HYDROLOGICAL PROCESSES
Hydrol. Process. **23**, 748–763 (2009)
Published online 15 December 2008 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/hyp.7174

Comparison of AnnAGNPS and SWAT model simulation results in USDA-CEAP agricultural watersheds in south-central Kansas

Prem B. Parajuli,[1]* Nathan O. Nelson,[1] Lyle D. Frees[2] and Kyle R. Mankin[3]
[1] *Department of Agronomy, Kansas State University, Manhattan, KS 66506*
[2] *USDA/NRCS, Salina, KS 67401*
[3] *Department of Biological and Agricultural Engineering, Kansas State University, Manhattan, KS 66506*

Abstract:

This study was conducted under the USDA-Conservation Effects Assessment Project (CEAP) in the Cheney Lake watershed in south-central Kansas. The Cheney Lake watershed has been identified as 'impaired waters' under Section 303(d) of the Federal Clean Water Act for sediments and total phosphorus. The USDA-CEAP seeks to quantify environmental benefits of conservation programmes on water quality by monitoring and modelling. Two of the most widely used USDA watershed-scale models are Annualized AGricultural Non-Point Source (AnnAGNPS) and Soil and Water Assessment Tool (SWAT). The objectives of this study were to compare hydrology, sediment, and total phosphorus simulation results from AnnAGNPS and SWAT in separate calibration and validation watersheds. Models were calibrated in Red Rock Creek watershed and validated in Goose Creek watershed, both sub-watersheds of the Cheney Lake watershed. Forty-five months (January 1997 to September 2000) of monthly measured flow and water quality data were used to evaluate the two models. Both models generally provided from fair to very good correlation and model efficiency for simulating surface runoff and sediment yield during calibration and validation (correlation coefficient; R^2, from 0·50 to 0·89, Nash Sutcliffe efficiency index, E, from 0·47 to 0·73, root mean square error, RMSE, from 0·25 to 0·45 m^3 s^{-1} for flow, from 158 to 312 Mg for sediment yield). Total phosphorus predictions from calibration and validation of SWAT indicated good correlation and model efficiency (R^2 from 0·60 to 0·70, E from 0·63 to 0·68) while total phosphorus predictions from validation of AnnAGNPS were from unsatisfactory to very good (R^2 from 0·60 to 0·77, E from −2·38 to 0·32). The root mean square error–observations standard deviation ratio (RSR) was estimated as excellent (from 0·08 to 0·25) for the all model simulated parameters during the calibration and validation study. The percentage bias (PBIAS) of the model simulated parameters varied from unsatisfactory to excellent (from 128 to 3). This study determined SWAT to be the most appropriate model for this watershed based on calibration and validation results.

KEY WORDS flow; sediment; total phosphorus; calibration; validation; R^2; E; RMSE; RSR; PBIAS

Received 31 March 2008; Accepted 17 September 2008

INTRODUCTION

The majority of the Kansas population (>70%) uses surface water for drinking water and other daily uses (Kansas Water Office [KWO], 2004); but 39% of assessed river and stream miles and 76% of assessed lake acreage in Kansas are impaired for one or more designated uses (KDHE, 2006). The Cheney Lake watershed has been identified as 'impaired waters' under Section 303(d) of the Federal Clean Water Act for sediments and total phosphorus. The TMDLs for the Cheney Lake have been established by Kansas Department of Health and Environment (KDHE) for phosphorus and sediments, which recommended a reduction of 45% for both pollutants (CLWMP, 2008). To improve water quality through conservation practices, it is essential to identify sources of contaminants and quantify their contributions to water quality impairment. Current water quality assessment techniques generally include two methods: (a) water quality field monitoring and (b) computer/mathematical modelling. Field monitoring is the most appropriate and reliable method to support water quality assessment. However, it is laborious and expensive. Computer/mathematical models provide an alternative to monitoring that can save time, reduce costs, and minimize the need for testing management alternatives (Shirmohammadi *et al*., 2006).

Currently, numerous watershed models with various capabilities and degrees of complexity are available. Several of these models are used to estimate runoff, sediment yield, and phosphorus loads, and many are applicable to water quality goal development and implementation (Borah *et al*., 2006). These models are being explored as tools for developing management strategies to reduce effects of non-point source pollution on water quality. Choosing the appropriate model for a given application and watershed is critical. Proper model calibration and validation is necessary to ascertain accuracy and reliability of results (Das *et al*., 2007a). Models are first

* Correspondence to: Prem B. Parajuli, 2011D Throckmorton Plant Science Center, Kansas State University, Manhattan, KS 66506. E-mail: parajpb@ksu.edu.

calibrated against historically available data collected from the study watershed. The calibrated model can then be validated by comparing model predictions with measured data in other watersheds in the same geographic region. Calibrating models in one watershed and validating them in another watershed enables evaluation of model capability in different conditions and management practices (Parajuli *et al.*, 2007, 2008).

Because of their proven applicability at the watershed scale in Kansas and other places; popularity and capability to predict flow, sediment, and nutrients, the Annualized AGricultural Non-Point Source (AnnAGNPS) and the Soil and Water Assessment Tool (SWAT) models were chosen for this study. Refereed AnnAGNPS applications are predominantly for sites in the USA (Yuan *et al.*, 2001, 2002; Polyakov *et al.*, 2007); however, applications in other countries also have been reported, e.g. Australia (Baginska *et al.*, 2003), China (Hong *et al.*, 2005), Canada (Das *et al.*, 2006), and Belgium (Licciardello *et al.*, 2007). AnnAGNPS has been calibrated, validated, and applied for runoff and sediment yield losses from watersheds in different geographic locations, conditions, and management practices (Gebremeskel *et al.*, 2005; Das *et al.*, 2007a; Sadeghi *et al.*, 2007, Licciardello *et al.*, 2007).

Refereed applications of the SWAT model are numerous, and SWAT has been implemented internationally, e.g. Greece (Gikas *et al.*, 2006), Ireland (Nasr *et al.*, 2007), and Switzerland (Abbaspour *et al.*, 2007). SWAT has been calibrated, validated, and applied for runoff, sediment yield and total phosphorus losses from watersheds in different geographic locations, conditions, and management practices (Saleh *et al.*, 2000; Spruill *et al.*, 2000; Santhi *et al.*, 2001; Kirsch *et al.*, 2002; Van Liew *et al.*, 2003; White *et al.*, 2004; Qi and Grunwald, 2005; White and Chaubey, 2005; Wang *et al.*, 2006; Jha *et al.*, 2007; Gassman *et al.*, 2007; Parajuli *et al.*, 2007).

Limited research has been performed comparing AnnAGNPS and SWAT models for predicting monthly flow, sediment yields, and total phosphorus. Sadeghi *et al.* (2007), applied AnnAGNPS and SWAT models in a coastal plain watershed, Choptank River, Maryland, which is also a USDA-CEAP watershed. Portions of the watershed were impaired due to high levels of nutrients and sediment. They used five years (1991–1995) of detailed observed flow and water quality data to provide baseline calibration and validation for the two models. In their study, the calibrated AnnAGNPS and SWAT models performed fair to good (Parajuli, 2007) in terms of model correlation coefficient (R^2) and Nash Sutcliffe efficiency index (E) when compared with observed data ($R^2 = 0{\cdot}51$, $E = 0{\cdot}49$ for AnnAGNPS and $R^2 = 0{\cdot}50$, $E = 0{\cdot}34$ for SWAT). They did not report model performance during model validation. The sediment yield and total phosphorus losses from the watershed were not evaluated.

Das *et al.* (2007a) compared performances of AnnAGNPS and SWAT models in a watershed in Ontario, Canada. They used ten years (1991–2000) of monthly observed flow data to provide baseline calibration (1991–1995) and validation (1996–2000) for the two models. In their study, the AnnAGNPS and SWAT model demonstrated good to very good (Parajuli, 2007) model efficiency for flow when compared with observed data during the calibration ($E = 0{\cdot}79$ for AnnAGNPS and $E = 0{\cdot}70$ for SWAT) and validation ($E = 0{\cdot}69$ for AnnAGNPS and $E = 0{\cdot}57$ for SWAT) period. For sediment yield they used monthly estimated sediment yield data to compare with models simulated results for the calibration and validation period. AnnAGNPS and SWAT showed fair to good (Parajuli, 2007) model efficiency during calibration ($E = 0{\cdot}53$ for AnnAGNPS and $E = 0{\cdot}41$ for SWAT) and validation ($E = 0{\cdot}35$ for AnnAGNPS and $E = 0{\cdot}24$ for SWAT) periods. They did not report model performance using other statistical parameters such as R^2, root mean square errors (RMSE), etc.

The USDA-Conservation Effects Assessment Project (CEAP) within the USA funded more than 13 individual CEAP projects using models to evaluate best management practices (BMPs). Most of the CEAP projects in the USA use the SWAT model for their watershed water quality studies. The Cheney Lake watershed (CLW) has records of using a single event based Agricultural Non Point Source (AGNPS) model (Bhuyan *et al.*, 2003) and AnnAGNPS model (Lyle F., personal communication, 2007) in some of the sub-watersheds, including Red Rock Creek. However, the SWAT model, which is one of the currently available and popularly applied watershed water quality modelling tools, has not been applied to evaluate model performances in any of the CLW sub-watersheds. It is important to test the AnnAGNPS and SWAT models to compare their performances to select an appropriate model for the conservation effects evaluation of the CLW because choosing an appropriate model might affect BMPs selection decision making.

The objectives of this research were to (i) compare AnnAGNPS and SWAT model simulation results for surface flow, sediment yield, and total phosphorus, and (ii) determine the most appropriate model for this watershed based on calibration and validation results. This study is unique in that it compares model performance in two separate, but similar, watersheds.

METHODS AND MATERIALS

Study area

Red Rock Creek watershed and Goose Creek watershed, two sub-watersheds within the Cheney Lake watershed located on the North Fork of the Ninnescah River (HUC 11 030 014) in south-central Kansas, were selected for this project due to their similar spatial and land use characteristics and history of water quality data. Cheney Lake, the primary receiving water in the watershed and primary drinking water source for the City of Wichita, KS, has been identified as 'impaired waters' under Section 303(d) of the Federal Clean Water Act due to high levels of sediment and total phosphorus transport to the lake. The Kansas Department of Health and

Environment (KDHE) has set TMDLs for eutrophication and silt for Cheney Lake (KDA, 2004; KDHE, 2004). In recent years, a significant number of state and federal incentive programmes have been implemented for water quality improvement in the Cheney Lake watershed. This watershed is part of the USDA Conservation Effects Assessment Project (CEAP), which seeks to quantify environmental benefits of conservation programmes.

