Using GIS for Processing, Analysis and Visualization of Hydrological Model Data - 16202

Angelique Lawrence*, Mehrnoosh Mahmoudi*, Shimelis Setegn*, Natalia Duque* (DOE Fellow), Awmna Rana*(DOE Fellow), Brian Looney**

*Applied Research Center - FIU
**Savannah River National Laboratory

ABSTRACT

The Applied Research Center at Florida International University (ARC-FIU) is supporting the soil and groundwater remediation efforts of the U.S. Department of Energy’s (DOE’s) Savannah River Site (SRS) by developing a surface water model to simulate the hydrology and the fate and transport of contaminants and sediment in the Tims Branch watershed. Advances in ArcGIS software through the development of spatial GIS database (geodatabase) technology, coupled with the development of data models such as ArcHydro which possesses a spatial relational database management system (RDMS) schema and relationship structure specific to hydrologic systems, provides modelers with tools and applications to assist in the processing, analysis and visualization of flow and contaminant transport data. In addition, the coupling of this type of geodatabase structure with a numerical modeling package such as MIKE SHE/MIKE 11 can serve as an efficient tool that significantly reduces the time needed for data preparation. The MIKE modeling system has a geographic information systems (GIS) user interface built into its system that can directly use GIS data for model inputs. GIS-based hydrologic models can provide a spatial element that other hydrologic models lack. GIS is important for processing and analysis of large spatiotemporal datasets derived from multiple sources that are being used for setting up of models, model calibration and simulation of flow and contaminant fate and transport. GIS-based approaches in hydrological modeling can also provide the benefit of combining different layers of geographic data to create new integrated information which can be quite useful for creating dependent or independent hydrological variables.

ARC-FIU has developed an ArcGIS geodatabase to support the hydrological modeling work being conducted at SRS. The SRS geodatabase was developed based on the ArcHydro and ArcGIS Base Map data models with modifications made for project specific input parameters. The significance of this approach was to ensure its replicability for potential application at other DOE sites. The SRS geodatabase assists in the storage and management of GIS and timeseries data being used for hydrological model development at Savannah River Site. Other GIS tools such as ArcGIS ModelBuilder coupled with Python scripting have enabled the automation of many of the repetitive geoprocessing tasks required for pre- and post-processing of the hydrological modeling data. Automation of many of the data processing tasks has facilitated faster and hence more complex analyses of field test data. The toolbox created is also a scalable and reusable application that can be implemented in other watersheds.
This paper describes the use of ArcGIS tools for the pre-processing, spatial delineation and analysis of hydrological model data; for automation of repetitive geoprocessing tasks; and for visualization of different hydrological components of the Tims Branch watershed.

INTRODUCTION

The A/M area of the Savannah River Site (SRS) in South Carolina has been part of a long-term clean-up strategy to address legacy contamination resulting from processing activities during the nuclear arms race of the Second World War. An innovative remediation technology to reduce mercury contamination was implemented by the U.S. Department of Energy’s Office of Environmental Management, which involved the injection of tin (II) chloride into mercury-contaminated groundwater of the Tims Branch watershed (TBW). Although tin is primarily deposited in the sediment, scientists speculate that there is potential for remobilization during episodic extreme events, such as storms or heavy rainfall. Understanding the distribution and the fate and transport of tin and its compounds within the overland and river sub-domains is therefore of primary importance as re-suspended sediment containing tin can be transported downstream where environmental conditions may be favorable for tin methylation, which is an environmental concern due to its toxicity to humans and animals. A hydrological model is therefore being developed using MIKE SHE and MIKE 11 to simulate the fate and transport of mercury (Hg), tin (Sn) and sediments in the Tims Branch watershed. As the MIKE modeling system has a built-in GIS user interface which enables direct input of spatiotemporal datasets required for model calibration and performing simulations, GIS technology is being utilized to support model development.

