Proceedings of Oklahoma Water 2005, Tulsa, OK, September 27 and 28, Paper # 11
Oklahoma Water Resources Research Institute, Stillwater, OK, 10 pgs.

STEWARDS: An Integrated Data System for ARS Watershed Research

Jean L. Steiner[1], Jin-Song Chen[1], and John Sadler[2]

Abstract

To provide support for Conservation Effects Assessment Project (CEAP), the Agricultural Research Service is implementing a web-based data system – STEWARDS (Sustaining the Earth's Watersheds – Agricultural Research Data System). The data system consists of a central database management system and a client application. The climate, water, soil, management, and economic data are compiled and annually updated into STEWARDS from the twelve watershed locations. The visualization and tabular interface tools allow the users to view and downloaded the data. The applications of the system in facilitating the CEAP watershed modeling and the benefits to increase the utility of data are discussed.

Introduction

Fast and platform-independent access over the Internet and the World Wide Web to large and heterogeneous scientific environmental data has prompted the need for data management systems that allow efficient data storage, data compression, interactive data browsing, data manipulation, data visualization, searching, and downloading. To support related needs such as environmental monitoring/control and decision-making/analysis, integration of modeling or other analytical tools into the data management system is desirable. As a result, advanced multi-functional environmental web-based data systems have been and continue to be developed. For example, the USGS' NWISWeb data system (http://waterdata.usgs.gov/nwis) provides access to water-resources data collected in the US. The ecological research community, led by the Long Term Ecological Research (LTER) Network, has established a network-wide information system to facilitate data exchange and integration to meet the needs for long-term ecological research (http://www.lternet.edu/informatics). The U.S. Army Corps of Engineers incorporated acquisition and analysis in the Corps Water Management System (http://www.hec.usace.army.mil/cwms/) to support their real-time water control mission (Fritz et al., 2002). In Europe, a hydrological, meteorological, and environmental data management system called WISKI (http://www.kisters.de/english) has been widely used for decision support and the management of resources or events. WISKI consists of three parts -- a data acquisition server, a central database, and a windows-based hydrological workbench (Natschke et al., 2003). Data retrieval and data analysis in WISKI are

[1] Grazinglands Research Laboratory, USDA-ARS, El Reno, OK
[2] Cropping Systems and Water Quality Research Unit, USDA-ARS, Columbia, MO

provided using powerful graphical and tabular interface tools, and internal or external statistical and modeling packages.

The USDA and the Agricultural Research Service (ARS) have conducted watershed research since early in the 20th century with additional sites added periodically to meet evolving research needs. Many of these research watersheds are still operational. However the ARS watershed data have been managed to address location-specific research needs and are disseminated independently at each research location, greatly reducing the accessibility and utility of these data for policy-relevant, multi-site analyses.

The recently established Conservation Effects Assessment Project (CEAP) was designed to quantify effects of USDA conservation programs and practices (Mausbach and Dedrick, 2004). Within CEAP, watershed assessment studies are conducted on twelve ARS benchmark watersheds at various scales that include measurements and hydrologic simulations. These studies require a variety of data that describe hydrology, soils, climate, topography, management practices, and land use. This is also true for other national programs conducting research at watershed scales, such as EPA's Total Maximum Daily Load (TMDL) program (Turner and Boner, 2004), and NOAA/USGS/USDA's climate and global change program (Cruise et al., 1999). Therefore, a centralized ARS watershed data system is needed to make these data easily accessible and comparable across multiple locations. Such a system would not only be a valuable resource to CEAP, but also to a wide array of stakeholders who are involved in watershed studies across the nation.

Objective

Our objective is to develop and implement a web-based data system to organize, document, manipulate, and compile climate, water, soil, land-management, and socio-economic data from ARS research watersheds for assessment of conservation practices and other hydrologic analyses. The system will be called STEWARDS, Sustaining the Earth's Watersheds, Agricultural Research Data System.