Models were calibrated in the Red Rock Creek watershed and validated in the Goose Creek watershed. Red Rock Creek watershed (Figure 1), located in Reno county, is of area 136 km^2 with an average elevation of 475 m. Land uses in the study area include grassland (32%), cropland (63%), woodland (4%), and 1% others (water, urban). Fine-loamy textured soils are predominant in this watershed. Goose Creek watershed (Figure 1), located in Kingman and Reno counties, is of area 136 km^2 with an average elevation of 505 m. Study area land uses include grassland (29%), cropland (64%), woodland (6%), and 1% others (water, urban). Fine-loamy textured soils are predominant in this watershed.

Model description

Both AnnAGNPS and SWAT are daily time step, watershed scale, pollutant-loading models developed to simulate long-term runoff, sediment, nutrients, and pesticide transport from agricultural watersheds (Table I) These models differ in structure. While the SWAT model uses hydrologic response units (HRUs), AnnAGNPS designates cells of various sizes; pollutants are routed from these cells into the associated reaches, and the model either deposits pollutants within the stream channel system or transports them out of the watershed (Geter and Theurer, 1998).

AnnAGNPS Model. The AnnAGNPS model was designed by the USDA Agriculture Research Service (USDA-ARS) and the USDA Natural Resources Conservation Service (USDA-NRCS) to evaluate NPS pollution from agriculture-dominated watersheds. It is a batch-process, continuous simulation, daily time step, pollutant-loading model developed to simulate long-term runoff, sediment, and chemical transport from agricultural watersheds (Cronshey and Theurer, 1998; Bingner and Theurer, 2003). It is a direct replacement for the single event model, Agricultural Non-Point Source (AGNPS) (Young *et al*., 1989), and retains many features of AGNPS (Yuan *et al*., 2001). Unlike AGNPS, AnnAGNPS divides the watershed into drainage areas with homogenous land use, soils, etc. and integrates these areas by simulated rivers and streams that route runoff and pollutants from each area downstream. The model uses and combines many modules of other widely used models, such as Revised Universal Soil Loss Equation (RUSLE) (Renard *et al*., 1997), Chemicals, Runoff, and Erosion from Agricultural Management Systems (CREAMS) (Knisel, 1980), Erosion Productivity Impact Calculator (EPIC) (Sharpley and Williams, 1990), and Groundwater Loading Effects on Agricultural Management Systems (GLEAMS) (Leonard *et al*., 1987).

AnnAGNPS allows users to select cell-based spatial representation, which is characterized by similar land and soil properties. Soil moisture conditions are calculated with the Soil Conservation Service curve numbers (SCS-CN) method, which serves as the basis for determining surface and subsurface runoff quantities. AnnAGNPS uses the RUSLE to calculate sediment delivered to a field edge as a result of runoff from any type of precipitation.

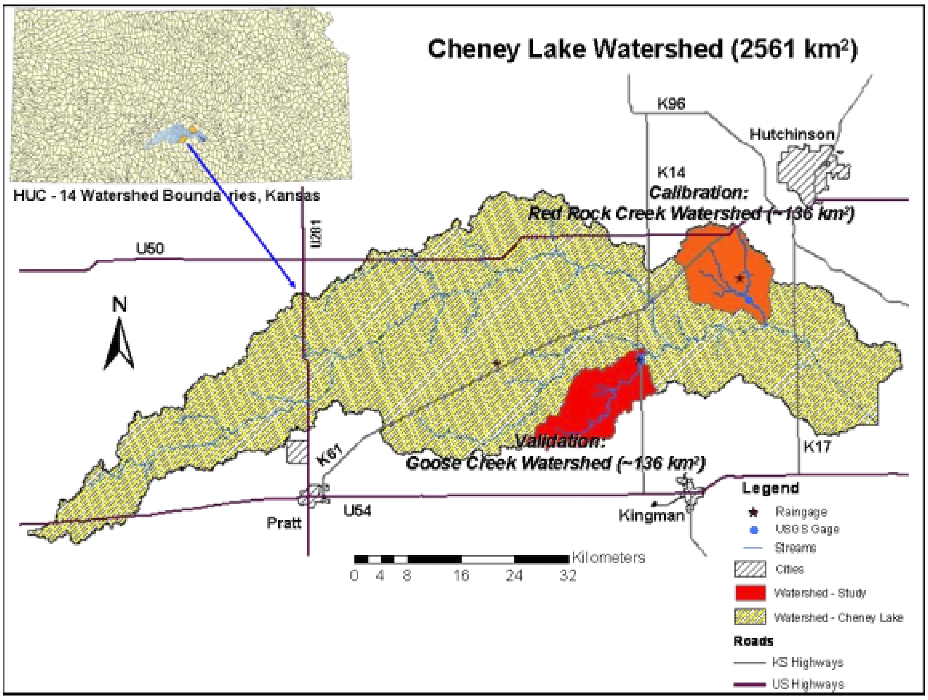

Figure 1. Location map of the calibrated (Red Rock Creek) and validated (Goose Creek) watersheds

Table I. Comparison of the hydrologic, sediment yield, chemical and BMP modeling of AnnAGNPS and SWAT models

Description	AnnAGNPS	SWAT	References
Model components and capabilities:	Hydrology, sediment yield, nutrients, pesticides snowmelt, rainfall, irrigation, and capabilities to generate cells and stream networks using TOPAZ	Hydrology, sediment yield, nutrients, pesticides, snow melt, channel and reservoir routing, crop growth, soil temperature, pathogen transport, stream processes with ArcView GIS plateform	Borah and Bera, 2003
Scale:	Long-term, daily, or sub-daily steps	Long-term, daily, or sub-daily steps	
Watershed representation:	Homogeneous landareas (cells), reaches, and impoundments	Sub-basins grouped based on climate, hydrologic response units, ponds, ground water, and main channel	
Input weather data:	Precipitation, maximum and minimum temperature, dew point temperature, sky cover, and wind speed	Precipitation, maximum and minimum temperature, relative humidity, wind speed, and solar radiation	Migliaccio and Srivastava, 2007
Weather simulator:	GEM and complete climate (Bingner *et al*., 2007)	WXGEN (Sharpley and Williams, 1990)	Migliaccio and Srivastava, 2007
PET method:	Penman (Jenson *et al*., 1990)	Penman-Monteith, Priestly-Taylor, and Hargreaves (Monteith, 1965; Allen, 1986; Allen *et al*., 1989; Priestley and Taylor, 1972; Hargreaves *et al*., 1985)	Migliaccio and Srivastava, 2007
Infiltration/surface runoff algorithms:	Modified SCS CN2	Modified SCS CN2 with daily time step or Green-Ampt Mein-Larson infiltration equation (Mein and Larson, 1973)	King *et al*., 1999
Peak runoff rate:	SCS TR-55 method (Bingner *et al*., 2007)	Modified rational and SCS TR-55 method s (Neitsch *et al*., 2005)	Borah and Bera, 2003
Groundwater and stream flow simulation:	Subsurface flow or Interflow - Yes	Subsurface flow or Interflow - Yes	Migliaccio and Srivastava, 2007
	Subsurface drainage or tile drainage - Yes	Subsurface drainage or tile drainage - Yes	
	Shallow aquifer or other water storage zone - No	Shallow aquifer or other water storage zone - Yes	
	Deep aquifer - No	Deep aquifer - Yes	
Subsurface flow:	Lateral subsurface flow using Darcey's equation or tile drain flow using Hooghoudt's equation and parallel drain approximation	Lateral subsurface flow using kinematic storage model (Sloan *et al*., 1983), and groundwater groundwater flow using empirical relations	Borah and Bera, 2003

Water routing/channel: ***streams/reaches:***	Manning's equation and channel shape relationships	Manning's equation and variable storage routing or Muskingum river routing method	Migliaccio and Srivastava, 2007
Channel for runoff:	Trapezoid channel cross-section, rectangular main channel, rectangular out-of-bank (floodplain) section	Trapezoidal channel cross-section and trapezoidal flood plain	Migliaccio and Srivastava, 2007
Sediment Yield: **Overland**	Uses RUSLE to generate sheet and rill erosion, HUSLE (Theurer and Clarke, 1991) for delivery ratio, and sediment deposition based on particle size distribution (Young *et al*., 1987) and particle fall velocity, ephemeral gully erosion model (Gordon *et al*., 2007)	Sediment yield based on Modified Universal Soil Loss Equation (MUSLE) expressed in terms of runoff volume, peak flow, and USLE factors (Neitsch *et al*., 2005)	Borah and Bera, 2003
Channel	Modified Einstein equation for sediment transport and Bagnold equation to determine transport capacity of flow. Bagnold, 1966; Theurer and Cronshey, 1998)	Bagnold's stream power concept for bed degradation and sediment transport, adjusted with USLE soil erodibility and cover factors, and deposition based on particle (Bagnold, 1977)	Borah and Bera, 2003
Chemical Simulation:	Soil moisture, nutrients, and pesticides in each cell are tracked using NRCS soil database and crop information and reach routing includes fate and transport of nitrogen, phosphorus, and individual pesticides, and organic carbon (Binger *et al*., 2007)	Runoff P based on partitioning factor daily organic N and sediment adsorbed functions, crop N and P use from supply and demand, and pesticides based on plant leaf-area index, application efficiency, wash off fraction, organic C adsorption, and half-life decay (Neitsch *et al*., 2005)	Borah and Bera, 2003
Best Management Practices (BMPs) Evaluation:	Agricultural managements: impact of watershed management practices on runoff and sediment BMPs include: vegetative filter strip, terracing, contouring, strip cropping, tillage operation, cover crops, crop rotations, tile drainage, irrigation, rice fish farming (Yuan *et al*., 2001; Yuan *et al*., 2006; Yuan *et al*., 2007)	Agricultural managements: impact of watershed management practices on runoff and sediment losses, nutrients, pesticides, pathogen transport BMPs include: grass waterway, grade stabilization structure, vegetative filter strip, terracing, contouring, strip cropping, tillage operations, cover crops, crop rotations, grazing operation, pesticide application, irrigation, nutrient management (Saleh *et al*., 2000; Gitau *et al*., 2003; Bosch *et al*., 2005; Chu *et al*., 2005; Bockhold *et al*., 2006; Bracmort *et al*., 2006; Gitau *et al*., 2007; Parajuli, 2007)	Borah and Bera, 2003

The Hydro-geomorphic Universal Soil Loss Equation (HUSLE) is used to estimate the total sediment yield leaving each field and entering the stream reach while accounting for deposition. Output is expressed on an event basis for selected stream reaches and as source accounting (contribution to outlet) from land or reach components over the simulation period.

