Integration of SSURGO maps and soil parameters within a geographic information system and nonpoint source pollution model system

M. Di Luzio, J.G. Arnold, and R. Srinivasan

ABSTRACT: Computer based frameworks, including geographical information systems (GIS), associated digital data sets and integrated hydrological models, have become essential tools for the practical deployment of watershed assessment projects and ultimately for providing support for water quality protection and conservation. Within these frameworks digital soil information and data sets play a key role in defining the spatial distribution of important hydraulic variables and consequently of fundamental hydrologic functions connected with NPSs (Nonpoint Sources) and their modeling. This paper illustrates background and results from the development of a GIS based tool required for the use of Soil Survey Geographic (SSURGO) data sets (an undergoing development and delivery effort of digital soil maps by the Natural Resources Conservation Service) and primarily designed in specific support of the Soil and Water Assessment Tool (SWAT) model and its applications in the United States.

The tool is developed as an addition or extension to the model companion, the ArcView GIS interface in its latest version (AVSWAT-X) which is designed to define watershed hydrologic features; store, organize, and manipulate the related spatial and tabular data; and analyze management scenarios. Within this framework the tool expedites the otherwise complex inclusion operations of the SSURGO data, such as: (1) downloading, via the Internet, up-to-date SSURGO data sets (version I or II); (2) processing and managing variously formatted data sets in order to create the needed digital soil maps; (3) generating and/or storing the required soil physical and hydraulic model input parameters derived from pedo-transfer functions; and (4) seamlessly including them in any watershed modeling framework. Besides the operational advantages, the tool addresses the trend of employing soil maps with increasing detail in order to improve the landscape representation and reduce the modeling drawbacks involved with the aggregation of the spatial input. Soil map features with unprecedented detail along with the elaborated attributes have a potential for enhancing modeling assessments in support of water quality programs (i.e. Total Maximum Daily Load) at the watershed as well as stream reach scale, as indicated in basic yet important highlights in the paper.

Keywords: ArcView, AVSWAT, geographical information systems, GIS, nonpoint source pollution, soil hydraulic properties, soil maps, SSURGO, SWAT

In the development of management plans for the remediation of water quality problems at the watershed scale, in addition to PS (Point Source) pollution, nonpoint source pollution (NPS) loading and its implications are broadly recognized as fundamental components (USEPA, 1999; Lovejoy et al., 2000). Due to the complex nature of the hydrologic processes involved, the recommended evaluation of alternative management, land use, and conservation practices is practically accomplished using mathematical nonpoint source pollution models coupled with geographical information system (GIS) software. One such system consists of the Soil and Water Assessment Tool (SWAT) model and the companion ArcView SWAT (AVSWAT) software.

SWAT (Arnold et al., 1998) is a hydrologic model developed to assist water resource managers in assessing the impact of management and climate on water supplies, and nonpoint source pollution in watersheds and large river basins. Main model components include: weather, hydrology, erosion/sedimentation, plant growth, nutrients, pesticides, agricultural management, stream routing, and pond/reservoir routing. The model was developed in the early 1990's and represents over thirty years of model improvement within the U.S. Department of Agriculture's Agricultural Research Service (USDA-ARS). The latest version, SWAT2003, has the potential to expand as well as enhance the model applications (i.e. support of TMDL programs), due to the recent inclusion of routines for bacteria transportation, tile drainage, urban pollution, Green and Ampt infiltration equations, and several other features (Arnold et al., 2002).

AVSWAT (Di Luzio et al., 2004) was developed as an extension of ArcView® 3.x GIS software (ESRI, 1996) in order to support SWAT2003. It provides a full range of user-friendly and interactive input/output manipulation capabilities designed to help the user in performing numerous tasks. These include: segmenting and dimensioning the watershed both using a digital description of topography of the territory (DEM, Digital Elevation Model) and a geomorphological approach, importing and formatting the supporting data (i.e. land use and soil maps), analyzing and displaying output data from the model simulation, formulating management scenarios and performing basic calibrations. AVSWAT applications are not limited to a particular geographic location, however for the United States, users are supported with predefined data sets. One of these data sets is part of the STATSGO (State Soil Geographic) database (USDA, 1994) at the 1:250,000 scale underlying map.

Mauro Di Luzio is an assistant research scientist at the Texas Agricultural Experiment Station, Blackland Research Center affiliated with Texas A&M University in Temple, Texas. **Jeffrey G. Arnold** is a hydraulic engineer with the U.S. Department of Agriculture's Agricultural Research Service in Temple, Texas. **Raghavan Srinivasan** is an associate professor in the Spatial Sciences Laboratory at Texas A&M University in College Station, Texas.

Table 1. Soil physical and chemical properties input data for SWAT model.

Lumped parameters	Layer (max 10) parameters	
Soil hydrologic group	Layer thickness	Silt content
Maximum rooting depth of soil profile	Moist bulk density	Sand content
Fraction of porosity (void space) from which anions are excluded	Available water capacity	Rock fragment content
Potential crack volume	Saturated hydraulic conductivity	Moist soil albedo
	Organic carbon content	Universal Soil Loss Equation soil erodibility
	Clay content	Electrical conductivity

The Soil Survey Geographic (SSURGO) database (USDA, 1995) is being developed by the Natural Resources Conservation Service (NRCS) at the map scale of between 1:12,000 and 1:63,360. Spatial differences between the two databases available (STATSGO and SSURGO) for small areas or stream reach watersheds could be substantial (Lathrop et al., 1995; Juracek and Wolock, 2002). In addition, the variation of detail in the spatial segmentation has a tremendous effect within nonpoint source pollution models (Brown et al., 1993; Vieux and Needham, 1993), particularly on erosion, transport capacity, and sediment yield results.

In the SWAT model, conceptualization of a watershed is preferably parameterized using a series of elementary hydrologic response units (HRUs), rather than adopting a lumped parameter representation. Each hydrologic response units corresponds to a particular combination of soil and land use-land cover categories superimposed within each subwatershed. A number of inputs to the model are hydrologic response units based (i.e. soil physical and chemical parameters, land use and soil management operations, and groundwater). Specifically the model requires as input the soil parameters listed in Table 1. The SWAT model, being physically based, is sensitive to these parameters which are involved in all of its simulation modules (Arnold et al., 1998).

Muttiah and Wurbs (2002) have shown scale-dependent effects of soil variability on watershed water balance with the SWAT model. In particular, improvement of simulation results from the reduction of the spatial aggregation of the input, or the increase in the number of computational units (HRUs), have been highlighted (Mamillapalli et al., 1996; Bingner et al., 1997; FitzHugh and MacKay, 2000 and 2001; Haverkamp et al., 2002).

SSURGO data sets to address watershed assessments. The development of methods allowing for the use of the SSURGO data sets within AVSWAT is primarily important for simulation implications and ultimately for the success of the addressed watershed applications. In order to pursue this specific objective several issues need to be addressed:

1. Soil hydraulic parameters. NRCS's soil geographic databases (Reybold and TeSelle, 1989) commonly consist of georeferenced digital map data and associated tables of soil interpretation data. Each polygon in the geographic database may contain a number of components (locations of the components are not specified, but their proportions are), equivalent to soil series or phases, each of which is characterized by a less than complete set of soil parameters. Several of these parameters are also provided as ranges instead of single values. Methods to estimate single, representative soil hydraulic properties, as required by the SWAT model, need to be applied from these generic soil data.

2. Soil data updates. Data updates occur because new information becomes available and because more advanced software solutions are adopted by NRCS. NRCS periodically revises soil series definitions based upon new levels of certification, information, and research. Tables contained in a county-level soil survey, such as SSURGO tables, may be updated more frequently than the survey is republished. This revised data is generally reviewed at the NRCS field office covering the area of interest. Through this process, as soil surveys are updated, the overall number of soil series exhibited on soil survey maps (and used to formulate map unit attributes) could be changed through the combination or reclassification of some series. In other areas, soil series may also be split to more accurately describe the predominant soil characteristics. Variation of the digital format of the database is also likely to happen over time as described later in this paper.

