JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION
APRIL AMERICAN WATER RESOURCES ASSOCIATION 2005

HYDROLOGICAL MODELING OF THE IROQUOIS RIVER WATERSHED USING HSPF AND SWAT[1]

Jaswinder Singh, H. Vernon Knapp, J.G. Arnold, and Misganaw Demissie[2]

ABSTRACT: The performance of two popular watershed scale simulation models – HSPF and SWAT – were evaluated for simulating the hydrology of the 5,568 km^2 Iroquois River watershed in Illinois and Indiana. This large, tile drained agricultural watershed provides distinctly different conditions for model comparison in contrast to previous studies. Both models were calibrated for a nine-year period (1987 through 1995) and verified using an independent 15-year period (1972 through 1986) by comparing simulated and observed daily, monthly, and annual streamflow. The characteristics of simulated flows from both models are mostly similar to each other and to observed flows, particularly for the calibration results. SWAT predicts flows slightly better than HSPF for the verification period, with the primary advantage being better simulation of low flows. A noticeable difference in the models' hydrologic simulation relates to the estimation of potential evapotranspiration (PET). Comparatively low PET values provided as input to HSPF from the BASINS 3.0 database may be a factor in HSPF's overestimation of low flows. Another factor affecting baseflow simulation is the presence of tile drains in the watershed. HSPF parameters can be adjusted to indirectly account for the faster subsurface flow associated with tile drains, but there is no specific tile drainage component in HSPF as there is in SWAT. Continued comparative studies such as this, under a variety of hydrologic conditions and watershed scales, provide needed guidance to potential users in model selection and application.
(KEY TERMS: watershed management; hydrologic cycle; nonpoint source pollution; surface water; modeling; agriculture.)

Singh, Jaswinder, H. Vernon Knapp, J.G. Arnold, and Misganaw Demissie, 2005. Hydrological Modeling of the Iroquois River Watershed Using HSPF and SWAT. Journal of the American Water Resources Association (JAWRA) 41(2): 343-360.

INTRODUCTION

Many computer models have been developed to simulate watershed hydrology and water quality processes. Watershed models are essential and effective tools for investigating the complex nature of processes that affect surface and subsurface hydrology, soil erosion, and the transport and fate of chemical constituents in watersheds and for assessing the impacts of land use changes, agricultural activities, and best management practices on these hydrologic processes. Two continuous simulation models that are commonly used for watershed management assessment are the Hydrological Simulation Program – FORTRAN (HSPF) (Bicknell *et al.,* 2001) and the Soil and Water Assessment Tool (SWAT) (Arnold *et al.,* 1998). Both of these models are included within Version 3.0 of the modeling framework developed by the U.S. Environmental Protection Agency (USEPA), referred to as Better Assessment Science Integrating Point and Nonpoint Sources (BASINS 3.0) (USEPA, 2001).

Major watershed restoration efforts are under way in Illinois to reduce sediment loads and nutrient concentrations and to improve the ecosystem along the Illinois River and its tributaries. As part of restoration efforts, hydrologic models for the Illinois River Basin are being applied and evaluated in the Illinois State Water Survey. Part of the overall process in applying models for the Illinois River Basin is a determination of which model(s) will perform best under varying watershed scales in simulating hydrology, sediments, and nutrients. The objective of the present

[1]Paper No. 04030 of the *Journal of the American Water Resources Association* (JAWRA) (Copyright © 2005). **Discussions are open until October 1, 2005.**
[2]Respectively, Assistant Professional Scientist, Illinois State Water Survey, 1320 SW Monarch Street, Peoria, Illinois 61602; Senior Hydrologist, Illinois State Water Survey, 2204 Griffith Drive, Champaign, Illinois 61820; Agricultural Engineer, USDA-ARS, 808 East Blackland Road, Temple, Texas 76502; and Director, Center for Watershed Science, Illinois State Water Survey, 2204 Griffith Drive, Champaign, Illinois 61820 (E-Mail/Singh: jsingh@sws.uiuc.edu).

study is to compare and assess the suitability of the HSPF and SWAT models for simulating the hydrology of one major tributary of the Upper Illinois River Basin, the Iroquois River watershed (IRW), which is representative of the land use and soils throughout much of the Illinois River Basin.

Qualitative information and general guidelines about models are most often passed from one model user to another or through Internet message boards such as the USEPA's BASINS list serve. But few studies are available to directly compare model performance that practitioners can use to determine which model is likely to be the best for a certain application. Several model reviews describe the relative capabilities and mathematical bases of various models (Deliman *et al.,* 1999; Franchini and Pacciani, 1999; Borah and Bera, 2003). However, to directly assess model performance, it is necessary to test models on real watershed applications.

Four recent studies have compared the hydrologic (streamflow) simulation capabilities of the HSPF and SWAT models: Im *et al.* (2003), Nasr *et al.* (2003), Van Liew *et al.* (2003), and Saleh and Du (2004). Two of these studies – Im *et al.* (2003) and Saleh and Du (2004) – also compared the sediment, nitrogen, and phosphorous simulation capabilities of the models, and Nasr *et al.* (2003) also evaluated phosphorous simulation. Van Liew *et al.* (2003) and Saleh and Du (2004) applied HSPF and SWAT to clusters of rural watersheds in Oklahoma and Texas, respectively, with watersheds ranging in size from less than 10 km^2 to as much as 922 km^2. Nasr *et al.* (2003) applied both models to a small (23 km^2) rural watershed in Ireland, and Im *et al.* (2003) applied the models to a 119 km^2 urbanizing watershed in Virginia.

These four comparative studies concluded that the HSPF model generally produced better results in streamflow simulation during model calibration. However, it is also worth noting that it is possible for HSPF to calibrate best in one watershed while SWAT calibrates best in a nearby watershed, as shown in the results by Van Liew *et al.* (2003) and Saleh and Du (2004). Additional factors other than model type that can affect calibration performance are the availability and accuracy of input data, including precipitation, and the variability of watershed characteristics such as land use and soils. Although HSPF generally performed best for calibration in the four studies, the performances of the models for validation periods are more mixed, as are the simulation results for estimating nutrient loadings. Of particular note, Van Liew *et al.* (2003) indicate that the SWAT model was more robust and gave better results when validating the models and transferring parameters for use with similar nearby watersheds; they also suggest that the SWAT model may be better suited for evaluating the impacts of climate variability on surface water resources.

From the studies cited, it is not always possible to identify and compare the extent to which each model was calibrated. Three of the four studies (Im *et al.,* 2003; Van Liew *et al.,* 2003; and Saleh and Du, 2004) indicated that the calibration of HSPF was less user-friendly and more difficult and time-consuming to learn, apply, and calibrate because of the numerous parameters to adjust and greater data preparation needs. Yet apparently for most watersheds the additional effort produced a more accurate hydrograph calibration.

The HSPF and SWAT comparative studies cited above were conducted for small to medium watersheds, of less than 1,000 km^2. Because no one model is best under all conditions, a complete understanding of comparative model performance requires applications under differing hydrologic conditions and watershed scales. The IRW is much larger (5,568 km^2), predominantly row cropped, and extensively tile drained. Thus, this study provides a valid example of comparative model performance evaluation for application to a large, tile drained agricultural watershed as typically found in the humid climatic region of the Midwestern United States. The suitability of HSPF and SWAT was evaluated for simulating the hydrology of the IRW for a long period representing a combination of dry, average, and wet years. HSPF (Version 12.0) was used within the BASINS 3.0 framework (USEPA, 2001), which facilitates data input and other processes within a geographic information system (GIS) framework. The version used for this study was SWAT 2000 (Neitsch *et al.,* 2002a), which was run within the ArcView SWAT (AVSWAT) interface (Di Luzio *et al.,* 2002) that also facilitates data inputs and other functions within a GIS framework.

MATERIALS AND METHODS

Iroquois River Watershed

The IRW is part of a larger study area of the Illinois River Basin that is a focus of the long term ecosystem restoration assessment study. The land use, physiography, and soils in this watershed represent conditions existing in most of the Illinois River Basin. The 151 km long Iroquois River drains about 5,568 km^2 in eastern Illinois and western Indiana (Figure 1). It flows west into Illinois from Indiana and drains into the Kankakee River at Aroma Park, Illinois. The Kankakee River flows farther northwest for 61 km until it merges with the Des Plaines River to

