CALIBRATION AND VALIDATION OF THE SOIL AND WATER ASSESSMENT TOOL ON THE CANNONSVILLE RESERVOIR WATERSHED

By Jennifer Benaman[1], Associate Member, ASCE, Christine A. Shoemaker[2], Fellow, ASCE, and Douglas A. Haith[3], Member, ASCE

In press for *ASCE Jn. Of Hydrologic Engineering*

ABSTRACT

Hydrologic and sediment loading calibration and validation were completed for a New York City water supply watershed, the Cannonsville Reservoir basin, located in upstate New York. An unusually large amount of data exists for this watershed, with over 20,000 data points consisting of climate, flow, and sediment measurements for calibration and validation. Hydrology results using the Soil and Water Assessment Tool (SWAT) displayed a Nash-Sutcliffe measure (NS) of 0.74 at the main monitoring station which drains 80% of the 1200 km^2 basin. Sediment results at this station came within 1% of average loading estimated from data for calibration. Two of the three smaller drainage areas all displayed calibrated hydrology NS values above 0.7. Changes in calibrated parameter values reflected northeastern conditions related to snowmelt events and fragipan soils. Validation of the hydrology and sediment results was based on an independent data set with no changes in parameter values. These validation results were for hydrology showed a NS = 0.76 and an 8% difference in average flow. Highlights on the limitations of SWAT are given, including the model's approach to snowmelt, sediment erosion, and sediment transport.

INTRODUCTION

The current regulatory trend of Total Maximum Daily Loads (TMDL) has moved the focus of water quality management from 'end of the pipe' or point source control to watershed scale analyses which incorporate point and non-point source pollution assessments. As a result, recent literature indicates that

[1] PhD Candidate, School of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14853; PH; Currently, Managing Engineer, Quantitative Environmental Analysis, 800 Brazos Street, Suite 1040, Austin, TX 78704 PH (512) 707 0090; FAX (512) 275 0915; email: jbenaman@qeallc.com

[2] Ripley Professor of Engineering, School of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14853; PH (607) 255-9233; FAX (607) 255-9004; email: cas12@cornell.edu

[3] Professor, Department of Biological and Environmental Engineering, Cornell University, Ithaca, NY 14853; PH (607) 255-2802; FAX (607) 255-4080; email: dah13@cornell.edu

modeling efforts have shifted to large-scale watershed simulations (e.g. greater than 1000 km^2) in an attempt to quantify the impact of all potential sources (Line et al. 1997). In addition, the use of distributed watershed models to evaluate the impact of a series of small subwatersheds on one water body has become important in watershed management and non-point source control. Initial constituents of concern for TMDLs are centered on suspended sediment, biochemical oxygen demand and nutrients.

For New York City (NYC), the ramifications of TMDL development and water quality standards affect areas beyond the city limits. NYC maintains and monitors 19 reservoirs and 3 controlled lakes in upstate New York, which provide drinking water to over 9 million NYC residents. In 1997, the United States Environmental Protection Agency (EPA) issued a Filtration Avoidance Determination (FAD), allowing the metropolitan area to not filter drinking water which originates from the upstate reservoirs (USEPA 1997). The review of the 1997 FAD ruling has included the EPA's evaluation of the city's overall watershed protection programs for these upstate reservoirs.

The New York City Watershed Memorandum of Agreement (MOA) grew out of this entire process (NYS 1996). The MOA is a landmark arrangement between NYC, EPA, local municipalities, and environmental organizations in order to ensure that NYC remains eligible for filtration avoidance through watershed management, without harming the economic welfare of the watershed communities (WRI and Delaware County 1999). The MOA provides funding for many environmental programs within the watersheds, including ongoing monitoring and modeling that aids in development of management practices to reduce pollution to the NYC's reservoirs.

One system of concern for the MOA is the Cannonsville Reservoir and its surrounding watershed. This agricultural watershed has been placed under a 'phosphorus restriction', which restricts future economic growth in the basin when the growth directly or indirectly increases phosphorus loadings (WRI and Delaware County 1999). Major sources of phosphorus into the reservoir are likely to be dairy pastures and manure/fertilizer application. Soluble phosphorus tends to sorb to sediment particles, resulting in sediment acting as its principle transport mechanism (Sharpley and Smith 1990). As a result, accurate simulation of sediment erosion and transport is important to help determine the fate of phosphorus in the Cannonsville system.

This paper focuses on the application of the distributed watershed model, the Soil and Water Assessment Tool (SWAT Version 2000 – SWAT2000), for describing flow and sediment transport on the Cannonsville Reservoir watershed. Ultimately, the goal of this research is to provide decision makers with a tool to evaluate management practices that would potentially reduce non-point source pollution coming from the reservoir's agricultural watershed. A significant amount of data exists on this basin to support model calibration on various spatial and temporal scales.

In our analysis, the model is calibrated, as well as validated. First, SWAT2000 is calibrated to 5 years, 7 months of stream flow data at four locations throughout the Cannonsville watershed (~ 1200 km^2). Because only two gages have sediment monitoring, sediment calibration occurs at these two locations in the watershed over the same period as hydrology. Then, we performed a validation test on an independent data set. The validation is a more stringent evaluation of a model because parameter values are not allowed to be adjusted to fit data. The validation assesses the ability of the model to forecast in periods and areas outside the calibration data. Validation for hydrology occurred at two monitoring sites, while sediment validation is performed at one sampling site. The fewer sites for validation compared to calibration are a reflection of data availability. Because current modeling needs to focus on more than just hydrology for TMDL analysis, careful consideration must be given to other constituents, especially those which are highly dependent on runoff events, such as suspended sediment.

In addition to calibration and validation results, we report on parameter values that were adjusted in order to model this Northeastern basin. The watershed for this SWAT2000 application is dominated by fragipan soils, hills and a climate of relatively cold, snowy winters and clear summers. Most importantly, the winters in this area are characterized by frequent snowpack/snowmelt periods (i.e. snowpack will develop, then temperatures will warm briefly, causing snowmelt to occur, and then the temperature goes below freezing again resulting in more snowfall). These characteristics differ from many of the Midwest and Southern areas in which SWAT has historically been applied, where land use tends to be more dominated by rangeland and snowfall is not an important component of the water balance. Off additional interest for this research is to investigate the possible limitations of applying SWAT to this watershed.

