JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

Vol. 53, No. 1 AMERICAN WATER RESOURCES ASSOCIATION February 2017

THE WATERSHED FLOW AND ALLOCATION MODEL: AN NHDPLUS-BASED WATERSHED MODELING APPROACH FOR MULTIPLE SCALES AND CONDITIONS[1]

Michele C. Eddy, Fekadu G. Moreda, Robert M. Dykes, Brandon Bergenroth, Aaron Parks, and James Rineer[2]

ABSTRACT: The Watershed Flow and Allocation model (WaterFALL®) provides segment-specific, daily streamflow at both gaged and ungaged locations to generate the hydrologic foundation for a variety of water resources management applications. The model is designed to apply across the spatially explicit and enhanced National Hydrography Dataset (NHDPlus) stream and catchment network. To facilitate modeling at the NHDPlus catchment scale, we use an intermediate-level rainfall-runoff model rather than a complex process-based model. The hydrologic model within WaterFALL simulates rainfall-runoff processes for each catchment within a watershed and routes streamflow between catchments, while accounting for withdrawals, discharges, and onstream reservoirs within the network. The model is therefore distributed among each NHDPlus catchment within the larger selected watershed. Input parameters including climate, land use, soils, and water withdrawals and discharges are georeferenced to each catchment. The WaterFALL system includes a centralized database and server-based environment for storing all model code, input parameters, and results in a single instance for all simulations allowing for rapid comparison between multiple scenarios. We demonstrate and validate WaterFALL within North Carolina at a variety of scales using observed streamflows to inform quantitative and qualitative measures, including hydrologic flow metrics relevant to the study of ecological flow management decisions.

(KEY TERMS: surface water hydrology; simulation; modeling; watershed management; water resources; NHDPlus; GWLF.)

Eddy, Michele C., Fekadu G. Moreda, Robert M. Dykes, Brandon Bergenroth, Aaron Parks, and James Rineer, 2017. The Watershed Flow and Allocation Model: An NHDPlus-Based Watershed Modeling Approach for Multiple Scales and Conditions. *Journal of the American Water Resources Association* (JAWRA) 53(1):6-29. DOI: 10.1111/1752-1688.12496

INTRODUCTION

The management, regulation, restoration, and conservation of flowing water systems for use by humans and biota require an understanding of water availability in terms of location, timing and duration, and volume (Jackson *et al.*, 2001; Bunn and Arthington, 2002; Magilligan and Nislow, 2005; Poff *et al.*, 2010). Because observational data are only available at select locations, the main method to satisfy the data needs of a surface water assessment is to perform hydrologic modeling. Through modeling, multiple streams, regions, and conditions can be assessed,

---

[1] Paper No. JAWRA-16-0022-P of the *Journal of the American Water Resources Association* (JAWRA). Received January 20, 2016; accepted October 24, 2016. © 2017 American Water Resources Association. **Discussions are open until six months from issue publication**.

[2] Environmental Engineer (Eddy), Hydrologic Modeler (Moreda), Director of Water and Ecosystem Management Center (Dykes), Research Programmer and Analyst (Bergenroth, Parks), and Senior GIS Architect (Rineer), RTI International, 3040 E. Cornwallis Rd., PO Box 12194, Research Triangle Park, North Carolina 27709 (E-Mail/Eddy: mceddy@rti.org).

which provides a means to estimating continuous streamflow and water quality predictions using limited observations of these same parameters under favorable levels of effort. When available at a scale commensurate to a management objective and validated against available observed data, a model becomes a tool that aids in the understanding of the underlying hydrological processes and the impacts of human modifications (*e.g.*, land use), climatic conditions, and management options.

Multiple studies have examined scale issues within the realm of hydrologic modeling. Blöschl and Sivapalan (1995) identify three types of scales that must be considered in hydrological modeling: process, observation, and modeling (working) scale. Processes such as precipitation events, runoff, infiltration, and channel flow are known to act at different scales (Blöschl and Sivapalan, 1995; Gentine *et al.*, 2012). Similarly, the observations that can be used to parameterize and calibrate the functions of a model representing these processes are available at many different scales, though oftentimes, not at the same scale as the process. Therefore, the modeling scale (*i.e.*, the scale at which hydrologic processes are simulated) must make the best possible use of and least amount of tradeoffs between the process scale and observation scale (Blöschl and Sivapalan, 1995). In addition, the management scale, or scale at which regulating agencies may take action resulting from any outcomes of the modeling may be considered.

In terms of modeling, Kircher (2009) concluded that a simple storage-discharge relationship (a first-order, nonlinear differential equation) applied at a catchment scale of a few square kilometers captures the behavior of hydrological systems that, at smaller scales (square meters), exhibit heterogeneities, such as complex hydraulic gradients. Jakeman and Hornberger (1993) found that a linear representation of a two-compartment system represents both slow- and quick-flow components of hydrology well over a range of catchment sizes (*i.e.*, <1 to $89.6\ \mathrm{km}^2$). However, these more simplistic representations of the hydrologic cycle do not provide the opportunity to consider watershed characteristics, such as land use and topography, needed for scenario-based management assessments. At the other end of the modeling spectrum, process-based models that allow for such scenario assessments use more complex and integrated algorithms to capture the hydrologic cycle across a range of typically larger ($\gg 1\ \mathrm{km}^2$) hydrologic response units derived from elevation changes or set based on a predetermined land surface grid. Examples of such models in use by the federal government and other state and local agencies for management include the Soil and Water Assessment Tool (SWAT) (Arnold *et al.*, 1998), the Hydrologic Simulation Program—FORTRAN (HSPF) (Bicknell *et al.*, 1997), the Precipitation-Runoff Modeling System (PRMS) (Markstrom *et al.*, 2015), and the Variable Infiltration Capacity Model (Liang *et al.*, 1994).

To provide the greatest flexibility of management opportunities, the working scale of a model should be at the smallest spatial scale that directly relates to management initiatives without creating an overburden of computation and parameterization efforts. Currently, this working scale is best approximated by the catchment scale (on the order of $1 \times 10^0\ \mathrm{km}^2$), which can be related to the stream reach level (Brakebill *et al.*, 2011), colocation of individual monitoring locations (U.S. Environmental Protection Agency [USEPA], STORET Water Quality Monitoring Stations NHDPlus Indexed Dataset, https://developer.epa.gov/epa-office-of-water-ow-storet-water-quality-monitoring-stations-nhdplus-indexed-dataset/, *accessed* June 17, 2016), and the USEPA's recently instituted catchment indexing for Integrated Reporting under the Clean Water Act (USEPA, Water Quality Framework, https://www.epa.gov/waterdata/water-quality-framework, *accessed* June 17, 2016). Although some of the complex process-based models that allow for scenario assessment have been applied at the catchment scale or smaller in select applications (Arnold *et al.*, 2012; Duda *et al.*, 2012), these models are not consistently and regularly applied to the catchment scale due to the number of model parameters used as input and in computations, as well as the resulting computer processing limitations (David Wells, USEPA, August 5, 2013, personal communication). Therefore, a need exists for a model that is designed to be applied at the management scale of the catchment while also simulating larger regions and balancing the complexity of hydrologic process simulation and available observation-based data inputs.

We sought to create such a model using a combination of methods and data already demonstrated and supported through peer-reviewed literature applied over a consistent hydrologic network and supported through a rigorous data management system. We built the model up from the simplest hydrologic representation that provides the desired output, given variation in the inputs necessary for scenario building, rather than requiring the use of a model as complex as possible to represent instream flows. Our main objectives for creating the model included: (1) scalability — catchment to full watershed; (2) portability — ability to model multiple ecoregions, climatic conditions, *etc.*; and (3) flexibility — ability to model different scenarios, such as a baseline, present day, human-altered conditions, or potential future conditions, with little additional effort.

The resulting Watershed Flow and Allocation model (WaterFALL®) simulates surface water flows using rainfall-runoff and intermediate-level subsurface processes at a daily time step across the topographically derived catchments defined within the enhanced National Hydrography Dataset (NHDPlus), or any similar digital representation of a connected stream network. WaterFALL is a versatile tool, capable of supporting a wide range of water resource management needs across diverse hydrologic regimes. The model is easy to set up. Input parameters are modest and can be filled by a wide variety of available national and local databases. The model is configured to perform extensive "what if?" analyses by substituting parameter values that reflect anticipated future climate conditions, land use patterns, or water withdrawal and discharge rates. The model is highly scalable. It can be run across a single NHDPlus catchment or any user-selected grouping of interconnected catchments that comprise larger hydrologic units. The granularity provided by georeferencing physical input parameters to small catchments enables the model to better capture the influence that spatial variability in key landscape and environmental conditions exerts on the magnitude and timing of surface water flows.

WaterFALL consists of five major components: (1) a hydrologic network; (2) a rainfall-runoff model; (3) routing mechanisms between assessment units; (4) underlying data parameterization; and (5) a server-based data management and analysis system. The resulting distributed modeling framework provides a daily time series of streamflow and its intermediate components (*i.e.*, base flow, runoff, soil moisture) for each assessment unit. Applications of WaterFALL are calibrated and validated against streamflow measurements corresponding to particular assessment units, typically from U.S. Geological Survey (USGS) gages where possible, using a weight-of-evidence approach for quantitative performance measures, qualitative depictions of the hydrologic regime, and consideration of the monitored streamflow source, method, and evaluation.

This study describes the five underlying technical model components within WaterFALL in detail using process equations and data source descriptions, where appropriate. We then describe the procedures used for model calibration, in terms of parameters and processes, and evaluation of model performance. To demonstrate WaterFALL's calibration and performance, the results derived from an application of the model across the state of North Carolina are presented as standard performance metric values and visualizations as well as through tabulated hydrologic metrics. Discussions within the study include direct comparisons between WaterFALL and other hydrologic models, potential model improvements, and strengths and limitations of WaterFALL in its current state.

