JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

Vol. 44, No. 6 AMERICAN WATER RESOURCES ASSOCIATION December 2008

HYDROLOGIC CALIBRATION AND VALIDATION OF SWAT IN A SNOW-DOMINATED ROCKY MOUNTAIN WATERSHED, MONTANA, U.S.A.[1]

Robert S. Ahl, Scott W. Woods, and Hans R. Zuuring[2]

ABSTRACT: The Soil and Water Assessment Tool (SWAT) has been applied successfully in temperate environments but little is known about its performance in the snow-dominated, forested, mountainous watersheds that provide much of the water supply in western North America. To address this knowledge gap, we configured SWAT to simulate the streamflow of Tenderfoot Creek (TCSWAT). Located in central Montana, TCSWAT represents a high-elevation watershed with ~85% coniferous forest cover where more than 70% of the annual precipitation falls as snow, and runoff comes primarily from spring snowmelt. Model calibration using four years of measured daily streamflow, temperature, and precipitation data resulted in a relative error (RE) of 2% for annual water yield estimates, and mean paired deviations (Dv) of 36 and 31% and Nash-Sutcliffe (NS) efficiencies of 0.90 and 0.86 for monthly and daily streamflow, respectively. Model validation was conducted using an additional four years of data and the performance was similar to the calibration period, with RE of 4% for annual water yields, Dv of 43% and 32%, and NS efficiencies of 0.90 and 0.76 for monthly and daily streamflow, respectively. An objective, regression-based model invalidation procedure also indicated that the model was validated for the overall simulation period. Seasonally, SWAT performed well during the spring and early summer snowmelt runoff period, but was a poor predictor of late summer and winter base flow. The calibrated model was most sensitive to snowmelt parameters, followed in decreasing order of influence by the surface runoff lag, ground water, soil, and SCS Curve Number parameter sets. Model sensitivity to the surface runoff lag parameter reflected the influence of frozen soils on runoff processes. Results indicated that SWAT can provide reasonable predictions of annual, monthly, and daily streamflow from forested montane watersheds, but further model refinements could improve representation of snowmelt runoff processes and performance during the base flow period in this environment.

(KEY TERMS: SWAT; simulation; runoff; snow hydrology; forests; watershed management.)

Ahl, Robert S., Scott W. Woods, and Hans R. Zuuring, 2008. Hydrologic Calibration and Validation of SWAT in a Snow-Dominated Rocky Mountain Watershed, Montana, U.S.A. *Journal of the American Water Resources Association* (JAWRA) 44(6):1411-1430. DOI: 10.1111/j.1752-1688.2008.00233.x

[1]Paper No. JAWRA-07-0151-P of the *Journal of the American Water Resources Association* (JAWRA). Received October 30, 2007; accepted March 21, 2008. **Discussions are open until June 1, 2009**.

[2]Respectively, Adjunct Professor, Department of Forest Management, College of Forestry and Conservation, The University of Montana, 32 Campus Drive., Missoula, Montana 59812; Associate Professor, Department of Ecosystem Sciences and Conservation, College of Forestry and Conservation, The University of Montana, Montana; and Chair, Department of Forest Management, College of Forestry and Conservation, The University of Montana, Montana (E-Mail/Ahl: rob.ahl@umontana.edu).

INTRODUCTION

Watershed simulation models are commonly used to investigate interactions of the hydrologic cycle. Once calibrated, watershed models make it possible to evaluate the impacts of natural or management-induced environmental changes in a way that cannot be done through field experiments and direct observation. With the use of simulations, management uncertainty may be reduced by evaluating scenarios before they are initiated. Following early conceptualizations (Crawford and Linsley, 1966), contemporary watershed models have become increasingly comprehensive tools (Singh, 1995) that are vital components of integrated environmental assessments (Jakeman and Lechter, 2003). To become versatile, however, models must be evaluated over a broad range of settings and applications. While many watershed models are currently available, the Soil and Water Assessment Tool (SWAT) is one of the more widely used models, having been applied in many countries around the world (Gassman *et al.*, 2005, 2007). SWAT is a physically based, semi-distributed, continuous time, watershed model. This model was developed to assess the impact of land management and climate patterns on water supply and nonpoint source pollution in large, complex watersheds with varying soil, landcover, and management conditions over long periods (Arnold *et al.*, 1993; Srinivasan and Arnold, 1994; Arnold *et al.*, 1998; Srinivasan *et al.*, 1998; Di Luzio *et al.*, 2002; Arnold and Fohrer, 2005).

SWAT represents the culmination of over 30 years of modeling efforts within the United States Department of Agriculture, Agriculture Research Service (USDA-ARS), and is an outgrowth from previous models such as the Simulator for Water Resources in Rural Basins (SWRRB) (Williams *et al.*, 1985), and Ground-Water Loading Effects of Agricultural Management Systems (GLEAMS) (Leonard *et al.*, 1987). By combining proven agricultural management models with simple, yet realistic routing components, SWAT expands the spatial domain of past field-scale models to that of watersheds. SWAT was designed with a flexible architecture that uses readily available data to describe the physical and climatic characteristics of a watershed. Using physically based inputs, SWAT can estimate reasonable results in remote ungaged basins where calibration is not possible. The code is computationally efficient, allowing continuous simulation over large areas and long time periods. By simulating watershed processes over the long-term, the impact of management practices can be evaluated (Arnold and Fohrer, 2005). In the United States, SWAT is currently being used in total maximum daily load (TMDL) assessments (Di Luzio *et al.*, 2002), to investigate the effectiveness of best management and conservation practices (Arabi *et al.*, 2006), and to evaluate the hydrologic effects of climate change (Jha *et al.*, 2006).

In the Rocky Mountain region of western North America, most of the streamflow and public water supply comes from snowmelt in forested, mountainous watersheds. Changing the extent, composition, and configuration of forest cover in these watersheds may alter the timing, magnitude, and spatial distribution of snow accumulation and snowmelt, thus altering water yield and runoff (Bosch and Hewlett, 1982; Stednick, 1996). Timber harvest, fires, and other stand-replacing disturbances are often associated with increased streamflow from reduced canopy interception and evapotranspiration in cleared and low density stands (Golding and Swanson, 1978; Troendle, 1983; Troendle and King, 1985, 1987; Pomeroy *et al.*, 2002). Understanding the interactions among forest structure, snow accumulation, snowmelt, and streamflow generation is therefore an integral component of effective water resources management in the Rocky Mountain region. Because the possibilities for evaluating management scenarios through experimental manipulation are limited, reliable modeling tools are essential for providing otherwise unattainable assessment data. To predict the streamflow hydrograph in these environments, models must adequately describe the effects of both forest cover and snowmelt dynamics on runoff.

Given its development philosophy and model architecture, SWAT may be useful for evaluating the effects of forest management on water resources in the Rocky Mountain region. SWAT's ability to simulate snowmelt processes is supported by algorithms that account for the effects of elevation on snowmelt (Fontaine *et al.*, 2002). While the effectiveness of these algorithms was assessed in a Minnesota watershed with low relief and mixed landcover (Wang and Melesse, 2005), SWAT has not been evaluated in the steep, high-elevation, mostly forested watersheds that are typical of the Rocky Mountain region. To address this research gap, we set up a SWAT model to simulate the streamflow of a forested upland watershed with snowmelt driven runoff. The research site is located in central Montana, and embodies biophysical characteristics that are representative of the high elevation, forested watersheds east of the Continental Divide in much of southwest Alberta, Montana and Wyoming.

The configured research watershed is situated within the Tenderfoot Creek Experimental Forest (TCEF), and is instrumented with a well-distributed array of climate and stream gauging stations (McCaughey, 1996). Our study thus provides an opportunity for unusually detailed calibration and validation of the SWAT model in an environment where little is known about its ability to simulate streamflow.

METHODS

Study Area

Tenderfoot Creek is a headwater tributary of the Missouri River that drains the Little Belt Mountains of central Montana in a westerly direction (Figure 1). Our watershed model encompasses the upper 2,251 ha of Tenderfoot Creek and is referred to as TCSWAT. The watershed contained within TCSWAT has a dendritic stream network that is bisected by a steep canyon along the main channel. A headwater reach and two major tributaries on each north and south aspect contribute flow to the watershed outlet. Geologically, TCSWAT is underlain by Precambrian meta-sedimentary rocks of the Belt Supergroup (Alt and Hyndman, 1986). Soils are loamy skeletal, mixed Typic Cryochrepts and clayey, mixed Aquic Cryoboralfs (Farnes and McCaughey, 1995). Slopes average 8°, although steeper 25° to 30° hillsides occur along the main canyon and upper headwalls. Elevations range from 1,991 m at the watershed outlet to 2,420 m at the highest point on the watershed boundary.

The forested vegetation mosaic of the study area, and Rocky Mountain region in general, is strongly influenced by periodic fires (Arno and Fiedler, 2005). In the past 400 years, most of the forested area within TCSWAT has burned at least once, but it has been nearly 120 years since the last major fire (Barrett, 1993). Approximately 65% of TCSWAT is composed of lodgepole pine (*Pinus contorta*), which represents the most recently initiated forest cover. Another 20% of the landcover in the watershed is comprised of older stands containing a mixture of shade tolerant subalpine fir (*Abies lasiocarpa*) and Englemann spruce (*Picea engelmannii*) as well as decadent lodgepole pine (*Pinus contorta*). Disturbed and low density conifer forest stands constitute 10% of the watershed. Shrubby meadows and small riparian areas (1%) surround many of the creek bottoms, while drier grassland openings (1%) tend to occur on higher ground. Talus slopes (2%) line the main canyon and some high, exposed ridges.

The TCSWAT watershed has a montane continental climate and the elevation-weighted mean annual precipitation is close to 800 mm. Almost 70% of the annual precipitation falls as snow, which accumulates in the watershed between November and May. The annual peak runoff occurs in response to snowmelt between May and early June. This creates a sharply peaked streamflow hydrograph with steep rising and falling limbs. Base flow begins in mid July and

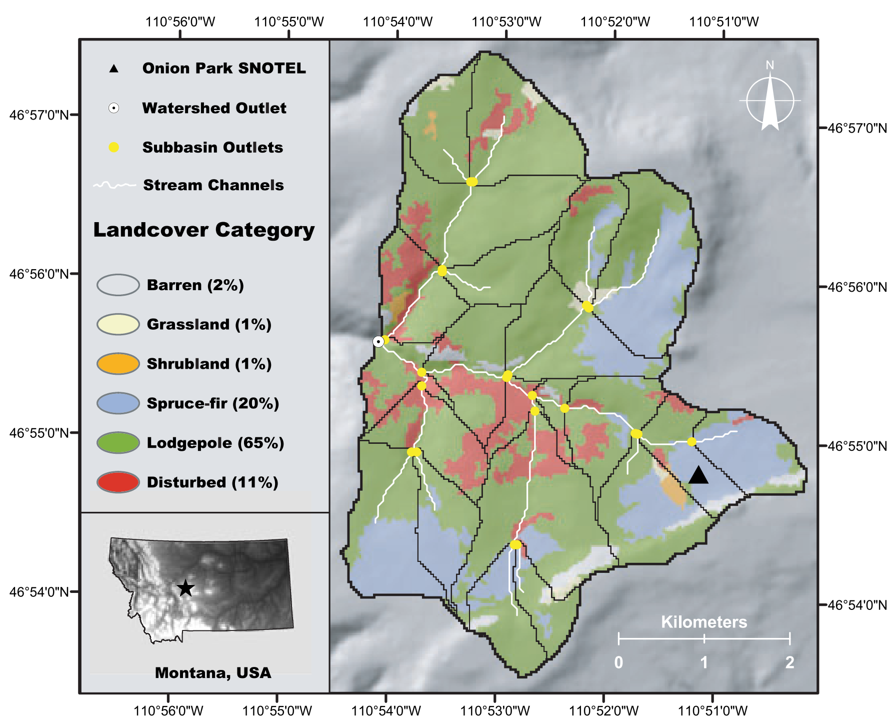

FIGURE 1. Delineation, Configuration, and Landcover Characteristics of the Tenderfoot Creek SWAT Model (TCSWAT), Located in the Little Belt Mountains of Central Montana, USA.

persists through April. Summer rainfall is limited, generally occurring as brief, high intensity thunderstorms, and contributes little to the total runoff.