SWAT model. The SWAT model is a product of USDA\ARS. SWAT is a physically based, semi-distributed parameter, watershed-scale model that operates on a continuous daily time step. It simulates hydrological processes, sediment yield, nutrient loss, and pesticide losses into surface/groundwater and the effects of agricultural management practices on water in large ungauged watersheds (Arnold *et al*., 1998). SWAT incorporates the effects of weather, surface runoff, evapotranspiration, crop growth, irrigation, groundwater flow, nutrient loading, pesticide loading, and water routing, as well as the long-term effects of varying agricultural management practices (Neitsch *et al*., 2002, 2005). In the hydrologic component, runoff is estimated separately for each sub-watershed of the total watershed area and routed to obtain the total runoff for the watershed. Runoff volume is estimated from daily rainfall using modified SCS-CN and Green–Ampt methods. Sediment yield is estimated from the Modified Universal Soil Loss Equation (MUSLE). SWAT has been applied extensively for streamflow, sediment yield, and nutrient modelling (Gosain *et al*., 2005; Vache *et al*., 2002; Varanou *et al*., 2002). The model requires input of DEM, land use, and soils, as well as time series of climate data such as daily precipitation and temperature.

In the SWAT model, the watershed is partitioned into small sub-basins that are further subdivided into HRUs based on unique land cover, soil, and topographic conditions. Such divisions are necessary to consider accurately possible effects of spatial variations in parameters on hydrological processes, sediment, and nutrient simulations. The hydrology component of the model calculates a soil water balance at each time step based on daily amounts of precipitation, runoff, evapotranspiration, percolation, and baseflow. Sediment yield from each sub-basin or HRU is computed with the MUSLE. The MUSLE approach of estimating sediment yield makes the sediment computation a non-linear function of the HRU area. Simulations are performed at the HRU level and summarized in each sub-watershed. The simulated variables (water, sediment, nutrients, and other pollutants) are routed through the stream network to the watershed outlet. SWAT evolved from the Simulator for Water Resources in Rural Basins (SWRRB) and Routing Outputs to Outlet (ROTO) models. Other models also influenced SWAT development including CREAMS (Knisel, 1980), GLEAMS (Leonard *et al*., 1987), and EPIC (Williams *et al*., 1984; Neitsch *et al*., 2002).

Model input

Both AnnAGNPS and SWAT use various sets of geospatially referenced data to create layers of information to satisfy the necessary input parameters. United State Geological Survey (USGS, 1999) 7·5-minute digital elevation data (DEM) was used to delineate watershed boundaries and topography. The Soil Survey Geographic Database (SSURGO) was used to create a soil database (USDA, 2005). Land use and land management were estimated by analysing Landsat 5 satellite imagery using stacked images from May and August of 1997 for major crop types and unsupervised classification techniques within ArcView Image Analysis with ground truth verification using Farm Service Agency records. Image Analysis in ArcView also is capable of performing the Normalized Difference Tillage Index (NDTI) (band 5-band 7)/band 5+band 7) function using Landsat 5 mid-infrared bands 5 and 7. Once the NDTI function was completed, results were separated into three crop residue covers: high, medium and low. Using this information paired with local knowledge, land use and land management were classified into 24 classes with major land uses including wheat, soybean, grain sorghum, corn, CRP, forestland, pastureland, rangeland, urbanland, and water (Lyle Frees, 1997, unpublished data). AnnAGNPS input accepts five types of land use identifiers (cropland, pasture, forest, rangeland and urban), and only the predominant land use and management are used to represent each AnnAGNPS cell. Hence, the land uses in the sub-watershed were reclassified according to the model requirement. Each land use type was included under a land use identifier during input data preparation.

For the AnnAGNPS model, a threshold critical source area (CSA) of 100 ha and a minimum source channel length (MSCL) of 130 m were used to generate spatially variable stream network parameters. AnnAGNPS divided the calibrated watershed into 169 cells. A MSCL of 130 m is a default value given in the model, but the CSA value was changed to 100 ha, which represents less than 1% of the calibration watershed area (13 600 ha or 136 km^2) as similar to the SWAT model threshold area. The dominant soil and land use for each cell were determined from the soil and land use shape files over the delineated sub-watershed. For the SWAT model, default threshold critical source area was used, which divided the calibration watershed into 12 sub-basins. The SWAT model further considered spatial variability of parameters within each sub-basin by means of HRUs.

Model calibration

Flow, and sediment yield-related model calibration parameters were selected (Table II) based on previous research (Das *et al*., 2007a; Parajuli *et al*., 2007; Sadeghi *et al*., 2007; Wang *et al*., 2007). Flow calibration was performed by adjusting the CN parameter. CN is a soil moisture balance parameter that allows the model to modify the soil moisture condition of the soil to estimate surface runoff. As the value of CN is reduced,

DOI: 10.1002/hyp

Table II. Model parameter test and adjustment during calibration

Parameters		Default value	Test range value	Final value
Flow:				
Curve Number (CN) in		73–83	73–83	77–79
AnnAGNPS and SWAT	Cropland	83	74–83	78
	Grassland	79	75–82	79
	Woodland	73	73–80	77
Sediment:				
USLE cover and management factor (C) in SWAT		0·03–0·20	0·03–0·50	0·03–0·20
		Winter wheat	0·03–0·30	0·03
		Grain Sorghum	0·20–0·50	0·2
		Soybean	0·20–0·50	0·2
		Corn	0·20–0·50	0·2
Manning'ss n EGs[1] in AnnAGNPS		0·04	0·04–1·00	0·04

[1] Ephemeral Gully.

the model allows less water to runoff from the surface. Other parameters than CN factor may also effect flow prediction but their effect is considered to be small; many studies ranked CN as the most sensitive parameter (Arabi *et al*., 2007; Das *et al*., 2007a; Feyereisen *et al*., 2007; Parajuli *et al*., 2007; Sadeghi *et al*., 2007; Wang *et al*., 2007). Red Rock Creek and Goose Creek contribute low flow in various seasons, although they are considered perennial streams. Baseflow separation analysis (Kyoung *et al*., 2005; Eckhardt, 2005) in the Red Rock Creek and Goose Creek watersheds indicated about 1·2% (weighted average) base flow of the total direct flow. Other studies confirmed similar results (Li *et al*., 2006) for these watersheds.

Models were calibrated for sediment yield by adjusting the Universal Soil Loss Equation crop cover management factor (USLE C), one of the most widely used sediment calibration factors (Parajuli, 2007). The USLE C factor is defined as the ratio of soil loss from land cropped under specified conditions to the corresponding loss from clean-tilled, continuous fallow land (Wischmeier and Smith, 1978).

Monthly flow, sediment yield, and total phosphorus data measured from the USGS gauge station for each watershed were used to calibrate and validate the model. Monthly measured data from January 1997 to September 2000 were used for model calibration and validation. Model predictions were evaluated statistically with the coefficient of determination (R^2) and Nash–Sutcliffe Efficiency Index (E) between measured values and model-predicted values after each model run with changed parameters. Model input parameters were continuously modified during the calibration phase until simulated flow and sediment yield gave results $R^2 \geq 0{\cdot}5$ and $E \geq 0{\cdot}5$ (Ramanarayanan *et al*., 1997; Parajuli *et al*., 2006). Flow calibration was determined first using all default parameters. The CN parameters were continuously modified within the specified range of values during the calibration phase to find the local optimum value. The CN range of 77–79 (77 for woodland, 78 for cropland, and 79 for grassland) resulted in the best calibration for both models. SWAT allows users to input initial SCS-CN for moisture condition II in the management data. AnnAGNPS allows users to input all CNs based on the soil hydrologic group (A, B, C, and D). Initial AnnAGNPS model predictions were made with the default RCN factor (retention calibration factor) of 1·00, after which the model was run automatically nine times using different RCN factors (1·000–2·789).

The USLE C, default was given as 0·03 for winter wheat and 0·20 for grain sorghum, soybean, and corn crops. During calibration, the C factor was tested in the range 0·03 to 0·30 for winter wheat and 0·20 to 0·50 for grain sorghum, soybean, and corn crops. Values above the default values were tested because the models were underpredicting sediment yield. However, increasing the C factors did not increase model efficiency for sediment yield prediction.

Therefore, the default C factor given in the model resulted in the best model efficiency, in this study. The default USLE support practice factor (*P*) given in the models (1·00) was tested in the range from 0·20 to 1·00 during model calibration. Models always underpredicted sediment yields using both lower and upper *P* factors. The *P* factor was fixed at 0·50, which generally represents the current conditions of the watersheds. Terraces are one of the conservation practices implemented in the watersheds. *P* is defined as the ratio of soil loss with a specific support practice to the corresponding loss with up-and-down slope culture. Support practices include contour tillage, strip cropping on the contour, and terrace systems. Stabilized waterways for the disposal of excess rainfall are a necessary part of each of these practices (Wischmeier and Smith, 1978). No calibration parameters were used to calibrate total phosphorus prediction. After model calibration, input parameters were not changed during the model validation process.

Management scenarios

Land in the CRP covers about 16% of the calibration watershed area. The CRP land (high, high terrace,

DOI: 10.1002/hyp

medium, medium terrace, low, low terrace) was simulated with five typical types of grass management: little bluestem, big bluestem, Indiangrass, side oats, and switchgrass, which represent the field conditions. These five types of grasses have about equal cover in the watershed. Grassland covers about 20% of the watershed area and typically includes rangeland big bluestem (high, medium, low). The CRP grasses are generally not fertilized (Lisa French, Cheney Lake Watershed Inc., 2007, personal communication).

A majority (~64%) of the land use area in both watersheds is cropland. Grain sorghum and soybean are major warm-season crops, and winter wheat is a primary cool-season crop grown in a four-year rotation (Lisa French, Cheney Lake Watershed Inc., 2007, personal communication). Typical planting and harvesting dates are 25 May and 20 October for warm-season crops and 20 October and 29 June for cool-season crops. Crop residue is left on the ground between the crop periods. Sorghum, soybean, and wheat are cultivated primarily with conventional and conservation tillage systems. Primary herbicides used for warm-season crops are Bicep II Magnum for sorghum and Roundup for soybean; Finesse was used for winter wheat. These methods apply to both calibrated and validated watersheds. Woodlands cover about 4% of the calibration watershed and about 6% of the validation watershed land-use area. Model default parameters were used for woodland areas assuming mixed forest trees in the watersheds.