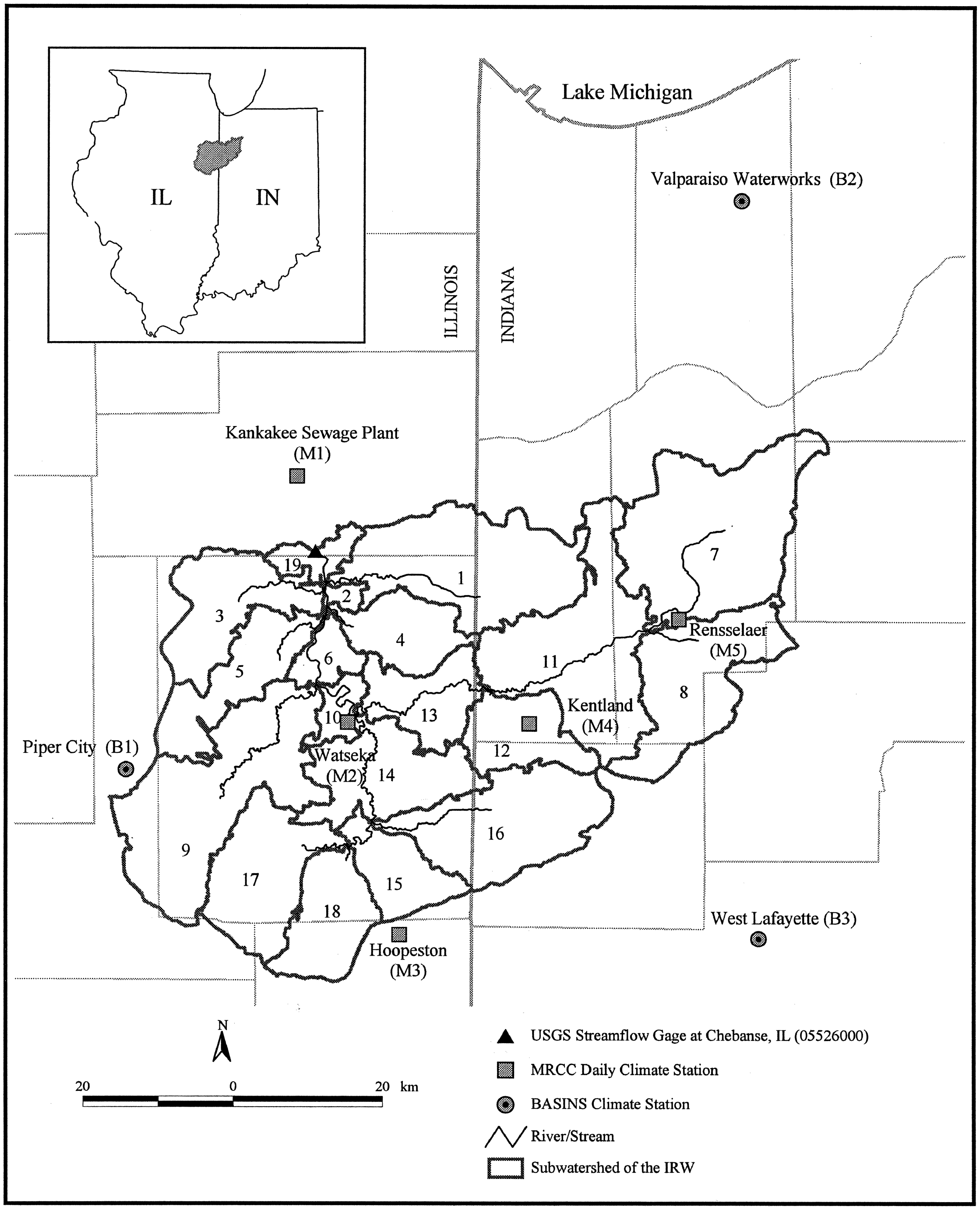

Figure 1. Iroquois River Watershed and Location of Climate and Streamflow Gaging Stations.

form the Illinois River. The average daily minimum and maximum temperatures for the IRW are 6°C and 16°C, respectively, and average annual precipitation is 990 mm.

Originally, a large portion of the IRW was prairie of nearly level to gently sloping topography and poor drainage (Knapp, 1992). Much of the region is an old glacial lake bed (Lake Watseka) and has predominantly flat topography (75 percent of the land area has slopes less than 2 percent). The soils are predominantly a heterogeneous mix of silts or clays with some local deposits of sand in the Indiana portion and the northern part of the watershed in Illinois. The average slope for the lower 129 km of the Iroquois River is less than 0.02 percent. A prominent rock outcrop near Chebanse, Illinois, maintains a nearly level pool for more than 32 km. In the western part of the watershed there are many artesian wells that contribute to the flow of the river (Page *et al.,* 1992). Agriculture accounts for 95 percent of land use in the watershed. Soybean and corn are commonly grown row crops, and subsurface tiles drain fields under predominantly silty-clay loam soils. Forest and urban land use cover 2.9 percent and 1.2 percent of the watershed area, respectively.

Brief Description of Models

HSPF is a comprehensive, conceptual, continuous simulation watershed scale model that simulates non-point source hydrology and water quality, combines it with point source contributions, and performs flow and water quality routing in the watershed reaches. Values of a large number of HSPF parameters cannot be obtained from field data and need to be determined through a model calibration exercise. However, many of these parameters were conceived to index properties of specific factors that influence events such as water storage and fluxes in the land phase of the hydrologic cycle (James, 1972). The model has three main modules, PERLND, IMPLND, and RCHRES, which simulate pervious land segments, impervious land segments, and free flow reaches/mixed reservoirs, respectively. HSPF estimates surface runoff using hourly time step as a function of infiltration computed using Philip's equation (Philip, 1957). The model uses a storage routing technique to route water from one reach to the next during stream processes. The hydraulic characteristics of reaches in the model are defined by parameters in the function tables (FTABLES) that represent volume discharge relations for reaches. The FTABLES can be modified based on observed hydraulic data. Actual evapotranspiration (ET) is a function of the PET (user input) and the amount of water available in the soil profile and on the land surface. There is no plant growth component in HSPF, and the effects of vegetation type, density, root growth, stage of development, and moisture characteristics of the soil layer are lumped into the parameter (LZETP) that controls actual ET from the root zone storage. There is no tile flow component in the HSPF. However, the efficient water removal effect from the field due to tiling is lumped in the parameters that control interflow inflow and discharge.

SWAT is a complex, continuous simulation conceptual model with spatially explicit parameterization (Arnold *et al.,* 2000). SWAT can predict, over long periods, the impact of land management practices on water, sediment, and agricultural chemical loads in large, complex watersheds with varying soils, land use, and management conditions. Major model components describe processes associated with water movement, sediment movement, soils, temperature, weather, plant growth, nutrients, pesticides, and land management. In each spatial subunit of the watershed, the water balance is represented by several storage volumes (e.g., canopy storage, snow, soil profile, shallow aquifer, and deep aquifer). Surface runoff is calculated using the modified Soil Conservation Service (SCS) curve number CN2 (USDA-SCS, 1972) technique when a daily time step is used or the Green and Ampt (1911) infiltration equation when an hourly or subdaily time step is used. CN2 is varied nonlinearly with the moisture content of the soil. Either variable storage or the Muskingum routing method is used for flow routing in the stream channels. Soil water processes include infiltration, evaporation, plant uptake, lateral flow, and percolation to deeper layers. Actual ET is computed as sum of actual soil evaporation and plant transpiration. Actual soil evaporation is estimated by using exponential functions of soil depth and water content. Plant transpiration is simulated as a linear function of potential ET (user input), leaf area index, and rooting depth and can be limited by soil water content. SWAT has a simple tile flow component in which the user specifies tile depth, the amount of time required to drain the soil to field capacity, and the time lag between the water entering the tile and discharging into the main channel. Tile drainage occurs when the soil water content exceeds the field capacity.

Model Preparation Using HSPF and SWAT

Based on its topography and existing stream network, the IRW was divided into 19 smaller, hydrologically connected subwatersheds and associated stream reaches using the automatic delineation tool of each model's GIS interface. A U.S. Geological Survey (USGS) elevation data (digital elevation model, DEM) layer and a predigitized stream network data layer (National Hydrography Dataset, NHD) were used to perform this task. A digitized soil information layer (NRCS-STATSGO soil database) and land use/land cover data layer (USGS-GIRAS database) were used for further subclassification of areas in the watershed. All the above GIS data layers for the IRW were taken from the BASINS 3.0 database obtained on compact discs from the USEPA in April 2002.

In HSPF, each subwatershed was partitioned into pervious and impervious areas based on land uses such as urban, agriculture, forest, barren, and wetland/water areas. Since BASINS-HSPF did not automatically create segments based on soil types, the dominant soil (Hydrologic Soil Group B) was assumed to be representative of the IRW soil conditions. Such an approach has been used in some previous HSPF studies (Donigian *et al.,* 1983; Jones and Winterstein, 2000). Because major hydrologic differences occur between pervious and impervious land use types and since agriculture is the major pervious land use in the IRW, all pervious land segments in the model were assigned the same hydrologic parameters. The storage routing scheme was used for channel routing. The FTABLES were not modified, and the default volume-discharge relationship was used as calculated by the model based on DEM and NHD data. The hourly time series of climate data required for hydrologic simulations using HSPF include precipitation, potential ET, potential surface evaporation, air temperature, dewpoint temperature, wind speed, and solar radiation. The BASINS-3.0 database contains complete sets of climate data for the period 1971 through 1995 for selected climate stations throughout the United States. Only one such station, at Piper City (Station B1, Figure 1), is located near the IRW. Five additional climate stations with complete daily precipitation and temperature data for 1971 through 1995 were identified in or near the watershed (Stations M1-M5, Figure 1) for use in this study. Data for these additional five stations were obtained from the Midwestern Regional Climate Center (MRCC). The daily precipitation data from these five stations were disaggregated into hourly data using the Data Disaggregation Tool in the HSPF Watershed Data Management Utility (WDMUtil). Hourly precipitation time series data from the Piper City Station and two additional BASINS stations in the vicinity of the IRW (Stations B2 and B3, Figure 1) were used as references for disaggregation. The hourly time series for the remaining climatic variables were obtained for the M1, M2, and M3 stations using data from the B1 station, for M4 using data from the B3 station, and for M5 using data from the B2 station. Potential ET in BASINS 3.0 climate stations is based on the standard class-A pan evaporation data adjusted using a regional coefficient.

To account for different land use and soils, functional modeling units called hydrologic response units (HRUs) were created in each subwatershed in SWAT based on the unique intersection of the land use and soils. All possible combinations of soil types and land use covering more than 1 percent of each subwatershed were included, resulting in 252 HRUs for the IRW. Row crop areas were equally split between soybean and corn. Other model parameters such as channel geometry and the length and slope of the overland flow path, which relate to physical dimensions of the watershed, were kept the same in both models. The Muskingum routing scheme was used for channel routing. The SCS curve number method was used for runoff simulation; thus, only the daily precipitation and maximum and minimum air temperature data were input from the six climate stations (Figure 1). Other daily time series of wind speed, solar radiation, and relative humidity were generated internally in SWAT, using monthly weather statistics for the nearest available climate station in the SWAT weather generator database. Potential ET was computed using the Penman-Monteith (Monteith, 1965) method. Climate stations were assigned to each subwatershed based on proximity.