Approach

The implementation of the data system consists of seven tasks (Fig. 1). Some of the tasks must be completed in sequence while others can proceed simultaneously.

The first task, System Requirements, lays out the architecture of the system which will consist of three main parts:

- Local Sites where the data managers will provide data, URLs, and metadata to the central site;
- Central Site which will host a central database management system for the receiving, storage, and management of data; and
- Clients where the users will access and interact with the central data system (Fig. 2).

The central data management system will also provide visualization and analytical tools to allow the users to view graphs, statistical report tables, and data summaries. A metadata search engine will be developed so that the users can search and determine if the dataset meets their needs based on the available metadata information.

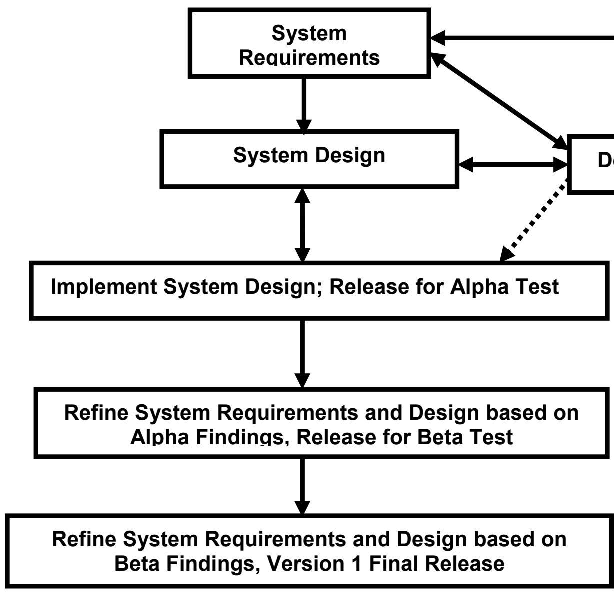

Figure 1. Task flow identified in the ARS STEWARDS data system.

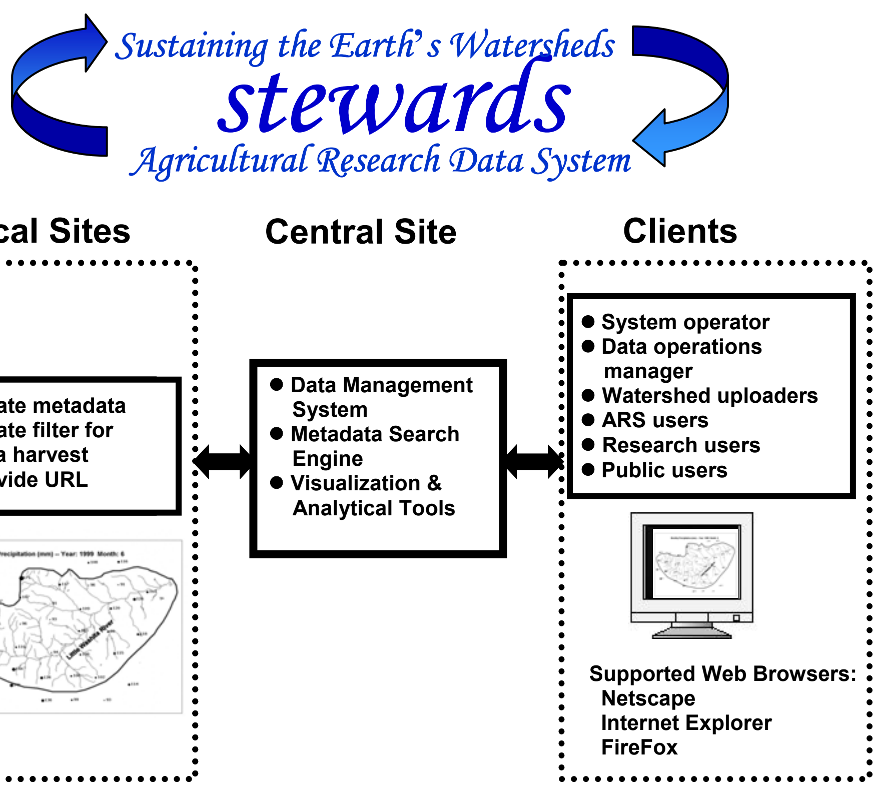

Figure 2. Architecture of the ARS STEWARDS data system.

