Modeling a Small, Northeastern Watershed with Detailed, Field-Level Data

T. L. Veith, A. N. Sharpley, J. G. Arnold

Abstract. *Time, resource, and replication constraints limit the practicality of conducting agricultural experimental studies at scales larger than plot-level. Thus, watershed-level models such as the Soil and Water Assessment Tool (SWAT) are increasingly used to forecast effects of land management changes on downstream water quality. With the generalization in scale, the question of effect of generalization in input data arises. That is, to what extent does having field-level, daily input data for a watershed model aid in the ability to predict watershed-scale, water quality impacts. The study site, FD-36, is a 39.5 ha agricultural subwatershed of a long-term USDA-ARS study watershed in south central Pennsylvania. FD-36 is characterized by loamy soils with a substantial near-stream fragipan. Fifty percent of FD-36 includes 24 row-cropped fields from three independently managed farms. Two SWAT scenarios were simulated on FD-36 and compared with each other as well as with measured data over two 4-year periods (1997-2000 and 2001-2004). The high-resolution scenario modeled seasonal crop, fertilizer, and tillage events of each row-cropped field continuously over the 8-year period. The low-resolution scenario treated all row-cropped fields as the same generic crop (AGRR in SWAT). Flow depth predictions at the outlet were similar for the two SWAT scenarios. While both scenarios showed higher levels of soil water in the fragipan soils than the surrounding soils, the high-resolution scenario was able to identify field-to-field distinctions due to the increased detail in input data. In general, model results were more defined at the field-level under the high-resolution scenario and followed patterns expected from knowledge about soil science, hydrology, P transport, and the characteristics of the study watershed. However, the time spent collecting, understanding, entering, and error-checking input data required for the high-resolution scenario was on the order of months, while full data collection for the low-resolution scenario took several days. Results suggest that while detailed input data can enable the model to provide valuable water quality information, research efficiency during exploratory and initial problem-solving efforts might be maximized by using more easily obtained, although more general, data.*

Keywords. *Data uncertainty, Model accuracy, Statistical analysis, SWAT, Water quality.*

Efforts by the USEPA to improve water quality through Total Maximum Daily Loads (TMDLs) have contributed to increased use of watershed-level water quality models to assess causes of impaired water bodies and provide remedial solutions (USEPA, 2007). Since 2003, the USDA Conservation Effects Assessment Project (CEAP) has promoted use of watershed-level models throughout the U.S. to help quantify effects of various conservation practices on improving and maintaining water quality nationwide (USDA-NRCS, 2006). Numerous TMDL and CEAP studies have used the Soil and Water Assessment Tool (SWAT), a physically based, watershed-level water quality model, on watersheds of diverse size and complexity. For example, Santhi et al. (2001) assessed the long-term effects of various BMPs on P export from a 4277 km^2 pasture-rangeland dominated watershed experiencing urban growth in Texas. They found P loads could be reduced 50% by implementing several measures, which included limiting dairy manure applications to crop P needs, exporting 38% of manure generated in the watershed, reducing P in livestock diet, and adopting a 1 mg L^{-1} wastewater treatment effluent P limit (Santhi et al., 2001). Gitau et al. (2004) used SWAT in combination with mathematical optimization to evaluate and improve BMP effectiveness for a 300 ha mixed-land use farm (corn silage, hay, pasture, forest) in New York. The study showed that careful placement and selection of BMPs based on soil type, land use, and stream presence could reduce long-term average annual dissolved P losses from the farm by 60% (Gitau et al., 2004). The wide applicability of SWAT is further demonstrated in reviews by Borah and Bera (2004) and Gassman et al. (2007).

Due to the inability to completely understand and recreate a natural system, various types of uncertainty are inherent in modeling agricultural systems. Sources of uncertainty, or discrepancy, between modeled outputs (predicted value) and natural system outputs (measured value) include but are not limited to input data, model parameter values, model processes, and process interactions (Harmel et al., 2006; van Griensven and Meixner, 2006). Plot-scale experiments

Submitted for review in February 2007 as manuscript number SW 6872; approved for publication by the Soil & Water Division of ASABE in February 2008.

The authors are **Tamie L. Veith, ASABE Member Engineer,** USDA-ARS Pasture Systems and Watershed Management Research Unit, Agricultural Engineer, University Park, Pennsylvania; **Andrew N. Sharpley,** Soil Scientist, Department of Crop, Soil and Environmental Sciences, University of Arkansas, Fayetteville, Arkansas; and **Jeffrey G. Arnold, ASABE Member Engineer,** Agricultural Engineer, USDA-ARS Grassland Soil and Water Research Laboratory, Temple, Texas. **Corresponding author:** Tamie L. Veith, USDA-ARS PSWMRU, Building 3702 Curtin Rd., University Park, PA 16802-3702; phone: 814-863-0888; fax: 814-863-0935; e-mail: tamie.veith@ars.usda.gov.

designed to elucidate local conditions as well as to characterize representative soil responses have greatly increased scientific knowledge of nonpoint-source pollution control at field and hillslope scales. At these scales, it is possible to collect detailed data and control various aspects of the system. With expansion to farm and watershed scales (the scales at which land management decisions are made and conservation practice placements determined), comprehensive, detailed data collection and control of the system become time-consuming, expensive, and increasingly impractical. Thus, many modeling efforts must use lower resolution input data than desired or adopt input data and calibration values from regions with similar characteristics. Some studies go a step further in customizing adopted data by using fuzzy logic to improve their location-specific studies based on nearby data (e.g., Mertens and Huwe, 2002; Schlüter and Rüger, 2007). The difficulties in obtaining comprehensive data for a particular representation of a watershed system, in conjunction with the numerous other sources of uncertainty inherent in modeling a natural system, raise the question: how spatially and temporally detailed can land management data become before the increase in resolution is no longer efficient?

Several studies have shown an effect of subbasin delineation on water and sediment yield. For example, White and Indrajeet (2005) compared effect of land use resolution on flow and sediment yield within 2,580 ha and 4,740 ha subbasins of the Illinois River watershed by adjusting SWAT's hydrologic response unit threshold. Among the modeled variations, they found minimal differences in prediction of flow but significant differences in sediment prediction, although not in clear correlation with degree of land use resolution. Others have modeled field-level (<1.5 ha) research plots with SWAT. Maski et al. (2006) confirmed that field-specific calibration based on crop-type, tillage, and soil improved comparisons with field-level measured flow and sediment, although sediment still tended to be underpredicted. Choi et al. (2005) successfully calibrated and validated SWAT to match measured flow, sediment, nitrogen (N), and P losses from two turfgrass plots. They used different USLE C-factors to represent different levels of turfgrass production between the two fields. Bosch et al. (2004) found that a higher resolution of input data improved SWAT hydrologic modeling of a 22.1 km^2 coastal plain watershed in Georgia. They considered three degrees of resolution: low resolution with default SWAT input parameters, high resolution with default SWAT input parameters, and high resolution with calibration. The low-resolution scenarios used 90 m DEMs with state-level land use and soil information. The high-resolution scenarios used 30 m DEMs, field-level land use surveys and digital orthophotos, and county-level soil data. They concluded that the higher resolution scenarios better represented nutrient mobilization, transport within, and export from the watershed.

Computer and internet technology is rendering increasingly detailed levels of public data (topography, climate, soils, land use) obtainable to researchers through government and university clearinghouses. Thus, modeling at the degree of detail used by Bosch et al. (2004) in their high-resolution scenarios is becoming reasonably efficient in the tradeoff between gathering spatially and temporally detailed data versus improving model prediction. However, modeling at an even more spatially and temporally explicit level than done by Bosch et al. (2004) requires in-depth familiarity of practices within the watershed and process knowledge gained through long-term, intensive research. Gathering and using such data in modeling is time and labor intensive. Is this additional effort beneficial for watershed-level, water quality control?

The current study uses a small agricultural watershed in which land management activities are controlled by farmers, not researchers. The study objective was to determine if using high-resolution crop management data in small-scale (<100 ha) watershed-level modeling provides a more realistic representation of the watershed than using low-resolution crop management data. Two resolutions of land use management were simulated in SWAT: a high-resolution scenario in which multi-year, field-specific management was used for each crop field; and a low-resolution scenario in which all crop fields were modeled as generic row crops, independent of year. Flow depth, sediment concentration, and P concentration at the outlet were compared across scenarios and with measured data to assess the ability of one SWAT scenario over the other to more accurately predict watershed outlet responses in time. Losses from individual fields within the watershed were evaluated between the SWAT scenarios, in consideration with known processes and characteristics internal to the watershed, to assess the apparent ability of either SWAT scenario over the other to more accurately portray spatial losses within the watershed.

Method

The study site, FD-36, is a 39.5 ha subwatershed of Mahantango Creek, a tributary of the Susquehanna River and ultimately the Chesapeake Bay (fig. 1). The encompassing Mahantango Creek watershed is a long-term study site of the USDA-ARS Pasture Systems and Watershed Management Research Unit, with 30 years of climatic and hydrologic data and a progression of research in runoff generation and nutrient movement for this region (Pionke et al., 1999). Watershed FD-36 is characterized by loamy soils with a substantial near-stream fragipan; average slope is 9% (fig. 2) (Needelman et al., 2004). Long-term climate in this region is temperate and humid, with average rainfall of 1100 mm $year^{-1}$ and stream flow of 450 mm $year^{-1}$. Growing season averages of rainfall and runoff from 1997 to 2000 were about 60% and 45%, respectively, of the annual averages (Srinivasan et al., 2005).