Model Calibration and Verification

The hydrologic components of HSPF and SWAT were calibrated to fit the observed daily streamflow data from a USGS streamflow gaging station (05526000) at Chebanse, Illinois (Figure 1) for the nine-year period of 1987 through 1995. This period was chosen because it represents a combination of dry, average, and wet years (annual precipitation ranged from 686 to 1,473 mm), with an average annual precipitation of 960 mm as compared to the 30-year average of 990 mm. Both models were run for the 11-year period of 1985 through 1995. The first two years were used to stabilize the simulation runs. The simulated streamflows for the remaining nine years (1987 through 1995) were compared with corresponding observed values to evaluate the accuracy of the models. Values of selected model parameters were varied iteratively within a reasonable range during the calibration runs until a satisfactory agreement between observed and simulated streamflow data was obtained. Both models were then verified using observed streamflow data from the same USGS gage at Chebanse but for a different 15-year period (1972 through 1986) that was not used for model calibration. The annual precipitation ranged from 813 to 1,219 mm during this period, and average annual precipitation was 998 mm. The model was run for the 17-year period of 1970 through 1986, but the first two years were again used to stabilize the simulation runs, and the simulated streamflows for 1972 through 1986 were compared with the observed data to verify the models. Performance of the two models was also compared for three distinct periods within the verification period: a drier than average period (1978 through 1980), an average period (1981 through 1986), and a wetter than average period (1972 through 1977). The average annual precipitation for these three periods, based on Stations B1 and M1-M5, was 879 mm, 998 mm, and 1,062 mm, respectively, which is a deviation of -11.3 percent, 0.7 percent, and 7.0 percent from the 30-year average of 990 mm. The standard error of daily streamflows at the Chebanse gage is considered to be less than 5 percent, as based on instantaneous discharge measurements by the USGS (Mades and Oberg, 1986).

The definitions of various HSPF model parameters calibrated in this study are given in Table 1. These parameters and the ranges within which their values were varied were selected based on other HSPF evaluation studies, specifically Chew *et al.* (1991); Laroche *et al.* (1996); Duncker and Melching (1998); Bergman and Donnangelo (2000); Jones and Winterstein (2000); and the BASINS Technical Note 6 (USEPA, 2000). Temporally varied values of several model parameters were used in this study (Table 1). The model was run on an hourly time step, and output was obtained on a daily basis. A stepwise approach was used for HSPF calibration in which an acceptable match was obtained between annual and monthly streamflows, and then the parameter values were further adjusted to obtain a satisfactory agreement between observed and simulated streamflow values. This approach was supported by the hierarchical structure of HSPF in which annual streamflows are affected by one set of parameters (e.g., LZETP, DEEPFR, LZSN, and INFILT parameters), monthly flows by another set (UZSN, BASETP, KVARY, AGWRC, and CEPSC), and stormflows by a third set (e.g., INFILT, INTFW, and IRC). Snowmelt and freezing processes in the watershed were simulated by changing the values of the SNOWCF, TSNOW, and CCFACT parameters.

SWAT was calibrated using a similar stepwise procedure. The six model parameters used for calibration and the ranges within which their values were varied (Table 1) were selected based on calibration guidelines provided in Arnold *et al.* (2000); Santhi *et al.* (2001); Neitsch *et al.* (2002a); and Van Liew and Garbrecht (2003). The model was run using a daily time step. Spatially varied values of soil physical properties (available water capacity, saturated conductivity, bulk density, texture, and organic matter) were assigned to different HRUs by SWAT based on NRCS-STATSGO database and were not calibrated. Surface runoff is extremely sensitive to parameter CN2, and decreasing the CN2 values results in decreased runoff and increased infiltration, baseflow, and recharge. The soils of Hydrologic Soil Group B may act as Group A soils when properly drained using

TABLE 1. List of HSPF and SWAT Calibration Parameters.

Parameter	Definition	Range	Calibrated Value
HSPF			
KVAR (per in)	Variable Ground Water Recession Flow	0.0 to 3.0	3.00
INFILT (in/h)	Index to Soil Infiltration Capacity	0.01 to 0.25	0.20
AGWRC (per d)	Basic Ground Water Recession Rate	0.92 to 0.99	0.98
LZSN (in)	Lower Zone Nominal Storage	3.0 to 8.0	5.00
UZSN (in)	Upper Zone Nominal Storage	0.05 to- 2.0	0.2 to 1.4*
BASETP	Baseflow Evapotranspiration	0.0 to 0.2	0.10
DEEPFR	Fraction of Inactive Ground Water	0.0 to 0.2	0.05
NSUR	Manning's n for Overland Flow	0.15 to 0.35	0.20
CEPSC (in)	Interception Storage Capacity	0.03 to 0.20	0 to 0.1*
INTFW	Interflow Inflow Parameter	1.0 to 3.0	1.2 to 1.8*
IRC	Interflow Recession Constant	0.3 to 0.85	0.6 to 0.8*
LZETP	Lower Zone Evapotranspiration	0.1 to 0.9	0.1 to 0.75*
TSNOW (°F)	Temp. at Which Precipitation is Snow	31 to 33	33.00
SNOWCF	Snow Gage Catch Correction Factor	1.1 to 1.5	1.20
SWAT			
CN2	Runoff Curve Number	64 to 76	67
ESCO	Soil Evaporation Compensation Factor	0.7 to 1.0	0.95
EPCO	Plant Uptake Compensation Factor	0.4 to 0.9	0.7
DDRAIN (mm)	Depth of Subsurface Drains	900 to 1,200	1,100
GW_DELAY (d)	Ground Water Delay	31 to 150	45
ALPHA_BF (d)	Baseflow Alpha Factor	0.018 to 1.0	0.07

*Monthly value range.

subsurface tile drains, as in the IRW. Therefore a smaller CN2 value of 67, which is applicable for Group A soils, was used in this study for corn and soybean. For HRUs of the same land use type, Van Liew *et al.* (2003) also used a single CN2 value. For urban and forest HRUs, SWAT's default CN2 values were used as such. Parameters ESCO and EPCO were varied to adjust the depth distribution for evaporation and plant uptake of water, respectively, from the soil profile. Similar to Santhi *et al.* (2001) and Saleh and Du (2004), single values of EPCO and ESCO were used for the entire IRW in this study. Parameter ALPHA_BF affects the recession limb of the simulated hydrograph, while GW_DELAY governs the watershed response in terms of time required for water leaving the bottom of the root zone to reach the shallow aquifer. The initial value of the ALPHA_BF for the entire IRW was determined using the baseflow filter method of Arnold and Allen (1999), and it was later calibrated. The depth of subsurface drains, DDRAIN, was adjusted between 900 and 1,200 mm. Sogbedgi and McIsaac (2002) recommended a value of 1,200 mm for central Illinois. Values of two other subsurface drainage-related parameters that affect the time to drain the soil profile and the time until water enters the channel network after entering the tiles were set to 24h and 48h, respectively, based on Neitsch *et al.* (2002b). Calibration using all six parameters initially focused on fitting the annual and monthly streamflows. Fine-tuning of parameters CN2, ESCO, GW_DELAY, and ALPHA_BF was then performed to obtain a reasonable match between observed and simulated daily flows.

Model Performance Evaluation

During model calibration and verification, agreement between observed and simulated streamflows on an annual, monthly, and daily basis was evaluated using both statistical and graphical measures. Statistical measures of agreement were the percent deviation (D_v) in the simulated streamflow volume with respect to the observed volume for long term, annual, and monthly basis; the Nash-Sutcliffe model efficiency (R) for monthly and daily flows (Nash and Sutcliffe, 1970); and the prediction efficiency (P_e) for daily flows (Van Liew *et al.,* 2003). Donigian *et al.* (1983) state that the annual and monthly fits between simulated and observed streamflows can be considered "very good" when the $|D_v|$ for these individual fits is 10 percent or less, "good" when it is between 10 and 15 percent, and "fair" when it is between 15 and 25 percent. These criteria were adopted for both HSPF and SWAT simulations. The R indicates how well the plot of observed versus simulated data fit the 1:1 line, and a value of R = 1.0 indicates perfect fit. In this study model calibration was considered satisfactory when $R \geq 0.75$ was obtained for daily streamflow comparison. Scatter plots of observed streamflow versus simulated/observed streamflow ratios (S/O) were used to determine if there were any systematic errors related to the magnitude of the observed monthly or daily flows. Daily streamflows were also compared by plotting the flow frequency (flow duration) curves. General agreement between the observed and simulated flow duration curves indicates adequate calibration over the range of the flow conditions simulated. The P_e is the quantitative measure of this agreement. It is essentially the coefficient of determination (r^2) between observed and simulated daily streamflows after sorting each series in ascending (or descending) order.

MODELING RESULTS

Model calibration statistics, comparing observed and simulated flows for annual, monthly and daily time intervals, are presented in Tables 2 and 3. Table 3 also shows statistical comparisons for the entire 15-year verification period as well as the three shorter periods representing varying climatic conditions. A comparison of observed flows to the S/O ratio is shown in Figure 2 for monthly values and in Figure 3 for daily values. Figure 4 shows the flow duration values of simulated and observed daily flows. For the 15-year verification period, a comparison of observed and simulated flows is shown in Figure 5 for annual values, in Figure 6 for monthly values, in Figures 7 for the daily S/O ratios, and Figure 8 for the daily flow duration values. Figures 2 and 6 also show the seasonal changes in observed monthly flow values over the entire calibration and verification periods, respectively. Similarly, the observed flow values for each day of a year over the calibration period are also shown in Figure 3. Figure 9 is a time-series plot of observed and simulated daily flows for the drier than average period.

Model Calibration Using 1987 Through 1995 Streamflows

Over the nine-year calibration period, HSPF overestimated the streamflow by 4.6 percent whereas SWAT underestimated it by 2.5 percent. The differences between annual observed and simulated streamflows were within 15 percent in seven of these years for HSPF and in six years for SWAT (Table 2).

TABLE 2. Watershed Water Balance During Calibration Period for HSPF and SWAT.

Indicator	1987	1988	1989	1990	Year 1991	1992	1993	1994	1995	Average
Precipitation (mm)	851	683	864	1229	795	902	1471	859	988	960
V_{obs} (mm)	193	160	246	490	363	272	833	284	330	351
HSPF										
PET (mm)	737	729	660	681	739	645	671	704	699	696
AET_{sim} (mm)	594	427	551	612	503	564	632	592	617	566
V_{sim} (mm)	246	236	279	503	356	277	716	320	376	368
D_v (percent)	28.9	49.4	13.5	2.5	-2.0	2.3	-13.9	12.4	14.0	4.6
SWAT										
PET (mm)	1,359	1,422	1,290	1,313	1,382	1,240	1,171	1,313	1,270	1,306
AET_{sim} (mm)	648	460	625	668	518	655	683	615	610	610
V_{sim} (mm)	160	221	267	434	368	211	772	277	376	343
D_v (percent)	-16.5	38.9	8.7	-11.7	1.6	-22.3	-7.2	-2.4	14.3	-2.5

V_{obs} = Observed streamflow volume; HSPF = Hydrological Simulation Program FORTRAN; PET = Potential evapotranspiration; AET_{sim} = Simulated actual evapotranspiration; V_{sim} = Simulated streamflow volume; D_v = Deviation of the streamflow volume; and SWAT = Soil and Water Assessment Tool.