Advances in ArcGIS software through the development of geodatabase technology, coupled with the development of data models such as ArcHydro which possesses a spatial relational database management system (RDMS) schema and relationship structure specific to hydrologic systems, provides modelers with tools and applications to assist in the processing, analysis and visualization of flow and contaminant transport data. In addition, the coupling of this type of geodatabase structure with a numerical model such as MIKE SHE/11 can serve as an efficient tool that significantly reduces the time needed for data preparation [11]. Gogu et al., 2001 [12] stress the benefits of putting large volumes of data into a structured, coherent and logical computer-supported system to ensure validity and availability for concurrent use by multiple users and provide a foundation for building GIS-based water resources applications. Additional benefits include the following:

- GIS-based hydrologic models can provide a spatial element that other hydrologic models lack.
- GIS enables hydrologists to pre-process and integrate data derived from multiple sources into a single manageable system.
- GIS-based approaches in hydrological modeling can provide the benefit of
combining different layers of geographic data to create new integrated information which can be quite useful for creating dependent or independent hydrological variables.

• GIS can be used for visualization of model-derived research results via maps, graphs and reports.

METHOD

Development of the Savannah River Site Geodatabase using ArcGIS

GIS technology has proven useful in supporting the Tims Branch modeling work being performed by FIU-ARC at Savannah River Site (SRS). The ArcGIS platform has been used as the basis for development of a geodatabase which is replicable for use at other DOE sites and provides an advanced spatial data structure for management, processing, and analysis of spatial and temporal numerical modeling data derived from multiple sources. ArcGIS tools such as ModelBuilder coupled with Python scripting has enabled the automation of many of the repetitive geoprocessing tasks required for pre- and post-processing of hydrological modeling data. Automation and batch processing of GIS data has facilitated faster model development, and the customized toolbox created is a scalable and reusable application that can be implemented for other geodatabases used to support modeling of surface water systems.

Development of the geodatabase has involved the following steps:

1. Import of an XML file generated from a pre-existing geodatabase that was created for the Oak Ridge Reservation in order to create a preconfigured database structure for the new SRS geodatabase.
2. Modification of the preconfigured SRS geodatabase based on specific model requirements.
3. Import of GIS and timeseries data.
4. Documentation of the geodatabase design using ArcGIS Diagrammer.

Import of ORR Geodatabase XML File

The SRS geodatabase was built upon a framework originally developed for modeling work being conducted by FIU-ARC researchers at the Oak Ridge Reservation (ORR). The ORR Geodatabase was created based on the ArcHydro and ArcGIS Base Map data models with modifications made for MIKE SHE/11 model-specific input parameters. The Arc Hydro data model is designed to support water resources applications within
the ArcGIS environment and possesses a structure that enables linkage with scalable hydrologic modeling tools and applications to model hydrologic systems [13].

An ArcGIS geodatabase is an XML-based GIS data exchange system which facilitates the export and import of preconfigured data as XML files which contain both the data definition and the data itself. The SRS geodatabase was therefore created by exporting the ORR geodatabase schema using ArcGIS tools as an XML file, and then importing the XML file into an empty file geodatabase to create the new preconfigured SRS geodatabase. Since the ORR database was primarily developed based on the ArcHydro data model, the new preconfigured SRS geodatabase possesses a spatial relational database management (RDMS) schema and relationship structure specific to hydrologic systems where spatial relationships between hydrological parameters and geographical features can be defined. The SRS geodatabase has a standardized data structure which helps in the organization of hydrologic features (e.g. channel cross sections, stream geometric networks and nodes, monitoring points, watersheds and subwatersheds, and other hydrographic and drainage files) and their relationships to each other, providing a common framework that can be utilized by various hydrologic models. The geodatabase structure also facilitates concurrent multi-user access, editing capability and management of spatial data within the ArcGIS framework and is comprised of a series of tables which contain feature, raster and attribute data, as well as metadata.