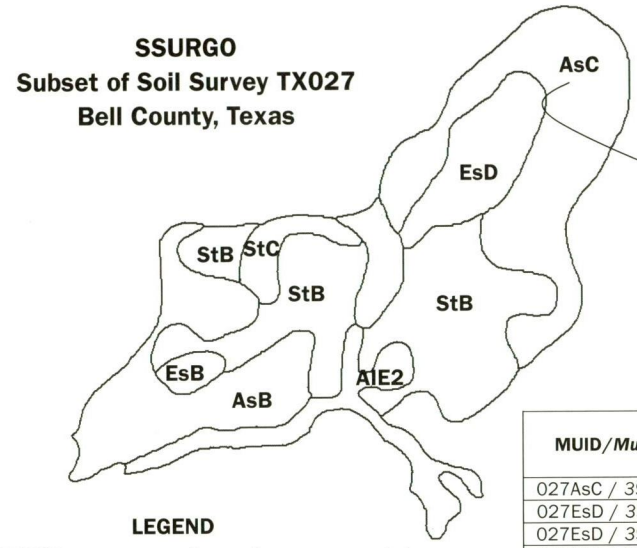
3. Delivery and management of spatial soil data. AVSWAT helps the user in dealing with several tasks required to set up and run the SWAT model. However the user still needs to cope with the collection of the necessary input data sets. Data delivery options are expanding the accessibility and reducing the cost of many digital data sets. Many of these changes are linked to the Internet. However the skill and the steep learning curve required for novices to manage GIS data sets, and to work with GIS-model systems (Potter et al., 2000), combined with the intrinsic complexity of the databases associated with soil maps (Lee et al., 1999) (i.e. dealing with issues such as handling various electronic formats and coordinate systems, and the proliferation and redundancy of data available through the Internet), pose significant challenges to users.

The objectives of this study are to: (1) describe the essential technical features of the SSURGO database and important aspects for its usage; (2) develop a tool for the use of this database within the AVSWAT framework, with special attention given to the issues highlighted above; and (3) highlight some specific benefits of the implemented tool as well as of the data sets in example applications.

Methods and Materials

SSURGO database. SSURGO is defined, according to the National Soils Handbook (USDA, 2002a), as the most detailed of the digital soil geographic databases prepared by NRCS. SSURGO consists of county-level maps, metadata, and tables, which define the proportionate extent of the component soils (phases) and their properties for each map unit. SSURGO is the primary product of the National Cooperative Soil Survey (NCSS) program, a joint effort of NRCS and other federal, state and local agencies, to place soil survey data in digital format in order to facilitate its usage within GISs. The mapping bases are normally orthophoto-quads, or 7.5-minute topo quads. The data are archived in 7.5-minute topo quad units and distributed as complete coverages for each survey area (which may consist of a county, multiple counties, or parts of multiple counties) with all adjoining 7.5-minute units matched within it. Digitizing of the county survey

Figure 1
Spatial units and tables linkage in the SSURGO I and SSURGO II data sets (the keys and numbers in italic following the "/" characters report the differences of the SSURGO II version).



Soil survey table

STSSAID	Soil survey properties
TX027	—

Map unit tables

STSSAID	MUSYM	MUID/*Mukey*	Map unit properties
TX027	AsC	027AsC / *391067*	—
TX027	EsD	027EsD / *391089*	—
TX027	StB	027StB / *391128*	—
—	—	—	—

Component tables

MUID/*Mukey*	SEQNUM/*Cokey*	COMPPCT/ *Comppct_r*	COMPNAME	Component properties
027AsC / *391067*	1 / *Mukey* : *519728*	100	AUSTIN	—
027EsD / *391089*	1 / *Mukey* : *519760*	56	EDDY	—
027EsD / *391089*	2 / *Mukey* : *519761*	44	STEPHEN	—
027StB / *391128*	1 / *Mukey* : *519819*	100 / *88*	STEPHEN	—
—	—	—	—	—

Layer/Horizon tables

MUID/*Mukey*	SEQNUM/*Cokey*	LAYERNUM/ *Chkey*	LAYDEPH/ *Hzdepb_r*	Layer properties
027AsC / *391067*	1/*Mukey* : *519728*	1/*Mukey* : *1142662*	36	—
027AsC / *391067*	1/*Mukey* : *519728*	2/*Mukey* : *1142663*	74	—
027AsC / *391067*	1/*Mukey* : *519728*	3/*Mukey* : *1142664*	127 / *203*	—
027EsD / *391089*	1/*Mukey* : *519760*	1/*Mukey* : *1142737*	13	—
027EsD / *391089*	1/*Mukey* : *519760*	2/*Mukey* : *1142738*	20	—
027EsD / *391089*	1/*Mukey* : *519760*	3/*Mukey* : *114739*	102	—
027EsD / *391089*	2/*Mukey* : *519761*	1/*Mukey* : *114740*	25 / *36*	—
027EsD / *391089*	2/*Mukey* : *519761*	2/*Mukey* : *114741*	36 / *47*	—
027EsD / *391089*	2/*Mukey* : *519761*	3/*Mukey* : *114742*	71 / *203*	—
—	—	—	—	—

LEGEND

STSSAID	State soil survey area symbol
MUSYM	Map unit symbol
MUID/*Mukey*	Map unit identifier/*map unit key*
SEQNUM/*Cokey*	Component ID = sequence number/ *component key*
COMPPCT/*Comppct_r*	Component percentage
COMPNAME	Component name
LAYERNUM/*Chkey*	Soil layer number/*horizon key*
LAYDEPH/*Hzepb_r*	Deepest layer depth/*distance top-bottom soil horizon (centimeters)*

information in accordance with NRCS-established specifications and standards is a complex and time consuming task (Beck et al., 2002). NRCS is in the process of working with the custodians of digitized surveys to review and check for standards certification (USDA, 1995). Once the attribute data associated with soil map units have been entered into the NRCS database management system, they are readily available for use with digitized soil maps. The status as of December 3, 2003 (USDA, 2003) was that 1708, out of 2519 soil surveys areas in the United States were digitized and archived.

Currently, two distinct and alternative data structures are deployed, the SSURGO first and SSURGO current versions, hereafter referred to as SSURGO I and SSURGO II.

SSURGO I has been in use since the 1990's and is the legacy of the basic database system previously adopted by NRCS and called the State Soil Survey Database (SSSD). The State Soil Survey Database system was based on the Soil Interpretation Record (SIR) and the attribute data, linked to the soil map units, which was identified as the map unit interpretation record (MUIR) relational database (USDA, 2002b). The map unit interpretation record database contains a collection of soil physical and chemical properties, interpretations, and performance data for a soil area, and its map unit components and component layers. These include available water capacity, soil reaction, soil erodibility factors, hydric soil ratings, ponding, flooding, water table depth and duration, bedrock, interpretations for sanitary facilities, building site development, engineering, cropland, woodland, recreational development, and yields for common crops, site indices of common trees, and potential production of rangeland plants. The tables in the database are linked to the data dictionary tables for definition of their codes and elements and are hierarchically and logically organized to contain related items. The tables are grouped in five sets: soil survey area tables, lookup tables, map unit tables, component tables, and layer tables. Details of the attributes included in the SSURGO database can be found in NRCS's publications (USDA, 1995). The map unit is digitally specified by a single or multiple georeferenced polygon(s). The unique specification of a soil map unit within each survey area is the map unit identifier (muid), created by the concatenation of the soil survey area symbol (ssaid) and the map unit symbol (musym), which links the graphic feature to attribute data in the MUIR. Each map unit represents an area dominated by up to three phases of soil series (components). For each component, there are 60 properties and interpretations in 84

different data elements (Component Tables) (i.e. flooding). Once again, there is no geographic delineation for the locations of the components within a map unit. The main component table reports the extent of a component as a percent of the map unit. For each component, one to six soil horizons (layers) are possible. For each layer, up to 28 soil properties are defined (i.e. percent clay) in the layer table. Figure 1 shows a simplified diagram of the soil data table structure for each map unit in SSURGO I.