The SWAT

SWAT has been in the public domain for over a decade, and several versions are currently available. In this study, we used SWAT2005 (Neitsch *et al.*, 2005) in conjunction with the ArcView SWAT interface (AVSWAT) (Di Luzio *et al.*, 2004). AVSWAT is a GIS-based graphical user interface that facilitates watershed delineation, subdivision, and initial parameterization. AVSWAT also incorporates sensitivity analysis, auto-calibration, and uncertainty assessment procedures (van Griensven *et al.*, 2006). These routines were formulated to automate SWAT calibration in a temperate, rainfall-dominated European watershed (van Griensven and Bauwens, 2003) and have not been applied in a mountainous, snowmelt driven watershed like TCSWAT.

Watershed and Subwatershed Delineation

Delineation of TCSWAT was based on a digital elevation model (DEM) with 30 m grid cell resolution (Gesch *et al.*, 2002), and coordinates of an outlet defined by a flume at the downstream extent of the Experimental Forest (Figure 1). Subwatersheds are the smallest areal unit that SWAT considers, and subdivision relative to watershed complexity is necessary to capture spatially explicit processes (Mangurerra and Engel, 1998; Haverkamp *et al.*, 2002; Arabi *et al.*, 2006). By specifying a critical source area delineation threshold with a relative area of 1.5% (30 ha), the watershed was configured with 22 subwatersheds (Figure 1). This level of segmentation created subwatersheds that each represent an average of 4.5% of the delineated basin, which is approximately what Arabi *et al.* (2006) suggested as the minimum watershed subdivision threshold required to detect the influence of basin-wide management practices.

Subwatershed Characterization

Model performance can only be assessed by comparison to known or observed data. To reduce model input uncertainty, field measurements were taken onsite whenever possible. In particular, measurements of temperature, precipitation, snow accumulation and melt, vegetation, and stream channel characteristics were all recorded within the research watershed itself. The observed streamflow record was also scrutinized before its use for calibration and validation.

Elevation Bands. Snow accumulation and melt processes in mountainous watersheds are strongly influenced by changes in elevation (Pomeroy and Burn, 2001). Fontaine *et al.* (2002) showed that definition of elevation bands within the model's subwatersheds can enhance simulation performance in watersheds with complex topography and large elevation gradients. To account for a 425 m elevation gain within the watershed, three elevation bands were defined in each of the 22 TCSWAT subwatersheds.

Climate Data and Lapse Rates. The Onion Park snow telemetry (SNOTEL 1008) site is located within TCSWAT and has a record of daily minimum and maximum temperature and precipitation from October 1, 1993 to the present (Figure 1). Data from this SNOTEL site and the National Weather Service cooperative station located approximately 40 straight-line kilometers southeast of the research site, at White Sulphur Springs (Station no. 248927) were used to estimate the local temperature and precipitation lapse rates of −7.14 °C/km and 765 mm/km, respectively (Table 1). Once established, these values were used to parameterize SWAT, and then maintained throughout the calibration.

Stream Channel Characteristics. The flow routing routines within SWAT require information concerning main channel dimensions within each subwatershed (Neitsch *et al.*, 2005). While geo-processing capabilities of the AVSWAT interface can be used to estimate channel dimensions based on the input DEM, we physically measured channel width and depth at three representative locations within each delineated subwatershed and used the arithmetic mean values as model input parameters.

Landcover Characteristics. SWAT uses a vegetation database to define parameters for mapped landcover categories, but options for representing

TABLE 1. Mean Annual Temperature (TMP), and Precipitation (PCP), Temperature and Precipitation Lapse Rates (TLPAS and PLAPS, respectively) for Two Recording Sites That Provided Climate Forcing Data for TCSWAT.

Climate Site	Elev. (m)	TMP (°C)	PCP (mm)	TLAPS (°C/km)	PLAPS (mm/km)
Onion Park	2,259	0.50	846	−7.14 (−6.0)	765 (0)
White Sulphur Springs	1,582	5.33	328		

Default lapse rates are shown in parentheses.

specific conifer forest types are limited. To account for the spatial variability in forest canopy and other vegetation characteristics, we used detailed stand-level and remotely sensed data (Redmond *et al.*, 2001) to map and classify the forested and nonforested landcover within TCSWAT. Multiple-attribute criteria organized stands into landcover categories similar to those of the Anderson *et al.* (1976) Level II classification, but with greater forest type diversity. Our landcover map was processed as raster data with 30 m resolution, and included six categories consisting of (1) barren ground, (2) grassland, (3) shrubland, (4) spruce-fir forest, (5) lodgepole pine forest, and (6) disturbed conifer forest. The vegetation map was finalized after field verification in each of the model's subwatersheds.

The characteristics of each landcover category were defined by editing the standard SWAT vegetation database (Table 2), with values derived from onsite measurements, remotely sensed data, and relevant literature (Burns and Honkala, 1990). The standard range-grass and range-brush parameter sets within SWAT were modified to describe the local grassland and shrubland categories. Similarly, the standard evergreen forest parameters were modified to better represent the characteristics of pure or mixed stands of Engleman spruce (*Picea engelmannii*), subalpine fir (*Abies lasiocarpa*), and whitebark pine (*Pinus albicaulis*), which were classified as spruce-fir forest in this watershed. Lodgepole pine (*Pinus contorta*) forest occurs in relatively homogeneous stands and was therefore defined as a distinct category. The barren landcover category, which includes bare ground, rock outcrops, and alpine conditions, was parameterized by modifying the dirt road transportation parameter set in SWAT. The fraction of total impervious area and connected impervious area was 0.98 and 0.95, respectively. The space between the impervious areas was represented by the generic Bermuda grass (*Cynodon dactylon*) parameters. The disturbed conifer forest category, representing decadent, burned, or regenerating forest stands, was based on the Anderson Level II (Anderson *et al.*, 1976) transitional classification. Like the other conifer forest categories, this category was based on a modification of the standard evergreen forest parameter set. The main difference between disturbed and mature forest cover characteristics were reduced canopy density, interception capacity, and ground surface roughness.

Leaf area index (LAI) is an indirect measure of canopy structure (White *et al.*, 1997; Hall *et al.*, 2003) and is used by SWAT as one of the variables that influences ET and canopy interception (Neitsch *et al.*, 2005). Although conifer forests maintain needles yearlong, the leaf area varies seasonally (Oliver and Larson, 1990; Lowman and Nadkarni, 1995; Waring and Running, 1998). Analysis of multi-temporal satellite-derived LAI (Holsinger *et al.*, 2005; USGS, 2005) indicated that winter values in the study area were approximately 40% lower than those of summer. This information was incorporated into the SWAT model by specifying both minimum (ALAI) and maximum (BLAI) values for LAI in the parameter set for each forest cover type. With these two parameters, LAI was allowed to vary seasonally without diminishing to zero during the winter months (Table 2). Canopy interception in SWAT is defined by a maximum canopy storage parameter (CANMX). Values for CANMX were defined for each landcover type by multiplying the mean precipitation event depth (6 mm) by an estimated annual interception fraction, based on field measurements within TCSWAT (Table 2).

Soil Characteristics. The State Soil Geographic (STATSGO) database characterized study area soils. STATSGO consists of a digital general soil association map and attribute databases developed by the National Cooperative Soil Survey (NRCS 1994). The Soil Survey Geographic (SSURGO) database is more comprehensive, but presently not complete for TCSWAT. So STATSGO is the most detailed soil map available. Wang and Melesse (2006) found only moderate and inconsistent improvement in SWAT performance when parameterized using the more detailed SSURGO rather than STATSGO data. STATSGO classifies all soils in TCSWAT as Stemple series,

TABLE 2. Estimated Characteristics of Landcover Within TCSWAT.

Landcover	Max Height	Max LAI	Min LAI	Int.	Base Temp.	OVN	CN2B
Barren	0	0	0	0	0	0.10	96
Grassland	0.75	1.50	0.75	3	10	0.12	70
Shrubland	3.50	2.00	1.00	5	10	0.13	65
Lodgepole	22	2.80	1.80	25	3	0.16	58
Spruce-fir	26	3.0	1.95	28	3	0.17	55
Disturbed	10	2.0	1.00	10	3	0.14	69

In order of appearance they are: maximum canopy height (m), maximum and minimum leaf area index (m^2/m^2), relative proportion of annual precipitation intercepted by the canopy (%), minimum mean daily temperature (°C) required for the onset of plant growth, Manning's coefficient for ground surface roughness(N), and SCS Curve Number for moisture condition II and soil group B.

which have four layers, a very gravelly loam texture, and belong to hydrologic group B (NRCS 1996). This indicates that they are moderately deep, and have average infiltration and water transmission rates.

Definition of Hydrologic Response Units. Unique combinations of soil, landcover, and management practices in SWAT are called hydrologic response units (HRUs), and each is assumed to have uniform hydrologic characteristics (Leavesley *et al.*, 1983; Maidment, 1993; Neitsch *et al.*, 2005). HRUs are the fundamental modeling unit within SWAT, and subwatersheds can be composed of one or several HRUs by specifying relative area thresholds for each defining component (Neitsch *et al.*, 2005). Due to the lack of variability in soil characteristics inferred from STATSGO, subwatersheds and landcover units were the only distinguishing factors for HRU definition in TCSWAT. The number and diversity of HRUs can influence model output (Bingner *et al.*, 1997; Mangurerra and Engel, 1998), and to ensure a high level of resolution, multiple HRUs were defined for each subwatershed. By specifying a minimum 5% of subwatershed area for landcover units, 54 HRUs were created.

Model Calibration

Calibration and validation of TCSWAT were based on a balanced, split-sample approach. Once configured and initially parameterized, SWAT was run on a daily basis from October 1, 1993 to December 31, 2002. The first two years of the simulation (1993 and 1994) allowed the model to equilibrate to ambient conditions. Calibration was based on the years 1997 to 2000, and the model was validated using data from the two years prior (1995, 1996) and two years following (2001, 2002) the calibration period. The period of record encompassed a wide range of climatic conditions including wet, dry and average years (Table 3). Although fairly short, the information richness provided by such variability is more valuable than a lengthy record alone (Gupta *et al.*, 1998).

Model Performance Criteria. Model performance was evaluated through visual interpretation of the simulated hydrographs, and commonly used statistical measures of agreement between measured and simulated streamflow (ASCE, 1993; Coffey *et al.*, 2004; White and Chaubey, 2005). Graphical evaluation focused first on matching the rising and recession limbs of the snowmelt driven runoff hydrograph, then second on maintaining base flow.