System Requirements. Following a systems engineering methodology, the system requirements have been defined based on the results from reviewing system requirements of other data management systems (Steiner et al., 2003); interviews with potential data system users; an inventory of data compiled from ARS watershed research programs; an inventory of data measurements and methodologies currently used in CEAP research; and input and output requirements for the Soil Water Assessment Tool (SWAT) and Annualized Agricultural Non-Point Source (AnnAGNPS) models to be used in CEAP. The system requirements guide the definitions of the system architecture and features, user classes, operating environment, design and implementation constraints, and user documentation. The system requirements have been completed within ARS and are currently in an external review process.

Metadata Requirements. In essence, metadata answer who, what, when, where, why, and how about every facet of the data and are an essential component of any scientific database. "Metadata" standards will be applied to allow users to understand the nature and quality of the data. Each dataset is required to have at least one metadata report.

Metadata sets will conform to the requirements of the Federal Geographic Data Committee (http://www.fgdc.gov/metadata/csdgm/) and may incorporate aspects of the National Biological Information Infrastructure's (NBII) Biological Data Profile (BDP) (http://www.nbii.gov/datainfo/metadata/standards/index.html). A metadata template, which is a file that documents the required data elements, will be developed by the end of 2005. The template will assist the data managers at each watershed site as they create metadata. The metadata creation tool will be selected from the existing metadata tools such as M^{3}Cat (http://www.intelec.ca/technologie_a.html) and Tkme (http://geology.usgs.gov/tools/metadata).

System Design. The goal of the system design is to generate a model of the data management system. The tasks underway are to identify graphical user interface (GUI) concepts and functionality, develop GUI layouts and usage storyboards, and develop system architecture diagrams and analysis. The system design will be refined according to prototype data system findings (described below).

Prototype Data System. The prototype data system has been developed in parallel with the overall system development. The goal of the prototype development is to explore options and answer key technical issues concerning the system architecture and minimize the potential for surprises during system development. The prototype findings will assist the system implementation in the later phase. The prototype is available for demonstration purposes but not for release. The prototype may be discarded as the operational system is developed or it may be transformed to become the "real" data system.

Data System (Alpha). In the alpha design phase, the software and contents of data system will be developed. The resulting system will be completely functional, but stability testing and distribution will be limited. In-house alpha testing will be performed by the system development team. Data from priority watershed sites (Walnut Creek, IA; Goodwin Creek, MS; Goodwater Creek, MO; Little River, GA; and Town Brook, NY) will be sequentially uploaded into the data system during the alpha development and testing.

Refined Data System (Beta). During the beta phase, the system design will be modified, as needed, based on alpha findings. After extensive in-house testing, the beta system will be released and distributed to beta testers within CEAP research and modeling teams.

Final Release of Version 1. After implementing and testing required fixes based on beta findings, the system will be fully implemented through the introduction of additional site data. After verifying the master copy for final release, STEWARDS will be made available to other users (initially other CEAP researchers and later public access). An on-line user's manual and technical documents will also be available with this release. Training will be provided to user groups. The data system will be publicized though presentations at conferences, publication in journals and establishing links to related database sites.

A Prototype Data System in El Reno, OK

A web-based prototype data system has been developed in El Reno, Oklahoma, using data from the Fort Cobb Reservoir Experimental Watershed (FCREW) and Little Washita River Experimental Watershed (LWREW, Fig. 3). Primary implementation and product application of this prototype are presented here to demonstrate the expected product and the usage of the ARS data system, STEWARDS.