An important issue is that most of the data in the database were stored as a range of soil properties, depicting the range for the soil survey area. In some parts of the country ranges were frequently quite large since they were designed to be inclusive of all variation that might be found within an entire soil series.

SSURGO II is the evolution of SSURGO I and is the outgrowth of the current development and adoption by NRCS of the National Soil Information System (NASIS) (USDA, 2000) to host and maintain the soil survey data. With the National Soil Information System, NRCS is improving several aspects of the delicate and complex itinerary of the soil survey data, starting with field data collection and interpretation, through effective storage methods, and ending with delivery to the public. The National Soil Information System includes generic software tools for managing data in the relational database system containing information about the soil survey database. Managing capabilities of the software include: editing data, querying the National Soil Information System database, generating reports, and exporting data. However, access to NASIS is restricted to authorized NRCS scientists and with the advent of NASIS 5.0, the data resides on a central server while the peripheral soil survey offices are provided with remote connections to view and edit the data they own. In addition, since 2001, the export function that provided the former SSURGO I tables has been modified to provide the new SSURGO II data tables. These tables are designed for GIS uses and contain selected data and interpretations that have been aggregated to a single value for each map unit. SSURGO II resembles the NASIS's database structure and is consistent with the one used in the Customer Service Toolkit and the Soil Data Viewer, two desk top solutions deployed by NRCS to allow developing and delivering technical assistance products such as maps and narrative documents related to conservation planning and resource assessment to customers. The SSURGO II data set is now a complete package of spatial, tabular, and metadata information for an individual survey area. The spatial data are certified to meet the map development standards; tabular data are certified to match the correlation document for the survey; and the spatial and tabular data are jointly certified to have internal referential integrity. The metadata complies with Federal Geographic Data Committee (FGDC) Version 2 metadata standards and provides a wealth of documentation describing the data structure of the SSURGO II, tables, columns, indexes, and relationships as well as a variety of attributes for each of these database objects (USDA, 2001a). Attributes include table and column descriptions and detailed domain information.

In the development of SSURGO II the existing SSURGO I data are being reviewed and archived in the newer format. Important differences between the two datasets are described in a National Soil Information System tutorial (USDA, 2001b). Essential issues for our goals are:

1. NRCS distributes SSURGO data to the general public, in either one of the two formats, via CD-ROM, and online primarily through the National SSURGO Website (http://www.ncgc.nrcs.usda.gov/branch/ssb/products/ssurgo/data/index.html) via file transmission protocol (FTP) (ftp://ftp.ftw.nrcs.usda.gov/pub/ssurgo/online98/data/). Currently, the Soil Data Mart (http://soildatamart.nrcs.usda.gov/) is being deployed by NRCS to improve data distribution. "Soil Data Mart will supersede the National SSURGO website, but this transition will be ongoing for most of 2004. During this period of transition, data for a particular survey area may reside at either site, but never at both sites simultaneously" (USDA, 2004).

2. Data sets are distributed in a similar, but not identical, compressed format, and the map component is slightly different. The tables, named differently, are still delivered in ASCII text files: in SSURGO I the columns are tab delimited while in SSURGO II they are pipe delimited.

3. Map units have an unlimited number of components, and components can have an unlimited number of horizons.

4. Some data elements in SSURGO II are physically and conceptually located in different tables and fields than in SSURGO I and some correspondent values have been updated. Tables are alternatively linked using the following keys: mukey vs. muid, cokey vs. seqnum, chkey vs. layernum).

5. SSURGO II includes representative values (RVs) for several properties, including derived soil hydraulic properties.

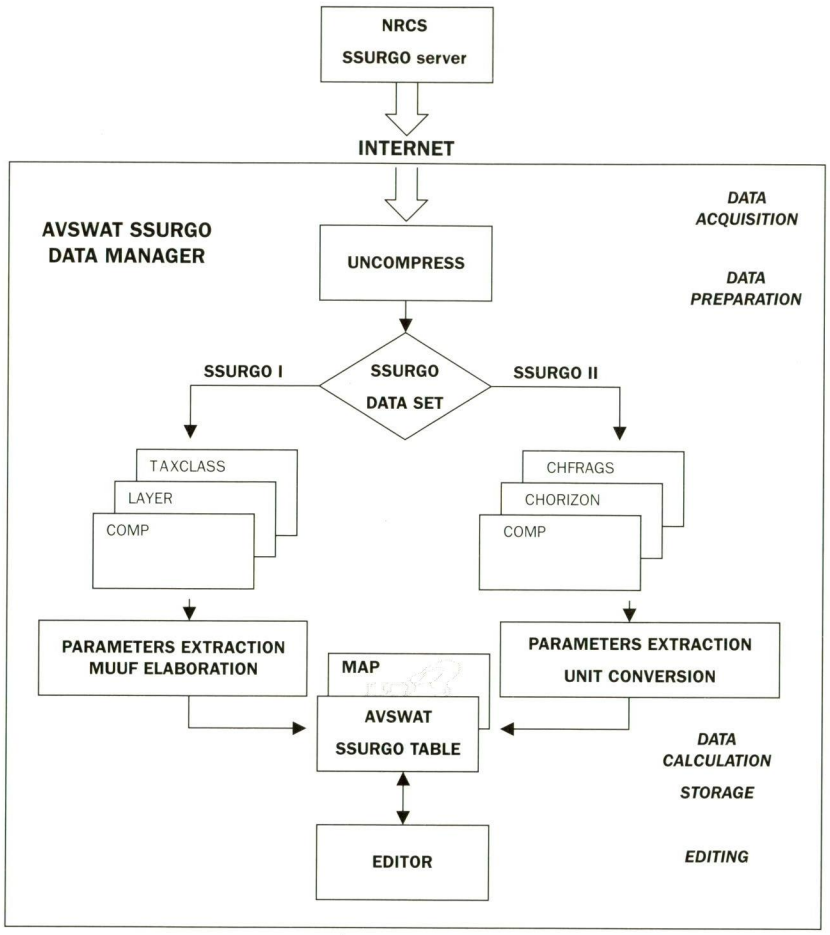
Figure 1 shows a simplified diagram of the soil data tables structure and linking keys for each map unit in SSURGO II side by side with SSURGO I. The figure also highlights some numerical updates (i.e., component percentage, soil layer depth) for the sub-set of the survey area data example (Bell County, Texas).

AVSWAT. The additional component of AVSWAT, the subject of this paper, has been designed and implemented as a tool for coping with these issues faced with the usage of SSURGO data sets. AVSWAT provides several dialog interfaces to help users accomplish various tasks, such as define the watershed hydrologic features; store, organize, and manipulate the related spatial and tabular data; and analyze management scenarios. ArcView's Dialog Designer and Spatial Analyst Extensions were used to create these dialog interfaces. Dialog Designer provides readily usable plug-in controls, such as menus, buttons, and tools. The specific tasks, activated using these controls, were developed with AVENUE, the ArcView 3.x's programming language. In addition, the Spatial Analyst extension provides spatial analysis procedures for raster data, whereas ArcView alone provides spatial analysis capabilities using vector data.

SSURGO extension for AVSWAT (SEA) was developed using the same basic materials and as a specific extension of AVSWAT. The following four major modules of SEA can be used to acquire, process and utilize SSURGO data sets: (1) data acquisition; (2) data preparation; (3) data calculation, storage and editing; and (4) data inclusion and merge, listed in the natural order in which these sections should be used. The first three units are grouped within a manager interface tool whose overall objective is the extraction of the map unit polygons and the creation of the derived table of attributes containing the SWAT soil inputs for each soil survey data set (Figure 2).

The data acquisition module contains procedures to download SSURGO data sets packaged for each soil survey area in a com-

Figure 2
The SSURGO data manager architecture in AVSWAT.



