BASINS/HSPF: MODEL USE, CALIBRATION, AND VALIDATION

P. B. Duda, P. R. Hummel, A. S. Donigian Jr., J. C. Imhoff

ABSTRACT. *This article presents recommendations by model developers and the authors about calibration and validation procedures for the Hydrological Simulation Program - Fortran (HSPF) as applied through BASINS. HSPF is a continuous simulation watershed model that simulates nonpoint-source runoff and pollutant loadings for a watershed and performs flow and water quality routing in stream reaches and well-mixed lakes and impoundments. HSPF can be used to estimate nonpoint-source loads from various land uses as well as fate and transport processes in streams and lakes. This article describes the ideal calibration and validation process for the full range of constituents modeled by HSPF, as well as the process for acceptable minimum calibration and validation of this model. The model information and guidance provided in this article may be used to help in determining the scope of the proposed ASABE Standard/Engineering Practice for model calibration and validation. Model calibration and validation are necessary and critical steps in any model application. For HSPF and most other watershed models, calibration is an iterative procedure of parameter evaluation and refinement, as a result of comparing simulated and observed values of interest. Model validation is in reality an extension of the calibration process. Its purpose is to ensure that the calibrated model properly assesses all the variables and conditions that can affect model results, and to demonstrate the ability to predict field observations for periods separate from the calibration effort. For HSPF calibration and validation, a "weight of evidence" approach is most widely used in practice when models are examined and judged for acceptance for assessment and regulatory purposes. This article explores the "weight of evidence" approach and the current practice of watershed model calibration and validation based on more than 30 years of experience with HSPF. Example applications are described and model results are shown to demonstrate the graphical and statistical procedures used to assess model performance. In addition, quantitative criteria for various statistical measures are discussed as a basis for evaluating model results and documenting the model application efforts.*

Keywords. *BASINS, Error analysis, HSPF, Model calibration, Model performance assessment, Model validation, Watershed models.*

The Hydrological Simulation Program—Fortran (HSPF) (Bicknell et al., 2005) is a continuous simulation time-step watershed model that simulates nonpoint-source runoff and pollutant loadings from upland areas within a watershed and routes flow and pollutant loadings in stream reaches and well-mixed lakes and impoundments. HSPF can be used to estimate nonpoint-source loads from various land uses as well as fate and transport processes in streams and lakes. HSPF can simulate any period from a few minutes to hundreds of years using a time step ranging from sub-hourly to daily. Typically, the model is run for a time span ranging from 5 to 20 years or more using an hourly time step.

The origins of HSPF can be traced back to the early 1960s and the Stanford Watershed Model (SWM) (Crawford, 1962; Crawford and Linsley, 1966), a tool that was instrumental in introducing the civil engineering profession to the concept of continuous hydrologic modeling. By the early 1970s, the developers of SWM expanded and refined SWM to create the Hydrocomp Simulation Program (HSP) (Hydrocomp, 1976, 1977), which also included general nonpoint-source loadings and water quality simulation capabilities. During the early 1970s, the U.S. Environmental Protection Agency (USEPA) sponsored the development of the Agricultural Runoff Management (ARM) (Donigian and Crawford, 1976b; Donigian et al., 1977) and the nonpoint-source (NPS) (Donigian and Crawford, 1976a) pollutant loading models to address pollution from agriculture, urban, and other land uses; the SWM approach was selected as the hydrologic foundation for an expanding suite of models of nonpoint-source pollution impacts (Donigian and Imhoff, 2002).

With wide distribution and application of the SWM in the late 1960s, civil engineers recognized the value of digital continuous simulation for hydrologic applications. By the early 1970s, Hydrocomp had demonstrated the utility of quantity/quality simulation by modeling a range of water quality constituents in a large basin in Washington. In the late 1970s, the USEPA recognized the need for a continuous process simulation approach to analyze and solve many complex water resource problems. Funding from the agen-

---

Submitted for review in January 2011 as manuscript number SW 9590; approved for publication by the Soil & Water Division of ASABE in June 2012.

The authors are **Paul B. Duda,** Senior Project Engineer, and **Paul R. Hummel, ASABE Member,** Vice-President, AQUA TERRA Consultants, Decatur, Georgia; **Anthony S. Donigian Jr.,** President and Principal Engineer, AQUA TERRA Consultants, Mountain View, California; and **John C. Imhoff,** Vice-President, AQUA TERRA Consultants, Ouray, Colorado. **Corresponding author:** Anthony S. Donigian Jr., AQUA TERRA Consultants, 2685 Marine Way, Suite 1314, Mountain View, CA 94043-1115; phone: 650-962-1864; e-mail: donigian@aquaterra.com.

cy to Hydrocomp resulted in the development of HSPF, a nonproprietary system of simulation modules in standard Fortran that handled essentially all the functions performed by HSP, ARM, and NPS and was considerably easier to maintain and modify. HSPF simulates the hydrologic and associated water quality processes on pervious and impervious land surfaces and in streams and well-mixed impoundments. Since the first public release (Release 5) of HSPF in 1980, the model has undergone a continual series of code and algorithm enhancements, producing a succession of new releases, leading up to the most recent Release 12 in 2001 (and Release 12.2 in 2005).

Since 1981, the U.S. Geological Survey has been developing software tools to facilitate watershed modeling by providing interactive capabilities for model input development, data storage and analysis, and model output analysis including hydrologic calibration assistance. The ANNIE (Flynn et al., 1995), HSPEXP (Lumb et al., 1994), and GenScn (Kittle et al., 1998) products developed by the USGS have greatly advanced and facilitated watershed model application, not only for HSPF, but also for many other USGS models.

HSPF is currently released as a core watershed model of the Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) environmental analysis system (USEPA, 2007; Duda et al., 2006), developed by the USEPA Office of Water. BASINS is a multipurpose environmental analysis system designed for use by regional, state, and local agencies performing watershed and water quality-based studies. It was developed to facilitate examination of environmental information, to support analysis of environmental systems, and to provide a framework for examining management alternatives. BASINS integrates environmental data, analytical tools, and modeling programs within a Geographic Information System (GIS) environment to support development of solutions to watershed management problems and environmental protection issues, including development of total maximum daily loads (TMDLs).

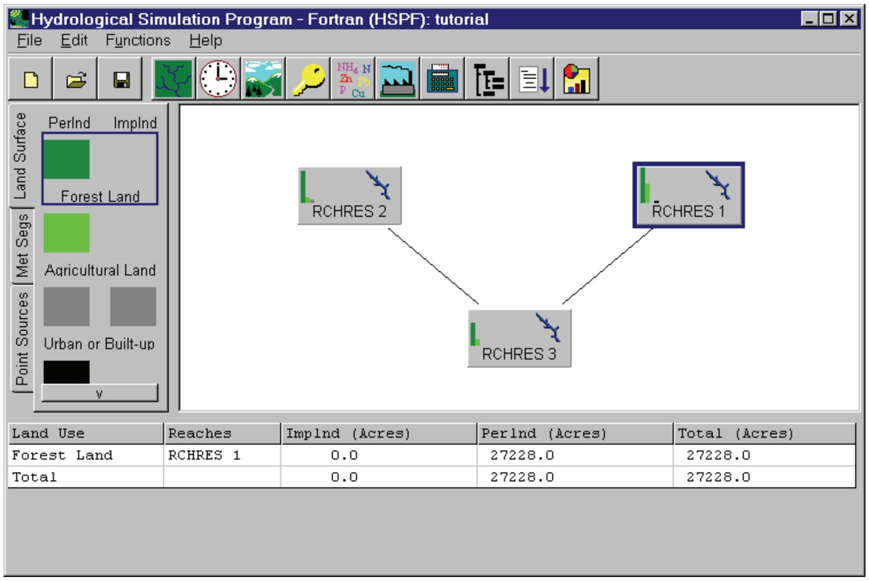
The Windows interface to HSPF, known as WinHSPF (Duda et al., 2001), was created for BASINS and works with the USEPA-supported HSPF model (fig. 1). WinHSPF supports the full suite of the HSPF model capabilities. BASINS contains an extension that allows the user to open WinHSPF directly from the BASINS user interface, extracting appropriate information for the preparation of HSPF input files.

While HSPF is fully integrated into BASINS through the WinHSPF interface, the code base of HSPF is maintained separately. This separation is accomplished by compiling the HSPF model as a dynamic link library (DLL), called by WinHSPF for running a simulation. Maintaining HSPF as a separate DLL means that it can be enhanced independently of WinHSPF and BASINS. A revised DLL can be copied into place on the user's computer, and the user will have access to the latest HSPF features. Input meteorologic data are provided to HSPF through the use of watershed data management (WDM) files, available through the BASINS Data Download tool.

HSPF DESCRIPTION

HSPF is a process-based, continuous simulation watershed model for quantifying runoff and addressing water quality impairments associated with combined point and nonpoint sources. HSPF simulates nonpoint-source runoff and pollutant loadings for a watershed and performs flow and water quality routing in stream reaches and well-mixed lakes and impoundments. HSPF can be used to estimate nonpoint-source loads from various land uses as well as fate and transport processes.

HSPF contains hundreds of process algorithms developed from theory, laboratory experiments, and empirical relations from instrumented watersheds. For pervious land segments, HSPF can compute air temperature as a function of elevation; snow accumulation and melting; hydrological cycle components (evapotranspiration, surface detention,



Land Use	Reaches	Implnd (Acres)	Perlnd (Acres)	Total (Acres)
Forest Land	RCHRES 1	0.0	27228.0	27228.0
Total		0.0	27228.0	27228.0

Figure 1. The WinHSPF interface available through BASINS.

surface runoff, infiltration, interflow, base flow, percolation to deep groundwater); sediment production and removal; soil temperature; surface water temperature and dissolved oxygen and carbon dioxide concentrations in overland flow; generalized water quality constituents modeled as accumulated storages removed by flow or potency factors associated with sediment; more detailed modeling of pesticide processes (runoff, leaching, adsorption/desorption, degradation), nutrient processes (transport by flow and sediment association, leaching, adsorption/desorption, denitrification, nitrification, plant uptake, immobilization, mineralization); and tracer elements.

For impervious land segments, HSPF can compute air temperature as a function of elevation; snow accumulation and melting; water budget (surface components only); solids accumulation and removal including methods that are independent of storm events; surface water temperature and gas concentrations; and generalized water quality constituents.

For channel segments, HSPF can compute hydraulic behavior using the kinematic wave assumption; longitudinal advection of dissolved and entrained constituents; water temperature using a heat balance approach (absorption of shortwave radiation, longwave radiation, emission of longwave radiation, conduction-convection and evaporation); inorganic sediment deposition, scour and transport by particle size; partitioning, hydrolysis, volatilization, oxidation, biodegradation, first-order decay, and parent chemical/metabolite transformations for generalized chemicals; dissolved oxygen and BOD processes (decay, settling, benthal sinks and sources, re-aeration, sinks and sources related to plankton metabolism); nitrogen processes (ammonia volatilization, ammonification, denitrification, ammonia adsorption/desorption with suspended sediment); phosphate adsorption/desorption with suspended sediment; phytoplankton processes (growth, respiration, sinking, zooplankton predation, death); zooplankton process (growth, respiration, death); benthic algae processes (growth, respiration, death); and carbon dioxide-bicarbonate system processes (carbon dioxide invasion, zooplankton respiration, BOD decay, net growth of algae, and benthal releases) that determine pH.

The model consists of a set of modules arranged in an organized structure, which permits the continuous simulation of a comprehensive range of hydrologic and water quality processes. HSPF's design incorporates a hierarchy of program subroutines, each of which performs a major task during the program's execution. The subroutines are grouped into different levels of operations in a hierarchical structure. The importance of this program structure lies in its modular design. This allows the addition or replacement of individual modules and allows HSPF to be more readily adapted to special applications designed by the user.

HSPF has been applied to watersheds ranging in size from the Chesapeake Bay, with roughly 160,000 km^2 (62,000 mi^2) of tributary area, down to a few acres/hectares. HSPF can simulate any period from a few minutes to hundreds of years. Typically, the model is run for a time span ranging from 5 to 20 years or more using an hourly time step.

HSPF is designed in a way that it can be readily applied to most watersheds in the world using existing meteorologic and hydrologic data; soils and topographic information; and land use, drainage, and system (physical and manmade) characteristics. Typical input time-series records include precipitation, potential evapotranspiration (and other meteorologic data), waste discharges, and calibration data such as streamflow and constituent concentrations. Physical measurements and related parameters are required to describe the land area, channels, and reservoirs. The model user is empowered to choose and represent the desired level of physical detail: HSPF applications have been developed to concurrently represent as many as 20 different land use types as well as hundreds of distinct subbasins and channel reaches.

The result of an HSPF simulation is a time history of the quantity and quality of water transported over the land surface and through various soil zones down to the groundwater aquifers. Runoff flow rate, sediment loads, nutrients, pesticides, toxic chemicals, and other quality constituent concentrations can be predicted. The model uses these results and stream channel information to simulate in-stream processes. Using the simulation output, HSPF can produce a time history of water quantity and quality at any point in the watershed channel network.

The integrity of HSPF is ensured by careful attention to version control and model maintenance. Software maintenance of HSPF, almost all of which has been performed by AQUA TERRA Consultants, has included correcting errors, implementing enhancements, adapting the code to new computer environments (hardware and operating system), testing, and providing new versions to EPA and USGS for distribution to users. At the same time, a continual flow of academic contributions have ensured that HSPF maintains a strong scientific basis.

HSPF CALIBRATION AND VALIDATION

Calibration and validation have been defined by the American Society of Testing and Materials as follows (ASTM, 1984):

- Calibration is a test of the model with known input and output information that is used to adjust or estimate factors for which data are not available.
- Validation is a comparison of model results with numerical data independently derived from experiments or observations of the environment.

Application of HSPF for predicting flow, sediment, and chemical loadings can be described as comprised of three phases, as shown in figure 2 (Donigian and Rao, 1990). Phase I includes data collection, model input preparation, and parameter evaluation, i.e., all the steps needed to set up a model, characterize the watershed, and prepare for model executions. Phase II is the model testing phase, which involves calibration, validation (or verification, as it is sometimes called), and, when possible, post-audit. This is the phase in which the HSPF model is evaluated to assess whether it can reasonably represent the watershed behavior, for the purposes of the study. Phase III includes the ultimate use of the model, as a decision support tool for man-

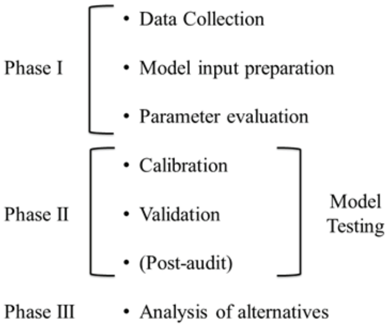


Figure 2. The modeling process.

agement and regulatory purposes.

Although specific application procedures for all watershed models differ due to the variations of the specific physical, chemical, and biological systems that they each attempt to represent, they have many steps in common. The calibration and validation phase is especially critical since the outcome establishes how well the model represents the watershed for the purpose of the study. Thus, this is the "bottom line" of the model application effort, as it determines if the model results can be relied upon and used effectively for decision-making.

Model validation is in reality an extension of the calibration process. Its purpose is to ensure that the calibrated model properly assesses all the variables and conditions that can affect model results. While there are several approaches to validating a model, perhaps the most effective procedure is to use only a portion of the available record of observed values for calibration; once the final parameter values are developed through calibration, simulation is performed for the remaining period of observed values, and goodness-of-fit between the recorded and simulated values is reassessed. This type of split-sample calibration/validation procedure is commonly used, and recommended, for many watershed modeling studies. Model credibility is based on the ability of a single set of parameters to represent the entire range of observed data. If a single parameter set can reasonably represent a wide range of events, then this is a form of validation.

In practice, the model calibration/validation process can be viewed as a systematic analysis of errors or differences between model predictions and field observations. Figure 3 schematically compares the model with the natural system, i.e., the watershed, and identifies various sources of potential errors to be investigated. These types of analyses require evaluation of the accuracy and validity of the model input data, parameter values, model algorithms, calibration accuracy, and observed field data used in the calibration/validation. Clearly, the model user must become a detective, searching for the causes of the errors or differences, and for potential remedies to improve the agreement and reduce the errors. A more complete discussion of these error sources is provided by Donigian and Rao (1990).

Watershed model performance, i.e., the ability to reproduce field observations, and calibration/validation are most often evaluated through both qualitative and quantitative measures, involving both graphical comparisons and statistical tests. For flow simulations where continuous records are available, all these techniques are employed, and the same comparisons will be performed, during both the calibration and validation phases. Comparisons of simulated and observed state variables will be performed for daily, monthly, and annual values, in addition to flow-frequency duration assessments. Statistical procedures include error statistics, correlation and model-fit efficiency coefficients, and goodness-of-fit tests.

