EXECUTIVE SUMMARY

MANAGEMENT OF NONPOINT SOURCE POLLUTION

AN INTEGRATIVE GIS-ANNAGNPS MODEL

SEPTEMBER 2003

CENTRAL PLAINS CENTER FOR BIOASSESSMENT
KANSAS BIOLOGICAL SURVEY
AND
KANSAS GEOLOGICAL SURVEY
UNIVERSITY OF KANSAS
AND
NATURAL RESOURCES CONSERVATION SERVICE
U.S. DEPARTMENT OF AGRICULTURE
AND
U.S. ENVIRONMENTAL PROTECTION AGENCY
REGION VII

Management of Nonpoint Source Pollution
An Integrative GIS-AnnAGNPS Model

1. Introduction

The water quality of reservoirs reflects the cumulative aspects of the physical, chemical, and biological conditions and stressors in the watersheds. Several common water quality problems have been noted in reservoirs of the Central Plains. There have been reports of elevated plant nutrient levels, with concurrent elevations in plant biomass (Carney, 2002; Dodds et al., 2002). Suspended solids and siltation have increased, and increases in these factors reduce light penetration, aesthetics, lake depth and volume, leading to alteration of aquatic habitats (deNoyelles et al., 1999). Water quality assessments have shown elevated levels of pesticides and other toxic chemicals (Scribner et al., 1996). Further, local and state regulatory agencies have fielded complaints regarding objectionable taste and odor conditions (e.g., KDHE, 1999). All these problems contribute to or are symptomatic of water quality degradation.

However, excess nutrients and siltation, both of which are considered the leading nonpoint pollution sources and result from intensive agricultural activities, are the water quality factors that contribute most to eutrophication (Huggins et al., 1990; Carpenter et al., 1998). Eutrophication is itself a serious and widespread problem in the mid-continent region. According to the National Water Quality Report to Congress, 50% of assessed U.S. lakes and a higher percentage of reservoirs in the agriculturally dominated central U.S. and plains region were considered eutrophic (USEPA, 2000). To manage and control eutrophication, watershed models are essential decision tools that can help water-quality planners assess the overall reservoir conditions, not only under current land-use conditions but also into the future.

Annualized Agricultural Nonpoint Source (AnnAGNPS) is a cell-based continuous-simulation model widely used to investigate surface runoff and nonpoint source pollution. The earlier versions of this model (e.g., AGNPS), which are event-related models, have been broadly and successfully used in the central U. S. (e.g., Young et al., 1989; Engel et

al., 1993; Mankin et al, 2003). AnnAGNPS requires over 400 parameters distributed over 34 data categories, including elevation, soil, land use, field management, operation and crop data. A user-friendly integrative system is greatly needed due to the complexity of these data and physical requirement necessary to import and populate input fields in AnnAGNPS.

Because of the inherent capabilities of data manipulation and visual demonstration ArcView® (ESRI, 1994) was selected as the platform for this user interface. The interface that organizes AnnAGNPS input and output were built as ArcView® extensions since this adoption is generally used and available to researchers and natural resources managers. In this design, a geographic information system (GIS) handles the spatial data like elevation, soil and land use while tabular data (e.g., field management and operation data) are incorporated in a Microsoft Access® database. The Access® database is simply used as a repository for the tabular data. As data management and manipulations data and model input and output program are within ArcView® the tabular data must be can be imported into ArcView®.

Typically, the manual preparation and importing of input data for a relatively large watershed may take several weeks. However, with this integrated system, these same tasks take just hours to days to accomplish. This GIS-integrated system is of great benefit in pre- and post-processing of data used in AnnAGNPS modeling as well as in the decision-making process.

The Hillsdale Reservoir Watershed in eastern Kansas was selected to illustrate the use and application of the interface that has thus far been developed. A brief description of the AnnAGNPS model and a discussion of the development of the integrated system are given in the following sections.