Weather and hydrologic data

Weather data, such as daily precipitation and daily ambient temperatures, were extracted from the National Climatic Data Center (NCDC). The SWAT model requires daily precipitation, daily maximum and minimum temperatures, daily solar radiation, daily wind speed, and daily relative humidity. The AnnAGNPS model requires daily precipitation, daily maximum and minimum temperatures, daily dew point temperatures, daily sky cover, and daily wind speed data. Daily precipitation data were used from only one weather station (Hutchinson South) for the calibration watershed and from two weather stations (USGS gauge 3 and Turon) for the validation watershed. SWAT uses weather data from the nearest weather station to its sub-watershed whereas AnnAGNPS uses a thiessen polygon average of the weather data. Daily precipitation data for the calibration watershed was obtained from the Hutchinson south weather station, which is located nearly in the middle of the watershed. The SWAT model fills in missing weather data with data from the Wichita airport weather station (Sedgwick County), which is located about 55 km southeast of the calibration watershed.

Daily precipitation data for the validation watershed were from USGS gauge 3 and Turon weather stations. The USGS gauge 3 weather station is located at the mouth of the Goose Creek watershed and Turon is located about 10 km from the watershed. To fill in missing data for the validation watershed, the SWAT model used data from the Pratt weather station (Pratt County), located about 31 km south-west of the watershed. The 1997 to 2000 average annual rainfall data measured for Hutchinson south was 813 mm (Figure 2a) and for USGS gauge 3 and Turon was about 662 mm (Figure 2b). USGS gauge 3 weather station had no rainfall data measured from October to December, 2000.

Statistical analysis

SWAT and AnnAGNPS model responses for flow were evaluated based on measured flow data from January 1997 to December 2000. This study used 45 months of rainfall–runoff measured/observed events for both watersheds. Statistical parameters used to evaluate the relationship between measured and predicted flow, sediment yield, and total phosphorus include correlation coefficient (R^2), Nash–Sutcliffe efficiency index (E), root mean square error (RMSE), RMSE–observations standard deviation ratio (RSR), and percentage bias (PBIAS). The R^2 value indicates the consistency with which measured versus predicted values follow a best fit line. If the R^2 values are less than or very close to zero, the model prediction is considered unacceptable or poor. If the values are one, then the model prediction is perfect (Santhi *et al.*, 2001). However, R^2 only describes how much of the observed dispersion is explained by the prediction, therefore R^2 is not used alone.

E indicates the consistency with which measured values match predicted values, or the fit of the data to a linear 1 : 1 measured versus predicted best-fit line (Nash

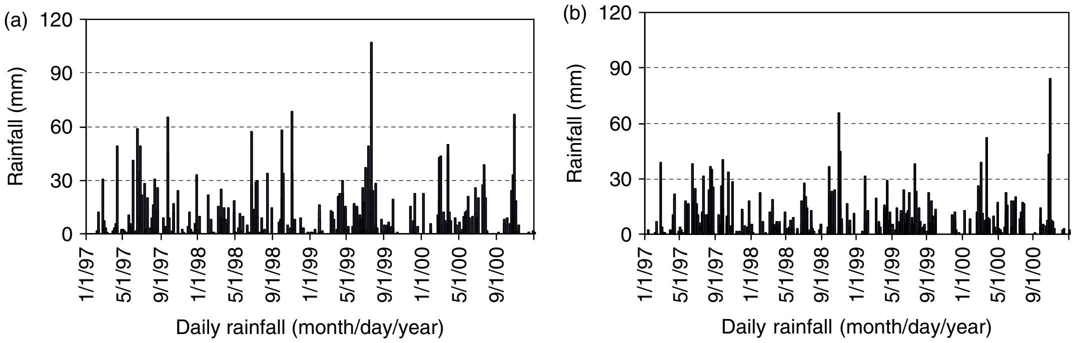

Figure 2. Daily rainfall data for (a) Red Rock Creek and (b) Goose Creek watersheds

DOI: 10.1002/hyp

and Sutcliffe, 1970) and estimated using Equation (1). E ranges from minus infinity (poor model) to 1·0 (perfect model). For example, if the square of the differences between the model predictions and the observations is as large as the variability in the observed data, then E = 0·0; if it exceeds it, then E < 0·0 (i.e. the observed mean is better than the predictor). Thus, a value of zero for E indicates that the observed mean, O, is as good as a predictor as the model, while negative values indicate that the observed mean is a better predictor than the model. E has been widely used to evaluate the performance of hydrologic models (Wilcox *et al.*, 1990). A limitation of E is the fact that the differences between the observed and predicted values are calculated as squared values. As a result larger values in a time series are strongly overestimated whereas lower values are neglected (Legates and McCabe, 1999).

$$E = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2} \quad (1)$$

where O = observed value and P = predicted value. The over-bar denotes the mean (observed or predicted) for the entire time period of the evaluation.

RMSE summarizes the average error between observed and predicted variates using the same units as those variates, and is estimated using Equation (2). The lower RMSE the better the model performance, and a value of zero represents perfect simulation of the observed data (Chu and Shirmohammadi, 2004). RMSE summarizes the mean difference in the units of observed and predicted values. The use of absolute error measures using RMSE provides an evaluation of the error in the units of the variable. The RMSE indicates the bias (deviation of the actual slope from the 1 : 1 line) compared with the random variation that may occur (Willmott, 1984). RMSE is biased when large outliers are present.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Q_{pi} - Q_{oi}\right)^2}{n}} \quad (2)$$

where $RMSE$ = root mean squared error, Q_{pi} = predicted value for event i, Q_{oi} = observed value for event i, and n = number of events.

RSR is estimated as the ratio of the RMSE and standard deviation of measured data using Equation (3). RSR varies from the optimal value of 0, which indicates zero RMSE or residual variation and therefore perfect model simulation, to a large positive value. The lower RSR, the lower the RMSE, and the better the model simulation performance (Moriasi *et al.*, 2007).

$$RSR = \frac{RMSE}{STDEV_{obs}} \quad (3)$$

where $RSR = RMSE-$observations standard deviation ratio, $RMSE$ = root mean squared error, and $STDEV_{obs}$ = standard deviation of measured data.

PBIAS measures the average tendency of the model-predicted values to be larger or smaller than their corresponding measured values. The optimal value of PBIAS is 0·0, with low-magnitude values indicating accurate model simulation. Positive values indicate model underestimation bias, and negative values indicate model overestimation bias (Gupta *et al.*, 1999). PBIAS is the deviation of data values being evaluated, expressed as a percentage, which is calculated from

$$PBIAS = \left[\frac{\sum_{i=1}^{n}\left(Y_i^{obs} - Y_i^{sim}\right) \times (100)}{\sum_{i=1}^{n}(Y_i^{obs})}\right] \quad (4)$$

where $PBIAS$ = percentage bias, Y_i^{obs} = observed value for event i, and Y_i^{sim} = predicted value for event i.

In this study, statistics using R^2, E, RMSE, RSE, and PBIAS values were considered. Model correlations and efficiencies, as modified by Parajuli (2007) from Moriasi *et al.* (2007), were classified as in Table III.

RESULTS AND DISCUSSION

Flow

Calibrated models for the Red Rock Creek watershed predicted mean monthly flow of the watershed

Table III. Classification of model efficiencies for the different pollutant parameters

Class	R^2, E Flow, sediment, TP	RSR Flow, sediment, TP	PBIAS Flow	PBIAS Sediment	PBIAS TP
Excellent	<0·90	0·00–0·25	$< \pm 10$	$< \pm 15$	$< \pm 25$
Very good	0·75–0·89	0·26–0·50	$\pm 11 \leq \pm 15$	$\pm 16 \leq \pm 30$	$\pm 26 \leq \pm 40$
Good	0·50–0·74	0·51–0·60	$\pm 16 \leq \pm 25$	$\pm 31 \leq \pm 50$	$\pm 41 \leq \pm 60$
Fair	0·25–0·49	0·61–0·70	$\pm 26 \leq \pm 30$	$\pm 51 \leq \pm 60$	$\pm 61 \leq \pm 70$
Poor	0·00–0·24	0·71–0·89	$\pm 31 \leq \pm 35$	$\pm 61 \leq \pm 70$	$\pm 71 \leq \pm 80$
Unsatisfactory	<0·00	>0·90	$\geq \pm 36$	$\geq \pm 71$	$\geq \pm 81$

R^2 = Coefficient of determination.
E = Nash sutcliffe efficiency index.
TP = Total phosphorus.
RSR = Root mean square error - observations standard deviation ratio.
PBIAS = Percent bias.

with very good correlation and good agreement ($R^2 = 0.80$, $E = 0.69$, $RMSE = 0.38$ for AnnAGNPS; $R^2 = 0.81$, $E = 0.56$, $RMSE = 0.45$ for SWAT) between mean monthly measured and mean monthly predicted flow values (Table IV, Figure 3). The SWAT model predicted monthly flow ($m^3\ s^{-1}$) estimated about 18% higher RMSE than AnnAGNPS in the calibration watershed. The calibrated models, when applied to the Goose Creek watershed for validation, predicted mean monthly flow with good correlation and fair agreement for both models ($R^2 = 0.50$, $E = 0.47$, $RMSE = 0.26$ for AnnAGNPS; $R^2 = 0.62$, $E = 0.48$, $RMSE = 0.25$ for SWAT) (Table IV). The AnnAGNPS model predicted monthly flow ($m^3\ s^{-1}$) estimated about 4% higher RMSE than SWAT in the validation watershed (Table IV). The estimated RSR values for both AnnAGNPS and SWAT model simulations during calibration and validation were excellent (0·09–0·12). The calculated PBIAS values for both models were unsatisfactory (PBIAS from 49 to −95), which means both models were biased to

Table IV. Estimated statistical parameters of model performance for calibration and validation watersheds

Class	Calibration							Validation						
	Slope	Intercept	R^2	E	RMSE	RSR	PBIAS	Slope	Intercept	R^2	NSE	RMSE	RSR	PBIAS
AnnAGNPS:														
Flow	0·54	−0·01	0·80	0·69	0·38	0·09	49·83	0·65	0·17	0·50	0·47	0·26	0·12	−40·35
Sediment	0·44	−3·64	0·83	0·60	230·00	0·10	57·70	0·70	43·67	0·62	0·64	312·00	0·19	−67·14
TP	0·90	385·08	0·60	0·32	704·00	0·13	−128·48	2·38	−55·50	0·77	−2·38	476·00	0·25	−117·33
SWAT:														
Flow	1·08	0·25	0·81	0·56	0·45	0·10	−95·06	0·86	0·13	0·62	0·48	0·25	0·11	−46·62
Sediment	0·56	−1·41	0·89	0·73	186·00	0·08	45·02	0·40	46·58	0·72	0·61	158·00	0·10	9·65
TP	0·51	126·88	0·70	0·68	487·00	0·09	3·13	0·64	33·34	0·60	0·63	178·00	0·09	13·44

R^2 = Coefficient of determination.
E = Nash sutcliffe efficiency index.
RSR = Root mean square error - observations standard deviation ratio.
TP = Total phosphorus.
PBIAS = Percent bias.