LEGEND

NRCS	Natural Resources Conservation Service
AVSWAT	ArcView GIS (geographical information system) SWAT model system
COMP – LAYER – TAXCLASS	Map unit component – soil layers – soil classification tables in SSURGO Version I
COMP – CHORIZON – CHFRAGS	Map unit component – soil horizons – horizon fragments tables in SSURGO Version II
MUUF	Map unit user files

pressed format, directly from the NRCS's servers through the Internet. Procedures are used for performing the following tasks:

1. Automatically log on to the data server.
2. Check the current availability of the data sets relative to a user's selected soil survey area.
3. Apply an Internet protocol through a software component (dynamic link library) customized to automatically download the required compressed files and control the transmission process.
4. Save these files within an organized framework.

Since data may reside now at the new Soil Data Mart, the user is driven through the steps necessary to check the data availability and download them via a web browser.

The data preparation module contains procedures to set up the required data by performing the following tasks:

1. Detect the structure of the data (SSURGO I or SSURGO II).
2. Extract the map unit polygons and the required ASCII tables with distinct procedures upon the type of data structure. The tables are listed within figure 2.
3. ASCII tables are converted into ArcView standard dBASE tables.

The data calculation, storage and editing module contains procedures that were developed in response to the need for single attribute values and water retention characteristics for the SWAT model.

When using the SSURGO I, a methodology developed by NRCS (Baumer and Rice, 1988) is applied. The soil hydraulic properties are estimated from surrogate soil characteristics for which data is more easily obtained during traditional soil surveys. The map unit user files computer program was built based on this method (Baumer et al., 1994). The map unit user files were written in FORTRAN to estimate soil hydraulic properties directly from the SIR. Unfortunately the map unit user files is a stand-alone program consisting of its own database, stored using the Soils-5 formats. The map unit user files have not been updated since December 1995 because the Soils-5 format is no longer used by NRCS. For AVSWAT a complete new program resembling the map unit user files has been developed using AVENUE. The program includes several procedures for performing the following tasks:

1. Access the SSURGO I derived tables.
2. Merge the input data set to predict certain values, which are not available in the soil databases but are needed as inputs into the next task calculation procedure.
3. Calculate the required soil properties at the map unit, component and layer level. Soil texture, organic matter, bulk density, clay activity and macroporosity are the main parameters used to make the basic soil-water retention estimates. The van Genuchten's model (van Genuchten and Nielsen, 1985) is used for modeling the water retention curve. The saturated hydraulic conductivity of the soil is calculated by this model. The Mualem (1976) method is used to estimate the unsaturated hydraulic conductivity as a function of water content or suction.
4. Adjust the unit of measure.

The SSURGO II data includes (unlike the SSURGO I) "representative values" along with attribute ranges. These values have been populated by NRCS scientists in response to the mapping and modeling need for a single soil property value specific to a map unit, a map unit component, or a horizon. In addition the applied methods are meant to supply those soil properties values that are not commonly available in soils databases (i.e. saturated hydraulic conductivity, unsaturated hydraulic conductivity, and soil-water retention characteristics). The

methods used to define the representative values are primarily based on field and lab data (USDA, 2002a), and pedo-transfer algorithms calculated from attribute ranges. NRCS scientist applied a simpler method than the map unit user files: a one-point procedure developed by Williams and Ahuja (1992), based on the equation for the ln-ln water retention function developed by Gregson et al. (1987).

SEA includes procedures performing the following tasks:

1. Read the SSURGO II tables.
2. Extract the appropriate RVs from the tables and adjust the unit of measure.

The soil data required by the SWAT model is listed in Table 1. SEA includes the storage of the derived soil data (either from SSURGO I or SSURGO II) required by SWAT in database tables (dBASE tables) referenced by the soil survey identifiers. Each of these tables can be further reviewed using a user-friendly dialog interface before the data are included in the study watershed.

The data inclusion and merge module contains procedures embedded in the Land Use and Soil Definition module within AVSWAT (Di Luzio et al., 2004) completing the hydrologic scheme of the watershed. This enhanced module allows these additional tasks:

1. Check the consistency of the derived SSURGO maps and the associated soil attribute tables for the survey areas superimposing the watershed.
2. Change the coordinate system, rasterize, and clip the original soil maps to the watershed area. If multiple survey areas are involved, the respective raster maps are merged and eventual concomitance of SSURGO I and SSURGO II are handled.
3. Control the reclassification of the SSURGO layer and the distribution of the newly referenced land use/soil combinations (HRUs) using the raster functionalit provided by the spatial analyst.
4. Extract (each single soil component record can be targeted) and format in the required ASCII SWAT model input within the watershed model simulation framework.
5. Review the data at the sub-basin and HRU level, part of the standard AVSWAT capability.

Results and Discussion

The SWAT model requires a number of soil hydraulic parameters in a spatially distributed venue as input. Although AVSWAT, the associated ArcView GIS interface, speeds up and supports the model's users, an additional tool was needed to expedite the collection of the necessary digital, large scale, soil maps, and to reduce the complexity, time, and labor involved in generating the associated parameters for a watershed simulation. In particular, applications of the model in support of watershed programs in the United States called for the implementation of the developing NRCS's SSURGO data sets.

A new architecture for AVSWAT has been developed to host extensions providing additional and optional functionality, features and data sets. The SEA is the first customized extension developed to provide large-scale soil maps and to derive soil map attributes as an enhanced alternative to the STATSGO input. Along with AVSWAT, this tool is almost entirely programmed in AVENUE, and therefore is platform independent and is transferable between Windows versions and UNIX. AVSWAT and SEA contain explicit alternative procedures for those functions that are platform dependent, such as the execution of the model access to the Internet.

A key requirement was the capability to derive single representative values for each soil type and profile for the model input. A parallel requirement was also the ability to use both the formats (SSURGO I and II) which are used by NRCS to deliver SSURGO data sets. We believe that providing for SSURGO I, the map unit user files program completely embedded within the SEA, gives the user a reliable method, still used officially in several states, while eliminating the need to learn and apply the stand alone software, and then collect and input manually the results. Also the SSURGO II data sets contain several tables (more than 50) from which the manual extraction of the parameter values could be a difficult task. In this case the implemented method provides a straightforward way for distilling the necessary inputs. These can also be revised using a customized tool before the parameters are actually included in the simulation flow and/or used as inputs for other models.

