Precision of the new field" , fprecl.asstring)
  fprecl = xxx4.asnumber
IF ( fprecl = nil) then
  exit
end
'check if the theFtab can be edited if not, set it editable
if(theFtab.isEditable.not) then
  theFtab.SetEditable(true)
end
' Create a new field
if(theFtab.CanAddFields) then
  FDload = field.Make(TXload,fenum,fwidthl,fprecl)
  theFtab.AddFields({FDload})
else
  msgBox.warning("Can not add a field to"++theFtab.asString,
    "Can not edit table")
  exit
end
else
  if(theFtab.isEditable.not) then
    theFtab.SetEditable(true)
  end
end
' the conversion factor and a multiplier
while(true)
  TXcf = msgbox.input("Enter a conversion factor"++NL++
    "(kg/yr -> g/s)", "Conversion Factor",
    conversf.asstring)
  if(TXcf = nil) then
    exit
  end
  if(TXcf.isnumber.not) then
    msgbox.error(TXcf++"is not a number","Wrong entry")
    continue
  end
  conversf = TXcf.asnumber
break
end
while(true)
  TXum = msgbox.input("Enter a multiplier of the"++NL++
    "chemical runoff from the field",
    "Runoff Multiplier", usercoef.asstring)
  if(TXum = nil) then
    exit

```

```

end
if(TXum.isnumber.not) then
    msgbox.error(TXum++"is not a number","Wrong entry")
    continue
end
usercoef = TXum.asnumber
break
end
'create lists
ltunit = list.make
ltnext = list.make
ltorder = list.make
'ltadd = list.make
'ltexnkt = list.make
'ltroff = list.make
unnx = theFTab.GetNumRecords
unnn = unnx - 1
dcload = dictionary.make(unnx)
dcexnkt = dictionary.make(unnx)
for each rec in theFTab
    Nunit = theFTab.ReturnValueNumber(FDunit, rec)
    ltunit.Add(Nunit)
    ltnext.Add(theFTab.ReturnValueNumber(FDnext, rec))
    ltorder.Add(theFTab.ReturnValueNumber(FDorder, rec))
    xk = theFTab.ReturnValueNumber(FDloss, rec)
    xt = theFTab.ReturnValueNumber(FDtrat, rec)
    exnkt = number.geteuler^( -1 * xk * xt )
    dcexnkt.add(Nunit, exnkt )
    xa = theFTab.ReturnValueNumber(FDarea, rec)
    xu = theFTab.ReturnValueNumber(FDuse, rec)
    xe = theFTab.ReturnValueNumber(FDexpt, rec)
    xroff = xa * xu * xe * usercoef * conversf
    dcload.add(Nunit, xroff )
end
' find maximum order
unmaxord = 1
for each i in 0..unnn
    av.ShowMsg("Finding maximum order ..."++i.asString)
    istatus = i * 100 / unnn
    av.SetStatus(istatus)
    if (ltorder.get(i) > unmaxord ) then
        unmaxord = ltorder.get(i)
    end
end
'calculate values of load
unmaxord1 = unmaxord - 1
for each k in 1..unmaxord1
    av.ShowMsg("Decaying and decaying ..."++k.asString)
    istatus = k * 100 / unmaxord
    av.SetStatus(istatus)
    for each i in 0..unnn
        if ( ltorder.get(i) <> k ) then
            continue
        end
        Nunit = ltunit.get(i)

```



```

    Nnext = ltnext.get(i)
    Nxvalue = dclload.get(Nunit) * dcexnkt.get(Nnext) + dclload.get(Nnext)
    dclload.set(Nnext, Nxvalue)
  end
end
av.clearMsg
av.clearStatus
'write results to Ftab
for each rec in theFtab
  Nunit = theFtab.ReturnValueNumber(FDunit, rec)
  Nload = dclload.get(Nunit)
  TheFtab.SetValue( FDload, rec, Nload )
  Nflow = theFtab.ReturnValueNumber(FDflow, rec)
  if (Nflow = 0 ) then
    Nconc = 0
  else
    Nconc = Nload / Nflow
  end
  TheFtab.SetValue( FDconc, rec, Nconc )
end
theFtab.Refresh
if(TheFtab.isEditable) then
  theFtab.SetEditable(false)
end
end

```

### B3 Scripts from the project “results”

#### **Cchart1 Draw charts of the concentration (12 months) at the center of selected features, property Click ( runs script "schart1")**

```
' Cchart1, One year of Concentrations from crwsd
chttitle = msgbox.input("Enter a legend's title",
    "Concentration Charts", "Conc. (m)g/m3")
if (chttitle <> nil ) then
    LTpar = {chttitle,"Conc",true,1978,1979, 1, 12}
    av.run("schart1",LTpar)
end
```

#### **Lchart Draw charts of the load (12 months) at the center of selected features, property Click ( runs script "schart1").**

```
' Lchart1, One year of load from crwsd
chttitle = msgbox.input("Enter a legend's title",
    "Load Charts", "Load (m)g/s")
if (chttitle <> nil ) then
    LTpar = {chttitle, "Load", true, 1978, 1979, 1, 12}
    av.run("schart1",LTpar)
end
```

#### **Qchar1 Draw charts of the flow rate (12 months) at the center of selected features, property Click ( runs script "schart1").**

```
' Qchart1, One year of flow from crwsd
chttitle = msgbox.input("Enter a legend's title",
    "Flow Charts", "Flow m3/s")
if (chttitle <> nil ) then
    LTpar = {chttitle, "Flow", true, 1978, 1979, 1, 12}
    av.run("schart1",LTpar)
end
```

## B4 Scripts from the project “flwprc”

**pmchar1**    **Draws charts of the precipitation depth at the center of selected features, for selected months, property: Click, ( executes script "schart1")**

```
' PMchar1, displays multiple years of precipitation
'           depth from gsprec or unprec
'=====
'' title, prefix, one_year, fy, ty, fm, tm,
'=====
sp = "xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx"
ln1 = "Legend title:"
ln3 = "From year (min 1960):"
ln4 = "From month:"
ln5 = "To year (max 1991)"
ln6 = "To month"
labls = {ln1,ln3,ln4,ln5,ln6}
defts = {"Precip. cm/d", "1990","1","1990","12"}
tytul = " Specify title and time interval"
while (true)
  xlist = msgBox.Multiinput(tytul,
    "Draw charts of monthly average precipitation", labls, defts )
  if (xlist.count = 0) then
    exit
  end
  for each i in 0..4
    if (i=0)then
      continue
    end
    if (xlist.get(i).isnumber.not) then
      msgbox.error(xlist.get(i)+"is not a number",
        "Error in line "++i.asstring)
      continue
    end
  end
  end
  fy = xlist.get(1).asnumber
  ty = xlist.get(3).asnumber
  fm = xlist.get(2).asnumber
  tm = xlist.get(4).asnumber
  if (fy < 1960) then
    msgbox.error(xlist.get(1)+" - year out of range ( < 1960 )",
      "Error in year field ")
    continue
  end
  if (ty > 1991) then
    msgbox.error(xlist.get(3)+" - year out of range ( 1991 < )",
      "Error in year field ")
  end
end
```

```

        continue
    end
    if ((fm < 1) or (fm>12)) then
        msgbox.error(xlist.get(3)++" - month out of range (1-12)",
            "Error in year field ")
        continue
    end
    if ((tm < 1) or (tm>12)) then
        msgbox.error(xlist.get(4)++" - month out of range (1-12)",
            "Error in year field ")
        continue
    end
    break
end
LTpar = {xlist.get(0), "PM", false, fy, ty, fm, tm}
av.run("schart1",LTpar)

```

### **pmchar1u** Draws charts of the precipitation depth, property: Update

```

'Pmchar1u.ave
'multiple years data
'update event for "draw charts of precipitation"
' Is enabled when a theme is active and first
theView = av.GetActiveDoc
if (theView.GetEditableTheme <> NIL) then
    SELF.SetEnabled(FALSE)
    exit
end
xx = theView.GetActiveThemes
if(xx.count <> 0) then
    visnm = xx.get(0).getname
    if (( visnm = "unprec") or (visnm = "gsprec")) then
        SELF.SetEnabled(TRUE)
        exit
    end
end
SELF.SetEnabled(FALSE)

```

### **qmchar1** Draws charts of the flow rate at the center of selected features, for selected months, property: Click, (runs script "schart1",)

```

' QMchar1, multiple years of flow from gsflow or unflow
'=====
'' title, prefix, one_year, fy, ty, fm, tm,
'=====
sp = "xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx"
ln1 = "Legend title:"
ln3 = "From year (min 1960):"
ln4 = "From month:"
ln5 = "To year (max 1991)"

```

```

ln6 = "To month"
labls = {ln1,ln3,ln4,ln5,ln6}
defts = {"Flow m3/s", "1990","1","1990","12"}
tytul = " Specify title and time interval"
while (true)
  xlist = msgBox.Multiinput(tytul,
    "Draw charts of monthly average flow rate", labls, defts )
  if (xlist.count = 0) then
    exit
  end
  for each i in 0..4
    if (i=0)then
      continue
    end
    if (xlist.get(i).isnumber.not) then
      msgbox.error(xlist.get(i)+"is not a number",
        "Error in line "+i.asstring)
      continue
    end
  end
  end
  fy = xlist.get(1).asnumber
  ty = xlist.get(3).asnumber
  fm = xlist.get(2).asnumber
  tm = xlist.get(4).asnumber
  if (fy < 1960) then
    msgbox.error(xlist.get(1)+" - year out of range ( < 1960 )",
      "Error in year field ")
    continue
  end
  if (ty > 1991) then
    msgbox.error(xlist.get(3)+" - year out of range ( 1991 < )",
      "Error in year field ")
    continue
  end
  if ((fm < 1) or (fm>12)) then
    msgbox.error(xlist.get(2)+" - month out of range (1-12)",
      "Error in year field ")
    continue
  end
  if ((tm < 1) or (tm>12)) then
    msgbox.error(xlist.get(4)+" - month out of range (1-12)",
      "Error in year field ")
    continue
  end
  end
  break
end
theView = av.GetActiveDoc
lstact = theView.GetActiveThemes
for each t in lstact
  if ( t.getname = "unflow") then
    LTpar = {xlist.get(0), "Qm", false, fy, ty, fm, tm}
    av.run("schart1",LTpar)
    exit
  end
  if ( t.getname = "gsflow") then

```

```

    LTpar = {xlist.get(0), "M", false, fy, ty, fm, tm}
    av.run("schart1",LTpar)
    exit
  end
end
end

```

### **qmchar1u** Draws charts of the flow rate, property: Update

```

'Qmchar1u.ave
'multiple years data
'update event for "draw charts of flow"
' Is enabled when the theme is active and first
theView = av.GetActiveDoc
if (theView.GetEditableTheme <> NIL) then
  SELF.SetEnabled(FALSE)
  exit
end
xx = theView.GetActiveThemes
if(xx.count <> 0) then
  visnm = xx.get(0).getname
  if (( visnm = "unflow") or (visnm = "gsflow")) then
    SELF.SetEnabled(TRUE)
    exit
  end
end
SELF.SetEnabled(FALSE)

```

### **schart1** Draws charts, is executed from such scripts as pmchar1 and qmchar1 (Adopted from ESRII examples supplied with Avenue)

```

' MakeBarChartSpotSymbols
' title, prefix, one_year, fy, ty, fm, tm,
ltitle = self.get(0)
prefix = self.get(1)
one_year = self.get(2)
fy = self.get(3)
ty = self.get(4)
fm = self.get(5)
tm = self.get(6)
' title for legend
'ltitle = "Flow m3/s"
' prefix (field name = prefix + 01..12 or prefix + yr + mt)
'prefix = "M"
' one year will be displayed ???
'one_year = false
'fy = 1988
'fm = 1
'ty = 1990
'tm = 12
' optional legend color
if(one_year) then

```

```

    fy = 1989
    ty = 1989
    fm = 1
    tm = 12
end
'Script.The.SetNumberFormat("d.ddddd")
'The max. bar height and width are calculated as perc. of height of
'displayed area. The largest bar is drawn at this height and the rest
'are scaled accordingly. A value of 0.1 means the largest bar will be
'0.1 of the height of the current display. Four sizes (small, medium,
'large and x-large) are programmed as examples.
bar_size = MsgBox.ListAsString ({"Small","Medium","Large","X-
Large"},"Select bar size:","Bar Symbols")
if (bar_size = "Small") then
    max_bar_height = 0.1
    bar_width = 0.01
    legend_gridlines = 2
elseif (bar_size = "Medium") then
    max_bar_height = 0.2
    bar_width = 0.02
    legend_gridlines = 3
elseif (bar_size = "Large") then
    max_bar_height = 0.3
    bar_width = 0.03
    legend_gridlines = 4
elseif (bar_size = "X-Large") then
    max_bar_height = 0.4
    bar_width = 0.04
    legend_gridlines = 5
else
    exit 'Cancel selected'
end
end
'Number of gridlines on legend bar. Uncomment the following line
'to override above settings.
legend_gridlines = 4
legend_color = Color.GetCyan
'draw_gridlines = MsgBox.YesNo("Draw gridlines on the bars?","Bar
Symbols",False)
draw_gridlines = true
'Point size of text used in legend
leg_text_size = 8
'To ignore specific data values (eg., those that mean "nodata") put
'the values in this list and they will not be charted. This list
'should always contain 0, to prevent drawing zero size bars.
ignore_values = {0}
'List of colors
'Make your own custom colors here and make sure to add to both
'the color_list and also color_names...
pink = Color.Make
pink.SetRGBList({255,105,180})
olive = Color.Make
olive.SetRGBList({188,238,104})
orange = Color.Make
orange.SetRGBList({255,69,0})
gold = Color.Make

```

```

gold.SetRGBList({255,215,0})
maroon = Color.Make
maroon.SetRGBList({255,52,179})
color_list = {Color.GetBlue, Color.GetCyan, gold,
              Color.GetGray, Color.GetGreen, Color.GetMagenta,
maroon,
              olive, orange, pink, Color.GetRed, Color.GetYellow }
color_names = {"Blue", "Cyan", "Gold", "Gray", "Green", "Magenta",
              "Maroon", "Olive", "Orange", "Pink",
              "Red", "Yellow"}

'Number of colors
num_colors = 12
theView = av.GetActiveDoc
theProjection = theView.GetProjection
project_flag = theProjection.IsNil.Not 'true if projected
theTheme = theView.GetActiveThemes.Get(0) 'Get first active theme
theFTab = theTheme.GetFTab
shpfield = theFTab.FindField("Shape")
gra_list = theView.GetGraphics
'This section prompts the user for the fields in the
'FTab to generate the bar symbols from.
f= " "
max_field_prec = 0 'Field precision of selected fields
selected_fields = {} 'Fields selected by user
numeric_fields = {} 'Numeric fields in the FTab
field_aliases = {} 'List of field aliases
all_fields = theFTab.GetFields
'Build list of numeric fields from all fields.
'Also list of field aliases to display to user.
for each f in all_fields
  if (f.IsTypeNumber) then
    numeric_fields.Add(f)
    field_aliases.Add(f.GetAlias)
  end
end 'for
'Select fields for bar symbols
user_colors = {}
done = FALSE
field_count = 0
for each yr in fy..ty
  for each mt in 1..12
    if (( yr = fy) and ( mt < fm)) then
      continue
    end
    if (( yr = ty) and ( mt > tm)) then
      continue
    end

    if (mt < 10 ) then
      txt = "0"+mt.AsString
    else
      txt = mt.AsString
    end
    if(one_year) then
      fname = prefix+txt

```



```

    else
        fname = prefix+yr.AsString+txt
    end
    if (fname <> nil) then
        field_count = field_count + 1
        f = theFTab.findfield(fname)
        selected_fields.Add(f)
        user_colors.Add(color_list.Get(field_count - 1 mod num_colors))
        'Store max field precision of selected fields for later use.
        max_field_prec = (max_field_prec Max f.GetPrecision)
    else
        'Cancel selected.
        done = TRUE
    end 'if
end 'month
end 'year
field_count = selected_fields.Count
if (field_count = 0) then
    'User did not select any fields to symbolize
    exit
end
'Set bar height and width in page display units
view_extent = theView.GetDisplay.ReturnVisExtent
max_bar_height = view_extent.GetHeight * max_bar_height
bar_width = view_extent.GetHeight * bar_width
'Iterate thru selection bitmap or all records in FTab if no
selection.
iter = theFTab.GetSelection
num_records = iter.Count
'If no selected set, set iteration to all records in FTab.
if (num_records = 0) then
    iter = theFTab
    num_records = theFTab.GetNumRecords
end
min_height = 0
max_height = 0
for each rec in iter
    for each f in selected_fields
        cur_height = theFTab.ReturnValueNumber(f,rec)
        if (ignore_values.FindByValue(cur_height) > -1) then
            'Ignore values user wants to ignore
            continue
        end
        min_height = (min_height Min cur_height)
        max_height = (max_height Max cur_height)
    end 'for
end 'for
'Since the max height of a bar could be negative, find abs value of
'both min and max values to determine a scale factor. When building
the
'actual bar for a field, divide field value by this scale factor.
That
'way the maximum bar height (max_bar_height) will = the max field
value,
```