Architecture. The prototype data system consists of localized watershed data, a central database management system and client application user interfaces (Fig. 2). The configuration includes:

- Microsoft Windows 2003;
- Microsoft Internet Information Services (IIS); and
- Microsoft SQL (Structured Query Language) 2000 server.

ASP.NET was used to develop the user interfaces. ASP.NET is a web development technology of Microsoft's .NET development platform. ASP.NET/MSSQL was used because this allows us to keep the system up to date with changes in the software industry, to increase speed, and to provide flexibility. The visualization tool is primarily driven by an open-source ZedGraph graphics library (http://zedgraph.sourceforge.net/). ZedGraph is a flexible charting library for .NET to create 2D line and bar graphs. It can handle multiple graphic objects on a page and allows an overlay of objects on a background images to facilitate spatial display.

Functionality.

- Web user interfaces (Fig. 4) provide interaction between users and the data system. Users can select multiple sites, time windows, data aggregations (daily, monthly, and annual), parameter types, and plot types (line chart, bar chart, line with symbol, and spatial display of measurements).
- Current tools only provide meteorological input data files for the SWAT model.

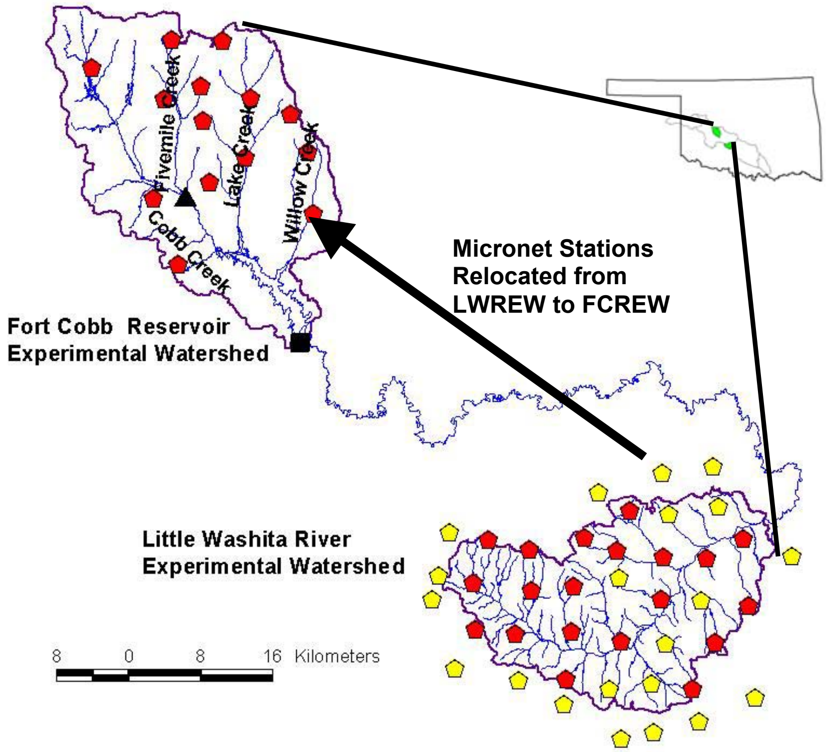

Figure 3. Research watersheds implemented in the prototype data system, showing configuration of the Micronet sites before and after relocation. (The triangle represents the USGS gage station near Eakly, OK, and the rectangle represents the OK Mesonet station at Fort Cobb Reservoir Experimental Watershed, used in Example 1 given below).

- Visualization tools – time trends, line/bar charts of multiple sites and up to two variables can be displayed for cross-site, cross-variable comparison, with outliers/abnormality identification. Multiple windows for different graphic objects can be opened so displaying various types of graph objects becomes flexible. The feature of multiple objects on one page will be implemented in the near future.
- Spatial display: Spatial distribution of climate and water quality parameters can be overlaid on a background map to facilitate examining spatial trends and identifying "hot spots".
- A utility to generate statistics reports will be developed.