MODEL DESCRIPTION

SWAT is a distributed watershed model developed by the Agricultural Research Service of the United States Department of Agriculture (USDA). SWAT requires a significant amount of data and parameters for development and calibration. The main purpose of SWAT is the computation of runoff and loadings from rural watersheds, especially those dominated by agriculture (Williams and Arnold 1993; Arnold et al. 1998). Because of the documented application of SWAT for agricultural watersheds (see "Past SWAT Applications") and its ease of use for simulating crop growth and agricultural management on the land surface, SWAT2000 was chosen as the watershed model for the Cannonsville Reservoir watershed. This research presents changes made to SWAT2000 parameters values in an attempt to accurately represent the unique characteristics of the Cannonsville watershed. Of additional interest are the possible limitations of the model, including its ability to accurately simulate snowmelt events for hydrology and sediment. Snowmelt is a major hydrological process in the Northeast and accounts for much of the sediment erosion in the winter and spring.

A modeled watershed is subdivided into a number of subwatersheds. These subwatersheds are further divided into hydrologic response units (HRUs), which are units of unique intersections of land use and soils. It is on the level of the HRU that most equations are solved in SWAT. For example, all the land growing corn every year on soil type, NY132, in the subwatershed 42, forms one HRU; however, this land need not be contiguous when designated in the same HRU. In this way, each HRU is modeled as a "lumped" area, meaning that if a given HRU exists in two different areas across the subwatershed, the impact of the HRU area that is closer to the receiving water is not differentiated from the impact the HRU area that is further away from the receiving water.

The foundation behind the hydrologic simulation in SWAT is a soil water balance (Neitsch et al. 2001), where the model tracks precipitation (measured data), and simulated soil water content, surface runoff, evapotranspiration, percolation, and return flow on a daily basis. The runoff is calculated using the Soil Conservation Service (SCS) curve number equation (SCS 1972). Potential evapotranspiration is estimated using the Penman-Monteith equation and corrected for land cover, based on simulated plant growth, to give actual evapotranspiration (Monteith 1965; Neitsch et al. 2001). The model calculates percolation when soil water content exceeds soil field capacity and determines the amount of water moving from one soil layer to the next using a storage routing method (Neitsch et al. 2001). SWAT simulates two groundwater aquifers in each subwatershed: an unconfined, shallow aquifer that contributes to stream flow and a deeper aquifer which does not add to stream flow within the modeled watershed (Arnold et al. 1993). Interflow or 'quick baseflow' is simulated using a kinematic storage model for subsurface flow developed by Sloan (Sloan et al. 1983; Neitsch et al. 2001). Finally, it is desired to understand the impact of different management practices, when implemented within different subwatersheds (i.e. those close to the reservoir as opposed to those far away). Hence, routing is incorporated to capture the important instream mechanisms that affect hydrology and sediment transport. SWAT routes flow by assuming a trapezoidal channel and using Manning's equation along with a variable storage routing method developed by Williams (1969; Neitsch et al. 2001).

Sediment erosion from each HRU is simulated using the Modified Universal Soil Loss Equation (MUSLE) (Williams and Berndt 1977). This equation replaces the traditional Universal Soil Loss Equation's (USLE) rainfall factor with a runoff factor. The MUSLE is solved for each HRU and final sediment yields are routed down the main channels using a stream power equation (Neitsch et al. 2001). This routing method assumes the maximum amount of sediment that can be transported in a given reach is a function of the peak channel velocity (Arnold et al. 1995).

PAST SWAT APPLICATIONS

Various versions of SWAT have been applied throughout the United States, primarily in Midwest and Southwest regions, including the Mississippi River basin, Illinois, Indiana, Texas, Oklahoma, and Wisconsin (Arnold and Allen 1996; Bingner 1996; Bingner et al. 1997; Arnold et al. 1998; Mamillapalli 1998; Manguerra and Engel 1998; Qiu and Prato 1998; Srinivasan et al. 1998; Arnold et al. 1999; Arnold et al. 2000; FitzHugh and Mackay 2000). Most of the previous research has focused on hydrology simulation for these regions, with only a few recent publications on additional constituents such as sediment or nutrients (Bingner 1996; Bingner et al. 1997; Qiu and Prato 1998; Srinivasan et al. 1998; FitzHugh and Mackay 2000; Santhi et al. 2001). The published articles found for Northeastern climates have centered primarily in northeast Pennsylvania (Cho et al. 1995; Peterson and Hamlett 1998). SWAT was applied to a

250 km^2 subwatershed of the Delaware River in Northeast Pennsylvania (the ultimate receiving water for the Cannonsville Reservoir outflow) in order to simulate hydrology and nitrogen (Cho et al. 1995). Peterson and Hamlett (1998) focused their work on modeling a catchment in Pennsylvania which was dominated by fragipan soils and experienced 'severe' snowmelt events throughout the winter and spring. These Pennsylvania studies most closely resemble the Cannonsville Basin in soil-types, geographic location, and climate. Both of these previously published papers are compared to our modeling effort in later discussion.

WATERSHED DESCRIPTION AND INPUT DATA

Watershed Description

The Cannonsville Reservoir in Delaware County, New York, is part of the New York City water supply system (Figure 1). The reservoir's 1178 km^2 watershed is dominated by dairy pastures (26% of land area) and forested land (59% of area). Urban areas comprise less than 1% of its land use while agriculture (corn and hay) and water make up the remaining 14%. The elevation of the watershed varies from 285 m above mean sea level in the lowland areas to around 995 m for the hilltops and the average land-surface slope from a digital elevation model (DEM) is 11%. The basin has ten point sources, primarily wastewater treatment plants for the communities along the West Branch Delaware River. In addition, two of the point sources within the watershed are industrial inputs, including a cooling water discharge.