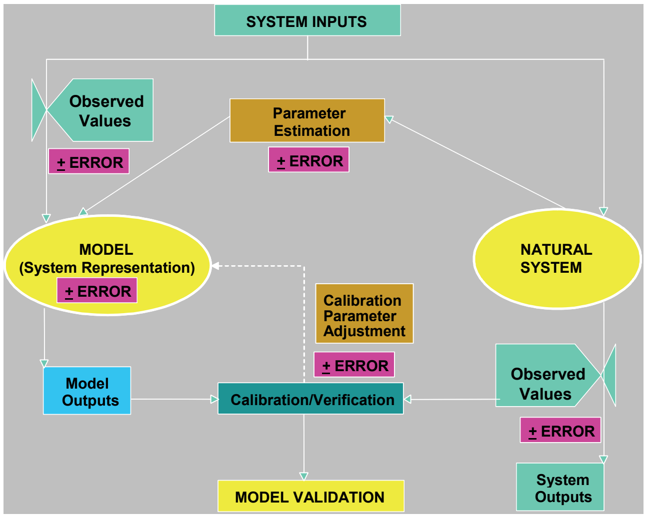


Figure 3. Model versus natural system: Inputs, outputs, and errors.

For sediment, water quality, and biotic constituents, model performance will be based primarily on visual and graphical presentations, as the frequency of observed data is often inadequate for accurate statistical measures. However, alternative model performance assessment techniques for water quality, e.g., error statistics and correlation measures, consistent with the population of observed data available are often used for model calibration and testing.

HSPF CALIBRATION AND VALIDATION PROCEDURES

Model application procedures for HSPF have been developed and described in the HSPF application guide (Donigian et al., 1984), in numerous watershed studies over the past 25 years (Donigian, 2000a), and more recently in HSPF applications to the Chesapeake Bay watershed (Donigian et al., 1994) and the Long Island Sound watersheds in Connecticut (Love and Donigian, 2002). In addition, Donigian (2002) and Donigian and Love (2003) summarized some recent experiences in model calibration and validation with HSPF. Model application procedures for HSPF include database development, watershed segmentation, and hydrology, sediment, and water quality calibration and validation.

As noted above, model calibration and validation are necessary and critical steps in any model application. For HSPF, calibration is an iterative process of parameter evaluation and refinement, as a result of comparing simulated and observed values of interest. It is required for parameters that cannot be deterministically, and uniquely, evaluated from topographic, climatic, edaphic, or physical/chemical characteristics of the watershed and compounds of interest. Fortunately, a large majority of HSPF parameters do not fall into this category. Calibration is based on several years of simulation (at least three to five years) in order to evaluate parameters under a variety of climatic, soil moisture, and water quality conditions. Calibration should result in parameter values that produce the best overall agreement between simulated and observed values throughout the calibration period.

Calibration includes the comparison of both monthly and annual values, and individual storm events, whenever sufficient data are available for these comparisons. All of these comparisons should be performed for a proper calibration of hydrology and water quality parameters. In addition, when a continuous observed record is available, such as for streamflow, simulated and observed values should be analyzed on a frequency basis and their resulting cumulative distributions (e.g., flow duration curves) compared to assess the model behavior and agreement over the full range of observations.

Calibration is a hierarchical process beginning with hydrology (runoff and streamflow), followed by sediment erosion and sediment transport, and finally calibration of nonpoint-source loading rates and instream fate and transport for water quality constituents. When modeling land surface processes, hydrologic calibration must precede sediment and water quality calibration, since runoff is the transport mechanism by which nonpoint-source pollution occurs. Likewise, adjustments to the in-stream hydraulics must be completed before in-stream sediment and water quality transport and processes are calibrated. Each of these steps is described briefly below.

Hydrologic Calibration

Hydrologic simulation combines the physical characteristics of the watershed and the observed meteorologic data series to produce the simulated hydrologic response. All watersheds have similar hydrologic components, but they are generally present in different combinations; thus, different hydrologic responses occur on individual watersheds. HSPF simulates runoff from four components: surface runoff from impervious areas directly connected to the channel network, surface runoff from pervious areas, interflow from pervious areas, and groundwater flow. Since the historic streamflow is not divided into these four units, the relative relationship among these components must be inferred from the examination of many events over several years of continuous simulation.

A complete hydrologic calibration involves a successive examination of the following four characteristics of the watershed hydrology, in the following order: (1) annual water balance, (2) seasonal and monthly flow volumes, (3) base flow, and (4) storm events. Simulated and observed values for each characteristic are examined and critical parameters are adjusted to improve or attain acceptable levels of agreement (discussed further below).

The annual water balance specifies the ultimate destination of incoming precipitation and is indicated as:

Precipitation
– Actual evapotranspiration
– Deep percolation
+/– Change in soil moisture storage
= Runoff

HSPF requires precipitation and potential evapotranspiration (PET) as input, which effectively drive the hydrology of the watershed; actual evapotranspiration is calculated by the model using the inputs for potential and ambient soil moisture conditions. Thus, both inputs must be accurate and representative of the watershed conditions; it is often necessary to adjust the input data derived from neighboring stations that may be some distance away in order to reflect conditions in the study watershed. HSPF allows the use of spatial adjustment factors that uniformly adjust the input data to watershed conditions, based on local isohyetal, evaporation, and climatic patterns. Fortunately, evaporation does not vary as greatly as precipitation with distance, and use of evaporation data from distant stations, e.g., 80 to 160 km (50 to 100 miles), is common practice.

In addition to climate input, the critical HSPF parameters that affect components of the annual water balance include soil moisture storages, infiltration rates, actual evapotranspiration, and losses to deep groundwater recharge (see the BASINS website, www.epa.gov/ost/basins/bsnsdocs/html, for information on HSPF parameters, including Technical Note 6, which provides parameter estimation guidance).

From the water balance equation, if precipitation is measured on the watershed, and if deep percolation to groundwater is small or negligible, actual evapotranspiration must be adjusted to cause a change in the long-term runoff component of the water balance. Changes in soil moisture storages, e.g., LZSN (lower zone soil moisture nominal) in HSPF, and vegetation characteristics affect the actual evapotranspiration by making more or less moisture available to evaporate or transpire. Both soil moisture and infiltration parameters also have a major impact on percolation and are important in obtaining an annual water balance. In addition, on extremely small watersheds, i.e., less than 80 to 200 ha (200 to 500 acres) that contribute runoff only during and immediately following storm events, surface detention and near-surface soil moisture storages can also affect annual runoff volumes because of their impact on individual storm events (described below). Whenever there are losses to deep groundwater, such as recharge or subsurface flow not measured at the flow gauge, the recharge parameters are used to represent this loss from the annual water balance.

In the next step in hydrologic calibration, after an annual water balance is obtained, the seasonal or monthly distribution of runoff can be adjusted with use of the infiltration parameter. This seasonal distribution is accomplished by dividing the incoming moisture among surface runoff, interflow, upper zone soil moisture storage, and percolation to lower zone soil moisture and groundwater storage. Increasing infiltration will reduce immediate surface runoff (including interflow) and increase the groundwater component; decreasing infiltration will produce the opposite result.

The focus of the next stage in the hydrologic calibration is the base flow component. This portion of the flow is often adjusted in conjunction with the seasonal/monthly flow calibration (previous step) since moving runoff volume between seasons often means transferring the surface runoff from storm events in wet seasons to low-flow periods during dry seasons; by increasing the infiltration parameter, runoff is delayed and occurs later in the year as an increased groundwater or base flow. The shape of the groundwater recession, i.e., the change in base flow discharge, is controlled by the groundwater recession rate, which controls the rate of outflow from the groundwater storage. Using hydrograph separation techniques, these values are often calculated as the slope of the receding base flow portion of the hydrograph; these initial values are then adjusted as needed through calibration.

In the final stage of hydrologic calibration, after an acceptable agreement has been attained for annual/monthly volumes and base flow conditions, simulated hydrographs for selected storm events can be effectively altered by adjusting surface detention and interflow parameters. These parameters are used to adjust the shape of the hydrograph to improve the agreement with observed data; both parameters are evaluated primarily from past experience and modeling studies, and then adjusted in calibration. In addition, minor adjustments to the infiltration parameter can be used to improve simulated hydrographs. Examination of both daily and short time interval (e.g., hourly or 15 min) flows may be performed depending on the purpose of the study and the available data.

In addition to the above comparisons, the water balance components (input and simulated) should be reviewed for consistency with expected literature values for the study watershed. This effort involves displaying model results for individual land uses for the following water balance components:

- Precipitation (plus irrigation and other gains/losses).
- Total runoff (sum of overland flow, interflow, and base flow).
- Total actual ET (sum of interception ET, upper zone ET, lower zone ET, base flow ET, and active groundwater ET).
- Deep groundwater recharge and losses.

Although observed values are not always available for each of the water balance components listed above, the average annual water balance produced by the model must be consistent with expected values for the region, as impacted by the individual land use categories. This is a separate consistency, or reality, check with data independent of the modeling (except for precipitation and irrigation) to ensure that the land use categories and the associated overall water balance reflect the local conditions. Studies by the USGS, Extension Service, local water agencies, and nearby universities are potential sources of such data to help establish the related water balance components for the model to emulate.

For many years, the hydrology calibration process has been facilitated with the aide of BASINS and the program HSPEXP, an expert system for hydrologic calibration specifically designed for use with HSPF, developed under contract for the U.S. Geological Survey (Lumb et al., 1994). BASINS provides data download capabilities for obtaining observed flow time series, post-processing tools for viewing model simulation results, as well as graphical user interfaces for interacting with the HSPF input files. The HSPEXP package gives calibration advice, such as which model parameters to adjust and/or input to check, based on predetermined rules, and allows the user to interactively modify the HSPF users control input (UCI) files, make model runs, examine statistics, and generate a variety of plots.

HSPEXP was developed by practiced modelers as an expert system to assist modelers with hydrology calibration. HSPEXP advises the user on which parameters can be meaningfully adjusted to reduce simulation error. It also provides explanations regarding the modifications so that the less experienced modeler gains practical understanding of calibration techniques. It is important to note that although HSPEXP can be an extremely useful tool in the calibration process, it has limitations. The user should have practical knowledge of the watershed and of the HSPF algorithms, and should be able to make good judgments regarding parameter adjustment. HSPEXP provides advice for hydrology calibration only (i.e., not water quality) and without consideration of snow simulation. It addition, it does not provide advice about how much to change a parameter, nor how much parameter values should differ among different model segments/land uses.

Snow Calibration

Since snow accumulation and snowmelt are important components of streamflow in cold climates, accurate simulation of snow depths and snowmelt processes is needed to successfully model the hydrologic behavior of the watershed. HSPF simulates snowfall and snow accumulation, maintaining a depth of snow on the land surface as a continuous variable throughout the model. Snow calibration is actually a part of the hydrologic calibration. It is usually performed during the initial phase of the hydrologic calibration since the snow simulation can impact not only winter runoff volumes but also spring and early summer streamflows.

Simulation of snow accumulation on the ground and snowmelt processes suffers from two main sources of modeling uncertainty: representative climatic input data, and parameter estimation. Additional climate data required for snowmelt runoff simulation (i.e., air temperature, solar radiation, wind, and dewpoint temperature) are often not available in the immediate vicinity of the watershed and consequently must be estimated or extrapolated from the nearest available weather station. Snowmelt simulation is especially sensitive to the air temperature and solar radiation, since these are the major driving forces in snowmelt processes. When additional nearby stations with air temperature data are available, the spatial adjustment factors, noted above, are used to adjust each of the required input climatic data types to precisely represent the hydrologic conditions in the watershed. The model also allows an internal correction for air temperature as a function of elevation, using a “lapse” rate that specifies the change in temperature for any elevation difference between the watershed and the temperature gauge.

In most applications, the primary goal of the snowmelt simulation is to adequately represent the total volume and timing of snowmelt to produce reasonable soil moisture conditions in the spring and early summer so that subsequent rainfall events can be accurately simulated. Where observed snow depth (and water equivalent) measurements are available, comparisons with simulated values should be made. However, a tremendous variation in observed snow depth values can occur in a watershed, as a function of elevation, exposure, topography, etc. Thus, a single observation point or location will not always be representative of the watershed. See BASINS Technical Note 6 (USEPA, 2000) for discussion and estimation of the snow parameters.

In many instances, it is difficult to determine if problems in the snowmelt simulation are due to the non-representative climate data or inaccurate parameter values. Consequently, the accuracy expectations and general objectives of snow calibration are not as rigorous as for the overall hydrologic calibration. Comparisons of simulated weekly and monthly runoff volumes with observed streamflow during snowmelt periods, and observed snow depth (and water equivalent) values are the primary procedures performed for snow calibration. Day-to-day variations and comparisons on shorter intervals (i.e., 2 hours, 4 hours, 6 hours, etc.) are usually not as important as representing the overall snowmelt volume and relative timing in the observed weekly or biweekly period.

Hydraulic Calibration

HSPF hydraulics are computed by combining the simulated inflows from the local drainage, inflows from any upstream reaches, and the physical data contained in the FTABLE, which is the stage-discharge relationship used for hydraulic routing in each stream reach. The FTABLE specifies values for surface area, reach volume, and discharge for a series of selected average depths of water within each reach. This information is a part of the required model input and is obtained from cross-section data, channel characteristics (e.g., length, slope, roughness), and flow calculations. Since the FTABLE is an approximation of the stage-discharge-volume relationship for relatively long reaches, calibration of the values in the FTABLE is generally not needed. However, if flows and storage volumes at high flow conditions appear to be incorrect, some adjustments may be justified. Since HSPF cannot represent bidirectional flow, e.g., estuaries, linkage with hydrodynamic models is often needed to simulate tidal conditions and flow in rivers and streams with extremely flat slopes.

Sediment Erosion Calibration

Sediment calibration must follow the hydrologic calibration and precede the water quality calibration. Calibration of the parameters involved in simulation of watershed sediment erosion is more uncertain than hydrologic calibration due to less experience with sediment simulation in different regions of the country. The process is analogous; the major sediment parameters are modified to increase agreement between simulated and recorded monthly sediment loss and storm event sediment removal. However, observed monthly sediment loss is often not available, and the sediment calibration parameters are not as distinctly separated between those that affect monthly sediment and those that control storm sediment loss. In fact, annual sediment losses are often the result of only a few major storms during the year. Donigian and Love (2003) provide a focused discussion of sediment calibration procedures and techniques that are recommended for HSPF.

Sediment loadings to the stream channel are estimated by land use category from literature data, local Extension Service sources, or by using procedures like the Universal Soil Loss Equation (USLE) (Wischmeier and Smith, 1972); the estimates are then adjusted for delivery to the stream with estimated sediment delivery ratios. Model parameters are adjusted so that model-calculated loadings are consistent with estimated loading ranges. The resulting loadings are further evaluated in conjunction with in-stream sediment transport calibration (discussed below) that extends to a point in the watershed where sediment concentration data are available. The objective is to represent the overall sediment behavior of the watershed, with knowledge of the morphological characteristics of the stream (i.e., aggrading or degrading behavior), using sediment loading rates that are consistent with available values and providing a reasonable match with in-stream sediment data.

In-Stream Sediment Transport Calibration

Once the sediment loading rates are calibrated to provide the expected input to the stream channel, the sediment calibration focuses on the channel processes of deposition, scour, and transport that determine both the total sediment load and the outflow sediment concentrations that will be compared with observations. Although the sediment load from the land surface is calculated in HSPF as a total sediment, it must be divided into sand, silt, and clay fractions for simulation of in-stream processes. Each sediment size fraction is simulated separately, and storages of each size are maintained for both the water column (i.e., suspended sediment) and the bed.

In HSPF, the transport of the sand (non-cohesive) fraction is commonly calculated as a power function of the average velocity in the channel reach in each time step. This transport capacity is compared to the available inflow and storage of sand particles; the bed is scoured if there is excess capacity to be satisfied, and sand is deposited if the transport capacity is less than the available sand in the channel reach. For silt and clay (cohesive) fractions, shear stress calculations are performed by the hydraulics module and are compared to user-defined critical, or threshold, values for deposition and scour for each size. When the shear stress in each time step is greater than the critical value for scour, the bed is scoured at a user-defined erodibility rate; when the shear stress is less than the critical deposition value, the silt or clay fraction deposits at a settling rate that is defined by the user for each size. If the calculated shear stress falls between the critical scour and deposition values, then the suspended material is transported through the reach. After all scour and/or deposition fluxes have been determined, the bed and water column storages are updated, and outflow concentrations and fluxes are calculated for each time step. These simulations are performed by the SEDTRN module in HSPF, complete details of which are provided in the HSPF user manual (Bicknell et al., 2005).