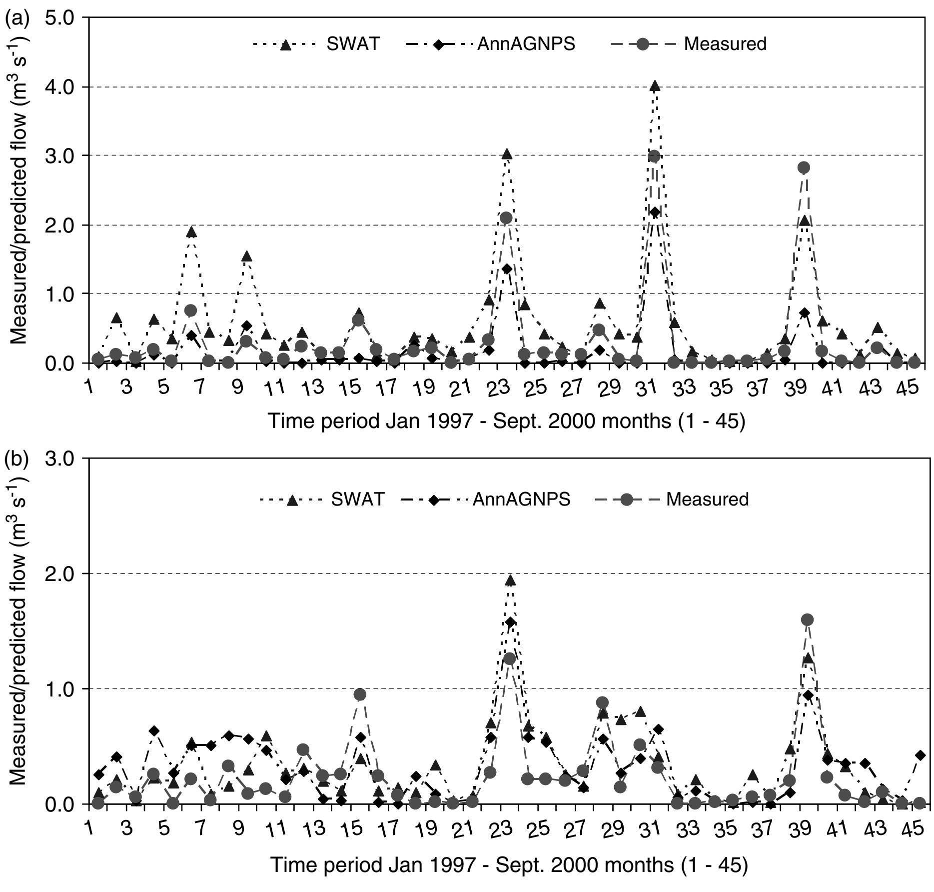

Figure 3. Time series measured and models predicted monthly average flow from January 1997 to September 2000 for (a) Red Rock Creek and (b) Goose Creek watersheds

underpredict or overpredict flow for different months of the simulation period. AnnAGNPS flow prediction was determined to be less biased than SWAT during model calibration (PBIAS: 50 versus −95) and validation (PBIAS: −40 versus −46) study.

Although, AnnAGNPS and SWAT models predicted surface runoff equally well in general for both calibration and validation watersheds, AnnAGNPS slightly underpredicted monthly surface flow compared with SWAT during model calibration (slope: 0·54 versus 1·08) and validation (slope: 0·65 versus 0·86). Model-predicted values of estimated slopes during model calibration and validation indicated no-significant (*P* values <0·001 at $\alpha_{0.05}$) difference.

Bhuyan *et al.* (2003) applied a single event-based AGNPS model in the Red Rock Creek watershed. The single event-based AGNPS model systematically overestimated surface runoff, and the model had to be adjusted for antecedent moisture condition (AMC) based on surface runoff prediction. Compared with predictions by Bhuyan *et al.* (2003), model predictions from the continuous version of AnnAGNPS used in this study had better correlation with observed values. Van Liew *et al.* (2003) applied the SWAT model in the Delaware Creek watershed in Oklahoma, which has land-use conditions similar to Red Rock Creek and Goose Creek watersheds, and found $R^2 = 0.68$ and $E = 0.84$ for mean monthly flow prediction. Spruill *et al.* (2000) applied the SWAT model in a small, central Kentucky watershed. The SWAT model prediction had E values for monthly flows between 0·58 and 0·89. They did not use other statistical parameters as used in this study, such as RMSE, RSR, and PBIAS.

Sediment yield

Calibrated AnnAGNPS and SWAT models for Red Rock Creek watershed predicted monthly sediment yield with very good correlation and good agreement ($R^2 = 0.83$, $E = 0.60$, $RMSE = 230$ for AnnAGNPS; $R^2 = 0.89$, $E = 0.73$, $RMSE = 186$ for SWAT) with measured mean monthly sediment yield data (Table IV, Figure 4). The AnnAGNPS model-predicted monthly sediment yield (Mg) estimated about 23% higher RMSE than SWAT in the calibration watershed.

Both models showed decreased, but still good, correlation and agreement ($R^2 = 0.62$, $E = 0.64$, $RMSE = 312$ for AnnAGNPS; $R^2 = 0.72$, $E = 0.61$, $RMSE = 158$ for SWAT) between mean monthly measured and mean monthly predicted sediment yield values during validation in the Goose Creek watershed (Table IV). The AnnAGNPS model predicted monthly sediment yield (Mg) estimated about 97% higher RMSE than SWAT

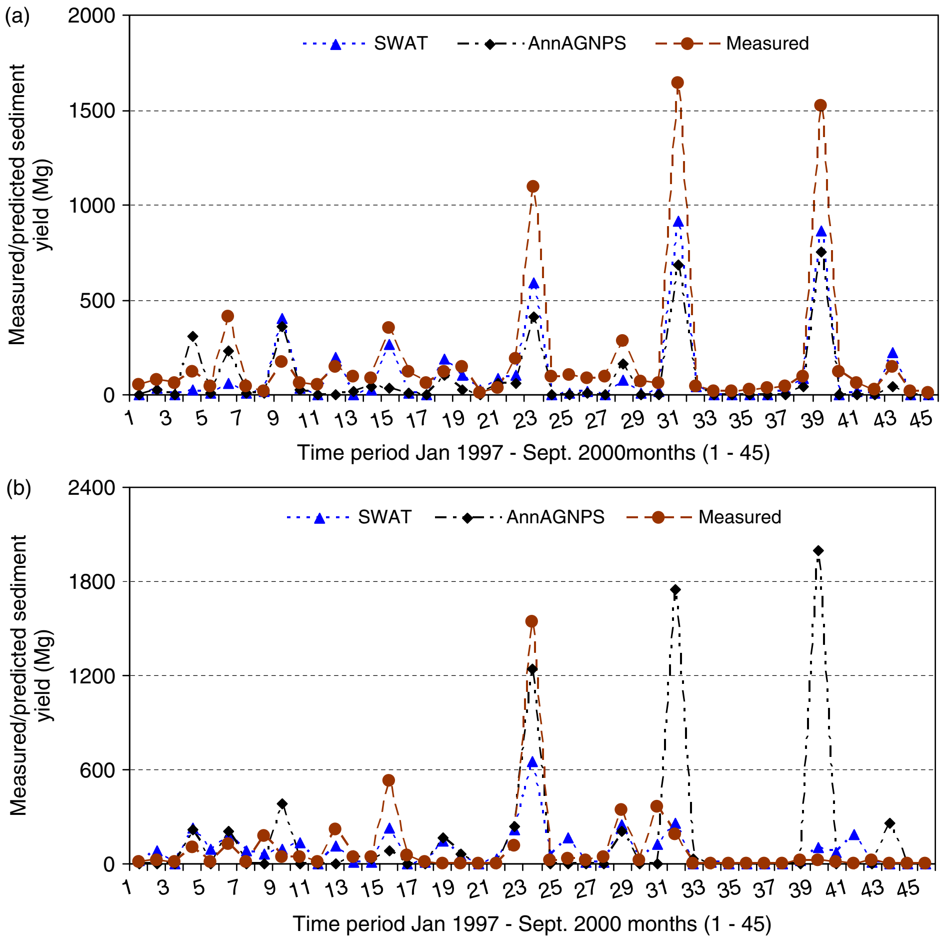

Figure 4. Time series measured and models predicted monthly average sediment yield from January 1997 to September 2000 for (a) Red Rock Creek and (b) Goose Creek watersheds

in the validation watershed (Table IV). The RSR values were estimated excellently when both AnnAGNPS and SWAT model simulation results for the calibration and validation watersheds were analysed. The estimated PBIAS values using AnnAGNPS model simulation ranged from fair to poor (PBIAS from 57 to −67) whereas SWAT performance ranged from excellent to good (PBIAS from 9 to 45) during model calibration and validation.

Although AnnAGNPS and SWAT models predicted sediment yield equally well in general, AnnAGNPS slightly underpredicted sediment yield compared with SWAT during calibration (slope: 0·44 versus 0·56), and SWAT slightly underpredicted sediment yield compared with AnnAGNPS during model validation (slope: 0·40 versus 0·70). However, model-predicted values for estimated slopes during model calibration and validation indicated no-significant (P values $<0{\cdot}001$ at $\alpha_{0{\cdot}05}$) difference.

Yuan *et al.* (2001) evaluated AnnAGNPS in the Mississippi Delta MSEA watershed. They used three years of measured data to compare model-predicted sediment yield from the watershed. The model-predicted monthly sediment yield showed good correlation ($R^2 = 0{\cdot}50$) with measured data. Das *et al.* (2007a), applied the AnnAGNPS model in the Grand River Basin in Canada. Their results demonstrated good agreement ($E = 0{\cdot}53$) between mean monthly observed and mean monthly model-predicted sediment yield values during the model calibration period. During the model validation period, model efficiencies decreased ($E = 0{\cdot}35$).