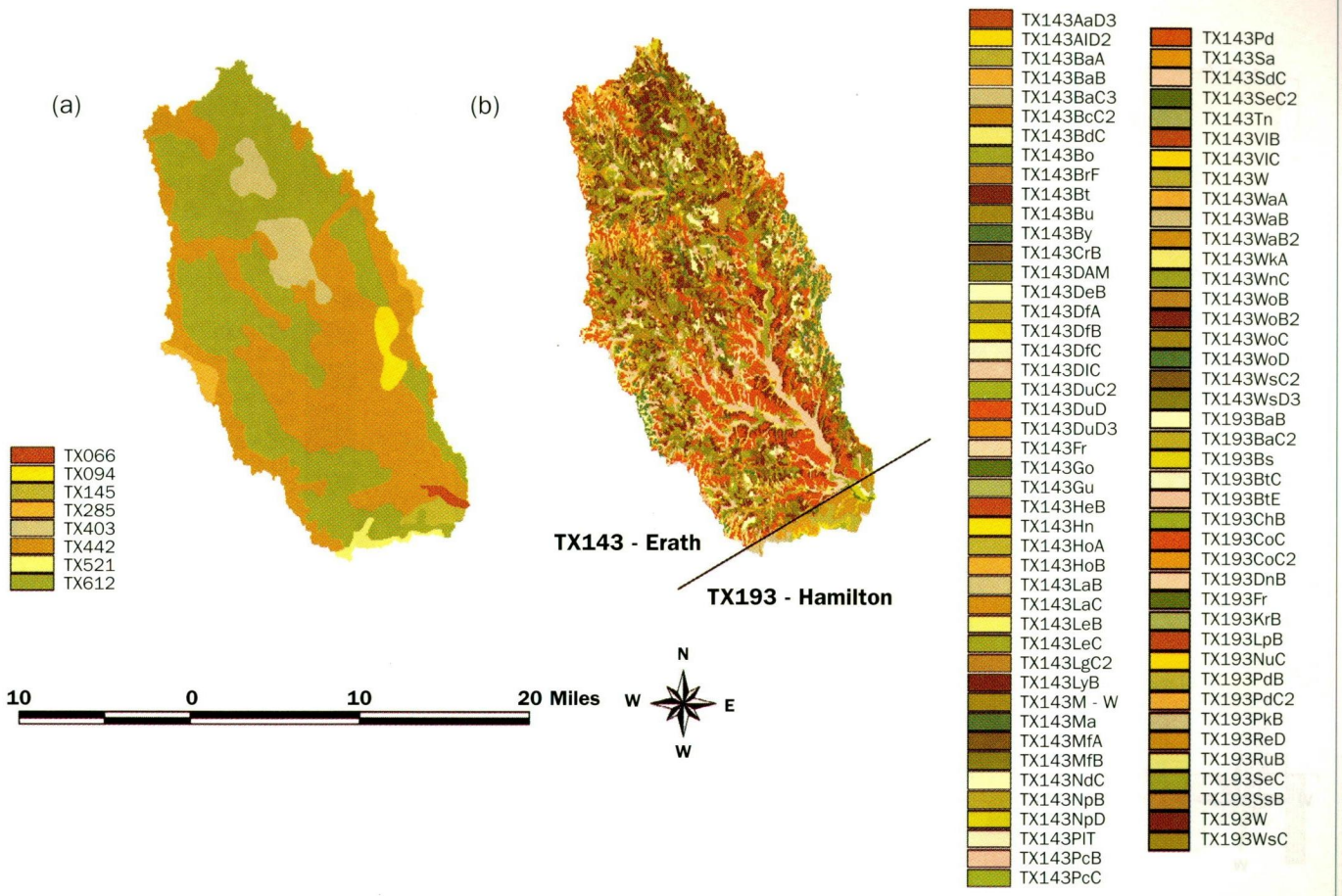
Another important requirement was to provide a way to overcome the challenging, time consuming and cumbersome task of collecting and managing raw digital data that need to be customized and formatted for model use. Without the developed tool the conventional method would require a remarkable amount of time. Several manual steps would be needed: find and access the source data server, register and login over the Internet onto the data server, browse the data location, determine data specification and formats, download selected files, determine the way to uncompress the files, uncompress the files, determine version of the data, determine the needed tables and digital map, extract the tables and the map, convert the tables in dBASE format and import them, import the map. These preliminary steps are accomplished by the Data Acquisition and Data Preparation modules of the developed tool using only a few mouse clicks. In addition the provided Internet connection to the main NRCS server gives the user access to the most updated version of the data sets. Since the data set files are quite large (approximately 10 to 25 MB for each survey area) the download time is restricted by the speed of the available access to the Internet. However this issue is expected to become less limiting over time as the speed of the Internet connections and the capability of computer equipment follow a general improving trend. In addition NRCS, complying to the FGDC directives, is planning the development of a centralized delivery point (Soil Data Warehouse) able to provide an around the clock official source of data, including SSURGOs, and more advanced protocols for the electronic data access through the Internet (USDA, 2001c).

A methodology developed for an earlier version of AVSWAT and SSURGO I data sets (Buland et al., 2000) would require around 30 manual steps employing around two days to accomplish the collection and management of a soil survey digital data set along with the derivation of the soil properties. The semiautomatic tools in SEA accomplish the same process in three steps and 30 minutes (including common 20 minutes of download time) expediting the same tasks and also eliminating chances of manual mistakes.

However, another important requirement of our tool was to provide a smooth inclusion of the soil data sets within the modeling scheme of the watershed. This is fully accomplished by the Data Inclusion and Merge section. Along with the AVSWAT functionality, the user is saved from other challenging and specific GIS tasks, including

Figure 3

The Upper North Bosque River Watershed (UNBRW) located in north central Texas: using (a) STATSGO and (b) SSURGO soil geographic data sets (SSURGO I data sets were available and used for the survey area TX143 and SSURGO II for the TX193).



the alignment of diverse coordinate systems, the joining of the soil attributes, and the overlay of the land use-land cover map, particularly difficult when data sets of the two different versions superimpose the same study watershed (see Figure 3). SEA includes routines specifically customized to manage the SSURGO map and attributes, leaving the user the freedom to set the distribution of the land use soils combinations (HRUs) as needed and to include specific soil components in the simulation.

The final goal of the implemented methodology was to integrate digital soil characteristics of the U.S. landscapes with the best available detail in AVSWAT. Figure 3 shows a comparison of the STATSGO and SSURGO data sets implementation in the Upper North Bosque River Watershed, covering about 932 square kilometers (360 sq mi) in the north central Texas. In the same watershed, while a total of eight soil map units are distinguished using STATSGO, a total of 86 units are defined using SSURGO. The watershed was segmented (62 sub-watersheds) using the National Elevation Dataset (NED) (Gesch et al., 2002) and the land cover map from the National Land Cover Data (NLCD) (Vogelmann et al., 2001) [both 30 meters (98 ft) resolution] and the number of HRUs defined were 1199 and 9635, respectively using STATSGO vs. SSURGO.

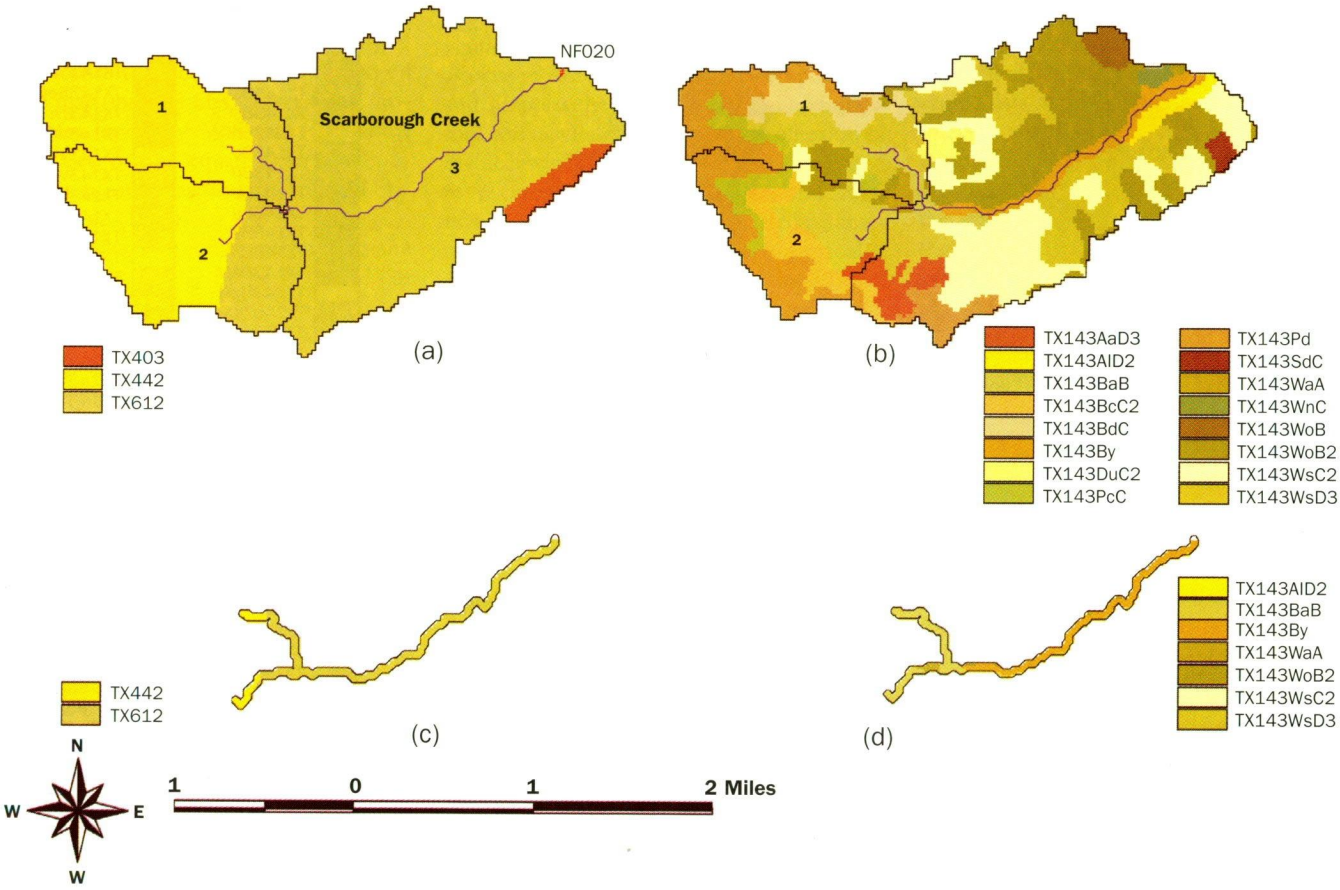
As mentioned above, it has been highlighted in the literature that the increase of the average number HRUs for each sub-watershed provides potential improvement to the simulation performances. However, simulation improvements are also dependent on scale, geomorphological and landscape settings, and complex hydrological interaction mechanisms.