**Modification of the SRS Geodatabase Configuration**

Once the preconfigured SRS geodatabase was generated, modifications were made with respect to the spatial domain as well as the feature dataset and raster catalog names and properties to configure the SRS geodatabase specific to the study area. For example, the GIS data in the ORR geodatabase was stored with the following spatial reference properties: North American Datum (NAD) 83, State Plane Projection (Zone 5301), Units Meters; so this had to be converted in the SRS geodatabase to: North American Datum (NAD) 83, UTM (Zone 17N), Units Meters.

**Import of GIS and Timeseries Data**

Modeling of hydrologic systems requires large amounts of historical data for development and calibration and includes, for example, GIS coverages/shapefiles of the delineated watersheds, surrounding buildings and man-made structures which may serve as sources of contamination, roads, stream gauge locations, monitoring wells, bore holes, land cover and soils; raster imagery; and observed/measured timeseries data such as flow rates, precipitation, evapotranspiration, contaminant concentration and surface and groundwater levels. Figure 1 below shows the SRS geodatabase system workflow design which depicts the various geospatial and timeseries data types stored in the geodatabase and the implementation of GIS tools for data geoprocessing to convert the files to compatible formats that can be used in the hydrological model.
The MIKE SHE/11 model uses GIS data inputs for many of its configuration parameters which contain spatial features within the model domain, such as points representing monitoring stations, lines representing rivers/stream networks, or polygons which outline areas such as the watershed and catchments (Figure 2).
Figure 2. MIKE-SHE model spatiotemporal (GIS) data inputs.

The significance of using GIS data is not just the spatial representation of hydrologic features, but their association with timeseries data attributes such as flow rates and directions, contaminant concentrations, water levels, precipitation, etc. Availability of data in this format shortens the time for data preparation and ultimately model development.

The majority of GIS data for the hydrological modeling work being conducted by FIU-ARC at SRS was provided by the Savannah River Nuclear Solutions (SRNS) Geotechnical Engineering Department at SRS in the form of an ArcGIS 10 map package. The supporting metadata for many of these files were provided in the form of XML files. ArcMap 10.2 was used to view the GIS data provided, which was stored in several small geodatabases. The ArcToolbox import utility was then used to consolidate the GIS data into the single SRS geodatabase created. The XML metadata files were then appended to their associated GIS feature classes using the metadata editor within ArcCatalog.

The accuracy and predictive forecasting ability of hydrological models largely depend on the availability of timeseries data (daily/monthly/annual) as well as the period of time this data covers. Some of the SRS site monitoring data sources used in this project include the U.S. Geological Survey (USGS) land cover and hydrography databases, the Natural Resources Conservation Service (NRCS) STATSGO or SSURGO soil databases, and the U.S. EPA’s MultiResolution Land Characteristics (MRLC)
Consortium or North American Land Cover Institute databases. Timeseries data (daily rainfall, stream flow and temperature) as well as several reports and journal publications from which several hydrological model parameters were derived, were also provided by Savannah River National Laboratory (SRNL). FIU-ARC researchers also conducted an extensive literature review in order to characterize the study area and retrieve additional data required for model development and calibration. The data derived from these multiple sources was also imported into the SRS geodatabase.

**Documentation of the SRS Geodatabase Design Using ArcGIS Diagrammer**

Documenting the geodatabase design can assist in representation of the map layers, metadata and other elements specific to the data model used to create the geodatabase. ArcGIS Diagrammer for ArcGIS 10.2 is a downloadable diagramming utility used to create, edit or analyze geodatabase schema. It generates diagrams and reports in the form of editable graphics within an interface similar to Microsoft Visual Studio and serves as a visual editor which accepts XML workspace documents that are created from ESRI’s ArcMap or ArcCatalog.
Geoprocessing of Model Data using ArcGIS