TABLE 3. Model Calibration and Verification Statistics for HSPF and SWAT.

Indicator			Model Calibration (1987 to 1995)	Model Verification Long Term (1972 to 1986)	Dry Period (1978 to 1980)	Average Period (1981 to 1986)	Wet Period (1972 to 1977)
AVERAGE ANNUAL							
Precipitation (mm)			960	999	879	998	1,061
V_{obs} (mm)			352	337	276	380	351
Vsim (mm)		HSPF	368	358	277	368	408
		SWAT	343	351	278	366	399
D_v (percent)		HSPF	4.6	5.3	0.3	-3.2	16.4
		SWAT	-2.5	4.0	0.8	-3.7	13.7
MONTHLY							
	R	HSPF	0.88	0.82	0.87	0.80	0.81
		SWAT	0.88	0.84	0.93	0.81	0.80
Number of Months With D_v Within	±10 percent	HSPF	22	28	6	9	13
		SWAT	27	34	5	10	19
	±15 percent	HSPF	36	40	8	14	18
		SWAT	44	48	6	19	23
	±25 percent	HSPF	49	71	10	28	33
		SWAT	64	72	15	26	31
DAILY							
	R	HSPF	0.81	0.69	0.69	0.69	0.71
		SWAT	0.79	0.74	0.83	0.72	0.70
	P_e	HSPF	0.99	0.99	0.98	0.99	0.98
		SWAT	0.99	0.99	0.99	0.99	0.99

V_{obs} = Observed streamflow volume; V_{sim} = Simulated streamflow volume; D_v = Deviation of the streamflow volume; HSPF = Hydrological Simulation Program FORTRAN; SWAT = Soil and Water Assessment Tool; R = Nash-Sutcliffe (1970) model efficiency; and P_e = Prediction efficiency.

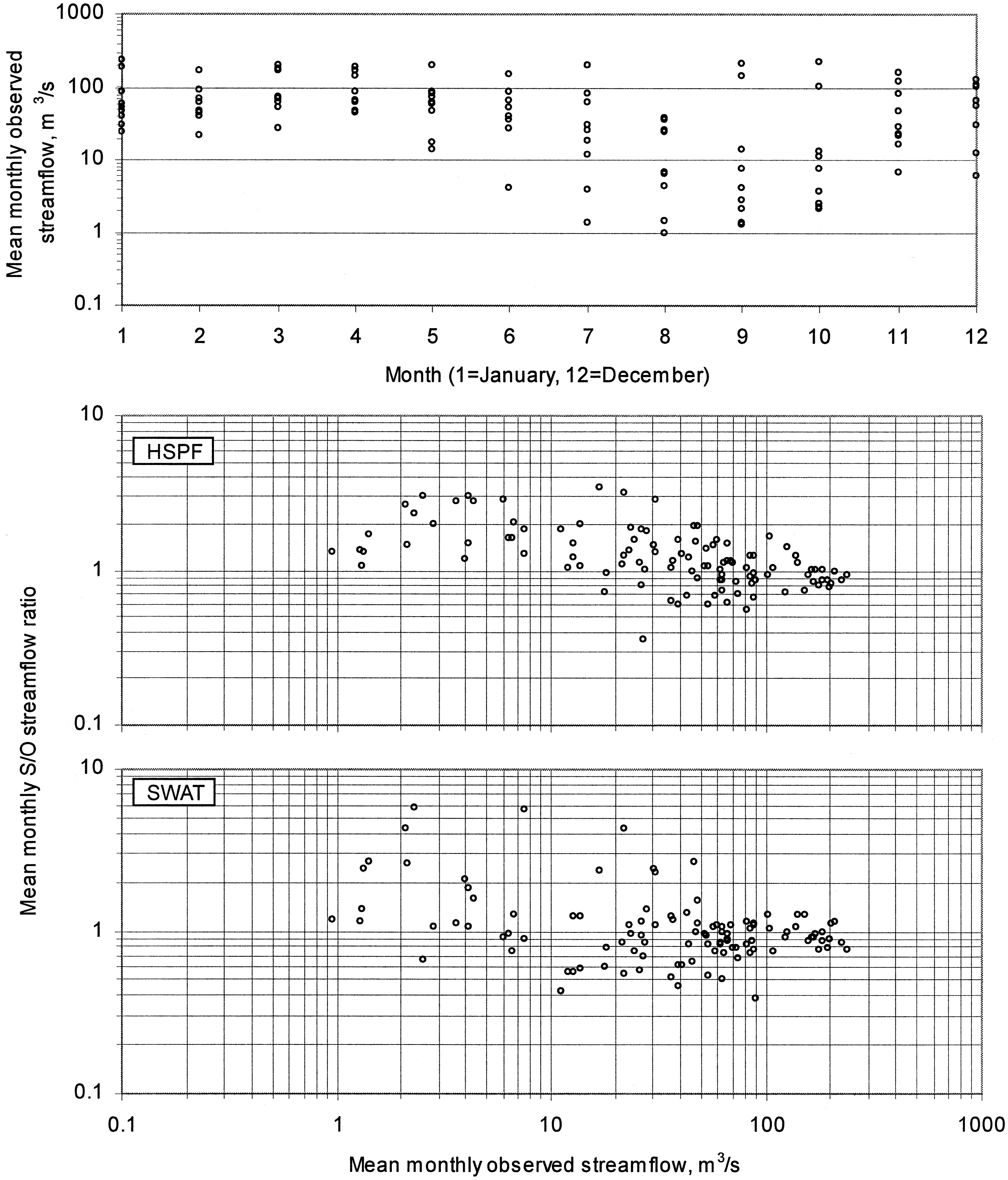

Figure 2. Scatter Plot of Months of a Year Versus Mean Monthly Observed Streamflows During Those Months (top) and Mean Monthly Observed Streamflow Versus Mean Monthly Simulated/Observed Streamflow Ratio for the Calibration Period for HSPF and SWAT.

Both models overestimated streamflow during the drought year of 1988 (HSPF by 49.4 percent and SWAT by 38.9 percent) and underestimated it during the wettest year, 1993 (HSPF by 13.9 percent and SWAT by 7.2 percent). Over the calibration period, 37 percent of the precipitation falling over the IRW was ultimately converted into streamflow. The simulated mean annual flow and actual ET values from HSPF over this period were 38 percent and 59 percent of the total precipitation amount and 36 percent and 63 percent from SWAT, respectively. The simulated monthly flows from both models had a high R value of 0.88 (Table 3). The simulated monthly flows from HSPF were within 15 percent from the observed values for

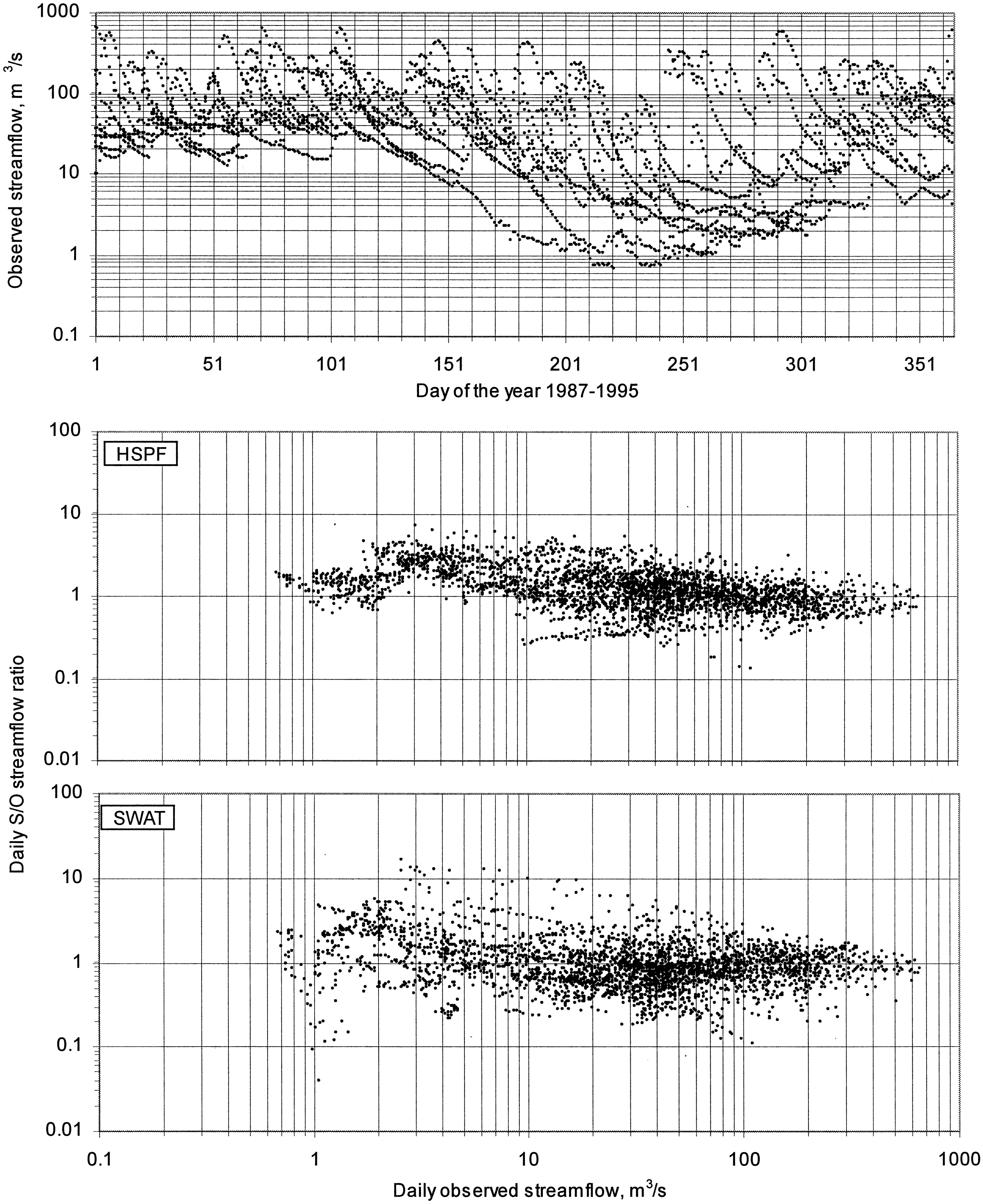

Figure 3. Scatter Plot of Days of a Year Versus Observed Streamflows During Those Days (top) and Daily Observed Streamflow Versus Daily Simulated/Observed Streamflow Ratio for the Calibration Period for HSPF and SWAT.