METHODS

Model Components

Hydrologic Network. The National Hydrography Dataset (NHD) contains digital line graphs and associated river reach attributes created from 1:100,000-scale (medium resolution) and 1:24,000-scale (high resolution; one inch of data equals 2,000 feet on the ground) topographic map series from the USGS (Simley and Carswell, 2009) for the contiguous United States (U.S.). The NHD dataset has been enhanced (NHDPlus) by incorporating the National Elevation Dataset (NED) and the Watershed Boundary Dataset (WBD) to define flowline (waterways) and catchment (drainage unit or subbasin) features. The NHDPlus catchments nest within the hydrologic units established within the WBD so that navigation upstream from a catchment at the pour point of a hydrologic unit will identify the catchments that encompass the full drainage area of the unit (McKay *et al.*, 2013). Therefore, central to the objective of WaterFALL, NHDPlus provides a consistent hydrologic network for the conterminous U.S. on which a hydrologic model may be applied. The spatially referenced flowline and elevation-derived catchment features within NHDPlus provide a basis for the spatial indexing, or georeferencing, of any defining characteristic that can be expressed within a spatial context (*e.g.*, land use distribution, dam locations, precipitation events), whereas the value-added attributes (*e.g.*, slope, stream order, average velocity) indexed to flowlines and/or catchments included within the NHDPlus database provide some of the data necessary to parameterize the rainfall-runoff model within WaterFALL. Version 1 of NHDPlus contains 2,595,196 catchments with an average area of 1.9 km^2 (1.2 mile^2), and 2,342,519 flowlines with an average length of 2.25 km (1.4 miles) (Brakebill *et al.*, 2011). This catchment scale is on the order of a few square kilometers, as opposed to the hundreds of square kilometers used in many lumped modeling applications. Use of the NHDPlus network allows for a distributed hydrologic model simulation that accounts for spatial variability in both the land surface and forcing functions within natural watershed units, rather than the arbitrarily defined grid cells commonly used for distributed hydrologic models.

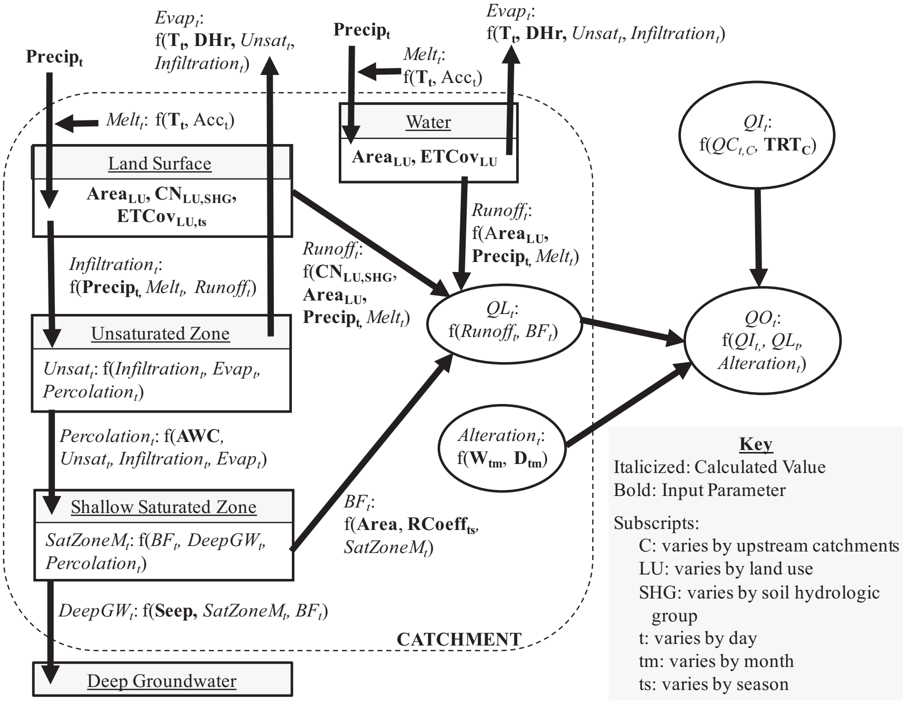


FIGURE 1. Schematic of the Model Components (italicized) and Input Parameters (bolded) Used to Simulate the Hydrologic Cycle within Each Enhanced National Hydrography Dataset (NHDPlus) Catchment (contained in dashed area) and then Routed throughout the Network within WaterFALL. Model components and parameters not fully spelled out are described within the text.

Rainfall-Runoff Model. The Generalized Watershed Loading Function (GWLF) is employed within WaterFALL to simulate rainfall-runoff processes within each NHDPlus catchment (Haith and Shoemaker, 1987; Haith *et al.*, 1992) (Figure 1). Numerous applications of GWLF exist throughout the world (Howarth *et al.*, 1991; Schneiderman *et al.*, 2002; Georgas *et al.*, 2009; Li *et al.*, 2010), including versions of the model that have been enhanced to suit different modeling needs (Dai *et al.*, 2005; Schneiderman *et al.*, 2007). GWLF simulates two components of surface water flow: (1) runoff (Runoff) generated through the curve number (CN) method and simple snowmelt (Melt) and accumulation (Acc) modules and (2) base flow (BF) that is released from the saturated zone over time through soil moisture (SatZoneM) and groundwater (DeepGW) accounting modules (Figure 1). Water moves through the different compartments represented by the model (*i.e.*, surface, unsaturated zone, shallow saturated zone, and deep groundwater) based on physical parameters and rate constants governed by process-based equations and mass balance. Rainfall can infiltrate into the unsaturated soil layer, or runoff to a surface water body. Infiltrated water resides in the unsaturated soil zone (Unsat) until it percolates to the saturated zone. The available water capacity (AWC) of the unsaturated zone controls the rate of percolation to the saturated zone. Evapotranspiration (Evap), as calculated within GWLF using the Hamon (1961) method based on temperature (T) and daylight hours (DHr) for potential evapotranspiration and cover coefficients (ETCov) based on land use type, in addition to percolation, can deplete water from the unsaturated layer. The saturated zone is depleted by local groundwater flow to surface water (*i.e.*, shallow base flow) controlled by the recession coefficient (RCoeff) and also by seepage (Seep) to a regional groundwater system. Both processes act as a linear function on the available saturated storage.

The WaterFALL application of GWLF relies on the original CN formulation for determination of runoff and infiltration (Haith and Shoemaker, 1987; Haith *et al.*, 1992), although with some modifications (Figure 1): (1) a two-stage linear recession based on the presence of percolation from the unsaturated zone and has been added to the base flow calculation and recession coefficients can now vary by season; (2) the hydrologic model is now run in a series over the catchments of the NHDPlus network with an embedded routing routine to accumulate and move water downstream; (3) methods to quantify the impacts of human interactions on streamflow have been added through withdrawals, discharges, and simple reservoir mass balance; and (4) model parameterization is

TABLE 1. Watershed and Hydrologic Models Relying on the Soil Conservation Service Curve Number (SCS-CN) Method for Runoff Estimation.

Model	Application	Time Step	Spatial Scale	Reference
AGNPS (Agricultural NonPoint Source pollution model)	Agricultural watershed simulation	Event-based	Uniform square areas (cells), some containing channels	Young *et al.* (1989)
AnnAGNPS (Annualized Agricultural NonPoint Source pollution model)	Agricultural watershed simulation	Daily or subdaily	Homogeneous land areas (cells), reaches, and impoundments	Bingner and Theurer (2001)
GWLF (Generalized Watershed Loading Function)	Water quantity and quality simulations	Daily	Subbasin network	Haith and Shoemaker (1987)
HEC-1/HEC-HMS	Urban watersheds, floods	Event-based	Distributed network; varying sizes	Feldman (1995) and USACE (2014)
SWAT (Soil and Water Assessment Tool)	Water quantity and quality simulations	Daily	Subbasin network	Arnold *et al.* (1998)
TR-20	Evaluation of flood events	Event-based	Watershed with land runoff and channel routing	NRCS (2015)
TR-55	Evaluation of flood events	Event-based	Small watershed consisting of subareas and reaches	NRCS (2009)

now automated through geospatial processing of datasets against the NHDPlus.

SCS-CN Formulation. The Soil Conservation Service Curve Number (SCS-CN) method approach to runoff estimation is a peer-reviewed and commonly applied method to simulate streamflow that falls in between simplistic and highly parameterized process-based models (Table 1). It is based on readily available land use and soils information, and therefore relates to the real-world engineering and planning needs of the professional community and provides the flexibility for scenario-based simulations at the catchment scale. The SCS-CN approach represents an intermediate level of complexity for hydrologic simulation and does not specifically identify water flow paths or runoff processes. Table 1 presents examples of the watershed and hydrologic models employing this method to represent both event and continuous simulation of streamflow. Borah and Bera (2003) included AGNPS, Ann AGNPS, HSPF, and SWAT in their model comparisons where they noted that none of these models sufficiently simulate flood waves, which were better simulated by event models using physically based flow-governing equations that require approximate numerical solutions of the equations that are subject to numerical instability problems and limited on space and time increments and watershed sizes.

Although the SCS-CN method garners wide governmental support in its use within a management context due to its real-world applicability, some members of the scientific community question the validity of its use (Garen and Moore, 2005; Walter and Shaw, 2005). Garen and Moore (2005) were critical of the SCS-CN method in relation to water quality simulations, largely due to the method's lack of specificity on the governing runoff process represented. However, with the focus on providing a catchment-scale water budget at the minimum of a daily time step, the objective of the current formulation of WaterFALL is to represent the general rainfall-runoff processes and water balance at the resolution of the NHDPlus catchment and stream segment, and then to compound the streamflow to larger watersheds. Therefore, the SCS-CN provides an appropriate and feasible level of simulation of the generalized runoff process of the hydrologic cycle, while allowing for the inclusion of additional functions and parameters to account for subsurface and base-flow processes.