Santhi *et al.* (2001) calibrated and validated the SWAT model in the Bosque River watershed in Texas. The calibrated SWAT model showed E values ranging from 0·69 to 0·80 for monthly sediment yield. However, the validated model had decreased E values (0·23 to 0·70) for monthly sediment prediction compared with measured data. Kirsch *et al.* (2002) calibrated the SWAT model in the Rock River Basin watershed in Wisconsin. The calibrated SWAT model for Yahara and Mendota subwatersheds in the Rock River Basin had E values of 0·75 for annual sediment prediction compared with measured sediment data. They did not have enough sediment data for model validation. Jha *et al.* (2007), applied the SWAT model in the Raccoon River watershed in Iowa. The SWAT model predicted sediment yield with good to very good correlation and agreement ($R^2 = 0{\cdot}55$, $E = 0{\cdot}53$) during the model calibration period based on the performance ratings of Parajuli (2007). SWAT predictions improved during the model validation period ($R^2 = 0{\cdot}80$, $E = 0{\cdot}78$). All of these studies calibrated and validated models in the same watershed using different periods of measured data using only two statistical parameters (R^2 and E). In the present study, which used separate calibration and validation watersheds but the same time period, model performance in the calibration watershed was similar to or better than in the validation watershed and used five statistical parameters.

Total phosphorus

The AnnAGNPS and SWAT models showed good correlation and fair to good agreement ($R^2 = 0{\cdot}60$, $E = 0{\cdot}32$, $RMSE = 704$ for AnnAGNPS; $R^2 = 0{\cdot}70$, $E = 0{\cdot}68$, $RMSE = 487$ for SWAT) between predicted and measured mean monthly total phosphorus during model calibration (Table IV, Figure 5). The AnnAGNPS model-predicted monthly total phosphorus (Kg) estimated about 44% higher RMSE than SWAT in the calibration watershed. When both models were applied to the Goose Creek watershed for validation, the AnnAGNPS model predicted total phosphorus with good correlation but unsatisfactory agreement ($R^2 = 0{\cdot}77$, $E = -2{\cdot}38$, $RMSE = 476$) for mean monthly total phosphorus. The SWAT model consistently predicted total phosphorus with good correlation and agreement ($R^2 = 0{\cdot}60$, $E = 0{\cdot}63$, $RMSE = 178$) with measured mean monthly total phosphorus values (Table IV). The AnnAGNPS model predicted monthly total phosphorus (kg) estimated about 167% higher RMSE than SWAT in the validation watershed (Table IV). The estimated RSR values using both AnnAGNPS and SWAT models were excellent (RSR from 0·09 to 0·25). The calculated PBIAS values for the AnnAGNPS model simulation results for both calibration and validation study were unsatisfactory with overprediction bias (PBIAS from −117 to −128) whereas the SWAT model demonstrated excellent performance with PBIAS values from 3 to 13.

Several previous studies determined that the AnnAGNPS model overpredicted total phosphorus loss (Baginska *et al.*, 2003; Yuan *et al.*, 2005; Das *et al.*, 2007b). Yuan *et al.* (2005), evaluated the AnnAGNPS model in the Deep Hollow watershed of the Mississippi Delta Management Systems Evaluation Area project. Sensitivity analyses of the phosphorus component in the model were evaluated. AnnAGNPS overpredicted dissolved phosphorus loss (121%), still the model correlation of the simulated monthly total phosphorus was very good compared with the observed total phosphorus (R^2 of 0·81), model agreement was not reported. Das *et al.* (2007b) evaluated the nutrient component of the AnnAGNPS model in a watershed in Ontario, Canada. The model overpredicted total phosphorus, which was reported to have very good correlation and fair agreement ($R^2 = 0{\cdot}82$, $E = 0{\cdot}23$). None of these studies validated phosphorus prediction of the AnnAGNPS model in either another watershed or for a different simulation period. Yuan *et al.* (2005) reported that all the forms and processes involved in the phosphorous cycle are not completely and scientifically simulated by the AnnAGNPS model. Although the model calibration might improve some statistics of model output for total phosphorus, further research and in-depth investigation on fate and transport of the phosphorus component of the AnnAGNPS model is necessary (Yuan *et al.*, 2005; Das *et al.*, 2007b). Our study found overprediction of total phosphorus, which was attributed mainly to soluble phosphorus loss.

DOI: 10.1002/hyp

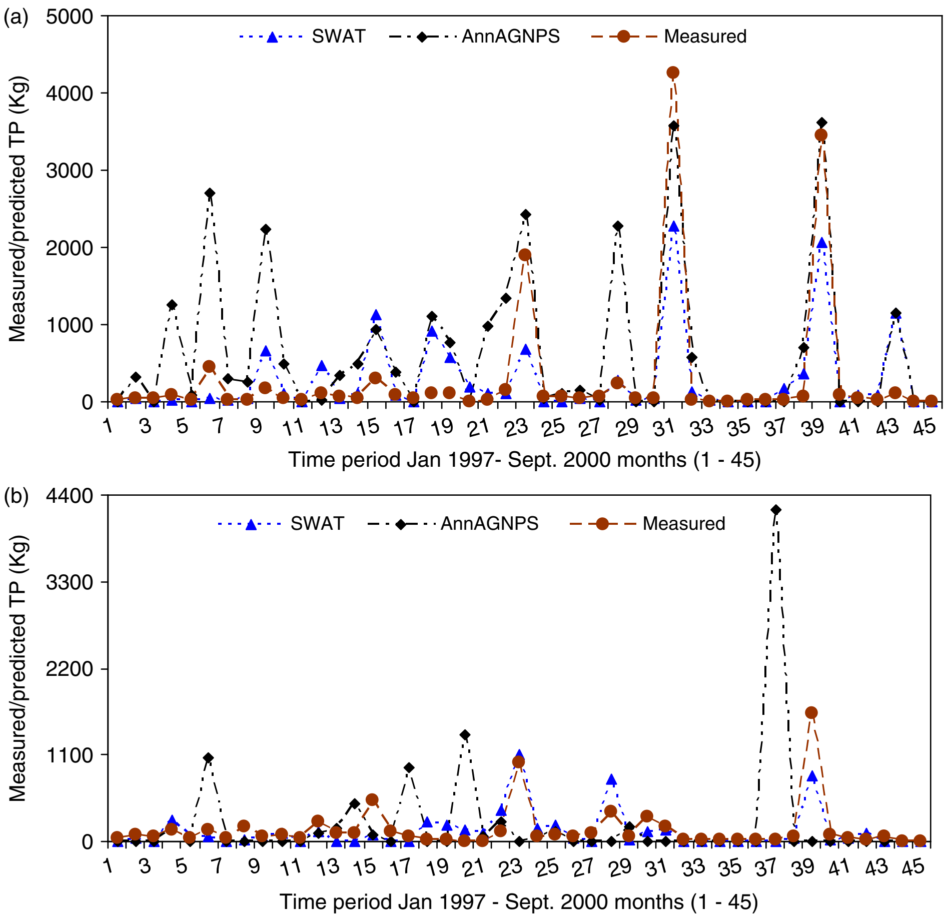

Figure 5. Time series measured and models predicted monthly average total phosphorus (TP) from January 1997 to September 2000 for (a) Red Rock Creek and (b) Goose Creek watersheds

Santhi *et al.* (2001) calibrated and validated the SWAT model in the Bosque River watershed in Texas. The calibrated SWAT model showed good agreement with E values ranging from 0·53–0·70, for monthly mean total phosphorus compared with mean monthly measured data. The validated model had fair to good agreement, with E values ranging from 0·39–0·72, for mean monthly total phosphorus prediction compared with mean monthly measured data. Several other studies successfully calibrated and validated the SWAT model for monthly total phosphorus prediction (Saleh and Du, 2004; White and Chaubey, 2005; Arabi *et al.*, 2006; Bracmort *et al.*, 2006; Cheng *et al.*, 2006; Tolson and Shoemaker, 2007). More applications of SWAT are described in Gassman *et al.* (2007).

Although no calibration parameters were used for phosphorus calibration, the SWAT model uses the QUAL2E stream flow process for model simulation, which reduces the amount of phosphorus leaving the watershed outlet or reach (Brown and Barnwell, 1987; Neitsch *et al.*, 2005). The SWAT model QUAL2E in-stream process with default parameters demonstrated fair to good model correlations and efficiencies in various outlets of the Upper Wakarusa watershed in Kansas (Parajuli, 2007). The AnnAGNPS model had no stream process routines in its Input Editor. AnnAGNPS assumes that organic phosphorus and insoluble inorganic phosphorus are associated with the clay fraction of the soil and that soil erosion is the mechanism transporting them from the soil profile to the water bodies. However, studies show that insoluble inorganic phosphorus, such as particulate phosphorus, could be lost through biological processes (Stuck *et al.*, 2001).

CONCLUSIONS

The objective of this research was to compare AnnAGNPS and SWAT model simulation results for surface flow, sediment yield, and total phosphorus using 45 months (January 1997 to September 2000) of measured data. The uniqueness of this study lies in its comparison of model performance in two separate yet similar watersheds. It is important to choose appropriate model to prioritize critical areas in the watershed.

According to the classifications of Parajuli (2007), this study concluded that both AnnAGNPS and SWAT models performed with fair to very good correlation (R^2 from 0·50 to 0·89) and fair to good agreement (E from 0·47 to 0·73) for surface flow and sediment yield when comparing model predictions with measured data during calibration and validation. SWAT also performed consistently well in terms of correlation (R^2 from 0·66 to 0·77)

DOI: 10.1002/hyp

and agreement (E from 0·63 to 0·68) for total phosphorus calibration and validation. Total phosphorus predictions during validation of AnnAGNPS were unsatisfactory to good (R^2 from 0·60 to 0·77, E from −2·38 to 0·32) with RMSE about 167% higher than that for SWAT for mean monthly total phosphorus prediction. AnnAGNPS over-predicted total phosphorus (PBIAS from −117 to −128), which was attributed mainly to soluble phosphorus loss. Because the KDHE has set TMDLs for eutrophication and silt for Cheney Lake, SWAT's ability to predict phosphorus in conjunction with BMPs is essential.

ACKNOWLEDGEMENTS

This material is based upon work supported by the Cooperative State Research, Education and Extension Services, US Dept. of Agriculture, under agreement no. 2006-51130-03703. We acknowledge the contributions of Ms Lisa French, project coordinator, and Mr Howard Miller, public relations coordinator, at Cheney Lake Watershed Inc., Hutchinson, KS; Mr Daniel S. Moore, hydraulic engineer, at USDA-NRCS, Portland, OR; Mr Vance Justice, Jr, hydrologist, and Dr Ron L. Bingner, agricultural engineer, at USDA-ARS, National Sedimentation Laboratory, Oxford, MS; and Mr Larry M. Pope, project chief, at USGS, KS. Contribution no. 08-249-J from the Kansas Agricultural Experiment Station.

REFERENCES

Abbaspour KC, Yang J, Maximov I, Siber R, Bogner K, Mieleitner J, Zobrist J, Srinivasan R. 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of Hydrology* **333**(2–4): 413–430.