Because the ultimate goal of this research was to evaluate the impact of different subwatersheds/tributaries on the watershed loadings to the reservoir, the full watershed was subdivided into a number of smaller subwatersheds. The model output can then be analyzed at each of the subwatershed outlets to determine the impact of that area on the entire catchment. The subwatersheds established for the Cannonsville watershed followed those designated by the NYDEP which were based on major tributaries entering the West Branch Delaware River (Figure 1) and Cannonsville Reservoir. These 43 subwatersheds (Figure 2) were delineated with the aid of Geographic Information Systems (GIS) using a 10 meter DEM and stream network (TIGER network obtained from the DEC) (Neitsch and DiLuzio 1999). Because the process of delineating subwatersheds in GIS, which entails the process of 'burning in' or designating the stream locations using the stream network data, is not full-proof, these GIS subwatershed delineations were compared to hand delineations performed by the NYDEP for quality control. The two subwatershed delineations were similar and therefore, the GIS delineations are used in this analysis. Each subwatershed is further divided into HRUs (see Model Description), which are determined by unique intersections of the land use and soils within the basins. These HRUs are the spatial level at which the model establishes management practices such as crop growth, fertilizer application, and livestock management. For the Cannonsville watershed, GIS data were utilized to determine 482 HRUs for the entire basin, resulting in an average of approximately 11 HRUs per subwatershed

Data Overview

A significant amount of data, both spatial and temporal, exists to aid in the development and calibration of a watershed model (Table 1, Figure 2). An unusually large amount of data exists for this watershed, with over 20,000 data measurements consisting of climate, flow and total suspended sediment observations. The data for the model development, calibration, and validation originated from numerous sources and include spatial information such as land use and soils, as well as time variable water quality and climate measurements.

Figure 2 shows the location of the five main climate stations used for input data. These stations provided daily point temperature (maximum and minimum) and precipitation data, which are the forcing functions for the SWAT model. The precipitation data were distributed over the watershed by assigning the data from closest gauge station to the geometric centroid of each subwatershed. In addition to the climate stations, a DEM was used for establishing subwatershed boundaries, elevations, and slopes. Finally, as discussed above, the model used spatial data of land use and soils to establish the 482 HRUs in the watershed.

HYDROLOGY CALIBRATION

For the Cannonsville Basin, SWAT2000 was calibrated over a 5 year, 7 month period from January 1994 to July 1999. First, the stream flow was calibrated on differing spatial and temporal scales. The longest-running flow gauge for the watershed drains approximately 80% of the watershed (Walton, Gauge 01423000, Figure 2). In addition, gauges located throughout the watershed which drain smaller subwatersheds and have shorter periods of record (about 2 years) were used during the calibration procedure (Gauges 01421618, 0142400103, and 01422500, Figure 2). The use of these smaller subwatersheds ensured that the model is accurately simulating the watershed on different spatial scales. Because the ultimate use of this model is to investigate long-term trends on the watershed, it is felt that the model needs to represent long-term average flows and sediment loads to understand the impacts of watershed management practices. In addition, because the model time step is daily, it is difficult to accurately capture daily results due to possible time shifts in the precipitation and flow data. Consequently, the flows were compared on an average monthly and average annual basis to determine if there were trends in model output or error.

Goodness-of-fit measures were evaluated to test the model accuracy. These measures included percent differences in averages and standard deviations over the simulation period, coefficient of correlations (R^2) and the Nash-Sutcliffe measure (NS) (ASCE 1993).

$$NS = 1 - \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(X_{oi} - X_{si}\right)^2}}{\frac{1}{n}\sum_{i=1}^{n}(X_{oi} - \overline{X})^2} \qquad (1)$$

where,

NS	=	Nash-Sutcliffe coefficient and
$\overline{X}$	=	Average measured value during simulation period.
X_{si}	=	Simulated monthly output on day *i*,
X_{oi}	=	Observed monthly output on day *i* and,
i	=	Month observed data was recorded.

NS can range from negative infinity to 1, with 1 being a perfect model to data agreement (ASCE 1993).

All of these goodness-of-fit measures were calculated for all four flow gauges draining various subwatershed sizes. In addition, model-to-data plots were analyzed for possible trends. The comparison of both averages and standard deviations investigates whether the model stream flow frequency distribution is similar to the measured one. Although R^2 values have been used often in the past to compare model results to data, the recommendations of ASCE (1993) indicate that the Nash-Sutcliffe measure is a better representative measure for model goodness-of-fit. As a result, these results will focus on NS, while only providing R^2 to compare this effort to previous applications of the SWAT model. The large amount of data provided confidence in the model calibration exercise because it supplied enough data on varying spatial and temporal scales. As a result, an analysis of the impact of parameter adjustments at a number of different calibration stations was possible.

Figure 3 shows the monthly calibration results (average daily flow for each month) from 1994 to July 1999 for flow at the four USGS stations. The Walton USGS flow station (Gauge 01423000, Figure 2) is just upstream of Beerston sediment monitoring station (WDBN, Figure 2), which is the primary calibration station for this research. The flow at Walton is prorated to Beerston using the ratio of the two stations' drainage areas, which is calculated to be 1.06. Inspection of Figure 3 indicates that the model tends to underestimate flow volumes and storm peaks more often in the winter and spring months. In addition, the flow volumes are over-estimated in the summer months. Plots of simulated average monthly flow versus observed monthly flows given in Figure 4 show R^2 values above 0.7 for all of the calibration stations. However, as mentioned earlier, the model has trouble hitting the high flow periods and the model overestimates in months with relatively lower flows. The model limitations in relation to these discrepancies are discussed further in the SWAT Limitations and Advantages section.

A summary of the statistical results for the hydrology calibration at all four flow stations is summarized in Table 2. The monthly NS values range from 0.63 to 0.78. The percent difference in averages range in magnitude from 4% for the main discharge point at Walton (Gauge 01423000) to 29% for the smallest and most upstream watershed, Town Brook (Gauge 01421618). The model has a lower standard deviation in monthly flow than is shown in the observation data at all stations, indicating that the model simulates less variability in the monthly average flow than what actually occurs. As shown in Table 2, the average simulated flow values are all within 5% of the data for three of the four stations, indicating that, in general, the volume of water is accurately simulated on large and small scales. The model does seem to have trouble with the smallest watershed (Town Brook, 37 km^2, 3% of watershed). This station experiences the smallest flow volumes of the four stations and small fluctuations in simulation results can cause large differences in summary statistics. In addition, the larger discrepancy at the Town Brook subwatershed also highlights a limitation of the model for simulating smaller subwatersheds, when integrated into larger systems.