GWLF Modifications. The original GWLF formulation relied on a linear storage-outflow model. However, Van de Griend *et al.* (2002) found that for groundwater discharge from a shallow unconfined aquifer, there are three main reasons that the assumption of a linear storage-outflow model may not hold: (1) a falling watertable continually decreases the effective thickness of the aquifer and decreases the ability to drain; (2) the hydraulic conductivity tends to decrease with depth; and (3) with prolonged drainage, the lower-order stream channels can run dry, leaving only the highest-order reaches receiving base flow. Brodie and Hostetler (2005) also noted that recession behavior for a stream can change through time due to factors such as catchment wetness, saturated aquifer thickness, or depth of stream penetration into the aquifer. Although the daily representation of the saturated soil zone moisture content (SatZoneM) across a distributed stream network addresses these points to an extent, particularly the differentiation between low- and high-order streams, additional accounting for the periods of decreasing watertable and for seasonality were needed after examination of low-flow period simulations using the original formulation within WaterFALL.

To account for slower drainage rates (*i.e.*, release to the stream channel) during periods of decreasing water table, a two-phase recession using the original RCoeff parameterization was instituted that reduces the recession rate during days without percolation into the saturated zone. The effect of this reduction is a lower rate of base flow on days when the saturated reservoir is not actively accumulating volume (Equation 1). The rate of reduction (Reduc) is currently set for an entire model simulation prior to the model application and held constant throughout a model simulation. The second modification to the linear storage-outflow model made within WaterFALL is to allow the RCoeff to vary by season. In testing the model formulation, the seasonal differences in model input parameters did not provide enough variation within the streamflow recessions to account for the potential processes highlighted by Brodie and Hostetler (2005). Equation (2) displays the simple modification made to minimize the number of model parameters needed to be estimated while allowing greater seasonality in the streamflow recession. Winter and summer RCoeffs (WinRCoeff and SumRCoeff, respectively) as well as a start and end month for the winter period become model input parameters, whereas the transition month RCoeff (TransRCoeff) is calculated as the average between the two rates. In watersheds where the seasonality in recession is accounted for through the seasonality in model inputs and soil moisture accounting, the start and end month for the winter period are set equal to one another and only a single RCoeff is used throughout the simulation period.

$$\mathrm{BF} = \begin{cases} \mathrm{RCoeff} \times \mathrm{SatZoneM}, & \mathrm{Percolation} > 0 \\ \mathrm{Reduc} \times \mathrm{RCoeff} \times \mathrm{SatZoneM}, & \mathrm{Percolation} = 0 \end{cases} \tag{1}$$

$$\mathrm{RCoeff} = \begin{cases} \mathrm{WinRCoeff}, & \text{Month Within Winter Period} \\ \mathrm{SumRCoeff}, & \text{Month Outside of Winter Period} \\ \mathrm{TransRCoeff}, & \text{Month=WinterStart or WinterEnd} \end{cases} \tag{2}$$

Streamflow Routing. WaterFALL employs a lag routing approach initially described by Linsley *et al.* (1975). This approach delays water from an upstream channel to the downstream network by a lag or travel time (TRT) and allows for the attenuation of the flow hydrographs. Upon initiating a model run, the routing algorithm with WaterFALL navigates all NHDPlus catchments upstream of the chosen pour point for the simulation (most downstream catchment). A TRT for each of the upstream catchments is then assigned to the pour point catchment based on the flowline length between the corresponding upstream catchment and the pour point, as well as the average velocity estimate for each upstream catchment provided within the NHDPlus Value Added Attributes (based on Jobson, 1996). Currently, WaterFALL uses this static TRT for each catchment to route the channel flows, and no additional attenuation is applied due to low- or high-flow days. The same process is used for any catchment within the network by calculating the TRT relative to the catchment of interest by subtracting that catchment's TRT to the pour point from any upstream catchment TRTs.

To create daily, routed streamflow throughout the NHDPlus network, WaterFALL completes the following operations moving from the most upstream catchment to the watershed pour point catchment: (1) complete the basic rainfall-runoff and base-flow hydrologic functions to calculate locally generated flow (QL) for all days of the model simulation; (2) pull out any relevant monthly values for withdrawals (W) or discharges (D) for the catchment; (3) calculate the cumulative outflow (QC) from the catchment considering lagged inflows, locally generated flow, and applicable withdrawals and discharges for each day of the model simulation; and (4) move to next downstream catchment and repeat Steps 1 through 3.

The routing of flows in Step 3 requires data storage of the daily values of QL, W, and D from the already processed upstream catchments within a database table of "n" columns (n = number of upstream catchments) by "d" rows (d = number of days in the model simulation). A database table of the TRT for each catchment to the pour point catchment is created upon each model simulation initiation. A TRT for each upstream catchment relative to the catchment of interest ($\mathrm{TRT}_{\mathrm{rel},n}$) in Step 3 is calculated and used to select the lagged upstream flows for each day of the simulation. Equations (3) and (4) illustrate this concept using matrix notation for catchment number four of a five catchment watershed (catchment 1 = headwater; catchment 5 = pour point), simulated for six days. For example, the streamflow from Catchment 3 takes 2.4 days to reach the outlet but less than one day to reach Catchment 4, so on Day 1 the streamflow out of Catchment 4 is the summation of local streamflows generated in Catchments 3 and 4.

$$\mathrm{TRT}=\begin{bmatrix}4.8\\3.2\\2.4\\1.5\\0\end{bmatrix},\mathrm{QL}=\begin{bmatrix}\mathrm{QL}_{1,1}&\mathrm{QL}_{2,1}&\mathrm{QL}_{3,1}&\mathrm{QL}_{4,1}&\mathrm{QL}_{5,1}\\\mathrm{QL}_{1,2}&\mathrm{QL}_{2,2}&\mathrm{QL}_{3,2}&\mathrm{QL}_{4,2}&\mathrm{QL}_{5,2}\\\mathrm{QL}_{1,3}&\mathrm{QL}_{2,3}&\mathrm{QL}_{3,3}&\mathrm{QL}_{4,3}&\mathrm{QL}_{5,3}\\\mathrm{QL}_{1,4}&\mathrm{QL}_{2,4}&\mathrm{QL}_{3,4}&\mathrm{QL}_{4,4}&\mathrm{QL}_{5,4}\\\mathrm{QL}_{1,5}&\mathrm{QL}_{2,5}&\mathrm{QL}_{3,5}&\mathrm{QL}_{4,5}&\mathrm{QL}_{5,5}\\\mathrm{QL}_{1,6}&\mathrm{QL}_{2,6}&\mathrm{QL}_{3,6}&\mathrm{QL}_{4,6}&\mathrm{QL}_{5,6}\end{bmatrix},$$

$$\mathrm{TRT}_{\mathrm{rel},4}=\begin{bmatrix}2.3\\1.7\\0.9\\0\\0\end{bmatrix} \tag{3}$$

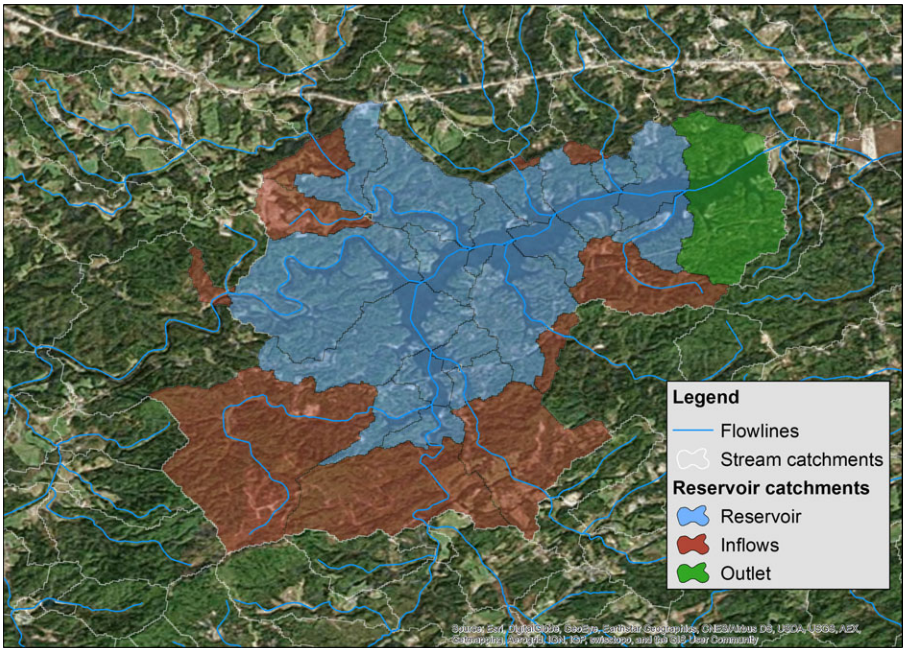
$$\begin{aligned}\mathrm{QC}_{4,1}&=\mathrm{QL}_{4,1}+\mathrm{QL}_{3,1}\\\mathrm{QC}_{4,2}&=\mathrm{QL}_{4,2}+\mathrm{QL}_{3,2}+\mathrm{QL}_{2,1}\\&\vdots\\\mathrm{QC}_{4,6}&=\mathrm{QL}_{4,6}+\mathrm{QL}_{3,6}+\mathrm{QL}_{2,5}+\mathrm{QL}_{1,4}\end{aligned} \tag{4}$$

This processing of lagged flows is completed for all catchments and for all days of the simulation resulting in a final set of daily cumulative streamflows (QC) for each catchment. When withdrawals and discharges are included within a simulation, they are tabulated by the same methods as shown in Equations (3) and (4) for QL for each day. Any upstream withdrawals (W) are subtracted (up to $W \leq \mathrm{QC}$) and then upstream discharges are added (D) before moving to the next downstream catchment in the navigation (Equation 5). If the $W \geq \mathrm{QC}$, then the available withdrawal is set equal to QC, QC is set equal to zero, and the available withdrawal is used in all subsequent routing calculations (see Appendix for further detail on withdrawals and discharges).

$$\mathrm{QC}=\sum \mathrm{QL}-\sum W+\sum D \tag{5}$$

The above procedure accounts for the water in the system through the catchments and conserves the volume of the water generated over the period of time by keeping a record of water that has not yet reached an outlet at a given time but adds the amount later. This routing algorithm is used in all WaterFALL simulations for all catchments, with exceptions only for control structures as described below.