Data Uploading/Updating. The ARS Micronet meteorological data (http://grl.ars.usda.gov/micronet/) for LWREW, Oklahoma Mesonet meteorological data for FCREW, USGS streamflow data for LWREW and FCREW, and water quality data for FCREW are uploaded and continually updated using a data filter and the Microsoft SQL system administration tools.

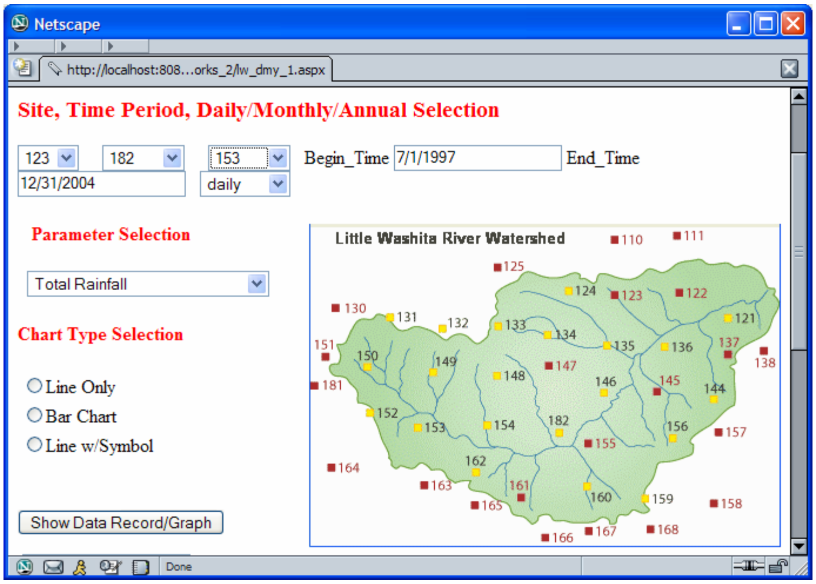

Figure 4. The user interface of the ARS prototype data system provides a variety of selection options.

Examples of Use of the Prototype Data System

Example 1. Time series of water quality (conductivity) of Fivemile Creek, Lake Creek, Willow Creek and Cobb Creek from FCREW (Fig. 5a) indicates that magnitudes of conductivity at Cobb Creek are significantly higher than the three other stream tributaries. In Fig. 5a, the sharp drops of conductivity observed on Jun 16, 2005 were associated with a peak in the discharge rate at the USGS gage station at Cobb Creek near Eakly, OK (Fig. 5b). This peak flow event was due to a storm occurred on June 10, 11, and 12, 2005 (Fig. 5c). This example depicts that the data system allows users to see data clearly and provides a basis for exploring cause-effect relationship among the water quality, hydrological, and meteorological data.

Example 2. To establish the watershed study at FCREW, a relocation of climate stations from LWREW was necessary. Twenty one Micronet climate stations were removed from LWREW and fifteen were relocated at FCREW (Fig. 3).

The question: *Does reduction in climate data density induce statistically dissimilar SWAT output?*

To answer this question, three scenarios were examined:

Scenario #1: original Micronet configuration of 41 stations in LWREW;
Scenario #2: 21 stations removed; and
Scenario #3: 38 stations removed (only 3 stations in entire watershed).