2. ANNAGNPS Model
AnnAGNPS (www.sedlab.olemiss.edu/agnps.html) is a watershed-scale computer model used to generate loading estimates for some constituents of agricultural nonpoint-source

pollution. It simulates runoff, sediment, and nutrient yields on a daily, monthly, or yearly basis upon user's request. Outputs can be specified at any desired source watershed locations such as specific cells, reaches, feedlots, and point sources. AnnAGNPS can be used as a screening tool or a watershed management application (e.g., evaluation of best management practices [BMP] or changes in land use) to address many total-maximum-daily-load (TMDL) and nutrient-criteria developments that are related to the identification and management of natural and/or anthropogenic nutrient levels and effects on the receiving water bodies.

Three types of data are required to running AnnAGNPS; topographic data, soil and land-use related data, and climate data. These data can be prepared and organized using the tools and/or models provided in the AnnAGNPS package. TOPAZ (TOpographic ParameteriZation) is a program that automates digital landscape analysis and was adopted by AnnAGNPS to evaluate topographic features, identify surface drainage, define watershed segmentation, and calculate subwatershed parameters based on digital elevation model data (DEM) (Garbrecht and Martz, 1995). This program is named TopAGNPS within AnnAGNPS and includes three major sub-programs. The first one is DEDNM (Digital Elevation Drainage Network Model), which is used to process the DEM data, identifies the drainage network, and subdivides the hydrographic watershed. This model requires three input files, DNMCNT.INP, DEDNM.INP and NTGCOD.INP. The first file is a control file and the last two are raster files. The critical source area (CSA) and the minimum source channel length (MSCL) are the important parameters in DNMCNT.INP associated with landscape segmentation and channel-network generation. They can be estimated either from topographic maps or field surveys to fit the scale and resolution of the study area.

The other sub-programs are RASPRO (RASter PROperties) and RASFOR (RASter FORmatting). RASPRO derives additional topographic data from the output produced by the program DEDNM while RASFOR formats the unformatted output files from both DEDNM and RASPRO so that they can be further processed. Soil and land-use related data can be organized using the AnnAGNPS Input Editor, a graphical user interface for

storing, retrieving, and managing parameters. Most of the information associated with soil and crop data, field operations, and feedlots must be obtained from various resource databases that need to be populated and are provided within the AnnAGNPS package. Climate data can be historical records, synthetic data, or a combination of the two. The synthetic data are generated using the climate-data generator (GEM) program (Johnson et al., 2000) based on climate stations surrounding the watershed.

3. Development of GIS-AnnAGNPS Modeling System

The pre- and post-processing modules were built as an ArcView extension called Spatial AnnAGNPS Extension (SAE), which was written in Avenue® script.
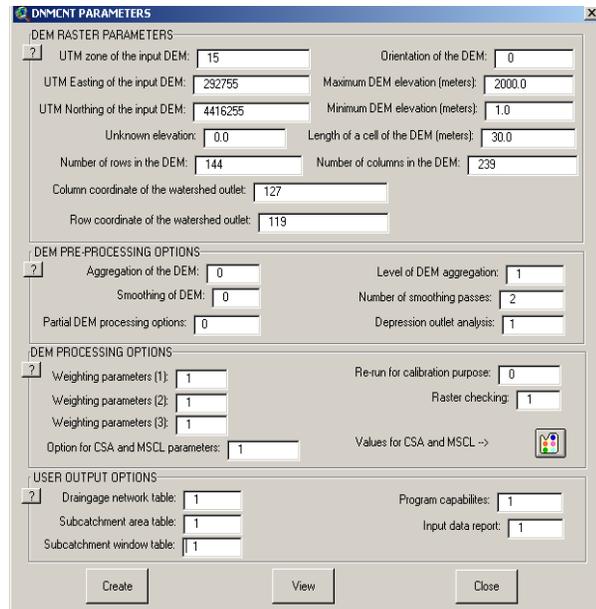
The ESR's Spatial Analyst is required in certain data operations. The interface consists of three modules: 1) inputs for topographic data, 2) inputs for soil and land-use and their associated data, and 3) model execution, output display, and scenarios. The first module was designed to assist in delineation of watershed boundary, outlet selection, TopAGNPS input-data preparation, and execution of TopAGNPS. The second module serves as an operating tool to organize and/or control soil data and land-use (field) information in appropriate formats. The last module initiates the AnnAGNPS program, manipulates output files, and provides graphical results in ArcView® as well as offers "what if" scenario analysis.