To simulate altered streamflow conditions, WaterFALL accounts for point withdrawals and discharges from human managed systems (*e.g.*, public water supplies, wastewater treatment, industry, agriculture). To account for onstream impoundments (*i.e.*, reservoirs) or other control structures, WaterFALL allows for two options: (1) replacement of simulated streamflow at a selected NHDPlus catchment with an input daily time series of streamflows that account for the managed flows or (2) simulation of reservoir operations through mass balance. Reservoir simulation methods are described within the Appendix, while Figure 2 illustrates the set up for the mass balance simulation method.



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community

FIGURE 2. Depiction of a Reservoir within WaterFALL Using the NHDPlus Network.

Model Parameterization. The original GWLF formulation relied on numerous look-up tables to parameterize the model based on location and watershed conditions. WaterFALL now automates the determination of all of these parameters, or the initial values of these parameters, by georeferencing spatially-based physical data. The processes used to parameterize each catchment seek to preserve as much spatial variability in the available physical data as possible to best represent the variability in physical properties of the watershed (Reed *et al.*, 2004), thereby reducing the noise introduced into model parameters (Kling and Gupta, 2009). Within each catchment, the rainfall-runoff processes are simulated for each type of land cover. The smallest unit for parameterization therefore becomes each combination of land use and underlying soil type (Table 2). The most generalized parameterization applies to the catchment level and includes temperature, precipitation, growing season start and end dates, daylight hours, and parameters for AWC, RCoeff, and Seep that are set to represent all combinations of land use and soils within the catchment (Table 3).

A description of the data sources and processes commonly used to parameterize WaterFALL follows for the categories of climate, land use, and soils. Alternative data may be used in specific applications and are not specified here.

Climate. WaterFALL relies on daily temperature and precipitation to drive the rainfall-runoff processes. A daily, 4-km gridded climate dataset originally created for the Conservation Effects Assessment Project and obtained from the U.S. Department of Agriculture (USDA) for the period of 1960-2001 (and supplemented by researchers from University of Texas for the years 2002-2006) provided a spatially explicit representation of precipitation and temperature (DiLuzio *et al.*, 2008). Each catchment is assigned the daily time series of precipitation and average temperature for the grid cell which it intersects. If a catchment lies within multiple grid cells, an area-weighted average is used to produce the daily time series for the catchment from the intersecting grid cells. Figure 3 depicts an example of the gain in spatial disaggregation of the observational climate data from using this gridded dataset in place of standardized climate station assignments. In this example higher precipitation is observed over the upstream portions of the HUC8 with the headwaters and downstream portions receiving lower precipitation totals for the day. If using a single, long-term climate station representation from the National Climatic Data Center with Thiessen polygon assignments (solid lines) the entire simulated watershed would have received a single precipitation value for the day and the spatial variability in the event would have been lost.

Land Use. Any number of land use/cover types can be considered within WaterFALL, provided a geospatial layer is available for the categorization. Within a catchment, each area of land use is overlain with soils data and constitutes a unit over which

TABLE 2. An Example Model Parameterization Showing the WaterFALL Hydrologic Simulation Distribution over Two Catchments within NHDPlus and across Each Land Use within the Catchments.

Catchment ID	Land Cover	Area (km^2)	Percent Area (%)	Soil Hydrologic Group	Curve Number
8896894	Developed, open space	0.15	5.1	B	69
	Developed, low intensity	0.18	6.0	B	68
	Developed, medium intensity	0.11	3.7	C	82
	Deciduous forest	1.47	50	B	60
	Evergreen forest	0.56	19	B	60
	Mixed forest	0.15	5.1	B	60
	Shrub/scrub	0.012	0.4	B	57
	Grassland/herbaceous	0.085	2.9	B	70
	Pasture/hay	0.24	8.2	B	64
	Total area:	2.96			
8896900	Open water	0.008	0.3	A	100
	Developed, open space	0.073	3.5	C	79
	Developed, low intensity	0.069	3.3	C	79
	Developed, medium intensity	0.053	2.5	C	82
	Deciduous forest	0.76	36	B	60
	Evergreen forest	0.109	5.2	B	60
	Mixed forest	0.105	5.0	B	60
	Grassland/herbaceous	0.12	5.6	C	80
	Pasture/hay	0.75	36	B	64
	Cultivated crops	0.008	0.5	C	82
	Woody wetlands	0.057	2.8	C	100
	Total area:	2.11			

TABLE 3. WaterFALL Model Parameters and Definition Methods.

Model Parameter	WaterFALL Determination	Applied to	Reference (if applicable)
Land use categories/areas (LU)	Georeferenced from geospatial layer	Catchment	Fry *et al.* (2011)[1]
Curve number (CN)	Look-up table based on land use and soil hydrologic group	Land use category	USDA-NRCS (1986)
Cover coefficient (ETCov)	Look-up table based on land use category and growing season	Land use category	Haith *et al.* (1992)
Soil hydrologic group (SHG)	Georeferenced to each land use category within a catchment	Land use category	USDA-NRCS (2014)
Temperature (T)/ Precipitation (Precip)	Georeferenced from 4-km grid to catchments using area weighting	Catchment by day	DiLuzio *et al.* (2008)[1]
Start and end dates of growing season (ts)	Georeferenced from national geospatial layer of first and last freeze dates	Catchment	NCDC (2002)
Number of daylight hours (DHr)	Calculated based on latitude of catchment centroid and day of the year	Catchment by day	Forsythe *et al.* (1995)
Available water capacity (AWC)	Georeferenced by land use type and soil hydrologic group; Calibrated	Catchment	Zhang *et al.* (2011)
Recession coefficient (RCoeff)	Georeferenced by land use type and soil hydrologic group; Calibrated	Catchment	Zhang *et al.* (2011)
Seepage rate (Seep)	Calibrated (starting value based on best professional judgment using watershed geophysical conditions)	Catchment	Calibration parameter only
Winter start (WinterStart) and end (WinterEnd) month	Calibrated (starting values based on best professional judgment using observed streamflows)	Watershed	Calibration parameter only
Alterations (W, D)	Monthly value from available data sources	Catchment	Dependent on study application

[1] The referenced dataset is typically used for current condition scenario simulations and can vary depending on the scenario definition.

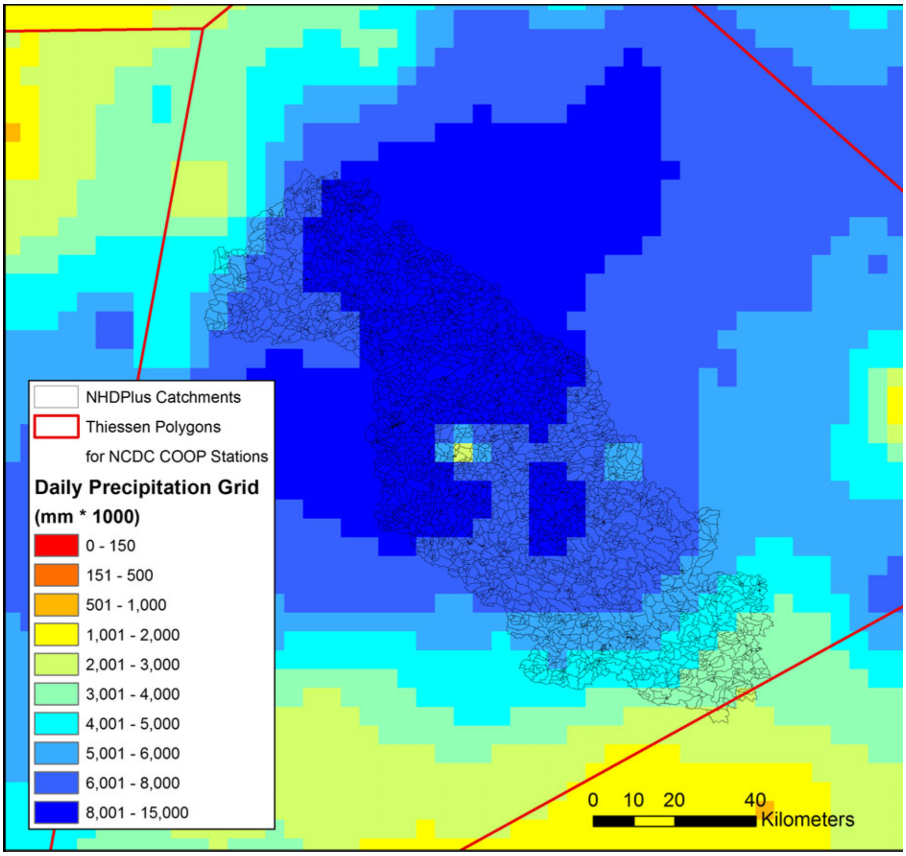


FIGURE 3. Depiction of a Single Day Precipitation Event over the NHDPlus Catchments That Make Up the Upper Neuse River 8-Digit Hydrologic Unit (HUC8). Gridded precipitation gradients are compared to single station values represented by Thiessen polygons.

runoff is simulated. Central to the SCS-CN method used for hydrologic simulation, a curve number is assigned to each land use type within each catchment based on the predominant underlying soil hydrologic group (USDA-NRCS, 1986). Additional characteristics required of land use, besides the basic type qualification, include the percent imperviousness for developed lands and the percent vegetative cover. Depending on the land use geospatial dataset used for a model run, these characteristics are available either as dataset attributes or qualifiers, or are estimated based on the land cover type. The majority of the WaterFALL simulations completed to date have been to simulate current, altered conditions within a watershed. Therefore, the most recent (at the time of model development) 2006 National Land Cover Database (NLCD) coverage was used (Fry *et al.*, 2011).