Long-term interactions and annual interviews with the farmers in the FD-36 watershed, experienced on-site technicians, and periodic sub-field research experiments have provided a unique wealth of spatial and temporal information about management and hydrologic process within the watershed. However, the watershed is not a component of a controlled, recorded, research project. While farmers in the watershed have cooperated readily with USDA-ARS researchers, their management methods are ultimately based on personal goals as opposed to being dictated by a research project.

The watershed encompasses 24 row-cropped fields from three contiguous farms. The fields range in size from 0.2 ha to 2.0 ha and account for about 50% of the land use within FD-36; the remainder is split between forest (30%) and

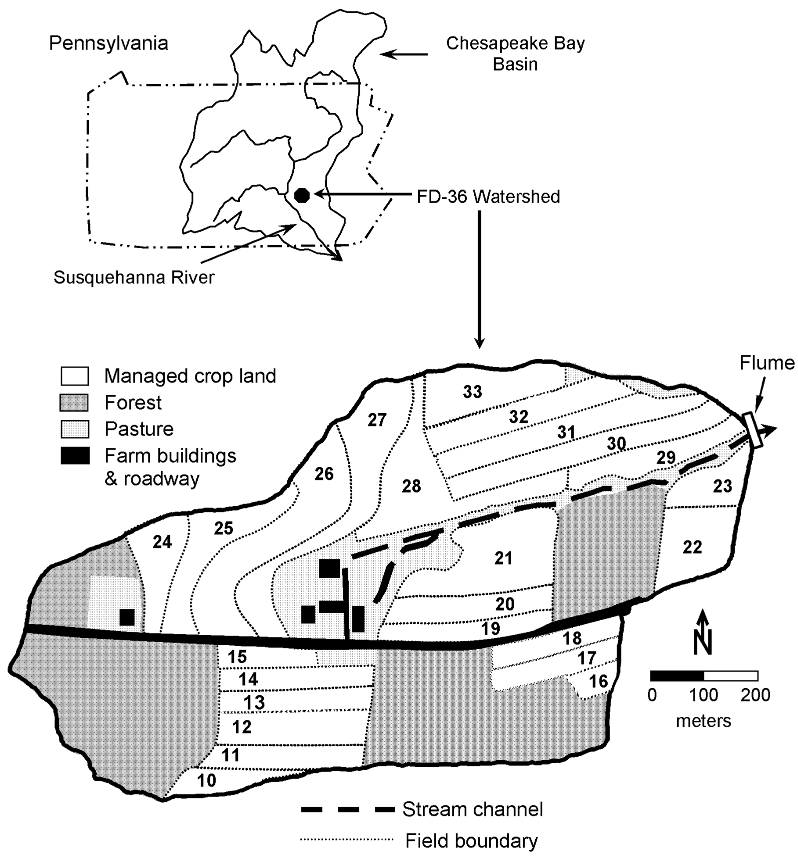

Figure 1. Watershed FD-36, with field boundaries and identification numbers.

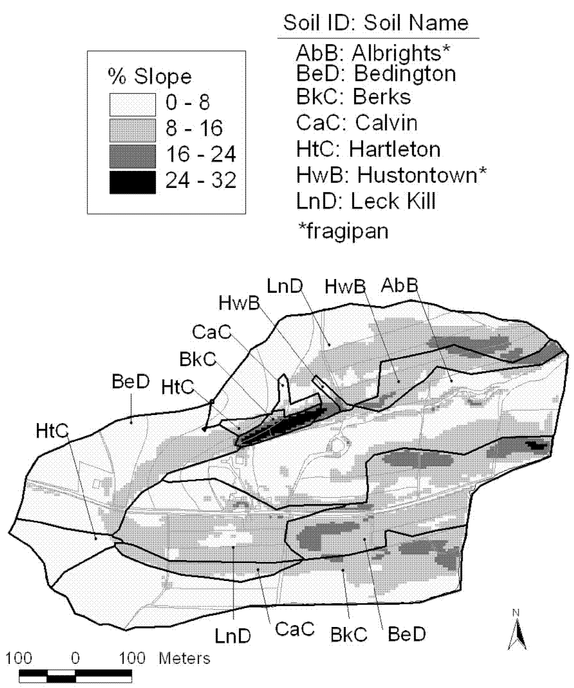

Figure 2. Placement of slopes and soils within watershed FD-36.

pasture (20%) (fig. 1). Management practices of individual fields over an 11-year period (1994-2004) were obtained from annual farmer surveys (table 1). The surveys included tillage, fertilizer, plant, and harvest dates and methods for each crop.

Watershed FD-36 was simulated in SWAT2003 (Arnold et al., 1998; Neitsch et al., 2005) from 1994 to 2004 in as much detail as provided by measured data. Detailed land use and management information was provided by the farmer surveys and the observations by USDA-ARS watershed technicians stationed on-site; the SWAT management files were individually modified to model field-specific practices and timings from year to year throughout the simulation period. Five-meter elevation grids, surveyed and digitized in July 1996, were used. Thirty-meter grid soil properties were taken from sampling and analysis completed in March 1998 and 2000, as described by Veith et al. (2005). Sub-hourly precipitation records and daily measurements for humidity, solar radiation, wind speed, and air temperature were collected from the FD-36 research weather station just north of the watershed outlet sampling site. Evapotranspiration was simulated by SWAT using the Penman-Monteith method (Singh, 1988), while runoff and infiltration separations were simulated using the curve number (USDA-SCS, 1972). Stream flow rates, sediment, and P measurements continued to be collected and analyzed, as described by Veith et al. (2005), throughout 2004.

SWAT employs physically based equations on a daily time step to predict water, sediment, and nutrient losses at the watershed outlet and from hydrologic response units within a watershed. A total of 123 hydrologic response units in FD-36, ranging in size from 25 m^2 to 3 ha, were defined based on unique soil groups within each field. These small unit divisions enable a one-to-one spatial mapping between each subfield unit of the watershed and the response of the associated SWAT hydrologic response unit. Soil and

Table 1. Land use and P management of the 24 row-cropped fields in watershed FD-36 for 2000.

Field No.[a]	Field Area (ha)	Crop	Tillage[b]	Fertilizer P Applied (kg ha^{-1})	Fertilizer Application (method/month)	Manure P Applied (kg ha^{-1})	Manure Application (method/month)	Mehlich-3 Soil P[c] (mg kg^{-1})
10	0.42	Wheat	MTM	34	Broadcast/April	0		328
11	0.70	Barley	MP/DH	7	Broadcast/March	0		220
12	0.93	Corn	MP/DH	24	Broadcast/April	0		222
13	0.62	Pasture	MP/DH	0		0		208
14	0.62	Corn	MP/DH	24	Broadcast/April	0		204
15	0.36	Corn	MP/DH	24	Broadcast/March	0		194
16	0.22	Corn	MP/DH	24	Broadcast/April	0		266
17	0.55	Barley	MP/DH	7	Broadcast/April	0		251
18	0.53	Corn	MP/DH	24	Broadcast/April	0		289
19	0.62	Corn	MTM	32	Broadcast/April	0		291
20	0.77	Oats	MP/DH	0		0		212
21	1.63	Corn	MP/DH	24	Broadcast/April	0		113
22	1.00	Soybean	C/DH	0		0		124
23	0.61	Soybean	C/DH	0		0		73
24	0.79	Corn	MTM	32	Broadcast/October	0		205
25	1.06	Wheat	MTM	66	Broadcast/April	0		267
26	2.00	Wheat	MTM	34	Broadcast/October	0		276
27	1.83	Soybean	MTM	34	Broadcast/April	0		147
28	1.65	Corn	C/DH	32	Broadcast/October	0		94
29	0.80	Corn	C/DH	0		86	Broadcast/May	225
30	1.26	Wheat	C/DH	0		51	Broadcast/April	181
31	1.24	Corn	C/DH	0		86	Broadcast/May	330
32	1.06	Soybean	C/DH	0		0		213
33	1.07	Corn	C/DH	0		86	Broadcast/May	350

[a] Refer to figure 2.
[b] MTM is minimum tillage mulchmaster; MP/DH is moldboard plow / disc harrow; and C/DH is chisel / disc harrow.
[c] Field averaged Mehlich-3 extractable soil P measured on a 30 m grid.

topographic properties were uniform within each hydrologic response unit but varied across units according to the respective input data. Changes in field management information remained at the field level, reflecting the reality of farmer operations.

SWAT was used to simulate hydrologic, sediment, and P response from the hydrologic response units and watershed on a daily basis. These responses were compared to corresponding measured data, which were compiled from sub-hourly into daily values. Manual calibration parameter values for FD-36 during 1997-2000, as detailed by Veith et al. (2005), were used. During calibration, three years were used as an initial state "warm-up" period (1994-1996). Due to record keeping protocols, these years contained detailed, but less complete, field management data than following years. In simulations shown in this article, SWAT is given the 1996 year for establishing state variables. Scenario evaluation and comparison results are drawn from the subsequent 8-year period (1997-2004).