```

'whether positive or negative; all other field values are scaled
accordingly.
ht_scale_factor = (min_height.Abs) Max (max_height.Abs)
'Figure out break points for legend to allow adding gridlines to
legend
'and (optionally) actual bars on map. This set of code figures out a
'round number larger than the maximum field value, then divides that
'number by the number of legend classes desired. Then it rounds the
break
'point to a "nice" round number. E.g., given a max field value of
123,
'and 3 classes, this set of code will create a legend with break
points
'at 45, and a top value of 135.
if (ht_scale_factor < 1) then
  theLog = ht_scale_factor.Log(10).Floor
else
  theLog = ht_scale_factor.Log(10).Truncate
end 'if
x = ht_scale_factor / (10^(theLog))
y = (10 * x + 0.5).round * 10.^(theLog) / 10
x = y / legend_gridlines
if (x < 1) then
  theLog = x.Log(10).Floor
else
  theLog = x.Log(10).Truncate
end 'if
x2 = x / (10^(theLog))
legend_break_val = (2 * x2 + 0.5).round * 10.^(theLog) / 2
'Map value to display page units
legend_break_height = legend_break_val / ht_scale_factor *
max_bar_height
'Start drawing bar symbols
gra_list.UnselectAll
av.ShowStopButton
av.ShowMsg("Creating bar symbols...")
counter = 0
gra_group = GraphicGroup.Make
'For each selected record make the bars
for each rec in iter
  acolor = -1 'initialize color index
  'find the center of the feature to place bar symbol at
  ctr = theFtab.ReturnValue(shpfield,rec).ReturnCenter
  if (project_flag) then
    'Project the centroid if the view has a projection
    ctr = ctr.ReturnProjected(theProjection)
  end
  xcen = ctr.GetX
  ycen = ctr.GetY
  'Add line signaling zero in bar chart
  line_start_pt = (xcen - ((field_count * bar_width) / 2))@ycen
  line_end_pt = line_start_pt + ((bar_width * field_count)@0)
  aline = Line.Make(line_start_pt,line_end_pt)
  gr = GraphicShape.Make(aline)
  gra_group.Add(gr)

```

```

'set starting point for drawing bars
bar_start_pt = line_start_pt
bar_count = 0 'number of bars drawn
for each f in selected_fields
  bar_count = bar_count + 1
  acolor = acolor + 1
  'Convert field value to bar height;divide by height scale factor.
  thevalue = theFTab.ReturnValueNumber(f,rec)
  aheight = (thevalue / ht_scale_factor) * max_bar_height
  if (ignore_values.FindByValue(thevalue) > -1) then
    'Don't draw values user wants to ignore
    bar_start_pt = bar_start_pt + (bar_width@0)
    continue
  end 'if
  'draw the bar for the field
  abox = rect.Make(bar_start_pt,bar_width@aheight)
  gr = GraphicShape.Make(abox)
  'the symbol of a poly is a rasterfill, thus SetStyle request
works.
  gr.GetSymbol.SetStyle(#RASTERFILL_STYLE_SOLID)
  gr.GetSymbol.SetColor(user_colors.Get(acolor))
  gra_group.Add(gr)
  'Draw reference lines on bar to indicate value
  if (draw_gridlines) then
    num_tics = (aheight.Abs / legend_break_height).Floor
    c = 0
    while (c <> num_tics)
      c = c + 1
      if (aheight < 0) then
        tic_start_pt = bar_start_pt - (0@(c * legend_break_height))
      else
        tic_start_pt = bar_start_pt + (0@(c * legend_break_height))
      end 'if
      tic_end_pt = tic_start_pt + (bar_width@0)
      aline = Line.Make (tic_start_pt,tic_end_pt)
      gr = GraphicShape.Make(aline)
      gra_group.Add(gr)
    end 'while
  end 'if
  bar_start_pt = bar_start_pt + (bar_width@0)
end 'for each f in selected_fields
'Clone bar chart and add it to the graphic list. Then reuse graphic
group.
'AddBatch used so that once ALL bar symbols for ALL records are
created
'they are drawn to the screen. Using Add request is much slower.
gg = gra_group.Clone
gra_list.AddBatch(gg)
'Add bar chart to the Theme's graphic list so that when the
'theme is not drawn, the bar symbols won't draw either.
theTheme.GetGraphics.Add(gg)
gra_group.Empty
counter = counter + 1
progress = (counter / num_records) * 100
more = av.SetStatus(progress)

```

```

    if (Not more) then
        'Break if user presses Stop button.
        break
    end
end ' for each rec in iter
gra_list.UnselectAll
if (not ((num_records = 1) and (aheight = 0))) then
    'Special case where only 1 record selected, and the total of the
    fields
    'of the record is zero; essentially no bars are drawn, so endbatch
    'should not be executed.
    gra_list.EndBatch
end
'Start building the legend.
'Set text size for legend
text_sym = av.GetSymbolWin.ReturnCurrentSymbol(#SYMBOL_TEXT).clone
text_sym.SetSize (leg_text_size)
'Add dummy text to get an idea for spacing of graphic elements
'Spacing of objects in legend = current point size of text.
gr = GraphicText.Make("ABC",0@0)
gr.SetSymbol (text_sym)
gr.SetSelected(TRUE)
gra_list.Add(gr)
spacing = gra_list.ReturnSelectedExtent.GetHeight
gra_list.RemoveGraphic(gr)
'Set number of decimal places for legend text equal to max field
precision
'This code builds a string like "d.ddd"
numformat = "d."
if (max_field_prec <> 0) then
    for each i in 1..max_field_prec
        numformat = numformat+"d"
    end
end
end 'if
'Draw legend at (0,0) then move it to the correct location
pnt = 0@0
abox = Rect.Make (pnt,bar_width@(legend_gridlines *
legend_break_height))
gr = GraphicShape.Make(abox)
gr.GetSymbol.SetStyle(#RASTERFILL_STYLE_SOLID)
gr.GetSymbol.SetColor(legend_color)
gra_group.Add(gr)
'Add "0" text in legend
gr = GraphicText.Make(0.SetFormat(numformat).AsString,
    (pnt + ((bar_width + spacing)*(-spacing / 2))))
gr.SetSymbol (text_sym)
gra_group.Add(gr)
'Draw a reference lines on bar.
for each i in 1..legend_gridlines
    line_start_pt = 0@(i * legend_break_height)
    line_end_pt = line_start_pt + (bar_width@0)
    aline = Line.Make (line_start_pt,line_end_pt)
    gr = GraphicShape.Make(aline)
    gra_group.Add(gr)
    'Add text indicating value represented by square

```

```

atext = (legend_break_val * i).SetFormat(numformat).AsString
'Add text string to right of line
gr = GraphicText.Make(atext,(line_end_pt + (spacing@(-spacing /
2))))
gr.SetSymbol (text_sym)
gra_group.Add(gr)
'gra_group.SetSelected (TRUE)
'Clone graphic group;otherwise it would get deleted on Empty
'gra_list.Add(gra_group.clone)
'gra_group.Empty
end 'for
gra_group.SetSelected (TRUE)
'Clone graphic group;otherwise it would get deleted on Empty
gra_list.Add(gra_group.clone)
gra_group.Empty
gra_list.UngroupSelected 'Workaround
gra_list.GroupSelected
extent = gra_list.ReturnSelectedExtent
x = extent.ReturnOrigin.GetX
y = extent.GetTop + spacing
gr = GraphicText.Make(ltitle,(0@y))
'gr = GraphicText.Make("Flow Rate",(x@y + (x@0)))
'gr = GraphicText.Make("TEXT EXAMPLE",(x@y + ((spacing * 2)@0)))
gr.SetSymbol (text_sym)
gra_group.Add(gr)
' acolor = acolor + 1 'Get next color
' y = y + (spacing * 2) 'Move y up for next box
'end 'for
gra_group.SetSelected(TRUE)
gra_list.Add(gra_group)
gra_list.UngroupSelected 'Workaround
gra_list.GroupSelected
'Draw white box to place legend in
extent = gra_list.ReturnSelectedExtent
origin = extent.ReturnOrigin - (spacing@spacing)
size = extent.ReturnSize + ((2*spacing)@(2*spacing))
arect = Rect.Make(origin,size)
gr = GraphicShape.Make(arect)
gr.GetSymbol.SetStyle(#RASTERFILL_STYLE_SOLID)
gr.GetSymbol.SetColor(Color.GetWhite)
gra_list.Add(gr)
gra_list.MoveSelectedToFront
gr.SetSelected(TRUE)
gra_list.GroupSelected
'Move legend to lower-left corner of View
alegend = gra_list.GetSelected.Get(0)
alegend.Invalidate
alegend.SetOrigin(theView.GetDisplay.ReturnVisExtent.ReturnOrigin)
alegend.Invalidate
'Merge the graphic legend into the theme's graphic list
theTheme.GetGraphics.Merge(theView.GetGraphics.GetSelected)
gra_list.UnselectAll
av.ClearMsg
av.ClearStatus

```

**movie2      Displays charts in a sentence, category: tools, property: apply (script  
qmchar1u is used in the property: Update)**

```
'movie2.ave
theView = av.GetActiveDoc
theThemes = theView.GetActiveThemes
p = theView.GetDisplay.ReturnUserPoint
t = theThemes.get(0)
if (t.CanSelect.not) then
    exit
end
t.SelectByPoint(p,#VTAB_SELTYPE_NEW )
recs = t.FindByPoint(p)
if(recs.count = 0) then
    exit
end
vv = t.getftab.updateselection
selchart = {"Horizontal", "Vertical"}
c = msgbox.choiceasstring(selchart, "Make a selection",
    "Chart type selection")
if (c = nil) then
    exit
end
theView = av.GetActiveDoc
visnm = theView.GetVisibleThemes.get(0).getname
if ( visnm = "Gsflow") then
    prefix = "M"
    nnote = "Available: 1940 - 1991"
    if(c = "Horizontal") then
        TheChart = av.getproject.FindDoc("MovieMh")
    end
    if(c = "Vertical") then
        TheChart = av.getproject.FindDoc("MovieMv")
    end
end
if ( visnm = "unflow") then
    prefix = "Qm"
    nnote = "Available: 1960 - 1991"
    if(c = "Horizontal") then
        TheChart = av.getproject.FindDoc("MovieQmh")
    end
    if(c = "Vertical") then
        TheChart = av.getproject.FindDoc("MovieQmv")
    end
end
end
CHtable = TheChart.GetVtab
sp = "xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx"
ln4 = "First year:"
ln5 = "Last year:"
labls = {sp, ln4,ln5}
fy = 1960
```

```

ty = 1991
fm = 1
tm = 12
defts = {" ", fy.asstring, ty.asstring}
while (true)
  xlist = msgBox.Multiinput(xnote,"Movie2", labls, defts )
  if (xlist.count = 0) then
    exit
  end
  if ((xlist.get(1).isnumber.not) or (xlist.get(1) < 1940)) then
    msgbox.error(xlist.get(1)+"is not a correct year", "Error in
field #1")
    continue
  end
  if ((xlist.get(2).isnumber.not) or (xlist.get(2) > 1991)) then
    msgbox.error(xlist.get(2)+"is not a correct year", "Error in
field #2")
    continue
  end
  break
end
av.showmsg(" Looking for the maximum value to set Y scale . . .")
fy = xlist.get(1).asnumber
ty = xlist.get(2).asnumber
bmap = Chtable.getselection
rec = -1
rec = bmap.getnextset(rec)
step = -1
lablst = list.make
'for each i in 0..11
' lablst.add("x")
'end
'==== find maximum value to set Y(or X) scale =====
xxmax = 0
for each yr in fy..ty
  for each mt in 1..12
    if (( yr = fy) and ( mt < fm)) then
      continue
    end
    if (( yr = ty) and ( mt > tm)) then
      continue
    end
    step = step + 1
    'if ( step < 12 ) then
    ' continue
    'end
    if (mt < 10 ) then
      txt = "0" + mt.AsString
    else
      txt = mt.AsString
    end
    name = prefix + yr.AsString + txt
    XFld=Chtable.FindField(name)
    xx22 = Chtable.returnvalueNumber(xfld,rec)
    if ( xx22 > xxmax ) then

```

```

        xxmax = xx22
    end
end
end
nomonths = step
msgbox.info("Maximum value is"++xxmax.asstring,"Maximum")
av.clearmsg
av.showstopbutton
if(thechart.getwin.isopen.not) then
    thechart.getwin.open
    thechart.getwin.minimize
end
TheList=TheChart.GetFields
TheFList = TheChart.getVtab.GetFields
if(c = "Horizontal") then
    thechart.getYaxis.setboundsmax(xxmax)
end
if(c = "Vertical") then
    thechart.getXaxis.setboundsmax(xxmax)
end
step = -1
for each yr in fy..ty
av.showmsg(" YEAR: "++yr.asstring)
    for each mt in 1..12
        if (( yr = fy) and ( mt < fm)) then
            continue
        end
        if (( yr = ty) and ( mt > tm)) then
            continue
        end
        step = step + 1
        if ( step = 12 ) then
            if(msgbox.miniYesNo("Start ?", True).not) then
                exit
            end
            thechart.getwin.restore
            if(thechart.getwin.isopen.not) then
                thechart.getwin.activate
            end
        end
        if (mt < 10 ) then
            txt = "0" + mt.AsString
        else
            txt = mt.AsString
        end
        xstatus = av.setstatus(100 * step / nomonths)
        if(xstatus.not) then
            rlt = msgbox.yesno("Do you really want to stop?","Stop movie",
true)
            if(rlt) then
                rlt2 = msgbox.yesno("Do you want to minimize the chart
window?",
                "Close Window ?", true)
                if(rlt2) then
                    thechart.getwin.minimize

```



```

        end
        exit
    else
        av.ClearStatus
        av.showstopbutton
    end
end
name = prefix + yr.AsString + txt
NewTitle = txt + " / " + yr.asstring
NewFld=TheChart.getVtab.FindField(name)
TheList.Remove(0)
Thelist.add(newfld)
TheChart.GetTitle.SetName(NewTitle)
theChart.SetSeriesFromRecords(true)
theChart.Getwin.Open
end
end
av.ClearStatus
r1t2 = msgbox.yesno("Do you want to minimize the chart?",
    "Close Window ?", true)
if(r1t2) then
    thechart.getwin.minimize
end

```

**aliaset**      **Sets the alias names for fields, that describe a chart axis.**

```

'aliaset.ave
xFtable = av.GetActiveDoc.getvtab
prefix = "QM"
fy = 1960
ty = 1992
fm = 1
tm = 9
for each yr in fy..ty
    for each mt in 1..12
        if (( yr = fy) and ( mt < fm)) then
            continue
        end
        if (( yr = ty) and ( mt > tm)) then
            continue
        end
        if (mt < 10 ) then
            txt = "0" + mt.AsString
        else
            txt = mt.AsString
        end
        name = prefix + yr.AsString + txt
        XFld=xFtable.FindField(name)
txalias = mt.asstring+"/"+yr.asstring
xxx = XFld.setalias(txalias)
    end
end

```

## B5 Scripts from the project “tools”

### order6 Determines the order of a stream/watershed in a flow system

```
'order6.ave  Adds a field that contains a numbering
'            system to specify the stream order
'            in a flow system.
TBunit = av.GetActiveDoc
xxx = TBunit.getGUI
if ( xxx <> "Table") then
    xxx2 = TBunit.asstring++"is active"
    msgbox.warning("A Table must be active", xxx2)
    exit
end
VTunit = TBunit.getVtab
'assume the default fields in UNIT table:
TXunit = "Unit"
TXnext = "Downstream"
FDunit = VTunit.FindField(TXunit)
FDnext = VTunit.FindField(TXnext)
'create list of fields that are in VTunit
LSFDunit = VTunit.GetFields
'if the default fields can not be found, ask user for required
fields:
if(FDunit=nil) then
    'Select current_unit id field
    FDunit = MsgBox.List(LSFDunit, "Select a field that contains"
        ++NL++" the watershed/stream ID",
        "Unit_id Selection")
end
'if not selected, quit
IF ( FDunit = nil) then
    exit
end
if(FDnext=nil) then
    'Select downstream_unit id field
    FDnext = MsgBox.List(LSFDunit, "Select a field that contains"++NL++
        "the downstream unit ID", "Next_id Selection")
end
IF ( FDnext = nil) then
    exit
end
TXorderField = msgbox.input("Enter the name of order field ",
    "Order Field Selection", "Order")
FDorder = vtunit.FindField(TXorderField)
xx = false
if (FDorder = nil) then
    xx = msgbox.YesNo( TXorderField++"field is not in the table "
        ++TBunit.asString+nl+"Do you want to create this field ?",
        "Field not Found !", true)
'end
if (xx.not) then
```