The American Society of Civil Engineers (ASCE, 1993) recommends the Nash-Sutcliffe (Nash and Sutcliffe, 1970) model efficiency (NS), and average runoff volume deviation (Dv) metrics for gauging hydrologic model performance. These statistics along with a measure of relative difference (RE) were employed to quantitatively evaluate model performance, and are described below with notation, where y is individual measured values, $\bar{y}$ is mean measured value, and $\hat{y}$ is individual simulated values.

TABLE 3. Climate and Hydrologic Variability Within TCSWAT.

Year	Mean Temp. °C	PCP (mm)	SWE (mm)	Water Yield (mm)	Peak Flow (m^3/s)
v 1995	0.4	978	457	511	7.5
v 1996	−0.1	879	399	495	3.9
c 1997	0.7	856	485	564	4.0
c 1998	1.6	831	318	430	2.4
c 1999	1.4	757	330	336	2.1
c 2000	1.0	714	373	357	2.0
v 2001	2.1	677	285	288	2.0
v 2002	1.2	727	351	339	2.4
Min	−0.1	677	285	288	2.0
Mean	1.0	803	375	415	3.3
Max	2.1	978	485	511	7.5

Presented are the mean annual temperature, precipitation (PCP), peak snow-water equivalent (SWE) measured at Onion Park (SNOTEL:1008), water yield, and peak stream discharge rate, measured at the watershed outlet. Years codes as ‘c’ and ‘v’ were used for calibration and validation, respectively.

NS is calculated as

$$\mathrm{NS} = 1 - \left[\frac{\sum (y-\hat{y})^2}{\sum (y-\bar{y})^2}\right] \tag{1}$$

Values of the NS coefficient can range from negative infinity to 1. NS coefficients greater than 0.75 are considered “good,” whereas values between 0.75 and 0.36 as “satisfactory” (Motovilov *et al.*, 1999; Wang and Melesse, 2006). When NS coefficients are 0 or less, model predictions are no better than the mean of the observed data over the evaluation period.

Dv is calculated as

$$\mathrm{Dv}(\%) = \frac{\sum |y-\hat{y}|}{\sum y}(100) \tag{2}$$

A Dv value of zero indicates a perfect model fit, and larger values describe greater average differences between measured and simulated values. Calibration studies using the Dv metric have reported satisfactory results when Dv is less than 40% (Coffey *et al.*, 2004).

RE is calculated as

$$\mathrm{RE} = \frac{|y-\hat{y}|}{y}(100) \tag{3}$$

Relative error (RE) is the percent difference between measured and simulated values over a specified time period, so smaller values are an indication of better model performance. When RE is <20%, model estimates are satisfactory. In this study, RE was used to quantify differences in total water yield over annual and seasonal periods.

Parameter Selection Through Sensitivity Analysis. Initial model simulations were conducted using default values for most model parameters. Potential evapotranspiration (PET) was modeled with the Penman-Monteith algorithm (Monteith, 1965) because it uses, in part, canopy height to estimate PET, which made it possible to impart different PET values for each landcover type. Surface runoff was modeled with the Curve Number approach (SCS, 1972) and the variable storage channel routing method.

Following initial parameterization, a suite of parameters commonly recommended for streamflow calibration was varied over their potential ranges with automated Latin Hypercube One-factor-At-a-Time (LH-OAT) global sensitivity analysis procedures (van Griensven *et al.*, 2006) (Table 4). Latin Hypercube (LH) sampling is computationally efficient, and One-At-a-Time procedures (OAT) ensure that a change in model output is unambiguously attributed to the change in an input parameter. Rankings associated with the LH-OAT analysis indicated that streamflow predictions were sensitive to variation in snow processes, surface runoff lag, ground water, soil, and Curve Number parameters. Given this, parameters within the above functional groups were selected for model calibration.

Parameter Adjustment for Calibration. The selected parameters were adjusted over a range of values through a stepwise process that utilized both automated methods (van Griensven and Bauwens, 2003), and manual refinement until an acceptable parameter set was obtained (Table 5). Snow parameters were estimated from previous field measurements at TCEF (Sappington, 2005; Woods *et al.*, 2006). Filter techniques provided initial estimates for the base flow recession constant, and runoff/base flow fractions (Arnold *et al.*, 1995; Arnold and Allen, 1999). Model development guidelines (Neitsch *et al.*, 2005) and literature reviews (Arnold *et al.*, 1998; Mangurerra and Engel, 1998; Santhi *et al.*, 2001a,b; Eckhardt and Arnold, 2001; White and Chaubey, 2005) provided guidance for values of the remaining surface runoff lag, soil, and ground-water parameters.

The uncalibrated model was initially tuned by minimizing the differences between measured and

TABLE 4. Parameter Ranges for Sensitivity Analysis of the Uncalibrated TCSWAT Model.

Component	Parameter Name	Description	Range
Basin	ESCO	Soil evaporation compensation factor	0-1
Basin	EPCO	Plant uptake compensation factor	±50%
Basin	SURLAG	Surface lag coefficient	0-4
Basin/snow	SMFMX	Maximum snowmelt rate (mm/C/day)	0-10
Basin/snow	SMFMN	Minimum snowmelt rate (mm/C/day)	0-10
Basin/snow	SFTMP	Snow fall temp (C)	0-5
Basin/snow	SMTMP	Snowmelt temp (C)	0-5
Basin/snow	TIMP	Snowpack temp lag factor	0.01-1
Basin/snow	TLAPS	Temp lapse rate (C/km)	±50%
Subbasin	SLOPE	Average Slope Steepness (m/m)	±50%
Subbasin	SLSUBBSN	Average slope length (m)	±50%
HRU	CN2	Initial SCS curve number II for moisture	±10%
HRU	CANMX	Maximum canopy storage (mm)	0-10
HRU	BLAI	Maximum potential leaf area index (LAI, m2/m2)	±50%
HRU	BIOMIX	Biological mixing efficiency	0-1
Channel	CH_N	Manning's "n" for main channel	±20%
Channel	CH_K2	Hydrologic conductivity for main channel (mm/hr)	0-150
Ground water	ALPHA_BF	Base flow recession constant	0-1
Ground water	GW_DELAY	Groundwater delay time (days)	0-100
Ground water	RCHRG_DP	Deep aquifer percolation fraction (%)	0-1
Ground water	GWQMN	Threshold depth in shallow aquifer for return flow (mm)	0-5000
Ground water	GW_REVAP	Groundwater re-evaporation coefficient	0.02-0.2
Ground water	REVAPMN	Threshold depth in shallow aquifer for re-evaporation (mm)	0-500
Soil	SOL_AWC	Soil available water capacity (mm)	±50%
Soil	SOL_K	Saturated hydraulic conductivity (mm/hr)	±50%
Soil	SOL_Z	Soil depth (mm)	±50%
Soil	SOL_ALB	Moist soil albedo	0-1

TABLE 5. Default Value, Calibration Range, and Final Calibration Estimate of Selected SWAT Model Parameters, by Watershed Model Component.

Component	Parameter Name	Default Value	Calibration Range	Final Estimate
Basin	SURLAG	4	0-4	0.05
Snow	SFTMP	1	0-5	1.0
	SMTMP	0.5	0-5	2.5
	SMFMX	4.5	0-10	3.0
	SMFMN	4.5	0-10	2.9
	TIMP	1	0.01-1	0.06
	SNOCOVMX	1	0-500	200
	SNO50COV	0.5	0-1	0.1
Ground water	ALPHA_BF	0.048	0-1	0.01
	GW_DELAY	31	0-100	1
	RCHRG_DP	0.05	0-0.20	0.15
Soil	SOL_AWC			
	Layer 1	0.09	0.05-0.20	0.18
	Layer 2	0.06	0.02-0.15	0.12
	Layer 3	0.06	0.02-0.15	0.12
	Layer4	0.05	0.02-0.15	0.10
	SOL_K			
	Layer 1	23	10-100	75
	Layer 2	11	5-75	50
	Layer 3	6	5-50	25
	Layer4	3.5	5-25	20
Landcover	CN2B			
	Barren	98	95-100	96
	Grassland	69	68-72	70
	Shrubland	61	63-67	65
	Spruce-fir	55	53-57	55
	Lodgepole pine	55	57-60	58
	Disturbed forest	55	67-74	69

simulated annual precipitation, snowmelt, and water yield. This was accomplished by adjusting the temperature and precipitation lapse rates (TLAPS and PLAPS, respectively) from default to locally derived values. With locally established lapse rates, the amounts of model predicted annual precipitation, snowmelt, and water yield closely matched observed values. Further model refinement then focused on matching the timing of simulated streamflow to measured monthly and daily values by making iterative changes in the selected calibration parameters. After parameter ranges were defined, calibrated values were estimated with automated methods based on the Shuffled Complex Evolution algorithm (Duan *et al.*, 1992, 1994; van Griensven and Bauwens, 2003). With only a single residual sum of squares objective function, this embedded algorithm strongly weighted the large values of snowmelt driven hydrograph peaks, giving small base flow values little influence on the calibration. This resulted in a failure to match low flow periods, and final parameter estimates were reached by manually refining parameter estimates based on the automated calibration procedures.

Model Validation

The reliability of hydrologic predictions generated by the calibrated model was assessed using two methods. First, using a traditional yet subjective approach, the RE, Dv, and NS performance statistics calculated from model output in the validation period (1995, 1996, 2001, and 2002) were compared to those generated during the calibration phase (1997-2000) for annual, seasonal, monthly, and daily time steps. Second, we used an objective, regression-based model invalidation procedure to compare the estimated and measured streamflow values over seasonal and monthly time steps of the validation period. With this approach we simultaneously tested the joint null hypothesis: H_0: $\beta_0 = 0$ and $\beta_1 = 1$ at $\alpha = 0.05$. When the modeled and measured values are closely matched, the y-intercept (β_0) should be close to zero and the slope (β_1), close to 1. Taken from Draper and Smith (1981), the test statistic, **Q**, is computed as

$$\mathbf{Q} = (\beta - b)' X' X(\beta - b) \sim pS^2 F_{p,v,1-\alpha}, \tag{4}$$

where β is hypothesized values for y intercept and slope, i.e., 0 and 1; b is vector of actual regression coefficients; $X'X$ is matrix term in independent variables (predicted y's); S^2 is residual mean square; p is number of regression coefficients − 1; $v = n - p$ is residual degrees of freedom (DF); and α is significance level (0.05).

This regression-based invalidation method could only be applied to monthly data for two reasons. First, as only four data points were available, annual data had insufficient DF. Second, daily streamflow data exhibited serial autocorrelation, as well as skewness and kurtosis that precluded assumptions of independence and normality (Coffey *et al.*, 2004).

Seasonal Streamflow Separation

While satisfactory overall performance could be achieved by the calibrated model, substantial differences in model performance existed between the runoff and base flow periods. Model performance was therefore evaluated when the runoff and base flow periods were considered in isolation. For separation, the months of September through April were categorized as base flow months, while May, June, and July were considered runoff months. Daily separation was based on interpretation of the flow duration curve for actual streamflow from 1993 to 2002, following a method similar to that of Hay *et al.* (2006). Daily mean discharge of < 0.2 m^3/s was classified as base flow, and values ≥0.2 m^3/s were assigned to runoff. The same quantitative measures used to evaluate overall streamflow predictions were also used in this analysis of individual hydrograph components.