Despite use of identical calibration parameter values, this study's 1997-2000 simulation results vary somewhat from those reported by Veith et al. (2005). This article uses SWAT2003 instead of SWAT2000. Additionally, several adjustments in daily precipitation files were made to more accurately represent multi-day storms, and all weather files were corrected for instrumentation errors. For example, when a storm spanned midnight, the storm was coded in the precipitation file as occurring entirely during the second day. This allowed SWAT to apply the entire storm volume to the watershed in a single day instead of over two separate days, more accurately capturing total volume and reflecting actual intensity of the storm. A similar technique was used by Maski et al. (2006). For the occasional instrumentation error, data from adjacent days with otherwise similar weather were used. All management files were carefully reviewed for accuracy, and necessary modifications were made to more accurately represent the timing of the management practices in conjunction with storm events. For example, a fertilization event reported by the farmer as occurring in mid-April may have initially been entered into the SWAT management file as occurring on April 15th. If storm events between April 13th and 15th resulted in excessive soil moisture, based on that field's soil type, fertilization was delayed in SWAT until the first realistic day for application.

To evaluate impacts of levels of input data, two scenarios were simulated in SWAT and compared to each other and to measured data. In the first SWAT scenario, the FD-36 watershed was represented in as much temporal and spatial detail as possible, using the data previously described in this section. In particular, crop-dependent field management practices were specified for each field throughout the simulation period. In the second scenario, the watershed was represented in as much detail as possible in all aspects except for the row-cropped fields. These fields were assigned a generic row-crop land use (designated AGRR in SWAT) throughout the simulation period with daily heat units, as determined by SWAT, governing crop growth. The first scenario required intensive information about the watershed; the second scenario was intended to represent a more typical modeling situation in which plant, fertilizer, and equipment data for specific fields and specific years are not known, but watershed topography, soils, climatic records, and general land use are known.

For the first scenario, this study added to the intensity of the high-level resolution scenario definition used by Bosch et al. (2004) in three ways. The hydrologic response unit threshold was set to 0% so that each unit's spatial location could be uniquely identified within the watershed. Land management activities were changed continuously throughout a single simulation as opposed to using multiple simulations to handle land use changes. Initial model input parameters were calibrated for the most detailed scenario.

Results were evaluated at the watershed outlet visually via daily time series graphs and statistically using Minitab's Individual Distribution Identification tool and Mann-Whitney nonparametric test for equality of medians (Minitab v15.1.1.0, www.minitab.com, State College, Pa.). Additionally, three statistical measures were calculated on a daily and monthly level: Nash-Sutcliffe (NS; Nash and Sutcliffe, 1970), coefficient of determination (R^2; Zar, 1984), and percent bias (Pbias; Martinec and Rango, 1989). Positive values of both NS and R^2 indicate positive correlations between predicted and measured values; higher values, up to the maximum value of 1, are preferred. In this study, the Pbias was calculated as 100 × (Σpredicted - Σmeasured) / (Σmeasured), such that positive percentages represent an overprediction by SWAT. In addition, average annual results for each hydrologic response unit for both the detailed SWAT project's initial calibration period (1997-2000) and a validation period (2001-2004) were displayed spatially on the watershed map and evaluated visually.

Results and Discussion

Flow Depth

High- and low-resolution predictions and measured values for FD-36 were evaluated across the 7-month growing periods (April-October) of each year, and analyses were calculated for two 4-year periods, 1997-2000 and 2001-2004, as well as the full 8-year period. During each 4-year period, the first year was the driest (table 2). In 1997, FD-36 received 523 mm of rainfall, 194 mm (27%) less than the 1998-2000 average of 717 mm. The 461 mm of 2001 rainfall were 318 mm (41%) less than the 2002-2004 average of 779 mm. The 8-year average growing period rainfall (684 mm) and measured runoff depth (167 mm) corresponded to the April-November averages of 694 mm and 169 mm, respectively, observed by Srinivasan et al. (2005), as did the stream-to-precipitation ratio of 0.24. Both measured and predicted runoff depths fluctuated yearly in correspondence with rainfall depths (table 2). Annual runoff predictions between the high- and low-resolution SWAT scenarios varied by less than 10 mm. Although daily predicted and measured peaks occurred on the same days, SWAT generally underpredicted measured values in the spring and early summer and overpredicted in the late summer and fall (fig. 3). Runoff-

Table 2a. Measured and SWAT-predicted rainfall and runoff from FD-36 over the 7-month sampling period (April through October).

Year	Total Rainfall (mm)	Total Runoff (mm) Measured	Predicted[a] High	Predicted[a] Low
1997	523	73	104	94
1998	805	211	276	275
1999	650	114	177	173
2000	696	167	236	239
2001	461	72	85	81
2002	740	131	144	145
2003	785	317	326	331
2004	813	247	316	322

[a] "High" indicates high-resolution scenario in which multi-year, field-specific management is used for each row crop field; "Low" indicates low-resolution scenario in which all crop fields are modeled as generic row crops, independent of year.

Table 2b. Measured and SWAT-predicted water quality concentration and losses from FD-36 over the 7-month sampling period (April through October).

Year	Sediment Measured	Sediment Predicted[a] High	Sediment Predicted[a] Low	Dissolved P Measured	Dissolved P Predicted[a] High	Dissolved P Predicted[a] Low	Particulate P Measured	Particulate P Predicted[a] High	Particulate P Predicted[a] Low	Total P Measured	Total P Predicted[a] High	Total P Predicted[a] Low
Mean Concentrations (mg L^{-1})[b]												
1997	355	168	193	0.03	0.02	0.02	0.12	0.24	0.09	0.15	0.25	0.11
1998	209	223	286	0.08	0.08	0.02	0.13	1.54	1.03	0.22	1.62	1.04
1999	95	197	220	0.06	0.04	0.02	0.06	0.51	0.27	0.12	0.56	0.28
2000	311	94	176	0.05	0.02	0.01	0.26	0.15	0.23	0.31	0.17	0.24
2001	520	81	138	0.07	0.01	0.02	0.24	0.10	0.08	0.31	0.11	0.11
2002	345	105	234	0.06	0.02	0.02	0.23	0.17	0.06	0.29	0.19	0.08
2003	206	76	155	0.08	0.02	0.01	0.12	0.14	0.16	0.20	0.17	0.17
2004	322	133	192	0.05	0.05	0.01	0.15	0.36	0.33	0.20	0.41	0.35
Losses (kg ha^{-1})												
1997	261	175	182	0.02	0.02	0.02	0.09	0.24	0.08	0.11	0.26	0.10
1998	441	615	785	0.18	0.23	0.05	0.28	4.25	2.82	0.45	4.48	2.87
1999	109	349	381	0.07	0.08	0.03	0.07	0.91	0.46	0.14	0.99	0.49
2000	519	223	419	0.09	0.05	0.03	0.44	0.36	0.54	0.52	0.41	0.57
2001	373	69	112	0.05	0.01	0.02	0.17	0.09	0.07	0.22	0.10	0.09
2002	452	151	340	0.08	0.03	0.03	0.30	0.25	0.09	0.38	0.27	0.12
2003	654	248	515	0.24	0.07	0.04	0.38	0.47	0.52	0.62	0.54	0.56
2004	796	420	618	0.12	0.15	0.05	0.36	1.14	1.07	0.48	1.29	1.11

[a] "High" indicates high-resolution scenario in which multi-year, field-specific management is used for each row crop field; "Low" indicates low-resolution scenario in which all crop fields are modeled as generic row crops, independent of year.
[b] Concentrations are flow-weighted using the "Flow" columns, which includes both storm and base flow.