Soils. The hydrologic condition of the soils underlying each land use within each catchment is required to properly apply a curve number for runoff calculations. In addition, the subsurface characterization plays a role in determining how fast and at what magnitude water will move through the soils and either enter the deeper groundwater or the stream channel. For WaterFALL, the Soil Survey Geographic Database (SSURGO) dataset (1:12,000 to 1:63,630 scale) is used to preserve local variances in soils conditions that are better suited to the NHDPlus catchment-scale (1:100,000) analysis (USDA-NRCS, 2014). From each intersection of land cover and soil map unit with corresponding components, attributes related to the soil hydrologic group; percent sand, silt, and clay; AWC; and slope were tabulated from SSURGO.

To further characterize the model parameters related to the rate of water movement through the subsurface, tabulated parameters created by the National Weather Service (NWS) for applications of the Sacramento Soil Moisture Accounting Model (SAC-SMA) were employed as a starting point for calibration of the RCoeff and the AWC of the unsaturated zone. The SAC-SMA parameterization processed soils (SSURGO) and land use (2001 NLCD) geospatial data layers on a 4.67-km grid scale across the contiguous U.S. (Anderson *et al.*, 2006; Zhang *et al.*, 2011). Similar to the climate parameters, the SAC-SMA gridded dataset was overlain on the NHDPlus catchments and geoprocessed to extract the necessary parameters by catchment. Within the SAC-SMA, a two-layer subsurface is depicted (Burnash, 1995), which is similar, although more complex, in terms of the simulated processes to the subsurface within GWLF. The related parameters within the SAC-SMA, the upper zone free water capacity (mm) and the depletion rate of the lower layer primary free water storage (day^{-1}) (Anderson *et al.*, 2006), are used as *a priori* estimates of the AWC and RCoeff, respectively, within WaterFALL. These *a priori* soils parameters are determined from the national dataset and are further adjusted with a regional scaling multiplier during calibration as discussed below.

Data Management. The key to WaterFALL's capability to perform, save, and analyze multiple-model runs across, in some areas, tens of thousands of NHDPlus catchments is a robust data management system. WaterFALL's infrastructure consists of the following data management elements: (1) a central database consisting of multiple tables to store input data and results from simulations; (2) a model code to perform time-varying calculations per catchment and to conduct routing procedures; (3) a web interface to allow visual navigation of the NHDPlus network and basic display of model results; (4) a graphical user interface (GUI) to perform model scenario set up; and (5) web services to communicate among the database, the model code, the web interface, and the GUI.

The benefit of having WaterFALL's infrastructure contained within a single, server-based environment is that one database contains all model data and simulations rather than having separate instances employed across various user desktop computers with files distributed on *ad hoc* basis. Each model simulation is tracked by a unique identifier that links to the metadata for the run and all results and intermediate data elements. The metadata saved for each model simulation indicates the model version, pour point catchment, inclusion of human alterations and/or control structures, the land use and climate datasets chosen, and the choice of user-identified or stored calibration parameters. This metadata allows for either recreation of the same model scenario, if needed, or for selection of the scenario for comparison to another during post-processing.

Model Calibration

Calibration Parameters. The simplest calibration process involves the adjustment of only three WaterFALL parameters (AWC, RCoeff, and Seep). When using seasonal recession coefficients, three additional parameters are included in the calibration (WinRCoeff, WinterStart, and WinterEnd) and the RCoeff becomes the SumRCoeff. The AWC and RCoeff parameters have initial values pre-processed by catchment within the WaterFALL database. These two parameters were found by Kling and Gupta (2009) to exhibit less noise in their parameterization across example models of varying system complexity because they could be easily identified through analysis of observed runoff signals. Because of the physical basis of these two parameters, *a priori* values for the parameters are indexed to individual catchments within the WaterFALL database based on data compiled by the NWS for the SAC-SMS (Table 3). When calibrating each of these parameters the actual value being adjusted is a multiplier that is applied to the *a priori* catchment-based values. This method of calibration attempts to avoid introducing noise that can cloud existing relationships between model parameters and underlying physical properties by including the original spatial heterogeneity in physically based parameters, but simplifying calibration to regional, watershed-based multipliers for AWC and RCoeff. This method facilitates the distributed modeling over the dense catchment network. The introduction of such noise with the regionalization of model parameters was identified by Kling and Gupta (2009) as a tradeoff between the simplicity of lumped models and the spatial variability but increased data needs of distributed modeling in simulating ungaged basins, which WaterFALL bridges by using available data at an appropriate scale over the combination of *a priori* values and the NHDPlus-based hydrologic network.

The third calibration parameter, Seep, controls the amount of water released from the saturated subsurface into the deep groundwater aquifer. This release constitutes a loss from the system, where the water is no longer available to reach the stream in the temporal context of daily rainfall-runoff modeling. Seepage is controlled in part by the geology within a region and the extent to which the groundwater and surface water are connected. Although related to the geology, the existing national-scale geologic characterization does not provide enough information to determine quantitative values on which to base this parameter. Therefore, a subbasin-specific Seep value is determined completely through calibration, although initial values are guided by the general geology of a region. For example, in areas underlain by unconfined aquifers or karst, Seep may be set at a high value of 0.1% to begin calibrations; whereas in areas with shallow bedrock or confining layers, Seep would be set at 0 to begin calibrations.

When including seasonality in the RCoeff within a WaterFALL simulation, the original RCoeff value included in the WaterFALL database is applied to the summer months and a different multiplier value is determined through calibration for the winter months. These two RCoeff values (RCoeff/SumRCoeff and WinRCoeff) can be determined through the automated calibration methods. The months in which to apply the WinRCoeff are set manually in a calibration/simulation run by providing a start (WinterStart) and end (WinterEnd) month. These selected months can typically be determined by examining mean monthly observed streamflows at calibration locations.

Calibration Process. The calibration process for the current WaterFALL system consists of selecting calibration locations based on available observed streamflow data, selecting validation locations, performing automated calibration model runs, performing secondary manual calibration when necessary, and extrapolating calibration parameters to uncalibrated (*i.e.*, typically ungaged) stream reaches. Description of each of these steps follows.

Location Selection. WaterFALL can be calibrated at any NHDPlus catchment where observed streamflow records exist. Parameter adjustments made based on comparison to observed data during a calibration run are used to estimate the parameter adjustments needed in ungaged regions of a watershed based on the major characteristics used in the original selection of the locations. Therefore, the overall goal for selecting calibration locations is to achieve representation of as many different aspects of the watershed as possible, while also maintaining a reasonable level of effort and maximizing available observed streamflow values reported to be of good quality.

Automated Calibration. A customized version of the Parameter Estimation Tool (PEST) (Doherty, 2010) has been set up to interact with WaterFALL and calibrate the parameters through an iterative process. PEST uses a nonlinear estimation technique known as the Gauss-Marquardt-Levenberg method (Doherty and Johnston, 2003). The strength of this method lies in the fact that it can generally estimate parameters using fewer model runs than any other estimation method. Once interfaced with WaterFALL, PEST's role is to minimize the weighted sum of squared differences between model-generated values and streamflow gage observations; this sum of weighted, squared, model-to-measurement discrepancies is referred to as the "objective function." Depending on the purpose of the model, different objective functions are available for use within WaterFALL's calibration process:

1. Minimize log-transformed differences in daily flows
2. Minimize differences in daily flows
3. Minimize log-transformed differences in monthly total streamflow
4. Minimize differences in monthly total streamflow.

The current calibration process is set up to examine all model parameters within a single calibration run using a single selected objective function. Based on a set of initial values and lower and upper bounds listed by parameter, PEST finds the optimum set of calibration parameters for a simulation. These parameters are applied to all catchments within the chosen calibration simulation. There may be multiple calibrations within a watershed. The interaction between WaterFALL and PEST does not currently allow for consideration of nested calibration parameters meaning that even if an upstream gage has already been calibrated and optimum parameters determined, calibration at a downstream gage is done independently and a potentially new set of parameters are determined for all catchments. When compiling the final set of parameters for the overall watershed the upstream parameter set would be combined with the downstream parameter set.

Manual Calibration. Because the WaterFALL calibration process relies on a single-stage automated evaluation of a single objective function, it is possible, based on the chosen function, to miss key elements of the hydrologic regime. Therefore, after initial automated calibration and evaluation, manual calibration of model parameters is undertaken as needed to improve the overall model fit. For instance, when using the minimization of log-transformed differences in daily flows the model tends to underpredict extreme high flows while better representing low flows. This underprediction is easily recognized by large percent bias measures. Manual adjustment may result in slight increases in the difference between daily flows but a larger improvement in the percent bias. The same model parameters are considered as in the automated calibration process.