3.1. Inputs for Topographic Data
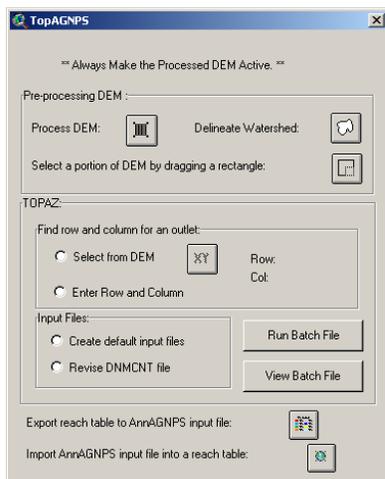
As a spatially distributed model, AnnAGNPS divides a watershed into sub-areas (cells) based on topographic characteristics so that the modeled cells homogeneously represents the characteristics of each cell. TOPAGNPS is used to automatically define watershed segmentation and calculate subwatershed parameters from a DEM. Together with AGFLOW, these two programs provide the topographic data required for AnnAGNPS (Bingner et al., 1997).

Identifying and delineating a watershed boundary that corresponds to a specific outlet location is visually difficult using DEMs. To more consistently and accurately delineate a watershed boundary, a tool was developed which directly derives the boundary from a DEM with a selected outlet (the *Delineate Watershed* button). This tool generates a polygon of the boundary by locating an outlet on a DEM and building the topographically correct drainage from that point to the top of the drainage region. The position of an outlet entered in DNMCNT.inp is defined as row and column numbers of a DEM. Accordingly, the position needs to be calculated based on the size of the DEM and a 30-meter grid space in order to get the row and column numbers of the outlet. The SAE interface calculates these numbers and display them on the dialog when a user locate an outlet on a DEM. Moreover, the associated inputs for DNMCNT.inp and DEDNM.inp can be generated from the DEM when a user clicks on the radio button *Create default input files*.   In addition, the parameters listed in DNMCNT.inp can be edited in the dialog **DNMCNT PARAMETERS** by clicking on the radio button *Revise DNMCNT file* to launch its dialog. This dialog is especially convenient for adjusting the size of subwatersheds by altering the values of CSA and MSCL parameters.

3.2. Inputs for Soil and Land-use and Their Associated Data

Polygon coverages of soil and land-use data are extracted using GIS techniques, while their associated reference data are generated from standard resource databases and field survey. In the model inputs, soil and/or land use data represented in subwatershed cells are characterized and represented by the dominant class (by area) within the total cell. With the integrated system, the user can very easily extract and assign the dominant soil and land use types to each cell within the AnnAGNPS model that is being populated.

The soil and/or land-use coverage is then clipped with the watershed segmentation coverage (Subwat.arc) created by TOPAGNPS described previously. Because the clipped soil and/or land-use coverage has several soil (or land use) types within each cell, the function of *Determine the dominant soil (or field/land use) ID in each cell* in the dialog **Import_Refernce_Data** can automatically extract the dominant type of soil and/or land use and prepare the data as a table in ArcView®.
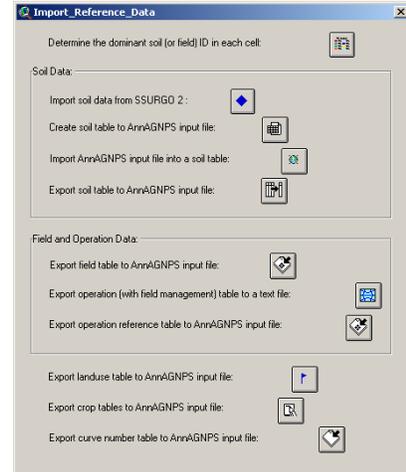
Soil characteristics and properties contain extremely important information that is used within AnnAGNPS to generated output estimates. The AnnAGNPS model requires about 30 soil variables for each soil map unit identified in a watershed. Most of these required soil variables are available in the Soil Survey Geographic (SSURGO) version 2.0 soil database exported from the National Soil Information System (NASIS). In order to use SSURGO in AnnAGNPS, it is necessary to calculate some derived variables from the available ones, to convert soil parameter units, and to process data for the AnnAGNPS required format. Due to the large amount of data involved, manual processing and entry of these required soil data for AnnAGNPS is very difficult, extremely time consuming, and prone to errors. An ArcView SSURGO to AnnAGNPS tool in the dialog completely automates these soil data processing efforts for AnnAGNPS in the required model format thus saving the user considerable time and effort while increasing data input precision.