```

        exit
    end
    'check if the VTunit can be edited if not, set it editable
    if(VTunit.isEditable.not) then
        VTunit.SetEditable(true)
    end
    if(VTunit.CanAddFields) then
        f1 = Field.Make( TXorderfield,#FIELD_SHORT, 5, 0 )
        'Add fields
        VTunit.AddFields({f1})
    else
        msgBox.warning("Can not add a field to"++VTunit.asString,
            "Can not edit table")
        exit
    end
else
    f1 = FDorder
    if(VTunit.isEditable.not) then
        VTunit.SetEditable(true)
    end
end
LSunit = list.make
LSnext = list.make
LSord = list.make
LSrecnxt = list.make
nx = VTunit.GetNumRecords
nn = nx - 1
nst = 100/nx    'variable used in status displaying
av.ShowMsg("Creating lists ...")
for each rec in VTunit
    av.ShowMsg("Creating lists ..."++rec.AsString)
    av.SetStatus(rec * nst)
    LSnext.Add(VTunit.ReturnValueNumber(FDnext, rec))
    LSunit.Add(VTunit.ReturnValueNumber(FDunit, rec))
    LSord.Add(1)
    LSrecnxt.Add(nx)
end
LSrecnxt.Add(nx)
'search for first order streams/watersheds and create list of the
'record numbers pointed by the "next/downstream unit" field
for each i in 0..nn
    av.ShowMsg("Searching for first order streams/watersheds
..."+i.asString)
    av.SetStatus(i * nst)
    for each j in 0..nn
        if ( LSnext.get(i) = LSunit.get(j) ) then
            LSord.set(j , 0 )
            LSrecnxt.set(i , j)
            Break
        end
    end
end
end
'write downstream record pointers into field f2
'for each rec in VTunit
' VTunit.SetValue( f2, rec, LSrecnxt.Get( rec ))

```

```

'end
'determine order of the rest of streams
for each i in 0..nn
  av.ShowMsg("Calculating order ..."++i.asString)
  av.SetStatus(i * nst)
  if ( LSord.get(i) <> 1 ) then
    continue
  end
  j = LSreclst.get(i)
  k = 2
  while ( j <> nx)
    if ( LSord.get(j) >= k ) then
      break
    end
    LSord.set(j , k )
    j = LSreclst.get(j)
    k = k + 1
  end 'while
end 'i
'write stream/watershed order into Table
for each rec in VTunit
  VTunit.SetValue( fl, rec, LSord.Get( rec ))
end
'refresh table and set it "not editable"
VTunit.Refresh
if(VTunit.isEditable) then
  VTunit.SetEditable(false)
end
end

```

## **upavg2      calculates weighted average over upstream units**

```

' upavg2.ave calculates weighted average
'   for upstream units:
'   output = upstr_sum(input*wght)/upstr_sum(wght)
'=====
' note: dbf table contains a dummy (first) record
' that can influence calculations !!!
TBunit = av.GetActiveDoc
xxx = TBunit.getGUI
if ( xxx <> "Table") then
xxx2 = TBunit.asstring++"is active"
msgbox.warning("A Table must be active", xxx2)
exit
end
TheFtab = TBunit.getVtab
' Ask user for required fields:
TXunit = "unit_id"
TXnext = "unit_nx"
TXorder = "order"
TXitem = "input"
TXadd = "output"
TXwgt = "weight"
FDunit = theFtab.FindField(TXunit)

```

```

FDnext = theFtab.FindField(TXnext)
FDorder = theFtab.FindField(TXorder)
FDadd = theFtab.FindField(TXadd)
FDitem = theFtab.FindField(TXitem)
FDarea = theFtab.FindField(TXwgt)
' Make a list of VTunit fields
LSFDunit = theFtab.GetFields
' If the default fields can not be found, ask user
' for the required fields:
if(FDunit=nil) then
  ' Select current_unit id field (from entity)
  FDunit = MsgBox.List(LSFDunit, "Select a field that contains"
    ++NL++"the watershed/stream ID", "From-unit Selection")
end
' If not selected, quit
IF (FDunit = nil) then
  exit
end
if(FDnext=nil) then
  ' Select downstream_unit id field
  FDnext = MsgBox.List(LSFDunit, "Select a field that contains"
    ++NL++" the downstream unit ID", "To-unit Selection")
end
' If not selected, quit
IF ( FDnext = nil) then
  exit
end
if(FDorder=nil) then
  ' Select order field
  FDorder = MsgBox.List(LSFDunit, "Select field that contains"
    ++NL++" the watershed/stream order","Order field Selection")
end
' If not selected, quit
IF ( FDorder = nil) then
  xxx = msgbox.info("Create a field of stream order",
    "Missing fields")
  exit
end
if(FDitem=nil) then
  ' Select field that contains values to be averaged
  FDitem = MsgBox.List(LSFDunit, "Select a field that contains"
    ++NL++"values to be averaged", "Feature Selection")
end
' If not selected, quit
IF ( FDitem = nil) then
  exit
end
if(FDarea=nil) then
  ' Select field that contains weight
  FDarea = MsgBox.List(LSFDunit, "Select a field that contains"
    ++NL++"weight", "Weight Selection")
end
' If not selected, quit
IF ( FDarea = nil) then
  exit

```

```

end
' Prompt for a name (name of the field, that will be created)
TXadd = msgbox.input("Enter the name of output field ",
"Output Field Selection", "output")
if (TXadd = nil) then
    exit
end
FDadd = theFtab.FindField(TXadd)
xx = false
if (FDadd = nil) then
    xx = msgbox.YesNo( TXadd++"field is not in the table "
    ++TBunit.asString+nl+"Do you want to create this field ?",
    "Field not Found !", true)
'end
if (xx.not) then
    exit
end
'check the setting of "input" field
fenum = FDitem.gettype
fprec = FDitem.getprecision
fwidth = FDitem.getwidth
xxx1 = msgbox.input("Please specify precision"++
    "(decimal part) of the output field",
    "Width of the new field" , fprec.asstring)
fprec = xxx1.asnumber
IF ( fprec = nil) then
    exit
end
xxx2 = msgbox.input("Please specify total width of the output
field",
    "Width of the new field" , fwidth.asstring)
fwidth = xxx2.asnumber
IF ( fwidth = nil) then
    exit
end
'check if the theFtab can be edited if not, set it editable
if(theFtab.isEditable.not) then
    theFtab.SetEditable(true)
end
' Create a new field
if(theFtab.CanAddFields) then
    FDadd = field.Make(TXadd, fenum, fwidth, fprec)
    theFtab.AddFields({FDadd})
else
    msgBox.warning("Can not add a field to"++theFtab.asString,
    "Can not edit table")
    theFtab.SetEditable(false)
    exit
end
else
if(theFtab.isEditable.not) then
    theFtab.SetEditable(true)
end
end
'create lists

```

```

ltunit = list.make
ltnext = list.make
ltorder = list.make
ltitem = list.make
ltadd = list.make
unnx = theFTab.GetNumRecords
unnn = unnx - 1
dcarea = dictionary.make(unnx)
dcmass = dictionary.make(unnx)
for each rec in theFTab
  Nunit = theFTab.ReturnValueNumber(FDunit, rec)
  ltunit.Add(Nunit)
  ltnext.Add(theFTab.ReturnValueNumber(FDnext, rec))
  ltorder.Add(theFTab.ReturnValueNumber(FDorder, rec))
  Narea = theFTab.ReturnValueNumber(FDarea, rec)
  if (Narea = 0 ) then
    msgbox.error("The weight can not be zero !"++
      NL++"record:"++rec.asstring++"field:"++
      FDarea.asstring, "Division by zero" )
    theFtab.SetEditable(false)
    exit
  end
  Nmass = theFTab.ReturnValueNumber(FDitem, rec) * Narea
  dcarea.add(Nunit, Narea )
  dcmass.add(Nunit, Nmass )
end
' find maximum order
unmaxord = 1
for each i in 0..unnn
  av.ShowMsg("Finding maximum order ..."++i.asString)
  istatus = i * 100 / unnn
  av.SetStatus(istatus)
  if (ltorder.get(i) > unmaxord ) then
    unmaxord = ltorder.get(i)
  end
end
'calculate cumulative values
unmaxord1 = unmaxord - 1
for each k in 1..unmaxord1
  av.ShowMsg("Calculating cumulative values ..."++k.asString)
  istatus = k * 100 / unmaxord
  av.SetStatus(istatus)
  for each i in 0..unnn
    if ( ltorder.get(i) <> k ) then
      continue
    end
    Nunit = ltunit.get(i)
    Nnext = ltnext.get(i)
    Nxmass = dcmass.get(Nunit) + dcmass.get(Nnext)
    Nxarea = dcarea.get(Nunit) + dcarea.get(Nnext)
    dcmass.set(Nnext, Nxmass)
    dcarea.set(Nnext, Nxarea)
  end
end
av.clearMsg

```

```

av.clearstatus
'write results to Ftab
for each rec in theFtab
  Nunit = theFtab.ReturnValueNumber(FDunit, rec)
  Xavg = dcmass.get(Nunit) / dcareas.get(Nunit)
  TheFtab.SetValue( FDadd, rec, Xavg )
end
theFtab.Refresh
if(TheFtab.isEditable) then
  theFtab.SetEditable(false)
end
end

```

## **cumul2      Accumulates values, going downstream (flow accumulation)**

```

' cumul2.ave Adds a field or writes to the existing
'           field the accumulated values. Summation
'           is performed downstream, along the flow
'           path.
' note: dbf table contains a dummy (first) record
' that can influence calculations !!!
TBunit = av.GetActiveDoc
xxx = TBunit.getGUI
if ( xxx <> "Table") then
xxx2 = TBunit.asstring++"is active"
msgbox.warning("A Table must be active", xxx2)
exit
end
TheFtab = TBunit.getVtab
' Ask user for needed fields:
TXunit = "unit_id"
TXnext = "unit_nx"
TXorder = "order"
TXitem = "item_x"
FDunit = theFtab.FindField(TXunit)
FDnext = theFtab.FindField(TXnext)
FDorder = theFtab.FindField(TXorder)
'FDadd = theFtab.FindField(TXadd)
FDitem = theFtab.FindField(TXitem)
' Make a list of VTunit fields
LSFDunit = theFtab.GetFields
' If the default fields can not be found, ask user for
' the required fields:
if(FDunit=nil) then
  ' Select current_unit id field (from entity)
  FDunit = MsgBox.List(LSFDunit, "Select a field that contains"
    ++NL++"the watershed/stream ID", "From-unit Selection")
end
' If not selected, quit
IF (FDunit = nil) then
  exit
end
if(FDnext=nil) then
  ' Select downstream_unit id field
  FDnext = MsgBox.List(LSFDunit, "Select a field that contains"

```



```

        ++NL++" the downstream unit ID", "To-unit Selection")
end
' If not selected, quit
IF ( FDnext = nil) then
    exit
end
if(FDorder=nil) then
    ' Select order field
    FDorder = MsgBox.List(LSFDunit, "Select field that contains"
        ++NL++" the watershed/stream order","Order field Selection")
end
' If not selected, quit
IF ( FDorder = nil) then
    xxx = msgbox.info("Create a field of stream order",
        "Missing fields")
    exit
end
if(FDitem=nil) then
    ' Select field that contains values to be accumulated
    FDitem = MsgBox.List(LSFDunit, "Select a field that contains"
        ++NL++"values to be summed", "Feature Selection")
end
' If not selected, quit
IF ( FDitem = nil) then
    exit
end
' Prompt for a name (name of the field, that will be created)
TXadd = msgbox.input("Enter the name of output field ",
    "Output Field Selection", "Accumulated")
FDadd = theFtab.FindField(TXadd)
xx = false
if (FDadd = nil) then
    xx = msgbox.YesNo( TXadd++"field is not in the table "
        ++TBunit.asString+nl+"Do you want to create this field ?",
        "Field not Found !", true)
'end
if (xx.not) then
    exit
end
'check the setting of "item" field
'ESRI gives no information how wide a field can be!
fenum = FDitem.gettype
fprec = FDitem.getprecision
fwidth = FDitem.getwidth
if(fwidth < 13) then
    fwidth = fwidth + 3
elseif (fwidth = 13) then
    fwidth = 15
elseif (fwidth > 13) then
    xxx = msgbox.input("Please specify a width of output field",
        "Width of the new field" , fwidth.asstring)
    fwidth = xxx.asnumber
end
IF ( fwidth = nil) then
    exit

```

```

end
'check if the theFtab can be edited if not, set it editable
if(theFtab.isEditable.not) then
    theFtab.SetEditable(true)
end
' Create a new field
if(theFtab.CanAddFields) then
    FDadd = field.Make(TXadd, fenum, fwidth, fprec)
    theFtab.AddFields({FDadd})
else
    msgBox.warning("Can not add a field to"++theFtab.asString,
        "Can not edit table")
    exit
end
else
    if(theFtab.isEditable.not) then
        theFtab.SetEditable(true)
    end
end
'create lists
ltunit = list.make
ltnext = list.make
ltorder = list.make
ltitem = list.make
ltadd = list.make
unnx = theFtab.GetNumRecords
unnn = unnx - 1
dcadd = dictionary.make(unnx)
for each rec in theFtab
    Nunit = theFtab.ReturnValueNumber(FDunit, rec)
    ltunit.Add(Nunit)
    ltnext.Add(theFtab.ReturnValueNumber(FDnext, rec))
    ltorder.Add(theFtab.ReturnValueNumber(FDorder, rec))
    Nitem = theFtab.ReturnValueNumber(FDitem, rec)
' ltitem.Add(Nitem)
    dcadd.add(Nunit, Nitem )
end
' find maximum order
unmaxord = 1
for each i in 0..unnn
    av.ShowMsg("Finding maximum order ..."++i.asString)
    istatus = i * 100 / unnn
    av.SetStatus(istatus)
    if (ltorder.get(i) > unmaxord ) then
        unmaxord = ltorder.get(i)
    end
end
'calculate cumulative values
unmaxord1 = unmaxord - 1
for each k in 1..unmaxord1
    av.ShowMsg("Calculating cumulative values ..."++k.asString)
    istatus = k * 100 / unmaxord
    av.SetStatus(istatus)
    for each i in 0..unnn
        if ( ltorder.get(i) <> k ) then

```

```

        continue
    end
    Nunit = ltunit.get(i)
    Nnext = ltnext.get(i)
    Nxvalue = dcadd.get(Nunit) + dcadd.get(Nnext)
    dcadd.set(Nnext, Nxvalue)
end
end
av.clearMsg
av.clearStatus
'write results to Ftab
for each rec in theFtab
    Nunit = theFtab.ReturnValueNumber(FDunit, rec)
    TheFtab.SetValue( FDadd, rec, dcadd.get(Nunit) )
end
theFtab.Refresh
if(TheFtab.isEditable) then
    theFtab.SetEditable(false)
end
end

```

**decom2**      **calculates the difference between the inflows and the outflow, the reverse process to the flow accumulation**

```

' decom2.ave for all modeling units calculates
' the difference between sum of inputs and the
' output ( deaccumulate values)
' note: dbf table contains a dummy (first) record
' that can influence the calculations !!!
TBunit = av.GetActiveDoc
xxx = TBunit.getGUI
if ( xxx <> "Table") then
xxx2 = TBunit.asstring++"is active"
msgbox.warning("A Table must be active", xxx2)
exit
end
TheFtab = TBunit.getVtab
' Ask user for required fields
TXunit = "unit_id"
TXnext = "unit_nx"
TXorder = "order"
TXitem = "input"
'TXadd = "decom_fd"
FDunit = theFtab.FindField(TXunit)
FDnext = theFtab.FindField(TXnext)
FDorder = theFtab.FindField(TXorder)
'FDadd = theFtab.FindField(TXadd)
FDitem = theFtab.FindField(TXitem)
' Make a list of VTunit fields
LSFDunit = theFtab.GetFields
' If the default fields can not be found, ask user for
' the required fields:
if(FDunit=nil) then
    ' Select current_unit id field (from entity)

```