These results for SWAT in the Cannonsville watershed compare to other SWAT applications in northeastern areas. For purposes of comparison, the reader is directed to the R2 values for this application in Table 2 (R^2 range = 0.72 to 0.80). The research performed in Northeastern Pennsylvania, which is an area similar in climate, land use, and soil type to the Cannonsville Basin, attained monthly hydrology R^2 values of 0.14 overall and 0.55 when snowmelt periods were 'removed' from the statistics (Peterson and Hamlett 1998). Another Northeastern study, Cho (1995) had R^2 which ranged from 0.57 to 0.83 for a small forested watershed in the Delaware River basin. Although a previous study in the Midwest (Srinivasan et al. 1998) obtained higher R^2 values (0.87 and 0.84), this midwestern area is not fully representative of the present watershed (since it has relatively dry soils with minimal snowfall). Another Midwest project applied SWAT to three Illinois watersheds, resulting in monthly R^2 values of 0.63, 0.78, and 0.95 for basin sizes of 122, 246, and 188 km^2, respectively (Arnold and Allen 1996). More recently, a modeling effort in Texas was completed and obtained calibrated hydrology R^2 values of 0.80 and 0.89 for two subbasins of 926 and 2997 km^2, respectively (Santhi et al. 2001). The Santhi et al. (2001) study assumed an 'acceptable calibration' for hydrology at an $R^2 > 0.6$.

The availability of climate data also plays a part in the accuracy of the model simulation. Arnold et al. (1998) cited spatial variability with precipitation data as one of the 'major limitations to large area hydrologic modeling.' Hoblit (1999) presented National Weather Service guidelines which recommend seven point climate stations to simulate a 1200 km^2 basin. However, the Cannonsville input for the model is based on just five point rain gauges (see Figure 2) from which precipitation must be distributed over all 43 subwatersheds. As a result, it is very possible that the rainfall over the entire watershed is being misrepresented for the model input. Corrections for this error are difficult without additional information. For example, radar imaging (i.e. NEXRAD data) could be used in place of point rain gauges to improve the rainfall representation (Fan et al. 1996; Peters and Easton 1996; Bedient et al. 2000).

SEDIMENT CALIBRATION

The United States Geological Survey (USGS) collects flow measurements at fifteen-minute intervals and reports the data as mean daily flow. Total suspended solids (TSS), which is the measurement used for sediment calibration, is not typically recorded on such short intervals. As a result, this research calibrated suspended sediment to monthly sediment loads, which were estimated by the New York Department of Environmental Conservation (DEC) using data at a station near the primary flow gauge (WDBN, Beerston, Figure 2).

Similar to the hydrology, the sediment results were desired to understand the long-term average trend of the system to management issues. As a result, the monthly sediment loads were a focus for calibration, as opposed to daily results. In addition, considerable uncertainty exists in the estimated measured sediment loads (see discussion below), therefore, it did not seem practical to compare daily model results to daily measured data. Finally, sediment erosion and transport modeling is highly uncertain and accurate simulation of sediment processes on the land surface is difficult to capture due to the heterogeneous nature of a watershed and the relatively unrefined equations used to explain certain processes (e.g. MUSLE). As a result, it is typically the case that a model that performs acceptably for hydrology, may still have limitations in fully capturing sediment loads. Consequently, this research will focus on the ability of the model to capture long-term trends on a percent difference basis, while further investigating the limitations SWAT may have in simulating different sediment processes.

The DEC developed estimated monthly loadings for sediment by combining approximately bi-weekly TSS sampling with high-intensity sampling during high flow events (Longabucco and Rafferty 1998). A high-flow event was defined as a rise in river stage of at least 0.152 m (0.5 ft) (Longabucco and Rafferty 1998). The frequency of sampling at Beerston was based on stage height at the USGS Walton flow gauge (Gauge 01423000, Figure 2). Longabucco and Rafferty (1998) details the protocol for sampling at Beerston, which ranges from one sample per week for stage height less than 0.91 m (3 ft) to hourly sample for stage height above 3.97 m (13 ft). Approximately 50% of the days contained one grab sample to characterize TSS concentration for that day. However, TSS concentration varies greatly with flow rate and as a result, one sample a day is not always representative of the TSS concentration occurring throughout an entire 24-hour period. Because the greatest variation in TSS concentration occurs during high-flow events (i.e., storms), more intensive sampling took place during these time periods

TSS concentrations were determined using the following approach: 1) use measurements, when available, 2) linearly interpolate days with no data between days with data, and 3) when a large data gaps occurred, assume a concentration based on flow and previous measurements. The TSS concentrations were then summarized into monthly estimated loadings (Longabucco and Rafferty 1998). These DEC estimated loadings were compared to model output using graphs and percent difference calculations, bearing in mind that the DEC estimated loads are most likely underestimates of sediment loading (e.g. some high flow events may be missed by the sampling effort). For example, because the sampling, for the most part, was manual, the sampling team may not have arrived at the station in time to capture a short storm which began unexpectedly. In addition to the primary station at Beerston, we also compared sediment observations and model results at an additional station (CTNBG, Town Brook, Figure 2), which has a limited amount of data (e.g. 10 months), and drains a smaller area (37 km^2). The DEC also developed estimated monthly loads for this Town Brook station (CTNBG, Figure 2) in a similar manner as described for Beerston (WDBN, Figure 2). As with hydrology, calibrating the sediment using two different gauges ensured the model accuracy on large and small spatial scales.

Sediment results at the West Branch Delaware River Beerston station (WDBN) and Town Brook (CTNBG) are shown in Figure 5. Figure 5 displays the model output along with the DEC estimated loadings. Overestimations in the model results could be explained by 'missed' storm events in the sampling effort. However, it is more of a concern when the model is underestimating the peak predicted from data. The largest discrepancy between modeled and the DEC estimated loadings occurs in January 1996. This event was a flood that resulted in a load of almost 60,000 metric tons (MT). The model is unable to capture sediment erosion during such a high flow event due to water movement into the flood plain, resulting in erosion and deposition in the flood plain, which is not simulated in SWAT. Because SWAT's equations do not capture the erosion processes occurring during this flood event, the model is compared to the data both including and excluding this storm.

Table 3 indicates that the calibrated monthly model averages come within approximately 1% of the measured data for the large station (WDBN), when excluding the January 1996 flood event. However, inclusion of this flood event results in an underestimate of the model to the data by 42%. This presents a limitation of the SWAT model for capturing flood plain erosion (see SWAT Limitations and Advantages section). The model sediment results at the smaller station of Town Brook (CTNBG) did not perform as well, which was also true for the hydrology results. However, this station only had 10 months of available data, which is likely not sufficient for a robust evaluation of the model. In addition, a storm event that is captured by the model in March 1999 may not have been captured by the monitoring effort, causing an underestimation in the estimated measured sediment load. The differences between DEC estimated loadings and those simulated may also be attributed in part to differences in modeled and observed flows.