The values of some soil variables provided in the SSURGO database are theoretical or default (e.g., organic nitrogen or phosphorous). These default values need to be adjusted to reflect actual watershed conditions. The functions of *Import soil data from AnnAGNPS input file* and *Export soil table to AnnAGNPS input file* facilitate data organization in the process of model calibration. The dialog reads in soil values from the AnnAGNPS input file as an ArcView® table so that the soil data and values can be easily edited. After editing, the updated data are exported back to the AnnAGNPS input file through the dialog.

Land use or field data needs to be identified, located, and quantified before can be used as model input variables. These data are used to determine imperviousness, runoff and other variables required in AnnAGNPS. The reference data associated with the land use/field data, which can be obtained from standard resources databases and field survey, are stored in an Access®
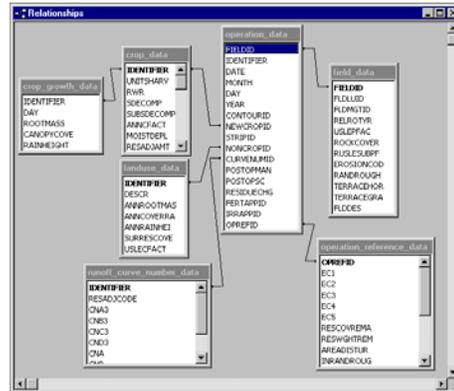
database. The database contains field management, operation, operation-reference, crop, crop growth, land use, and curve number data.   AnnAGNPS uses the Revised Universal Soil Loss Equation (RUSLE) to calculate soil erosion. RUSLE is the successor to the Universal Soil Loss Equation (USLE), which has been widely used throughout the world since the early 1960's to guide conservation planning at the local field office level. RUSLE version 1.05 (http://www.sedlab.olemiss.edu/rusle) is the current version used in AnnAGNPS v2.2.

Like USLE, RUSLE is a set of mathematical equations that calculate a value for the average annual erosion caused by sheet and rill erosion, the type of erosion caused by raindrop impact and overland flow on hillslope profiles. Land use has more effect on soil loss than any other single factor. Of the major factors affecting soil loss, land use is generally the only one that can be easily changed to control soil loss. The effects of climate, soil, and topography are fixed for any specific site, at least for land uses like agriculture.

Cultural practices involve the type of vegetation grown and how the vegetation is grown and managed to reduce soil loss. A second consideration, referred to as supporting practices, uses structural type measures like terraces, contouring, and strips of close growing vegetation to control erosion. Frequently a combination of cultural and structural practices is used. RUSLE uses a sub-factor approach to estimate the main effects of cultural practices on soil loss. This sub-factor approach is very powerful and gives

RUSLE its capability of being land use independent. The major sub-factors that RUSLE considers are canopy, ground cover, roughness, ridge height, buried residue, live and dead roots, and degree of consolidation. These variables are a function of the type of vegetation and the level of production of the vegetation and how soil and vegetative components are manipulated with tillage and other operations like grazing. An important consideration in RUSLE is how these sub-factors vary through the year in relation to temporal variations of rainfall erosivity.

RUSLE is the best available technology for erosion prediction of average annual soil loss for local conservation planning of specific field sites. The results of RUSLE are linked together through the use unique identifications codes. These reference data can be exported as AnnAGNPS input format through the dialog from the data tables that are imported from the Access® database.