Model Performance Decomposition

The relative influence of individual and grouped parameters associated with the final calibrated model was assessed through a stepwise decomposition. Following calibration, partial model sensitivities were quantified by systematically resetting parameter groups, and individual parameters within them, to their default values, and running the model with the other parameter groups in their calibrated state. Output from each partially calibrated model run was then compared to the fully calibrated model, and performance differences were noted as relative changes in the NS statistic. Assessment focused on the snow, surface runoff lag, ground water, soil, and SCS Curve Number parameter groups.

RESULTS

Model Calibration

A visual comparison of the hydrographs obtained using the default and calibrated model parameters in 1999, one of the four years used for model calibration, indicates that calibration substantially improved the fit of modeled *vs*. observed daily streamflow (Figure 2). The 1999 hydrograph produced by the default parameterization was extremely flashy, with several short but large runoff events between Julian days 50 and 140 that were presumably caused by snowmelt. These events were followed by a more prolonged runoff period, the timing of which coincided with the observed runoff. However, the predicted peak flow was 75% less than the observed value. The default model also underestimated base flows, predicting zero flow for periods of up to 54 days. In contrast, the calibrated 1999 model hydrograph closely matched the observed data throughout most of the year. The best fit was observed in the spring runoff period, particularly on the rising limb of the hydrograph. Summer recession values were overestimated, while winter and early spring base flow components were underestimated. Similar patterns between default and calibrated hydrographs were also evident across the remaining calibration period (Figure 3).

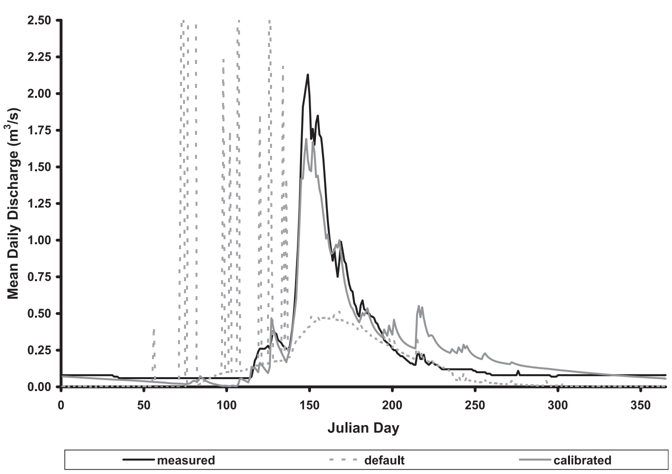

FIGURE 2. Measured, Default, and Calibrated Daily Streamflow Hydrographs of 1999.

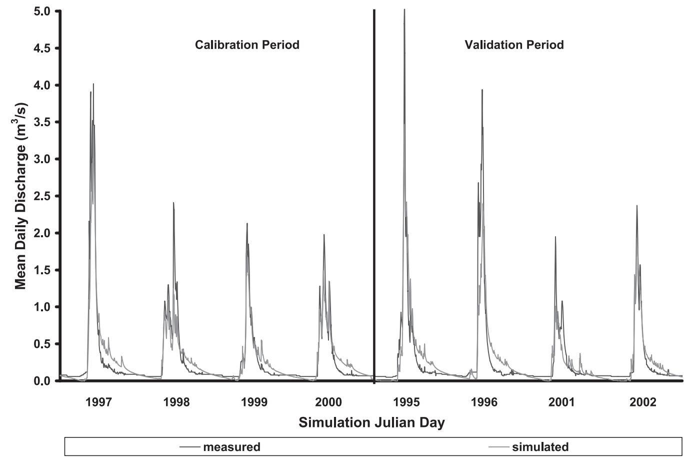

FIGURE 3. Measured and Simulated Daily Streamflow in During Calibration and Validation.

Annual Water Yield. Over the four year calibration period (1997-2000), the default model consistently overestimated the annual water yield, but the overall default difference was within 5% of the recorded value (Table 6). Calibration brought the relative error of estimate (RE) for the same period down to 2% but reversed its direction so that the model predicted slightly less than the measured water yield.

Monthly Water Yield. Calibration also improved model performance at the monthly time step (Table 6). The average difference between monthly estimates (Dv_m) produced by the default model was approximately 80%. After calibration, average monthly deviations were reduced to ~30%. The monthly Nash-Sutcliffe (NS_m) model efficiency coefficient increased from 0.23 for the default model to 0.90 after calibration (Figure 4).

TABLE 6. Model Performance Statistics for Streamflow Predictions in TCSWAT.

Simulation Type	Year	Obs (mm)	Sim (mm)	RE (%)	Dv_m	Dv_d	NS_m	NS_d
Default	1997	564	609	8	71	123	0.53	−1.24
	1998	430	408	−5	73	91	0.35	−1.45
	1999	336	347	3	92	114	−0.48	−3.39
	2000	357	411	15	99	109	−0.47	−3.18
Calibration	1997	564	563	<1	34	37	0.90	0.88
	1998	430	375	−13	31	36	0.82	0.75
	1999	336	337	<1	27	30	0.92	0.92
	2000	357	374	5	30	37	0.92	0.86
Validation	1995	511	517	1	45	32	0.95	0.78
	1996	495	460	−7	50	39	0.83	0.74
	2001	288	234	−19	42	31	0.83	0.70
	2002	339	354	4	31	22	0.97	0.94
Overall Default		1,688	1,775	5	82	110	0.23	−1.76
Overall Calibration		1,688	1,649	−2	31	36	0.90	0.86
Overall Validation		1,632	1,565	−4	43	32	0.90	0.76

Shown are observed (Obs) and simulated (Sim) water yields, relative error (RE), mean pair-wise deviation (Dv), Nash-Sutcliffe model efficiency (NS) over monthly (m) and daily (d) time-steps for default and calibrated models in calibration period 1997-2000, and for the calibrated model during the validation period (1995-996, 2001-2002).

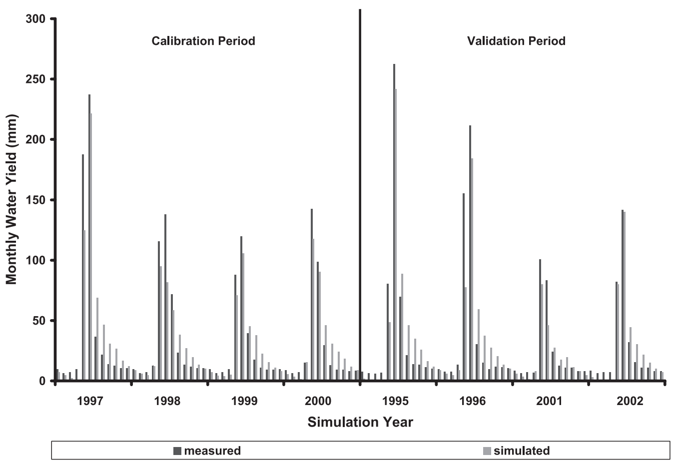

FIGURE 4. Measured and Simulated Monthly Streamflow Over the 1997-2000 Calibration and 1995, 1996, 2001, and 2002 Validation Time Periods.

Daily Streamflow. Calibration had its greatest impact on performance at the daily time step (Table 6). The average percent deviation in daily estimates (Dv_d) produced by the default model was 110%. After calibration, average daily deviations were reduced to 36%. With default parameters the overall Nash-Sutcliffe coefficient for simulation at the daily time step (NS_d) was −1.76, whereas the calibrated model yielded an NS_d coefficient of 0.86 for the same time period.

Model Validation

Model performance in the four year validation period was similar to that achieved during calibration at annual, monthly, and daily time steps (Table 6).

Annual Water Yield. The RE for annual water yields in the four year validation period was 4%, a reduction of 2% compared to the calibration period but still a 1% increase in accuracy over predictions from the default model. In a year-by-year comparison, the greatest errors during calibration and validation occurred in 1998 (−13%), and 2001 (−19%), respectively, but because the relative errors for annual estimates were all less than 20%, annual water yield predictions were validated.

Monthly Water Yield. At the monthly time step, the calibrated model performed well when applied to the validation period. The average percent deviation statistic (Dv_m) changed 12%, increasing from 31% for the calibration period to 43% for the validation period. Deviation between measured and simulated monthly water volumes was lowest during the spring runoff period in May and June, and greatest during recession flows of August and September. In a cyclical manner, the model consistently underestimated water yield from January through June, yet overestimated the water yield from July through November (Figure 5). In spite of these discrepancies, monthly NS_m values ranged from 0.83 to 0.97, and at 0.90 the overall validation efficiency was identical to that for the calibration period, and the model was therefore validated.

Daily Streamflow. Model performance measures were lower for daily flows in the validation compared to the calibration period, but still indicated acceptable results (Table 6). The mean relative deviation between simulated and observed daily streamflows (Dv_d) remained relatively unchanged, from 36% in the calibration to 32% during the validation period.

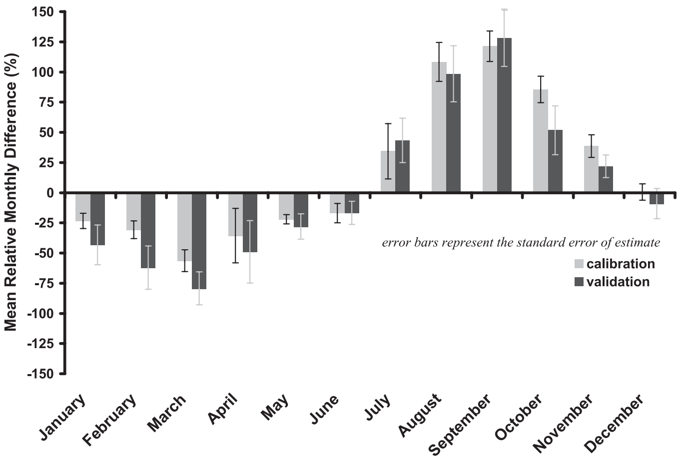

FIGURE 5. Relative Mean Monthly Deviation From Measured Streamflow (%) and Their Associated Standard Errors for the Calibration and Validation Simulation Periods.

TABLE 7. Model Performance Statistics for Base Flow and Runoff Hydrograph Components Predicted by TCSWAT.

		Base Flow Period		Runoff Period	
Monthly simulation	Year	Dv_m	NS_m	Dv_m	NS_m
Calibration	1997-2000	62	−4.73	24	0.80
Validation	1995-96, 2001-02	66	−6.90	24	0.83
		Base flow (<0.2 m^3/s)		**Runoff (≥ 0.2 m^3/s)**	
Daily simulation	Year	Dv_d	NS_d	Dv_d	NS_d
Calibration	1997	80	−7.15	29	0.81
	1998	51	−4.86	32	0.42
	1999	59	−5.59	19	0.89
	2000	86	−11.99	23	0.69
Validation	1995	93	−9.09	34	0.70
	1996	70	−14.65	44	0.48
	2001	52	−4.36	38	0.16
	2002	73	−5.17	15	0.89
Overall calibration		69	−6.93	27	0.78
Overall validation		72	−8.58	34	0.63

At the monthly time step, base flow was defined as January-April, and August-December, while the runoff component was defined by the months of May, June, and July. Daily base flow was separated from runoff by a 0.2 m^3/s threshold, where runoff was captured by values equal to or greater than the threshold.

The efficiency of daily streamflow predictions (NS_d) dropped from 0.86 in the calibration period to 0.76 during validation. The closest match between simulated and observed daily flows occurred during the runoff period between May and early July (Figure 3). Later in the year, the model underestimated rates of decline in falling limbs of snowmelt-induced hydrograph peaks, so that predicted flows were consistently higher than the observed values during recession periods. Conversely, the calibrated model consistently underpredicted base flow in winter and early spring.