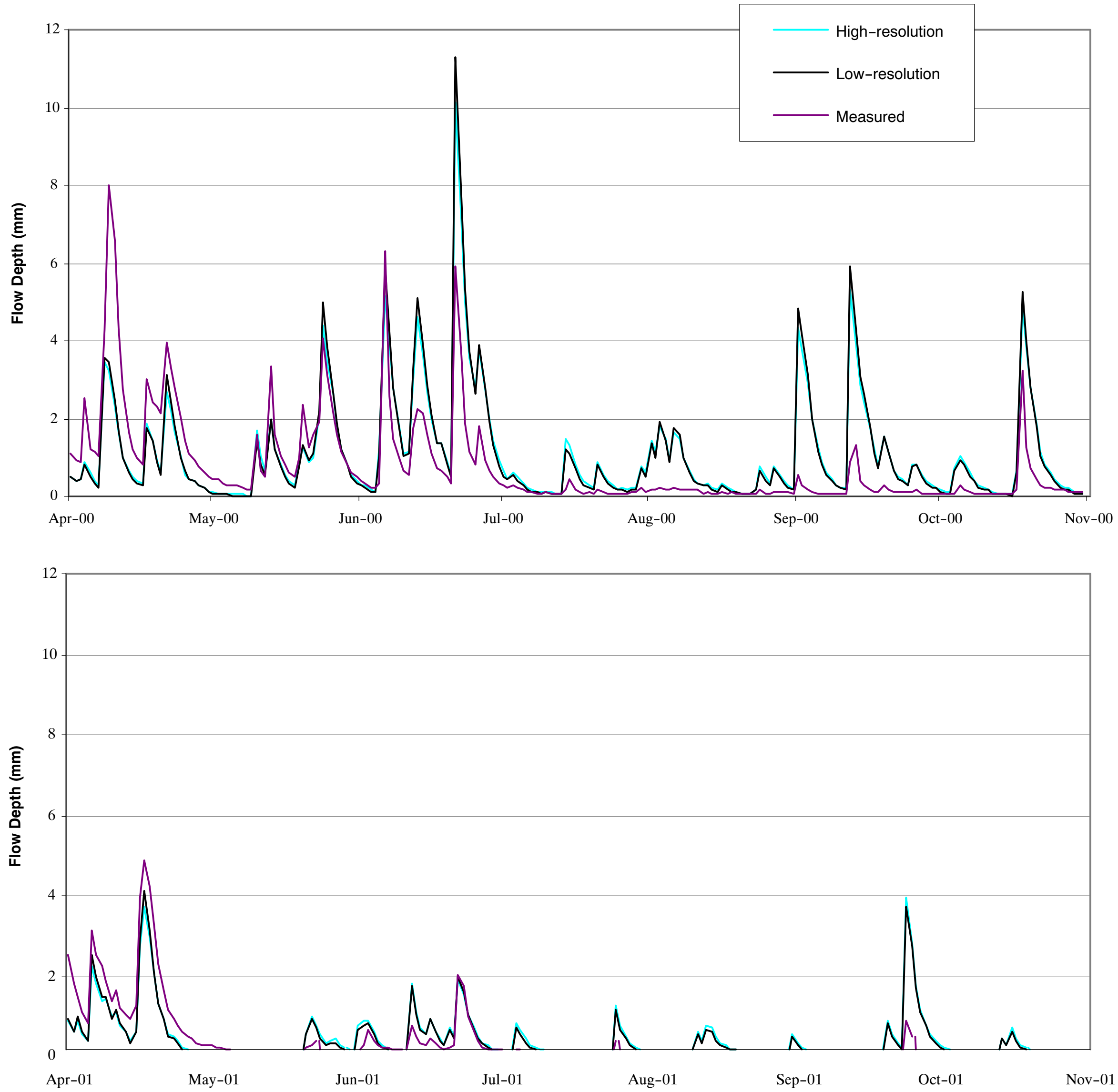

Figure 3. Daily measured and SWAT-predicted stream flow for two years at the outlet of FD-36. SWAT predictions are shown using high- and low-resolution data. High-resolution uses detailed temporal management information for each crop field; low-resolution depicts all crop fields as general row crop.

generation studies in this region have found that runoff responses for larger rainfall depths (>13 mm) tend to depend more on pre-storm watershed moisture conditions than on storm intensities (Gburek et al., 2002; Srinivasan et al., 2002). This suggests that SWAT may not represent soil moisture as accurately as needed to predict peak flows throughout the year.

Statistically, monthly flow depths for both the high- and low-resolution SWAT scenarios followed a similarly shaped gamma distribution with $p = 0.100$, where the larger the value of p the greater the likelihood that they fit that distribution. Both scenarios fit the Weibull distribution slightly more closely with $p = 0.125$. The SWAT scenarios fit these two distributions better than 12 other distributions included in Minitab's Individual Distribution Identification tool. The monthly measured data fit a gamma distribution more strongly than a Weibull distribution ($p = 0.199$ vs. $p = 0.056$). The properties of both the gamma and Weibull distributions lend their practicality for use with storm flow data (Haan, 1977). None of the three flow data sets were normally distributed. Using the Mann-Whitney non-parametric test, the two SWAT scenarios can be considered to be from the same distribution with $p = 0.88$. The probabilities that low- and high-resolution scenarios came from the same distribution as the measured data were $p = 0.13$ and $p = 0.11$, respectively. In contrast to evaluation of flow data on a monthly basis, when daily flow data for the 8-year period was evaluated, Minitab was not able to fit a distribution to any of the scenarios (no p-values were > 0.01), and the Mann-Whitney test determined no pair of scenarios to be significantly different (all p-values were < 0.04).

Supporting these statistical findings, monthly NS and R^2 values between the two SWAT scenarios were 1.00 for all three time-period groupings (1997-2000, 2001-2004, 1997-2004). There was a slight bias towards higher monthly flow predictions by the high-resolution scenario during 1997-2000 (Pbias = -1.5) and by the low-resolution scenario during 2001-2004 (Pbias = 0.9); these biases are confirmed by the annual prediction patterns between scenarios (table 2). Compared to the measured data, the two SWAT scenarios

Table 3. Nash-Sutcliffe (NS), R^2, and percent bias (Pbias) values between measured and SWAT-predicted values at the outlet of FD-36 over the 7-month sampling period (April through October).

	Data Resolution	Flow Volume: Monthly			Flow Volume: Daily			Sediment Conc.: Monthly		
		NS	R^2	Pbias	NS	R^2	Pbias	NS	R^2	Pbias
1997-2000	High[a]	0.4	0.56	40.2	0.49	0.62	40.2	-0.04	0.15	-44.5
	Low	0.39	0.55	38.2	0.47	0.62	38.2	-0.02	0.04	-20.1
2001-2004	High[a]	0.43	0.58	13.6	0.58	0.64	13.6	-1.14	0	-72.8
	Low	0.42	0.59	14.7	0.57	0.65	14.7	-0.46	0.01	-44.9
1997-2004	High[a]	0.42	0.56	24.9	0.54	0.62	24.9	-0.51	0.04	-60.9
	Low	0.41	0.56	39.6	0.52	0.64	24.6	-0.19	0.02	-34.5

[a] "High" indicates high-resolution scenario in which multi-year, field-specific management is used for each row crop field; "Low" indicates low-resolution scenario in which all crop fields are modeled as generic row crops, independent of year.

performed similarly with regard to NS, R^2, and Pbias (table 3) at both daily and monthly levels.

The greatest impact between high- and low-resolution scenarios was seen spatially, at the field level. Both scenarios showed higher levels of soil water in the fragipan soils than in the surrounding soils (figs. 2 and 4). This is an important verification of the SWAT processes at the hydrologic response unit level. Surface and shallow subsurface hydrologic process within FD-36 have been shown to be dominated by soil type (colluvial versus residual) even more so than by land management (Needelman et al., 2004) with runoff generation controlled by saturation-excess over the fragipans (Gburek et al., 2006). Correspondingly, predicted soil water depth was relatively uniform by soil type, higher over the fragipan and lower in the well-drained soils (fig. 4). Defining varied management impacted cropland percolation more widely than did specifying a single general land use (fig. 5). For example, field 28 (see figs. 1 and 2) flows toward the stream with a long, narrow sloped region of 30% at the field's lower edge, paralleling the stream. Site observation has confirmed that this field contributes notable surface runoff and erosion to the stream, leaving little soil water for evapotranspiration and percolation (McDowell et al., 2001). This response was represented under the high-resolution scenario but only for the intensely sloped portion of the field under the low-resolution scenario (figs. 5, 6, and 8). Variations in total surface and shallow subsurface flow leaving each field (fig. 6) showed distinct dependence on land management. In the high-resolution scenario, runoff from each field varied from neighboring fields, clearly defining field boundaries when mapped using a few more legend intervals than shown. This mimics the observed situation, as every field follows a unique cropping pattern throughout the 8-year period.