```

    FDunit = MsgBox.List(LSFDunit, "Select a field that contains"
        ++NL++"the watershed/stream ID", "From-unit Selection")
end
' If not selected, quit
IF (FDunit = nil) then
    exit
end
if(FDnext=nil) then
    ' Select downstream_unit id field
    FDnext = MsgBox.List(LSFDunit, "Select a field that contains"
        ++NL++" the downstream unit ID", "To-unit Selection")
end
' If not selected, quit
IF ( FDnext = nil) then
    exit
end
if(FDorder=nil) then
    ' Select order field
    FDorder = MsgBox.List(LSFDunit, "Select field that contains"
        ++NL++" the watershed/stream order","Order field Selection")
end
' If not selected, quit
IF ( FDorder = nil) then
    xxx = msgbox.info("Create a field of stream order",
        "Missing fields")
    exit
end
if(FDitem=nil) then
    ' Select a field that contains values to be decomposed
    FDitem = MsgBox.List(LSFDunit,
        "Select a field that contains"++NL++
        "the values to be decomposed",
        "Input Field Selection")
end
' If not selected, quit
IF ( FDitem = nil) then
    exit
end
' Prompt for a name (name of the field, that will be created)
TXadd = msgbox.input("Enter the name of the output field ",
    "Output Field Selection", "Decomposed")
FDadd = theFtab.FindField(TXadd)
xx = false
if (FDadd = nil) then
    xx = msgbox.YesNo( TXadd++"field is not in the table "
        ++TBunit.asString+nl+"Do you want to create this field ?",
        "Field not Found !", true)
'end
    if (xx.not) then
        exit
    end
    'check the setting of "item" field
    'width and the precision of the input field are assumed
    fenum = FDitem.gettype
    fprec = FDitem.getprecision

```

```

fwidth = FDitem.getwidth
' Create a new field
if(theFtab.isEditable.not) then
  theFtab.SetEditable(true)
end
if(theFtab.CanAddFields) then
  FDadd = field.Make(TXadd, fenum, fwidth, fprec )
  theFtab.AddFields({FDadd})
else
  msgBox.warning("Can not add a field to"
    ++theFtab.asString, "Can not edit table")
  exit
end
else
  if(theFtab.isEditable.not) then
    theFtab.SetEditable(true)
  end
end
end
'create lists
ltunit = list.make
ltnext = list.make
ltorder = list.make
ltitem = list.make
ltadd = list.make
unnx = theFtab.GetNumRecords
unnn = unnx - 1
dcadd = dictionary.make(unnx)
dcdec = dictionary.make(unnx)
for each rec in theFtab
  Nunit = theFtab.ReturnValueNumber(FDunit, rec)
  ltunit.Add(Nunit)
  ltnext.Add(theFtab.ReturnValueNumber(FDnext, rec))
  ltorder.Add(theFtab.ReturnValueNumber(FDorder, rec))
  Nitem = theFtab.ReturnValueNumber(FDitem, rec)
  dcadd.add(Nunit, Nitem )
  dcdec.add(Nunit, Nitem )
end
' find maximum order
unmaxord = 1
for each i in 0..unnn
  av.ShowMsg("Finding maximum order ..."++i.asString)
  istatus = i * 100 / unnn
  av.SetStatus(istatus)
  if (ltorder.get(i) > unmaxord ) then
    unmaxord = ltorder.get(i)
  end
end
'calculate decomposed values
unmaxord1 = unmaxord - 1
for each k in 1..unmaxord1
  av.ShowMsg("Decumulatng values ..."++k.asString)
  istatus = k * 100 / unmaxord
  av.SetStatus(istatus)
  for each i in 0..unnn
    if ( ltorder.get(i) <> k ) then

```

```
        continue
    end
    Nunit = ltunit.get(i)
    Nnext = ltnext.get(i)
    Nxvalue = dcdec.get(Nnext) - dcadd.get(Nunit)
    dcdec.set(Nnext, Nxvalue)
end
end
av.clearMsg
av.clearstatus
'write results to Ftab
for each rec in theFtab
    Nunit = theFtab.ReturnValueNumber(FDunit, rec)
    TheFtab.SetValue( FDadd, rec, dcdec.get(Nunit) )
end
theFtab.Refresh
if(TheFtab.isEditable) then
    theFtab.SetEditable(false)
end
```

## **Appendix C      Arc/Info macros--AMLs**

This Appendix contains the following Arc/Info procedures written in Arc/Info Macro Language (AML):

- 1) WSHGS.AML - Creates the following grids: stream network, watersheds, and watershed outlets;
- 2) NEXTWSH.AML - Creates an info file that contains two items: ID and the downstream watershed ID. Also creates a grid of watershed outlets that contains the ID of downstream watershed;
- 3) RAININFO.AML- Based on the time series stored in the attribute table of the point coverage, and the grid that specify the spatial division into zones, this AML calculates zonal averages and stores them in Arc/Info INFO table;
- 4) RAINMAP.AML - Creates grids of average monthly precipitation;
- 5) RAINM2.AML - Creates an INFO table from the grids of average monthly precipitation;
- 6) SELDATA.AML - Creates point coverage of stations with complete record of flow rate/precipitation depth in selected months;
- 7) FD4Y.AML - Supports the process of redistribution of the flow rate recorded in theUSGS stations;
- 8) SLOPE3.AML - Calculates the slope along a flow path;
- 9) DEMALB - Converts a grid into Albers Equal Area coordinate system;

## C1 WSHGS.AML

```
/* -----
/* Program: WSHGS.AML
/* Purpose: Creates the following grids: stream network,
/*          watersheds, and watershed outlets.
/*          The watershed outlet grid contains:
/*          the most upstream cell of the first order stream,
/*          the most downstream cell of each reach, and the cells
/*          created from a point coverage of USGS stations.
/*          Runs only in GRID
/* -----
/* Usage: &r WSHINFO <fdir> <facc> <fbas> <trsh> <fout> <item>
/*          <fwsh> <flpp> <fstr>
/* Arguments: (input)
/*          fdir = (grid) flow direction
/*          facc = (grid) flow accumulation
/*          fbas = (grid) area under investigation (basin, mask)
/*          trsh = (value) threshold for stream delineation
/*          ( if grids are in meters, then trsh is in km2 )
/*          fout = (point) selected flow gauging stations
/*          (additional pour points for watersheds delineation)
/*          item = (text) name of the item from <fout> which will
/*          be used in "pointgrid" conversion.
/*
/*          (output, grids that will be created)
/*          fstr = (grid) stream network
/*          fwsh = (grid) drainage areas
/*          flpp = (grid) outlets
/*
/* Temporary: lso, lsl, lmx, lpd, lmn, lpu, lgs, lpx, (grids)
/* Globals: none (see Temporary)
/* Locals: tresh (variable)
/* -----
/* Notes: The first order watersheds (and the respective outlets)
/*          have ID number in the range 100000 < ID < 200000.
/*          Cellsize and Window of <fdir> are assumed.
/*          This AML checks neither for the existence and
/*          correctness of the input files nor for the existence
/*          of files that have the same names as the temporary
/*          files and the files to be created.
/*          If an error occurs then all grids and all info files that
/*          have the same name as output grids and temporary grids are
/*          erased !!!
/* -----
/* History: Author Pawel Mizgalewicz 11/21/93
/*          updated for point coverage 05/15/95
/*          working name net4.aml ( tested for the Cedar River
/*          - Iowa River basin, grid = 3000^2 cells, basin =
/*          3.2*10e6 cells of size 100m*100m).
/*          edited 6/13/95, 6/21/95, 2/19/96
/*=====
```



```

&args fdir facc fbas trsh fout item fwsh flpp fstr
&severity &error &routine bailout

&if [SHOW PROGRAM] <> GRID &then ; &return This only runs in GRID.
&if [null %fstr%] &then
  &do
    &call usage
    &return
  &end

&messages &off
/* GRID settings:
&s xxcell = [show setcell]
&s xxwindow = [show setwindow]
setwindow %fdir%
setcell %fdir%
/* -----
/* Creating a stream network
/* -----
/* calculate threshold in number of cells
&sv trsh = %trsh% * 1000000 /* conversion into meters (map units)
&sv tresh = %trsh% / ( $$cellsize * $$cellsize )
  &type stream delineation ... [date -vfull]
%fstr% = con ( %facc% >= %tresh% AND %fbas% > 0 , 1 )
/* -----
/* Finding the most downstream cell in each reach (watershed outlet)
/* -----
  &type streamorder [date -vfull]
lso = streamorder ( %fstr% , %fdir% )
  &type streamlink [date -vfull]
lsl = streamlink ( %fstr% , %fdir% )
  &type zonalmax [date -vfull]
lmx = zonalmax ( lsl , %facc% )
  &type lpd condition [date -vfull]
lpd = con ( lmx == %facc% , lsl, 0 )
kill lmx all
/* lpd contains cells that have been selected by finding cell with
/* the largest value of the flowaccumulation within each partial
/* drainage area.
/* lpd cells have value equal to the stream_ID and the watershed-ID.

/* -----
/* Finding the most upstream cell in each first order reach (wshed
outlet)
/* -----
  &type zonalmin [date -vfull]
lmn = zonalmin ( lsl , %facc% )
  &type lpu condition [date -vfull]
lpu = con ( lso == 1 AND lmn == %facc% , lsl + 100000 , 0 )
  &type lfs focalsum for first order wshds [date -vfull]
kill lso all ; kill lsl all ; kill lmn all
/* lpu contains cells that have been selected by finding cell with
/* the smallest value of the flowaccumulation within first order
/* partial watersheds.

```

```

/* lpu cells have value equal to the stream_ID (and the watershed-ID)
/* increased by 100 000. Numbers >100 000 and <200 000 are used to
/* identify first order watersheds
/* (flow gauging station_ids are large numbers, for example 5448500)

/* -----
/* Converting point coverage into grid.
/* Cell value = value from %item%
/* -----
lgs = pointgrid ( %fout% , %item%, #, #, #, ZERO )

/* -----
/* Creating grid with all potential outlets
/* -----
      &type all outlets [date -vfull]
lpx = con ( lpu == 0 , con ( lgs == 0 , lpd, lgs ) , lpu )
kill lpu all      /* upper end of first order reach
kill lgs all      /* usgs gauging stations
kill lpd all      /* down (lowest) end of the reach
      &type setnull cells == 0 [date -vfull]
%flpp% = setnull ( lpx < 1, lpx )
&if not [ exists %flpp%.vat -info ] &then ; buildvat %flpp%
kill lpx all
/* all watershed outlets must be located on the stream network.
/* Gauging stations, which are located on the first cell of the first
/* order stream will not affect the value of the cell e.g. node
number
/* will not be changed.
/* -----
/* Delineating watersheds
/* -----
      &type watershed [date -vfull]
%fwhsh% = watershed ( %fdir% , %flpp% )
&if not [exists %fwhsh%.vat -info] &then ; buildvat %fwhsh%
      &type watershed end [date -vfull]
/* restore GRID settings:
setcell %xxcell%
setwindow %xxwindow%
&messages &on
&return
/*-----
&routine USAGE
/*-----
&type Usage: &r WSHGS <fdir_grid> <facc_grid> <basin_grid>
<tresh_num>
&type          <fout_point> <item> <watershed> <pour_points>
<streams>
&return &inform
/*-----
&routine EXIT
/*-----
/* delete all temporary files:
&if [exists lso -grid ] &then ; kill lso all.
&if [exists lsl -grid ] &then ; kill lsl all
&if [exists lmx -grid ] &then ; kill lmx all

```

```

&if [exists lpd -grid ] &then ; kill lpd all
&if [exists lmn -grid ] &then ; kill lmn all
&if [exists lpu -grid ] &then ; kill lpu all
&if [exists lgs -grid ] &then ; kill lgs all
&if [exists lpx -grid ] &then ; kill lpx all
/* on error erase grids that exist or that have been created
&if [exists %fwsh% -grid ] &then ; kill %fwsh% all
&if [exists %flpp% -grid ] &then ; kill %flpp% all
&if [exists %fstr% -grid ] &then ; kill %fstr% all
/* restore GRID settings:
setcell %xxcell%
setwindow %xxwindow%
&messages &on
&return
/*-----
&routine BAILOUT
/*-----
&severity &error &ignore
&call exit
&return &warning An error has occurred in WSHGS.AML
/* -----
/*      EXAMPLE OF APPLICATION
/* -----
/* grid
/*      /* input
/* &s fdir = $HOME/iowa/data/crfdr      /* crfdr =
flowdirection(elev_grid)
/* &s facc = $HOME/iowa/data/crfac      /* crfac = flowaccumulation
(crfdr)
/* &s fbas = $HOME/iowa/data/crbas      /* basin or watershed
/* &s trsh = 25                          /* units in km2
/* &s fout = $home/iowa/test90/stations/gsflow /* point coverage
/* &s item = station_id                  /* item in %fout% coverage
/*      /* output
/* &s fwsh = crwsh                        /* watersheds
/* &s flpp = crlpp                        /* pour points
/* &s fstr = crstr                        /* stream network
/*
/* &r wshgs %fdir% %facc% %fbas% %trsh% %fout% %item% %fwsh% %flpp%
%fstr%
/*
/* q                                     /* quit from GRID
/*
/* &return                               /* return from example AML
/* -----

```

## C2 NEXTWSH.AML

```
/* -----
/* Program: NEXTWSH.AML
/* Purpose: Creates an info file that contains two items: ID and
/*           the downstream watershed ID. Also creates a grid of
/*           watershed outlets that contains the ID of downstream
/*           watershed.
/*           Runs only in GRID
/* -----
/* Usage: &r NEXTWSH <fdir> <fwsh> <flpp> <infn> <idin> <nxin> <fnxt>
/* Arguments: (input)
/*           fdir = (grid) flow direction
/*           fwsh = (grid) watersheds (partial drainage areas)
/*                   value in VAT = unit_id number
/*           flpp = (grid) pour points ( watershed outlets)
/*                   value in VAT = unit_id number
/*           (output)
/*           infn = (info file) name of the info file to be created
/*           idin = (item) name of the item in which the unit_id
/*                   number will be stored
/*           nxin = (item) name of the item in which the downstream
/*                   unit_id will be stored
/*           fnxt = (grid) similar to %flpp% (watershed outlets):
/*                   the values in VAT are not equal to the unit_id
/*                   number, they equal to the downstream unit
/*                   ID number.
/* Temporary: xxxtmp, xxxcomnx, (GRIDS)
/*           xxxtmp1 (INFO)
/* Globals:
/* Locals: see Temporary
/* -----
/* Notes: The ID of the next watershed for the most downstream
/*         watershed equals 0.
/*         If the flowdirection in watershed pour point can't be
/*         determined, the next watershed ID = -1
/*         This AML checks neither for the existence and correctness
/*         of the input files nor for the existence of files that have
/*         the same names as the temporary files and the files to be
/*         created.
/*         If an error occurs then all grids and all info files that
/*         have the same name as output grids/info files and temporary
/*         grids/info are erased !!!
/* -----
/* History: coded by Pawel Mizgalewicz 12/20/93
/*           and converted into a stand alone procedure
/*           6/10 - 6/21/95, 2/19/96
/* =====
&args fdir fwsh flpp infn idin nxin fnxt
&severity &error &routine bailout

&if [SHOW PROGRAM] <> GRID &then ; &return This only runs in GRID.
```

```

&if [null %fnxt%] &then
  &do
    &call usage
    &return
  &end

&messages &off
/* GRID settings:
&s xxcell = [show setcell]
&s xxwindow = [show setwindow]
setcell %fwsh%
setwindow %fwsh%
/* -----
&type searching for next (downstream) watershed ... [date -vfull]
/* -----
&sv wsh = xxxtmp
%wsh% = con ( isnull ( %fwsh%), 0, %fwsh% )
%fnxt% = con (%flpp% > 0, con (%fdir% == 1, %wsh%(1,0), ~
  con (%fdir% == 2, %wsh%(1,1), ~
  con (%fdir% == 4, %wsh%(0,1), ~
  con (%fdir% == 8, %wsh%(-1,1), ~
  con (%fdir% == 16, %wsh%(-1,0), ~
  con (%fdir% == 32, %wsh%(-1,-1), ~
  con (%fdir% == 64, %wsh%(0,-1), ~
  con (%fdir% == 128, %wsh%(1,-1), -1)))))))))
kill xxxtmp all
&if not [exists %fnxt%.vat -info] &then ; buildvat %fnxt%
xxxcomnx = combine ( %flpp% , %fnxt% )
/* the %fnxt% grid is similar to the %flpp%, watershed outlets grid.
/* The cell value is equal to the next watershed ID number
/* -----
&type Creating the info file ... [date -vfull]
/* -----
arc additem xxxcomnx.vat xxxtmp1 %idin% 4 8 B
arc additem xxxtmp1 xxxtmp1 %nxin% 4 8 B
cursor xxnext declare xxxtmp1 info rw
cursor xxnext open
&do &while %:xxnext.aml$next%
  &s :xxnext.%idin% = [value :xxnext.%flpp%]
  &s :xxnext.%nxin% = [value :xxnext.%fnxt%]
  cursor xxnext next
&end
cursor xxnext close
cursor xxnext remove
arc pullitems xxxtmp1 %infn%
%idin%
%nxin%
end
kill xxxcomnx all
&s x = [delete xxxtmp1 -info]
/* restore GRID settings:
setcell %xxcell%
setwindow %xxwindow%
&messages &on
&return

```