In HSPF, sediment transport calibration involves numerous steps in determining model parameters and appropriate adjustments needed to ensure a reasonable simulation of the sediment transport and behavior of the channel system. These steps are usually as follows:

1. Divide input sediment loads into appropriate size fractions.
2. Run HSPF to calculate shear stress in each reach and use this information to estimate critical scour and deposition values.
3. Estimate initial parameter values and storages for all reaches.
4. Adjust scour, deposition, and transport parameters to impose scour and deposition conditions at appropriate times, e.g., scour at high flows, deposition at low flows.
5. Analyze sediment bed behavior and transport in each channel reach.
6. Compare simulated and observed sediment concentrations, bed depths, and particle size distributions, where available.
7. Repeat steps 1 through 5 as needed.

Rarely are there sufficient observed local data to accurately calibrate all parameters for each stream reach. Consequently, model users focus the calibration on sites with observed data and review simulations in all parts of the watershed to ensure that the model results are consistent with field observations, historical reports, and expected behavior from past experience. Ideally, comprehensive datasets available for storm runoff should include both tributary and main stem sampling sites. Observed storm concentrations of TSS should be compared with model results, and the sediment loading rates by land use category should be compared with the expected targets and ranges, as noted above.

Nonpoint-Source Loading and Water Quality Calibration

The essence of watershed water quality calibration is to obtain acceptable agreement of observed and simulated concentrations (i.e., within defined criteria or targets), while maintaining the in-stream water quality parameters within physically realistic bounds, and the nonpoint-source loading rates within the expected ranges from the literature. The following steps are usually performed at each of the calibration stations, following the hydrologic and sediment (for sediment-associated constituents) calibration and validation, and after the completion of input development for point-source and atmospheric contributions:

1. Estimate all model parameters, including land use specific accumulation and depletion/removal rates, washoff rates, and subsurface concentrations.
2. Superimpose the hydrology and sediment (for sediment-associated constituents) and tabulate, analyze, and compare simulated nonpoint loadings with expected range of nonpoint loadings from each land use, and adjust loading parameters when necessary to improve agreement and consistency.
3. Calibrate in-stream water temperature.
4. Compare simulated and observed in-stream concentrations at each of the calibration stations.
5. Analyze the results of comparisons in steps 3 and 4 to determine appropriate in-stream and/or nonpoint parameter adjustments, and repeat those steps as needed until calibration targets are achieved. Watershed loadings are adjusted when the in-stream simulated and observed concentrations are not in full agreement, and in-stream parameters have been adjusted throughout the range determined reasonable.

Calibration procedures and parameters for simulation of nonpoint-source pollutants will vary depending on whether constituents are modeled as sediment-associated or flow-associated. This refers to whether the loads are calculated as a function of sediment loadings or as a function of the overland flow rate. Due to their affinity for sediment, contaminants such metals, toxic organics, and phosphorous are usually modeled as sediment-associated, whereas BOD, nitrates, ammonia, and bacteria are often modeled as flow-associated.

Calibration of sediment-associated pollutants begins after a satisfactory calibration of sediment washoff has been completed. At this point, adjustments are performed in the contaminant potency factors, which are user-specified parameters for each contaminant. Potency factors are used primarily for highly sorptive contaminants that can be assumed to be transported with the sediment in the runoff. Generally, monthly and annual contaminant loss are not available, so the potency factors are adjusted by comparing simulated and recorded contaminant concentrations, or mass removal, for selected storm events. For nonpoint-source pollution, mass removal in terms of contaminant mass per unit time (e.g., g min^{-1}) is often more indicative of the washoff and scour mechanisms than are instantaneous observed contaminant concentrations.

Calibration procedures for simulation of contaminants associated with overland flow are focused on the adjustment of parameters related to daily accumulation rates (lb acre^{-1} d^{-1}), accumulation limits (lb acre^{-1}), and washoff parameters (in. h^{-1}). As was the case for sediment-associated constituents, calibration is performed by comparing simulated and recorded contaminant concentrations, or mass removal, for selected storm events. In most cases, proper adjustment of corresponding parameters can be accomplished to provide a good representation of the washoff of flow-associated constituents. The HSPF application guide (Donigian et al., 1984) includes guidelines for calibration of these parameters, and the HSPFParm database (www.epa.gov/waterscience/ftp/basins/HSPFParm/hspfparm.pdf) includes representative values for selected model applications for most conventional constituents.

In study areas where pollutant contributions are also associated with subsurface flows, contaminant concentration values are assigned for both interflow and active groundwater. The key parameters are simply the user-defined concentrations in interflow and groundwater/base flow for each contaminant. HSPF includes the functionality to allow monthly values for all nonpoint loading parameters in order to better represent seasonal variations in the resulting loading rates.

In studies requiring detailed assessment of agricultural or forested runoff water quality for nutrients or pesticides, the mass balance soil module within HSPF, referred to as AGCHEM, may need to be applied. Model users should consult the HSPF user's manual, the application guide, and studies by Donigian et al. (1998a, 1998b, 2011), Donigian (2002), and Donigian and Love (2002, 2007) that discuss application, input development, and calibration procedures. In addition, the Chesapeake Bay Program has developed excellent documentation on its landmark application of HSPF, including AGCHEM for agriculture, forest, and urban areas; new potential users of AGCHEM will find a wealth of information and guidance on the Chesapeake Bay Program website (www.chesapeakebay.net/about/programs/modeling/53/).

In-stream HSPF water quality calibration procedures are highly dependent on the specific constituents and processes that are represented, and in many ways, water quality calibration is equal parts art and science. Calibration of in-stream water quality is complicated by two factors. First, the interrelationships of the various constituents result in changes in simulated concentrations for numerous constituents when the user adjusts a parameter value that is specific to only one constituent. For example, if the user increases the value for the algal respiration rate parameter in order to reduce simulated plankton populations, the modification will also result in increased values for nutrients and inorganic carbon and a decreased value for dissolved oxygen. Thus, the final calibration of any one water quality constituent cannot be completed until all adjustments have been made to associated constituents. The calibration is complete when the best overall fit to the data is achieved for all constituents which are simulated.

The second factor that complicates in-stream water quality calibration is the wide range of values that have been reported for certain model parameters. The variability of literature values for many of these parameters results from the complexity of the physical, chemical, and biological factors that influence the ultimate biochemistry of each individual stream. Given the potential complexity of in-stream water quality calibration, as well as the flexibility allowed in constituents and processes simulated by HSPF, the calibration procedures described here can only be a general guide for model users, with a need for the model user to define more detailed calibration procedures at the onset of each specific model application.

As noted above, the goal is to obtain acceptable agreement of observed and simulated concentrations (i.e., within defined criteria or targets) while maintaining the in-stream water quality parameter values within physically realistic bounds and the nonpoint loading rates within the expected ranges from the literature. The specific model parameters to be adjusted depend on the model options selected and constituents being modeled (e.g., BOD decay rates, re-aeration rates, settling rates, algal growth rates, temperature correction factors, coliform die-off rates, adsorption/desorption coefficients, etc.). Part of the art of water quality calibration is assessing the interacting effects of modeled quantities, e.g., algal growth on nutrient uptake, and being able to analyze multiple time series plots jointly to determine needed parameter adjustments. The HSPF application guide and other model application references noted above are useful sources of information on calibration practices; restraining the calibration process to realistic parameter values can be achieved by referring to available parameter value compendiums that are published in the literature (e.g., Bowie et al., 1985).

Extent of Calibration

A common dilemma for model users is determining how much calibration is needed for an acceptable model application. Stated differently, what is the ideal level of calibration, and at the opposite extreme, what would be considered a minimum level? This is a universal modeling issue, and not specific to HSPF. There are no definitive answers to these questions that are applicable to all modeling applications. Ideally, a watershed model application, for moderate-size watersheds of 260 km^2 (100 mi^2) or more, would include multiple gauging sites on the main stem, supplemented with monitoring sites on each major tributary, along with small-scale single land use sites for the dominant land uses in the watershed. This level of monitoring is rarely available, and as a result, most applications are performed with considerably less data. A minimum level of support data might be a gauge at the watershed outlet, plus another upstream main stem gauge, along with a smaller-scale tributary gauge. Multiple years of data are needed and should span a minimum of three to five years for calibration and a comparable record for validation. In most cases, water quality data will be much less frequent than flow data (which should be continuous) and may be limited to biweekly or monthly grab samples; some storm event sampling should be included if nonpoint sources are a significant contributor.

In reality, all models can be applied without calibration; they do not require the observed monitoring data to operate. The purpose of calibration (and validation) is to develop and demonstrate confidence that the model is adequately representing the dynamic watershed behavior with sufficient accuracy so that the model can be used for subsequent planning and decision-making. This level of accuracy (discussed further under Model Performance Criteria) is highly specific to the types of water resources and water quality issues being addressed, and the types of decisions to be made. Thus, the acceptable level of data for support of such model applications is also a function of the decisions to be made. The general guidelines noted above provide some range of what might be considered acceptable for many applications.

Model Parameterization

From the above discussion on model calibration, and the hierarchical nature of calibration with HSPF, it should be evident that a comprehensive model like HSPF includes literally thousands of model parameters; a detailed listing of all the parameters, along with definitions, units, evaluation methods, etc., is well beyond the scope of this paper. However, there are considerable resources available to assist and guide users in estimating model parameters and in calibration. Fortunately, the vast majority of model parameters can be estimated from the watershed setting, including climate, topography, soils, land use, channel dimensions, etc. A much more limited set of parameters are commonly used in calibration of the various components of HSPF (i.e., hydrology, sediment erosion, nonpoint loadings, hydraulics, sediment transport, water quality fate and transport), and these resources also address initial estimation of the calibration parameters, which would then be adjusted as part of the calibration process.

The BASINS website (www.epa.gov/ost/basins/bsnsdocs/html) is the primary repository of HSPF documentation, computer codes, and resources for parameter estimation and assessment. The primary documents and databases supporting HSPF parameter estimation and calibration are as follows:

- BASINS Technical Note 6: Estimating hydrology and hydraulic parameters for HSPF (USEPA, 2000).
- BASINS Technical Note 8: Sediment parameter calibration guidance for HSPF (USEPA, 2006).
- HSPFParm: An interactive database of HSPF model parameters, Version 1.0 (Donigian et al., 1999).
- Application guide for HSPF (Donigian et al., 1984) (Somewhat outdated operational details, but useful calibration information for new users).
- Parameter lists for the major modules of HSPF (PERLND, IMPLND, RCHRES) (Donigian, 2010).

As evidenced by the above list, a wide range of useful parameter and calibration information and guidance is available, but no single source covers all the myriad types of parameter data that a user might need. This is one of the major weaknesses of the supporting information for HSPF. The HSPFParm database, published in 1999, includes complete parameter information for approximately 45 model applications (both quantity and quality) across the U.S. and provides interactive search, analysis, and export capabilities to extract parameter sets for users. Although it is still available on the BASINS website and is provided as part of BASINS/HSPF workshop materials, the USEPA has not had resources in recent years to support further populating the database with the hundreds of additional applications of HSPF across the U.S. since that time. Under support from the Minnesota Pollution Control Agency, HSPFParm is currently being upgraded and customized for HSPF applications within Minnesota, with an expected completion of late 2012 (C. Regan, MPCA, personal communication, January 2012).

MODEL PERFORMANCE CRITERIA

Model performance criteria, sometimes referred to as calibration or validation criteria, have been contentious topics for more than 30 years (Thomann, 1980, 1982; Donigian, 1982; ASTM, 1984). These issues have been thrust to the forefront in the environmental arena as a result of the need for and use of modeling for exposure/risk assessments, TMDL determinations, and environmental assessments. Although no consensus on model performance criteria is apparent from past and recent model-related literature, a number of basic truths are evident and are likely to be accepted by most modelers in modeling natural systems:

- Models are approximations of reality; they cannot precisely represent natural systems.
- There is no single, accepted statistic or test that determines whether or not a model is validated.
- Both graphical comparisons and statistical tests are required in model calibration and validation.
- Models cannot be expected to be more accurate than the errors (confidence intervals) in the input and observed data.

All of these basic truths must be considered in the development of appropriate procedures for model performance and quality assurance of modeling efforts. Despite a lack of consensus on how they should be evaluated, in practice, environmental models have long been applied, and their results continue to be used for assessment and regulatory purposes. A “weight of evidence” approach is most widely used and accepted when models are examined and judged for acceptance for these purposes. Simply put, the “weight of evidence” approach embodies the above truths and demands that multiple model comparisons, both graphical and statistical, be demonstrated in order to assess model performance, while recognizing inherent errors and uncertainty in the model, the input data, and the observations used to assess model acceptance.

Although calibration and validation of all watershed models uses different types of graphical and statistical procedures, they will generally include some of the following:

Graphical comparisons:

- Time series plots of observed and simulated values for fluxes (e.g., flow) or state variables (e.g., stage, sediment concentration, biomass concentration).
- Observed vs. simulated scatter plots, with a 1:1 line displayed, for fluxes or state variables.
- Cumulative frequency distributions of observed and simulated fluxes or state variable (e.g., flow duration curves).

Statistical tests:

- Error statistics, e.g., mean error, absolute mean error, relative error, relative bias, standard error of estimate, etc.
- Correlation tests, e.g., linear correlation coefficient, coefficient of model-fit efficiency, etc.
- Cumulative distribution tests, e.g., Kolmogorov-Smirnov (KS) test.

These comparisons and statistical tests are fully documented in a number of comprehensive references on applications of statistical procedures for biological assessment (Zar, 1999), hydrologic modeling (McCuen and Snyder, 1986), and environmental engineering (Berthouex and Brown, 1994).

Time series plots are generally evaluated visually as to the agreement, or lack thereof, between the simulated and observed values. Scatter plots usually include calculation of a correlation coefficient, along with the slope and intercept of the linear regression line; thus, the graphical and statistical assessments are combined. For comparing observed and simulated cumulative frequency distributions (e.g., flow duration curves), the KS test can be used to assess whether the two distributions are different at a selected significance level. Unfortunately, the reliability of the KS test is a direct function of the population of the observed data values that define the observed cumulative distribution. Except for

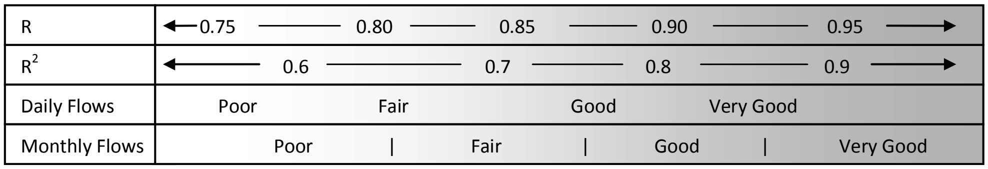


Figure 4. R and R^2 value ranges for model performance.

flow comparisons at major USGS gauge sites, there is unlikely to be sufficient observed data (i.e., more than 50 data values per location and constituent) to perform this test reliably for most water quality and biotic constituents. Moreover, the KS test is often quite easy to pass, and a visual assessment of the agreement between observed and simulated flow duration curves, over the entire range of high to low flows, may provide a comparable or even more demanding means of assessing agreement in many situations.

In recognition of the inherent variability in natural systems and unavoidable errors in field observations, the USGS provides the following characterization of the accuracy of its streamflow records in all its surface water data reports (e.g., Socolow et al., 1997):

Excellent = 95% of daily discharges are within 5% of the true value.

Good = 95% of daily discharges are within 10% of the true value.

Fair = 95% of daily discharges are within 15% of the true value.

Records that do not meet these criteria are rated as “poor.” Clearly, model results for flow simulations that are within these accuracy tolerances can be considered acceptable calibration and validation results, since these levels of uncertainty are inherent in the observed data.

Table 1 lists general calibration/validation tolerances or targets that have been provided to model users as part of HSPF training workshops over the past ten years (e.g., Donigian, 2000b). The values in the table attempt to provide some general guidance, in terms of the percent mean errors or differences between simulated and observed values, so that users can gauge what level of agreement or accuracy (i.e., very good, good, fair) may be expected from the model application.

Certain caveats apply to table 1. The tolerance ranges should be applied to mean values, and individual events or observations may show larger differences and still be acceptable. In addition, the level of agreement to be expected depends on many site and application-specific conditions, including the data quality, purpose of the study, available resources, and available alternative assessment procedures that could meet the study objectives.

Table 1. General calibration/validation targets or tolerances for HSPF applications (Donigian, 2000b).