Allen RG. 1986. A penman for all seasons. *Journal of Irrigation and Drainage Engineering ASCE* **112**(4): 348–368.

Allen RG, Jensen ME, Wright JL, Burman RD. 1989. Operational estimates of evapotranspiration. *Agronomy Journal* **81**: 650–662.

Arabi M, Govindaraju RS, Hantush MM. 2006. Cost-effective allocation of watershed management practices using a genetic algorithm. *Water Resource Research* **42**: W10 429. DOI:10,1029/2006WR004 931.

Arabi M, Frakenberger JR, Engel BA, Arnold JG. 2007. Representation of agricultural conservation practices with SWAT. *Hydrological Processes* **22**(16): 3042–3055. DOI:10·1002/hyp.6890.

Arnold JG, Srinivasan R, Muttiah RS, Williams JR. 1998. Large area hydrologic modelling and assessment part I: model development. *Journal of American Water Resources Association* **34**(1): 73–89.

Baginska B, Milne-Home W, Cornish PS. 2003. Modelling nutrient transport in Currency Creek, NSW with AnnAGNPS and PEST. *Environmental Modelling & Software* **18**: 801–808.

Bagnold RA. 1966. An approach to the sediment transport problem from general physics. Prof. Paper 422-J. US Geological Survey, Reston, VA.

Bagnold RA. 1977. Bedload transport in natural rivers. *Water Resources Research* **13**: 303–312.

Bhuyan SJ, Mankin KR, Koelliker JK. 2003. Watershed-scale AMC selection on for hydrologic modelling. *Transactions of the ASAE* **46**(2): 303–310.

Bingner RL, Theurer FD. 2003. AnnAGNPS technical processes documentation, Version 3·2. USDA-ARS, National Sedimentation Laboratory: Oxford, MS.

Bingner RL, Theurer FD, Yongping Y. 2007. *AnnAGNPS technical processes documentation, Version 4·0*. USDA-ARS, National Sedimentation Laboratory: Oxford, MS.

Bockhold AR, Thompson AL, Baffaut C, Sadler EJ. 2006. Evaluating BMPs in a Claypan watershed. ASABE Meeting Paper number: 062 114. ASABE, St. Joseph, MI.

Borah DK, Bera M. 2003. Watershed-scale hydrologic and nonpoint-source pollution models: Review of mathematical bases. *Transactions of the ASAE* **46**(6): 1553–1566.

Borah DK, Yagow G, Saleh A, Barnes PL, Rosenthal W, Krug EC, Hauck LM. 2006. Sediment and nutrient modelling for TMDL development and implementation. *Transactions of the ASAE* **49**(4): 967–986.

Bosch DD, Strickland T, Sullivan DG, Wauchope D, Lowrance R, Potter T. 2005. SWAT Application for Conservation Effects Assessment in the Southeastern Coastal Plain. ASAE Meeting paper number: 052 040. ASAE, St. Joseph, MI.

Bracmort KS, Arabi M, Frankenberger JR, Engel BA, Arnold JG. 2006. Modelling long-term water quality impact of structural BMPs. *Transactions of the ASAE* **49**(2): 367–374.

Brown LC, Barnwell TO Jr. 1987. The enhanced water quality models QUAL2E and QUAL2E-UNCAS documentation and user manual. EPA document EPA/600/2-87/007. USEPA, Athens, GA.

Cheney Lake Watershed Management Plan (CLWMP). 2003–2007, 2008. Cheney Lake Watershed Inc. 18 east 7th street—South Hutchinson, KS, 67 505.

Cheng H, Ouyang W, Hao F, Ren X, Yang S. 2006. The nonpoint-source pollution in livestock-breeding areas of the Heihe River basin in Yellow River. *Stochastic Environmental Research and Risk Assessment* **21**(3): 213–221. DOI:10·1007/s00 477-006-0057-2.

Chu TW, Shirmohammadi A. 2004. Evaluation of the SWAT model's hydrology component in the piedmont physiographic region of Maryland. *Transaction of the ASAE* **47**(4): 1057–1073.

Chu TW, Shirmohammadi A, Abbot L, Sadeghi A. 2005. Watershed level BMP Evaluation with SWAT model. ASAE Meeting Paper number: 052 098. ASAE, St. Joseph, MI.

Cronshey RG, Theurer FG. 1998. AnnAGNPS-non point pollutant loading model. In *Proceedings First Federal Interagency Hydrologic Modelling Conference*. Las Vegas, NV.

Das S, Rudra RP, Gharabaghi B, Goel PK, Singh A, Ahmed I. 2007a. Comparing the Performance of SWAT and AnnAGNPS Model in a Watershed in Ontario. Watershed Management to Meet Water Quality Standards and TMDLS (Total Maximum Daily Load) Proceedings. ASABE Publication number: 701P0207. ASABE, St. Joseph, MI.

Das S, Rudra RP, Gharabaghi B, Singh A, Ahmed SI, Goel PK. 2007b. Evaluation of nutrient component of AnnAGNPS model in a watershed in Ontario. Watershed Management to Meet Water Quality Standards and TMDLS (Total Maximum Daily Load) Proceedings. ASABE Publication number: 701P0207. ASABE, St. Joseph, MI.

Das S, Rudra RP, Goel PK, Gharabaghi B, Gupta N. 2006. Evaluation of AnnAGNPS in cold and temperate regions. *Water Science and Technology* **53**(2): 263–270.

Eckhardt K. 2005. How to construct recursive digital filters for baseflow separation. *Hydrological Processes* **19**(2): 507–515.

Feyereisen GW, Strickland TC, Bosch DD, Sullivan DG. 2007. Evaluation of SWAT manual calibration and input parameter sensitivity in the Little River watershed. *Transactions of the ASABE* **50**(3): 843–855.

Gassman PW, Reyes MR, Green CH, Arnold JG. 2007. The Soil and Water Assessment Tool: historical development, applications, and future research directions. *Transactions of the ASABE* **50**(4): 1211–1250.

Gebremeskel S, Rudra RP, Gharabaghi B, Das S, Singh A, Bai H, Jiang G. 2005. Assessing the performance of various hydrological models in the Canadian Great lakes basin. Watershed Management to Meet Water Quality Standards and Emerging TMDL (Total Maximum Daily Load) Proceedings. ASAE Publication number: 701P0105. ASABE, St. Joseph, MI.

Geter WF, Theurer FD. 1998. AnnAGNPS-RUSLE sheet and rill erosion. In *Proceedings of 1st Federal Interagency Hydrologic Modelling Conference*. Interagency Advisory Committee on Water Data, Subcommittee on Hydrology. Washington, DC.

Gikas GD, Yiannakopoulou T, Tsihrintzis VA. 2006. Modelling of non-point source pollution in a Mediterranean drainage basin. *Environmental Modelling & Assessment* **11**(3): 219–233.

Gitau MW, Chaubey I, Nelson MA, Pennigton JH. 2007. Analysis of BMP and land use change effects in a Nrthwest Arkansas agricultural watershed. ASABE Meeting Paper number: 072 244. ASABE, St. Joseph, MI.

Gitau MW, Veith TL, Gburek WJ. 2003. Optimizing Best Management Practice Selection to Increase Cost-effectiveness. ASAE Meeting Paper number: 032 110. ASAE, St. Joseph, MI.

Gordon LM, Bennett SJ, Bingner RL, Theurer FD, Alonso CV. 2007. Simulating ephemeral gully erosion in AnnAGNPS. *Transactions of the ASABE* **50**(3): 857–866.

Gosain AK, Rao S, Srinivasan R, Reddy NG. 2005. Return-flow assessment for irrigation command in the Palleru river basin using SWAT model. *Hydrological Processes* **19**: 673–682.

Gupta HV, Sorooshian S, Yapo PO. 1999. Status of automatic calibration for hydrologic models: comparison with multi-level expert calibration. *Journal of Hydrologic Engineering* **4**(2): 135–143.

Hargreaves GH, Samani ZA. 1985. Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture* **1**: 96–99.

Hong HS, Huang JL, Zhang LP, Du PF. 2005. Modelling pollutant loads and management alternatives in Jiulong River watershed with AnnAGNPS. *Huan Jing Ke Xue* **26**(4): 63–69.

Jensen ME, Burman RD, Allen RG (eds). 1990. Evapotranspiration and irrigation water requirements. ASCE Manuals and Reports on Engineering Practice No. 70, ASCE: NY.

Jha MK, Gassman PW, Arnold JG. 2007. Water quality modelling for the Raccoon river watershed using SWAT. *Transactions of the ASABE* **50**(2): 479–493.

Kansas Department of Agriculture. 2004a. Kansas Agricultural Statistics. Kansas Farm Facts 2004. Available at http://www.nass.usda.gov/ks/ffacts/2004/pdf/ffpdf.html as of 30 August 2007.

Kansas Department of Health and Environment. 2004b. Bureau of water. Kansas Surface Water Quality Standards and Supporting Documents. Available at http://www.kdhe.state.ks.us/water/download/kwqs_plus_supporting.pdf as of 21 December 2007.

Kansas Department of Health and Environment. 2006. Bureau of Environmental Field Services. Kansas Water Quality Assessment (305(b) Report). Topeka, KS.

Kansas Water Office. 2004. Reports and Publications. Fact Sheets. Kansas Water Office Fact Sheets. Available at http://www.kwo.org/as of 29 August 2007.

King KW, Arnold JG, Bingner RL. 1999. Comparison of Green-Ampt and curve number methods on Goodwin Creek watershed using SWAT. *Transactions of the ASAE* **42**(4): 919–925.

Kirsch K, Kirsch A, Arnold JG. 2002. Predicting sediment and phosphorus loads in the Rock River basin using SWAT. *Transactions of the ASABE* **45**(6): 1757–1769.

Knisel WG (Ed.). 1980. A field scale model for chemical, runoff, and erosion from agricultural management systems. Conservation Service Report 26, US Department of Agriculture: Washington DC.

Kyoung JL, Engel BA, Tang Z, Choi J, Kim KS, Muthukrishnan S, Tripathy D. 2005. Automated Web GIS based Hydrograph Analysis Tool, WHAT. *Journal of American Water Resources Association* **41**(6): 1407–1416.

Legates DR, McCabe Jr. GJ. 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research* **35**(1): 233–241.

Leonard RA, Knisel WG, Still DA. 1987. GLEAMS: Groundwater Loading Effects of Agricultural Management Systems. *Transactions of the ASAE* **30**(5): 1403–1418.

Li X, Frees L, Moore DS, Wang S. 2006. Calibrating the AnnAGNPS model in the Red Rock Creek watershed. Unpublished report. Department of Geography, University of Kansas, Lawrence, KS.