The sediment results obtained here compare to previous research on sediment transport simulation. One article discussing an application of SWAT in the Texas Gulf basin compared SWAT simulated annual sediment loads to loads estimated from data using sediment rating curves. The annual sediment load results for six different basins displayed percent differences ranging from –35% to 130%. The smallest difference was for a 13,000 km^2 basin at 6% (Arnold et al. 1999). Srinivasan (1998) also performed a sediment calibration in Texas and came within 2% of the measured sediment loads. More recently, additional work in Texas calibrated sediment yields (i.e. MT/ha) at two large subbasins and obtained percent differences of –16% and –20% (Santhi et al. 2001).

There have also been a few other journal publications using other models to simulate sediment transport. One long-term study applied sediment transport routines very similar to those in SWAT and obtained model-to-data percent differences ranging from -30% to 5% over five different simulation periods for a 257 km^2 watershed in Northern Texas (Arnold et al. 1995). A calibrated storm-event model (i.e. hourly time-step) developed in the United Kingdom was found to underestimate total sediment loading by 6 to 23% for the three storms chosen for the calibration (Moore 1984). Finally, a series of calibrated 'box-models' developed for a specific urban watershed in Massachusetts simulated sediment transport on average, within -6.8% of the average daily measured loads (Solo-Gabriele and Perkins 1997).

PARAMETER CHANGES

Table 4 presents an overview of the SWAT2000 parameter changes for the model calibration. These changes reflect, in a large part, the special characteristics of this Northeastern watershed (i.e. fragipan soils with large snowmelt events). The default curve numbers set by the SWAT user interface (i.e. the values recommended by the SCS Handbook - (Service 1972)) were reduced by 15%, indicating that the Cannonsville watershed has better soil/drainage than the assumed conditions in the default SWAT database. The runoff lag coefficient was increased to its maximum value (0.99) so that storm recessions would be less steep. The soil evaporation compensation factor was decreased from the default value of 0.95 to 0.85, resulting in more evapotranspiration, especially during summer months. In addition, the available water capacity in all soils was reduced by 25% (e.g. multiplied by 0.75) to increase baseflow volumes. This change in soil water content followed an approach taken by researchers in Northeastern Pennsylvania where a SWAT modeling exercise in similar soils as the Cannonsville Basin (fragipans) found that the model underestimated baseflow due to the soil types in the watershed (Peterson and Hamlett 1998). The groundwater coefficient that controls the amount of water that moves from the shallow aquifer to the root zone (i.e., the SWAT 'revap' parameter) was changed from 0.02 to 0.1 to allow more movement of water from the shallow aquifer to the unsaturated root zone. Finally, in an attempt to accurately simulate snowmelt the threshold snowpack temperature for snowmelt to begin was increased from 0.5 degrees Celsius to 5 degrees Celsius in order to simulate snowmelt only when warmer air temperatures persisted.

For the sediment calibration, only two model parameters were changed from the SWAT defaults (see Table 4). The peak rate adjustment factor for sediment routing in the tributary channels was changed from 1 to 0.6. In addition, the average slope length used in the MUSLE and time of concentration calculations was adjusted (reduced) from the values set during the initial model set-up in GIS (Table 3). Neitch (2000) states that the average slope length parameter is one of the more uncertain values for a subwatershed and can be adjusted for calibration of SWAT.

MODEL VALIDATION

Model validation is the process of re-running the simulation, using a different time-series for input data, without changing any parameter values which may have been adjusted during calibration. Validation can also occur during the same time-period as calibration, but at a different spatial location. In this case, data over a four year period from 1990 to 1993 were used to validate the model for a different time-period. One flow station and one water quality station were available for data-to-model comparisons during the 1990 to 1993 time-period. In addition, one flow station, which had not been used during the calibration exercise, was available for data-to-model comparisons, with data from 1997-1999. This 1997-1999 station served as a ‘consistency-check’ of the calibration results. As with the calibration, the three above-mentioned goodness-of-fit measures were calculated and model-to-data plots were inspected.

The purpose of model validation is to establish whether the model has an ability to estimate output (hydrology and sediment in this case) for other locations, time periods or conditions than those for which the parameter values were adjusted to fit. Model validation involves re-running the model using input data independent of data used in calibration (e.g. differing time period), but keeping the calibrated parameters unchanged. Flow data from Walton (Gauge 01423000) was available from 1991 to the end of 1993, providing hydrology validation in a different time period. In addition, flow data approximately 27 km upstream from Walton on the West Branch Delaware River at a station upstream of Delhi, NY (Gauge 01421900, Figure 2) existed from 1997 to 1999. Although this station contained data which fell within the calibration time period and its flow was correlated to Walton, this gauge was not used during the calibration exercise, providing location for a check of the hydrology calibration results.. Finally, one water quality station (WDBN, Figure 2) contained sediment loading data from late 1991 to the end of 1993 for validation analysis.

The results of the model validation are shown in Figures 6 through 8. The plotted hydrology results in Figures 6 and 7 indicate the model performed as well in the validation period of 1990-1993, as for the calibration period of 1994-1999 at the Walton gauge (validation NS = 0.76; calibration NS = 0.74 – see Table 2). The Delhi gauge provided an analysis of the model performance during the calibration period but at a different location. Delhi’s flow is correlated with Walton’s (i.e. it is directly upstream of the Walton gauge), indicating this gauge is not an independent check on the model. Although the Delhi gauge does not fully satisfy the requirements for validation, it does provide us with a consistency check on the calibration results. This gauge displayed an NS above 0.6, which was similar to the calibration NS for the Town

Brook watershed and lower than the calibration results at Trout Creek, and Little Delaware. The parameter values used in the 1990-1993 simulation were not changed from the values calibrated to the 1994-1999 data in order to insure that the validation was done using an independent dataset. This validation/calibration check illustrates the accuracy of the model for simulating time-periods and locations outside of the calibration time period.

In addition, the sediment validation results displayed a 31% difference between the observed and simulated average monthly load (Table 3, Figure 8). Although this difference is larger than for the sediment calibration results that excluded January 1996, it should be noted that the high sediment loading event in April 1993 was the result of snowmelt erosion, which is difficult to model. As with the calibration results, these validation results illustrate the difficulty in capturing sediment loading results when using equations that do not fully describe the system processes (i.e. the use of a stream power equation for sediment transport) and those that have parameters that can not be determined directly from field data (i.e. the MUSLE equation parameters).