Development of hydrological models requires data that may include thousands of groundwater monitoring wells, boreholes, stream reaches with gauges, weather stations, land cover, vegetation, soil type, topography, geology, water quality and satellite imagery. The application of GIS tools enables hydrologists to pre-process and integrate this data that was likely derived from multiple sources, have different spatial references, are at different scales, and are from different time periods, into a single manageable system. As previously mentioned, the MIKE SHE model uses an extensive amount of GIS data inputs (Figure 2) for many of its configuration parameters. ESRI’s ArcGIS 10.2 geoprocessing tools were used for pre-processing the hydrological modeling data.
Development of Process Flow Models using ArcGIS ModelBuilder

Simple tasks such as retrieving data from the SRS geodatabase; geoprocessing the data and exporting for use in the hydrological model; subsequent import and post-processing of model results; data analysis; and data visualization via graphs, charts and maps can be very repetitive and time consuming. ArcGIS ModelBuilder was therefore used to develop a reusable tool which can iterate over the set of GIS-based model parameters, perform geoprocessing actions, calculate statistical outputs and display results in the form of maps and graphs. The use of ArcGIS ModelBuilder assists in automating these tasks which saves time and can facilitate batch processing of this data. Customization of the geoprocessing tools is also possible using Python scripts if there are hydrological model-specific requirements.

ArcGIS ModelBuilder has built-in ArcGIS tools that were used to automate several repetitive model-specific geoprocessing tasks, for example, clipping all of the GIS feature classes to the study domain being used in the MIKE SHE model and then projecting them to UTM coordinates. ModelBuilder also generates workflow diagrams (Figure 4) which document and visually represent tools and scripts (if any) that have been incorporated in the process flow model.

![Figure 4. ArcGIS ModelBuilder workflow diagram for clipping GIS data to the study domain.](image)

Hydrological models developed in the past by FIU-ARC researchers at DOE sites have involved the development of customized Python scripts to conduct specific geoprocessing tasks that were not already integrated into the ArcGIS system. This required additional programming of the built-in automated geoprocessing tools. A toolbox was created for the specified study area but was developed as a scalable and reusable application to enable its implementation at other DOE sites. This toolbox will be used to conduct several of the geoprocessing tasks required during development of the Tims Branch hydrological model.
**Development of Model-Specific Input Files**

The MIKE SHE model requires a number of spatial and temporal parameters such as topography, vegetation, land use, precipitation, soil composition, geologic formation, etc. The built-in GIS user interface of the MIKE modeling system enables direct input of these spatiotemporal datasets in the form of GIS shapefiles for model calibration and performing simulations, however, there are instances where GIS input file modification is necessary either for use of the data at smaller scales or to modify appended timeseries or attribute data to generate compatible MIKE SHE/11 files. Although MIKE SHE directly accepts GIS shapefiles, at times the attribute field with the relevant data required is not read by the model due to an incompatible field type. For example, a non-integer numeric field may be “single”, however, the field type accepted by the model is “double”. As a result, modification of the attribute table is necessary to create a new field with “double” as the field type into which the required numeric data can be copied.

**Model Domain**

For this project, the Tims Branch watershed boundary was used as the model domain, defining the areal extent of surface water drainage to a point. The data was downloaded from the USGS Watershed Boundary Dataset (WBD) in the form of a GIS coverage. This digital vector dataset was determined solely upon science-based hydrologic principles and is used by geographic information systems for general mapping and in the analysis of surface water systems. Hydrologic units (HU) establish a baseline drainage boundary framework, accounting for all land and surface areas. The Tims Branch 12-digit Hydrologic Unit Code (HUC) is 030601060504. The watershed boundary was delineated and georeferenced to the USGS 1:24,000 scale topographic base map meeting National Map Accuracy Standards (NMAS). The downloaded GIS polygon was added to the MIKE SHE model to define the model domain as viewed in Figure 5.