Licciardello F, Zema DA, Zimbone SM, Bingner RL. 2007. Runoff and soil erosion evaluation by the AnnAGNPS model in a small Mediterranean watershed. *Transactions of the ASABE* **50**(5): 1585–1593.

Mein RG, Larson CL. 1973. Modelling infiltration during a steady rain. *Water Resources Research* **9**(2): 384–394.

Migliaccio KW, Srivastava P. 2007. Hydrologic components of watershed scale models. *Transactions of the ASABE* **50**(5): 1695–1703.

Monteith JL. 1965. *Evapotranspiration and the environment. The state and movement of water in living organisma, XIXth symposium*. Society for Experimental Biology, Swansea, Cambridge University Press.

Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE* **50**(3): 885–900.

Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models: Part I. A discussion of principles. *Journal of Hydrology* **10**(3): 282–290.

Nasr A, Bruen M, Jordan P, Moles R, Kiely G, Byrne P. 2007. A comparison of SWAT, HSPF, and SHETRAN/GOPC for modelling phosphorus export from three catchments in Ireland. *Water Research* **41**(5): 1065–1073.

Neitsch SL, Arnold JG, Kiniry JR, Williams JR, King KW. 2002. Soil and Water Assessment Tool Theoretical Documentation (Version 2000). Grassland, soil and Water Research Laboratory, Agricultural Research Service, Temple, TX.

Neitsch SL, Arnold JG, Kiniry JR, Williams JR. 2005. Soil and Water Assessment Tool (SWAT), Theoretical Documentation. Blackland Research Center, Grassland, Soil and Water Research Laboratory, Agricultural Research Service: Temple, TX.

Parajuli PB. 2007. SWAT bacteria sub-model evaluation and application. PhD dissertation. Department of Biological and Agricultural Engineering, Kansas State University, Manhattan, KS.

Parajuli PB, Mankin KR, Barnes PL. 2006. Calibration and Validation of SWAT/Microbial sub-model 2005 for Fecal Coliform Bacteria Prediction on a Grazed Watershed. ASABE Paper No. 062 196. St. Joseph, MI.

Parajuli PB, Mankin KR, Barnes PL. 2007. New Methods in Modelling Source-Specific Bacteria at Watershed Scale Using SWAT. Watershed Management to meet Water Quality Standards and TMDLs (Total Maximum Daily Load) Proceedings. ASABE Publication No. 701P0207. ASABE: St. Joseph, MI.

Parajuli PB, Mankin KR, Barnes PL. 2008. Source specific fecal bacteria modelling using SWAT model. *Bioresource Technology* **100**(2): 953–963. DOI:10·1016/j.biortech.2008·06·045.

Polyakov V, Fares A, Kubo D, Jacobi J, Smith C. 2007. Evaluation of a non-point source pollution model, AnnAGNPS, in a tropical watershed. *Environmental Modelling & Software* **22**(11): 1617–1627.

Priestley CHB, Taylor RJ. 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review* **100**: 81–92.

Qi C, Grunwald S. 2005. GIS-Based Hydrologic Modelling in the Sandusky Watershed Using SWAT. *Transactions of the ASAE* **48**(1): 160–180.

Ramanarayanan TS, Williams JR, Dugas WA, Hauck LM, McFarland AMS. 1997. Using APEX to identify alternative practices for animal waste management. ASAE Paper No. 972 209. ASAE, St. Joseph, MI.

Renard KG, Foster GR, Weesies GA, McCool DK, Yoder DC, coordinators. 1997. Predicting Soil Erosion by Water: A Guide to Conservation Planning with the Revised Universal Soil Loss Equation (RUSLE). U.S. Department of Agriculture, Agriculture Handbook No. 703.

Sadeghi A, Kwang Y, Graff C, McCarty G, McConnell L, Shirmohammadi A, Hively D, Sefton KA. 2007. Assessing the Performance of SWAT and AnnAGNPS Models in a Coastal Plain Watershed, Choptank River, Maryland, U.S.A. ASABE Meeting, Paper No. 072 032. ASABE, St. Joseph, MI.

Saleh A, Du B. 2004. Evaluation of SWAT and HSPF within BASINS program for the upper North Bosque River watershed in central Texas. *Transactions of the ASAE* **47**(4): 1039–1049.

Saleh A, Arnold JG, Gassman PW, Hauck LM, Rosenthal WD, Williams JR, MacFarland AMS. 2000. Application of SWAT for the Upper North Bosque River watershed. *Transactions of the ASAE* **43**(5): 1077–1087.

Santhi C, Arnold JG, Williams JR, Hauck LM, Dugas WA. 2001. Application of a watershed model to evaluate management effects on point and non-point source pollution. *Transactions of the ASAE* **44**(6): 1559–1570.

Sharpley AN, Williams JR, eds. 1990. EPIC-Erosion Productivity Impact Calculator, 1. model documentation. U. S. Department of Agriculture, Agricultural Research Service, Technical Bulletin, 1768.

Shirmohammadi A, Chaubey I, Harmel RD, Bosch DD, Muñoz-Carpena R, Dharmasri C, Sexton A, Arabi M, Wolfe ML, Frankenberger J, Graff C, Sohrabi TM. 2006. Uncertainty in TMDL Models. *Transactions of the ASAE* **49**(4): 1033–1049.

Sloan PG, Moore ID, Coltharp GB, Eigel JD. 1983. Modelling surface and subsurface stormflow on steeply sloping forested watersheds. Water Resources Institute Report No. 142. University of Kentucky, Lexington, KY.

Spruill CA, Workman SR, Taraba JL. 2000. Simulation of daily and monthly stream discharge from small watersheds using the SWAT model. *Transactions of the ASAE* **43**(6): 1431–1439.

Stuck JD, Izuno FT, Campbell KL, Bottcher AB, Rice RW. 2001. Farm-level studies of particulate phosphorus transport in the Everglades Agricultural Area. *Transactions of the ASAE* **44**(5): 1105–1116.

Theurer FD, Clarke CD. 1991. Wash load component for sediment yield modelling. *Proceedings of the Fifth Federal Interagency Sedimentation Conference*, Las Vegas, NV.

Theurer FD, Cronshey RG. 1998. AnnAGNPS-reach routing processes. *Proceedings First Federal Interagency Hydrologic Modelling Conference*, Las Vegas, NV.

Tolson BA, Shoemaker CA. 2007. Cannonsville reservoir watershed SWAT2000 model development, calibration, and validation. *Journal of Hydrology* **337**(1–2): 68–86.

US Department of Agriculture (USDA), Natural Resources Conservation Service. 2005. Soil Data Mart. Available at http://soildatamart.nrcs.usda.gov/Default.aspx as of 29 october 2007.

US Geological Survey (USGS). 1999. National Elevation Dataset. Available at http://gisdasc.kgs.ukans.edu/as of 29 October, 2007.

Vache KB, Eilers JM, Santelmann MV. 2002. Water quality modelling of alternative agricultural scenarios in the U.S. Corn Belt. *Journal of American Water Resources Association* **38**(3): 773–787.

Van Liew MW, Arnold JG, Garbrecht JD. 2003. Hydrologic Simulation on agricultural watersheds choosing between two models. *Transactions of the ASABE* **46**(6): 1539–1551.

Varanou E, Gkouvatsou E, Baltas E, Mimikou M. 2002. Quantity and quality integrated catchment modelling under climate change with user of soil and water assessment tool model. ASCE *Journal of Hydrological Engineering* **7**(3): 228–244.

Wang X, Melesse AM, Yang W. 2006. Influences of potential evapotranspiration estimation methods on SWAT's hydrologic simulation in a northwestern Minnesota watershed. *Transactions of the ASABE* **49**(6): 1755–1771.

Wang S, Kang S, Zhang L, Li F. 2007. Modelling hydrological response to different land-use and climate change scenarios in the Zamu River basin of northwest China. *Hydrological Processes* **22**(14): 2502–2510. DOI: 10·1002/hyp.6846.

White KL, Chaubey I. 2005. Sensitivity analysis, calibration, and validations for a multisite and multivariable SWAT model. *Journal of American Water Resources Association* **41**(5): 1077–1089.

White KL, Chaubey I, Haggard BE, Matlock MD. 2004. Comparison of two methods for modelling monthly TP yield from a watershed. ASABE Meeting Paper No. 042 162. ASABE: Joseph, MI.

Wilcox BP, Rawls WJ, Brakensiek DL, Wight JR. 1990. Predicting runoff from rangeland catchments: A comparison of two models. *Water Resources Research* **26**: 2401–2410.

Williams JR, Jones CA, Dyke PT. 1984. A modelling approach to determining the relationship between erosion and soil productivity. *Transactions of the ASAE* **27**(1): 129–144.

Willmot CJ. 1984. On the evaluation of model performance in physical geography. In *Spatial Statistics and Models*. Gaile GL (ed). D. Reidel; 443–460.

Wischmeier WH, Smith DD. 1978. Predicting rainfall erosion losses: a guide to conservation planning. Agriculture Handbook 282. United States Department of Agriculture (USDA), Agricultural Research Service (ARS). US Govt Printing Office: Washington, DC.

Young RA, Onstead CA, Bosch DD, Anderson WP. 1987. AGNPS, Agricultural Non-Point-Source Pollution Model. A Watershed Analysis Tool. USDA Conservation Research Report 35, Washington DC.

Young RA, Onstead CA, Bosch DD, Anderson WP. 1989. AGNPS: a nonpoint source pollution model for evaluating agricultural watersheds. *Journal of Soil Water Conservation* **44**(2): 168–173.

Yuan Y, Bingner RL, Rebich RA. 2001. Evaluation of AnnAGNPS on Mississippi Delta MSEA watersheds. *Transactions of the ASAE* **44**(5): 1183–1190.

Yuan Y, Bingner RL, Theurer FD. 2006. Subsurface Flow Component for AnnAGNPS. *Transactions of the ASABE* **22**(2): 231–241.

Yuan Y, Bingner RL, Theurer FD, Kolian S. 2007. Water quality simulation of Rice/Crawfish field ponds within Annualized AGNPS. *Transactions of the ASABE* **23**(5): 585–595.

Yuan Y, Bingner RL, Theurer FD, Moore PA, Rebich RA. 2005. Phosphorus component in AnnAGNPS. ASAE Meeting Paper number: 052 169. ASAE, St. Joseph, MI.

Yuan Y, Dabney S, Bingner RL. 2002. Cost/benefit analysis of agricultural BMPs for sediment reduction in the Mississippi Delta. *Journal of Soil and Water Conservation* **57**(5): 259–267.