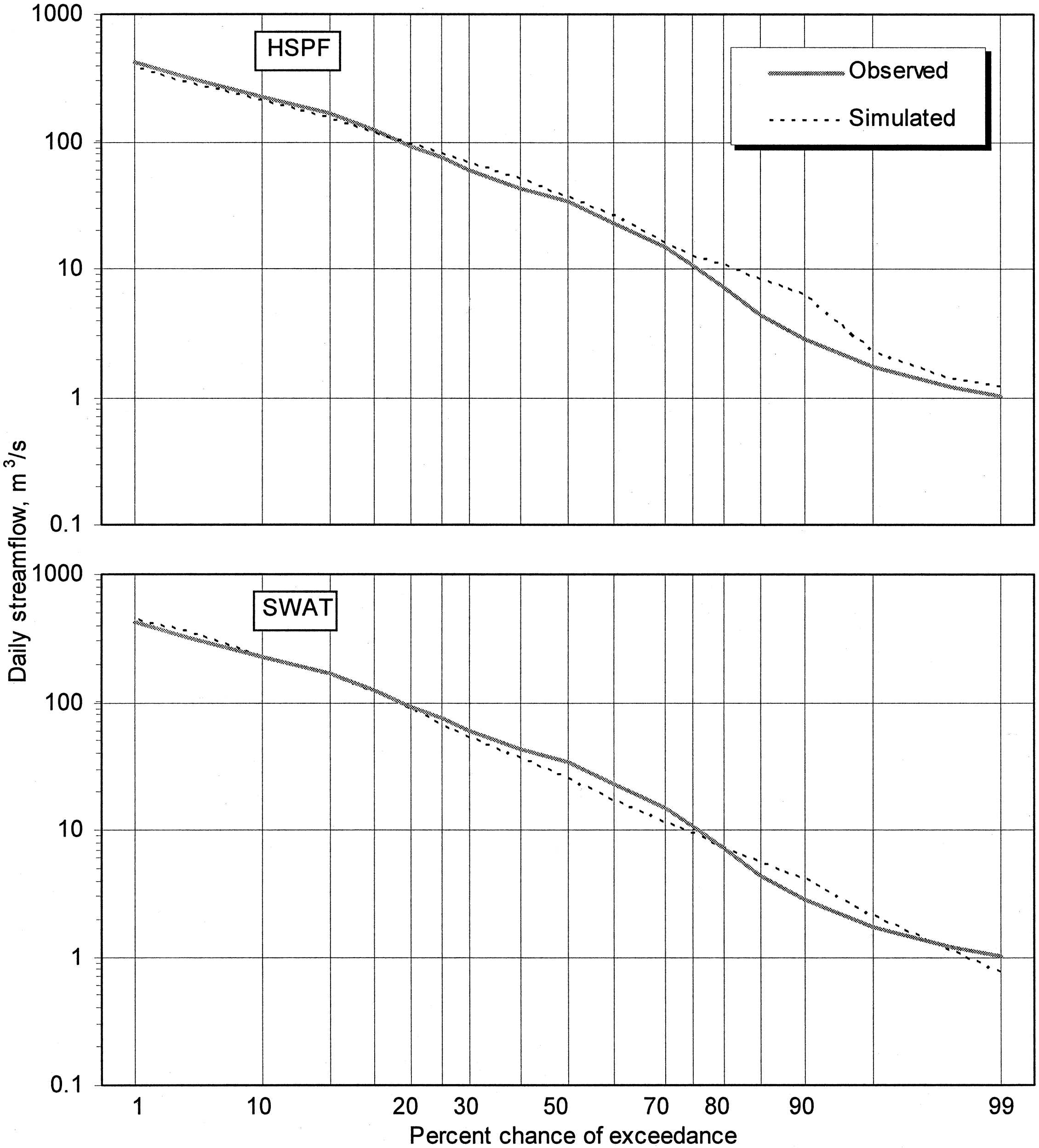

Figure 4. Observed and Simulated Daily Streamflow Duration Curves for the Model Calibration Period for HSPF and SWAT.

36 of the 108 months during the calibration period, while SWAT monthly flows were within 15 percent for 44 months. Mean monthly flows greater than 25 m^3/s were estimated well by both models as indicated by the uniform scatter of points around the S/O equal to 1 line in Figure 2. HSPF was biased in overestimating flows less than 25 m^3/s. SWAT results show a wider scatter of data points around S/O equal to 1 line for flows less than 25 m^3/s but generally less bias than that shown for HSPF (Figure 2). SWAT also has a wider scatter of data points than HSPF in the daily comparison of observed flow versus the S/O ratio (Figure 3). HSPF generally overestimated flows less than 12 m^3/s, while SWAT showed the tendency to largely underestimate some flows less than 3 m^3/s (Figure 3). For other flows, the data points were evenly scattered around the S/O equal to 1 line for both models. The daily flow duration curves (Figure 4) also confirm the HSPF tendency to overestimate low flows. Overall, both models estimated the daily flows satisfactorily, as indicated by high daily R value of 0.81 and 0.79 for HSPF and SWAT, respectively, and

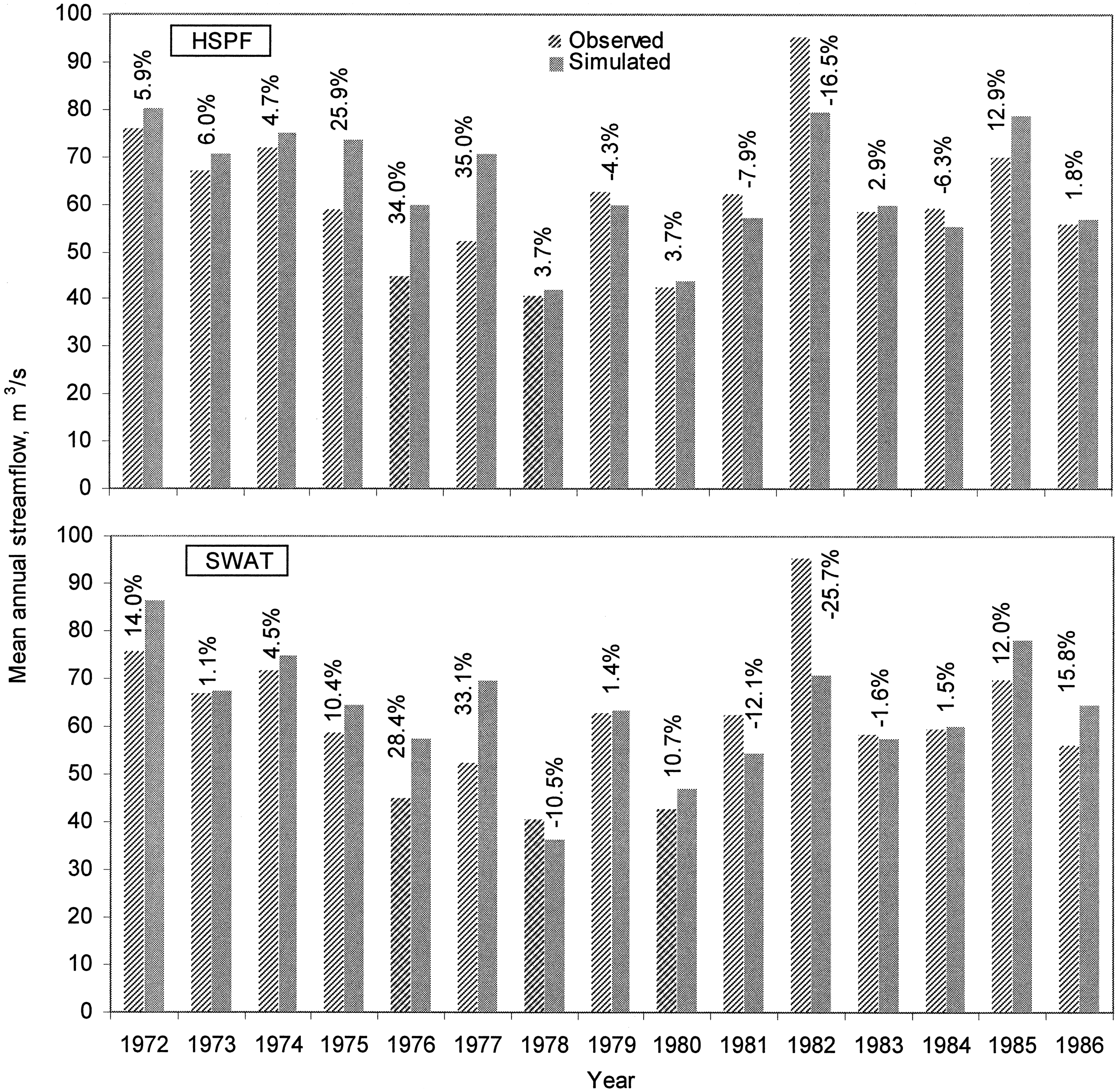

Figure 5. Observed and Simulated Mean Annual Streamflows and Their Percent Errors for HSPF and SWAT for the Entire Model Verification Period (1972 to 1986).

high P_e value of 0.99 for both (Table 3). These results show that both models were calibrated equally well to satisfactorily simulate the streamflow for the IRW.