Nevertheless the implementation of soil datasets with increased detail brings an objective advantage, such as the opportunity to formulate more precise and diversified management strategies (at the HRU level) related to the soil types within the study watershed and its composing sub-watersheds. Two examples are briefly discussed:

(1) The potential of animal wastes to recycle nutrients, build soil quality, and increase crop productivity is well established. A concern with land application of animal wastes is that excessive applications may damage the water body through excessive accumulation (and subsequent release) of nutrients. Therefore the main objective of the best management practices is the utilization of the nutrients in the animal manures to meet crop nutrient removal rates during harvesting and utilization of the biomass. As an example, the management of dairy wastes is a key objective for the Upper North Bosque River Watershed where dairy farming is the predominant agricultural practice: 70 dairies with over 30,000 milking heads of cattle are

Figure 4

The Scarborough Creek Watershed in the Upper North Bosque River Watershed: (a) and (c) STATSGO , (b) and (d) SSURGO soil units, in the watershed, (a) and (b), and the riparian area, (c) and (d).



included in this area (Kiesling et al., 2001). This environment creates a particular water quality concern since the Bosque River discharges into Lake Waco, which is used as the public drinking water supply for the city of Waco and several adjoining communities. Current and previous programs to improve water quality in the watershed include extensive sampling as well as the use of computer simulation models to identify sources of nutrients and to determine how they are partitioned throughout the watershed to improve water quality.

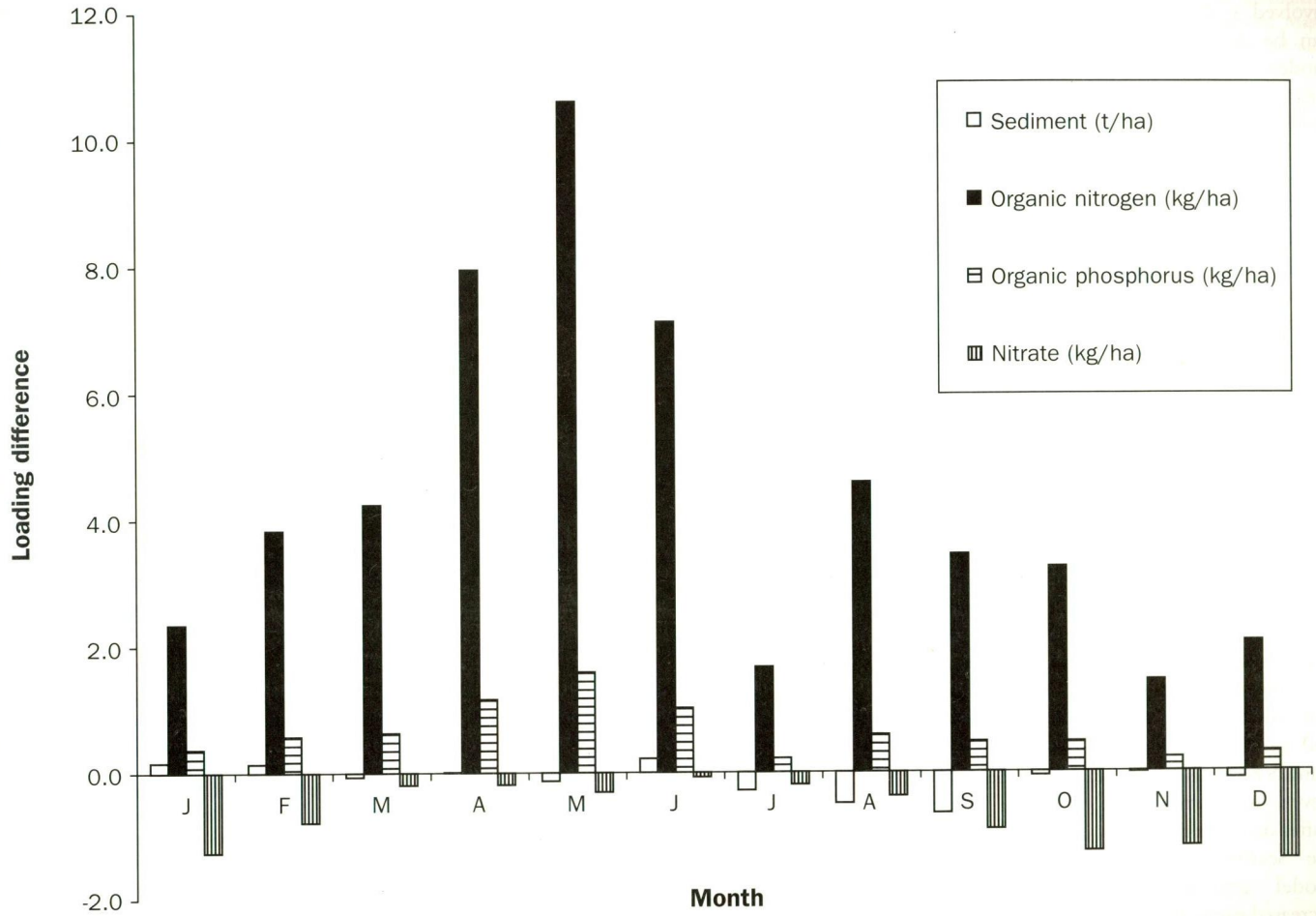
In one application to the Upper North Bosque River Watershed, Saleh et al. (2000) have shown the advantages of using the SWAT model to evaluate varying conditions within the watershed. In particular the evaluation of pollutant loading in the presence vs. absence of dairy operations was realized. A critical 7.9 ha (19.5 ac) Upper North Bosque River Watershed's sub-watershed, draining into the Scarborough Creek, was highlighted in this study. This sub-watershed consisted of an estimated milking heard size of 1256, dairy cow density of 1.6 (cows/ha or 0.65 cows/ac) and 45.4 percent of the area with waste application fields.

We applied the AVSWAT system to the Scarborough watershed. Figure 4 (a) and (b) shows the sub-watershed, split into three subunits, and identified at the outlet location (labeled NF020) obtained with the data described above. Using the STATSGO data sets, three soil map units and 41 HRUs were defined. Using the SSURGO data sets, 16 soil map units and 151 HRUs were defined. The SWAT model was applied with no regard to the waste application location, only allowing automatic fertilization (this is an option in the model to meet the needs of the growing

Table 2. Land use classes in the Scarborough Creek Watershed.