The data system prepares a meteorological input file at each station for SWAT model runs for each scenario. SWAT-simulated streamflow were examined at daily, monthly and annual time scales. The results indicated 1) no significant differences for simulated streamflow (Fig. 6) were observed between Scenario #1 and #2 (not shown); 2) slight visual differences were observed in simulated streamflow between Scenario #1 (or #2) and #3 (Fig. 6). It was concluded that the reduction in climate data density due to the relocation of stations does not induce statistically dissimilar SWAT output. Although the number of meteorological stations used in the SWAT simulation is not necessarily the same as what the user provided, the purpose of this example is to depict the impact of the number of weather stations available (not the number of stations used by SWAT) on SWAT simulation results. This example illustrates the usefulness of the data system for SWAT model support and how SWAT modeling with the data system support could be used in decision-making about reconfiguration of the Micronet stations.

Summary

The implementation of the ARS data system STEWARDS is underway and will provide an efficient way to access, view, download, and visualize measured data and model simulations to diverse internal and external users. A prototype data system developed in El Reno, OK, provides flexibility to the user to access watershed data for two experimental watersheds in Oklahoma. Visualization tools enable users to see the data clearly and perform preliminary data summarization and analysis. STEWARDS, when completed, will facilitate watershed assessments to support the Conservation Effects Assessment Program (CEAP) and support research to improve natural resource decision-making. STEWARDS will increase the impact of ARS watershed research by increasing the efficiency of researchers working within and across locations with diverse collaborators and will make the publicly funded research data available to a wide array of partners and stakeholders.

References

Cruise, J.F., A.S. Limaye, and N. Al-Abed. 1999. Assessment of impacts of climate change on water quality in the southeastern United States. Journal of the American Water Resources Association, 35(6):1539-1550.

Fritz, J.A., W.J. Charley, D.W. Davis, and J.W. Haines. 2002. New water management system begins operation at US projects. Hydropower & Dams, Issue 3, 2002. p 49-53.

Mausbach, M.J., and A.R. Dedrick. 2004. The length we go: Measuring environmental benefits of conservation practices. *J. Soil Water Conserv*. 59(5):96A-103A.

Natschke, M., R. Funke, and A. Matamala. 2003. The hydrological workbench WISKI: A modern tool for integrated hydrological data management from acquisition to quality assurance, presentation and dissemination. Progress in Water Resources, v 8, Water Resources Management III, p 377-385.

Steiner, J.L., D.C. Goodrich, S. Hardegree, M.R. Burkhart, T.C. Strickland, and M.A. Weltz. 2003. Information Technology Applications in the ARS Watershed Network. In Renard, K. G. eds. First Interagency Conference on Research in the Watersheds, October 27-30, 2003. U.S. Department of Agriculture, Agricultural Research Service.

Turner, B.G., and M.C. Boner. 2004. Watershed monitoring and modeling and USA regulatory compliance. Water Science and Technology, 50(11):7-12.