```

/*-----
&routine USAGE
/*-----
&type Usage: &r NEXTWSH <fdir_grid> <wsh_grid> <pourpt_grid>
<newINF_name>
&type          <id_item_name> <next_id_item_name> <nextwsh_grid>
&return &inform
/*-----
&routine EXIT
/*-----
/* delete all temporary files:
&if [exists xxxtmp1 -info ] &then ; &s x = [delete xxxtmp1 -info]
&if [exists xxxtmp -GRID] &then ; kill xxxtmp all
&if [exists xxxcomnx -GRID] &then ; kill xxxcomnx all
&if [exists %fnxt% -GRID] &then ; kill %fnxt% all
/* restore GRID settings:
setcell %xxcell%
setwindow %xxwindow%
&messages &on
&return
/*-----
&routine BAILOUT
/*-----
&severity &error &ignore
&call exit
&return &warning An error has occurred in NEXTWSH.AML
/* -----
/*      EXAMPLE OF APPLICATION
/* -----
/* grid
/*      /* input
/* &s fdir = $HOME/iowa/data/crldr/* crldr = flowdirection(elev_grid)
/* &s fwsh = crwsh                /* watersheds
/* &s flpp = crlpp                /* pour points
/*      /* output
/* &s fnxt = crnxt                /* outflow,value=downstream wshd ID
/* &s infn = %fnxt%.dat          /* new info file
/* &s idin = unit_id             /* item (unit ID)
/* &s nxin = next_id             /* item ( downstream unit ID)
/*
/* &run nextwsh %fdir% %fwsh% %flpp% %infn% %idin% %nxin% %fnxt%
/*
/* q                             /* quit from GRID
/*
/* &return                       /* return from example AML
/* -----

```

### C3 RAININFO.AML

```
/* -----  
/* Program: RAININFO.AML  
/* Purpose: Based on the time series stored in the attribute table  
/*           of the point coverage, and the grid that specify the  
/*           spatial division into zones, this AML calculates  
/*           zonal averages and stores them in Arc/Info INFO table  
/* Method: For a given month of the selected period this AML  
/*           creates a grid of spatially distributed feature  
/*           such as rainfall, temperature, or evaporation.  
/*           The Inverse Distance Weighting IDW procedure is  
/*           applied. To use different method only one line  
/*           need to be changed. In the next step this AML  
/*           calculates average values for all zones, specified by  
/*           and stores them in INFO file. The procedure is  
/*           repeated for all months from the specified time  
/*           Runs only in GRID  
/* -----  
/* Usage: &r RAININFO <fy> <fm> <ty> <tm> <p_data> <zone>  
/*           <outinf> {id_itm} {prc} {kcell} {prx}  
/* Arguments: (input)  
/*           fy, fm = from year, from month  
/*           ty, tm = to year, to month  
/*           p_data = (coverage) point coverage that contains time  
/*                   series  
/*           zone = (grid) defines zones and their IDs  
/*           outinf = (info) output info file to be created  
/*           id_itm = (item) name of the item in which the zone ID  
/*                   will be stored, default {id_itm} = zone_id  
/*           prc = (number) the values must be converted  
/*                   into integer numbers to create VAT.  
/*                   Before conversion they are multiplied by  
/*                   <prc> to preserve decimal part of value.  
/*                   default {prc} = 1000 (three decimal places)  
/*           kcell = (number) cell size multiplier.  
/*                   The grid that is created from point coverage  
/*                   will assume the size of the cell from  
/*                   <zone> grid multiplied by {kcell}  
/*                   default {kcell} = 1  
/*           prx = (prefix) Items that are created  
/*                   in <outinf> info file will named as follows:  
/*                   <prx> + item name from input coverage <p_data>  
/*                   default {prx} = "p".  
/* Temporary: xxx, xxx2, xxxcom (grid), xxxtmp1 (info)  
/* Locals: ab, bb, cc, nitem, prfnitem, yr,mt (variables)  
/* -----  
/* Notes: The following convention is used for naming the items  
/*         in PAT (table that stores original time series)  
/*         myyyymm, where: m indicates monthly values,  
/*         yyyy = year, mm = month, for example item m197603  
/*         carries data for March 1976.
```

```

/* -----
/*   History: Coded by Pawel Mizgalewicz  05/05/95
/*           Modified 06/18/1995 - 6/21/95,
/*           02,20,96 - defaults added for id_itm prc kcell.
/* -----

&args fy fm ty tm p_data zone outinf id_itm prc kcell prx

&severity &error &routine bailout

&if [SHOW PROGRAM] <> GRID &then ; &return This only runs in GRID.
&if [null %outinf%] &then
  &do
    &call usage
    &return
  &end
&if [null %prx%] &then ; &s prx = p
&if [null %kcell%] &then ; &s kcell = 1
&if [null %prc%] &then ; &s prc = 1000
&if [null %id_itm%] &then ; &s id_itm = zone_id

&messages &off
/* GRID settings:
&s xxcell = [show setcell]
&s xxwindow = [show setwindow]
&describe %zone%
&sv size = [calc %grd$dx% * %kcell%]
&type cell size = %size%
/* -----
/*           Make a copy of %zone%.VAT file and add item %zone%
/* -----
arc additem %zone%.vat xxxtmp1 %zone% 4 5 B
cursor xxcurl declare xxxtmp1 info rw
cursor xxcurl open
&do &while %:xxcurl.aml$next%
  &s :xxcurl.%zone% = %:xxcurl.VALUE%
  cursor xxcurl next
&end
cursor xxcurl close
cursor xxcurl remove
/* -----
/*           Do it for all selected years and months
/* -----
&DO yr = %ty% &to %fy% &by -1
  &DO mt = 12 &to 1 &by -1
    &s ab = [calc %yr% = %fy%] AND [calc %mt% lt %fm%]
    &s bb = [calc %yr% = %ty%] AND [calc %mt% gt %tm%]
    &s cc = %ab% OR %bb%
    &IF NOT %cc% &THEN
      &do
        &s templ = mxxxx0y
        &s nitem = [subst %templ% xxxx %yr% ]
        &if %mt% lt 10 &then
          &sv nitem = [subst %nitem% y %mt% ]
        &else

```



```

        &sv nitem = [subst %nitem% 0y %mt% ]
&type processing item %nitem% ... [date -vfull]
setcell %size%
setwindow MAXOF
xxx = idw ( %p_data% , %nitem% )
setcell %zone%
setwindow %zone%
xxx2 = int ( %prc% * zonalmean ( %zone% , xxx ) )
kill xxx all
xxxcom = combine ( %zone% , xxx2 )
kill xxx2 all
&s prfnitem = %prx%%nitem%
arc additem xxxcom.vat xxxcom.vat %prfnitem% 4 12 F 4 xxx2
cursor xxcurl declare xxxcom.vat info rw
cursor xxcurl open
&do &while %:xxcurl.aml$next%
    &s :xxcurl.%prfnitem% = [calc %:xxcurl.xxx2% / %prc%]
    cursor xxcurl next
&end
cursor xxcurl close
cursor xxcurl remove
arc joinitem xxxtmp1 xxxcom.vat xxxtmp1 %zone% %zone%
arc dropitem xxxtmp1 xxxtmp1 xxx2
kill xxxcom all
&end
&end
/* -----
&type Adding ID item ... [date -vfull]
/* -----
arc additem xxxtmp1 xxxtmp1 %id_itm% 4 8 B # %zone%
cursor xxcur2 declare xxxtmp1 info rw
cursor xxcur2 open
&do &while %:xxcur2.aml$next%
    &s :xxcur2.%id_itm% = [value :xxcur2.%zone%]
    cursor xxcur2 next
&end
cursor xxcur2 close
cursor xxcur2 remove
/* -----
&type Cleaning INFO table (dropping redundant items) .. [date -vfull]
/* -----
arc dropitem xxxtmp1 xxxtmp1 value
arc dropitem xxxtmp1 xxxtmp1 count
arc dropitem xxxtmp1 %outinf% %zone%
&s x = [delete xxxtmp1 -info]
/* restore GRID settings:
setcell %xxcell%
setwindow %xxwindow%
&messages &on
&return
/*-----
&routine USAGE
/*-----
&type Usage: &r RAININFO <fy> <fm> <ty> <tm> <pnt_data> <zone>

```

```

&type <out_inf> <ID_item> <prc> <kcell> {prefix}
&return &inform
/*-----
/* &routine CHECK
/*-----
/* IF temporary file exists, inform and exit
/* If file to be build exists , inform and exit
/* If input file is not correct or does not exist, inform and exit
/* return
/*-----
&routine EXIT
/*-----
/* delete all temporary files:
&if [exists xxxtmp1 -info ] &then; &s x = [delete xxxtmp1 -info]
&if [exists xxx -GRID] &then ; kill xxx all
&if [exists xxx2 -GRID] &then ; kill xxx2 all
&if [exists xxxcom -GRID] &then ; kill xxxcom all
/* restore GRID settings:
setcell %xxcell%
setwindow %xxwindow%
&messages &on
&return
/*-----
&routine BAILOUT
/*-----
&severity &error &ignore
&call exit
&return &warning An error has occurred in RAININFO.AML
/* -----
/* EXAMPLE OF APPLICATION
/* -----
/* grid
/* &s fy = 1990
/* &s fm = 1
/* &s ty = 1990
/* &s tm = 12
/* &s pdat = ../stations/gsrain
/* &s zone = crwsh
/* &s id unit_id
/*
/* &r raininfo %fy% %fm% %ty% %tm% %pdat% %zone% prcinf2 %id% 10000
10 r
/*
/* q
/* &return
/* -----

```

## C4 RAINMAP.AML

```
/* -----
/* Program: RAINMAP.AML
/* Purpose: Creates GRIDS of average monthly precipitation
/*           (10e-3 cm/d -- value from PAT is multiplied by 1000,
/*           assumed cell size = 1000)
/*           Stations that have complete record of precipitation for
/*           a given month (from specified time period) are used.
/*           rainm2.aml uses output to put values into an INFO file
/*           Runs only in GRID
/* -----
/* Usage: &r RAINMAP <fy> <fm> <ty> <tm> <xxxdat> {YES|NO}
/* Arguments: (input)
/*           fy, fm = from year, from month
/*           ty, tm = to year, to month
/*           xxxdat = info file that contains precipitation depth
/*           yes|no = yes - include records with values = 0
/*                   no - exclude records and stations that have 0
/*                   (default = no)
/* Temporary: xxxxx1
/* Locals: templ, nitem, nit, yr,mt, selec (variables)
/* -----
/* Notes: Data are stored in PAT (point). Each record is related
/*         to gauging station. The following naming convention is
/*         used for items: myyyymm, where:
/*         m = indicates monthly values,
/*         yyyy = year, mm = month, for example, item m197603
/*         contains records for March 1976
/*         Info file is created by rainm2.aml
/* -----
/* History: Coded by Pawel Mizgalewicz 08/08/95
/*           edited 02/20/96
/* -----
&args fy fm ty tm xxxdat zero

&severity &error &routine bailout

&if [SHOW PROGRAM] <> GRID &then ; &return This only runs in GRID.
&if [null %xxxdat%] &then
  &do
    &call usage
    &return
  &end
&if [null %zero%] &then ; &s zero = YES
&s xx = [translate %zero%]
&select %xx%
  &when NO
    &s klm = gt
  &when YES
    &s klm = ge
  &otherwise
```

```

        &do
            &call usage
            &return
        &end
&end /* select

&messages &off
setwindow maxoff
setcell 1000

&DO yr = %fy% &to %ty%
    &DO mt = 1 &to 12
        &s ab = [calc %yr% = %fy%] AND [calc %mt% lt %fm%]
        &s bb = [calc %yr% = %ty%] AND [calc %mt% gt %tm%]
        &s cc = %ab% OR %bb%
        &IF NOT %cc% &THEN
            &do
                /* &s nit = %nit% + 1
                &s templ = mxxxx0y
                &s nitem = [subst %templ% xxxx %yr% ]
                &if %mt% lt 10 &then
                    &sv nitem = [subst %nitem% y %mt% ]
                &else
                    &sv nitem = [subst %nitem% 0y %mt% ]
                &type %nitem%
                /* &s namit%nit% = %nitem%
                &sv selec = res %nitem% %klm% 0
                &type selecting complete records %selec%

arc reselect %xxxxdat% xxxxx1 point
%selec%
~
n
n

        &end

        arc build xxxxx1 point
        p%nitem% = int ( 1000 * idw ( xxxxx1, %nitem% ) )
        kill xxxxx1 all

    &end

&end
&end
&messages &on

&return
/*-----
&routin USAGE
/*-----
&type Usage: &r RAINMAP <from_yr> <from_mt> <to_yr> <to_mt>
<data_point>
&type                {YES|NO}
&return &inform
/*-----
&routin EXIT

```

```

/*-----
/* delete all temporary files:
&if [exists xxxxx1 -COVER] &then ; kill xxxxx1 all
&messages &on
&return
/*-----
&routine BAILOUT
/*-----
&severity &error &ignore
&call exit
&return &warning An error has occurred in RAINMAP.AML
/* -----
/*      EXAMPLE OF APPLICATION
/* -----
/*      /* input
/* &s data = $HOME/iowa/data/gsprec      /* point coverage and data
/* grid
/* &run rainmap 1940 1 1993 6 %data% YES
/* q
/* &return                                /* return to Arc/Info prompt
/* -----

```

## C5 RAINM2.AML

```
/* -----  
/* Program: RAINM2.AML  
/* Purpose: Creates an INFO table from the grids of average  
/*           monthly precipitation (needs grid of zones)  
/*           (10e-3 cm/d -- value from PAT was multiplied by 1000,  
/*           in original rainmap.aml assumed cell size = 1000 )  
/*           (rainmap.aml creates grids of precipitation)  
/*           Runs only in GRID, output: integer values !!!  
/* -----  
/* Usage: &r RAINM2 <fy> <fm> <ty> <tm> <ppath> <zones> <idname>  
/*           <frominfo> <toinfo>  
/* Arguments: (input)  
/*           fy, fm = from year, from month  
/*           ty, tm = to year, to month  
/*           ppath = path to directory that contains grids of  
/*                   precipitation  
/*           zones = zones for calculatuing average prec. depth  
/*           idname = same as zones, contains identification  
/*                   number for zones (unit_id, modeling units)  
/*           frominfo = info file that contains station id numbers  
/*           toinfo = output info file  
/*                   copy of frominfo + precipitation items  
/* Temporary: xxxcom  
/* Locals: templ, nitem, nit, yr,mt,(variables)  
/* -----  
/* Notes: Post processor for RAINMAP.AML  
/*         RAINMAP.aml changes units, here means are converted to  
/*         integer values !!!  
/*         The following naming convention is  
/*         used for items: myyyymm, where:  
/*         m = indicates monthly values,  
/*         yyyy = year, mm = month, for example, item m197603  
/*         contains records for March 1976  
/*         Info file is created by rainm2.aml  
/* -----  
/* History: Coded by Pawel Mizgalewicz 95  
/*           edited 02/20/96  
/* -----  
&args fy fm ty tm ppath zones idname frominfo toinfo  
  
/* &s ppath = ../rain25/  
/* &s zones = crwsh  
/* &s idname = unit_id  
/* &s frominfo = crwshc.dat  
/* &s toinfo = crwshc.datp  
  
/* &sv fy = 1990  
/* &sv fm = 1  
/* &sv ty = 1990  
/* &sv tm = 12
```

```

copyinfo %frominfo% %toinfo%

&type ...aml, start [date -vfull]
grid
&messages &off
setwindow %zones%
setcell %zones%

/* %idname% = %zones%
/* repeat in reverse order
&DO yr = %ty% &to %fy% &by -1
  &DO mt = 12 &to 1 &by -1
    &IF %mt% ge %fm% and %mt% le %tm% &THEN
      &do
        &s templ = mxxxx0y
        &s nitem = [subst %templ% xxxx %yr% ]
        &if %mt% lt 10 &then
          &sv nitem = [subst %nitem% y %mt% ]
        &else
          &sv nitem = [subst %nitem% 0y %mt% ]

&type p%nitem% processing ... [date -vfull]
p%nitem% = zonalmean ( %zones% , %ppath%p%nitem%
xxxcom = combine ( %idname% , p%nitem% )
kill p%nitem% all
&sys arc joinitem %toinfo% xxxcom.vat %toinfo% %idname% %idname%
kill xxxcom all