These validation results are comparable to previous studies. Two other SWAT studies approached validation in the same way as this study (i.e. model was run with an independent dataset and the parameter values remained unchanged from the calibration). Srinivasan et al. (1998) obtained hydrology calibration R^2 of 0.87 and 0.84 for two basins in Texas. The hydrology validation of the model in the same subwatersheds produced R^2 of 0.65 and 0.82, respectively (Srinivasan et al. 1998). Santhi et al. (2001) reported hydrology validation R^2 of 0.92 and 0.80 for two other basins in Texas (calibration R^2 = 0.80 and 0.89, respectively.

The sediment validation for the Srinivasan (1998) research compared estimates of cumulative loads developed from a sediment survey to simulated results. They found the model underestimated total estimated load by about 14% (calibration difference = 2%). Santhi (2001) obtained differences in sediment yields for validation at -44% and –70%, while the calibration differences for the same two basins were only –16% and –20%. This marked difference in the validation versus the calibration results highlights the difficulty in simulating sediment erosion and transport, even after the attainment of an acceptable calibration.

There were few sediment validation studies for any watershed model, presumably because sediment forecasting and continuous TSS data collection is difficult. Some literature was found on validation results for models other than SWAT. Moore (1984) performed a sediment loading validation using a storm-event based model and attained total sediment load differences of -32.8% to 15.5% (calibration range: -6.3% to 23.3%). Finally, the watershed-specific model developed by Solo-Gabriele (1997) came within 19% of the measured average daily sediment loads for its validation period (calibration: -6.8%). It should be noted, that these two models vary from the SWAT model. The Moore (1984) model was a storm-event model specifically designed to model sediment transport and the Solo-Gabriele model was developed with a specific urban watershed in Massachusetts in mind (1997). Therefore, it is difficult to make exact comparisons between these models and a more general calibrated model, such as SWAT.

SWAT LIMITATIONS AND ADVANTAGES

Perhaps one of the more important results from this research is the investigation of the limitations of the SWAT model, when applied to the Cannonsville system, or one of similar characteristics. Our research indicates that there are modeling limitations, primarily in modeling snowmelt, the simulation of evapotranspiration, and the capturing of sediment erosion processes. In addition, a limitation seems to exist in modeling small subwatersheds, when they have been integrated and calibrated into a larger system. This section discusses these weaknesses and advantages to the SWAT mode as well as presents some areas where further research could occur to improve them.

An overall weakness of the SWAT model is the use of equations that have parameters that are not directly measured using data. For example, the curve number equation, although used often to estimate runoff volumes, is highly uncertain due to the use of a parameter (i.e., the curve number) that has been determined empirically using different land uses. In addition, the MUSLE, which is used for soil erosion simulation, is also uncertain because of the number of parameters in the equation that are set from qualitative information (e.g., soil type and ground cover). Although efforts have been made to incorporate more process-based equations, there is still room for improvement in some of the basic processes modeled by SWAT.

The first specific shortcoming is the model's simulation of snowmelt for hydrology. The underestimation of flows in the winter and spring in Figure 3 are most likely due to the simulation of snowmelt. The difficulty in simulating snowmelt has been documented in previous work with SWAT. The Pennsylvania study mentioned above, located about 100 km southwest of the Cannonsville watershed, applied SWAT in an area analogous to the Cannonsville Basin. This Pennsylvania research found that the model had 'difficulty' accurately capturing snowmelt periods (Peterson and Hamlett 1998).

In addition, our model's overestimation of flows in the summer may be due to the underestimation of evapotranspiration. The Penman-Monteith equation, which is used to estimate evapotranspiration, requires significant data, including, but no limited to, solar radiation, wind speed, soil characteristics, and canopy cover characteristics. Because only precipitation and temperature were available as input data, the other meteorological data needed for this calculation were generated using a weather generator in SWAT. Considerable uncertainty exists in weather generation, even though the parameters for the weather generator were set for an area in upstate New York. As a result, there is considerable uncertainty in the final evapotranspiration values determined by SWAT.

In terms of sediment erosion modeling, the SWAT model can not capture flood plain erosion, which is shown in the inability to simulate the January 1996 flood event in the Cannonsville that experienced both floodplain erosion and snowmelt erosion. The MUSLE algorithm is designed to simulate erosion occurring due to the runoff produced during a storm event (Williams and Berndt 1977). Arnold et al. (1998) stated that 'SWAT does not simulate detailed event-based flood and sediment routing'. They felt the model was best developed to evaluate management impacts on long-term erosion and sedimentation (Arnold et al. 1998). As a result, it would be inaccurate to apply the model in an attempt to evaluate particular flooding events, especially those in which flood plain erosion occurred.

In addition to floodplain erosion, during the spring and winter months, snowmelt erosion is difficult to model using the MUSLE which is employed by SWAT (Williams and Berndt 1977). The original USLE was developed to simulate erosion caused by sheet flow runoff. As a result, even the MUSLE, which was adapted by replacing the rainfall energy used in the USLE with a runoff factor, has difficulty with snowmelt erosion. This is shown by the model's inability to capture the sediment load at

Beerston in January 1996. The January 1996 event was characterized by heavy rains falling on an existing snowpack. Currently, the SWAT erosion algorithms are designed to reduce sediment erosion calculated by the MUSLE when snowpack exists (Neitsch et al. 2001). These predictions are a function of snowpack depth and are unrelated to surface runoff or soil temperature. However, in the case of heavy rains and high amounts of snowmelt such as those experienced in January 1996, it may not be accurate to assume that erosion was minimal. Although the snowpack existed, the heavy rains most likely accelerated the snowmelt and caused considerable soil erosion that was not captured by the SWAT algorithms.

Another cited limitation in SWAT is the sediment routing routine, which has 'relatively simplistic' equations (Arnold et al. 1998). Currently, the sediment transport routine uses a stream power equation that calculates the maximum transportable sediment load, given the stream velocity. However, this equation does not consider important sediment transport characteristics, such as bottom shear stress, which determines whether erosion or deposition will occur, given flow velocities and resulting shears. It may be beneficial to investigate the refinement of these equations to incorporate more representative processes in the stream channel.

Finally, the simulation of the Town Brook watershed illustrated a possible limitation in the model when modeling small watersheds. In this case, SWAT did not perform as well on Town Brook, which was one of the smallest subwatersheds in the Cannonsville system (3% of the area), as it did on other, larger watersheds. This may be due to the fact that parameters had to be adjusted on a basinwide basis. While these parameter adjustments seemed to be accurate for the system as whole, they may not have been representative for a small area such as Town Brook. Because of Town Brook's size, the response of the subwatershed to precipitation events is more sensitive and timing of the events is crucial for model performance. Consequently, errors that may be 'averaged-out' on larger basins would be quite apparent on small basins because the input data and parameters have been set on a large scale. In this case, it may be more accurate to calibrate Town Brook, separately and account for time shifts in the precipitation data or changes on the land surface in order to accurately capture the processes on such a small scale.