Extrapolation to Uncalibrated Reaches. After each individual calibration is complete, the final calibrated parameters are extrapolated to the remaining uncalibrated (*i.e.*, typically ungaged) catchments of the watershed based on the major characteristics used in the original selection of the locations informally following the various techniques available in the published literature on transference of parameters (*e.g.*, Merz and Blöschl, 2004; Wagener *et al.*, 2004; Parajka *et al.*, 2005; Hrachowitz *et al.*, 2013). Characteristics considered include drainage area, coincident subbasin, major land use classifications, topology, and/or condition (*i.e.*, reference or highly altered). Extrapolation of the AWC and RCoeff parameters, which are actually multipliers on physical data available for each catchment as previously described, reduces the uncertainty for ungaged streams and preserves the heterogeneity of the physical basis of the parameters across the NHDPlus catchments spanning different soils and land use combinations as shown by Anderson *et al.* (2006). Extrapolation of parameters is possible based on (1) the way the underlying data within WaterFALL have been parameterized and (2) the design of the calibration process and parameters. As a result of WaterFALL's distributed framework, the number of catchment units modeled in a full watershed is very large. Calibration of parameters for each catchment individually would be computationally prohibitive. The balance between the spatial granularity of the modeling and the computational requirements for calibration is accomplished through the use of the intermediate-level hydrologic model.

Model Performance

Classic hydrologic goodness of fit statistics used in model performance evaluation include percent bias (PBIAS) and the Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) (Equations 6 and 7). The PBIAS quantifies the percent difference in total (summed) daily/monthly/annual volume of observations versus model estimates. The NSE ranges from $-\infty$ to 1, where a value of 0 indicates that the model predictions are as accurate as the mean of the observed data. A negative NSE value indicates that the residual variance is larger than the data variance. When evaluating model performance on a daily basis, both of these measures are disproportionately impacted by large storm events, where the residual (*i.e.*, difference between observation and model) for a single day with peak flow will cause a larger reduction in these quantitative metrics than a difference in a day with low flow. Therefore, a modified version of the NSE, rNSE, described in Krause *et al.* (2005) is also included in the evaluation (Equation 8). In rNSE the differences between the observed and modeled values are quantified as relative deviations which reduces the influence of the absolute differences during high flows, although absolute differences in low flows are enhanced. This metric is therefore sensitive to systematic over- or underprediction.

$$\mathrm{PBIAS} = \frac{\sum_{t=1}^{n} S_t - \sum_{t=1}^{n} O_t}{\sum_{t=1}^{n} O_t} \times 100 \tag{6}$$

$$\mathrm{NSE} = 1.0 - \frac{\sum_{t=1}^{n} (S_t - O_t)^2}{\sum_{t=1}^{n} (O_t - \mu_o)^2} \tag{7}$$

$$\mathrm{rNSE} = 1.0 - \frac{\sum_{t=1}^{n} \left(\frac{S_t - O_t}{O_t}\right)^2}{\sum_{t=1}^{n} \left(\frac{O_t - \mu_o}{\mu_o}\right)^2} \tag{8}$$

where S_t is the model-simulated flow time series; O_t is the observed flow time series; and μ_o is the mean (average) of observed flow.

Numeric thresholds for evaluation of model performance for the SWAT model were suggested by Moriasi *et al.* (2007). These thresholds state that monthly NSE values greater than 0.50, 0.65, and 0.75 are indicative of satisfactory, good, and very good model performance, respectively, whereas PBIAS in mean streamflow within ±25, 15, and 10% were considered indicative of satisfactory, good, and very good hydrological model performance, respectively (Moriasi *et al.*, 2007). The thresholds established for the SWAT model are suggested for use as a generalized guidance for the traditional performance parameter evaluation of WaterFALL given the similarity of runoff mechanisms between the two models.

A secondary evaluation of model performance relies on bias in individual hydrologic metrics that are the basis for multiple management efforts. Hydrologic metrics, which have been described by a number of hydrologic and ecologic studies and applications (Richter *et al.*, 1997; McManamay *et al.*, 2012; Knight *et al.*, 2013; Murphy *et al.*, 2013), describe different aspects of the magnitude, frequency, duration, timing, and rate of change in streamflow. Suggested performance bounds on these metrics typically used in similar studies are ±30% (Murphy *et al.*, 2013; Caldwell *et al.*, 2015).

APPLICATION TO NORTH CAROLINA

In support of efforts of the North Carolina Ecological Flows Science Advisory Board (EFSAB), WaterFALL was parameterized and calibrated to simulate statewide streamflows. The statewide application covered 13 major river basins (22 separate watersheds) that are either entirely within the state or have some portion of their headwaters within the state. The total drainage area modeled was 101,888 km^2 (39,339 mile^2), which was represented by 58,702 NHDPlus catchments. The pour points for the basins were determined as the point closest to the border of the state that allowed for coverage of the watershed within the state or, for coastal watersheds, the point at which tidal influences begin. The available simulation period included the years 1960-2006. A one-year spin-up period was included in all calibration and validation model runs to eliminate the effect of the initial conditions of water storage within the subsurface and initialize watershed routing. A five-year calibration period was used based on the most recent five years of model input data. The five years preceding that period were used for model performance validation at calibration locations. A 30-year model simulation period was used to evaluate model performance based on hydrologic metrics.

Model Application

The model simulation was set up using the model inputs described in Table 3. Specifically, land use from the NLCD 2006 and climate data ranging from 1960-2006 were used to simulate recent hydrologic conditions. Water uses in the form of withdrawals and discharges to the stream network were compiled from state databases provided by the North Carolina Department of Environment and Natural Resources' (NCDENR's) Division of Water Resources characterizing: public water supply (PWS) withdrawals; nonpublic water supply (NPWS) withdrawals and discharges (*e.g.*, industrial uses, electricity generation); coastal nonpublic water supply withdrawals (Coastal NPWS); and National Pollutant Discharge Elimination System (NPDES) permitted discharges. These data account for the permitted human alterations to the natural flow regime from industry, public water supply, wastewater treatment, and agriculture. In most of North Carolina, nonagricultural water users must report their withdrawal amount if it is >100,000 gallons per day (379 m^3 per day), and agricultural users are required to report water use of above 1 million gallons per day (3,785 m^3 per day). In the 15 coastal counties that make up the Central Coastal Plain Capacity Use Area, users that withdraw over 10,000 gallons per day (38 m^3 per day) must register and report annual water use (NCDWR, 2013). These datasets represent the best available data on human alterations in North Carolina. The data are limited by permit regulations; therefore, small water withdrawals, especially for agricultural irrigation, may be underrepresented.

Human alterations from major dams were also included in the simulation as control structures (the reservoir functionality was not available at the time of these simulations). In the state of North Carolina, there are 11 major dams directly upstream from USGS gages with time-series flow data. The major control structures were selected if there was a downstream gage <1.21 km (0.75 miles) away and if there were no other large tributaries entering between the control structure and the gage (with the exception of very minor flows from 1 or 2 small tributaries). Minor dams that were not on a major waterway or that had <10,000 acre-ft (12.3 Mm^3) of normal storage capacity were not considered control structures and were modeled as run of river water bodies. The spatial location of the control structures was derived from the National Inventory of Dams dataset and checked for accuracy against USGS topographic maps. Each selected control structure was assigned to a unique NHDPlus catchment. Time-series flow data below the control structures were used in place of WaterFALL-derived streamflow data for the corresponding NHDPlus catchment during routing.

The GAGESII dataset lists 272 USGS streamflow monitoring gages within North Carolina, 221 of which were active in 2009 (Falcone *et al.*, 2010). Twenty-eight gages throughout the state, with watersheds ranging in size from 38.3 to 2,727 km^2, were selected for model calibration (Figure 4). Gage selection was based on available period of record of observations (at least 20 years of data through 2006), size distribution, watershed distribution, and inclusion within the stream classification systems developed by Hendriksen and Heasley (2010) and McManamay *et al.* (2012). Inclusion within these classification systems was necessary as one of the intended applications of this statewide application is an evaluation of such systems. The calibration locations were also selected to capture the range of hydrologic conditions expected within each watershed; therefore, large and small watersheds and reference and nonreference condition gages were chosen where possible within each major river basin simulated. Corresponding validation locations, where the model was not calibrated and instead calibrated parameters from nearby or physically similar gages were applied, were chosen using the same criteria (Figure 4), resulting in 20 comparison locations for model validation.

The objective function used during calibration minimized the differences in log-transformed daily flows, which gives equal weight to differences in streamflows at the low end of the hydrograph compared to the high end of the hydrograph. This type of model calibration objective often results in better representation of low flows at the expense of potentially underestimating peak streamflows. The water years of 2002 through 2006 were used for calibration as those years represented the best correspondence to model input data related to land use and human water uses. These years also represented a range of wet and dry years throughout the state. For these calibration gages, a validation of model performance in the uncalibrated period of 1997-2001 (water years) was used to examine the model performance through time. For the validation locations, the entire period of 1997-2006 was used to evaluate model performance for uncalibrated locations. A 40-year period (water years 1967-2006) was used to calculate hydrologic metrics described by Hendriksen and Heasley (2010) and McManamay *et al.* (2012) as a secondary set of model performance measures. For these measures the bias computed as the percent difference between the observed and model hydrologic metric was calculated and compared to hydrologic uncertainty bounds of ±30% (Murphy *et al.*, 2013; Caldwell *et al.*, 2015).