      &end
    &end
  &end

&type rain series - point cover created, exiting ... [date -vfull]

&return

```

## C6 SELDATA.AML

```
/* -----
/* Program: SELDATA.AML
/* Purpose: Selects the items that contain data for specified
/*          time period. Then, selects gauging stations that
/*          have complete record (no missing data).
/*          Creates point coverage of selected stations with data
/*          for specified months
/*          Runs only in ARC
/* -----
/* Usage: &r SELDATA <xxxdat> <xxxhed> <jointit> <xxxout> {YES|NO}
/* Arguments: (input)
/*            fy, fm = from year, from month
/*            ty, tm = to year, to month
/*            xxxdat = info file that contains flow database
/*            xxxhed = point coverage that represents location of
/*                   observation stations. xxxdat and xxxhed must
/*                   contain common item to relate these tables
/*                   (for example item station_id)
/*            jointit = name of the common item for both tables
/*                   (output)
/*            xxxout = point coverage which contains selected flow
/*                   records or precipitation measurements,
/*                   as well as selected gauging stations.
/*            yes|no = yes - include records with values = 0
/*                   no - exclude records and stations that have 0
/*                   (default = no)
/* Temporary: %xxxout%tmp (cover), xxxxx1 (info)
/* Locals: temp1, nitem, nit, yr,mt, selec (variables)
/* -----
/* Notes: Data file is an INFO file. Each record is related to
/*        gauging station. The following naming convention is
/*        used for items:
/*        myyyymm, where m indicates monthly values,
/*        yyyy = year, mm = month, for example item m197603
/*        contains records for March 1976
/*        Due to the Arc/Info restrictions only about 65-month
/*        period can be selected in single run of this AML
/* -----
/* History: Coded by Pawel Mizgalewicz 05/05/95
/*          Modified 06/15/1995, edited 2/20/96
/* -----
&args fy fm ty tm xxxdat xxxhed jointit xxxout zero

&severity &error &routine bailout

&if [SHOW PROGRAM] <> ARC &then ; &return This only runs in ARC.
&if [null %xxxout%] &then
&do
&call usage
&return
```



```

&end
&if [null %zero%] &then ; &s zero = no
&s xx = [translate %zero%]
&select %xx%
  &when NO
    &s klm = gt
  &when YES
    &s klm = ge
  &otherwise
    &do
      &call usage
    &return
  &end
&end /* select

/* -----
/*      Make a copy of point coverage
/* -----
copy %xxxhed% %xxxout%tmp    /* create temporary file
&sv pultmp = xxxxx1         /* temporary info file
/* -----
&type Selecting years and months ...      [date -vfull]
/* -----
PULLITEMS %xxxdat% %pultmp%
%joinit%
&DO yr = %fy% &to %ty%
  &DO mt = 1 &to 12
    &s ab = [calc %yr% = %fy%] AND [calc %mt% lt %fm%]•
    &s bb = [calc %yr% = %ty%] AND [calc %mt% gt %tm%]
    &s cc = %ab% OR %bb%
    &IF NOT %cc% &THEN
      &do
        &s templ = mxxxx0y
        &s nitem = [subst %templ% xxxx %yr% ]
        &if %mt% lt 10 &then
          &sv nitem = [subst %nitem% y %mt% ]
        &else
          &sv nitem = [subst %nitem% 0y %mt% ]
        &type selecting item %nitem%
%nitem%
      &end
    &end
  &end
&end

end

/* join selected items with header file (point coverage)
joinitem %xxxout%tmp.pat xxxxx1 %xxxout%tmp.pat %joinit% %joinit%
&sv st = [delete xxxxx1 -info]
/* -----
&type Selecting gauging stations that have full record ... [date -
vfull]
/* -----
reselect %xxxout%tmp %xxxout% point
&DO yr = %fy% &to %ty%

```

```

&DO mt = 1 &to 12
  &s ab = [calc %yr% = %fy%] AND [calc %mt% lt %fm%]
  &s bb = [calc %yr% = %ty%] AND [calc %mt% gt %tm%]
  &s cc = %ab% OR %bb%
  &IF NOT %cc% &THEN
    &do
      /* &s nit = %nit% + 1
      &s templ = mxxxx0y
      &s nitem = [subst %templ% xxxx %yr% ]
      &if %mt% lt 10 &then
        &sv nitem = [subst %nitem% y %mt% ]
      &else
        &sv nitem = [subst %nitem% 0y %mt% ]
      &type %nitem%
      /* &s namit%nit% = %nitem%
      &sv selec = res %nitem% %klm% 0
      &type selecting complete records %selec%
%selec%
~
n
y

      &end
    &end
  &type end
%selec%
~
n
n

build %xxxout% point
kill %xxxout%tmp all
&return
/*-----
&routine USAGE
/*-----
&type Usage: &r SELDATA <from_yr> <from_mt> <to_yr> <to_mt>
<data_info>
&type          <station_cover> <join_item> <out_cover> {YES|NO}
&return &inform
/*-----
&routine EXIT
/*-----
/* delete all temporary files:
&if [exists %pultmp% -info ] &then ; &s x = [delete %pultmp% -info]
&if [exists %xxxout%tmp -COVER] &then ; kill %xxxout%tmp all
&messages &on
&return
/*-----
&routine BAILOUT
/*-----
&severity &error &ignore
&call exit
&return &warning An error has occurred in SELDATA.AML

```

```
/* -----  
/*      EXAMPLE OF APPLICATION  
/* -----  
/*      /* input  
/* &s data = $HOME/iowa/data/gsflow.pat /* Info file of monthly flows  
/* &s head = $HOME/iowa/data/crfhd    /* coverage of gauging stations  
/* &s cname = station_id             /* common item  
/*      /* output  
/* &s fout = mflow91                 /* output coverage  
/*  
/* &run seldata 1991 1 1991 12 %data% %head% %cname% %fout% NO  
/*  
/* &return                           /* return to Arc/Info prompt  
/* -----
```

## C7 FD4Y.AML

```
/* -----
/* fd4y.aml
/*      1) Creates ASCII files that contain data required
/*          for C program fd4y. ( fd4y calculates flow rate in
/*          all modeling units based on the flow recorded in
/*          USGS gauging stations).
/*      2) Stores the results obtained from fd4y program in
/*          PAT--polygon attribute table

&s prc = unprec.pat
&s flw = gsfl128.pat
&s out = unflow.pat
&s fy = 1960
&s fm = 1
&s ty = 1992
&s tm = 9

/* input1 xxxgsin, ASCII file
/* input2 xxxunin, ASCII file
/* output xxxunout, ASCII file

tables

&DO yr = %ty% &to %fy% &by -1
  &DO mt = 12 &to 1 &by -1
    &s ab = [calc %yr% = %fy%] AND [calc %mt% lt %fm%]
    &s bb = [calc %yr% = %ty%] AND [calc %mt% gt %tm%]
    &s cc = %ab% OR %bb%
    &IF NOT %cc% &THEN
      &do
        &s templ = mxxxx0y
        &s nitem = [subst %templ% xxxx %yr% ]
        &if %mt% lt 10 &then
          &sv nitem = [subst %nitem% y %mt% ]
        &else
          &sv nitem = [subst %nitem% 0y %mt% ]
        &type processing item %nitem% ... [date -vfull]
/* -----
select %flw%
unload xxxgsin station_id station_nx %nitem% DELIMITED INIT
select %prc%
unload xxxunin unit_id unit_nx gswsh area_km2 order p%nitem%
DELIMITED INIT
&sys fdy4 xxxgsin xxxunin xxxunout

&s i = 0
&do &until [exists xxxunout -file]
  &s i = %i% + 1
```

```

&end

define %nitem%.dat
unit_id,4,8,B
q%nitem%,4,12,F,4
~
add from xxxunout
select %nitem%.dat

&s x = [delete xxxunout -file]
&if %x% eq 0 &then
  &type file xxxunout has been deleted
&else
  &type ERROR: could not delete file xxxunout

&s x = [delete xxxgsin -file]
&if %x% eq 0 &then
  &type file xxxgsin has been deleted
&else
  &type ERROR: could not delete file xxxgsin

&s x = [delete xxxunin -file]
&if %x% eq 0 &then
  &type file xxxunin has been deleted
&else
  &type ERROR: could not delete file xxxunin

&sys arc joinitem %out% %nitem%.dat %out% unit_id order
kill %nitem%.dat

  &end
&end
&end

q
q
&return

```

## C8 SLOPE3.AML

```
/* -----
/* Program: FSLOPE.AML
/* Purpose: Calculates slope along the flow path. Calculates the
/*           difference between the elevation of the current cell
/*           and the elevation of the next cell on the flow path and
/*           divides the result by the distance between these cells
/*           Runs only in GRID
/* -----
/* Usage: &r FSLOPE <fdir> <dem> <fslp>
/* Arguments: (input)
/*            fdir = (grid) flow direction
/*            dem  = (grid) elevation
/*            (output)
/*            fslp = (grid) slope grid
/* Temporary: none
/* -----
/* Notes: If flow direction is
/*         different than 1, 2, 4, 8, 16, 32, 64, and 128,
/*         the slope 0 is assumed.
/*         Vertical and horizontal units should be the same
/* -----
/* History: Coded by Pawel Mizgalewicz 10/02/94 (slope3.aml)
/*          edited 16/06/95, 21/06/95, 02/24/96
/*          ( version with DOCELL -> slope3.aml)
/*          ( version with CONDition -> slope2.aml )
/* =====

&args fdir dem fslp
&severity &error &routine bailout

&if [SHOW PROGRAM] <> GRID &then ; &return This only runs in GRID.
&if [null %fslp%] &then
  &do
    &call usage
    &return
  &end

&messages &off
/* GRID settings:
&s xxcell = [show setcell]
&s xxwindow = [show setwindow]
setcell minof
setwindow maxof
/* -----
&type calculating slope ... [date -vfull]
/* -----
&describe %dem%
&sv diag = 1.414213562 * %grd$dx%
DOCELL
  if (%fdir% == 1)
```

```

    %fslp% = ( %dem% - %dem%(1,0) ) / %grd$dx%
else
  if (%fdir% == 2)
    %fslp% = ( %dem% - %dem%(1,1) ) / %diag%
  else
    if (%fdir% == 4)
      %fslp% = ( %dem% - %dem%(0,1) ) / %grd$dx%
    else
      if (%fdir% == 8)
        %fslp% = ( %dem% - %dem%(-1,1) ) / %diag%
      else
        if (%fdir% == 16)
          %fslp% = ( %dem% - %dem%(-1,0) ) / %grd$dx%
        else
          if (%fdir% == 32)
            %fslp% = ( %dem% - %dem%(-1,-1) ) / %diag%
          else
            if (%fdir% == 64)
              %fslp% = ( %dem% - %dem%(0,-1) ) / %grd$dx%
            else
              if (%fdir% == 128)
                %fslp% = ( %dem% - %dem%(1,-1) ) / %diag%
              else
                %fslp% = 0
END
&type end of fslope [date -vfull]
/* restore GRID settings:
setcell %xxcell%
setwindow %xxwindow%
&messages &on
&return
/*-----
&routine USAGE
/*-----
&type Usage: &r FSLOPE <fdir_grid> <elevation_grid> <slope_grid>
&return &inform
/*-----
&routine EXIT
/*-----
/* on error delete slope grid:
&if [exists %fslp% -GRID] &then ; kill %fslp% all
/* restore GRID settings:
setcell %xxcell%
setwindow %xxwindow%
&return
/*-----
&routine BAILOUT
/*-----
&severity &error &ignore
&call exit
&return &warning An error has occurred in FSLOPE.AML
/* -----
/*      EXAMPLE OF APPLICATION slope of stream reaches
/* -----
/* grid                                     /* start GRID

```

```

/*      /* input
/* &s fdir = $HOME/iowa/data/crfd
/*                                     /* crfd = flowdirection(elev_grid)
/* &s dem = $HOME/iowa/data/crdem      /* elevation grid
/* &s fstr = crstr
/*      /* output
/* &s fslp = crslp                      /* slope grid
/* &s str = strslp                      /* stream slope
/*
/* /* create slope grid
/* &run fslope %fdir% %dem% %fslp%
/*
/* /* assign unique number to each reach
/* xxxlsl = streamlink(%fstr%)
/*
/* /* calculate average slope
/* %strs% = zonalmean ( xxxlsl , %fslp% )
/*
/* kill xxxlsl all                      /* delete temporary grid
/* kill %fslp% all                      /* delete slope grid
/* q                                    /* quit from GRID
/*
/* &return                              /* return from example AML
/* -----

```



## C9 DEMALB.AML

```
/* DEMALB.AML - converts a grid into Albers Equal Area coordinates
&args input output csize
```