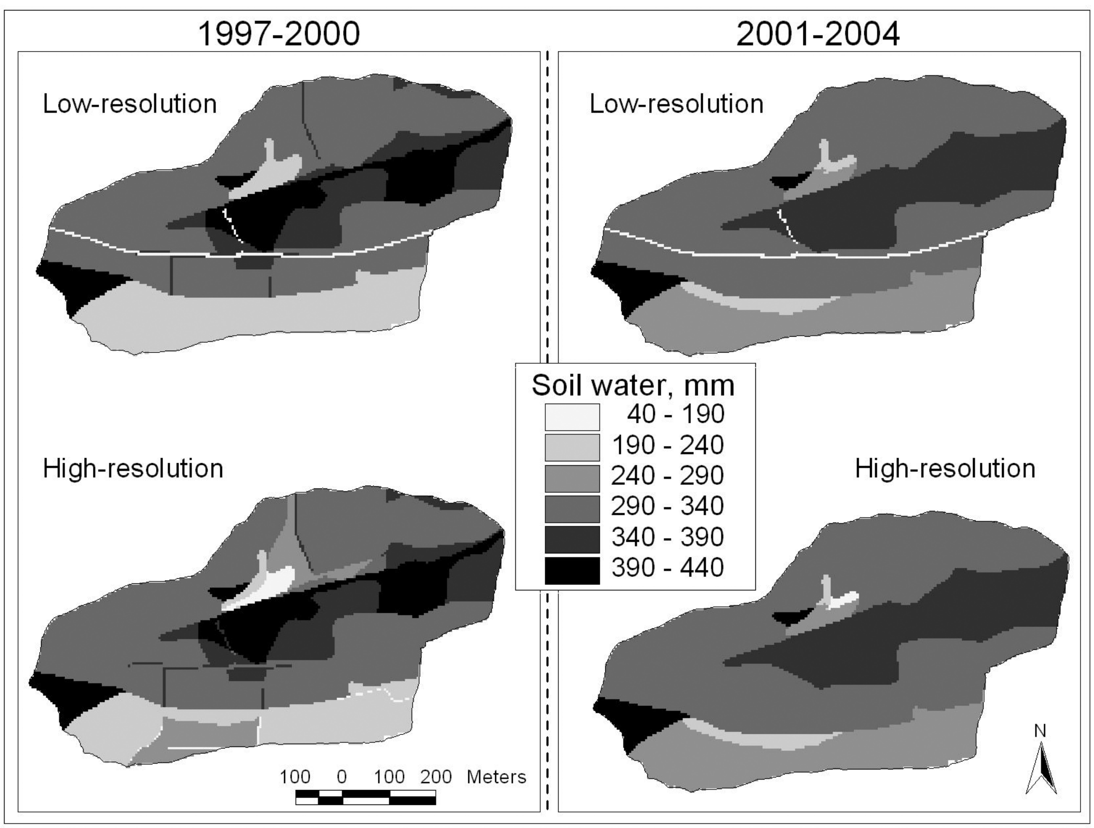

Figure 4. Average annual soil water depth (mm) for two levels of land management distinction in SWAT, split by the SWAT calibration (1997-2000) and validation periods (2001-2004).

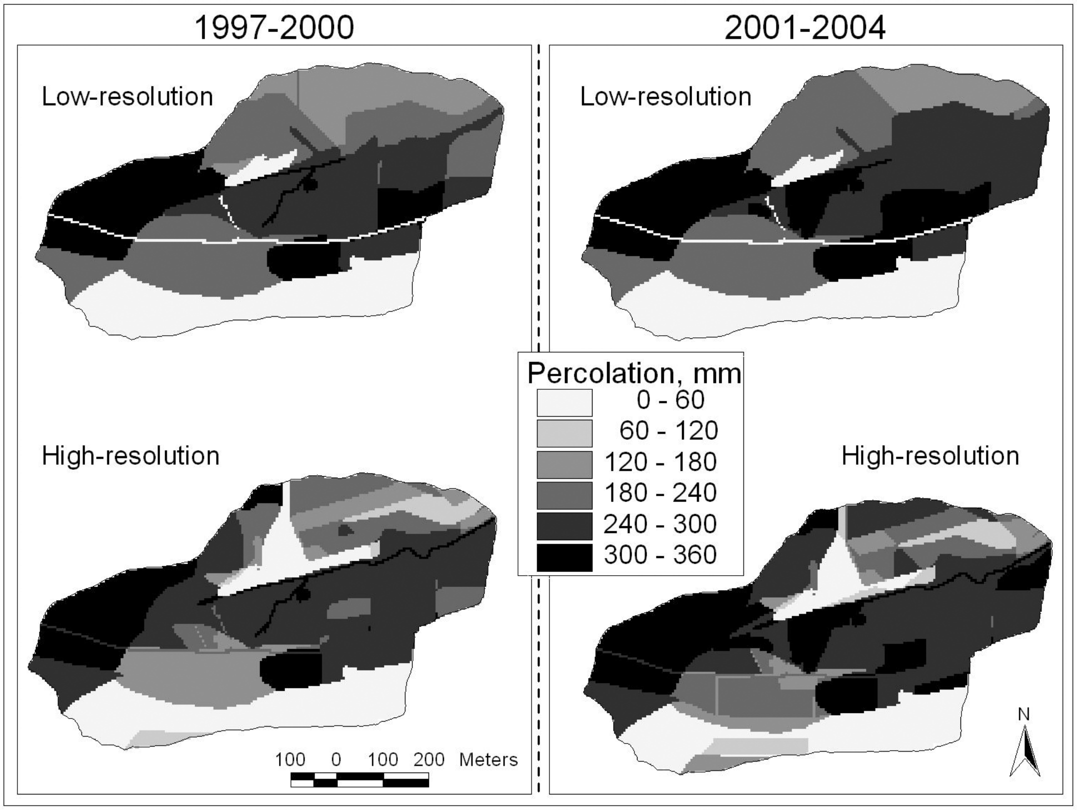

Figure 5. Average annual amount of water percolating out of the root zone (mm H_2O) for two levels of land management distinction in SWAT, split by the SWAT calibration (1997-2000) and validation periods (2001-2004).

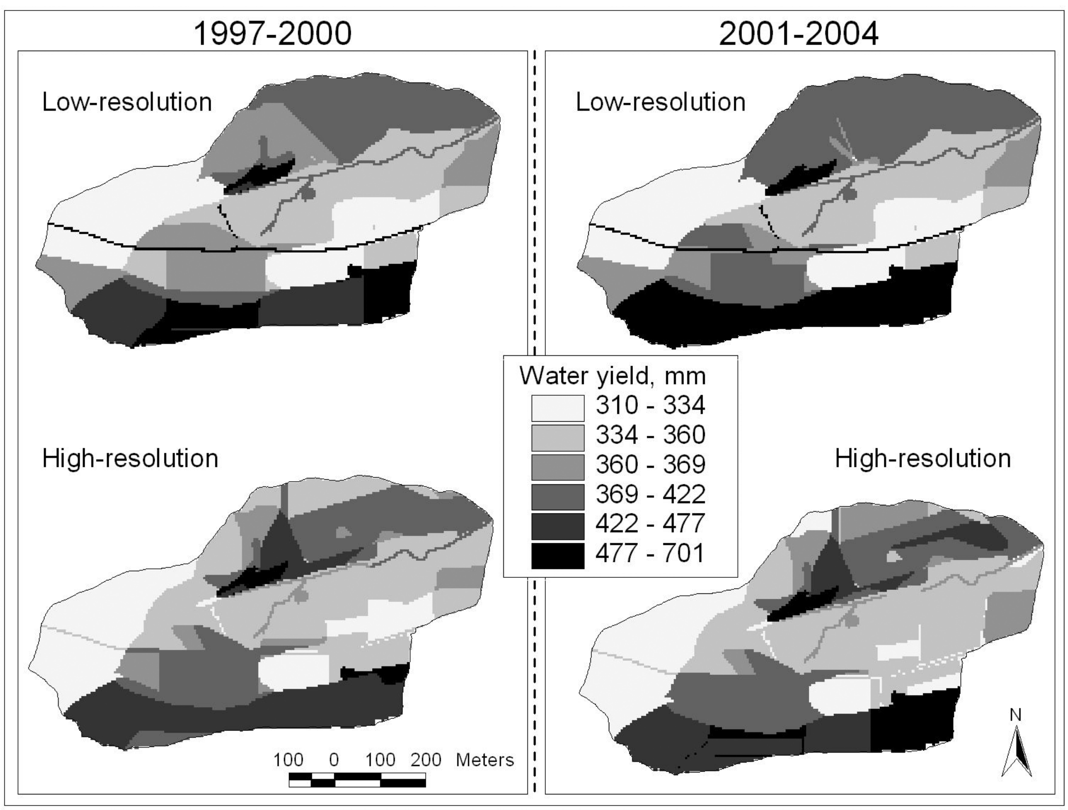

Figure 6. Average annual net water yield to stream (mm) for two levels of land management distinction in SWAT, split by the SWAT calibration (1997-2000) and validation periods (2001-2004).