	% Difference Between Simulated and Recorded Values		
	Very Good	Good	Fair
Hydrology/flow	<10	10 to 15	15 to 25
Sediment	<20	20 to 30	30 to 45
Water temperature	<7	8 to 12	13 to 18
Water quality/nutrients	<15	15 to 25	25 to 35
Pesticides/toxics	<20	20 to 30	30 to 40

Figure 4 provides value ranges for both correlation coefficients (R) and coefficient of determination (R^2) for assessing model performance for both daily and monthly flows. The figure shows the range of values that may be appropriate for judging how well the model is performing based on the daily and monthly simulation results. As shown, the ranges for daily values are lower to reflect the difficulties in exactly duplicating the timing of flows, given the uncertainties in the timing of model inputs, mainly precipitation.

Given the uncertain state-of-the-art in model performance criteria, the inherent errors in input and observed data, and the approximate nature of model formulations, absolute criteria for watershed model acceptance or rejection are not generally considered appropriate by most modeling professionals. And yet most decision makers want definitive answers to the questions: How accurate is the model? Is the model good enough for this evaluation? And how uncertain or reliable are the model predictions? Consequently, we propose that targets or tolerance ranges, such as those shown above, be defined as general targets or goals for model calibration and validation for the corresponding modeled quantities. These tolerances should be applied to comparisons of simulated and observed mean flows, stage, concentrations, and other state variables of concern in the specific study effort, with larger deviations expected for individual sample points in both space and time. The values shown above have been derived primarily from HSPF experience and selected past efforts on model performance criteria; however, they reflect common tolerances accepted by many modeling professionals (Donigian and Imhoff, 2009).

CASE STUDIES

This section presents results from two HSPF applications as examples of the types of graphical and statistical comparisons recommended for model calibration and validation. The first example is an application of HSPF in Connecticut for nutrient loadings to Long Island Sound, and the second is an application of HSPF to the Housatonic River in Massachusetts for hydrology modeling. As noted earlier, for an excellent application of AGCHEM for nonpoint load simulations, refer to the Chesapeake Bay Program landmark application of HSPF and AGCHEM (www.chesapeakebay.net/about/programs/modeling/53/).

Connecticut Watershed Model (CTWM)

The Connecticut Watershed Model (CTWM), based on HSPF, was developed to evaluate nutrient sources and loadings within each of six nutrient management zones that lie primarily within the state of Connecticut, and assess their delivery efficiency to Long Island Sound (LIS). The CTWM evolved by first performing calibration and validation on three small test basins across the state (Norwalk, Quinnipiac, and Salmon, fig. 5) representing a range of land uses, including urban, forest, and agricultural. The model was then extended to three major river calibration basins (Farmington, Housatonic, and Quinebaug) and subsequently expanded to a statewide model by using the most spatially applicable set of calibrated watershed parameters in non-calibrated areas. The user-friendly interface and framework of the CTWM was specifically designed to promote continuing use to assess multiple BMPs, implementation levels, and relative impacts of point-source controls for nutrient reductions to LIS. Complete details of the study and the model development and application are provided in the final report (AQUA TERRA Consultants and HydroQual, 2001). Love and Donigian (2002) summarized the techniques and methods used in the CTWM model development and the "weight of evidence" approach used in the calibration and validation, while Donigian and Love (2002) discussed and presented the model results of alternative growth and management scenarios on nutrient loads to LIS.

The hydrologic calibration for the test watersheds and the major basins was performed for the time period of 1991-1995, while the period of 1986-1990 was used for validation. The available flow data include continuous flow records at the USGS gauge sites shown in figure 5 for the entire time period. Consistent with the calibration procedures discussed above, comparisons of simulated and observed flow were performed during the calibration and validation at daily, monthly, and yearly time steps, as well as flow-frequency duration assessments. In addition, the input and simulated water balance components (e.g., precipitation, runoff, and evapotranspiration) were reviewed for the individual land uses.

Calibration of the CTWM was a cyclical process of making changes to parameter values, running the model, producing the aforementioned comparisons of simulated and observed values, and interpreting the results. This process was greatly facilitated with the use of HSPEXP, an expert system for hydrologic calibration, specifically designed for use with HSPF (Lumb et al., 1994). This package gives calibration advice, such as which model parameters to adjust and/or input to check, based on predetermined rules, and allows the user to interactively modify the HSPF UCI files, make model runs, examine performance statistics, and generate a variety of plots. The post-processing capabilities of the BASINS component program GenScn were used extensively in the calibration/validation effort.

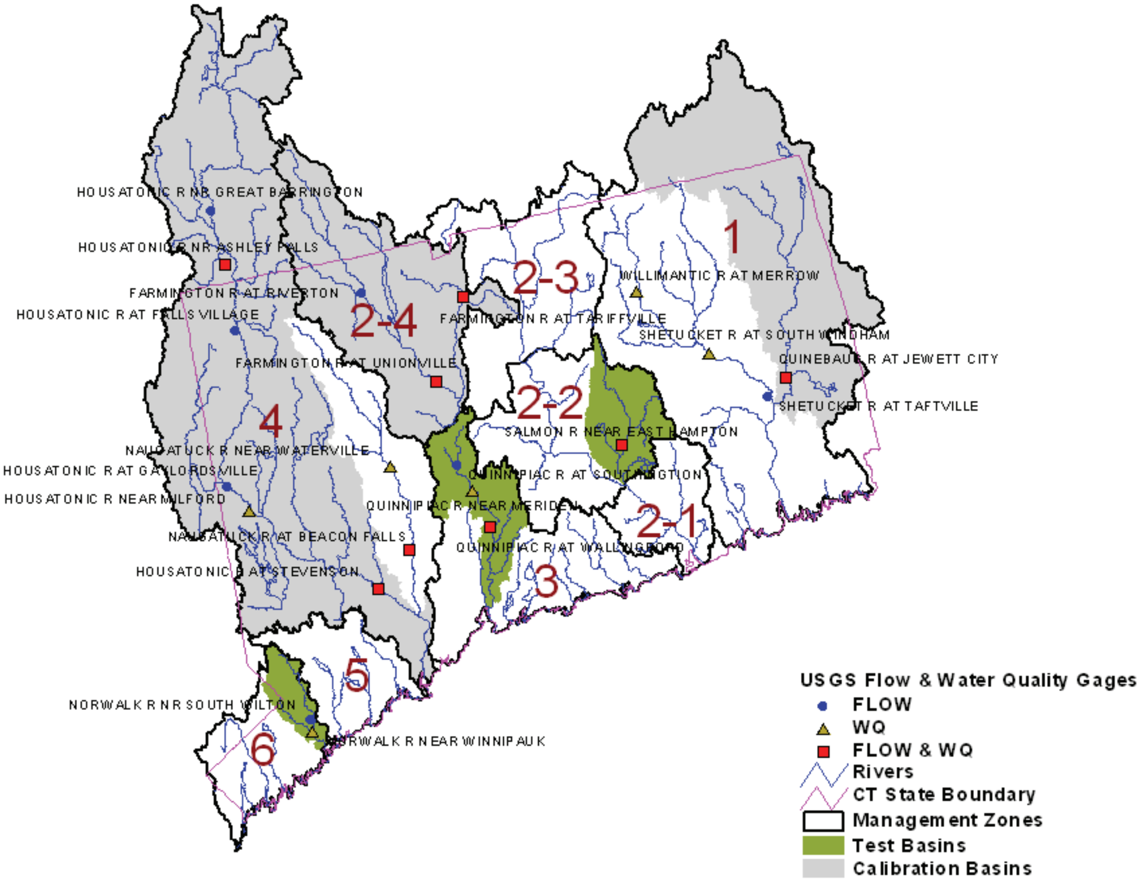


Figure 5. USGS flow and water quality gauges for the CTWM.

The hydrology calibration focused primarily on the monthly agreement of simulated and observed values, as opposed to individual storm events, due to the greater sensitivity of LIS to long-term versus short-term nutrient loads (HydroQual, 1996).

The time period of the water quality calibration coincided with the hydrology calibration period, i.e., 1991-1995. However, sufficient water quality data to support a validation were not available, the primary limitation being the lack of adequate point-source data for the earlier period. In addition, both resource and data limitations precluded modeling sediment erosion and in-stream sediment transport and deposition processes, and their impacts on water quality. The calibration followed the steps discussed above for nonpoint and water quality calibration. The results presented here are a summary of the complete modeling results presented in the final report (AQUA TERRA Consultants and HydroQual, 2001).

Table 2 shows the mean annual runoff, simulated and observed, along with daily and monthly correlation coefficients for the six primary calibration sites. The CTWM hydrology results consistently show a good to very good agreement based on annual and monthly comparisons, as defined by the calibration/validation targets discussed above. The monthly correlation coefficients are consistently greater than 0.9, and the daily values are greater than 0.8. The annual volumes are usually within the 10% target for a very good agreement, and always within the 15% target for a good agreement.

Figures 6 and 7 present graphical comparisons of simulated and observed daily flows for the Quinnipiac River at Wallingford and the Farmington River at Tariffville, respectively. Figures 8 and 9 show flow duration plots for the same sites. Figures 10 and 11 show the scatterplots for daily flows at the Farmington gauge for the calibration and validation periods. Based on the general "weight of evidence" involving both graphical and statistical tests, the hydrology component of the CTWM was confirmed to be both calibrated and validated, and provides a sound basis for the water quality and loading purposes of the study.

WATER QUALITY RESULTS

As noted above, the essence of watershed water quality calibration is to obtain acceptable agreement of observed and simulated concentrations (i.e., within defined criteria or targets), while maintaining the in-stream water quality parameters within physically realistic bounds, and maintaining the nonpoint loading rates within the expected ranges from the literature. The nonpoint loading rates, sometimes referred to as "export coefficients," are highly variable, with value ranges sometimes up to an order of magnitude depending on local and site conditions of soils, slopes, topography, climate, etc. Although a number of studies on export coefficients have been done for Connecticut, the values developed by Frink (1991) and shown in table 3 along with a standard error term, appear to have the widest acceptance.

Table 3. Frink's export coefficient (lb acre^{-1} year^{-1}).

	Total Nitrogen	Total Phosphorus
Urban	12.0 ±2.3	1.5 ±0.2
Agriculture	6.8 ±2.0	0.5 ±0.13
Forest	2.1 ±0.4	0.1 ±0.03

Table 4. CTWM loading rates (lb acre^{-1} year^{-1}).

	Mean (Range)	
	Total Nitrogen	Total Phosphorus
Urban, pervious	8.5 (5.6 to 15.7)	0.26 (0.20 to 0.41)
Urban, impervious	4.9 (3.7 to 6.6)	0.32 (0.18 to 0.36)
Agriculture	5.9 (3.4 to 11.6)	0.30 (0.23 to 0.44)
Forest	2.4 (1.4 to 4.3)	0.04 (0.03 to 0.08)
Wetlands	2.2 (1.4 to 3.5)	0.03 (0.02 to 0.05)

The above loading rates were used for general guidance, to supplement our past experience, in evaluating the CTWM loading rates and imposing relative magnitudes by land use type. No attempt was made to specifically calibrate the CTWM loading rates to duplicate the export coefficients noted above. The overall calculated mean annual loading rates and ranges for total N and total P for 1991-1995 are summarized in table 4. Considering the purposes of the study, and the assumptions in the model development (e.g., sediment not simulated), these loading rates were judged to be consistent with Frink's values and the general literature, and thus acceptable for the modeling effort (see the final report for details and breakdown of TN and TP into components).

Tables 5 and 6 display the mean simulated and observed quality stations where calibration was performed. The comparison of mean concentrations, and the ratios of simulated to observed values, demonstrate that simulated values are generally within 20% of observed, i.e., the ratios are mostly between 0.8 and 1.2, and often between 0.9 and 1.1. The biggest differences are for the phosphorus compounds, where the ratios range from 0.91 to 1.9. Considering all the sites (table 6), the mean value for the ratios for DO, TOC-concentrations for the five-year period for all of the water and nitrogen forms are within a range of 0.89 to 0.99, while the phosphorus ratios are within 1.33 to 1.40. Comparing these ratios to the proposed calibration targets indicates a "very good" calibration of nitrogen and a borderline "fair" calibration of phosphorus.

Table 2. Summary of CTWM hydrologic calibration/validation annual flow and correlation coefficients.

		Calibration Period (1991-1995)				Validation Period (1986-1990)			
		Mean Annual Flow (in.)		R Value		Mean Annual Flow (in.)		R Value	
Station Name	Station Number	Obs.	Sim.	Avg. Daily	Avg. Monthly	Obs.	Sim.	Avg. Daily	Avg. Monthly
Test watershed gauges									
Salmon River near East Hampton	01193500	23.6	24.4	0.83	0.92	26.3	25.8	0.79	0.92
Quinnipiac River at Wallingford	01196500	26.3	26.4	0.82	0.94	29.0	28.3	0.71	0.91
Norwalk River at South Wilton	01209700	21.4	21.7	0.84	0.93	25.9	25.2	0.75	0.91
Major basin gauges									
Quinebaug River at Jewett City	01127000	23.8	23.6	0.82	0.93	27.2	24.7	0.86	0.95
Farmington River at Tariffville	01189995	26.2	26.0	0.85	0.92	26.2	29.1	0.87	0.94
Housatonic River at Stevenson	01205500	31.7	31.9	0.88	0.98	34.6	31.5	0.87	0.96

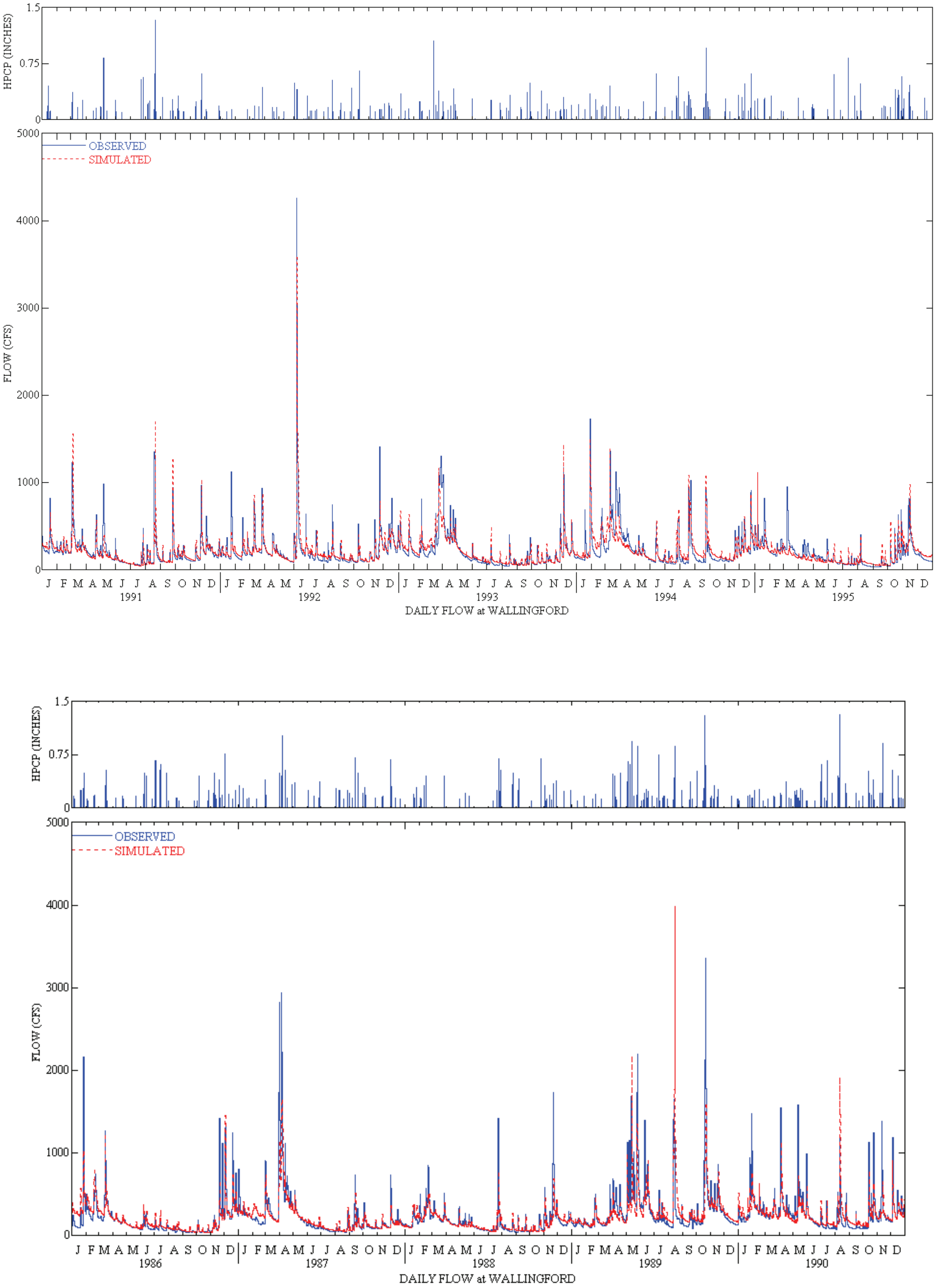


Figure 6. Observed and simulated daily flow for the Quinnipiac River at Wallingford, Connecticut: (top) calibration and (bottom) validation.