Seasonal Streamflow Separation

Model performance statistics for the runoff period indicated a generally good fit between observed and predicted flows (Figure 3). The mean differences between monthly measured and simulated runoff estimates (Dv_m) in calibration and validation runs were both less than 25% at the monthly time step. Monthly calibration and validation efficiency scores (NS_m) were also equal to or greater than 0.80. Relative errors (Dv_d) and model efficiencies (NS_d) for daily flows were slightly lower than what was achieved for the monthly values, but also indicated satisfactory model performance (Table 7).

In contrast with the runoff period, base flows were not well represented by SWAT. Base flows on the lower part of the hydrograph recession limb were overestimated, while the winter and early spring base flows were underestimated. Mean monthly differences between base flow estimates were more than twice as large as those generated during runoff, with relative errors that were nearly 65% for both the calibration and validation periods (Table 7). Results were similar for the daily flow simulations, with relative errors of approximately 70% for both calibration and validation periods. Negative NS_d efficiency scores for the calibration and validation periods indicate that model performance was no better than an average of the observed values. Given these performance measures, SWAT seasonal predictions of base flow could not be validated over either monthly or daily time steps.

Regression-Based Model Invalidation

Results of the regression-based model invalidation procedure were generally consistent with those obtained using traditional performance criteria. In three out of the four validation years the p-value was greater than 0.05, indicating a failure to reject the joint null hypotheses of a zero y-intercept and slope equal to 1 (Table 8). In other words, model predictions closely matched the observed values. The results were even better when only flows during the runoff period were considered, with all p-values > 0.05. Consistent with the traditional evaluation techniques, validation of the base flow component was unsuccessful. All p-values were less than 0.05 indicating a rejection of the joint null hypotheses of a zero y-intercept and slope equal to 1 (Table 8).

TABLE 8. Results of Regression-Based Model Invalidation Procedure for TCSWAT.

	All Flows			Base Flows			Runoff		
Months of Year	*F*-Stat.	DF	*p*-Value	*F*-Stat.	DF	*p*-Value	*F*-Stat.	DF	*p*-Value
1995	0.56	10	0.47	1,015.97	7	0.00	0.29	1	0.69
1996	2.24	10	0.17	214.09	7	0.00	0.37	1	0.65
2001	9.26	10	0.01	103.94	7	0.00	1.77	1	0.41
2002	0.44	10	0.84	574.13	7	0.00	2.02	1	0.39
All Months	3.88	46	0.06	1,413.86	34	0.00	2.41	10	0.15

Base flow months are January, February, March, April, August, September, October, November, December. Runoff months are May, June, and July. Validation is assumed when $p > 0.05$.

Model Performance Decomposition

Calibration of the snow parameter set had the greatest effect on model performance followed in decreasing order of influence by the surface runoff lag factor (SURLAG), ground water, soil, and Curve Number parameter sets (Table 9).

Snow Parameters. Setting the snow parameters to their default values reduced the NS efficiency from 0.92 to −0.06. With default snow values the snowmelt driven runoff peak occurred 75-80 days earlier than the calibrated and observed peaks, and the recession limb of the hydrograph was extended by a similar number of days (Figure 6a). The parameter with the greatest impact on model calibration was the snow pack temperature lag factor (TIMP), followed by the snowmelt temperature (SMTMP), and the snow cover depletion curve (SNCOV50). Use of default values for the maximum and minimum snowmelt rate factors (SMFMX and SMFMN) had only a minimal effect on model performance, while setting the snow covered area parameter (SNOCOVMX) back to the default value of 1 improved model efficiency by 1%.

TABLE 9. Relative Influence of Factors Affecting TCSWAT Calibration.

	Default Value	Calibrated Value	NS_d	NS_d Change (%)
Calibrated model			**0.92**	
Composite snow			**−0.06**	**106**
TIMP	1.0	0.06	0.52	43
SMTMP	0.5	1	0.71	23
SNOCOV50	0.5	0.1	0.73	20
SMFMX	4.5	3	0.89	3
SNOCOVMX	1.0	200	0.93	−2
SMFMN	4.5	2.9	0.91	1
Composite SURLAG	4.0	0.05	**0.19**	**79**
Composite GW			0.80	**13**
ALPHA_BF	0.05	0.01	0.81	12
GW_DELAY	31	1	0.85	7
Composite soil			**0.88**	**4**
SOL_K	23	75	0.88	4
SOL_AWC	0.09	0.18	0.91	1
Composite CN2			**0.89**	**3**
Lodgepole pine	55	58	0.90	2
Disturbed forest	55	69	0.91	1
Shrubland	61	65	0.92	0
Grassland	69	70	0.92	0
Spruce-fir	55	55	0.92	0

Nash-Sutcliffe (NS_d) efficiencies for daily flows with default and calibrated snow, surface lag, ground water, soil, and curve number parameters. Only values for the primary layer are given for soil parameter sets. Values given in bold indicate NS scores when the otherwise calibrated model was run with the given parameter(s) reset to their default values and the subsequent relative change in NS compared with the fully calibrated model.

Surface Runoff Lag Factor. Setting the surface runoff lag factor (SURLAG) from the calibrated value of 0.05 to the default value of 4.0 reduced the NS_d from 0.92 to 0.19. Comparison of the hydrographs obtained using the default and calibrated values for SURLAG shows that the primary effect of surface runoff lag factor calibration was to smooth the hydrograph by slowing the rate of runoff during the snowmelt period (Figure 6b).

Ground-Water Parameters. Setting the groundwater parameters to their default values had only a moderate effect on model performance. With the default parameters, model efficiency was only 12% lower than with the calibrated parameter set (Table 9). Calibration of the ground-water parameter set, however, improved model fit during the streamflow recession period and made more water available for base flow (Figure 6c). Of the calibrated ground-water parameters, adjustment of the base flow recession constant (ALPHA_BF) yielded the greatest improvement in model performance. Reducing ALPHA_BF from the default value of 0.048-0.01 slowed the shallow aquifer response to recharge, causing a reduction in the annual runoff peak during snowmelt but making more water available for streamflow later in the year. Reducing the value of the ground-water delay parameter (GW_DELAY) from the default of

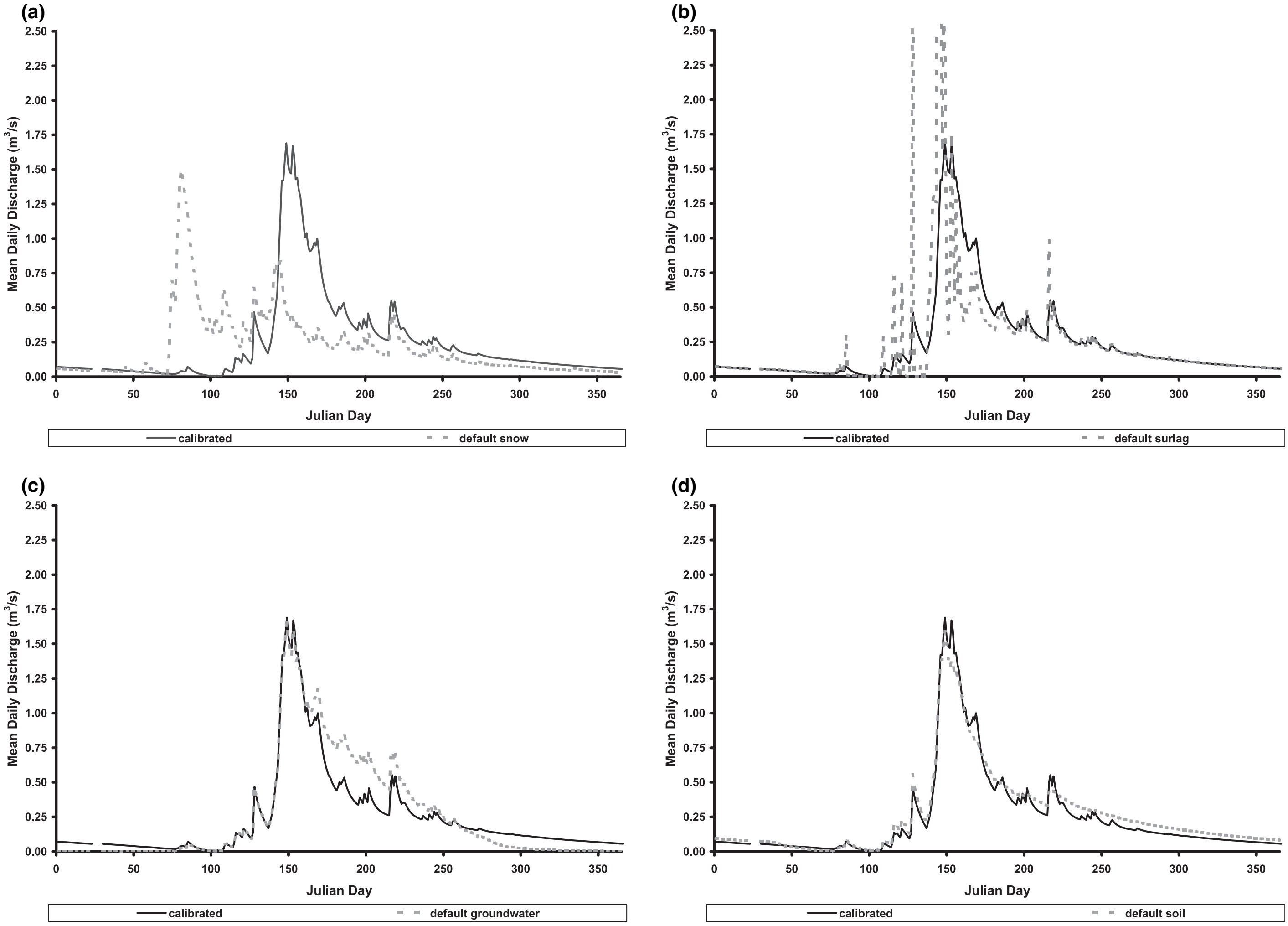

FIGURE 6. (a) Hydrographs Obtained Using Calibrated Model, and Otherwise Calibrated Model With Default Snow Parameters in 1999. (b) Hydrographs obtained using calibrated model, and otherwise calibrated model with default value for SURLAG in 1999. (c) Hydrographs obtained using calibrated model, and otherwise calibrated model with default ground-water parameters in 1999. (d) Hydrographs obtained using calibrated model, and otherwise calibrated model with default soil parameters in 1999.

31 to 1 day affected both the width of the peak discharge and the quantity of water available for base flow.

Soil Parameters. Use of the default soil parameters reduced the overall model efficiency by just 4%. Calibration of the soil parameters primarily improved the fit to the observed daily streamflows on the recession limb of the hydrograph. Of the two soil parameters adjusted, the soil hydraulic conductivity (SOL_K) had the greatest influence on model fit. Increasing SOL_K from the default value of 23 mm/h to the calibrated value of 75 mm/h increased the modeled peak flows during the snowmelt season. Increasing the available water holding capacity parameter (SOL_AWC) made more water available for streamflow in the base flow period, but the improvement in model efficiency was less than 1% (Figure 6d).

SCS Curve Numbers. When Curve Numbers for all represented landcover types were set to their default values, the decrease in model efficiency was just 3%. The only detectable effect on model fit was a reduction in the early runoff peaks. Lodgepole pine forest covers the majority of the watershed, and changing the default Curve Number for that landcover from 55 for standard evergreen forest to 58 yielded the greatest increase in model performance. Following lodgepole pine, introduction of the disturbed conifer forest landcover produced the second largest improvement in efficiency. While this landcover type occupies only 10% of the watershed, the change in Curve Number from 55 to 69 had a noticeable effect on model fit.