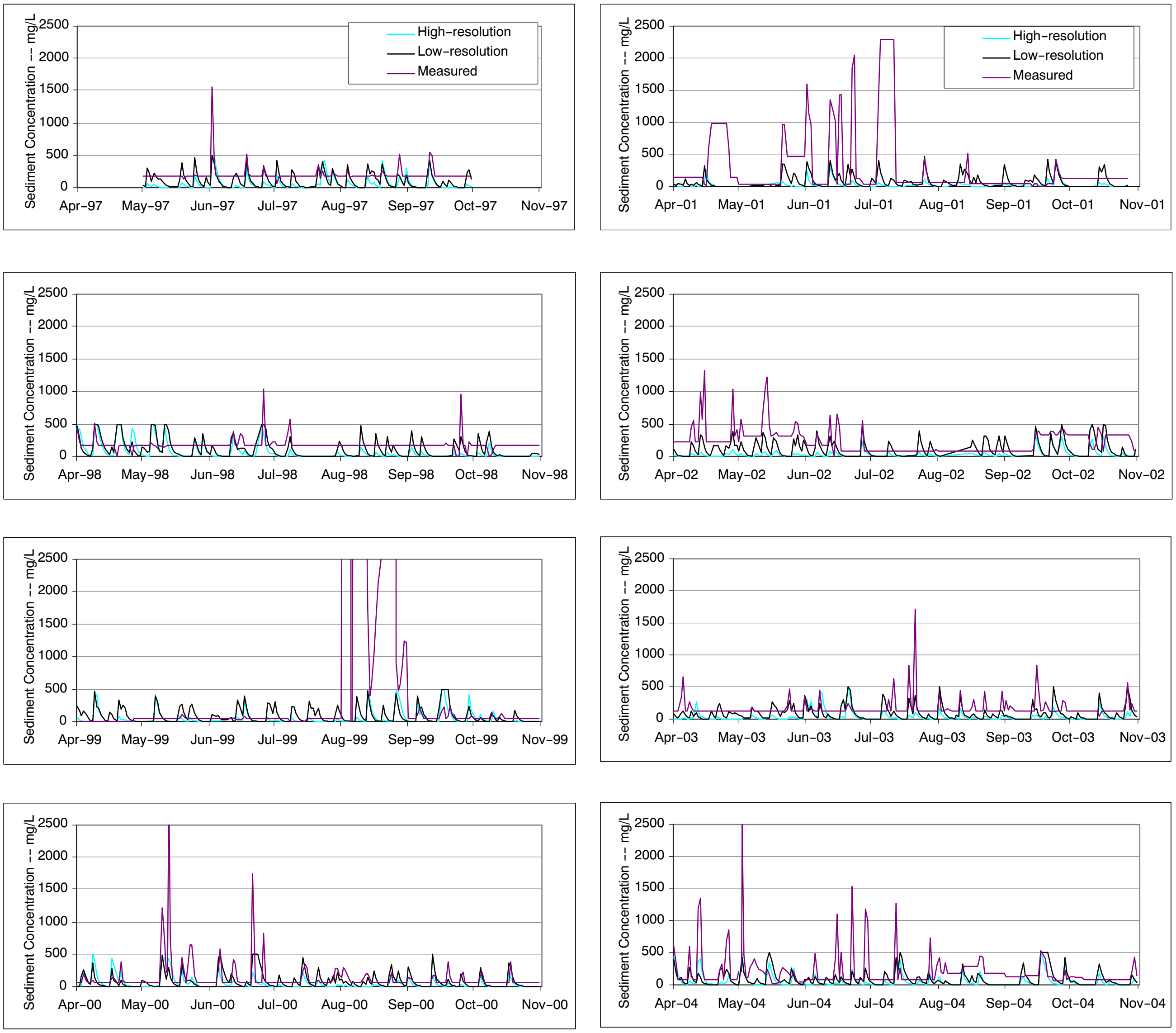

Figure 7. Daily measured and SWAT-predicted sediment concentration for the study period at the outlet of FD-36. SWAT predictions are shown using high- and low-resolution data.

Sediment Concentration

Measured sediment loss over the 8-year period averaged 451 kg ha^{-1} year^{-1} compared with a high-resolution predicted loss of 281 kg ha^{-1} year^{-1} and a low-resolution predicted loss of 419 kg ha^{-1} year^{-1} (table 2). Thus, high- and low-resolution scenarios estimated sediment losses of 6.7 and 1.3 t year^{-1} less, respectively, from the 39.5 ha watershed than was measured. Accordingly, event peaks for sediment concentration were generally lower for the high- than low-resolution scenario (fig. 7). Simulated values were much lower than the measured values during base flow (measured monthly), large runoff events, and spring flush events. During some years, SWAT scenarios predicted the mid-flow-range sediment-concentration event graphs accurately (1997, 2000), whereas in other years (1999, 2001), the difference between measurement and prediction was substantial (fig. 7). Sediment predictions appeared to match measured data more closely during the 1997-2000 calibration period than the 2001-2004 validation period, as expected.

The two SWAT scenarios differed enough in magnitude of event peaks that they could not be considered to share a common statistical distribution with each other or with the measured data at either the monthly or daily level (Mann-Whitney $p < 0.05$ in all pairwise comparisons). Accordingly, NS, R^2, and Pbias values for sediment concentration were not helpful in determining goodness-of-fit among the three scenarios (table 3). Considering the daily time series, timing of event peaks was typically well matched between the SWAT and measured scenarios (fig. 7). However, a difference of a few days in predicted versus actual management can make an immediate and extreme impact on erosion predictions despite careful use of detailed land management and climate information. This was particularly the case for spring and fall tillage operations, which was when the most notable differences in predicted and measured sediment loss occurred.

The high-resolution scenario predicted larger sediment losses from fields with greater slopes and fields on the

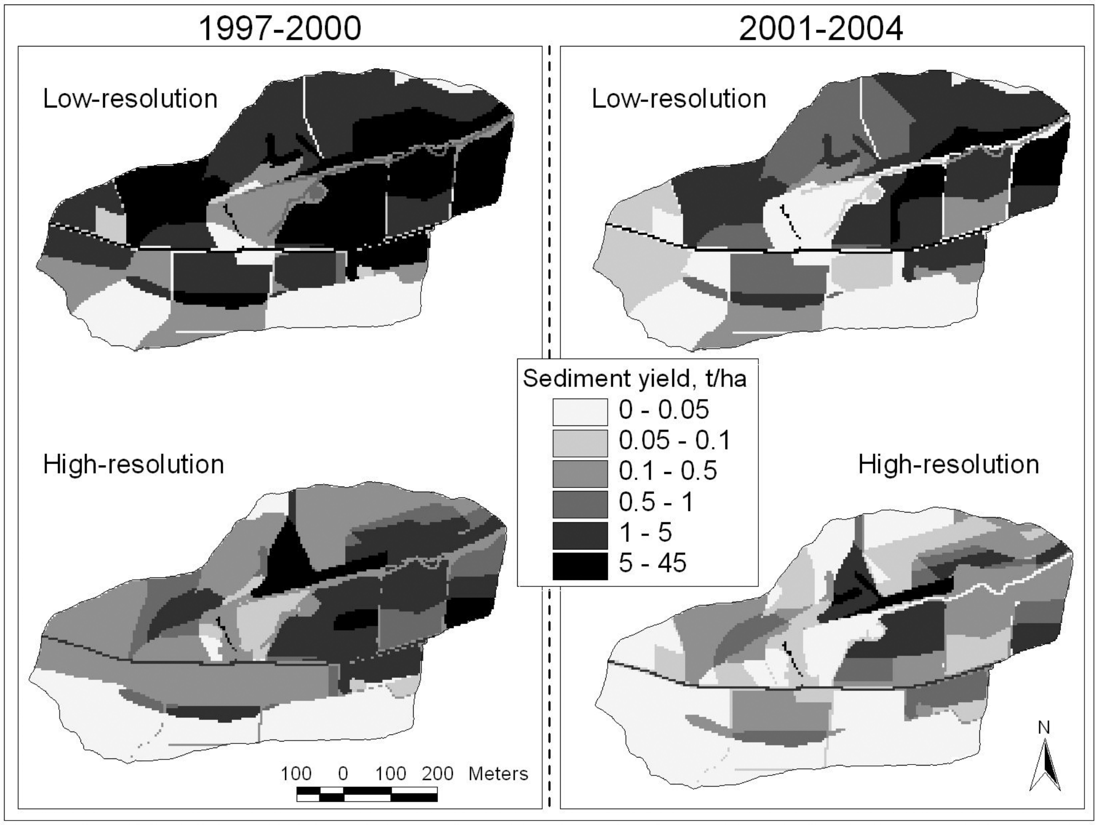

Figure 8. Average annual sediment yield to stream (t ha^{-1}) for two levels of land management distinction in SWAT, split by the SWAT calibration (1997-2000) and validation periods (2001-2004).

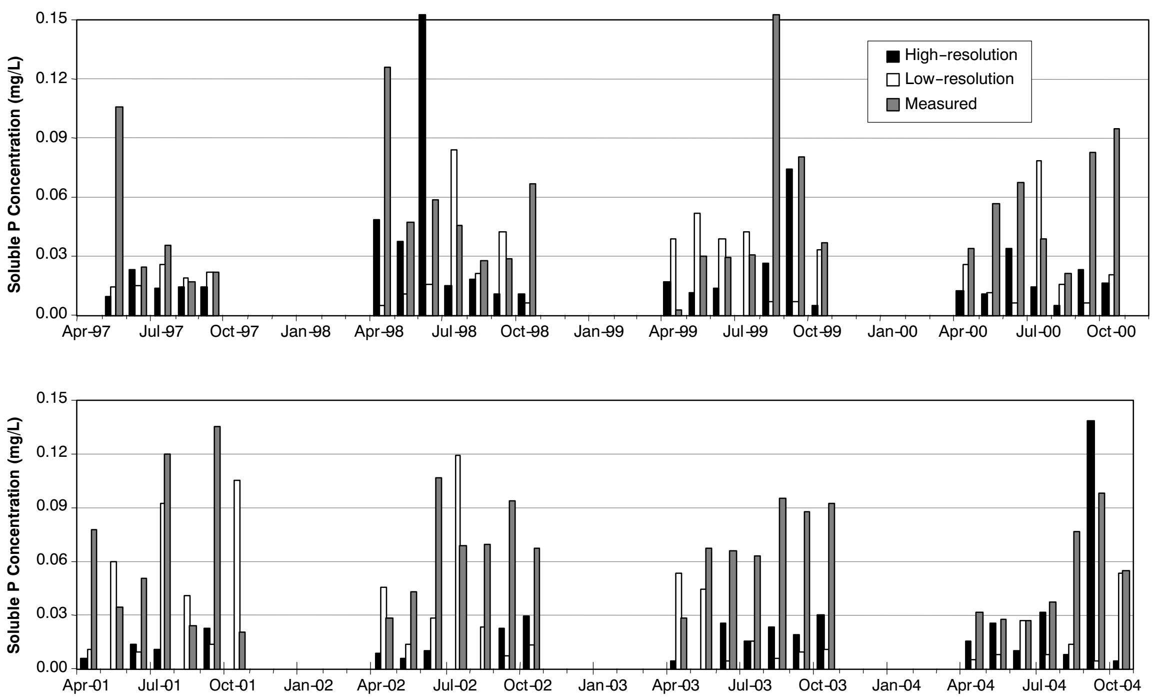

Figure 9. Monthly measured and SWAT-predicted soluble phosphorus concentrations for the study period at the outlet of FD-36. SWAT predictions are shown using high- and low-resolution data.

fragipan than from fields outside the fragipan, which have a higher infiltration rate (fig. 8). The low-resolution scenario generally predicted higher sediment losses than did the high-resolution scenario but showed the same trend of larger sediment losses from fields on versus outside the fragipan.