```
%output% = project (%input% , #, NEAREST, %csize% )
```

```
OUTPUT
```

```
Projection      ALBERS
```

```
Zunits          NO
```

```
Units           METERS
```

```
Datum           NAD83
```

```
Spheroid        GRS1980
```

```
Xshift          0.0000000000
```

```
Yshift          0.0000000000
```

```
Parameters
```

```
 29 30 0.000 /* 1st standard parallel
```

```
 45 30 0.000 /* 2nd standard parallel
```

```
-96 0 0.000 /* central meridian
```

```
 23 0 0.000 /* latitude of projection's origin
```

```
0.00000 /* false easting (meters)
```

```
0.00000 /* false northing (meters)
```

```
END
```

```
&return
```

## BIBLIOGRAPHY

- Al-Abed, N. and H. R. Whiteley, 1995. Modeling Water Quality and Quantity in the Lower Portion of the Grand River Watershed, Ontario. Proc. of the International Symposium on Water Quality Modeling, April 2-5, 1995, Orlando FL., ASAE (American Society of Agricultural Engineers), St. Joseph, MI. pp. 213-222.
- Arnold, J. G., J. R. Williams, A. D. Nicks, and N. B. Sammons, 1990. *SWRRB A Basin Scale Simulation Model for Soil Water Resources Management*, Texas A&M University Press, College Station.
- Arnold, J. G., B. A. Engel, and R. Srinivasan, 1993. Continuous time grid cell watershed model. In Heatwole, C. D. (ed.), *Application of Advanced Information Technologies : Effective Management of Natural Resources*. Information and Electrical Technologies Division of American Society of Agricultural Engineers, ASAE, St. Joseph, Michigan.
- Bailey, G. W. and R. R. Swank, Jr., 1983. Modeling Agricultural Nonpoint Source Pollution: A Research perspective. In: Shaller, F. W. and G. W. Bailey, (eds.), *Agricultural Management and Water Quality*, Iowa State University, Ames, pp. 27-47.
- Battaglin, W. A., D. A. Goolsby, and R. H. Coupe, 1993, Annual Use and Transport of Agricultural Chemicals in the Mississippi River, 1991-92. In: Goolsby, D. A., L .L. Boyer, and G. E. Mallard, (eds.), *Selected Papers on Agricultural Chemicals in Water Resources of the Midcontinental United States: US Geological Survey Open-File Report 93-418*, p. 26-40.

- Battaglin, W. A. and D. A. Goolsby, 1995a. Spatial Data in Geographic Information System Format on Agricultural Chemical Use, Land Use, and Cropping Practices in the United States - prepared in cooperation with the US Environmental Protection Agency. US Geological Survey, Water-Resources Investigations Report 94-4176, 87 pp.
- Battaglin, W. A. and D. A. Goolsby, 1995b. Spatial Data in Geographic Information System Format on Agricultural Chemical Use, Land Use, and Cropping Practices in the United States;  
<http://h2o.usgs.gov/public/pubs/bat/bat000.html>
- Battaglin, W. A., 1995. Location of sites sampled during 1989 and 1990 in the midwestern United States and their drainage areas, Arc/Info maps Sta\_recon and Bas\_recon (communication via Internet).
- Beasley, D. B., L F. Huggins, and E. J. Monke, 1982. ANSWERS: A model for watershed planning. Transactions of American Society of Agricultural Engineers, ASAE 23 (4), pp. 938-944.
- Beasley, D. B. and L. F. Huggins, 1991. An excerpt from the ANSWERS User's Manual, second edition, 1991.  
<http://soils.ecn.purdue.edu/~aggrass/models/answers.html>
- Bian, L., H. Sun, C. Blodgett, S. Egbert, W. Li, L. Ran., and A. Koussis, 1996. An Integrated Interface System to Couple the Swat Model and Arc/Info. Third International NCGIA Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, January, 21-25.
- Bond, L. E., 1986. Topographic partition of watersheds with digital elevation data. *Water Resources Research*, 22(1): 15-24.

- Bondelid, T. R., S. A. Hanson, and L. T. Taylor, 1991, Technical Description of the Reach File. Horizon Systems Corporation, Herndon, Virginia, 18 pp.
- Bouzaher, A. and A. Manale, 1993. Agricultural Atrazine Use and Water Quality: A CEEPES Analysis of Policy Options. Staff Report 93-SR 66, Center for Agricultural and Rural Development, Iowa State University, Ames, IA.
- Burt, T. P., A. L. Heathwaite, and S. T. Trudgill (eds.), 1993. *Nitrate: Processes, Patterns and Management*. John Wiley & Sons, Chichester, pp. 342 - 367.
- Camara, A. S., and C. W. Randall, 1984. The QUAL II Model. *Journal of Environmental Engineering*, ASCE, 110, pp. 993-996)
- Chow, V. T., D. R. Maidment, and L. W. Mays, 1988. *Applied Hydrology*. McGraw-Hill, New York.
- Comfort, S. D. and F. W. Roeth, 1996. Questions and Answers about Atrazine. World Wide Web, <http://www.unl.edu/ianr/pubs/extnpuhs/pesticid/g1158.htm>
- Coote, D. R., E. M. MacDonald, and R. DeHaan, 1978. Relationships Between Agricultural Land and Water Quality. In: Loehr, R. C., D. A. Haith, M. W. Walter, and C. S. Martin, (eds.), *Best Management Practices for Agriculture and Silviculture. Proc., 1978 Cornell Agricultural Waste Management Conference*. Ann Arbor Science, Ann Arbor, Mich., pp. 79-93.
- Cressie, N. and J. J. Majure, 1994. Use of GIS in Observational Studies. Preprint 94-3, Dept. of Statistics, Iowa State University, Ames, IA.
- Cronshey, R. G., F. D. Theurer, and R. L. Glenn, 1993. GIS - Water Quality Computer Model Interface: A Prototype. NCGIA Second International Conference/Workshop on Integrating Geographic Information Systems and

Environmental Modeling, September 26-30, 1993, Brackenridge, Colorado USA. 1993.

- Davis J. S. and H. M. Keller, 1983. Dissolved loads in streams and rivers - discharge and seasonally related variations. In: Webb, B. W. (ed.) *Dissolved Loads of Rivers and Surface Water Quantity/Quality Relationships*. Proceedings of the Hamburg Symposium, August 1983. IAHS Publ. no. 141, pp. 79 - 89.
- Deetz H., Ch. and O. S. Adams, 1969. *Elements of map projection with applications to map and chart construction*, Greenwood, New York, 226 pp.
- DeVries, J. J. and T. V. Hromadka, 1993. Streamflow. In: Maidment, D. R., (ed.), *Handbook of Hydrology*, McGraw-Hill, New York, pp. 21.1-21.39.
- Donigian, A. S. Jr. and W. C. Huber, 1991. *Modeling of Nonpoint Source Water Quality in Urban and Non-urban Areas*. EPA/600/3-91/039, U.S. Environmental Protection Agency, Athens, Georgia, 72 pp.
- Drungil, C. E. C., W. F. Geter, and R. G. Shepherd, 1995. Watershed Scale Water Quality Modeling in the NRCS HUWQ Project. Proc. of the International Symposium on Water Quality Modeling, April 2-5, 1995, Orlando FL., ASAE (American Society of Agricultural Engineers), St. Joseph, MI. pp. 177-185.
- Engel, B. R., 1993. Methodologies for Development of Hydrologic Response Units Based on Terrain, Land Cover, and Soils Data. NCGIA Second International Conference/Workshop on Integrating Geographic Information Systems and Environmental Modeling, September 26-30, 1993, Brackenridge, Colorado USA.
- Engel, B. R., R. Srinivasan, and C. Rewerts, 1993a. A Spatial Decision Support System for Modeling and Managing Agricultural Non-Point-Source Pollution. In: Goodchild, M. F., B. O. Parks, and L. T. Steyaert, (eds.) *Geographic*

*Information Systems and Environmental Modeling*, Oxford University Press, New York, pp. 231-237.

Engel, B. R., R. Srinivasan, J. Arnold, C. Rewerts, and S. J. Brown, 1993b. Nonpoint Source (NPS) Pollution Modelling Using Models Integrated with Geographic Information Systems (GIS). *Water Science and Technology*. Vol. 28, No. 3-5, pp. 685-690.

Engel, B., 1996. AGNPS-GRASS Interface.

<http://soils.ecu.purdue.edu/~aggrass/models/agnps/Index.html>

ESRI, 1992. *Cell-based Modeling with Grid*. Environmental Systems Research Institute, Inc., Redlands, California.

ESRI, 1995. ArcView 2.1--computer software. Environmental Systems Research Institute, Inc., Redlands, California.

Evans, B. M. and D. A. Miller, 1988. Modeling Nonpoint Pollution at the Watershed Level with the Aid of a Geographic Information System. In: *Nonpoint Pollution: 1988 - Policy, Economy, Management, and Appropriate Technology*. American Water Resources Association, pp. 283-290.

EWG, 1996. Environmental Working Group. Information published on World Wide Web, <http://www.ewg.org/>

Geter, W. F., P. Smith, C. Drungil, R. Shepherd, and B. Kuenstler, 1995. Hydrologic Unit Water Quality Model GIS Interface to Four ARS Water Quality Models for Use by Soil Conservation Service. Proc. of the International Symposium on Water Quality Modeling, April 2-5, 1995, Orlando FL., ASAE (American Society of Agricultural Engineers), St. Joseph, MI. pp. 341-347.

- Gianessi, L. P., and C. M. Puffer, 1988. *Herbicide Use in the United States: Washington, D. C., Resources for the Future, Inc., Quality of the Environment Division, 489 p.*
- Goodchild, M. F., B. O. Parks, and L. T. Steyaert, (eds.) 1993. *Geographic Information Systems and Environmental Modeling*, Oxford University Press, New York, pp. 231-237.
- Goolsby, D. A., W. A. Battaglin, J. D. Fallon, D. S. Aga, and E. M. Thurman, 1993a. Persistence of Herbicides in Selected Reservoirs in the Midwestern United States: Some Preliminary Results. In: Goolsby, D. A., L. L. Boyer, and G. E. Mallard, (eds.), *Selected Papers on Agricultural Chemicals in Water Resources of the Midcontinental United States: U.S. Geological Survey Open-File Report 93-418 pp. 51-63.*
- Goolsby, D. A., E. M. Thurman, M. L. Pomes, M. Meyer, and W. A. Battaglin, 1993b. Occurrence, Deposition, and Long Range Transport of Herbicides in precipitation in the Midwestern and Northeastern United States. In: Goolsby, D. A., L. L. Boyer, and G. E. Mallard, (eds.), *Selected Papers on Agricultural Chemicals in Water Resources of the Midcontinental United States: U.S. Geological Survey Open-File Report 93-418 pp. 75-89.*
- Goolsby, D. A, W. A. and Battaglin, and E. M. Thurman. 1993c. Occurrence and Transport of Agricultural Chemicals in the Mississippi River Basin, July Through August 1993, U.S. Geological Survey Circular 1120-C pp. 22.
- Goolsby, D. A, and W. A. Battaglin, 1993. Occurrence, Distribution and Transport of Agricultural Chemicals in Surface Waters of the Midwestern United States. In: Goolsby, D. A., L. L. Boyer, and G. E. Mallard, (eds.), *Selected Papers on*

*Agricultural Chemicals in Water Resources of the Midcontinental United States*: U.S. Geological Survey Open-File Report 93-418 pp. 1-25.

Goolsby, D. A., L. L. Boyer, and G. E. Mallard, (eds.), 1993. *Selected Papers on Agricultural Chemicals in Water Resources of the Midcontinental United States*. U.S. Geological Survey Open-File Report 93-418, 90 pp.

Goring, C. A. I., D. A. Laskowski, J. W. Hamaker, and R. W. Meikle, 1975. Principles of Pesticide Degradation in Soil. In: Haque, R., and V. H. Freed, (eds.), *Environmental Dynamics of Pesticides*. Plenum Press, New York, pp. 135-172.

Gruessner, B., N. C. Shambaugh, and M. C. Watzin. 1995, Comparison of an Enzyme Immunoassay and Gas Chromatography/Mass Spectrometry for the Detection of Atrazine in Surface Waters. *Environmental Science and Technology*, Vol. 29, No. 1, pp. 251-254.

Hall, J. K., M. Pawlus, and E. R. Higgins, 1972. Losses of Atrazine in Runoff Water and Soil Sediment. *Journal of Environmental Quality*, vol. 1, no. 2: pp. 172-176.

Hamaker, J. W., 1975. The Interpretation of Soil Leaching Experiments. In: Haque, R., and V. H. Freed, (eds.). *Environmental Dynamics of Pesticides*. Plenum Press, New York, pp. 115 - 133.

Haque, R., and V. H. Freed (eds.), 1975. *Environmental Dynamics of Pesticides*. Plenum Press, New York,

Harlin, J. M., and K. J. Lanfear (eds.) 1993. *Geographic Information Systems and Water Resources*, Proc. of the Symposium, American Water Resources Association AWRA, Bethesda, Maryland.



- Hatfield, J. L., J. L. Anderson, E. E. Alberts, T. Prato, D. G. Watts, A. Ward, G. Delin, R. Swank, 1993. Management Systems Evaluation Areas -- an Overview. Proceedings of the Conference on Agricultural Research to Protect Water Quality, Feb. 21-24, 1993, Minneapolis, Minnesota, Soil and Water Conservation Society, Ankeny, Iowa, pp. 1-15.
- He, C., F. J. Riggs, and Y. Kang, 1993. Integration of Geographic Information Systems and a Computer Model to Evaluate Impacts of Agricultural Runoff on Water Quality. In: J. M. Harlin and K. J. Lanfear, (eds.), *Geographic Information Systems and Water Resources*, American Water Resources Association AWRA Proc., Bethesda, Maryland, pp. 79-87.
- Heathwaite, A. L., 1993. Nitrogen Cycling in Surface Waters and Lakes. In: Burt, T. P., A. L. Heathwaite, and S. T. Trudgill, (eds.), *Nitrate: Processes, Patterns and Management*. John Wiley & Sons, Chichester, pp. 99 - 144.
- Helsel,, D. R. and R. M. Hirsh, 1995. *Statistical Methods in Water Resources*. Elsevier, The Netherlands, 530 pp.
- Hession, W. C. and V. O. Shanholtz, 1988. A Geographic Information System for Targeting Nonpoint-Source Agricultural Pollution. *Journal of Soil and Water Conservation*, Vol. 43, no. 3, pp. 264-266.
- Hydrosphere, 1993a. Hydrodata, USGS Daily and Peak Values; User Manual and CD-Roms, Hydrosphere Data Products, INC., 1002 Walnut, Suite 200, Boulder, Colorado 80302.
- Hydrosphere, 1993b. Climatedata, TD 3200 Summary of the Day Cooperative Observer Network; User Manual and CD-Roms, Hydrosphere Data Products, INC., 1002 Walnut, Suite 200, Boulder, Colorado 80301.

- Jacobson, B. M., J. Feng, G. D. Jennings, and K. C. Stone, 1995. Watershed Scale Non-Point Source Model Evaluation for the North Carolina Coastal Plain. Proc. of the International Symposium on Water Quality Modeling, April 2-5, 1995, Orlando FL., ASAE (American Society of Agricultural Engineers), St. Joseph, MI. pp. 186-191.
- James, A., and D. J. Elliot, 1993. Models of Water Quality in Rivers. In: James, A. (ed.) *An Introduction to Water Quality Modelling*. John Wiley & Sons, Chichester, pp. 141-181.
- James, A., (ed.) 1993. *An Introduction to Water Quality Modelling*. John Wiley & Sons, Chichester, 311 pp.
- Jankowski, P. and G. Haddock, 1993. Integrated Nonpoint Source Pollution Modeling System. NCGIA Second International Conference/Workshop on Integrating Geographic Information Systems and Environmental Modeling, September 26-30, 1993, Brackenridge, Colorado USA.
- Johanson, R. C., J. C. Imhoff, and H. Dana, Jr., 1980. Users Manual for Hydrological Simulation Program-FORTRAN (HSPF). EPA/9-80-015, Environmental Protection Agency, Athens, Georgia.
- Jones, P. J. and P. T. Burt, 1993. Nitrate in Surface Waters. In: Burt, T. P., A. L. Heathwaite, and S. T. Trudgill, (eds.), *Nitrate: Processes, Patterns and Management*. John Wiley & Sons, Chichester, pp. 269 - 317.
- Kalkhoff, S. J. and R. L. Kuzniar, 1991. Hydrologic Data for the Big Spring Basin, Clayton County, Iowa, Water Year 1989. U. S. Geological Survey, OpenFile Report 91-63, Iowa City, Iowa, 66 pp.

- Keeney, D. R., 1983. Transformation and Transport of Nitrogen. In Shaller, F. W. and G. W. Bailey, (eds.), *Agricultural Management and Water Quality*, Iowa State University, Ames, pp. 48-64.
- Kenaga, E. E., 1975. Partitioning and Uptake of Pesticides. In: Haque, R., and V. H. Freed, (eds.), *Environmental Dynamics of Pesticides*. Plenum Press, New York, p. 229.
- Klaghofer, E., W. Birnbaum, and W. Summer, 1993. Linking Sediment and Nutrient Export Models with a Geographic Information System. In: *HydroGIS: Application of Geographic Information System in Hydrology and Water Resources*, Proc., Viena April 1993. IAHS no. 211, pp. 501-506,
- Knisel, W. G., G. R. Foster, and R. A. Leonard, 1983. CREAMS: A System for Evaluating Management Practices. In: Shaller, F. W. and G. W. Bailey, (eds.), *Agricultural Management and Water Quality*, Iowa State University, Ames, pp. 178-199.
- Knisel, W. G. (ed.) 1984 (1980). *CREAMS: A Field-Scale Model for Chemicals, Runoff, and Erosion from Agricultural Management Systems*. U.S. Department of Agriculture, Conservation Research Report No. 26, 640 pp.
- Kolpin, D. W., M. R. Burkart, and E. M. Thurman, 1991. Herbicides and Nitrate in Near-Surface Aquifers in the Midcontinental United States, 1991. U. S. Geological Survey Water-Supply Paper 2413, 34 pp.
- Kreuger, J. K., and N. Brink, 1988. Losses of Pesticides from Agriculture. In: *Pesticides: Food and Environmental Implications*. Proc. Int. Symp. on Changing Perspectives in Agrochemicals, jointly organized by the International Atomic Energy Agency and the Food and Agriculture Organization of the United Nations, Neuheberg, 24-27 Nov. 1987. IAEA, Vienna, pp. 101-112.