Land use	Percent watershed area
Range-brush	9.8
Pasture	19.08
Range grasses	51.82
Water	0.96
Wetland mixed	0.01
Urban - commercial	1.2
Forest deciduous	6.19
Forest evergreen	2.42
Agricultural land	8.52

Figure 5
Average of the monthly output differences obtained for the Scarborough Creek Watershed simulating the period of 1988-1996 (simulation output using STATSGO was subtracted from the SSURGO derived output).



crops) in the less than 10 percent agriculture areas of the sub-watershed (see land use distribution in Table 2) and using the remaining default values provided by AVSWAT.

We compared the simulations for the two cases, obtained using the same precipitation and temperature inputs used by Saleh et al. (2000) but in the nine year period from January 1988 to December 1996. We verified the absence of water flow changes to the provided soil data sets at the outlet of the watershed. We also verified a slight sensitivity on the sediment yield, and sensitive loading differences in some nutrient components, particularly the organic nitrogen. This is shown in Figure 5, where the difference of the monthly outputs for the two methods (STATSGO outputs are subtracted from the SSURGO outputs) were averaged and plotted. It is not the objective of this study to fully investigate the reasons of these differences, but it is worth noting that the average soil organic carbon content (percent of soil weight) in the two top soil layers of the watershed are 0.73 and 0.52 using STATSGO, and 0.85 and 0.27 using SSURGO. This is an important issue since organic matter is an indicator of soil health or quality, which in turn can reflect the soil's ability to cycle nutrients, both in the field as well as in the simulation context.

In summary, while precision agriculture (Vanden Heuvel, 1996) and soil quality protection (Eghball, 2002) call for site-specific information about variability (spatial and temporal) of key landscape variables, modeling of the same or alternative management practices (i.e. waste nutrient applications) meets the same requirements. In this perspective, while the proliferation of remote sensing technology promises to provide fast and large availability of topographic and land cover information, digital soil data sets, such as SSURGO, show their importance in the watershed modeling of wide areas of the United States.

(2) The management and preservation of riparian corridors have a remarkable potential to help in controlling the transport of sediment and chemicals to stream channels (Lowrance et al., 1985). In fact, given its streamside nature, the riparian ecosystem is traversed by the movement of water and water-born chemicals and sediments from the upslope agricultural fields and forests. Vegetation processes (i.e., sedimentation, streambank stabilization, and water temperature reduction) and soil processes (i.e., gaseous loss of nitrogen via microbial denitrification and trapping of adsorbed pesticides) provide a basis for nonpoint source pollution control by riparian corridors. The complexity of the riparian ecosystem, which is highly

connected to the areas adjacent to the stream as well as the entire watershed, and the involved hydrological fluxes and processes, can be described by watershed simulation models. Initial attempts to include riparian areas assessments in the SWAT model are underway (Arnold et al., 2002). The model would help in understanding the observed effectiveness of managed and unmanaged riparian buffers (Riley, 2000) in removing nutrients from surface water and groundwater flowing within a watershed. In particular the potential for the movement of nutrients off-site still exists, even when animal wastes are applied at agronomic rates. The modeling tool would help in defining animal waste disposal that would minimize environmental degradation while avoiding unnecessary land use restrictions. In this case landscape position (i.e. distance from the stream and land use) and soil information (i.e. permeability, organic matter, clay content) are linked to hydrological processes like the water residence time, denitrification and the absorption of phosphorus (P) and nitrogen (N) forms, and can play a key role for optimizing manure application using the SWAT model. For the watershed cited in example (1), Figure 4 (c) and (d) show the soil map units obtained using STATSGO (two units) and SSURGO (seven units) respectively, in the riparian buffer delineated 100 feet around each side of the Scarborough Creek. For the SWAT model application, this result provides increased simulation insights and potential for the derived planning strategies.

Ultimately these examples show that an effective use of the SSURGO vs. STATSGO data sets within AVSWAT has the potential to enhance evaluations at the watershed as well as the small reach watershed scale. The knowledge of the landscape with increased precision (i.e. soil type information) will allow a better assessment of management alternatives to meet agronomic, soil, and water quality protection objectives.

Summary and Conclusion

A methodology was developed to fully integrate SSURGO data sets with SWAT and AVSWAT. SEA (SSURGO Extension for AVSWAT) was created as an optional addition to the newly extendable AVSWAT framework. SEA provides access to soil survey data sets semi automatically downloaded via the Internet, and shields users from the complexity of these soil databases.

Single representative values (calculated using pedo-transfer functions) required by SWAT for a number of soil parameters, are either generated, when the SSURGO data sets are still distributed in their earlier format (MUIR scheme in SSURGO I), or extracted from the tables distributed by NRCS (resembling the NASIS's format in SSURGO II). SEA stores the soil survey derived data in tables and customized data dialog boxes to provide editing and error checking capabilities. Finally the soil maps and the attributes are seamlessly clipped over any study watershed framework in the AVSWAT system. SEA's practical advantages are the reduction of the time for data acquisition and organization of up-to-date data sets, as well as increasing the input detail. This has the potential to improve model simulation and play a key role in planning management strategies (i.e. TMDL programs) requiring adequate spatial precision.

Future developments of SEA are related to the method of distribution of data by NRCS, currently deploying and updating a natural resources gateway (including the Soil Data Mart) to facilitate the location and distribution of spatial and non-spatial natural resources. Updates of the SSURGO data set location would require only the update of the SEA's Data Acquisition component, including the application of a new Internet based direct data assess protocol that NRCS is planning to provide to the end users.

List of acronyms

AVSWAT = ArcView GIS - SWAT model system

DEM = Digital Elevation Model

FGDC = Federal Geographic Data Committee

FTP = File Transmission Protocol

GIS = Geographic Information System

HRU = Hydrologic Response Unit

MUIR = Map Unit Interpretation Record

MUUF = Map Unit User Files

NASIS = National Soil Information System

NCSS = National Cooperative Soil Survey

NED = National Elevation Dataset

NLCD = National Land Cover Data

NPS = Nonpoint Sources

NRCS = Natural Resources Conservation Service

RV = Representative Values

SEA = SSURGO Extension for AVSWAT

SIR = Soil Interpretation Record

SSSD = State Soil Survey Database

SSURGO = Soil Survey Geographic

STATSGO = State Soil Geographic

SWAT = Soil and Water Assessment Tool

TMDL = Total Maximum Daily Load

UNBRW = Upper North Bosque River Watershed

WWW = World Wide Web

References Cited

Arnold, J.G., R. Srinivasan, R.S. Muttiah, and J.R. Williams. 1998. Large area hydrologic modeling and assessment part I: Model development. Journal of American Water Resources Association 34(1):73-89.

Arnold, J.G., K.W. King, R. Srinivasan, and M. Di Luzio. 2002. SWAT2000: Current capabilities and research opportunities in applied watershed modeling. American Society of Agricultural Engineers annual international meeting, Chicago, Illinois, July 28-31, 2002. Paper No. 022137.

Baumer, O. and J. Rice. 1988. Methods to predict soil input data for DRAINMOD. 1988 American Society of Agricultural Engineers meeting, December 13-16, 1988. Paper No. 882564.

Baumer, O., P. Kenyon, and J. Bettis. 1994. Map Unit User Files (MUUF) v2.14 user's manual. Lincoln, Nebraska: Natural Resources Conservation Service, National Soil Survey Center.

Beck, J.M., J.N. Shaw, P.L. Chaney, and J.E. Hairston. 2002. Image rectification for recompiling and digitizing soil survey maps. Journal of Soil and Water Conservation 57(2):95-100.

Bingner, R.L., J. Garbrecht, J.G. Arnold, and R. Srinivasan. 1997. Effect of watershed subdivision on simulation runoff and fine sediment yield. Transactions of the American Society of Agricultural Engineers 40(5):1329-1335.

Brown, D.G., L. Bian, and S.J. Walsh. 1993. Response of a distributed watershed erosion model to variations in input data aggregation levels. Computers and Geosciences 19(4):499-509.