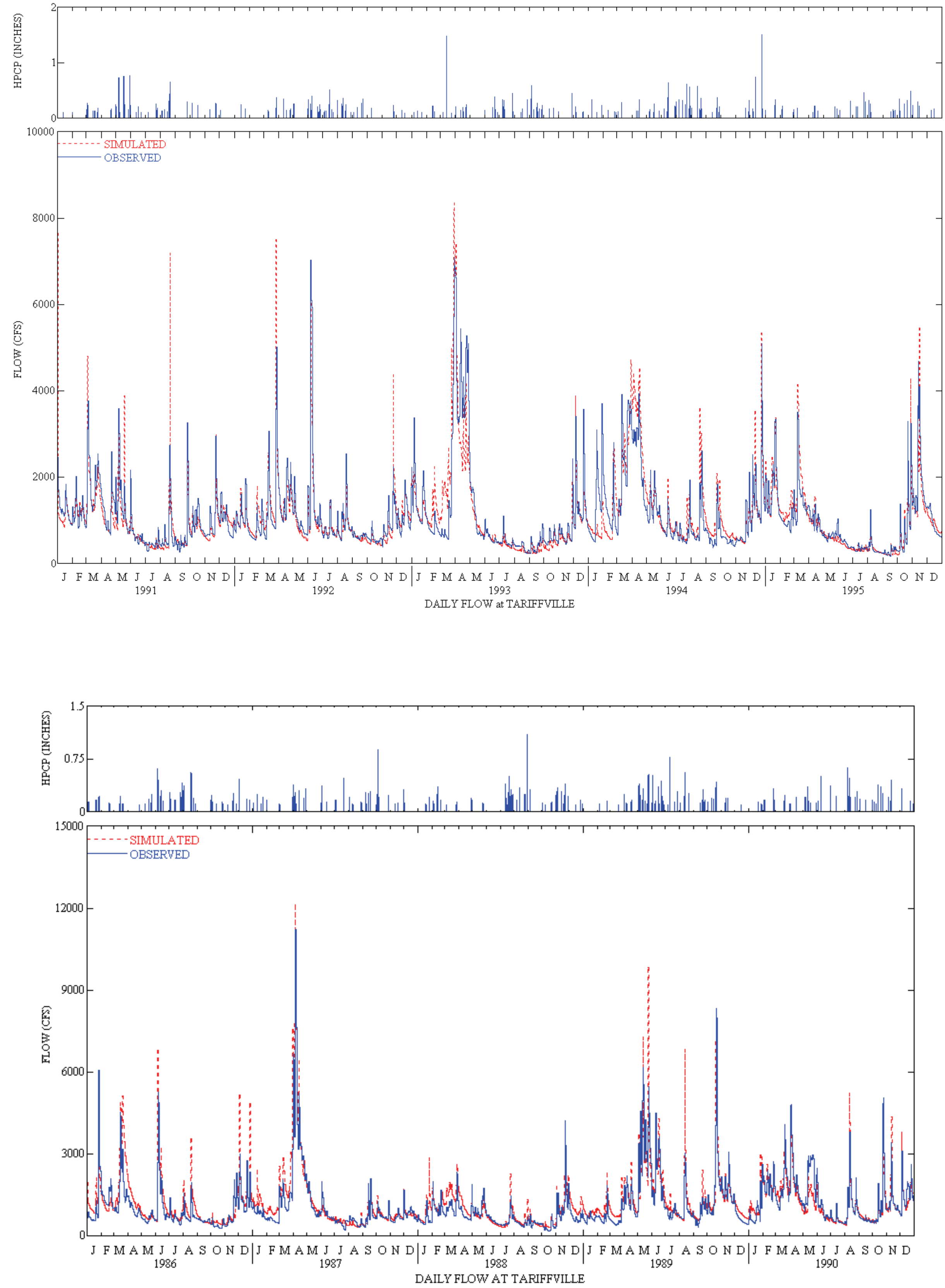
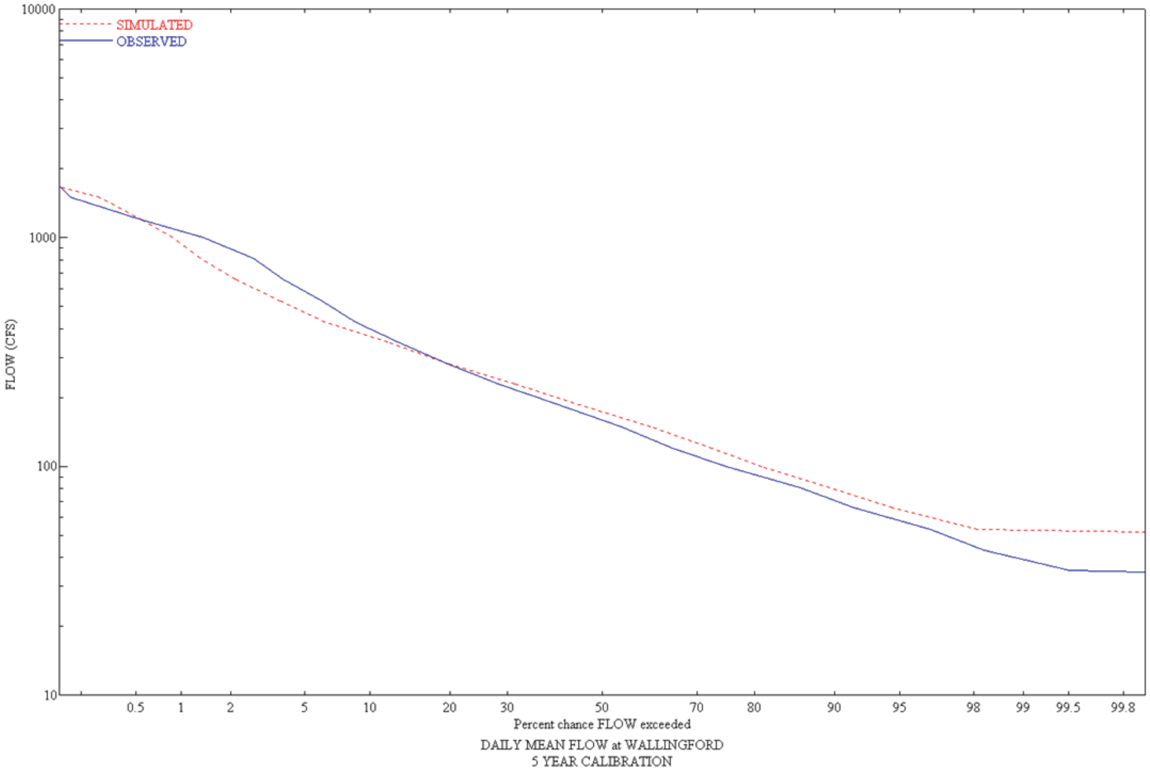


Figure 7. Observed and simulated daily flow for the Farmington River at Tariffville, Connecticut: (top) calibration and (bottom) validation.



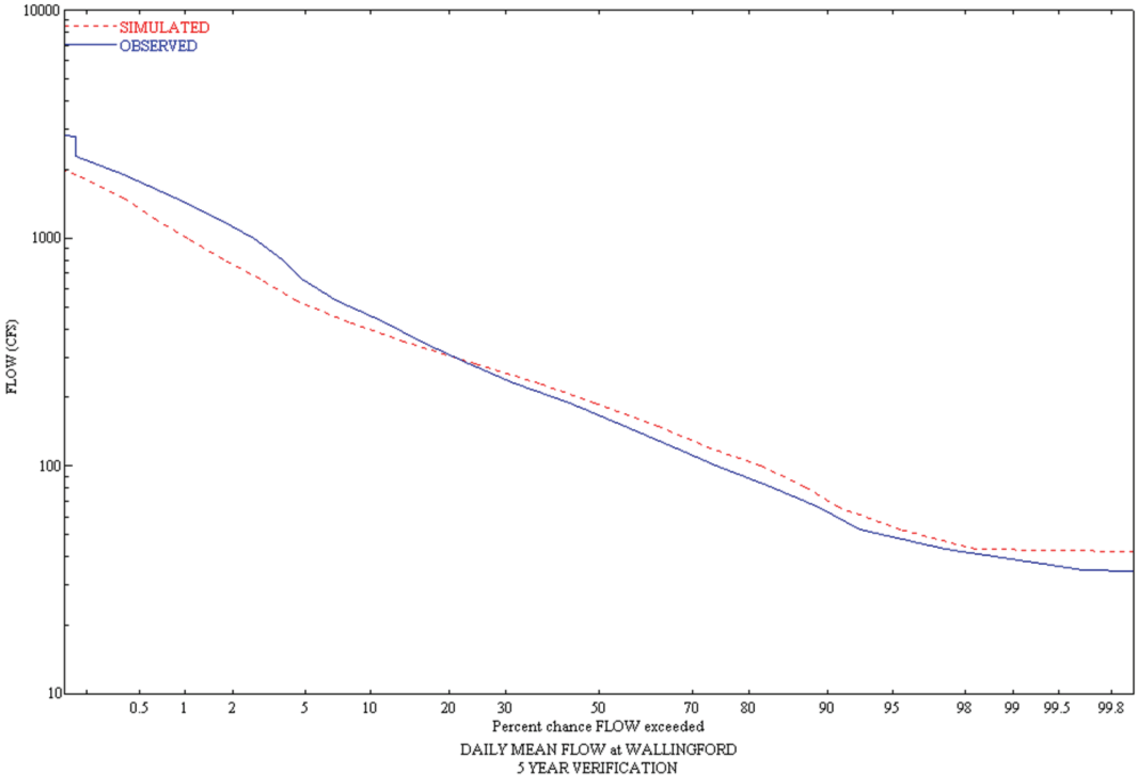
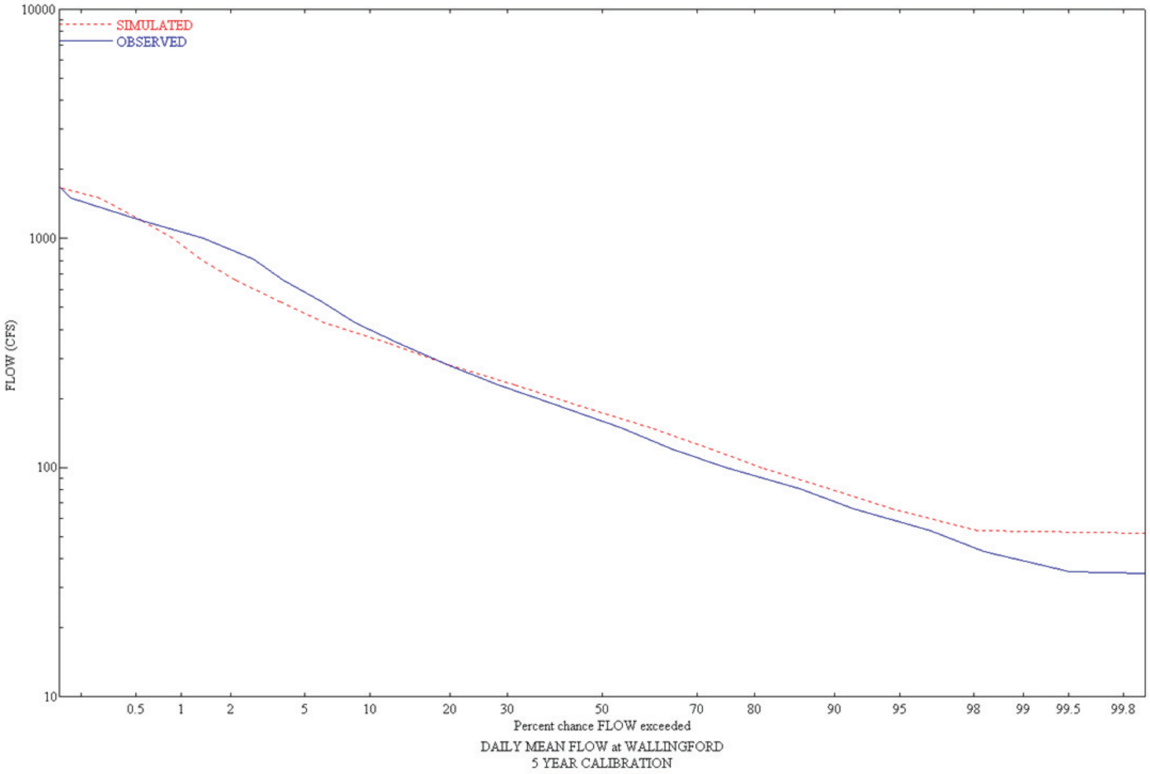


Figure 8. Observed and simulated daily flow duration curves for the Quinnipiac River at Wallingford, Connecticut: (top) calibration and (bottom) validation.



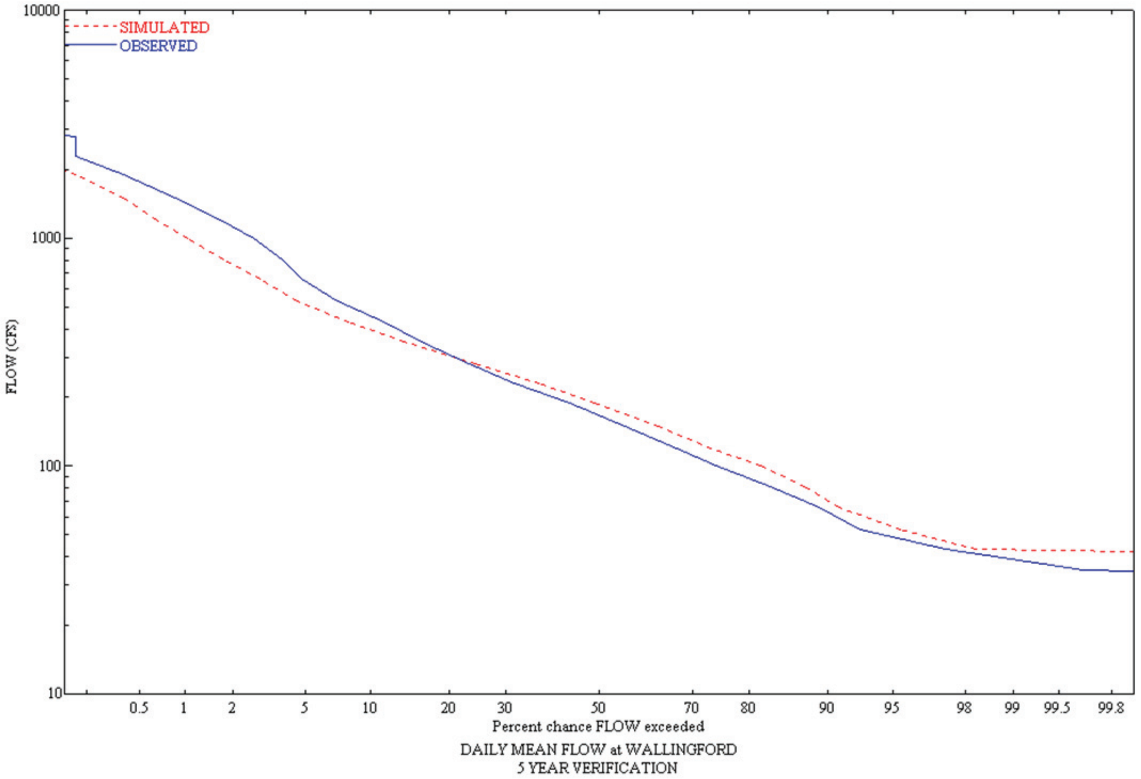
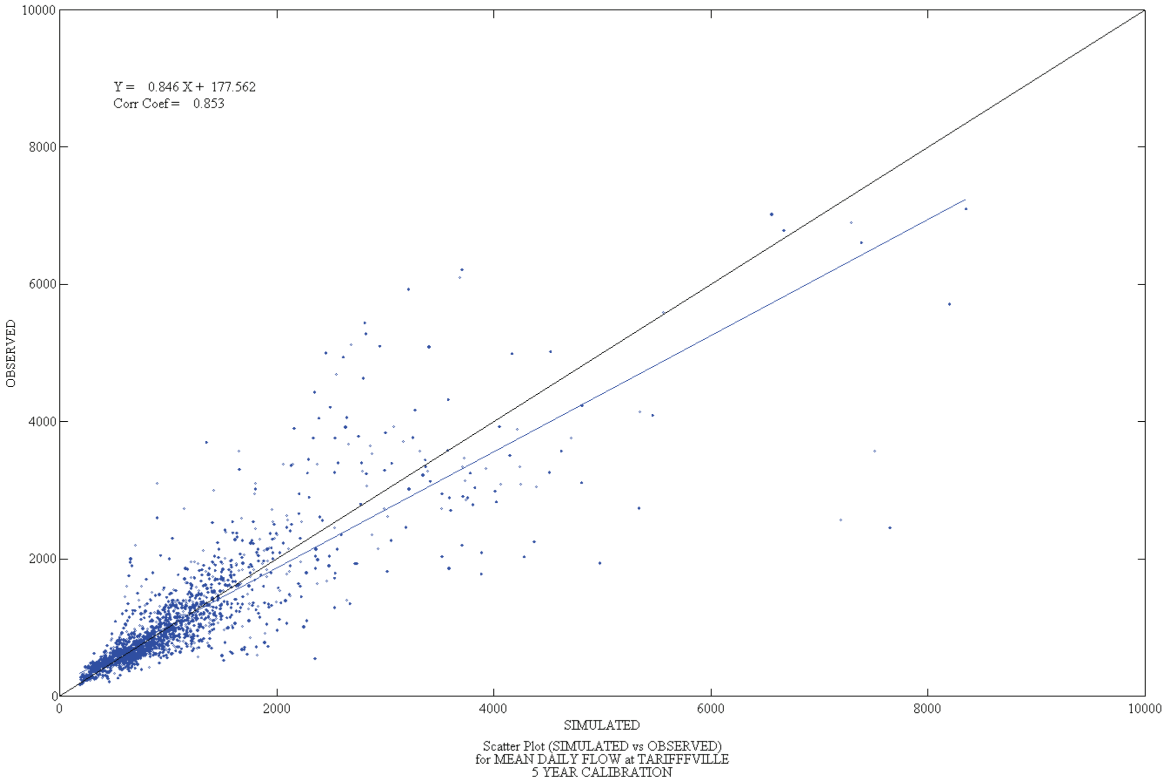