Comparative Model Performance During the 1972 Through 1986 Verification Period

Comparison of average annual flows for the entire 15-year model verification period revealed that both models estimated the streamflow to within 15 percent in 11 years (Figure 5). Also, both models underestimated or overestimated flows in the same year in 11 of the 15 years. Similar to the calibration period, HSPF was biased in overestimating the mean monthly flows less than 25 m^3/s, while SWAT was less biased (Figure 6). Larger flows were equally well estimated by both models, since the data points were evenly scattered around S/O equal to 1 line (Figure 6). The simulated monthly flows by HSPF and SWAT were within 15 percent of the observed flows, respectively, in 40 and 48 of the 180 months during the verification period. Both models results had a similar high monthly value for R (0.82 for HSPF and 0.84 for SWAT; Table 3). The plot of observed daily flows versus S/O ratio showed a wider scatter of data points

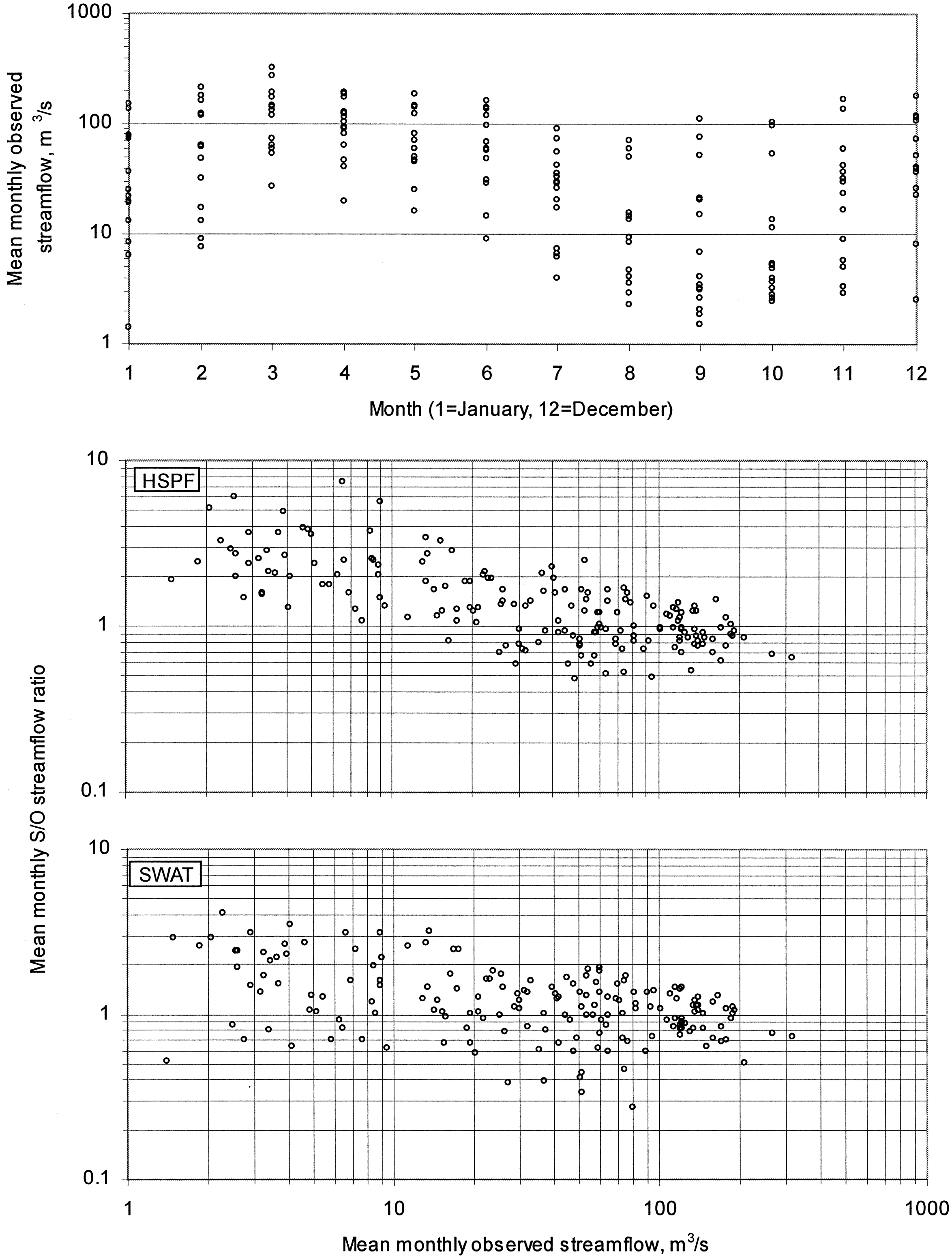

Figure 6. Scatter Plot of Months of a Year Versus Mean Monthly Observed Streamflows During Those Months (top) and Mean Monthly Observed Streamflow Versus Mean Monthly Simulated/Observed Streamflow Ratio for the Entire Model Verification Period (1972 to 1986) for HSPF and SWAT.

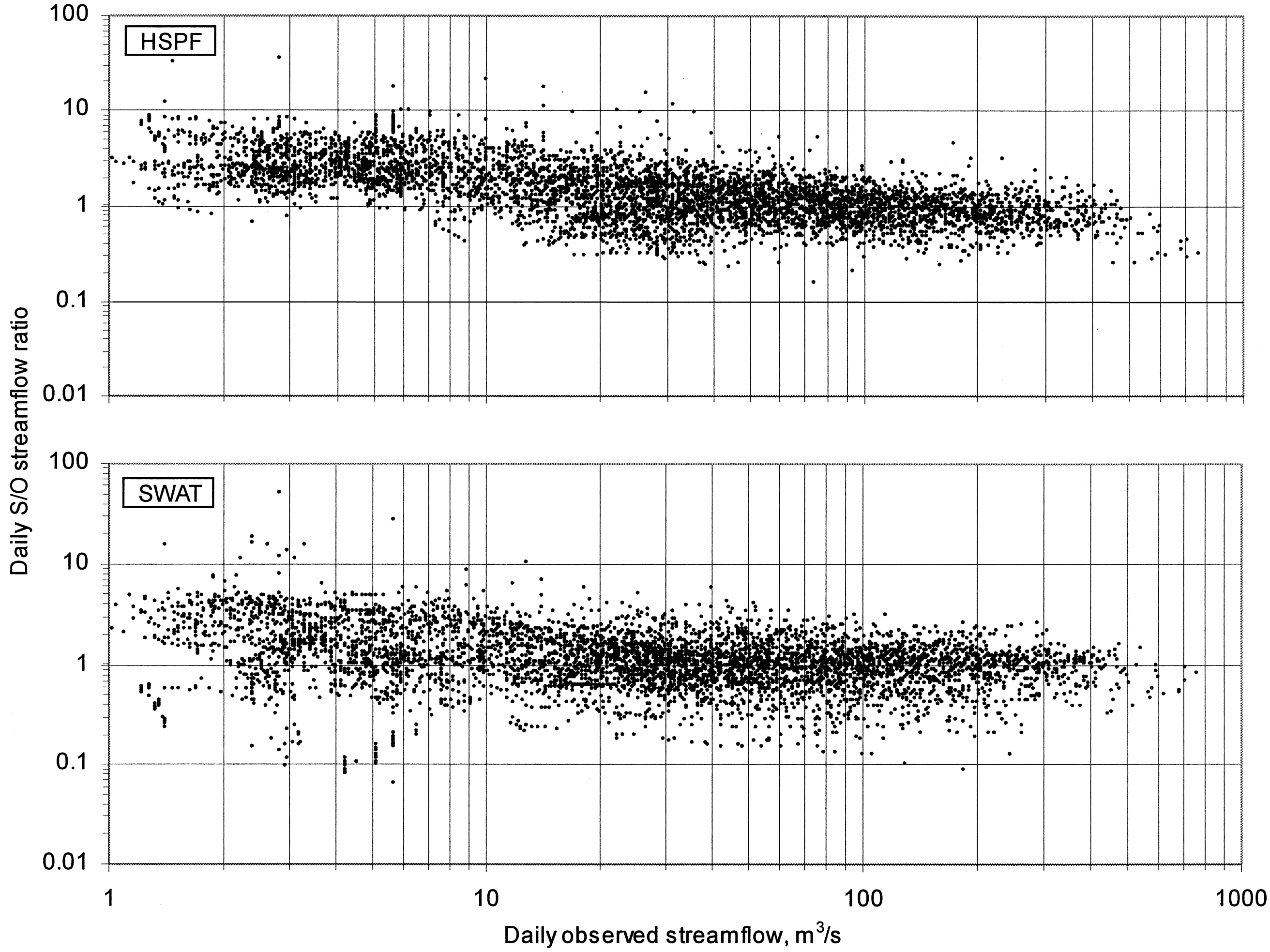

Figure 7. Scatter Plot of Daily Observed Streamflow Versus Daily Simulated/Observed Streamflow Ratio for the Entire Model Verification Period (1972 to 1986) for HSPF and SWAT.

around S/O equal to 1 line for the SWAT than for HSPF (Figure 7). HSPF was biased in overestimating daily flows less than 12 m^3/s (Figure 7), while SWAT was less biased but showed a tendency to underestimate some flows smaller than 6 m^3/s. The daily R value was 0.69 and 0.74 for HSPF and SWAT, respectively (Table 3). The overestimation of low flows by HSPF is also evident in the daily flow duration curves (Figure 8). A high P_e value of 0.99 was shown by both models. These results show that both models performed satisfactorily; however, the overall performance of SWAT is slightly superior to that of HSPF for this 15-year verification period.