DISCUSSION

The goal of this study was to evaluate the performance of SWAT in previously untested conditions to determine if it would be suitable for streamflow prediction in a forested snow-dominated mountain environment at a scale commonly used in natural resource planning and management: the watershed scale.

Few examples of calibration in upland watersheds exist because reliable climate and streamflow data are rare. First, not many mountain watersheds are gauged for streamflow over meaningful time periods, and furthermore, rugged terrain makes it difficult to accurately extrapolate climate data from distant locations.

The research watershed within Tenderfoot Creek Experimental Forest is instrumented with climate and hydrologic monitoring sites that provided an unusually comprehensive dataset to calibrate SWAT. Climate and ecological characteristics of our site are similar to those found throughout the eastern slope of the Rocky Mountains, and should make results from this study applicable to comparable watersheds.

Model Calibration

While simulations based on default parameters produced acceptable annual water yield estimates, the timing of hydrologic processes was mismatched at monthly and daily intervals. Through a process that combined automated and manual calibration of snowmelt, surface runoff lag, ground water, soil, and Curve Number parameter groups, model performance at the monthly, and most noticeably the daily, time steps was improved substantially.

Automated calibration procedures offer many advantages over manual methods. However, it appears that an auto-calibration algorithm with only a single sum of squares objective function cannot be used effectively in watersheds with snowmelt dominated hydrology like TCSWAT that have a strongly seasonal, unimodal hydrograph. With such a format, very little weight is given to model performance during low flow periods because the errors associated with those values are an order of magnitude smaller than those of the runoff period, and this leads to an inability to simultaneously calibrate the model for both high and low flows. Base flows, nonetheless, can be just as important as snowmelt driven peak flows when considering the effects of management on fish populations and riparian condition in mountainous watersheds (Tennant, 1976; Binns and Eiserman, 1979). Instead, an algorithm with multiple, rather than a single objective function(s), may yield more favorable results. One function could be used to calibrate the rising limb and peak of the hydrograph, another for the recession limb focusing on the inflection between runoff and ground-water dominated flow regimes, and a third to account for the base flow period.

Without that capacity to match both high and low flow values simultaneously with automated procedures, the final parameter set was obtained by manually refining parameter estimates produced by the auto-calibration algorithm. After calibration, model performance was very good, with Nash-Sutcliffe efficiencies of 0.90 and 0.86 for monthly and daily streamflow predictions.

Model Validation

We validated model performance with two distinct techniques. Using the traditional validation approach, model performance measures derived over annual, monthly, daily, and seasonal calibration periods were compared to those generated when SWAT was applied to an independent dataset. This traditional assessment is subjective because the final conclusion is in the eyes of the beholder (i.e., there is no test to confirm whether the lack of agreement between performance measures is significant or not). To provide an alternative that is objective, a regression-based invalidation procedure was applied. Satisfying both the regression assumptions and sample size requirements, analysis of monthly data provided the most reliable measure of model performance, which was evaluated for the whole validation period and also over the two hydrograph components.

Conclusions derived from the objective test were similar to those obtained from the traditional assessment. Results indicated no significant differences between simulated and observed values over whole simulation periods, and model output could therefore be validated at the monthly time step. Objective evaluation of separate hydrograph components with monthly data also supported traditional comparisons, where the model preformed well during runoff periods, but failed to represent base flow periods adequately.

In the final analysis, both the traditional and objective validation procedures indicated that SWAT produced favorable results over whole year, monthly, and daily time steps. However, when performance was evaluated seasonally, both evaluation methods also pointed out that SWAT produced better runoff than base flow estimates in this watershed. The overall validity of the model should therefore be judged against its intended use. For example, estimates produced from the current form of SWAT may not be appropriate for evaluating the impacts of forest management on base flows. But, as most concerns associated with forestry in the Rocky Mountains center around the effect of management on the snowmelt runoff period (Troendle, 1983; Stednick, 1996), it may be valid to use SWAT to evaluate these effects because it has been shown to produce reasonable runoff rates during this period.

Model Performance Decomposition

The snowmelt-induced peak is the most prominent feature of the annual streamflow hydrograph of the Tenderfoot Creek research watershed, and matching the large peak by adjustment of snowmelt parameters increased the calibration efficiency of TCSWAT more than any other parameter group. Once snow process and surface runoff lag parameters were calibrated, we achieved acceptable model performance statistics even when base flow estimates were poor because the large annual snowmelt peak was well matched.

Nash-Sutcliffe (NS) efficiency statistics reported for the Tenderfoot Creek site for both the calibration and validation periods compare favorably with others reported in the literature (Table 10). This confirms that SWAT can predict annual, monthly, and daily hydrologic processes in forested snow-dominated mountain watersheds with efficiency levels that are similar to those obtained in other regions where it has been successfully applied. Despite generating favorable statistics, two aspects of SWAT performance in the research watershed warrant further examination. These areas of interest are the representation of snowmelt infiltration and runoff processes and the difficulty associated with matching the recession curve and subsequently poor representation of base flow periods.

Snowmelt Infiltration and Runoff Processes

Of all the snowmelt parameters, adjusting the snowpack temperature lag factor parameter (TIMP), and the threshold temperature for the onset of snowmelt (SMTMP) yielded the greatest improvement in model performance during calibration (Table 9). Snowpack of TCSWAT tends to be light and dry, and to reflect these characteristics TIMP was set to a low value nearly one-tenth that of the default, while SMTMP was increased 2.5 times over the default to account for large diurnal temperature ranges during the snowmelt season. Both of these

TABLE 10. TCSWAT Calibration and Validation Period Statistics Compared to Published Studies.

Reference	**Location**	**Area (sq. km)**	**Calibration NS_m**	**Calibration NS_d**	**Validation NS_m**	**Validation NS_d**
Cao *et al.* (2006)	New Zealand	2,075		0.78		0.72
Eckhardt and Arnold (2001)	Germany	81		0.70		0.73
Eckhardt *et al.* (2002)	Germany	81		0.76		0.81
Muleta and Nicklow (2005)	IL	133		0.74		
Arabi *et al.* (2006)	IA	6	0.84		0.73	
		7	0.73		0.63	
Fontaine *et al.* (2002)	WY	4,999	0.86			
Santhi *et al.* (2001a)	TX	926	0.79		0.87	
		2,997	0.83		0.62	
Srinivasan *et al.* (1998)	TX	238	0.84		0.82	
White and Chaubey (2005)	AR	362	0.89		0.85	
		684	0.89		0.72	
		1,020	0.86		0.87	
Jha *et al.* (2006)	IL	447,500	0.69	0.58	0.81	0.65
Van Liew and Garbrecht (2003)	OK	160	0.75	0.60	0.83	0.71
Wang and Melesse (2005)	MN	2,419	0.86	0.64	0.90	0.62
		4,040	0.86	0.67	0.83	0.50
Wang and Melesse (2006)	MN	515	0.88	0.51	0.50	0.31
		515	0.92	0.49	0.49	0.26
Wu and Xu (2006)	LA	3,435	0.93	0.92	0.85	0.78
		662	0.94	0.90	0.81	0.71
		1,682	0.96	0.83	0.87	0.69
NS_m = monthly	min	6	0.69	0.49	0.49	0.26
NS_d = daily	**TCSWAT**	**23**	**0.90**	**0.86**	**0.90**	**0.76**
	max	447,500	0.96	0.92	0.90	0.81

Bold type indicates values derived from TCSWAT.

changes had the effect of decreasing or eliminating snowmelt during isolated warm days in the early spring, and is consistent with the observed behavior of the snowpack at TCEF, which is initially dry and does not become ripe and produce melt until mid to late May each year (Sappington, 2005). In the translation to streamflow, adjustment of the snowmelt parameters caused delayed snowmelt-generated runoff and shifted the annual peak from winter to spring (Figure 6a).

Although calibration of snow parameters substantially improved the timing of snowmelt conversion to streamflow, the decomposition analysis showed the calibrated hydrograph would still have been extremely flashy without adjustment of the surface runoff lag parameter, SURLAG. This parameter dampens streamflow response by delaying the delivery of surface runoff produced each day to the stream. The fact that SURLAG played such an important role in the model calibration, suggests that simulated streamflow was routed to the stream primarily as surface runoff. This indication is inconsistent with field observations at TCEF and in other forested environments in western North America. In these areas, little or no overland flow is generated even during the period of maximum snowmelt because infiltration rates are extremely high (Stadler *et al.*, 1996). As such, snowmelt is routed to the stream by a combination of lateral flow and shallow ground-water flow. The divergence between modeled and observed hydrologic processes is likely because SWAT assumes an infiltration rate of zero in frozen soils (Neitsch *et al.*, 2005). Soil temperatures in SWAT are calculated as a lagged function of the air temperature, adjusted for the effects of surface litter cover and snow cover (Neitsch *et al.*, 2005). During the calibration and validation periods, SWAT-modeled near-surface soil temperatures were below freezing during 95% of the snowmelt period (Figure 7). Much of that snowmelt was therefore routed to the stream as overland flow, and SURLAG was set to a value well below that used in comparable studies (Eckhardt and Arnold, 2001). Thus, simulated surface runoff rates were similar to the slower, shallow subsurface flows that predominate in this environment. The absence of infiltration into frozen soils also explains why, in contrast with many other studies, the calibrated SWAT model was relatively insensitive to the Curve Number parameter set. In situations where the infiltration rate is zero, SWAT ignores the Curve Number parameter defined for the HRU.

Data from the Onion Park SNOTEL site indicate that mean daily winter temperatures at TCEF are well below freezing for extended periods, so we assume that soils within the study area are frozen for much of the winter. The model assumption that there

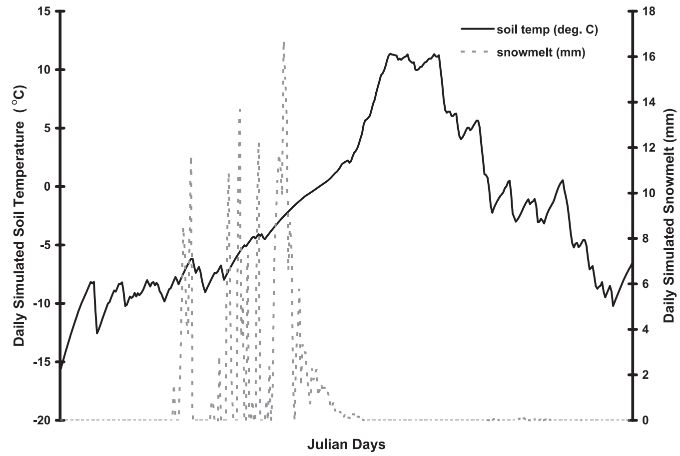

FIGURE 7. The Relationship Between Daily Soil Temperature and Snowmelt in a Lodgepole Pine Forest HRU, Within a Subwatershed of TCSWAT, Over the Year 1999.

is no snowmelt infiltration into these frozen soils may not apply in this environment for two reasons. First, in the relatively dry summer environment of central Montana, soils are usually well drained immediately prior to the first snowfall. This limits the extent to which frozen interstitial water could constrict the matrix flow of infiltration during snowmelt (Gray *et al.*, 2001). Second, like most forest soils, those at TCEF contain an abundance of macropores that typically facilitate infiltration rates in frozen soils comparable to those in unfrozen soils of the same type, regardless of the extent to which the soil matrix is clogged by ice (Popov, 1972; Roberge and Plamondon, 1987; Stadler *et al.*, 1996). Modification of SWAT to allow infiltration into frozen forest soils, using the approaches outlined by Gray *et al.* (2001) could lead to improved model performance and more realistic representation of runoff processes in forested upland watersheds.