![Figure 5](image.png)

Figure 5. The Tims Branch watershed boundary which defines the model domain as viewed in the MIKE SHE model.
**Tims Branch Stream Network**

The Tims Branch stream network was derived from the USGS National Hydrography Dataset (NHD), which is a GIS-based digital vector dataset that represents drainage network features such as rivers, streams, canals, lakes, ponds, coastline, dams, and streamgages. Like the Watershed Boundary Dataset, the NHD was delineated and georeferenced to the USGS 1:24,000 scale topographic base map meeting National Map Accuracy Standards (NMAS). The NHD files contain flow network attributes imbedded in the data that allow GIS scientists to trace flow directions and facilitate delineation of features such as nodes, cross sections and cross section profiles required for developing the stream flow model using MIKE 11.

![Figure 6. An ArcMap view of the Tims Branch delineated cross sections (left and center); the cross section profile of the cross section #PG9 (right).](image)

The NHD and WBD can be used by a GIS with other data themes such as elevation, boundaries, transportation, and control structures to produce general reference maps, however, because they both use a linked addressing system based on reach codes and other basic NHD and WHD features, in-depth geospatial analyses are possible to study cause and effect relationships, such as how a source of poor water quality upstream might affect a fish population downstream.
Topography

In MIKE SHE, topography defines the upper boundary of the model. The topography is used as top elevation in both the Unsaturated Zone (UZ) and Saturated Zone (SZ) modules. It also defines the drainage surface of Overland (OL) flow. For this reason, topography is one of the most important parameters in the MIKE SHE model setup and the spatial resolution of the file from which it is derived is very significant. The MIKE SHE model input for topography was generated by using ArcGIS tools to convert a 10 ft (~3m) resolution digital elevation model (DEM) (Figure 7) to a point shapefile which contained XY coordinate data with associated elevation values. The GIS shapefile was then imported into the MIKE SHE model which interpolates the point data via a triangular interpolation method into a gridded surface. This was then exported as a .dfs2 file, which is a native MIKE SHE grid file format. The .dfs2 file was then used to replace the point shapefile in the model. The USGS DEM format used was generated from 7.5 minute DLG hypsography data and was downloaded from the South Carolina Department of Natural Resources (SCDNR) GIS Data Clearinghouse.

Figure 7. Digital elevation model (DEM) of Tims Branch watershed.

Land Use/Land Cover

Gridded land use/land cover data within Aiken County, SC was downloaded from the US DOI/USGS National Land Cover Database (NLCD) for the years 1992, 2001, 2006 and 2011. The grid files were all converted to GIS polygon shapefiles and clipped to the model domain so that they can be incorporated in the MIKE SHE model. ArcMap was also used to create thematic maps of the downloaded land cover data which depict the land cover types and their distribution within the Tims Branch watershed. Significant changes observed include (1) the introduction of high intensity developed land in 2011; (2) a large increase in shrubs between 1992 and 2011; (3) the
disappearance of transitional vegetation between 1992 and 2011; and (4) a decrease in cultivated crops between 1992 and 2011.

Figure 8. Geospatial distribution of land cover existing in 1992 and 2011.

The land use/land cover data was also used to create new shapefiles to represent the Manning’s roughness and the paved runoff coefficients. Table 1 shows the values of Manning’s M that were assigned to each land use classification in the land cover shapefile previously described.