Figure 9. Observed and simulated daily flow duration curves for the Farmington River at Tariffville, Connecticut: (top) calibration and (bottom) validation.



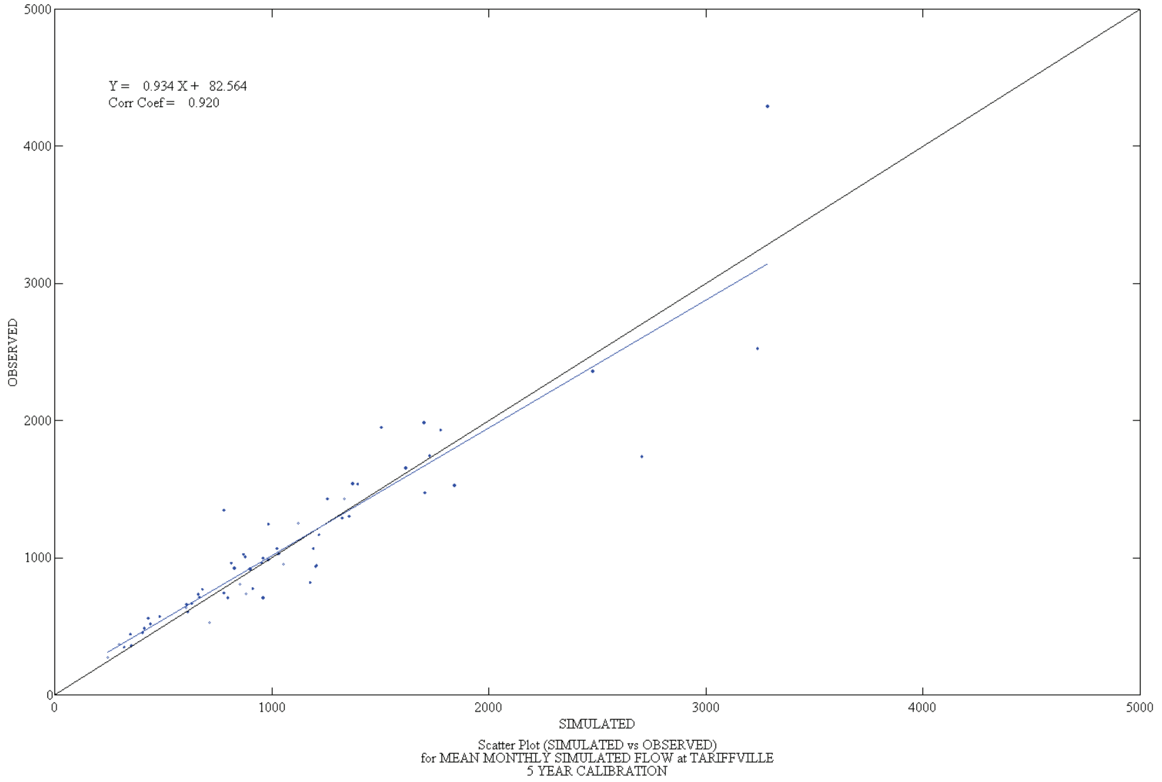
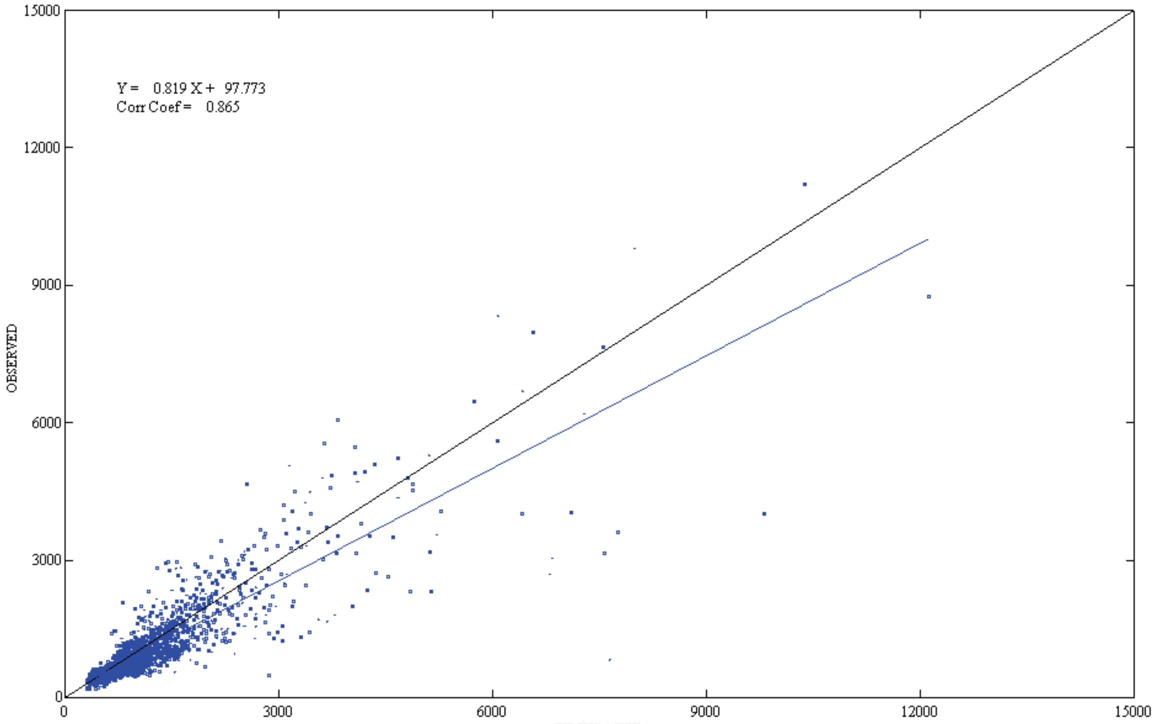
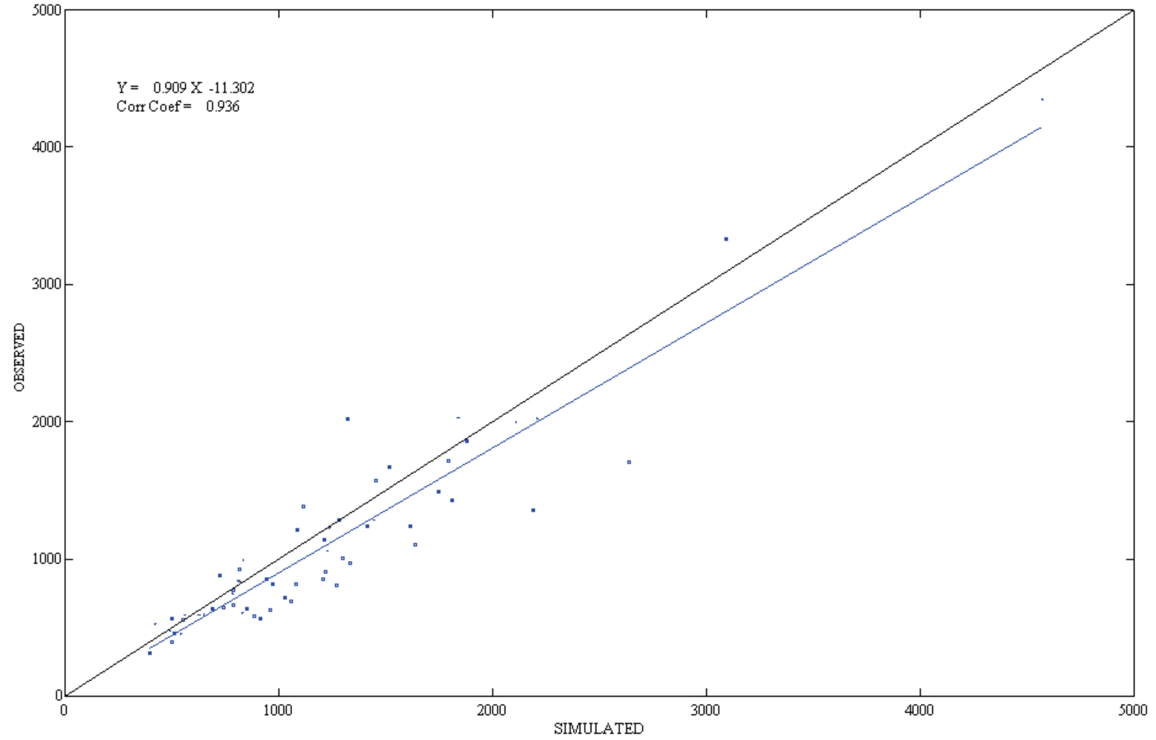


Figure 10. Scatterplots of observed and simulated (top) daily and (bottom) monthly flow for the Farmington River at Tariffville, Connecticut.



Scatter Plot (SIMULATED vs OBSERVED)
for MEAN DAILY SIMULATED FLOW at TARIFFVILLE
5 YEAR VERIFICATION



Scatter Plot (SIMULATED vs OBSERVED)
for MEAN MONTHLY SIMULATED FLOW at TARIFFVILLE
5 YEAR VERIFICATION

Figure 11. Scatterplots of observed and simulated (top) daily and (bottom) monthly flow for the Farmington River at Tariffville, Connecticut.

Table 5. Average annual concentrations (mg L^{-1}) for the calibration period (1991-1995).

Constituent	Obs.	Sim.	Ratio[a]	Obs.	Sim.	Ratio[a]	Obs.	Sim.	Ratio[a]
	Salmon River near East Hampton			Quinnipiac River at Wallingford			Norwalk River at Winnipauk		
Dissolved oxygen	10.9	10.5	0.96 (48)	10.4	10.3	0.99 (46)	11.6	10.4	0.90 (97)
Ammonia as N	0.03	0.02	0.82 (43)	0.19	0.18	0.92 (46)	0.04	0.04	1.18 (80)
Nitrite-nitrate as N	0.22	0.27	1.21 (46)	2.82	2.45	0.87 (46)	0.39	0.40	1.03 (93)
Organic nitrogen	0.31	0.25	0.80 (30)	0.50	0.60	1.20 (44)	0.33	0.28	0.86 (70)
Total nitrogen	0.53	0.51	0.97 (30)	3.64	3.29	0.90 (44)	0.73	0.69	0.94 (70)
Orthophosphate as P	0.01	0.01	0.91 (48)	0.32	0.36	1.10 (46)	0.02	0.02	0.93 (94)
Organic phosphorus	0.02	0.02	1.30 (48)	0.07	0.11	1.62 (46)	0.02	0.03	1.18 (94)
Total phosphorus	0.02	0.03	1.35 (48)	0.39	0.47	1.19 (46)	0.04	0.05	1.10 (94)
Total organic carbon	3.9	2.8	0.71 (45)	4.5	4.8	1.06 (44)	4.0	3.2	0.81 (28)
	Quinebaug River at Jewett City			Farmington River at Tariffville			Housatonic River at Stevenson		
Dissolved oxygen	10.4	10.3	0.99 (43)	10.2	10.8	1.06 (49)	9.5	9.5	1.01 (41)
Ammonia as N	0.08	0.06	0.73 (42)	0.10	0.09	0.82 (48)	0.06	0.06	1.10 (33)
Nitrite-nitrate as N	0.44	0.37	0.84 (42)	0.71	0.59	0.83 (49)	0.36	0.41	1.15 (40)
Organic nitrogen	0.45	0.39	0.86 (40)	0.31	0.28	0.90 (45)	0.33	0.28	0.84 (38)
Total nitrogen	0.96	0.80	0.83 (40)	1.15	0.97	0.85 (45)	0.77	0.75	0.97 (38)
Orthophosphate as P	0.02	0.04	1.67 (43)	0.07	0.13	1.90 (49)	0.01	0.02	1.49 (32)
Organic phosphorus	0.03	0.04	1.23 (43)	0.03	0.05	1.59 (49)	0.02	0.03	1.19 (33)
Total phosphorus	0.06	0.08	1.44 (43)	0.10	0.18	1.82 (49)	0.03	0.05	1.47 (40)
Total organic carbon	5.6	4.9	0.86 (41)	3.9	3.3	0.84 (45)	3.8	2.9	1.06 (49)

[a] Ratios calculated from simulated and observed concentrations prior to rounding (sample size shown in parentheses).

Table 6. Average and range of simulated and observed concentration ratios for all sites.

Constituent	Average	Range
Dissolved oxygen	0.99	0.90 to 1.06
Ammonia as N	0.93	0.73 to 1.18
Nitrite-nitrate as N	0.99	0.83 to 1.21
Organic nitrogen	0.91	0.80 to 1.20
Total nitrogen	0.91	0.83 to 0.97
Orthophosphate as P	1.33	0.91 to 1.90
Organic phosphorus	1.35	1.18 to 1.62
Total phosphorus	1.40	1.10 to 1.82
Total organic carbon	0.89	0.71 to 1.06

Figures 12 and 13 present typical graphical comparisons made for simulated and observed water quality constituents. Figure 12 presents a comparison of simulated and observed total phosphorus for the Quinnipiac River at Wallingford. Figure 13 presents a similar comparison for total nitrogen at the Tariffville gauge on the Farmington River.