Recession Curve and Base flow Periods

The shape of the recession limb of the hydrograph peak was strongly influenced by adjustment of the ALPHA_BF and GW_DELAY parameters, indicating that ground-water discharge was the primary runoff flow path. This is consistent with studies performed in other snowmelt dominated watersheds (Flerchinger *et al.*, 1992, 1994; Ohte *et al.*, 2004) and with the results of ongoing research within TCSWAT (B.L. McGlynn, unpublished data), all of which have shown that ground-water discharge is the primary mechanism for the transfer of snowmelt runoff to mountain streams during the recession limb of the hydrograph. With the current model structure, however considerable difficulty was associated with matching the rapid transition between shallow subsurface flow during and shortly after the snowmelt period and the base flow regime of the remaining portions of the year dominated by ground-water flow.

SWAT consistently underestimated monthly water yields from January through June and overestimated the monthly water yields from July through November, and the same trends were observed in the daily flow simulations. Both of these trends may be explained in terms of the effect ground-water parameters have on the timing and delivery of ground water to the stream. In general, the model apportioned little water to base flow during the snowmelt runoff period, partly because of the frozen soil effect previously described. The model also routed water out of the ground-water reservoir faster than occurs under field conditions. Lowering the value of the base flow runoff coefficient (ALPHA_BF) dampened the model response to snowmelt-induced recharge, increasing the amount of runoff available for base flow later in the year. However, lowering ALPHA_BF also decreased the snowmelt runoff peak by reducing the ground-water contribution, and increased the runoff during the snowmelt recession period when runoff is rapidly changing from a shallow subsurface flow dominated regime to one dominated by ground-water flow. Reducing the value of GW_DELAY partially offset the effect of limited base flow on the runoff peak, but improvements in model performance during the base flow period were largely achieved at the expense of reduced model performance during the runoff period and vice-versa. It may be possible to obtain a better fit between observed and simulated flows in the recession period of the snowmelt hydrograph and during base flow by incorporating two base flow recession constants into the SWAT ground-water parameter set – one for the recession limb of snowmelt driven runoff and the other for the maintenance of base flow.

CONCLUSIONS

Although annual water yield predictions based on default SWAT parameterization were acceptable, calibration increased model performance in all time steps over which streamflow predictions were evaluated, especially at the daily level of prediction. Calibration was most strongly influenced by snow process, and surface runoff lag parameters in the Tenderfoot Creek watershed, because they were most responsible for matching the large annual hydrograph peak and smoothing the streamflow response to snowmelt. Once runoff peaks were well represented, reasonable performance statistics could be generated regardless of base flow accuracy. Evaluation of model performance against a validation dataset provides an independent check on the robustness of the calibrated parameter set. With an objective regression-based hypothesis testing procedure model performance can be evaluated objectively, and add rigor to validation statements.

A relatively small drop in performance from the calibration period to independent validation period also suggests that the calibrated parameter set is robust across the 1995-2002 period. Future studies may be able to use SWAT with these calibrated parameters to investigate various biophysical interactions at the watershed scale. Note, however, that relationships between frozen soil and snowmelt infiltration, along with ground-water recession and base flow maintenance, still require refinement.

Recommendations

While the existing model produces acceptable results, the assumption that no infiltration exists in frozen soils means that the model does not properly represent the snowmelt infiltration and runoff process. In addition, the model does not perform well during base flow periods. Refinement of the model components representing snowmelt infiltration and runoff and ground-water discharge during base flow would further improve model performance and the physical representation of hydrologic processes within this and similar Rocky Mountain watersheds. Model calibration using the automated procedures in SWAT could also be improved by incorporating multiple objective functions to simultaneously calibrate the three major components of the annual streamflow hydrograph: the runoff, recession, and base flow periods.

ACKNOWLEDGMENTS

Financial support for this research was provided by the USDA, Forest Service, Rocky Mountain Research Station, Forestry Sciences Laboratory, and by the University of Montana, College of Forestry and Conservation. Three anonymous reviewers provided comments that improved the paper.

LITERATURE CITED

Alt, D. and D.W. Hyndman. 1986. Roadside Geology of Montana. Mountain Press, Missoula, Montana.

Anderson, J.R., E.E. Hardy, J.T. Roach, and R.E. Witmer. 1976. A Land Use and Land Cover Classification System for use With

Remote Sensor Data. Geological Survey Professional Paper 964. U.S. Government Printing Office, Washington, D.C.

Arabi, M., R.S. Govindaraju, M. Hantush, and B.A. Engel, 2006. Role of Watershed Subdivision on Modeling the Effectiveness of Best Management Practices With SWAT. Journal of the American Water Resources Association 42(2):513-528.

Arno, S.F. and C.E. Fiedler. 2005. Mimicking Nature's Fire: Restoring Fire-Prone Forests in the West. Island Press, Washington, D.C.

Arnold, J.G. and P.M. Allen, 1999. Automated Methods for Estimating Baseflow and Groundwater Recharge From Streamflow Records. Journal of the American Water Resources Association 35(2):411-424.

Arnold, J.G., P.M. Allen, and G. Bernhardt, 1993. A Comprehensive Surface-Groundwater Flow Model. Journal of Hydrology 142(1):47-69.

Arnold, J.G., P.M. Allen, R. Muttia, and G. Bernhardt, 1995. Automated Base Flow Separation and Recession Analysis Techniques. Ground Water 33(6):1010-1018.

Arnold, J.G. and N. Fohrer, 2005. SWAT2000: Current Capabilities and Research Opportunities in Applied Watershed Modeling. Hydrological Processes 19(3):563-572.

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

ASCE, 1993. Criteria for Evaluation of Watershed Models. Journal of Irrigation and Drainage Engineering 119(3):429-442.

Barrett, S.W., 1993. Fire History of Tenderfoot Creek Experimental Forest, Lewis and Clark National Forest. Final Report RJVA No. INT-92679 for U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Forestry Sciences Laboratory, Bozeman, Montana..

Bingner, R.L., J. Garbrecht, J.G. Arnold, and R. Srinivasan, 1997. Effect of Watershed Subdivision on Simulation Runoff and Fine Sediment Yield. Transactions of the ASAE 40(5):1329-1335.

Binns, N.A. and F.A. Eiserman, 1979. Quantification of Fluvial Trout Habitat in Wyoming. Transactions of the American Fisheries Society 108:215-228.

Bosch, J.M. and J.D. Hewlett, 1982. A Review of Catchment Experiments to Determine the Effect of Vegetation Changes on Water Yield and Evapotranspiration. Journal of Hydrology 55:3-23.

Burns, R.M. and B.H. Honkala, tech. coords. 1990. Silvics of North America: 1. Conifers. Agriculture Handbook 654. U. S. Department of Agriculture, Forest Service, Washington, DC. Vol. 1, 675 pp.

Cao, W., W.B. Bowden, T. Davie, and A. Fenemor, 2006. Multi-Variable and Multi-Site Calibration and Validation of SWAT in a Large Mountainous Catchment With High Spatial Variability. Hydrological Processes 20:1057-1073.

Coffey, M.E., J.L. Workman, A.W. Taraba, and A. W. Fogel, 2004. Statistical Procedures for Evaluating Daily and Monthly Hydrologic Model Predictions. Transactions of the ASCE 47(1):59-68.

Crawford, N.H. and R.S. Linsley, 1966. Digital Simulation in Hydrology: Stanford Watershed Model IV. Technical Report 39, Stanford University, Department of Civil Engineering , Palo Alto, California.

Di Luzio, M., R. Srinivasan, and J.G. Arnold, 2002. Integration of Watershed Tools and SWAT Model Into BASINS. Journal of the American Water Resources Association 38(4):1127-1141.

Di Luzio, M., R. Srinivasan, and J.G. Arnold, 2004. A GIS-Coupled Hydrological Model System for the Watershed Assessment of Agricultural Nonpoint and Point Sources of Pollution. Transactions in GIS 8(1):113-136.

Draper, N.R. and H. Smith, 1981. Applied Regression Analysis (Second Edition). John Wiley and Sons, New York, 709 pp.

Duan, Q., S. Sorooshian, and V. Gupta, 1992. Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models. Water Resources Research 28(4):1015-1031.

Duan, Q., S. Sorooshian, and V. Gupta, 1994. Optimal use of the SCE-UA Global Optimization Method for Calibrating Watershed Models. Journal of Hydrology 158(3-4):265-284.

Eckhardt, K. and J.G. Arnold, 2001. Automatic Calibration of a Distributed Catchment Model. Journal of Hydrology 251(1-2):103-109.

Eckhardt, K., S. Haverkamp, N. Fohrer, and H.G. Frede, 2003. SWAT-G, a Version of SWAT99.2 Modified for Application to low Mountain Range Catchments. Physics and Chemistry of the Earth 27(9-10):641-644.

Farnes, P.E. and W.W. McCaughey, 1995. Hydrologic and Geologic Characterization of Tenderfoot Creek Experimental Forest, Montana. Unpublished Report. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Forestry Sciences Laboratory, Bozeman, Montana, 200 pp.

Flerchinger, G.N., K.R. Cooley, and Y. Deng, 1994. Impacts of Spatially and Temporally Varying Snowmelt on Subsurface Flow in a Mountainous Watershed: 1. Snowmelt Simulation. Hydrological Sciences Journal 39(5):507-520.

Flerchinger, G.N., K.R. Cooley, and D.R. Ralston, 1992. Groundwater Response to Snowmelt in a Mountainous Watershed. Journal of Hydrology 133(3-4):293-311.

Fontaine, T.A., T.S. Cruickshank, J.G. Arnold, and R.H. Hotchkiss, 2002. Development of a Snowfall-Snowmelt Routine for Mountainous Terrain for the Soil Water Assessment Tool (SWAT). Journal of Hydrology 262(1-4):209-223.

Gassman, P.W., M.R. Reyes, C.H. Green, and J.G. Arnold, 2005. Review of Peer-Reviewed Literature on the SWAT Model. *In:* Proceedings of the 3rd International SWAT Conference, R. Srinivasan, J. Jacobs, D. Day, and K. Abbaspour (Editors). Swiss Federal Institute for Environmental Science and Technology (EAWAG), Zurich, Switzerland, pp. 1-8.

Gassman, P.W., M.R. Reyes, C.H. Green, and J.G. Arnold, 2007. The Soil and Water Assessment Tool: Development History, Applications, and Future Research Directions. Transactions of the ASABE 50(4):1211-1250.

Gesch, D., M. Oimoen, S. Greenlee, C. Nelson, M. Steuck, and D. Tyler, 2002. The National Elevation Dataset. Photogrammetric Engineering and Remote Sensing 68(1):5-11.

Golding, D.L. and R.H. Swanson, 1978. Snow Accumulation and Melt in Small Forest Openings in Alberta. Canadian Journal of Forest Research 8(4):380-388.