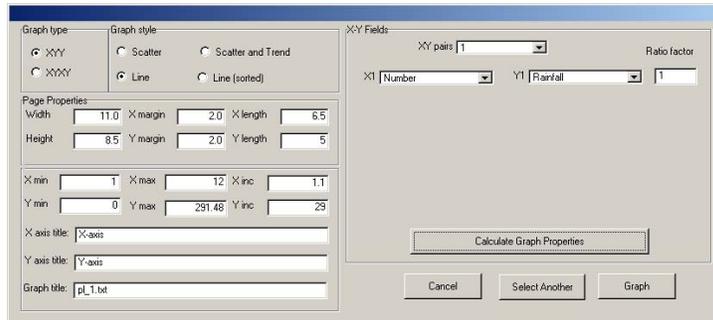### 3.3. Model Execution, Output Display, and Scenarios
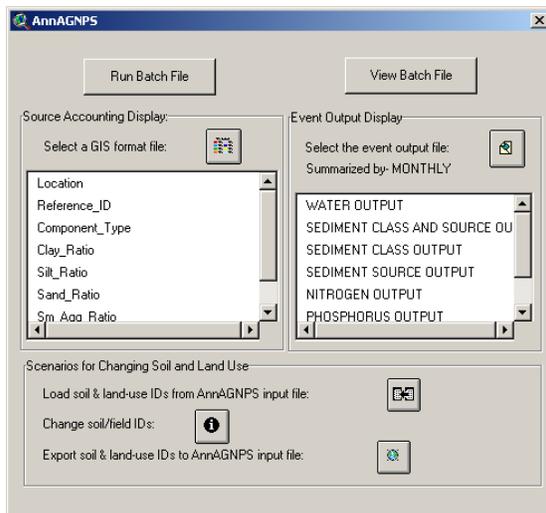
In addition to the AnnAGNPS input-data file, a climate data file needs to be prepared in order to execute the AnnAGNPS simulation. The climate data consists of six main elements; daily precipitation, maximum and minimum temperature, dew point temperature, sky cover, and wind speed. Daily precipitation is the principal variable that defines the hydrologic portion of the model while the temperature data are used to indicate freezing conditions. The remaining climatic variables are used to compute potential evapotranspiration. The GEM program is able to generate represented climatic data from the historical data of several weather stations located in or near the watershed. The generated data may be better to represent the conditions within a relatively large watershed as the distribution of rainfall could dramatically vary in the watershed. As the AnnAGNPS input and climatic data are prepared, the AnnAGNPS

program can be executed either by selecting AnnAGNPS in button *Run Batch File* or through the original program structure.

Two types of output can be displayed through the interface; source accounting and event output displays. For displaying a source accounting GIS map, an item of interest is selected from the list box and a new map coverage will then be created and added to the active View. To display an event output, select an item of interest from the list box of the dialog and then enter options in the graphing dialog. A plot for the selected options of the interest item will be displayed. Scenario analyses can be conducted in the **AnnAGNPS** dialog so that decision makers can ask "what-if" questions that focus on potential changes in current soil and/or land use condition.
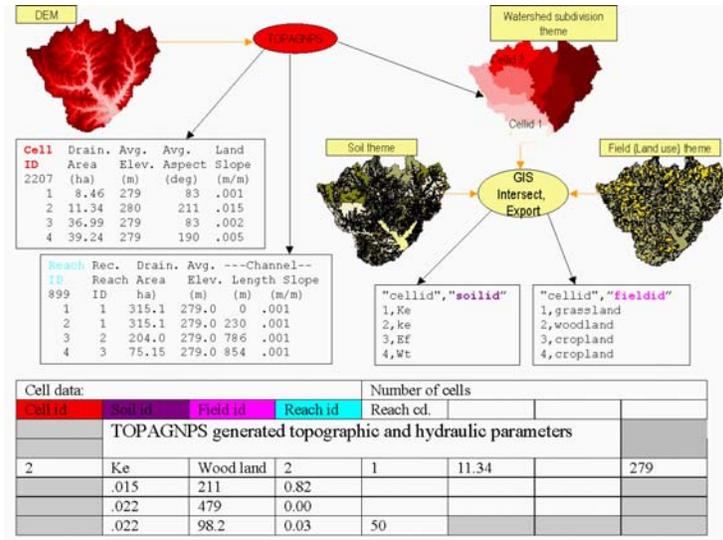
4. An Example Application