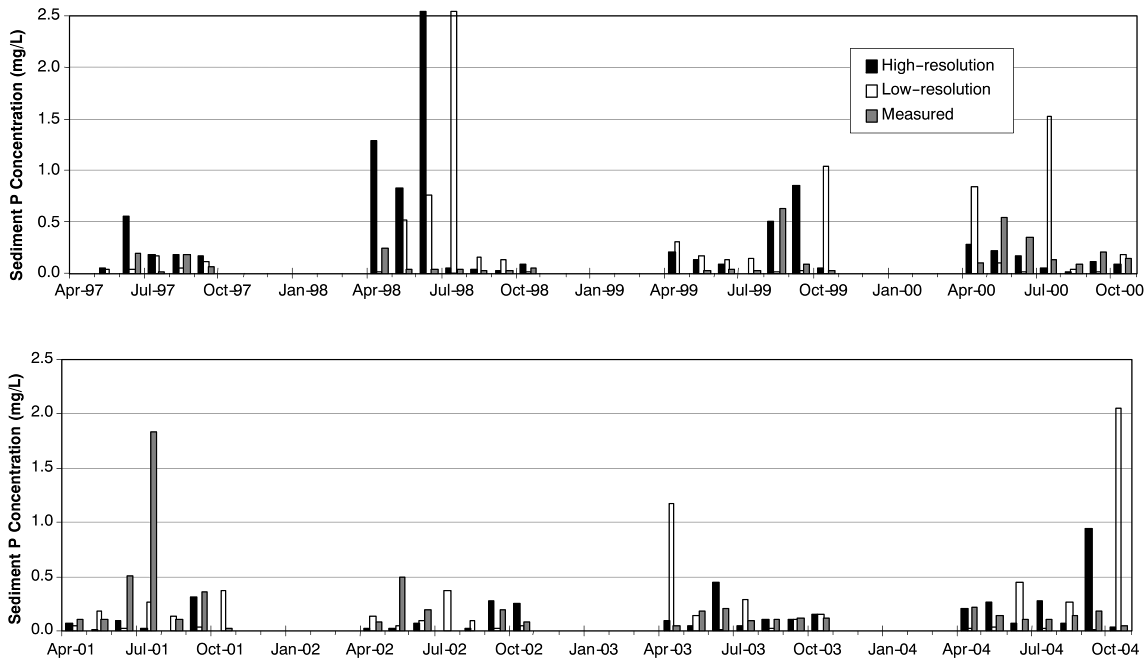

Figure 10. Monthly measured and SWAT-predicted sediment phosphorus concentrations for the study period at the outlet of FD-36. SWAT predictions are shown using high- and low-resolution data.

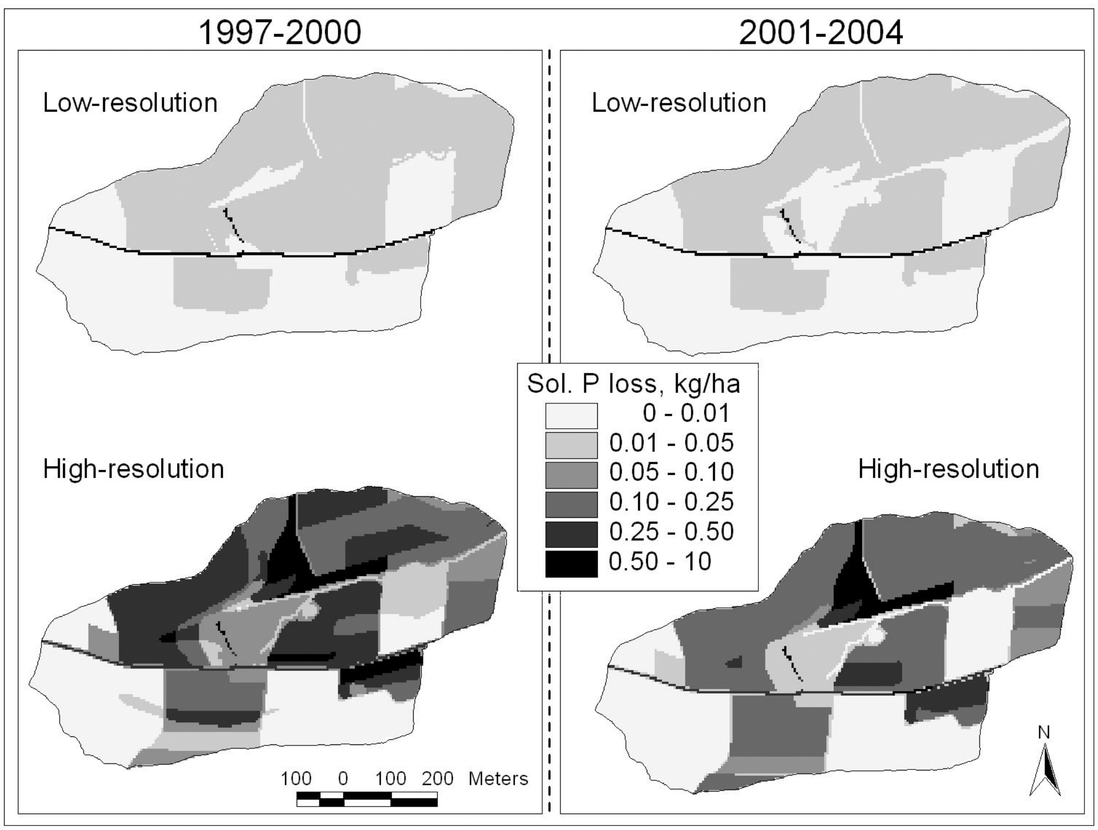

Figure 11. Average annual soluble P loss to stream (kg ha^{-1}) for two levels of land management distinction in SWAT, split by the SWAT calibration (1997-2000) and validation periods (2001-2004).

Phosphorus Loss

Due to file size computational constraints, P output was evaluated only on a monthly level. For soluble P concentrations, the two SWAT scenarios differed enough in magnitude of event peaks that they could not be considered to share a common statistical distribution with each other or with measured data on a monthly basis. However, for sediment P concentrations, all three scenarios (high-resolution, low-resolution, and measured) fit a three-parameter Weibull distribution and, according to the Mann-Whitney nonparametric test, were not likely to be from different distributions (p = 0.92 for high and low, p = 0.82 for high and measured, p = 0.73 for low and measured). Total P, which is highly biased by sediment P, was also not likely to be from different distributions, although with less strong p-values and without a common distribution determined through Minitab's Individual Distribution Identification tool. Despite statistical findings, there were no clear visual patterns in monthly time series (figs. 9 and 10), and the monthly NS, R^2, and Pbias values (table 3) did not indicate strong correlation between SWAT scenarios and measured data.

As expected, the majority of soluble P was lost from the cropped land, with the forests and pastures contributing little (fig. 11). Almost no distinction in soluble P with fragipan or slope was predicted. Soil P concentrations in the FD-36 watershed have been determined to be a factor of land management and not landscape position, surface soil texture, or other natural factors (Needelman et al., 2001).

Lack of clear trends in SWAT prediction of P is likely due to the majority of applied P on this watershed being in manure form. SWAT currently incorporates all applied P imme–diately into the soil as is appropriate for chemical fertilizer P. However, differences in transport mechanisms of manure-P versus chemical P have demonstrated the need to upgrade models accordingly (Vadas et al., 2005, 2007).

Conclusions

Due to the natural accumulation of error in modeling a water-based system in which sediment loss depends on water movement and P loss depends on both water and sediment movement, it is not surprising to find much more accurate predictions in flow depth than in P losses. However, while detailed input data frequently aid model accuracy, overall uncertainty associated with input data can quickly multiply as single data inputs are broken down into multiple inputs of finer temporal and spatial resolution (Chaplot, 2005; Harmel et al., 2006). The study's findings provide insight on the level of confidence in model response that is gained from using detailed data compared with using general field management data. Results were more defined at the field-level under the high-resolution scenario and followed patterns expected from knowledge about soil science, hydrology, P transport, and the characteristics of the study watershed. Spatial information provided by the field-level data helped evaluate management impacts within the watershed. However, the time spent collecting, understanding, entering, and error-checking input data required for the high-resolution scenario was on the order of months, while full data collection for the low-resolution scenario took several days. Additionally, in-depth calibration of the high-resolution scenario was substantially more complex than for the low-resolution scenario due to crop management distinctions between fields and across seasons.