Buland, D., E. Wang, B. Dayton, G. Lawson, and M. Di Luzio. 2000. Using SSURGO county-level soil data with SWAT watershed model. Presentation at the 2000 Soil and Water Conservation Society annual conference, St. Louis, Missouri, July 8-12, 2000.

Di Luzio, M., R. Srinivasan, and J.G. Arnold. 2004. A GIS-Hydrological Model System for the Watershed Control of Agricultural Nonpoint and Point Sources of Pollution. Transactions in GIS 8(1):113-136.

Eghball, B. 2002. Soil Properties as Influenced by Phosphorus- and Nitrogen-Based Manure and Compost Applications. Agronomy Journal 94:128-135.

ESRI.1996.What's New in ArcView GIS 3.0. Environmental Systems Research Institute. Redlands, California.

FitzHugh, T.W. and D.S. Mackay. 2000. Impacts of input parameter spatial aggregation on an agricultural nonpoint source pollution model. Journal of Hydrology 236:35-53.

FitzHugh, T.W. and D.S. Mackay. 2001. Impacts of subwatershed partitioning on modeled source- and transport-limited sediment yields in an agricultural nonpoint source pollution model. Journal of Soil and Water Conservation 56(2):137-143.

Gesch, D.B., M.J. Oimoen, S.K. Greenlee, C.A. Nelson, M. Steuck, and D.J. Tyler. 2002. The National Elevation Dataset. Photogrammetric Engineering and Remote Sensing 63(1):5-11.

Gregson K., D.J. Hector, and M. McGowan. 1987. A one-parameter model for the soil water characteristic. Journal of Soil Science 38:483-486.

Haverkamp, S., R. Srinivasan, H.G. Frede, and C. Santhi. 2002. Subwatershed spatial analysis tool: Discretization of a distributed hydrologic model by statistical criteria. Journal of the American Water Resources Association 38(6):1723-1733.

Juracek, K.E. and D.M. Wolock. 2002. Spatial and statistical differences between 1:250,000- and 1:24,000-scale digital soil databases. Journal of Soil and Water Conservation 57(2):89-94.

Kiesling, R.L., A.M.S. McFarland, and L.M. Hauck. 2001. Stream Community Responses to Eutrophication from Nonpoint Source Nutrient Loading. Proceedings of American Water Resources Association Spring Specialty Conference. Water quality monitoring and modeling, San Antonio, Texas, April 30- May 2, 2001.

Lathrop, R.G., J.D. Aber, and J.A. Bognar. 1995. Spatial variability of digital soil maps and its impact on regional ecosystem modeling. Ecological Modeling 82:1-10.

Lee, B.D., J.A. Wald, and L.J. Lund. 1999. Introducing students to online county soil surveys and the STATSGO database using GIS. Journal of Natural Resources and Life Sciences Education 28:93-96.

Lovejoy, S.B., J.G. Lee, and B.A. Engel. 2000. Managing watersheds: Improving the decisions with science and values. Journal of Soil and Water Conservation 55(4):434-436.

Lowrance, R., R. Leonard, and J. Sheridan. 1985. Managing riparian ecosystems to control nonpoint pollution. Journal of Soil and Water Conservation 40(1):87-91.

Mamillapalli, S., R. Srinivasan, J.G. Arnold, and B.A. Engel. 1996. Effect of spatial variability on basins scale modeling. Proceedings of the Third International Conference/Workshop Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, January 21-26, 1996. Available from: www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/main.html.

Mualem, Y. 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. Water Resources Research 12(3):522-524.

Muttiah, R.S. and R.A. Wurbs. 2002. Scale-dependent soil and climate variability effects on watershed water balance of the SWAT model. Journal of Hydrology 256:264-285.

Potter, S.R., D.C. Yoder, and J.C. King. 2000. A GIS analysis of the topography, soil type, and land use impacts on the water quality of a small suburban watershed. Oral presentation at the 2000 Soil and Water Conservation Society annual conference, St. Louis, Missouri, July 8-12, 2000. Also abstract in Journal of Soil and Water Conservation 55(3):394-395.

Reybold, W.U. and G.W. TeSelle. 1989. Soil geographic databases. Journal of Soil and Water Conservation 44(1):28-29.

Riley, R.C. 2000. Effectiveness of riparian buffers for nitrate removal. Presentation at the 2000 Soil and Water Conservation Society annual conference, St. Louis, Missouri, July 8-12, 2000.

Saleh, A., J.G. Arnold, P.W. Gassman, L.W. Hauck, W.D. Rosenthal, J.R. Williams, and A.M.S. McFarland. 2000. Application of SWAT for the upper north Bosque watershed. Transactions of the American Society of Agricultural Engineers 43(5):1077-1087.

U.S. Department of Agriculture (USDA). 1994. State soil geographic (STATSGO) Database: Data use information. Natural Resources Conservation Service Miscellaneous Publication 1492.

U.S. Department of Agriculture (USDA). 1995. Soil survey geographic (SSURGO) Database: Data Use information. Natural Resources Conservation Service Miscellaneous Publication 1527.

U.S. Department of Agriculture (USDA). 2000. About NASIS (National Soil Information System). Natural Resources Conservation Service. Available from http://nasis.nrcs.usda.gov/intro/. Accessed March 19, 2003.

U.S. Department of Agriculture (USDA). 2001a. SSURGO 2.0 Static Tabular Metadata. Natural Resources Conservation Service. Available from http://nasis.nrcs.usda.gov/documents/metadata/ssurgo2_0/. Accessed March 19, 2003.

U.S. Department of Agriculture (USDA). 2001b. NASIS 5.0 Tutorials. Natural Resources Conservation Service. Available from http://nasis.nrcs.usda.gov/documents/tutorial5/. Accessed March 19, 2003.

U.S. Department of Agriculture (USDA). 2001c. Soil Data Delivery And Distribution: Outline Physical Design. Natural Resources Conservation Service. Available from http://nasis.nrcs.usda.gov/documents/warehouse/WarehouseOPD.pdf. Accessed March 19, 2003.

U.S. Department of Agriculture (USDA). 2002a. National Soil Survey Handbook, title 430-VI. Natural Resources Conservation Service. Available from http://soils.usda.gov/procedures/handbook/main.htm. Accessed March 19, 2003.

U.S. Department of Agriculture (USDA). 2002b. Map Unit Interpretation Database (MUIR). Natural Resources Conservation Service. Available from http://ortho.ftw.nrcs.usda.gov/muir/factsheet.html. Accessed March 19, 2003.

U.S. Department of Agriculture (USDA). 2003. Status of soil survey digitizing (SSURGO). Natural Resources Conservation Service. Available from http://www.ftw.nrcs.usda.gov/jpg/ssurgo.jpg. Accessed December 17, 2003.

U.S. Department of Agriculture (USDA). 2004. National SSURGO Data. Natural Resources Conservation Service. Available from http://www.ncgc.nrcs.usda.gov/branch/ssb/products/ssurgo/data/index.html. Accessed March 22, 2004.

U.S. Environmental Protection Agency (USEPA). 1999. Protocol for developing nutrient TMDLs. Office of Water (4503F) Report EPA-841-B-99-007.

Vanden Heuvel, R.M. 1996. The promise of precision agriculture. Journal of Soil and Water Conservation 51(1):38-40.

van Genuchten, M.T. and D.R. Nielsen. 1985. On describing and predicting the hydraulic properties of soils. Annales Geophysical 3(5):615-628.

Vieux, B.E. and S. Needham. 1993. Nonpoint-pollution model sensitivity to grid-cell size. Journal of Water Resources Planning and Management 119(2):141-157.

Vogelmann, J.E., S.M. Howard, L. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel. 2001. Completion of the 1990s National Land Cover Dataset for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. Photogrammetric Engineering and Remote Sensing 67(6):650-662.

Williams, R.D. and L.R. Ahuja. 1992. Evaluation of similar-media scaling and one-parameter model for estimating the soil water characteristics. Journal of Soil Science 43(3):237-247.