Although there are limitations to the SWAT model, it is still a viable model for simulating water balance and sediment transport on large watersheds. Past efforts have illustrated calibrated models that were used to investigate management options and understand crop growth (e.g., Santhi et al. 2001). The model is capable of the continuous simulation of flow and sediment transport in a semi-distributed fashion for analyzing the impact of different subwatersheds on a receiving water body. In addition, the model developers should be recognized for the model's ease of use and high quality of documentation, which is often lacking in many 'off the shelf' model packages.

CONCLUSIONS

With the associated changes in parameter values for a Northeastern region, the SWAT2000 model was calibrated and validated. Validation indicates the ability of the model to forecast in time periods and locations outside of the calibration area, which is important since common use of a watershed model includes forecasting. In addition, this study illustrated the application of the new version of SWAT (SWAT2000) to a unique watershed in which interflow plays an important role in stream flow due to fragipan soils and relatively steep slopes.

At Walton, the station which drains 80% of the entire watershed, the NS = 0.74 for calibration (simulation years 1994 to 1999) and NS = 0.76 for validation (simulation years 1990 to 1993) for hydrology. In terms of average flow at Walton, the model came within 3.8% of the data for the calibration and 7.8% for the validation. The standard deviation of the Walton flow in the model is significantly lower than in the observed data (10.5 versus 16 for calibration and 10.9 versus 15.2 in validation). The calibration also included fitting data for three smaller drainage areas. These fits had NS ranging from 0.63 to 0.78. Finally, a smaller station, which was not used in the calibration procedure but contained data from 1997 to 1999 was analyzed for addition model validation. This station displayed a validation NS = 0.62, again indicating the model was accurate at differing spatial scales. These hydrology results indicate the model is effectively simulating both large and small spatial scales in the watershed.

In general, sediment is more difficult to model than hydrology. The calibration results for sediment predicted mean values were only 1% different from the data, when an extreme flood event was omitted from the comparison. The sediment validation results had a 31% underestimation of sediment load, which is within the range of error presented for other sediment validation studies. These results were reasonable, considering the difficultly most regional models have with sediment transport representation. Future research will attempt to identify ways to improve sediment loading simulation.

Parameter values were changed in the calibration process from defaults set from the SWAT database. These changes attempted to reflect northeastern conditions, including long lasting periods of deep snow pack and fragipan soils. Parameters modified during calibration included reductions in curve numbers, available water content, and average slope length and were primarily related to the accurate representation of land surface and soils (i.e. fragipans).

Limitations for the model calibration existed in both the available data and the model formulation. The greatest data limitations are perhaps the availability of climate data to distribute over a large area and the infrequency of TSS concentration samples. For the Cannonsville watershed, there exist only five gauges for a 1200 km^2 basin. Underestimation of sediment transport could also be due to lack of data when developing the DEC estimated sediment loadings, which were developed through bi-weekly sampling, intensive sampling during high flow events, and linear interpolation when data was unavailable. Model limitations included the simulation of snowmelt, evapotranspiration, snowmelt erosion, floodplain erosion, and sediment transport routines. Snowmelt erosion, in particular, is not modeled by the MUSLE, which is the equation employed by SWAT for sediment yield calculations.

The lack of a more robust NS or percent differences indicates limitations in the predictive capability of the calibrated model, especially for sediment. There are many possible sources of these limitations, including a) errors in input data (precipitation, soil type, rock cover, etc.), b) over-simplification of various factors in the model equations, c) values of the calibrated parameters which are not optimal, and d) errors in the observed output data (e.g. sediment loads). However, based on current information, the model demonstrates its utility as a useful tool to understand processes in the watershed and as a basis for effective management in the Cannonsville and other Northeastern watersheds.

ACKNOWLEDGEMENTS

J. Benaman’s tuition and stipend were funded under the EPA’s Science to Achieve Results (STAR) fellowship, Grant #U915552. Delaware County, the DEC, Cornell University’s WRI, the USGS, and the NYDEP all provided data and insight which was vital to this project’s completion. Particular thanks are given to Steve Pacenka (Cornell WRI), Keith Porter (Cornell WRI), Patricia Bishop (DEC), Carol Stepcheck (NYDEP), and Dean Fraiser (Delaware County) for their dedication, support, and advice throughout the duration of the project. The advice of the researchers at the Blackland Research Center (Texas A&M), especially Drs. Jeff Arnold, Susan Neitch, and Nancy Sammons is greatly appreciated. Finally, the authors would like to acknowledge the work of Bryan Tolson (Cornell), who aided in the initial model development and calibration.

APPENDIX I. REFERENCES

American Society of Civil Engineers Task Committee on Definition of Criteria for Evaluation of Watershed Models of the Watershed Management Committee, (ASCE) (1993). "Criteria for Evaluation of Watershed Models." *Journal of Irrigation and Drainage Engineering* 119(3): 429-442.

Arnold, J. G. and P. M. Allen (1996). "Estimating Hydrologic Budgets for Three Illinois Watersheds." *Journal of Hydrology* 176: 57-77.

Arnold, J. G., P. M. Allen and G. Bernhardt (1993). "A Comprehensive Surface-Groundwater Flow Model." *Journal of Hydrology* 142: 47-69.

Arnold, J. G., R. S. Muttiah, R. Srinivasan and P. M. Allen (2000). "Regional Estimation of Base Flow and Groundwater Recharge in the Upper Mississippi River Basin." *Journal of Hydrology* 227: 21-40.

Arnold, J. G., R. Srinivasan, T. S. Ramanarayanan and M. DiLuzio (1999). "Water Resources of the Texas Gulf Basin." *Water Science and Technology* 39(3): 121-133.

Arnold, J. G., R. Srinivason, R. R. Muttiah and J. R. Williams (1998). "Large Area Hydrologic Modeling and Assessment Part I : Model Development." *Journal of the American Water Resources Association* 34(1): 73-89.

Arnold, J. G., J. R. Williams and D. R. Maidment (1995). "Continuous-Time Water and Sediment-Routing Model for Large Basins." *Journal of Hydraulic Engineering* 121: 171-183.