Comparative Model Performance During Dry, Average, and Wet Climatic Conditions

The model results for the three climatic periods representing dry, average, and wet conditions are presented in Table 3. For the drier than average condition, the monthly and daily R values obtained from SWAT (0.93 and 0.83) were noticeably higher compared to HSPF (0.87 and 0.69). Compared to SWAT, HSPF was able to simulate more monthly flow values (8 of 36 total months) to within 15 percent of the observed flows, but fewer months (10 of 36) to within 25 percent. The satisfactory agreement between the shapes of the observed and simulated streamflow hydrographs for this three-year period is shown in Figure 9. For the average climatic conditions, both HSPF and SWAT underestimated the streamflow by 3.2 and 3.7 percent, respectively, and both showed a P_e value of 0.99. However, the monthly and daily R values were slightly better for the SWAT (0.81 and 0.72, respectively) compared to HSPF (0.80 and 0.69), and SWAT also estimated monthly streamflows to within 15 percent in 19 of 72 months, compared to 14 months by HSPF. For the wetter-than-average period, HSPF and SWAT overestimated the streamflow by

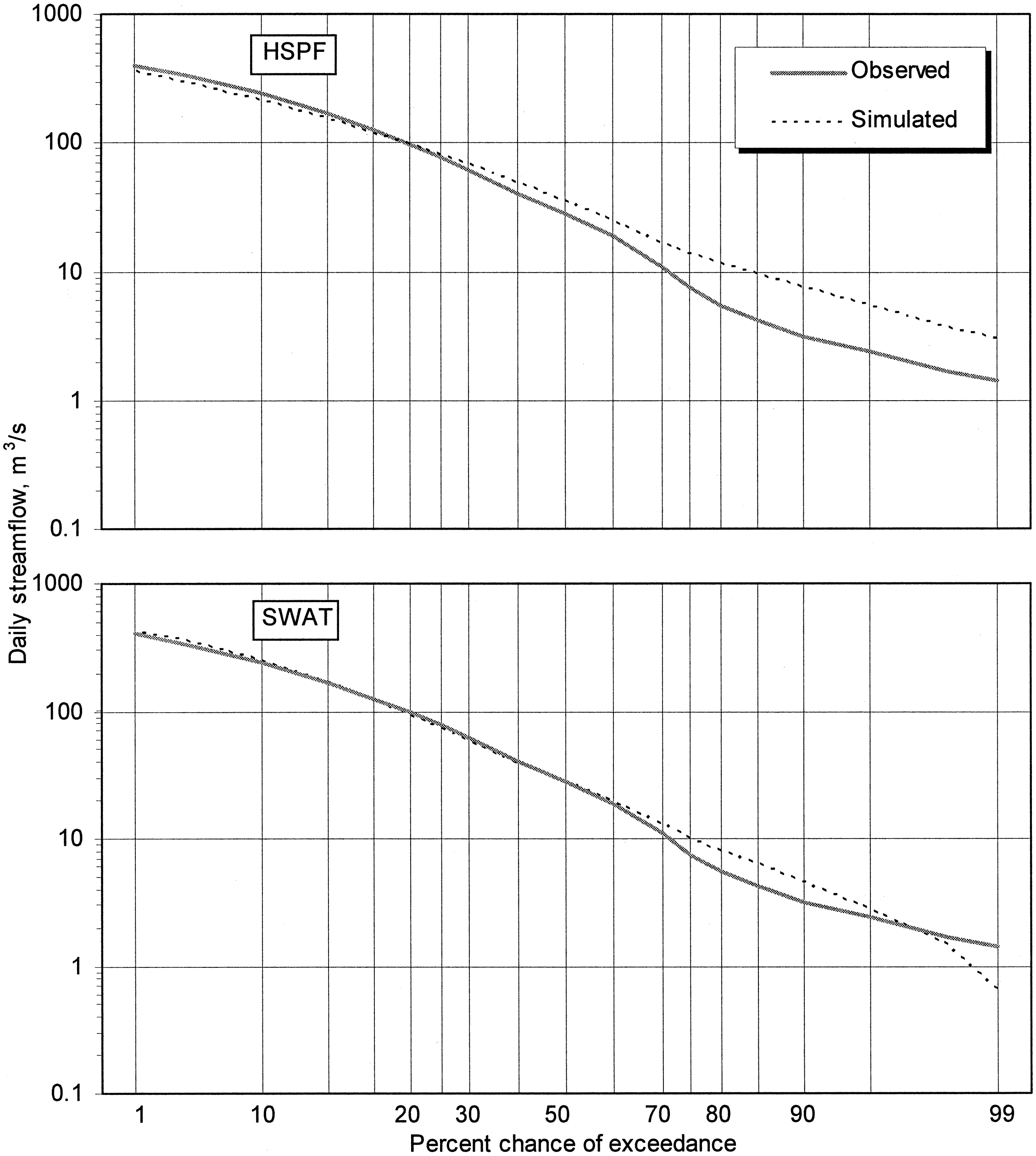

Figure 8. Observed and Simulated Daily Streamflow Duration Curves for the Entire Model Verification Period (1972 to 1986) for HSPF and SWAT.

16.4 and 13.7 percent, respectively. The monthly and daily R values for the HSPF (0.81 and 0.71, respectively) were slightly better than those for SWAT (0.80 and 0.70); however, SWAT showed a slightly higher daily value of P_e. SWAT also estimated monthly streamflows to within 15 percent in 23 of 72 months, compared to 18 months by HSPF. These overall results indicate that both models performed satisfactorily for all climate conditions, but the SWAT model results were generally better than HSPF for the drier than average period.

DISCUSSION

One of the substantial differences between the two model applications relates to the estimate of PET in which the PET input to HSPF had an average annual value of 696 mm compared to the 1,306 mm estimated by SWAT (Table 2). The PET values calculated by SWAT used the Penman-Monteith method (Monteith, 1965), and the PET input included in the BASINS 3.0 database for use by HSPF used pan evaporation data and a multiplication factor. Potential ET is not

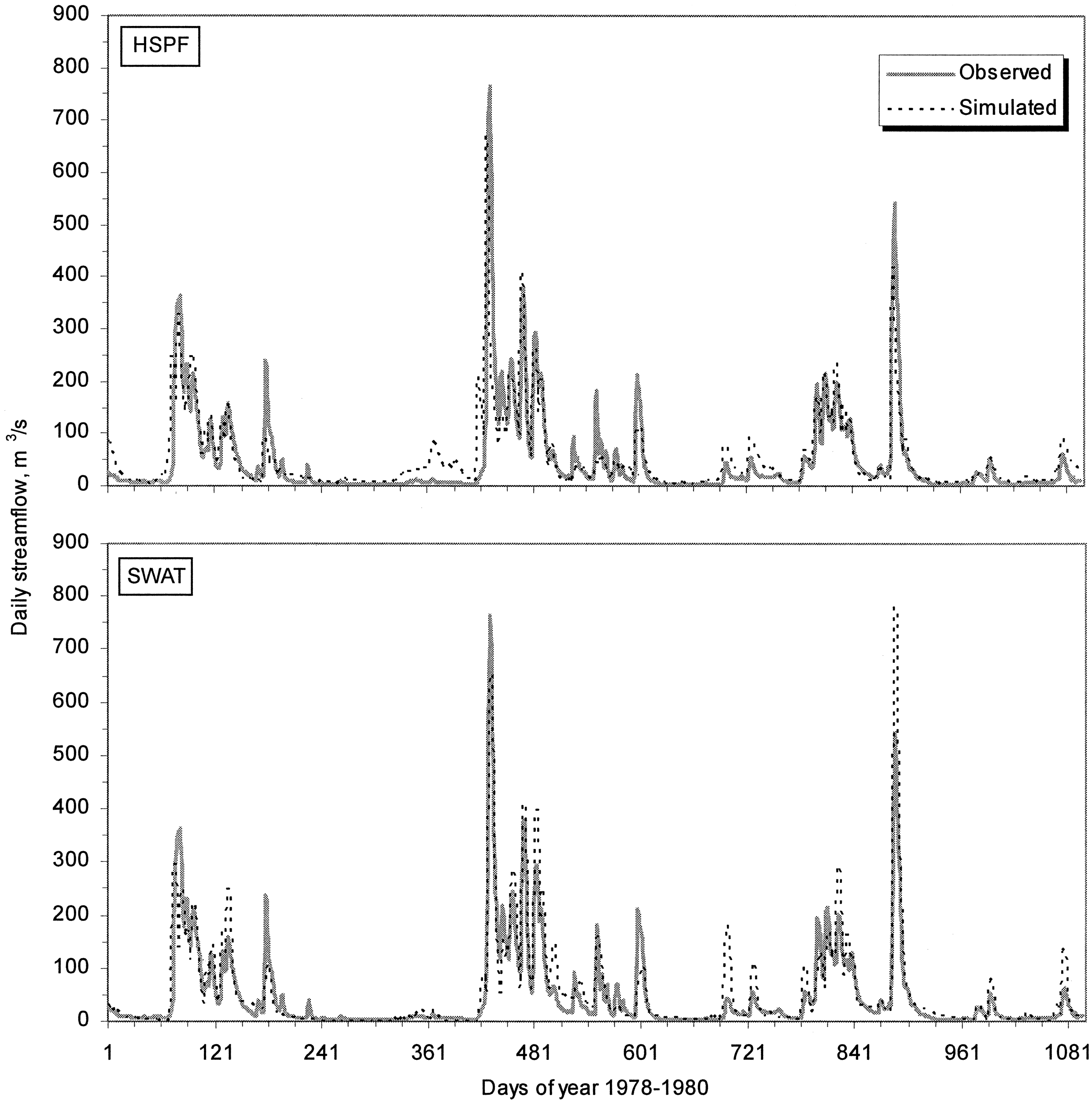

Figure 9. Observed Versus HSPF and SWAT Simulated Daily Streamflows for the Drier Than Average Period (1978 to 1980).

directly measurable, and estimates can vary depending on methodology; however, the differential given here between the two methods seems to be particularly great. The model calibration process allows parameters in each model to adjust to the PET amount, resulting in estimates of actual evapotranspiration that are comparatively similar given the wide differential in PET. However, the actual ET values estimated by SWAT were consistently higher than those from HSPF except for the year 1995 (Table 2).

HSPF showed a consistent bias in overestimating mean monthly streamflows less than 25 m^3/s (Figures 2 and 6) and daily streamflows less than 12 m^3/s (Figures 3 and 7). These lower flows mostly occurred from July to November, as shown in Figures 2 and 6 for monthly flows and Figure 3 for daily flows. It is possible that a comparatively low value of the actual ET used by HSPF resulted in the overestimation in the streamflow, particularly for these low flow conditions that typically occur in the middle to later part of the growing season, when evapotranspiration and its influence on soil moisture are greatest.

SWAT is capable of performing hourly or subdaily simulations. However, a daily time step was used in this study and in previous model comparison studies cited earlier. The difference in time steps between hourly HPSF and daily SWAT simulations and its effect on runoff production and other processes may have greater influence on model comparison studies on smaller watersheds. But in the case of a larger watershed application such as that presented here, the differences between hourly and daily simulations are minimized. SWAT may also be well suited to the IRW because of the extensive tile drainage and predominance of agriculture in the watershed. HSPF parameters may be adjusted to represent the faster response of tile drains to discharge the subsurface flow (Duncker and Melching, 1998), but there is no specific tile drainage component in HSPF as there is in SWAT.