Table 1. Land Use Classifications and Corresponding Manning’s M Number Assigned

<table>
<thead>
<tr>
<th>Land Use</th>
<th>Area (m$^2$)</th>
<th>%</th>
<th>Manning’s M (1/n) Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural</td>
<td>170,975</td>
<td>0.34</td>
<td>41</td>
</tr>
<tr>
<td>Barren Land</td>
<td>58,151</td>
<td>0.12</td>
<td>81</td>
</tr>
<tr>
<td>Forest</td>
<td>35,267,379</td>
<td>70.83</td>
<td>21</td>
</tr>
<tr>
<td>Rangeland</td>
<td>7,287,896</td>
<td>14.64</td>
<td>25</td>
</tr>
<tr>
<td>Urban/Built-up Land</td>
<td>6,816,222</td>
<td>13.69</td>
<td>90</td>
</tr>
<tr>
<td>Water</td>
<td>76,866</td>
<td>0.15</td>
<td>11</td>
</tr>
<tr>
<td>Wetland</td>
<td>115,658</td>
<td>0.23</td>
<td>23</td>
</tr>
</tbody>
</table>
Manning’s $n$ values were obtained from standard civil engineering Manning’s tables available online as well as $n$ values derived from the technical report by Tachiev et al, 2014, “Remediation and Treatment Technology Development and Support for DOE Oak Ridge Office: EFPC Model Update, Calibration and Uncertainty Analysis”. The land cover shapefile attribute table was then modified to include a new field of Manning’s M (i.e., $1/n$) numbers. This added field was then used as the basis for generating a new polygon shapefile to represent the Manning’s Roughness Coefficients within the SRS/Tims Branch study area. As the MIKE SHE model only accepts point/line shapefiles for spatially distributed Manning’s M, ArcGIS tools were used to convert the polygon shapefile to a point shapefile which was then input into the model. The model then interpolated the values to generate a gridded surface which was saved as a MIKE (.dfs2) grid file. This grid file was then used to replace the shapefile in the model configuration.

Paved runoff coefficient values were derived from the Clean Water Team Guidance Compendium for Watershed Monitoring and Assessment State Water Resources Control Board 5.1.3 FS-(RC) 2011, which specifies the runoff coefficient (C) as a dimensionless coefficient relating the amount of runoff to the amount of precipitation, with larger values for areas with low infiltration and high runoff (pavement, steep gradient), and lower values for permeable, well vegetated areas (forest, flat land). This data is required by the MIKE SHE model and can be a significant parameter indicating flooding areas during storm events as water moves fast overland on its way to a river channel or a valley floor. Paved runoff coefficient values were assigned to the land use classifications outlined in Table 1 above. A value of 0.7 was given to the Urban/Built-up Land and a value of zero assigned to all other land use types. In the same manner as described above for development of the Manning’s Coefficient GIS shapefile, the land cover shapefile attribute table was modified to include a new field of runoff coefficients. This added field was then used as the basis for generating a new polygon shapefile to represent the Paved Runoff Coefficients within the SRS/Tims Branch study area.

Development of Maps & Reports

GIS can also serve as a useful tool in visually displaying research data and modeling results via maps, graphs and reports which help to enhance the understanding and interpretation of model-derived data and to obtain a perception closer to reality. Several maps of the Tims Branch watershed were created using the ArcGIS ArcMap interface as seen in Figure 9.

---

1 Source: Clean Water Team Guidance Compendium for Watershed Monitoring and Assessment State Water Resources Control Board 5.1.3 FS-(RC) 2011 is a factsheet prepared by the California Environmental Protection Agency State Water Resources Control Board that can be accessed online at the following URL: http://www.waterboards.ca.gov/water_issues/programs/swamp/docs/cwt/guidance/513.pdf.
CONCLUSION

The application of GIS technology has proven to be a significant benefit in supporting and speeding up the development of the Tims Branch hydrological model. FIU-ARC researchers will continue to apply GIS tools for pre- and post-processing as well as visualization of hydrological model data for the Tims Branch model as the project continues into the next phase which involves simulation of the stream flow and contaminant transport. In addition, features such as land cover or land use that may change over time, together with other significant timeseries data such as rainfall and stream flow/discharge, may be contributing factors or indicators of hydrological changes in the watershed, so geospatial analyses of these parameters will be conducted to provide additional information that may support the hydrological modeling results.

REFERENCES


ACKNOWLEDGMENTS

Funding for this work was provided by the U.S. Department of Energy Office of Environmental Management under grant DE-EM0000598.