Gray, D.M., B. Toth, L. Zhao, J.W. Pomeroy, and R.J. Granger, 2001. Estimating Areal Snowmelt Infiltration Into Frozen Soils. Hydrological Processes 15:3095-3111.

van Griensven, A. and W. Bauwens, 2003. Multiobjective Autocalibration for Semidistributed Water Quality Models. Water Resources Research 39(12):1348.

van Griensven, A., T. Meixner, S. Grunwald, T. Bishop, M. Di Luzio, and R. Srinivasan, 2006. A Global Sensitivity Analysis Tool for the Parameters of Multi-Variable Catchment Models. Journal of Hydrology 324(1-4):10-23.

Gupta, H.V., S. Sorooshian, and P.O. Yapo, 1998. Toward Improved Calibration of Hydrologic Models: Multiple and Noncommensurable Measures of Information. Water Resources Research 34(4):751-764.

Hall, R.J., D.P. Davidson, and D.R. Peddle, 2003. Ground and Remote Estimation of Leaf Area Index in Rocky Mountain Forest Stands, Kananaskis, Alberta. Canadian Journal of Remote Sensing 29(3):411-427.

Haverkamp, S., R. Srinivasan, H.G. Frede, and C. Santhi, 2002. Subwatershed Spatial Analysis Tool: Discritization of a Distributed Hydrologic Model by Statistical Criteria. Journal of the American Water Resources Association 38(6):1723-1733.

Hay, L.E., G.H. Leavesley, M.P. Clark, S.L. Markstrom, R.J. Viger, and M. Umemoto, 2006. Step Wise, Multiple Objective Calibration of a Hydrologic Model for a Snowmelt Dominated Basin. Journal of the American Water Resources Association 42(4):877-890.

Holsinger, L., R.E. Keane, R. Parsons, and E. Karau, 2005. Development of Biophysical Gradient Layers. *In*: The LANDFIRE Prototype Project: Nationally Consistent and Locally Relevant Geospatial Data for Wildland Fire Management, M.G. Rollins and C.K. Frame (Technical Editors). 2006. Gen. Tech. Rep. RMRS-GTR-175, U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado, pp. 99-122.

Jakeman, A.J. and R.A. Lechter, 2003. Integrated Assessment and Modelling: Features, Principles and Examples for Catchment Management. Environmental Modelling & Software 18(6):491-501.

Jha, M., J.G. Arnold, P.W. Gassman, F. Giorgi, and R.R. Gu, 2006. Climate Change Sensitivity Assessment on Upper Mississippi River Basin Streamflows Using Swat. Journal of the American Water Resources Association 42(4):997-1015.

Leavesley, G.H., R.W. Litchey, M.M. Troutman, and L.G. Saindon, 1983. Precipitation-Runoff Modeling System: User's Manual. Available from Books and Open File Report Section, U.S. Geological Survey Water-Resources Investigations Report 83-4238, Denver, Colorado, 207 pp.

Leonard, R.A., W.G. Knisel, and D.A. Still, 1987. Gleams: Groundwater Loading Effects on Agricultural Management Systems. Transactions of ASAE 30(5):1403-1428.

Lowman, M.D. and N.M. Nadkarni, 1995. Forest Canopies. Academic Press, San Diego, California.

Maidment, D.R. 1993. Handbook of Hydrology. McGraw-Hill, New York.

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

McCaughey, W.W. 1996. Tenderfoot Creek Experimental Forest. *In:* Experimental Forests, Ranges, and Watersheds in the Northern Rocky Mountains: A Compendium of Outdoor Laboratories in Utah, Idaho, and Montana, W.C. Schmidt and J.L. Friede (Compilers). U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, Forestry Sciences Laboratory, Bozeman, Montana, Gen. Tech. Rep. INT-GTR-334. pp. 101-108.

Monteith, J.L., 1965. Evapotranspiration and the Environment. *In:* The State and Movement of Water in Living Organisms, 19th Symposia of the Society for Experimental Biology. Cambridge University Press, London. pp 205-234.

Motovilov, Y.G., L. Gottschalk, K. Engeland, and A. Rodhe, 1999. Validation of a Distributed Hydrological Model Against Spatial Observations. Agricultural and Forest Meteorology 98-99:257-277.

Muleta, M.K. and J.W. Nicklow, 2005. Sensitivity and Uncertainty Analysis Coupled With Automatic Calibration for a Distributed Watershed Model. Journal of Hydrology 306 (1-4):127-145.

Nash, J.E. and J.V. Sutcliffe, 1970. River Flow Forecasting Through Conceptual Model Part 1 – A Discussion of Principles. Journal of Hydrology 10(3):282-290.

Neitsch, S.L., J.G. Arnold, J. R Kiniry, and J. R Williams, 2005. Soil and Water Assessment Tool Theoretical Documentation, Version 2005. Grassland, Soil, and Water Research Laboratory, Agriculture Research Service, Temple, Texas.

NRCS (U.S. Department of Agriculture, Natural Resource Conservation Service), 1994. State Soil Geographic Data Base (STATSGO): Data Use Information. U.S. Department of Agriculture, National Soil Survey Center, Lincoln, Nebraska, Miscellaneous Publication Number 1492. 113 pp.

NRCS (U.S. Department of Agriculture, Natural Resource Conservation Service), 1996. National Soil Survey Handbook, Title 430-VI. U.S. Government Printing Office, Washington, D.C.

Ohte, N., S.D. Sebestyen, J. B Shanley, D.H. Doctor, C. Kendall, S.D. Wankel, and E.W. Boyer, 2004. Tracing Sources of Nitrate in Snowmelt Runoff Using a High-Resolution Isotopic Technique. Geophysical Research Letters 31 (l21506) 1-4, doi:10.1029/2004gl020908.

Oliver, C.D. and B.C. Larson, 1990. Forest Stand Dynamics. McGraw-Hill, New York.

Pomeroy, J.W. and E. Burn, 2001. Physical Properties of Snow. *In:* Snow Ecology: An Interdisciplinary Examination of Snow-Covered Ecosystems, H. Jones, J.W. Pomeroy, D.A. Walker, and R.W. Hoham (Editors). Cambridge University Press, London, pp. 45-118.

Pomeroy, J.W., D.M. Gray, N.R. Hedstrom, and J.R. Janowicz, 2002. Prediction of Seasonal Snow Accumulation in Cold Climate Forests. Hydrological Processes 16(18):3543-3558.

Popov, E.G., 1973. Snowmelt Runoff Forecasts – Theoretical Problems. The Role of Ice and Snow in Hydrology, Proceedings Banff Symposium, Vol. 2 UNESCO-WMO-IASH, pp. 829-839.

Redmond, R.L., C. Winne, and C. Fisher, 2001. Existing Vegetation and Landcover of West-Central Montana. Final Scene Reports Submitted to U.S. Department of Agriculture, Forest Service, Northern Regional Office, Missoula, Montana.

Roberge, J. and A.P. Plamondon, 1987. Snowmelt Runoff Pathways in a Boreal Forest Hillslope, the Role of Pipe Throughflow. Journal of Hydrology 95:39-54.

Santhi, C., J.G. Arnold, J.R. Williams, W.A. Dugas, R. Srinivasan, and L.M. Hauck, 2001a. Validation of the SWAT Model on a Large River Basin With Point and non-Point Sources. Journal of the American Water Resources Association 37(5):1169-1187.

Santhi, C., J.G. Arnold, J.R. Williams, L.M. Hauck, and W.A. Dugas, 2001b. Application of a Watershed Model to Evaluate Management Effects on Point and Nonpoint Source Pollution. Transactions of ASCE 44(6):1559-1570.

Sappington, J., 2005. Snow Accumulation and Snowmelt in Thinned Lodgepole Pine Stands, West Central Montana. M.S. Thesis. The University of Montana, Missoula, Montana, 54 pp.

SCS (U.S. Department of Agriculture, Soil Conservation Service), 1972. SCS National Engineering Handbook, Section 4, Hydrology. Chapter 10, Estimation of Direct Runoff From Storm Rainfall. U.S. Department of Agriculture, Soil Conservation Service, Washington D.C, pp. 10.1-10.24.

Singh, V.P., 1995. Computer Models of Watershed Hydrology. Water Resources Publications, Highlands Ranch, Colorado.

Srinivasan, R. and J.G. Arnold, 1994. Integration of a Basin-Scale Water Quality Model With GIS. Water Resources Bulletin 30(33):453-462.

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

Stadler, D., H. Wunderli, A. Auckenthaler, and H. Flühler, 1996. Measurement of Frost-Induced Snowmelt Runoff in a Forest Soil. Hydrological Processes 10: 1293-1304.Stednick, J.D. 1996. Monitoring the Effects of Timber Harvest on Annual Water Yield. Journal of Hydrology 176:79-95.

Stednick, J.D., 1996. Monitoring the Effects of Timber Harvest on Annual Water Yield. Journal of Hydrology 176:79-95.

Tennant, D.L., 1976. Instream Flow Regimens for Fish, Wildlife, Recreation and Related Environmental Resources. Fisheries 1:6-10.

Troendle, C.A., 1983. The Potential for Water Yield Augmentation From Forest Management in the Rocky Mountains. Water Resources Bulletin 19(3):359-373.

Troendle, C.A. and R.M. King, 1985. The Effect of Timber Harvesting on the Fool Creek Watershed, 30 Years Later. Water Resources Research 21(12):1915-1922.

Troendle, C.A. and R.M. King, 1987. The Effect of Partial and Clearcutting on Streamflow at Deadhorse Creek, Colorado. Journal of Hydrology 90(1/2):145-157.

USGS (U.S. Department of the Interior, Geological Survey), 2005. MODIS/Terra Leaf Area Index/FPAR 8-day L4 Global 1km ISIN Grid. United States Geological Survey, Center for Earth Observation and Science (EROS), Sioux Falls, South Dakota.

Van Liew, M.W. and J. Garbrecht, 2003. Hydrologic Simulation of the Little Washita River Experimental Watershed Using SWAT. Journal of the American Water Resources Association 39(2):413-426.

Wang, X. and A.M. Melesse, 2005. Evaluation of the SWAT Model's Snowmelt Hydrology in a Northwestern Minnesota Watershed. Transactions of the ASAE 48(4):1359-1376.

Wang, X. and A.M. Melesse, 2006. Effects of STATSGO and SSURGO as Inputs on SWAT Model's Snowmelt Simulation. Journal of the American Water Resources Association 42(5):1217-1236.

Waring, R.H. and S.W. Running, 1998. Forest Ecosystems: Analysis at Multiple Scales. Academic Press, San Diego, California.

White, K.L. and I. Chaubey, 2005. Sensitivity Analysis, Calibration, and Validations for a Multisite and Multivariable SWAT Model. Journal of the American Water Resources Association 41(5):1077-1089.

White, J.D., S.W. Running, R. Nemani, R.E. Keane, and K.C. Ryan, 1997. Measurement and Remote Sensing of LAI in Rocky Mountain Montane Ecosystems. Canadian Journal of Forest Research 27:1714-1727.

Williams, J.R., A.D. Nicks, and J.G. Arnold, 1985. Simulator for Water Resources in Rural Basins. Journal of Hydraulic Engineering 111(6):970-986.

Woods, S.W., R. Ahl, J. Sappington, and W. McCaughey, 2006. Snow Accumulation in Thinned Lodgepole Pine Stands, Montana, USA. Forest Ecology and Management 235:202-211.

Wu, K. and Y.J. Xu, 2006. Evaluation of the Applicability of the SWAT Model for Coastal Watersheds in Southern Louisiana. Journal of the American Water Resources Association 42(5):1247-1260.