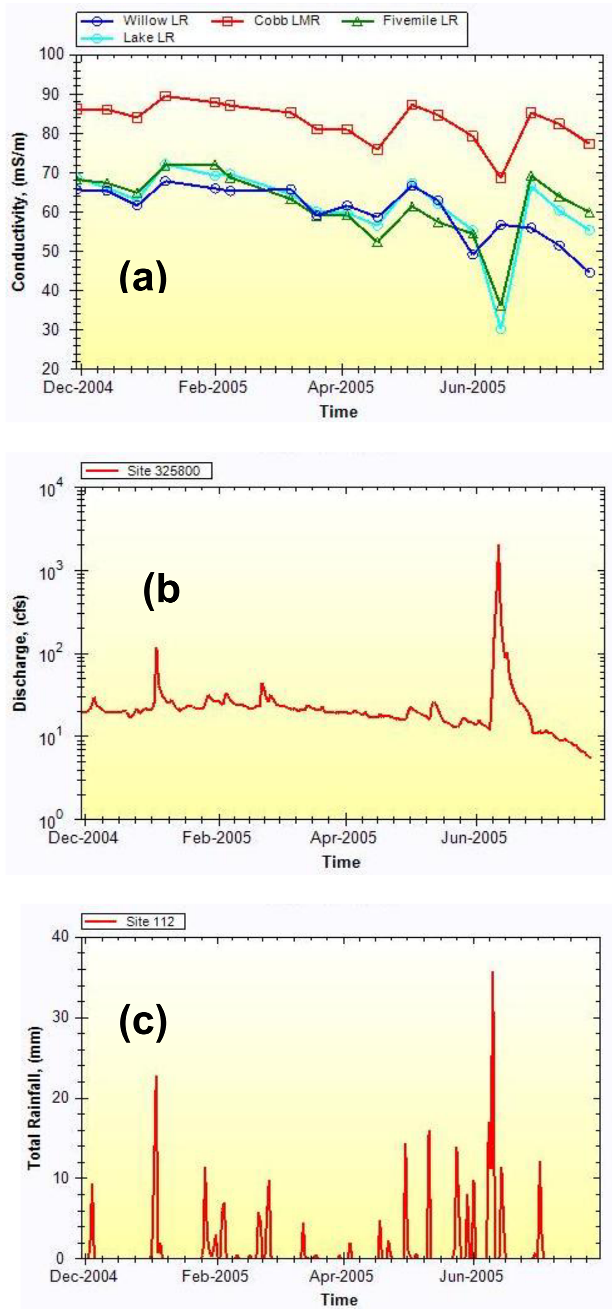

Figure 5. Time series of a) conductivity for four streams in the Fort Cobb Reservoir Experimental Watershed, b) stream discharge on Cobb Creek, and c) precipitation at Fort Cobb Reservoir Experimental Watershed.

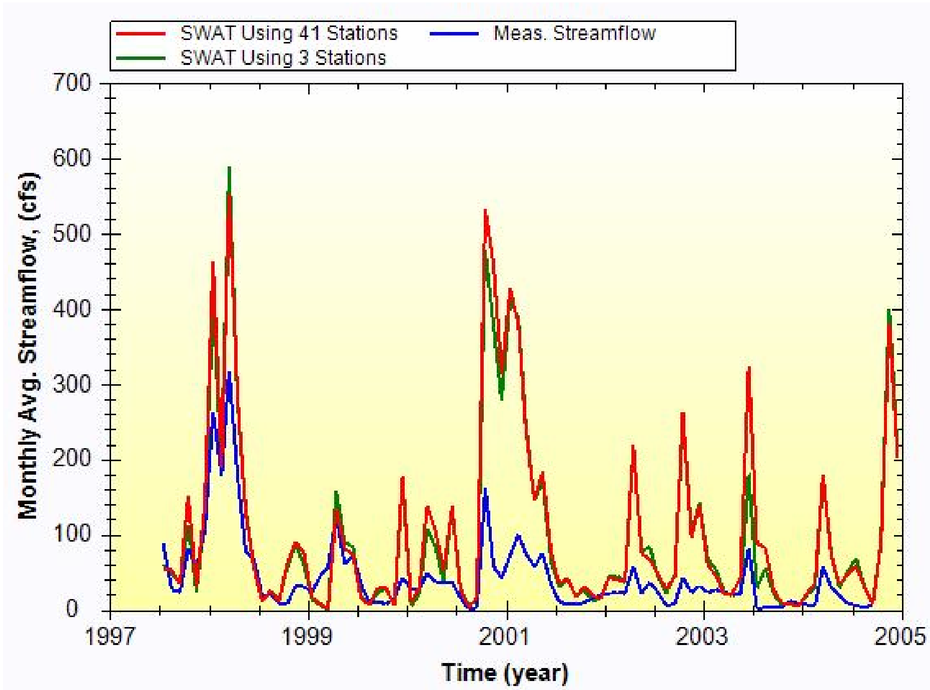

Figure 6. Comparison of simulated monthly streamflow for Scenario #1, 41 rain gages, and #3, 3 rain gages, to measured streamflow. (Results for Scenario #2, 21 rain gages, were not distinguishable from those of scenario #1).