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 $2.4*10^6$ km^2 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^2 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^2 (threshold area for stream delineation = 25 km^2) whereas excluding stream starting points from the set of watershed outlets points produced units of average area 46.7 km^2.

The threshold area of 25 km^2 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^2 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^2 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 foundings 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^2);
  simulation region (area 32,000 km^2);
  gauged zone (average area 1,100 km^2); and
  modeling unit (average area 30 km^2).

- **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.