The Hillsdale Reservoir Watershed was used to demonstrate application and use of the integrated system to generate an AnnAGNPS simulation. Calibration and determination of the accuracy of the simulated results are not presented here but will be addressed later in the project. Hillsdale Reservoir and its watershed occupy about 144 square miles of land in east central Kansas and span four counties. The reservoir receives most of its water from Big Bull, Little Bull, Rock, and Wade Creeks, and supplies drinking water to more than 30,000 residents in Johnson and Miami Counties. The predominate land use within the watershed is agriculture.

The DEM, soil coverage, and land use coverage were downloaded from the DASC website (mapster.kgs.ukans.edu/dasc/catalog/coredata.html) at the Kansas Geological Survey. The DEM of Hillsdale watershed was processed through TOPAGNPS to generate a watershed-subdivision theme and topographical parameters for AnnAGNPS. The GIS-integrated system clipped this theme with soil and land use coverages and then

decided the dominant soil or land use. Two tables for soil and land-use IDs were linked with subwatershed-cell IDs.

The table in the adjacent figure shows the input format of cell data in the AnnAGNPS input file. It was generated using Input Editor in the AnnAGNPS package. Soil variables were obtained from the SSURGO database (www.ftw.nrcs.usda.gov/ssur _data.html).


The SSURGO database contains two data sets: the spatial data that depicts the soil boundaries, and ancillary data for soil map-unit composition, soil properties, and interpretations (USDA-NRCS, 1995). Most of the AnnAGNPS input parameters were directly obtained from SSURGO, but some were derived from the data found within this database. Two files *comp* and *layer* in SSURGO were processed to get the required parameters. The *comp* file covers general soil-type information such as muid, compname, seqnum and others. The *layer* file includes specific soil information by layers and has the most important data of generating the soil-input file for AnnAGNPS. These required data were all generated using the method discussed early.
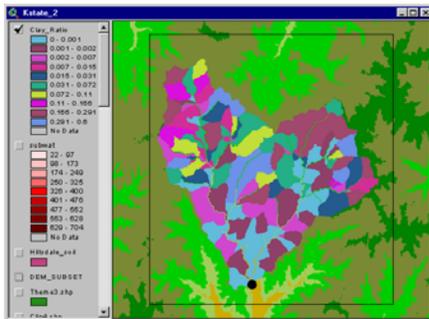
The spatial land use data were obtained from two sources; the USGS National Land Cover Data (NLCD) (landcover.usgs.gov/ftpdownload.html) and Digital Orthophoto Quarter Quadrangles (DOQQ) while some of the reference data were compiled in the Access database that was generated as past of this effort. The 10-year synthetic climatic data generated from GEM was used for this demonstration. Only the modeling results for the Little Bull Creek watershed, which

constitutes only a portion of the Hillsdale watershed, are presented. A new coverage was generated for the source accounting of clay ratio. A plot of rainfall and runoff for the 10-year simulation was created and presented here.
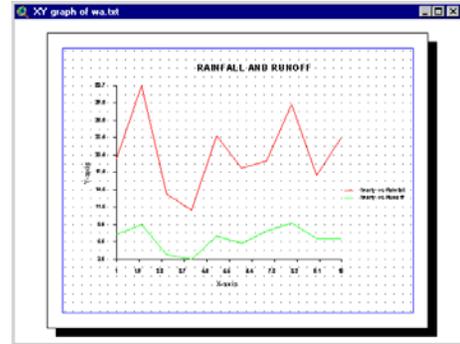
5. Conclusions

AnnAGNPS is an effective tool for watershed assessment and management. However, the complex modeling procedures and data preparation keep it from being an efficient modeling tool and thus limits its actual application. The development of this modeling interface simplifies data manipulation, model operation, and result interpretation such that the model can be populated and run by users with limited experience and knowledge of the model and its use. New features such as sensitivity analysis will be added into this integrated interface system.