Calibration Results

Calibration results for the 28 gage locations are presented in Table 4 using the hydrologic performance measures of NSE, rNSE, and PBIAS for the 5-year calibration period, 5-year validation period, and 40-year validation period used to calculate hydrologic metrics. In addition to calculation of the NSE and rNSE at a daily time step, these two evaluation parameters were also examined for a three-day moving average to assess the impact of timing of

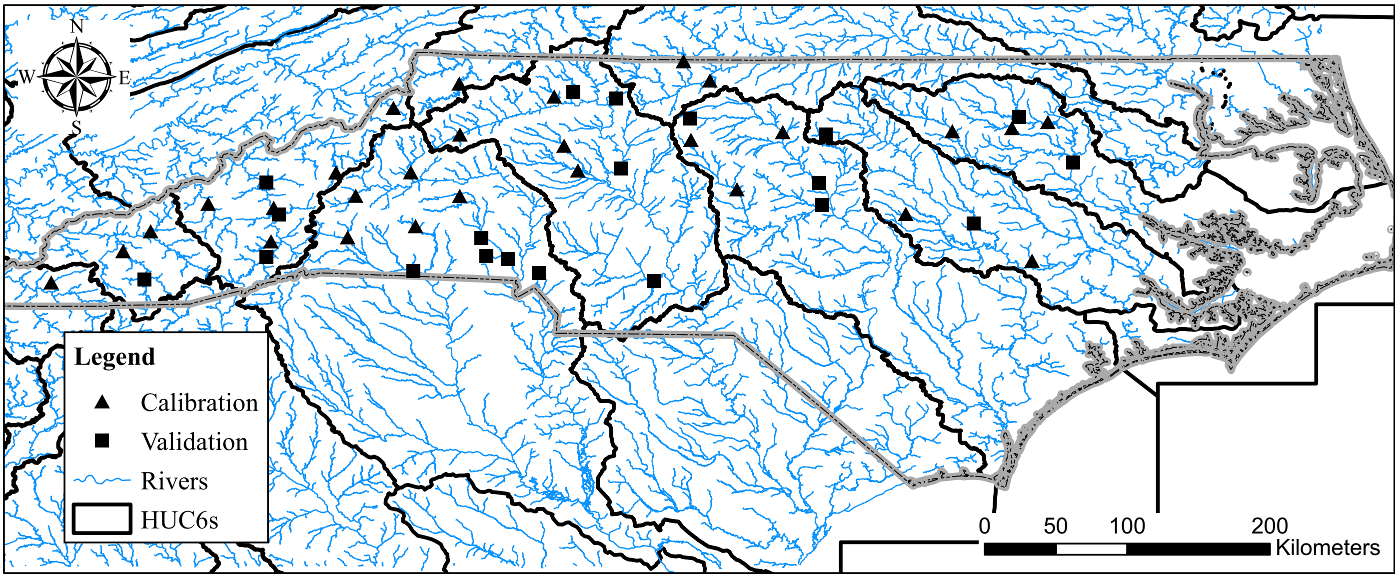


FIGURE 4. Location of Calibration and Validation Gages for the North Carolina Statewide Application of WaterFALL.

TABLE 4. Model Performance at Calibration Gages across North Carolina.

				Calibration (WY 2002-2006)					Validation (WY 1997-2001)					40 Years (WY 1967-2006)	
				All Daily Flows			Three-Day Average		All Daily Flows			Three-Day Average		Monthly Median	
Watershed	**Gage ID**	**Status**	**Drainage Area (km^2)**	**NSE**	**rNSE**	**PBIAS**	**NSE**	**rNSE**	**NSE**	**rNSE**	**PBIAS**	**NSE**	**rNSE**	**NSE**	**PBIAS**
Roanoke	02070500	Ref	627	0.52	0.77	2.7	0.62	0.69	0.47	0.68	18.1	0.54	0.65	−0.05	23.2
	02071000	Nonref	2,727	0.54	0.6	−8	0.65	0.47	0.48	0.78	−3.1	0.61	0.75	0.51	21.1
Tar	02081747	Nonref	1,106	0.4	0.73	−30	0.45	0.67	0.39	0.29	−21.4	0.5	0.54	0.31	−20.6
	02082770	Nonref	430	0.38	0.55	−21	0.43	0.51	0.28	0.86	−19.2	0.36	0.85	0.78	−5.5
	02083000	Nonref	1,362	0.48	0.7	−19.7	0.53	0.73	0.38	0.84	−19.8	0.44	0.92	0.76	−6
Neuse	02088000	Nonref	216	0.21	0.86	−16	0.33	0.84	0.33	0.58	−3.6	0.51	0.54	0.12	49.3
	0208925200	Nonref	149	−0.35	0.45	0.9	0.22	0.59	0.1	0.74	−5.3	0.4	0.88	0.62	−3.6
Cape Fear	02096500	Nonref	1,570	0.51	0.81	−27.1	0.65	0.83	0.45	0.86	−23.9	0.57	0.86	0.82	7.4
	02099000	Nonref	38	0.56	0.94	−39.5	0.59	0.91	0.22	0.93	−51.3	0.25	0.92	0.55	5.6
	02100500	Nonref	904	0.58	0.9	−22.7	0.68	0.91	0.38	0.82	−17.7	0.54	0.81	0.82	6.6
Yadkin-Pee Dee	02111180	Ref	125	0.3	0.8	−11.2	0.57	0.86	0.4	0.78	2.7	0.66	0.8	0.76	−4.1
	02112360	Ref	204	0.12	0.62	−9.9	0.5	0.74	0.41	0.76	−5.5	0.61	0.81	0.52	−18.7
	02118000	Nonref	793	0.47	0.09	−1.8	0.58	−0.09	0.65	0.84	−5.8	0.72	0.83	0.74	−4.1
	02118500	Ref	401	0.45	0.7	−0.8	0.59	0.59	0.57	0.79	3.4	0.7	0.77	0.39	24.1
Catawba-Wateree	02137727	Ref	326	0.22	0.9	−4.1	0.55	0.87	0.36	0.79	15.3	0.61	0.78	0.75	−0.2
	02140991	Ref	521	0.54	0.81	8.3	0.64	0.72	0.34	0.37	42.3	0.35	0.25	0	32.5
	02143000	Ref	215	0.57	0.87	−6.2	0.71	0.84	0.56	0.84	−4.2	0.68	0.81	0.76	−4.1
Broad	02149000	Ref	205	0.2	0.78	−2.2	0.59	0.76	0.13	0.66	9.5	0.6	0.69	0.68	−7
	02152100	Ref	157	0	0.85	−13.1	0.48	0.81	0.44	0.82	−11.2	0.61	0.79	0.75	−2.8
New	03161000	Ref	531	0.44	0.9	−2.5	0.63	0.91	0.67	0.87	3.8	0.78	0.88	0.82	2.6
French Broad	03446000	Nonref	173	0.14	0.63	0.1	0.56	0.76	0.34	0.46	10	0.61	0.53	0.69	−7.2
	03451500	Nonref	2,448	0.69	0.87	−2.9	0.75	0.87	0.83	0.89	1.4	0.87	0.9	0.83	−1.3
	03459500	Nonref	906	0.57	0.9	−14.2	0.62	0.88	0.74	0.86	−3.9	0.81	0.89	0.79	−3.8
	03463300	Ref	112	0.31	0.9	−12.2	0.4	0.93	0.09	0.63	2.6	0.65	0.68	0.68	−9.9
Watauga	03479000	Ref	239	0.33	0.54	−1.3	0.57	0.75	0.38	0.69	7.6	0.58	0.71	0.67	2.3
Little Tennessee	03503000	Nonref	1,129	0.5	0.64	25.2	0.5	0.63	0.4	0.49	34.3	0.43	0.53	0.48	24.4
	03512000	Nonref	477	0.05	0.2	23.9	0.39	0.48	0.05	0.2	23.9	0.39	0.48	0.64	9.6
Hiwassee	03550000	Nonref	269	0.24	0.68	−11.5	0.56	0.81	−1.08	−0.01	16.5	−0.2	0.47	0.8	−6.4

Note: NSE, Nash-Sutcliffe efficiency; rNSE, modified NSE; PBIAS, percent bias.

streamflow on the model performance. Because of the reliance on the calibration objective function to minimize the difference in log-transformed daily streamflows, and therefore give more weight to predictions at low-flow conditions over high-flow conditions, there was an expectation that some extreme high flows would be underpredicted, impacting the performance measure of NSE. These impacts were in fact seen as evidenced by the differences in NSE and rNSE where the median daily values among all gages during the calibration period were 0.42 and 0.78, respectively. Daily rNSE ranged from 0.09 to 0.94 for the calibration period. As anticipated, PBIAS shows an underprediction of flow volume on average across the calibration gages with a range from −40 to 25% and a median value of −7.1%. Use of the three-day average to compare NSE and rNSE provides evidence that the model performance is impacted by timing and magnitude of peak flows — in most instances of the calibration period, when a significant improvement was seen in the NSE from the daily to the three-day average value, the rNSE value remained relatively unchanged.

Validation of the model performance through time was completed by comparing performance measures from the calibration period to the validation period. Across the set of gages there were no major differences in the daily fit statistics of NSE and rNSE (median values of 0.39 and 0.78, respectively) and the median PBIAS improved to −0.9%, although the range in PBIAS increased due to two gages with significant changes in PBIAS (02099000 further decreased in prediction from −39.5% to −51%; 02140991 increased from 8.3% during calibration to 42% during validation). Figure 5 displays example monthly hydrographs and fit statistics between the two time periods for Henry Fork, a mid-range sized watershed (215 km^2), within the Catawba River watershed.

The long-term model validation for the calibration sites necessary for calculating hydrologic metrics was completed at a monthly time step for comparison with the thresholds set by Moriasi *et al.* (2007). For the median monthly NSE, the percentage of sites ranked as very good, good, and satisfactory was 43, 18, and 21%, respectively. Evaluation of PBIAS over the 40-year period resulted in 71% ranked as very good performance, no sites in the ±10 to ±15% range for a ranking of good, and 21% ranked as satisfactory. In each case, the majority of sites met a performance ranking of good to very good.

Model Validation at Uncalibrated Locations

Validation results for the 20 gage locations that were uncalibrated and received calibration parameters from nearby and physically similar calibrated gages are presented in Table 5 using the same hydrologic performance measures over the 10-year validation period for daily and three-day average streamflows.

TABLE 5. Model Performance at Uncalibrated Gages across North Carolina.