Ideally, statistical and graphical measures of model performance during verification should be similar to those displayed during calibration. Results of the verification process provide a better measure than calibration of the model's ability to accurately predict flows. In this study, the SWAT results were more consistent from calibration to verification as compared with HSPF results, for which there is a sizable reduction in R between calibration and verification (Table 3) along with greater separation between observed and simulated low flows (Figures 3, 4, 7, and 8).

As is also observed in other comparison studies (Im *et al.,* 2003; Van Liew *et al.,* 2003; Saleh and Du, 2004), HSPF requires comparatively more effort to apply and thus is considered less user friendly. The calibration parameters in HSPF were numerous and therefore required many more iterative calibration runs in comparison to SWAT, which used fewer parameters in calibration. Data preparation for the HSPF model also required a greater effort as compared to that for SWAT.

SUMMARY AND CONCLUSIONS

The performances of the HSPF and SWAT models were evaluated for simulating the hydrology of the 5,568 km^2 Iroquois River watershed, a predominantly tile drained, agricultural watershed in Illinois and Indiana. The preparation of both model applications was performed within the BASINS 3.0 modeling system, and all data were provided by the BASINS database except additional climate data that were used by both models. Each model was calibrated for a nine-year period (1987 through 1995) and verified using an independent 15-year period (1972 through 1986) by comparing the simulated and observed daily, monthly, and annual streamflows. Verification results were further compared for three shorter periods representing drier than average, average, and wetter than average climatic conditions.

The HSPF results showed a consistent bias toward overestimating low flows, those being roughly the lowest 25 percent of the observed monthly and daily flows. This bias was particularly apparent within the model verification period. Based on the statistical and graphical comparisons, the HSPF and SWAT models performed similarly during model calibration, but SWAT predicted low flow conditions noticeably better during the model verification period. There was a significant difference between the two models in the estimation of PET. The HSPF estimate of PET provided by the BASINS database was roughly half of the estimate computed within SWAT. Part of the difference in low flow simulation between the two models may be attributed to the difference in PET and its effect on the simulation of actual evapotranspiration and soil moisture during the latter portion of the growing season when low flows typically occur. However, despite these and other differences in the model simulations, the results indicate that the characteristics of simulated flows from both models are good and generally similar to each other and to the observed flows.

As noted in other studies, HSPF requires comparatively more effort to apply than SWAT. The preparation of climate data for the HSPF model required a greater effort as compared to that for SWAT, and the calibration of HSPF required many more iterations because of the larger number of parameters used in calibration. The comparative effort in model application may be a deciding factor in model selection for situations in which there is little expected difference in the accuracy of streamflow simulations.

Various factors, including the scale of the watershed and the occurrence of subsurface tile drains, may have contributed to the somewhat better performance of SWAT for the Iroquois River watershed. For this study there was a drop off in the performance of HSPF compared to SWAT when moving from the calibration to verification periods, and additional comparative studies are needed to determine if this is a systematic characteristic. No one model is best under all conditions, and a more complete understanding of model performance will require continued comparative studies under differing hydrologic conditions and watershed scales.

LITERATURE CITED

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 (JAWRA) 35(2):411-424.

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

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

Bergman, M.J. and L.J. Donnangelo, 2000. Simulation of Freshwater Discharges From Ungaged Areas to the Sebestian River, Florida. Journal of the American Water Resources Association (JAWRA) 36(5):1121-1132.

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, Jr., T.H. Jobes, and A.S. Donigian, Jr., 2001. Hydrological Simulation Program – FORTRAN (HSPF), User's Manual for Version 12.0. USEPA, Athens, Georgia.

Borah, D.K. and M. Bera, 2003. Watershed-Scale Hydrologic and Nonpoint-Source Pollution Models: Review of Mathematical Bases. Trans. ASAE 46(6):1553-1566.

Chew, C.Y., L.W. Moore, and R.H. Smith, 1991. Hydrologic Simulation of Tennessee's North Reelfoot Creek Watershed. Res. J. Water Pollution Control Federation 63:10-16.

Deliman, P.N., R.H. Glick, and C.E. Ruiz, 1999. Review of Watershed Water Quality Models. Technical Report W-99-1, U.S. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, Mississippi.

Di Luzio, M., R. Srinivasan, J.G. Arnold, and S. Neitsch, 2002. ARCVIEW Interface for SWAT 2000 User's Guide. Blackland Research Center, Texas Agricultural Experiment Station, Temple, Texas. *Available at* http://www.brc.tamus.edu/swat/downloads/doc/swatav2000.pdf. *Accessed in* March 2003.

Donigian, A.S., J.C. Imhoff, and B.R. Bicknell, 1983. Predicting Water Quality Resulting From Agricultural Nonpoint Source Pollution Via Simulation – HSPF. *In:* Agricultural Management and Water Quality, F.W. Schaller and G.W. Bailey (Editors). Iowa State University Press, Ames, Iowa, pp. 200-249.

Duncker, J.J. and C.S. Melching, 1998. Regional Rainfall-Runoff Relations for Simulations of Streamflow for Watersheds in Du Page County, Illinois. USGS, WRIR 98-4035. Urbana, Illinois.

Franchini, M. and M. Pacciani, 1991. Comparative Analysis of Several Conceptual Rainfall Runoff Models. J. Hydrol. 122:161-219.

Green, W.H. and G.A. Ampt, 1911. Studies on Soil Physics, 1. The Flow of Air and Water Through Soils. J. Agric. Sciences 4:11-24.

Im, S., K. Brannan, S. Mostaghimi, and J. Cho, 2003. A Comparison of SWAT and HSPF Models for Simulating Hydrologic and Water Quality Responses From an Urbanizing Watershed. 2003 ASAE Annual International Meeting, American Society of Agricultural Engineers, Las Vegas, Nevada, ASAE Paper No. 03-2175. St. Joseph, Michigan.

James, L.D., 1972. Hydrologic Modeling, Parameter Estimation, and Watershed Characteristics. J. Hydrol. 17:283-307.

Jones, P.M. and T.A. Winterstein, 2000. Characterization of Rainfall-Runoff Response and Estimation of the Effect of Wetland Restoration on Runoff, Herone Lake Basin, Southwestern Minnesota, 1991-97. USGS, WRIR 00-4095. Mounds View, Minnesota.

Knapp, H.V., 1992. Kankakee River Basin Streamflow Assessment Model: Hydrologic Analysis. Illinois State Water Survey Contract Report 541, Champaign, Illinois.

Laroche, A-M., J. Gallichand, R. Lagace, and A Pesant, 1996. Simulating Atrazine Transport With HSPF in an Agricultural Watershed. J. Environmental Eng. 122(7):622-630.

Mades, D.M. and K.A. Oberg, 1986. Evaluation of the U.S. Geological Survey's Gaging-Station Network in Illinois. U.S. Geological Survey Water-Resources Investigation Report 86-4072. Urbana, Illinois.

Monteith, J.L., 1965. Evaporation and the Environment. *In:* The State and Movement of Water in Living Organisms. XIXth Symposium of the Society for Experimental Biology, Cambridge University Press. Swansea, United Kingdom, pp. 205-234.

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

Nasr, A., M. Bruen, G. Parkin, S. Birkinshaw, R. Moles, and P. Byrne, 2003. Modelling Phosphorous Loss from Agricultural Catchments: A Comparison of the Performance of SWAT, HSPF, and SHETRAN for the Clarianna Catchment. *In:* Diffuse Pollution and Basin Management, M. Buren (Editor). Proc., 7th International Specialised IWA Conference, Dublin, Ireland, pp. 3-181 to 3-185.

Neitsch, S.L., J.G. Arnold, J.R. Kiniry, R. Srinivasan, and J.R. Williams, 2002a. Soil and Water Assessment Tool User's Manual: Version 2000. GSWRL Report 02-02, BRC Report 02-06, Publ. Texas Water Resources Institute, TR-192, College Station, Texas, 438 pp. *Available at* http://www.brc.tamus.edu/swat/downloads/doc/swatuserman.pdf. *Accessed in* March 2003.

Neitsch, S.L., J.G. Arnold, and R. Srinivasan, 2002b. Pesticide Fate and Transport Predicted by the Soil and Water Assessment Tool (SWAT): Atrazine, Metachlor and Trifluralin in the Sugar Creek Watershed. BRC Publ. No. 2002-03, Blackland Research Center,Texas Agric. Experiment Station, Temple, Texas.

Page, L.M., K.S. Cummings, S.L. Post, and M.E. Retzer, 1992. Biologically Significant Illinois Streams: An Evaluation of the Streams of Illinois Based on Aquatic Biodiversity. Technical Report 1992(1), Illinois Natural History Survey, Center for Biodiversity, Champaign, Illinois.

Philip, J.R., 1957. The Theory of Infiltration: 1. The Infiltration Equation and Its Solution. Soil Science 83:345-357.

Saleh, A. and B. Du., 2004. Evaluation of SWAT and HSPF Within BASINS Program for the Upper North Bosque River Watershed in Central Texas. Trans. ASAE 47(4):1039-1049.

Santhi, C., J.G. Arnold, J.R. Williams, W.A. Dugas, R. Srinivasan, and L.M. Hauck, 2001. Validation of the SWAT Model on a Large River Basin With Point and Nonpoint Sources. JAWRA 37(5):1169-1188.

Sogbedgi, J.M. and G.F. McIsaac, 2002. Modeling Streamflow From Artificially Drained Agricultural Watersheds in Illinois. Journal of the American Water Resources Association (JAWRA) 38(6):1753-1765.

USDA-SCS (U.S. Department of Agriculture-Soil Conservation Service), 1972. Hydrology, Section 4, Chapters 4-10. *In:* National Engineering Handbook. USDA Soil Conservation Service, Washington D.C.

USEPA (U.S. Environmental Protection Agency), 2000. BASINS Technical Note 6: Estimating Hydrology and Hydraulic Parameters for HSPF. USEPA, Office of Water, EPA-823-R-99-013, Washington, D.C.

USEPA (U.S. Environmental Protection Agency), 2001. Better Assessment Science Integrating Point and Nonpoint Sources – BASINS Version 3.0, User's Manual. EPA-823-B-01-001, Washington, D.C.

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 (JAWRA) 39(2):413-426.

Van Liew, M.W., J.G. Arnold, and J.D. Garbrecht, 2003. Hydrologic Simulation on Agricultural Watersheds: Choosing Between Two Models. Trans. ASAE 46(6):1539-1551.