Bedient, P. B., B. C. Hoblit, D. C. Gladwell and B. E. Vieux (2000). "NEXRAD Radar for Flood Prediction in Houston." *Journal of Hydrologic Engineering* 5(3): 269-277.

Bingner, R. L. (1996). "Runoff Simulated from Goodwin Creek Watershed using SWAT." *Transactions of the American Society of Agricultural Engineers* 39(1): 85-90.

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.

Cho, S., G. D. Jennings, C. Stallings and H. A. Devine (1995). "GIS-Based Water Quality Model Calibration in the Delaware River Basin." *American Society of Agricultural Engineers Meeting Presentation*: Microfiche No. 95-2404.

Eckhardt, K., S. Haverkamp, N. Fohrer and H.-G. Frede (2002). "SWAT-G, a Version of SWAT99.2 Modified for Application to Low Mountain Ranges Catchments." *Physics and Chemistry of the Earth (accepted).*

Fan, Y., E. E. Wood, M. L. Baeck and J. A. Smith (1996). "Fractional Coverage of Rainfall Over a Grid: Analyses of NEXRAD Data Over the Southern Plains." *Water Resources Research* 32(9): 2787-2802.

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.

Hoblit, B. C. (1999). *Spatial Scale Data Requirements Using NEXRAD (WSR-88D) for Accurate Hydrologic Prediction in Urban Watersheds.* Master of Science. Department of Environmental Science and Engineering. Rice University.

Line, D. E., D. L. Osmond, S. W. Coffey, R. A. McLaughlin, G. D. Jennings, J. A. Gale and J. Spooner (1997). "Nonpoint Sources." *Water Environment Research* 69(4): 844-860.

Longabucco, P. and M. R. Rafferty (1998). "Analysis of Material Loading to Cannonsville Reservoir: Advantages of Event-Based Sampling." *Journal of Lakes and Reservoir Management* 14(2-3): 197-212.

Mamillapalli, S. (1998). *Effect of Spatial Variability on Stream Flow Modeling.* Ph.D. Thesis. Purdue University.

Manguerra, H. B. and B. A. Engel (1998). "Hydrologic Parameterization of Watersheds for Runoff Prediction Using SWAT." *Journal of the American Water Resources Association* 34(5): 1149-1162.

Monteith, J. L. (1965). "Evaporation and the Environment". *In The State of Movement of Water in Living Organisms*. Swansea, XIXth Symposium Soc. for Exp. Biol. 205-235.

Moore, R. J. (1984). "A Dynamic Model of Basin Sediment Yield." *Water Resources Research* 20(1): 89-103.

Neitsch, S., J. G. Arnold and J. R. Williams (2000). *Soil and Water Assessment Tool User's Manual - Draft.* Temple, TX. USDA Agricultural Research Service and Texas A&M Blackland Research Center.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry and J. R. Williams (2001). *Soil and Water Assessment Tool Theoretical Documentation: Version 2000.* Temple, TX. USDA Agricultural Research Service and Texas A&M Blackland Research Center. 506.

Neitsch, S. L. and M. DiLuzio (1999). *ArcView Interface for SWAT 99.2*, USDA Agricultural Research Center and Texas A&M University Texas Agricultural Experiment Station.

New York State. (1996). *New York City Watershed Memorandum of Agreement.* http:/www.state.ny.us/watershed

Peters, J. C. and D. J. Easton (1996). "Runoff Simulation Using Radar Rainfall Data." *Water Resources Bulletin* 32(4): 753-760.

Peterson, J. R. and J. M. Hamlett (1998). "Hydrologic Calibration of the SWAT Model in a Watershed Containing Fragipan Soils." *Journal of the American Water Resources Association* 34(3): 531-544.

Qiu, Z. and T. Prato (1998). "Economic Evaluation of Riparian Buffers in an Agricultural Watershed." *Journal of the American Water Resources Association* 34(4): 877-890.

Santhi, C., J. G. Arnold, J. R. Williams, W. A. Dugas, R. Srinivasan and L. M. Hauck (2001). "Validation of the SWAT Model on a Large River Basin with Point and Nonpoint Sources." *Journal of American Water Resources Association* 37(5): 1169-1188.

Sharpley, A. N. and S. J. Smith (1990). "Phosphorus Transport in Agricultural Runoff: The Role of Soil Erosion". *In Soil Erosion on Agricultural Land*. J. Boardman, I. D. L. Foster and J. A. Dearing. New York, John Wiley & Sons Ltd. 687 pp.

Sloan, P. G., I. D. Moore, G. B. Coltharp and J. D. Engel (1983). *Modeling Surface and Subsurface Stormflow on Steeply-Sloping Forested Watershed.* Lexington. University of Kentucky. 142 pp.

Soil Conservation Service, (SCS). (1972). "Section 4: Hydrology". *In National Engineering Handbook*. SCS.

Solo-Gabriele, H. M. and F. E. Perkins (1997). "Watershed-Specific Model for Streamflow, Sediment, and Metal Transport." *Journal of Environmental Engineering* 123(1): 61-70.

Srinivasan, R., T. S. Ramanarayanan, J. G. Arnold and S. T. Bednarz (1998). "Large Area Hydrologic Modeling and Assessment Part II : Model Application." *Journal of the American Water Resources Association* 34(1): 91-101.

United States Environmental Protection Agency, (EPA). (1997). *New York City Filtration Avoidance Determination.*

Water Resources Institute, (WRI). and N.Y. Delaware County (1999). *Delaware County Comprehensive Strategy for Phosphorus Reductions*. Cornell University, Center for the Environment, Prepared for the Delaware County Board of Supervisors.

Williams, J. and J. Arnold (1993). *A System of Hydrologic Models.* U.S. Geological Survey. Water-Resources Investigations Report 93-4018.

Williams, J. R. (1969). "Flood Routing with Variable Travel Time or Variable Storage Coefficients." *Transactions of the American Society of Agricultural Engineers* 12(1): 100-103.

Williams, J. R. and H. D. Berndt (1977). "Sediment Yield Prediction Based on Watershed Hydrology." *Transactions of the ASAE* 20(6): 1100-1104.

KEYWORDS

Watershed modeling, northeastern climate, SWAT, calibration, validation, sediment transport

NOMENCLATURE

NS	=	Nash-Sutcliffe coefficient and
$\overline{X}$	=	Average measured value during simulation period.
X_{si}	=	Simulated monthly output on day *i*,
X_{oi}	=	Observed monthly output on day *i* and,
i	=	Month observed data was recorded.