Based on the general "weight of evidence" of the hydrology and water quality simulation results, including the CTWM loading rates, the mean concentrations and ratios, and the time series comparisons of observed and simulated values, the CTWM was determined to be an acceptable representation of the Connecticut watersheds providing load-

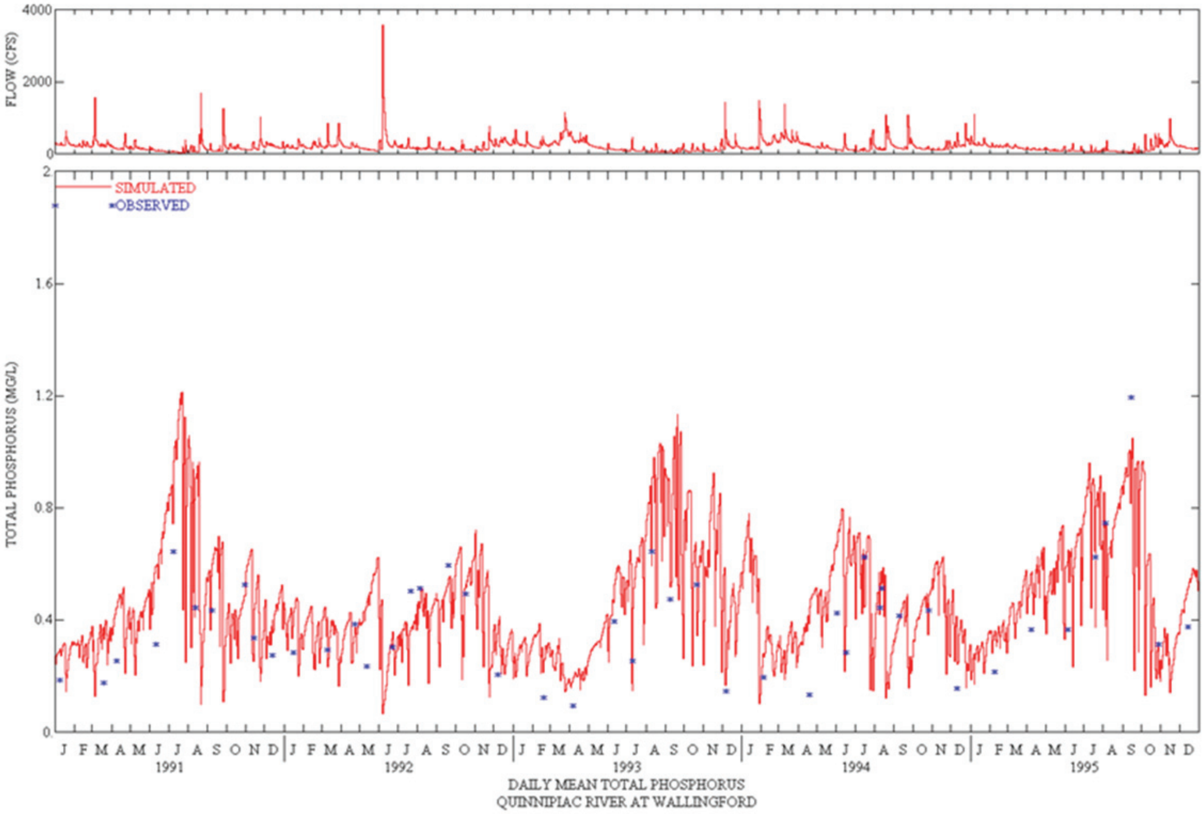


Figure 12. Observed and simulated daily total phosphorus concentrations for the Quinnipiac River at Wallingford, Connecticut.

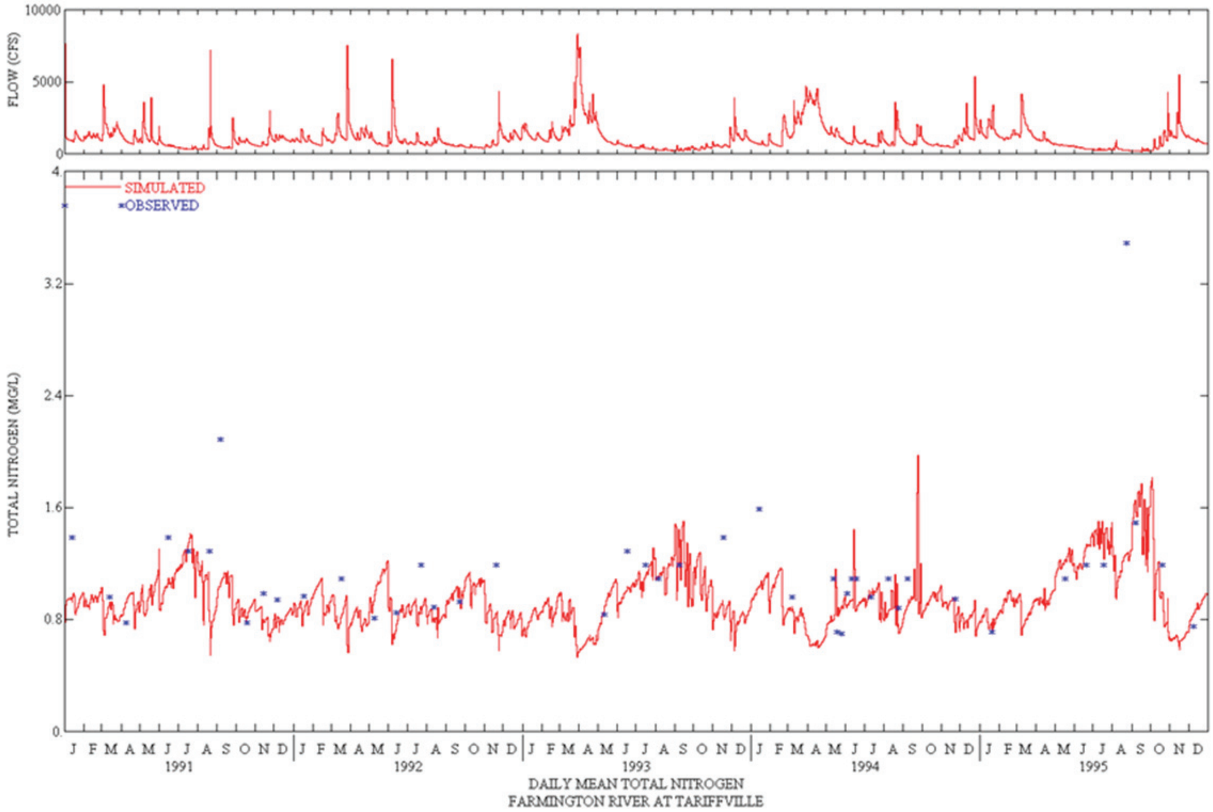


Figure 13. Observed and simulated daily total nitrogen concentrations for the Farmington River at Tariffville, Connecticut.

Table 7. Annual simulated and observed runoff (in.) for the Housatonic River watershed.[a]

Year	Precipitation	Simulated Flow	Observed Flow	Error (%)
1990	58.9	35.1	35.6	-1.4
1991	47.0	23.3	22.8	2.1
1992	45.7	23.7	20.1	15.2
1993	47.6	27.6	26.0	5.8
1994	46.3	25.9	25.5	1.5
1995	44.0	20.7	21.0	-1.4
1996	62.0	39.4	41.5	-5.3
1997	42.2	21.4	23.2	-8.4
1998	42.2	22.0	23.9	-8.6
1999	46.9	21.6	24.8	-14.8
Total	482.7	260.7	264.4	-1.4
Average	48.3	26.1	26.4	-1.4

[a] Source: Weston Solutions (2006).

Table 8. Annual flow statistics from HSPEXP.[a]

	Upstream Tributary		Watershed Outlet	
Statistic	Sim.	Obs.	Sim.	Obs.
Average runoff (in.)	27.12	26.23	26.07	26.44
Total of highest 10% flows (in.)	10.88	10.72	8.56	8.94
Total of lowest 50% flows (in.)	4.22	4.19	5.09	5.13
Evapotranspiration (in.)	23.77	25.55[b]	23.41	26.09[b]
Total storm volume (in.)[c]	47.07	51.91	38.72	42.36
Average of storm peaks (cfs)[c]	710.84	791.88	2310.38	2287.19
	Calc.	Crit.	Calc.	Crit.
Error in total volume (%)	3.40	10.00	-1.40	10.00
Error in 10% highest flows (%)	1.50	15.00	-4.20	15.00
Error in 50% lowest flows (%)	0.60	10.00	-0.60	10.00
Error in storm peaks (%)	-10.20	15.00	1.00	15.00

[a] Source: Weston Solutions (2006). Sim. = simulated, Obs. = observed, Calc. = calculated, and Crit. = criterion.

[b] PET (estimated by multiplying observed pan evaporation by 0.73).

[c] Based on 31 storms occurring between 1990 and 1999.

ings to LIS. This evidence indicates that the predicted nitrogen and carbon loadings are a "very good" representation of the observed data, based on the established calibration targets, and that the phosphorus loadings are a "fair" representation. Clearly, improvements can be made to better represent these loadings, especially for phosphorus, but the CTWM in its current form is a sound tool for examining loadings to LIS and providing the basis for developing and analyzing alternative watershed scenarios designed to improve the water quality of LIS.

HOUSATONIC RIVER WATERSHED MODEL

HSPF was applied to the almost 780 km^2 (300 mi^2) Housatonic River watershed in Massachusetts. The tables in this section demonstrate some additional types of comparisons for evaluating the hydrologic simulation results, in comparison with the targets shown in table 1. Table 7 shows the annual simulated and observed runoff, along with annual precipitation and percent error or difference for each year of the ten-year calibration. The total difference for the ten years is less than 2%, while the annual differences are within about 15%, indicating a good to very good calibration (Weston Solutions, 2004, 2006).

Table 8 shows the statistical output available from HSPEXP for both the watershed outlet and an upstream tributary of about 155 km^2 (60 mi^2), while table 9 shows a variety of statistics for both daily and monthly comparisons at the watershed outlet. The storm statistics in table 8 are based on a selection of 31 events throughout the ten-year period, distributed to help evaluate seasonal differences. The correlation statistics in table 9 indicate a "good" cali-

Table 9. Daily and monthly average flow statistics for the Housatonic River watershed.[a]

	Daily		Monthly	
	Sim.	Obs.	Sim.	Obs.
Count	3652	3652	120	120
Mean(cfs)	539.85	547.65	540.46	547.56
Geometric mean(cfs)	376.61	380.86	424.39	428.44
Correlation coefficient (R)	0.86		0.93	
Coeff. of determination (R^2)	0.74		0.87	
Mean error(cfs)	-7.80		-7.10	
Mean absolute error(cfs)	152.97		101.22	
RMS error(cfs)	284.09		140.26	
Model fit efficiency (1.0 = perfect)	0.73		0.87	

[a] Source: Weston Solutions (2006).

Table 10. Average observed monthly runoff and residuals for the Housatonic River watershed.[a]

Month	Average Observed (in.)	Average Simulated (in.)	Average Residual[b] (in.)	Error (%)
Jan.	2.94	2.71	-0.24	-8.09
Feb.	2.01	2.34	0.33	16.46
Mar.	3.61	3.85	0.23	6.42
Apr.	4.25	4.16	-0.09	-2.07
May	2.86	2.28	-0.58	-20.19
June	1.44	1.26	-0.18	-12.55
July	1.07	0.97	-0.10	-9.03
Aug.	0.95	1.13	0.18	18.66
Sept.	0.85	0.98	0.14	16.39
Oct.	1.75	1.66	-0.08	-4.80
Nov.	2.15	2.05	-0.09	-4.38
Dec.	2.56	2.70	0.13	5.03
Total	26.46	26.08	-0.35	-1.32%

[a] Source: Weston Solutions (2006).

[b] Average residual = simulated – observed.

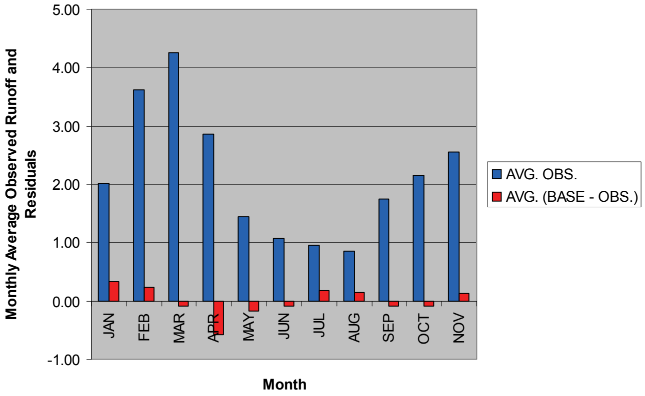


Figure 14. Housatonic watershed observed runoff and residuals (inches) (Weston Solutions, 2006).

bration for daily values, and a "very good" calibration of monthly flows, when compared to the value ranges in figure 3.

Table 10 shows the mean monthly observed and simulated runoff, along with their differences (or residuals) and error (%), as another assessment of the seasonal representation of the model. Figure 14 graphically shows the mean observed and the residuals from table 10. This demonstrates a need to improve the spring and early summer results, where the model underestimated the monthly observations.

Table 11 shows the simulated and expected water balance for the watershed, and table 12 shows the separate water balances for each land use simulated by the model. As noted earlier, these comparisons are consistency checks to compare the overall simulation with the expected values from the literature, and to evaluate how well the model represents land use differences.

Table 11. Average annual expected and simulated water balance.

Component	Expected Ranges	Simulated
Moisture supply	43 to 53	48
Total runoff	23 to 27	24
Total ET	20 to 23	23
Deep recharge	1 to 4	1

Table 12. Simulated water balance components by land use.[a]

Component	Forest	Agri-culture	Urban Pervious	Wetland	Urban Impervious
Moisture supply	48.6	48.4	48.5	48.5	48.3
Total runoff	22.6	25.8	26.5	21.3	42.8
Surface runoff	1.0	4.6	4.6	0.3	42.7
Interflow	7.9	8.8	8.8	4.8	0.0
Base flow	13.6	12.3	13.1	16.2	0.0
Total ET	24.6	22.1	21.2	24.2	5.5
I/R ET[b]	9.6	6.1	6.3	4.6	5.5
Upper zone ET	7.8	6.5	9.2	11.1	0.0
Lower zone ET	6.6	9.2	5.3	4.6	0.0
Active GW ET	0.0	0.0	0.0	2.9	0.0
Base flow ET	0.6	0.3	0.3	1.0	0.0
Deep recharge	1.4	0.5	0.8	3.0	0.0

[a] Source: Weston Solutions (2006).

[b] I/R ET = interception/retention ET.

DISCUSSION

The model summary, the discussion of calibration/validation procedures, and the case studies that have been presented for HSPF present a consistent picture of the suitability and versatility of the model to address a broad scope of water resources and water quality issues. The value of HSPF to water resource planners is enhanced by the model's comprehensive nature, which features all of the following:

- Flexibility in addressing a wide range of water quantity and quality problems.
- Comprehensive representation of multiple pollutant sources (e.g., point and nonpoint) using a single model.
- Convenient data management features that save time and money.
- Modular program structure that facilitates program changes and additions for special applications.

The introductory section of this article provides a high-level list of the many processes and physical settings that can be effectively simulated using HSPF. However, every model has its limitations related to process representation, and HSPF is no exception. The most notable limitations for the model include:

- Capability to model agricultural tile drainage processes is not included in the model.
- Selected agricultural conditions, such as crop rotations and certain BMPs, are difficult to represent explicitly.
- Process detail useful for representing certain urban storm water BMPs is not fully developed.
- Capability to model wetlands processes is not explic-

itly included in the model.

- Channel hydraulics are simulated using a simplified routing technique that does not allow consideration of backwater flows and tidally influenced conditions.
- Reservoir and lake simulations are restricted to fully mixed, single compartment representations.

It is important that modelers consider the potential impact of these limitations in HSPF (or in any other model) on the eventual success or failure of their modeling efforts. Before selecting and applying a model, the user must first gain an understanding of the critical processes that dominate the environmental issues of concern, and then select a model that provides the flexibility and process detail necessary to reproduce the endpoint impacts of these processes and settings, given the limitations of the available data. In many instances, a flexible model such as HSPF can provide multiple options and approaches for evaluating a particular environmental issue. However, when HSPF cannot offer the desired level of process detail, particularly for channel processes or ecosystems effects, it has become a common practice to link the land surface loadings and/or riverine processes and output generated by HSPF to more complex models. This is especially common for coastal or tidally influenced systems, complex, deep reservoirs or lakes needing multi-dimensional analyses, or when assessing watershed impacts on ecosystem functions is needed.

The calibration/validation discussion in this article has focused on presenting a “weight of evidence” approach to watershed model calibration and validation based on experience with the HSPF model. Examples have been provided to demonstrate some of the graphical and statistical comparisons that should be performed whenever model performance is evaluated. Although not all models will employ the identical procedures described above, it is clear that multiple tests and evaluations, not reliance on a single statistic, should be part of all watershed modeling studies.

While continuous simulation models like HSPF are the most powerful tools for assessing watershed loadings, they have some significant disadvantages. These models require large amounts of input data, including observations over periods of many years. The learning process involved in using these models is significant, and like all currently available models, there is uncertainty inherent in input data, algorithms, and modeling assumptions.

BASINS reduces the disadvantages of using continuous simulation models by addressing each of these issues. BASINS provides a tremendous amount of input data so that the data gathering process is much less daunting. BASINS includes graphical user interfaces that make the models easier to use, as well as analysis tools that help make model output easier to understand. BASINS also provides a suite of watershed models with a broad range of sophistication and complexity, so that the user can choose the model most appropriate for a given study or assessment.