Evaluating predicted results temporally on a monthly and daily basis using multiple methods (probability distributions, goodness-of-fit statistics, and graphs) reduced the tendency for analysis error due to generalization of a particular method or time unit. For example, grouping data by month or year can aid in identifying seasonal and long-term trends. In conjunction, a daily flow graph of FD-36 reflected that SWAT predicted the timing and trends of the measured data well but differed in magnitudes. Daily information, while often difficult to present legibly in journal format, can contribute important insights to daily statistics and monthly results. Similarly, comparing several levels of spatial detail (e.g., field and watershed) can improve model calibration and analysis of watershed response to management practices. Unfortunately, high levels of spatial detail in measured water quality data are typically even more difficult and unrealistic to obtain than those of land management data. Thus, the use of models with physically based processes and the ability for subfield distinction is quite useful in evaluating the impacts of management practices. Regardless, evaluation of various scenarios of a watershed, relative to a baseline scenario, is only realistic to the extent that the baseline and alternate scenarios are correctly portrayed and are characteristic of the region.

Obtaining data as specific to the time and location of a particular research question as possible has been shown to provide valuable information. However, to maximize research efficiency, the cost and time concerns related to data collection should be considered in a study's purpose. When appropriate, modelers must use next-best alternatives and make or obtain expert judgments from those knowledgeable in both the watershed-specific data processes and in how to best represent these in the specific model being used.

References

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams. 1998. Large-area hydrologic modeling and assessment: Part 1. Model development. *J. American Water Resources Assoc.* 34(1): 73-89.

Borah, D. K., and M. Bera. 2004. Watershed-scale hydrologic and nonpoint-source pollution models: Review of applications. *Trans. ASAE* 47(3): 789-803.

Bosch, D. D., J. M. Sheridan, H. L. Batten, and J. G. Arnold. 2004. Evaluation of the SWAT model on a coastal plain agricultural watershed. *Trans. ASABE* 47(5): 1493-1506.

Chaplot, V. 2005. Impact of DEM mesh size and soil map scale on SWAT runoff, sediment, and NO_3-N loads predictions. *J. Hydrol.* 312(1-4): 207-222.

Choi, I., C. L. Munster, D. M. Victor, R. H. White, G. A. Stewart, and C. E. Richards. 2005. Calibration and validation of the SWAT model on a field-scale watershed for turfgrass sod establishment. ASAE Paper No. 701P0105. In *Proc. 3rd Conference of Watershed Management to Meet Water Quality Standards and Emerging TMDL.* P. W. Gassman, ed. St. Joseph, Mich.: ASAE.

Gassman, P. W., M. R. Reyes, C. H. Green, and J. G. Arnold. 2007. The Soil and Water Assessment Tool: Historical development, applications, and future research directions. *Trans. ASABE* 50(4): 1211-1250.

Gburek, W. J., C. C. Drungil, M. S. Srinivasan, B. A. Needelman, and D. E. Woodward. 2002. Variable-source-area-controls on phosphorus transport: Bridging the gap between research and design. *J. Soil and Water Cons*. 57(6): 534-543.

Gburek, W. J., B. A. Needelman, and M. S. Srinivasan. 2006. Fragipan controls on runoff generation: Hydropedological implications at landscape and watershed scales. *Geoderma* 131(3-4): 330-344.

Gitau, M. W., T. L. Veith, and W. J. Gburek. 2004. Farm-level optimization of BMP placement for cost-effective pollution reduction. *Trans. ASAE* 47(6): 1923-1931.

Haan, C. T. 1977. *Statistical Methods in Hydrology.* Ames, Iowa: Iowa State University Press.

Harmel, R. D., R. J. Cooper, R. M. Slade, R. L. Haney, and J. G. Arnold. 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Trans. ASABE* 49(3): 689-701.

McDowell, R. W., A. N. Sharpley, and G. J. Folmar. 2001. Phosphorus export from an agricultural watershed: Linking source and transport mechanisms. *J. Environ. Qual.* 30(5): 1587-1595.

Martinec, J., and A. Rango. 1989. Merits of statistical criteria for the performances of hydrological models. *Water Resources Bulletin* 25(2): 421-432.

Maski, D., K. R. Mankin, P. Tuppad, S. Anand, K. A. Janssen, and G. M. Pierzynski. 2006. Calibration and validation of SWAT for field-scale sediment-yield prediction. ASABE Paper No. 062236. St. Joseph, Mich.: ASABE.

Mertens, M., and B. Huwe. 2002. FuN-Balance: A fuzzy balance approach for the calculation of nitrate leaching with incorporation of data imprecision. *Geoderma* 109(3-4): 269-287.

Nash, J. E., and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models: Part 1. A discussion of principles. *J. Hydrol.* 10(3): 282-290.

Needelman, B. A., W. J. Gburek, A. N. Sharpley, and G. W. Petersen. 2001. Environmental management of soil phosphorus: Modeling spatial variability in small fields. *SSSA J.* 65(5): 1516-1522.

Needelman, B. A., W. J. Gburek, G. W. Petersen, A. N. Sharpley, and P. J. A. Kleinman. 2004. Surface runoff along two agricultural hillslopes with contrasting soils. *SSSA J.* 68(3): 914-923.

Neitsch, S. L., J. G. Arnold, J. R. Kiniry, and J. R. Williams. 2005. Soil and Water Assessment Tool theoretical documentation: Version 2005. College Station, Tex.: Texas Water Resources Institute. Available at: www.brc.tamus.edu/swat/doc.html. Accessed 8 Dec. 2007.

Pionke, H. B., W. J. Gburek, R. R. Schnabel, A. N. Sharpley, and G. Elwinger. 1999. Seasonal flow, nutrient concentrations, and loading patterns in stream flow draining an agricultural hill-land watershed. *J. Hydrol.* 220(1-2): 62-73.

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. *J. American Water Resources Assoc.* 37(5): 1169-1188.

Schlüter, M., and N. Rüger. 2007. Application of a GIS-based simulation tool to illustrate implications of uncertainties for water management in the Amudarya river delta. *Environ. Modeling and Software* 22(2): 158-166.

Singh, V. P. 1988. *Hydrologic Systems: Vol. II. Watershed Modeling*. Englewood Cliffs, N.J.: Prentice Hall.

Srinivasan, M. S., W. J. Gburek, and J. M. Hamlett. 2002. Dynamics of storm flow generation: A field study in east-central Pennsylvania. *Hydrol. Proc.* 16(3): 649-665.

Srinivasan, M. S., P. Gerard-Marchant, T. L. Veith, W. J. Gburek, and T. S. Steenhuis. 2005. Watershed-scale modeling of critical source areas of runoff generation and phosphorus transport. *J. American Water Resources Assoc.* 41(2): 361-375.

USEPA. 2007. The Total Maximum Daily Load (TMDL) program. Washington, D.C.: U.S. Environmental Protection Agency, Office of Wetlands, Oceans, and Watersheds (4503T). Available at: www.epa.gov/owow/tmdl/ Accessed 10 December 2007.

USDA-NRCS. 2006. Conservation Effects Assessment Project (CEAP). Washington, D.C.: USDA Natural Resources Conservation Service. Available at: www.nrcs.usda.gov/Technical/nri/ceap/index.html Accessed 10 December 2007.

USDA-SCS. 1972. Sect. 4: Hydrology. In *Soil Conservation Service National Engineering Handbook.* Revised 1985. Washington, D.C.: USDA Soil Conservation Service.

Vadas, P. A., B. E. Haggard, and W. J. Gburek. 2005. Predicting dissolved phosphorus in runoff from manured field plots. *J. Environ. Qual.* 34(4): 1347-1353.

Vadas, P. A., W. J. Gburek, A. N. Sharpley, P. J. A. Kleinman, P. A. Moore, Jr., M. L. Cabrera, and R. D. Harmel. 2007. A model for phosphorus transformation and runoff loss for surface-applied manures. *J. Environ. Qual.* 36(1): 324-332.

Van Griensven, A., and T. Meixner. 2006. Methods to quantify and identify the sources of uncertainty for river basin water quality models. *Water Sci. and Tech.* 53(1): 51-59.

Veith, T. L., A. N. Sharpley, J. L. Weld, and W. J. Gburek. 2005. Comparison of measured and simulated P losses with indexed site vulnerability. *Trans. ASAE* 48(2): 557-565.

White, K. L., and C. Indrajeet. 2005. Influence of hydrologic response unit (HRU) distribution on SWAT model flow and sediment predictions. ASAE Paper No. 701P0105. In *Proc. 3rd Conference of Watershed Management to Meet Water Quality Standards and Emerging TMDL.* P. W. Gassman, ed. St. Joseph, Mich.: ASAE.

Zar, J. H. 1984. *Biostatistical Analysis.* 2nd ed. Englewood Cliffs, N.J.: Prentice Hall.