- Kruger, E. L., H. U. Baluch, L. Somasundaram, R. S. Kanvar, and J. R. Coats, 1993. Fate of atrazine, deethylatrazine, deisopropylatrazine, and hydroxyatrazine as affected by soil depth, soil characteristics, and saturated/unsaturated soil moisture conditions. Proceedings of the Conference on Agricultural Research to Protect Water Quality, Feb. 21-24, 1993, Minneapolis, Minnesota, Soil and Water Conservation Society, Ankeny, Iowa, pp. 567-570.
- Krysanova, V., D. Muller-Wohlfeil, A. Becker, 1996. Mezoscale Integrated Modelling of Hydrology and Water Quality with GIS Interface. Third International NCGIA Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, January, 21-25.
- Lanfear, K. J., 1994. Conversion of the USEPA Reach File 1 to an Arc/INFO coverage. U.S. Geological Survey, Reston, Virginia.  
<http://nsdi.usgs.gov/nsdi/wais/water/rf1.HTML>
- Loehr, R. C., D. A. Haith, M. W. Walter, and C. S. Martin, (eds.), 1978. *Best Management Practices for Agriculture and Silviculture*. Proc., 1978 Cornell Agricultural Waste Management Conference, Ann Arbor Science, Ann Arbor, Michigan.
- Maguire, D. J., M. F. Goodchild, and D. W. Rhind, (eds.), 1991. *Geographical Information Systems; Principles and Applications* (2 vol.), London, Longman, and New York: Wiley.
- Maidment, D. R. and E. Parzen, 1984. Cascade Model of Monthly Municipal Water Use. *Water Resources Research*, Vol. 20, No. 1, pp. 15-23.
- Maidment, D. R., 1992. A Grid-Network Procedure for Hydrologic Modeling. Report Prepared for the Hydrologic Engineering Center, U. S. Army Corps of Engineers, Davis, California, October, 1992.

- Maidment, D. R., 1993a. Developing a Watershed Structure. Report Prepared for the Hydrologic Engineering Center, U. S. Army Corps of Engineers, Davis, California, April, 1993.
- Maidment, D. R., (ed.), 1993b. *Handbook of Hydrology*, McGraw-Hill, New York.
- Maidment, D. R., 1996. GIS and Hydrologic Modeling - an Assessment of Progress. Third International NCGIA Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, January, 21-25.
- Mamillapalli, S., R. Srinivasan, J. G. Arnold, B. A. Engel, 1996. Effect of Spatial Variability on Basin Scale Modeling. Third International NCGIA Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, January, 21-25.
- MathSoft, 1993. *A Crash Course in S-PLUS, version 3.2*. StatSci a Division of MathSoft, Inc., Seattle, Washington.
- McCutcheon, S. C., J. L. Martin, And T. O. Barnwell, Jr., 1993. Water Quality. In: Maidment, D. R., (ed.), *Handbook of Hydrology*, McGraw-Hill, New York, pp. 11.3-11.73.
- Mendel, O. and J. Repa, 1994. Evaluation of surface water pollution by nitrate in Northeast Slovakia. *Hydrological, Chemical and Biological Processes of Transformation and Transport of Contaminants in Aquatic Environments*, Proceedings of the Rostov-on-Don Symposium, May 1993). IAHS Publ. no. 219, pp. 375-385.
- Mitchell, J. K., B. A. Engel, R. Srinivasan, and S. S. Y. Wang, 1993a. Validation of AGNPS for Small Watersheds Using an Integrated AGNPS/GIS System. In J. M. Harlin and K. J. Lanfear (eds.) *Geographic Information Systems and Water Resources*, AWRA Proc., Bethesda, Maryland, pp. 89-100.

- Mitchell, J. K., B. A. Engel, R. Srinivasan, and S. S. Y. Wang, 1993b. Validation of AGNPS for Small Watersheds Using an Integrated AGNPS/GIS System. *Water Resources Bulletin*, Vol. 29, No. 5, AWRA, pp. 833-842.
- Mizgalewicz, P. J., 1991. An Analysis of Monthly Water Demand in Phoenix, Arizona. Master Thesis, University of Texas, Austin, 170 pp.
- Moody, J. A., and D. A. Goolsby, 1993. Spatial Variability of Triazine Herbicides in the Lower Mississippi River, *Environmental Science and Technology*, Vol. 27, no 10, pp. 2120-2126.
- Moore W. J., 1991. *Inorganic Contaminants of Surface Water - Alberta, Canada: Research and Monitoring Priorities*. Springer-Verlag, New York, 334 pp.
- Morse, G., A. Eatherall, and A. Jenkins, 1994. Managing Agricultural Pollution Using a Linked Geographical Information System and Non-Point Source Pollution Model. *Journal of the Institution of Water and Environmental Management*, Vol. 8, No. 3, pp. 277-286.
- Mosley, M. P. and A. I. McKerchar, 1993. Streamflow. In: Maidment, D. R., (ed.), *Handbook of Hydrology*, McGraw-Hill, New York, pp. 8.1-8.39.
- Mueller, D. K., B. C. Ruddy, and W. A. Battaglin, 1993. Relation of Nitrate Concentrations in Surface Water to Land Use in The Upper-Midwestern United States, 1989-90. In: Goolsby, D. A., L.L. Boyer, and G. E. Mallard, (eds.), *Selected Papers on Agricultural Chemicals in Water Resources of the Midcontinental United States*: U.S. Geological Survey Open-File Report 93-418 pp. 41-50.
- Mullins, J. A., R. F. Carsel, J. E. Scarbrough, and A. M. Ivery, 1993. PRZM-2, A Model for Predicting Pesticide Fate in the Crop Root and Unsaturated Soil Zones: Users Manual. US Environmental Protection Agency, Athens, GA.

- Oberle, S. L. and M. R. Burkart, 1994. Water Resource Implications of Midwest Agroecosystems. *Journal of Environmental Quality*, 23: pp. 4-8.
- O'Connor, D. J., R. V. Thomann, J. L. Schnoor, and J. L. Manciani, 1983. Modeling Toxic Substances in Natural Water Systems, Manhattan College.
- O'Connor, D. J., 1988a. Models of Sorptive Toxic Substances in Freshwater Systems. I: Basic Equations. *Journal of Environmental Engineering*, ASCE, Vol. 114, No. 3, pp. 507-532.
- O'Connor, D. J., 1988b. Models of Sorptive Toxic Substances in Freshwater Systems. II: Lakes and Reservoirs. *Journal of Environmental Engineering*, ASCE, Vol. 114, No. 3, pp. 533-551.
- O'Connor, D. J., 1988c. Models of Sorptive Toxic Substances in Freshwater Systems. III: Streams and Rivers. *Journal of Environmental Engineering*, ASCE, Vol. 114, No. 3, pp. 552-574.
- OECD, 1986. *Water Pollution by Fertilizers and Pesticides*. Organization for Economic Co-Operation and Development, Paris, 146 pp.
- Plimmer, J. R., 1988. Movement of Pesticides from the Site of Application. In: Pesticides: Food and Environmental Implications. Proc. Int. Symp. on Changing Perspectives in Agrochemicals, jointly organized by the International Atomic Energy Agency and the Food and Agriculture Organization of the United Nations, Neuheberg, 24-27 Nov. 1987. IAEA, Vienna, pp. 61-77.
- Ramanarayanan, T. S., R. Srinivasan, and J. G. Arnold, 1996. Modeling Wister Lake Watershed Using a GIS-Linked Basin-Scale Hydrologic/Water Quality Model. Third International NCGIA Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, January, 21-25.

- Rea, A. and J. R. Cederstrand, 1995. GCIP Reference Data Set (GREDS), GEWEX Continental-scale International Project, U.S. Geological Survey, Open-File Report 94-388, CD-ROM.
- Rewerts, C. C. and B. A. Engel, 1991. ANSWERS on GRASS: Integrating a Watershed Simulation with a GIS. ASAE Paper No. 91-2621, American Society of Agricultural Engineers, St. Joseph, MI, 8 pp.
- De Roo, A. P. J., 1993. Validation of the ANSWERS catchment model for runoff and soil erosion simulation in catchments in The Netherlands and the United Kingdom. In: *HydroGIS: Application of Geographic Information System in Hydrology and Water Resources*, Proc., Viena, April 1993. IAHS no. 211, pp. 465-474.
- Schoellhamer, D. H., 1988. Lagrangian Transport Modeling with QUAL II Kinetics. *Journal of Environmental Engineering*, ASCE, 114(2), pp. 368-381.
- Scribner, E. A., E. M. Thurman, D. A. Goolsby, M. Meyer, M. S. Mills, and M. L. Pomes, 1993. Reconnaissance Data for Selected Herbicides, Two Atrazine Metabolites, and Nitrate in Surface Water of the Midwestern United States, 1989-90. U.S. Geological Survey Open-File Report 93-457, Lawrence, Kansas, 77 pp.
- Scribner, E. A., D. A. Goolsby, E. M. Thurman, M. Meyer, and M. L. Pomes, 1994. Concentrations of Selected Herbicides, Two Triazine Metabolites, and Nutrients in Storm Runoff From Nine Stream Basins in the Midwestern United States, 1990-1992. U.S. Geological Survey Open-File Report 94-396, Lawrence, Kansas, 144 pp.
- Shaller, F. W. and G. W. Bailey, (eds.), 1983. *Agricultural Management and Water Quality*, Iowa State University, Ames, 472 pp.



- Sharpley, A. N. and J. R. Williams, (eds.), 1990. EPIC-Erosion/Productivity Impact Calculator. Technical Bulletin 1768, US Department of Agriculture, Agricultural Research Service, Temple, Texas.
- Smith, J. A., 1993. Precipitation. In: Maidment, D. R. (ed.), *Handbook of Hydrology*, McGraw-Hill, New York, pp. 3.1-3.47.
- Smith, R. A., R. B. Alexander, G. D. Tasker, C. V. Price, K. W. Robinson, D. A. White, 1993. Statistical Modeling of Water Quality in Regional Watersheds. Conference Proceedings, Watershed'93, pp. 751-754.
- Soenksen, P. J., J. L. Hatfield, and D. J. Schmitz, 1993. Chemical Loads of Nitrate, Atrazine, and Metolachlor in Walnut Creek Watershed near Ames, Iowa, 1991-1992. USGS, Iowa City, IA and Agricultural Research Service, Ames, IA.
- Soenksen, P. J., 1994. Personal Communication and transfer of the Walnut Creek Watershed data.
- Srinivasan, R., and J. G. Arnold, 1994. Integration of Basin-Scale Water Quality Model with GIS. *Water Resources Bulletin*, 30(3), pp. 453-462.
- Srinivasan, R., J. G. Arnold, and R. S. Muttiach, 1995. Continental Scale Hydrologic Modeling Using GIS. Proc. of the International Symposium on Water Quality Modeling, April 2-5, 1995, Orlando FL., ASAE (American Society of Agricultural Engineers), St. Joseph, MI. pp. 231-241.
- Squillace, P. J. and R. A. Engberg, 1988. Surface-Water Quality of the Cedar River basin, Iowa-Minnesota, With Emphasis on the Occurrence and Transport of Herbicides, May 1984 Through November 1985. U.S. Geological Survey, Water Resources Investigations Report 88-4060, Iowa City, IA, 81 pp.

- Tim, U. S. and R. Jolly, 1994. Evaluating Agricultural Nonpoint-Source Pollution Using Integrated Geographic Information Systems and Hydrologic/Water Quality Model. *Journal of Environmental Quality*, 23, pp. 25-35.
- Tisdale, S. L., W. L. Nelson, J. D. Beaton, and J. L. Havlin, 1993 (1956). *Soil Fertility and Fertilizers*. (5th ed.) Macmillan, New York, 634 pp.
- Thomann, R. V. and J. A. Mueller, 1987. *Principles of Surface Water Quality Modeling and Control*. Harper, New York, NY, 644 pp.
- Thurman, E. M., M. T. Meyer, M. L. Pomes, C. A. Perry, and A. P. Schwab, 1990. Enzyme-linked immunosorbent assay compared with gas chromatography/mass spectrometry for the determination of triazine herbicides in water. *Analytical Chemistry*, v. 62, pp. 2043-2048.
- Thurman, E. M., D. A. Goolsby, M. T. Meyer, and D. W. Kolpin, 1991. Herbicides In Surface Waters of the Midwestern United States: The Effect of Spring Flush. *Environmental Science and Technology*, October 1991, pp. 1794-1796.
- Thurman, E. M., D. A. Goolsby, M. T. Meyer, M. S. Mills, M. L. Pomes, and D. W. Kolpin, 1992. A Reconnaissance Study of Herbicides and Their Metabolites in Surface Water of the Midwestern United States Using Immunoassay and Gas Chromatography/Mass Spectrometry. *Environmental Science and Technology*, Vol 26, No 12, pp. 2440-2447.
- University of Mississippi, 1990. Concepts and methods for assessing solute dynamics in stream ecosystems, Stream Solute Workshop, University of Mississippi, Feb. 1-5, 1989. *J. of the North American Benthological Society*, 9(2), pp. 95-119.
- U. S. Environmental Protection Agency, 1990. County-level fertilizer sales data. Washington D.C., U. S. EPA, Office of Policy Planning and Evaluation, September 1990, computer disk.

- USGS 1990. Digital Elevation Models, Data Users Guide, National Mapping Program, Technical Instructions, Data Users Guide 5, Department of the Interior, U.S. Geological Survey, Reston, Virginia.
- USGS 1991. Metadata for CLIMATE\_DIV, NOAA Climate Divisions. U.S. Geological Survey, [http://nsdi.usgs.gov/nsdi/wais/water/climate\\_div.HTML](http://nsdi.usgs.gov/nsdi/wais/water/climate_div.HTML).
- Venables, W. N. and B. D. Ripley, 1994. *Modern Applied Statistics with S-Plus*. Springer-Verlag, New York, 462 pp.
- Vieux, B. E. and S. Needham, 1993. Nonpoint-Pollution Model Sensitivity to Grid-Cell Size. *Journal of Water Resources Planning and Management*, Vol. 119, No. 2, March/April 1993, pp. 141 - 157.
- Warn, A. E., 1984. Calculating Future Levels of Nitrate in Rivers and Pumped Storage Reservoirs. *Wat. Sci. Tech.*, IAWPRC, Vol. 16, York, pp. 635-642.
- Wauchope, R. D., 1978. The Pesticide Content of Surface Water Draining from Agricultural Fields-A Review. *Journal of Environmental Quality*, vol. 7, no. 4: pp. 459-472.
- Webb, B. W. (ed.), 1983. Dissolved Loads of Rivers and Surface Water Quantity/Quality Relationships. Proceedings of the Hamburg Symposium, August 1983. IAHS Publ. no. 141.
- Weber, J. B., 1972. Interaction of Organic Pesticides with Particulate Matter in Aquatic and Soil System. In: Faust S. D. (chair) *Fate of Organic Pesticides in the Aquatic Environment*, American Chemical Society Symposium, Los Angeles, CA, March 29-31, 1971.). *Advances in Chemistry Series III*, American Chemical Society, Washington D. C., pp. 55 - 88.

- Weber, J. B., 1988. Pesticide Dissipation in Soils as a Model for Xenobiotic Behavior. In: *Pesticides: Food and Environmental Implications*. Proc. Int. Symp. on Changing Perspectives in Agrochemicals, jointly organized by the International Atomic Energy Agency and the Food and Agriculture Organization of the United Nations, Neuheberg, 24-27 Nov. 1987. IAEA, Vienna, pp. 45-60.
- White, D. A. and M. Hofschien, 1993. A Spatial Model for Assessing Nutrient Loads in Rivers and Streams. NCGIA Second International Conference/Workshop on Integrating Geographic Information Systems and Environmental Modeling, September 26-30, 1993, Brackenridge, Colorado USA.
- Williams, J. R., P. T. Dyke, W. W. Fuchs, V. W. Benson, O. W. Rice, and E. D. Taylor, 1990. EPIC--Erosion/Productivity Impact Calculator. U. S. Department of Agriculture, Agricultural Research Service, Technical Bulletin 1768.
- Wilson, E. M., 1991. *Engineering Hydrology*. Macmillan, London. 348 pp.
- Wolfe, M. L., W. D. Batchelor, T. A. Dillaha, S. Mostaghimi, and C. D. Heatwole, 1995. A Farm Scale Water Quality Planning System for Evaluating Best Management Practices. Proc. of the International Symposium on Water Quality Modeling, April 2-5, 1995, Orlando FL., ASAE (American Society of Agricultural Engineers), St. Joseph, MI. pp. 324-332.
- Wu, T. L., D. L. Correll, and H. E. H. Remenapp, 1983. Herbicide runoff from experimental watersheds. *Journal of Environmental Quality*, vol. 12, no. 3, pp. 330-336.
- Yoon, J., G. Padmanabhan, and L. H. Woodbury, 1993. Linking Agricultural Nonpoint Source Pollution Model (AGNPS) to a Geographic Information System (GIS). In J. M. Harlin and K. J. Lanfear (eds.) *Geographic Information*

*Systems and Water Resources*, AWRA Proc., American Water Resources Association, Bethesda, Maryland. pp. 79-87.

Young, R. A., C. A. Onstad, D. D. Bosh, and W. P. Anderson, 1989. AGNPS: A Non-Point Source Pollution Model for Evaluating Agricultural Watersheds. *J. Soil Water Conserv.*, 44(2), pp. 168-173.

Zollweg, J. A., W. J. Gburek, H. B. Pionke, and A. N. Sharpley 1995. GIS-Based Modeling of Phosphorus Output from a Northeastern Upland Agricultural Watershed. Proc. of the International Symposium on Water Quality Modeling, April 2-5, 1995, Orlando FL., ASAE (American Society of Agricultural Engineers), St. Joseph, MI. pp. 251-258.

## VITA

Pawel Jerzy Mizgalewicz was born in Krakow, Poland, on January 4, 1956, the son of Czeslawa Mizgalewicz and Jerzy Mizgalewicz. After completing his work at XIV Grammar School, Krakow, Poland, in 1975, he entered the University of Technology at Krakow. He received the degree of Master of Science from the University of Technology in June, 1980. During the following years he was employed as a research-scientist (1980-1984) and a graduate teacher (1984-1993) at the University of Technology, Krakow. Most of his research and publications are related to the urban water distribution systems, sewage networks, and drainage systems. He received the Polish Ministry Award for research in 1986.

In September, 1988, he entered The Graduate School of The University of Texas at Austin. He received the degree of Master of Public Affairs and the degree of Master of Science from the University of Texas in December 1991. From 1989 to present he is employed as a graduate research assistant at the Center for Research in Water Resources, University of Texas at Austin.

Permanent address: ul. Jaremy 14B/43, 31-318 Krakow, Poland

This dissertation was typed by Pawel J. Mizgalewicz.