FUTURE DEVELOPMENTS

The current resurgence of government concern for nonpoint-source issues and problems and the focus on watershed-scale assessment and management, as catalyzed by various sections and amendments to the Clean Water Act, has renewed interest in nonpoint-source and comprehensive watershed modeling. The comprehensive nature of HSPF, and its flexibility in allowing consideration of the combined impacts of both point-source and nonpoint-source pollutants at the watershed scale, has led to unprecedented interest in model applications. In addition, the model’s use within a multimedia framework, such as that used in the Chesapeake Bay Program, and linkage with numerous estuarine and multidimensional hydrodynamic/water quality models, has further advanced its utility for sophisticated environmental analyses. To support this increased interest and usage, there will be a need for HSPF and supporting software to continue to grow. Improvements in process algorithms, enhanced and broadened capabilities to interact with a wide variety of environmental data, and more powerful user interaction will all be required.

In addition to representing natural processes, modeling systems such as HSPF must provide process algorithms that represent the effects of human-induced sources or processes on environmental state variables. Models must include algorithms that can be used to represent any environmental disturbance that could influence the behavior of the natural watershed system. Examples of such phenomena include nutrient and pesticide application, tillage practices, crop harvest and residue practices, tile drainage, livestock grazing, feedlot runoff, highway drainage, urban development, stormwater detention structures, stream channelization, combined sewers, construction practices, mine drainage, silvicultural practices, municipal and industrial discharges, etc.

Many of these conditions and effects can be represented directly by HSPF, and others can be approximated by adjusting values for parameters contained in existing HSPF algorithms; selected conditions and/or practices (e.g., BMPs) may require development of enhanced algorithms. We envision that considerable work will be done to develop additional sets of HSPF parameter value changes (i.e., model scenarios) that reflect our best understanding of the physical and chemical changes resulting from a particular modification or activity. This may be the most critical area of model development activity, as it directly affects our ability to use models like HSPF for environmental management and decision-making.

HSPF was developed prior to the proliferation of a new generation of data and data generation techniques that offer refined spatial detail for a number of parameters critical to watershed modeling. In some cases, these new data are best used to support existing process algorithms that are solved for a higher-resolution grid. However, the potential also exists to replace or enhance certain process algorithms to improve the simulation of natural processes by taking advantage of new data. For example, satellite data, GIS, and digital elevation models (DEMs) have made it possible to compute the aspect (i.e., the direction toward which a slope faces) for watersheds or watershed segments at a high level of detail. The availability of techniques to reliably compute aspect invites the incorporation of improved process algorithms for snowmelt, soil temperature, and water temperature in areas of significant topographical relief.

The two technologies that offer the greatest body of new data that could be used to refine process algorithms are satellite remote sensing data and the transformation of remote sensing data, by use of GIS and related capabilities, to derive other useful data types. The remote sensing data available from current and future satellites offer an opportunity to develop new process algorithms that could offer improved representation of precipitation, surface runoff, soil moisture, groundwater, and water quality variables, including thermal pollution, erosion, sediment load, and trophic state of receiving waters. An immediate need of watershed-scale models are algorithms using radar imaging data to better represent spatially varying inputs, such as intense localized thunderstorms.

ACKNOWLEDGEMENTS

The current HSPF model code is the result of almost 40 years of model development efforts beginning in about 1975. Those efforts were preceded by nearly 20 years of antecedent pioneering model development work on the Stanford Watershed Model at Stanford University and on the Hydrocomp Simulation Program (HSP) at Hydrocomp, Inc., by Ray Linsley, Norman Crawford, Delbert Franz, and others.

The original development of HSPF, and predecessor nonpoint-source models, was sponsored by the USEPA Environmental Research Laboratory in Athens, Georgia. David Duttweiler was the Laboratory Director, and Robert Swank was the head of the Technology Development and Applications Branch, which supervised the project during the code development period. Jim Falco was the Project Officer initially on the HSPF development work; he was succeeded by Tom Barnwell, who oversaw HSPF activities for EPA from 1979 until 1995.

The initial HSPF model and user manual development work was performed by Hydrocomp, Inc., where Robert Johanson was the Project Manager on the HSPF development effort. Other members of the Hydrocomp project team are acknowledged in the original (Release 5.0) version of the user manual (EPA Publication EPA-600/9-80-015) published in April 1980. Subsequent revisions and extensions to the HSPF code and user manual were performed by Anderson-Nichols & Co., Inc., and beginning in 1985, by AQUA TERRA Consultants, which has been the sole HSPF maintenance contractor since that time. During the post-1980 period, particularly significant code development and support efforts have been provided by Jack Kittle, Brian Bicknell, and Tom Jobes through EPA contracts with AQUA TERRA.

As noted earlier in this article, the USGS Office of Surface Water has also played a significant role in the HSPF development efforts. USGS funded a series of both code development and applications projects from 1980 through the early 2000s, producing a number of companion and complementary programs to HSPF (e.g., ANNIE, ANNIE/WDM, HSPEXP, GenScn) to support data management, data analyses, calibration, and scenario analyses with HSPF. Alan Lumb was the primary lead on all the HSPF-related USGS work until his retirement in 2002; since that time, Kate Flynn has been the primary lead.

REFERENCES

AQUA TERRA Consultants and HydroQual. 2001. Modeling nutrient loads to Long Island Sound from Connecticut watersheds, and impacts of future buildout and management scenarios. Prepared for the Connecticut Department of Environmental Protection, Hartford, Conn.

ASTM. 1984. E978-84: Standard practice for evaluating environmental fate models of chemicals. Philadelphia, Pa.: American Society of Testing Materials.

Berthouex, P. M., and L. C. Brown. 1994. *Statistics for Environmental Engineers.* Boca Raton, Fla.: CRC Press.

Bicknell, B. R., J. C. Imhoff, J. L. Kittle Jr., T. H. Jobes, and A. S. Donigian Jr. 2005. Hydrological Simulation Program—Fortran (HSPF). User's Manual for Release 12.2, Athens, Ga.: U.S. EPA National Exposure Research Laboratory in cooperation with the U.S. Geological Survey, Reston, Va.

Bowie, G. L., W. B. Mills, D. B. Porcella, C. L. Campbell, J. R. Pagenkopf, G. L. Rupp, K. M. Johnson, P. W. H. Chan, and S. A. Gherini. 1985. *Rates, Constants, and Kinetics Formulations in Surface Water Quality Modeling.* 2nd ed. EPA-600/3-85-040. Washington, D.C.: U.S. Environmental Protection Agency.

Crawford, N. H. 1962. The synthesis of continuous streamflow on a digital computer. PhD diss. Palo Alto, Cal.: Stanford University.

Crawford, N. H., and R. K. Linsley. 1966. Digital simulation on hydrology: Stanford watershed model IV. Tech. Report 39. Palo Alto, Cal.: Stanford University.

Donigian, A. S., Jr. 1982. Field validation and error analysis of chemical fate models. In *Modeling Fate of Chemicals in the Aquatic Environment*, 303-323 K. L. Dickson, A. W. Maki, and J. Cairns, eds. Ann Arbor, Mich.: Ann Arbor Science.

Donigian, A. S., Jr. 2000a. Bibliography of HSPF and related references. Mountain View, Cal.: AQUA TERRA Consultants.

Donigian, A. S., Jr. 2000b. Lecture 19: Calibration and verification issues, Slide L19-22. HSPF training workshop handbook and CD. Presented and prepared for the U.S. EPA Office of Water and Office of Science and Technology, Washington, D.C.

Donigian, A. S., Jr. 2002. Watershed model calibration and validation: The HSPF experience. In *Proc. WEF Specialty Conf.: National TMDL Science and Policy 2002*. Alexandria, Va.: Water Environment Federation.

Donigian, A. S., Jr. 2010. BASINS/HSPF training workshop CD. Presented and prepared for the U.S. EPA Office of Water and Office of Science and Technology, Washington, D.C.

Donigian, A. S., Jr., and N. H. Crawford. 1976a. Modeling pesticides and nutrients on agricultural lands. EPA-600/3-76-043. Durham, N.C.: U.S. EPA Office of Research and Development.

Donigian, A. S., Jr., and N. H. Crawford. 1976b. Modeling nonpoint pollution from the land surface. EPA-600/3-76-083. Durham, N.C.: U.S. EPA Office of Research and Development.

Donigian, A. S., Jr., and J. C. Imhoff. 2002. From the Stanford model to BASINS: 40 years of watershed modeling. Reston, Va.: ASCE Task Committee on Evolution of Hydrologic Methods through Computers.

Donigian, A. S., Jr., and J. C. Imhoff. 2009. Evaluation and performance assessment of watershed models. In *Proc. WEF Specialty Conf.: TMDL 2009*. Alexandria, Va.: Water Environment Federation.

Donigian, A. S., Jr., and J. T. Love. 2002. The Connecticut watershed model: A tool for BMP impact assessment in Connecticut. In *Proc. WEF Watershed 2002.* Alexandria, Va.: Water Environment Federation.

Donigian, A. S., Jr., and J. T. Love. 2003. Sediment calibration procedures and guidelines for watershed modeling. In *Proc. WEF Specialty Conf.: TMDL 2003*. Alexandria, Va.: Water Environment Federation.

Donigian, A. S., Jr., and J. T. Love. 2007. The Housatonic River watershed model: Model application and sensitivity/uncertainty analysis. In *Proc. 7th Intl. IWA Symp. on System Analysis and Integrated Assessment in Water Management (Watermatex 2007*). London, U.K.: International Water Association.

Donigian, A. S., Jr., and P. S. C. Rao. 1990. Selection, application, and validation of environmental models. In *Proc. Intl. Symposium on Water Quality Modeling of Agricultural Nonpoint Sources*, Part 2: 577- 604. Report ARS-81. D. G. Decoursey, ed. Washington, D.C.: USDA-ARS.

Donigian, A. S., Jr., D. C. Beyerlein, H. H. Davis Jr., and N. H. Crawford. 1977. Agricultural Runoff Management (ARM) model - Version II: Testing and refinement. EPA-600/3-77-098. Athens, Ga.: U.S. EPA Environmental Research Laboratory.

Donigian, A. S., Jr., J. C. Imhoff, B. R. Bicknell, and J. L. Kittle Jr. 1984. Application guide for the Hydrological Simulation Program - Fortran. EPA 600/3-84-066. Athens, Ga.: U.S. EPA Environmental Research Laboratory.

Donigian, A. S., Jr., B. R. Bicknell, A. S. Patwardhan, L. C. Linker, C. H. Chang, and R. Reynolds. 1994. Chesapeake Bay program - Watershed model application to calculate bay nutrient loadings: Final findings and recommendations (Final report). Prepared for the U.S. EPA Chesapeake Bay Program, Annapolis, Md.

Donigian, A. S., Jr., B. R. Bicknell, R. V. Chinnaswamy, and P. N. Deliman. 1998a. Refinement of a comprehensive watershed water quality model (Final report). Tech. Report EL-98-6. Prepared for the U.S. Army Corps of Engineers Waterways Experiment Station, Vicksburg, Miss.

Donigian, A. S., Jr., R. V. Chinnaswamy, and P. N. Deliman. 1998b. Use of nutrient balances in comprehensive watershed water quality modeling of Chesapeake Bay (Final report). Tech. Report EL-98-5. Prepared for the U.S. Army Corps of Engineers Waterways Experiment Station, Vicksburg, Miss.

Donigian, A. S., Jr., J. C. Imhoff, and J. L. Kittle Jr. 1999. HSPFParm: An interactive database of HSPF model parameters. Version 1.0. EPA-823-R-99-004. Prepared for the U.S. EPA Office of Science and Technology, Washington, D.C.

Donigian, A. S., Jr., B. R. Bicknell, and M. Bandurraga. 2011. Watershed modeling for the Santa Clara River in southern California. In *Proc. World Environ. and Water Resources Congress: Bearing Knowledge for Sustainability*. Reston, Va.: ASCE Environmental and Water Resources Institute.

Duda, P. B., J. L. Kittle Jr., M. H. Gray, P. R. Hummel, and R. A. Dusenbury. 2001. WinHSPF - An interactive Windows interface to HSPF: User's manual. Washington, D.C.: U.S. EPA Office of Water.

Duda, P. B., J. L Kittle Jr., A. S. Donigian Jr., and R. Kinerson. 2006. Chapter 11: Better Assessment Science Integrating Point and Nonpoint Sources (BASINS). In *Watershed Models*, 273-289. V. P. Singh and D. K. Frevert, eds. Boca Raton, Fla.: CRC Press.

Flynn, K. M., P. R. Hummel, A. M. Lumb, and J. L. Kittle Jr. 1995. User's manual for ANNIE, Version 2: A computer program for interactive hydrologic data management. USGS Water-Resources Investigations Report 95-4085. Reston, Va.: U.S. Geological Survey.

Frink, C. R. 1991. Estimating nutrient exports to estuaries. *J. Environ. Qual.* 20(4): 717-724.

Hydrocomp. 1976. *Hydrocomp Simulation Programming Operations Manual.* Palo Alto, Cal.: Hydrocomp, Inc.

Hydrocomp. 1977. *Hydrocomp Water Quality Operations Manual.* Palo Alto, Cal.: Hydrocomp, Inc.

HydroQual. 1996. Water quality modeling analysis of hypoxia in Long Island Sound using LIS 3.0. Conducted by direction of the Management Committee of the Long Island Sound Study through a contract with the New England Interstate Water Pollution Control Commission, Lowell, Mass.

Kittle, J. L., A. M. Lumb, P. R. Hummel, P. B. Duda, and M. H. Gray. 1998. A tool for the generation and analysis of model simulation scenarios for watersheds (GenScn). WRI Report 98-4134. Reston Va.: U.S. Geological Survey.

Love, J. T., and A. S. Donigian Jr. 2002. The Connecticut watershed model: Model development, calibration, and validation. In *Proc. WEF Watershed 2002*. Alexandria, Va.: Water Environment Federation.

Lumb, A. M., R. B. McCammon, and J. L. Kittle Jr. 1994. User's manual for an expert system (HSPEXP) for calibration of the Hydrological Simulation Program - Fortran. Water-Resources Investigations Report 94-4168. Reston, Va.: U.S. Geological Survey.

McCuen, R. H., and W. M. Snyder. 1986. *Hydrologic Modeling: Statistical Methods and Applications.* Englewood Cliffs, N.J.: Prentice-Hall.

Socolow, R. S., C. R. Leighton, J. L. Zanca, and L. R. Ramsey. 1997. Water resources data, Massachusetts and Rhode Island: Water year 1997. Water-Data Report MA-RI-97-1. Marlborough, Mass.: U.S. Geological Survey, Water Resources Division.

Thomann, R. V. 1980. Measures of verification. In *Workshop on Verification of Water Quality Models*, 37-59. R. V. Thomann and T. O. Barnwell, eds. EPA-600/9-80-016. Athens, Ga.: U.S. EPA, Office of Research and Development.

Thomann, R. V. 1982. Verification of water quality models. *J. Environ. Eng. Div. ASCE* 108(EE5): 923-940.

USEPA. 2000. BASINS Technical Note 6: Estimating hydrology and hydraulic parameters for HSPF. EPA-823-R00-012. Washington, D.C.: U.S. EPA, Office of Water. Available at: www.epa.gov/waterscience/basins/.

USEPA. 2006. BASINS Technical Note 8: Sediment parameter and calibration guidance for HSPF. Washington, D.C.: U.S. EPA, Office of Water. Available at: ww.epa.gov/waterscience/basins/.

USEPA. 2007. Better Assessment Science Integrating Point and Nonpoint Sources - BASINS Version 4.0. EPA-823-C-07-001. Washington, D.C.: U.S. EPA, Office of Water. Available at: www.epa.gov/waterscience/basins/.

Weston Solutions. 2004. Appendix A: Watershed model calibration. In *Model Calibration: Modeling Study of PCB Contamination in the Housatonic River, Vol. 2*. Prepared for the U.S. Army Corps of Engineers and U.S. EPA. DCN GE-100504-ACJS. West Chester, Pa.: Weston Solutions, Inc.

Weston Solutions. 2006. Section 6.1: HSPF watershed model validation. In *Model Calibration: Modeling Study of PCB Contamination in the Housatonic River, Vol. 1*. Prepared for the U.S. Army Corps of Engineers and U.S. EPA. DCN GE-030706-ADBR. West Chester, Pa.: Weston Solutions, Inc.

Wischmeier, W. H., and D. D. Smith. 1972. *Predicting Rainfall Erosion Losses: A Guide to Conservation Planning*. Agriculture Handbook 537. Washington, D.C.: USDA.

Zar, J. H. 1999. *Biostatistical Analysis*. 4th ed. Upper Saddle River, N.J.: Prentice Hall.