				All Daily Flows (WY 1997-2006)			Three-Day Average (WY 1997-2006)		Monthly 40 Years (WY 1967-2006)	
Watershed	**Gage ID**	**Status**	**Drainage Area (km^2)**	**NSE**	**rNSE**	**PBIAS**	**NSE**	**rNSE**	**Median NSE**	**PBIAS**
Tar	02082950	Ref	458	−0.03	0.1	−25.9	0.35	0.05	0.76	0
	02083500	Nonref	5,654	−36.92	0.51	−7.9	−20.28	0.72	0.66	−26.6
Neuse	02085000	Nonref	171	0.23	−51.37	−3.8	0.49	−21.32	0.1	50.6
	02088500	Nonref	601	0.37	−29.11	−4.6	0.51	−179.8	0.73	10.3
Cape Fear	02093800	Nonref	53	0.34	0.88	−40.3	0.41	0.81	0.3	−27.3
	02096960	Nonref	3,302	0.53	0.87	−19.5	0.64	0.86	0.81	4.8
	02102000	Nonref	3,714	0.5	0.71	−17.5	0.56	0.73	0.61	34.1
Yadkin-Pee Dee	02113000	Nonref	332	0.46	0.9	−15.3	0.58	0.87	0.68	−8.4
	02114450	Nonref	111	0.08	0.64	10.4	0.42	0.63	−0.13	42.9
	02116500	Nonref	5,905	0.48	0.06	6.1	0.69	0.88	0.81	8.4
	02126000	Nonref	3,553	0.56	0.88	−1.2	0.57	−0.24	0.19	81.9
Catawba-Wateree	02143500	Nonref	179	0.57	0.17	13.4	0.66	−0.14	0.13	25.6
	02144000	Nonref	82	0.45	−0.63	4	0.59	−1.3	0.39	16.2
	02145000	Nonref	1,627	0.71	0.73	6.5	0.76	0.71	0.36	20.3
	02146300	Nonref	80	0.45	0.87	−10.2	0.49	0.79	−0.66	62.3
Broad	02151500	Nonref	2,266	0.72	0.75	8.9	0.8	0.73	0.82	1.3
French Broad	03443000	Nonref	767	0.58	0.76	−1.5	0.67	0.84	0.76	−9.8
	03451000	Nonref	337	0.53	0.69	24	0.66	0.59	0.64	23.3
	03453000[1]	Nonref	409	1	1	3.4	1	1	1	3.7
Little Tennessee	03500240	Ref	148	−0.3	−0.14	20	0.27	0.36	0.49	16.8

[1] This gage represents a control structure within the model.

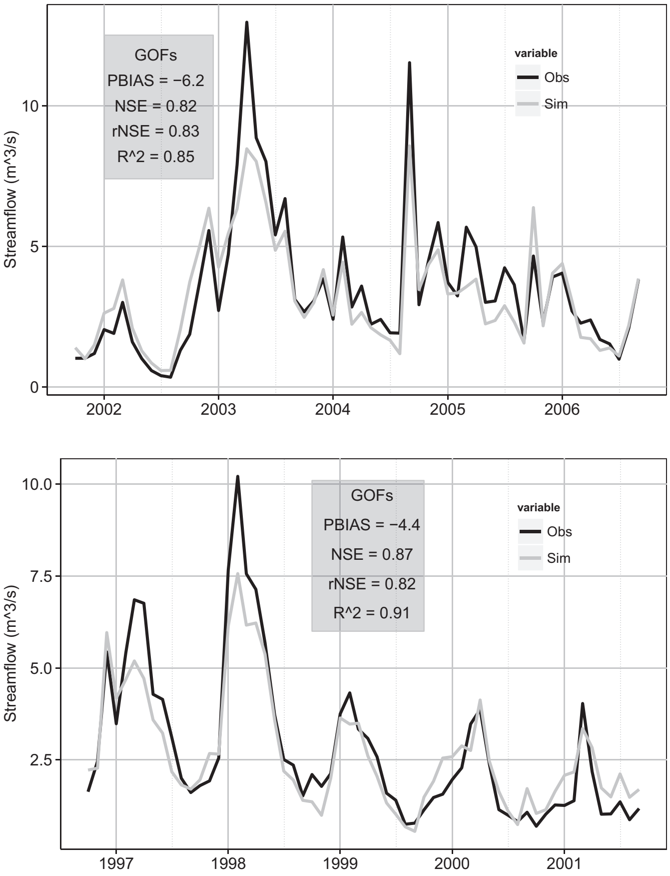


FIGURE 5. Example Model Performance during Calibration (top) and Validation (bottom) Periods at Calibration Gage 02143000 Henry Fork near Henry River, North Carolina. PBIAS, percent bias; NSE, Nash-Sutcliffe efficiency; rNSE, modified NSE.

Monthly NSE and PBIAS are presented for the 40-year validation period used to calculate hydrologic metrics. Included within the validation locations is gage 03453000 (Ivy River near Marshall, North Carolina), which was used as a control structure input into the model to represent power plant regulation at Ivy Dam approximately 0.64 km upstream from the gage. The validation statistics for this location are used to verify the control structure function within the model (NSE = 1 for all evaluations and minor PBIAS due to routing of gaged flows from the dam to the actual gage location). These statistics are excluded from further performance summaries.

Model performance at validation locations is highly site specific. There are gage locations at which the validation performance measures are comparable to the highest performing calibration locations; however, there are also locations with poor performance measures. The median rNSE across daily values for the 10-year period was 0.69 with a range from −51 to 0.9. Similarly, for PBIAS the median value across validation locations was −1.5% with a range from −40 to 24%. Similar to calibration locations, the three-day average NSE was typically an improvement over the daily NSE, while the three-day rNSE remained unchanged or even degraded indicating an issue with the model representation of high flows. A monthly hydrograph for gage 02143500 (Indian Creek near Laboratory, North Carolina), which received calibration parameters from calibration gage 02143000 as both locations were small tributaries in the same subwatershed of the Catawba River, is depicted in Figure 6.

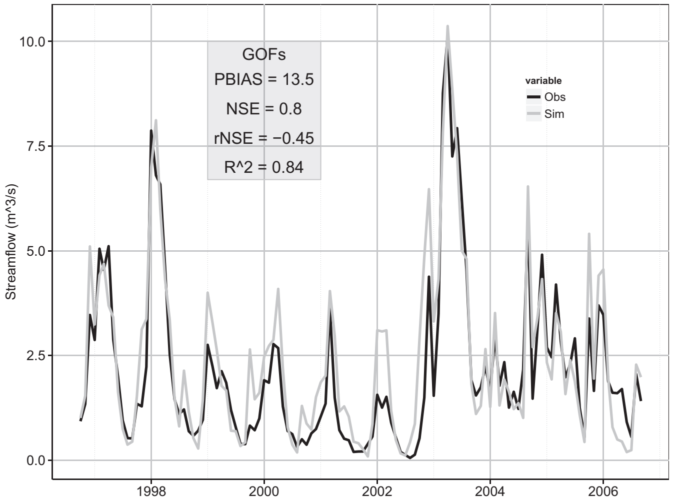


FIGURE 6. Example Model Performance during Validation Period at Uncalibrated Gage 02143500.

The long-term model simulation performance for calculating hydrologic metrics at the validation locations was again evaluated at a monthly time step against thresholds set by Moriasi *et al.* (2007). For the median monthly NSE the percentage of sites ranked as very good, good, and satisfactory were 30, 20, and 5%, respectively. Evaluation of PBIAS over the 40-year period resulted in 40% ranked as very good performance, no sites in the ±10 to ±15% range for a ranking of good, and 20% ranked as satisfactory. For the uncalibrated sites, the majority met a performance ranking of satisfactory.

Validation through Hydrologic Metrics

In support of environmental flow development, McManamay *et al.* (2012) developed a river classification for eight states in the Southeastern U.S. using nine flow metrics to identify eight stream classes, although two stream classes had two or less members. To provide a simpler method of classification the authors created a classification tree using five variables, which was able to correctly classify 85% of the gages in the study. The metrics and corresponding thresholds used within the classification tree include 0.41 and 0.63 for mean September flow, 0.32 for mean minimum July flow, 1.99 for mean maximum November flow, and 219 for coefficient of variation (all are normalized values). Figure 7 displays a comparison between the observed and modeled values for the mean flow, median flow (used in normalization), and the four unique (September mean flow is considered twice) metrics used within the classification tree. Mean, median, and September mean flows show model predictions within the ±30% uncertainty bounds for almost all calibration and validation sites. Some variation in predictive capability is seen in the July minimum, November maximum, and coefficient of variation modeled values. However, for both the November maximum and coefficient of variation, the majority of the values for the streams in question are well above and below the thresholds used in classification, respectively. The uncertainty in July minimum flows is more centered around the threshold value and may impact the determination of stream classes between stable base-flow designations and other more varied classes.

Hendriksen and Heasley (2010) developed a stream classification system for North Carolina using hydrologic data from 185 stream gages within the state containing at least 18 years of unaltered flows. They used 22 flow metrics to create six perennial stream classes and one seasonal stream class. Figure 8 displays the percent difference for each of the 22 metrics across the calibration and validation locations for the 40-year simulation period. The vast majority of sites and metrics fall within acceptable uncertainty ranges of ±30%. The two metrics for which model performance indicates poor prediction ability are for the rise and fall rates (RR and FR, respectively). These components of the hydrologic regime have been noted as controversial metric choices because of the inability of most models to recreate the trends considered with these metrics. The median absolute bias across the sites ranges from 0% (Calibration Sites 03503000 and 03512000; Validation Site 02145000) to 19% (Calibration Site 02114450), whereas the median absolute bias across metrics ranges from 0% (VLFSR) to 29% (AFM) excluding RR and FR.

DISCUSSION

Model Performance Relative to Similar Models

A recent model comparison study conducted to quantify differences in streamflow predictions through hydrologic metric calculation included WaterFALL, HSPF, PRMS (two versions), and SWAT

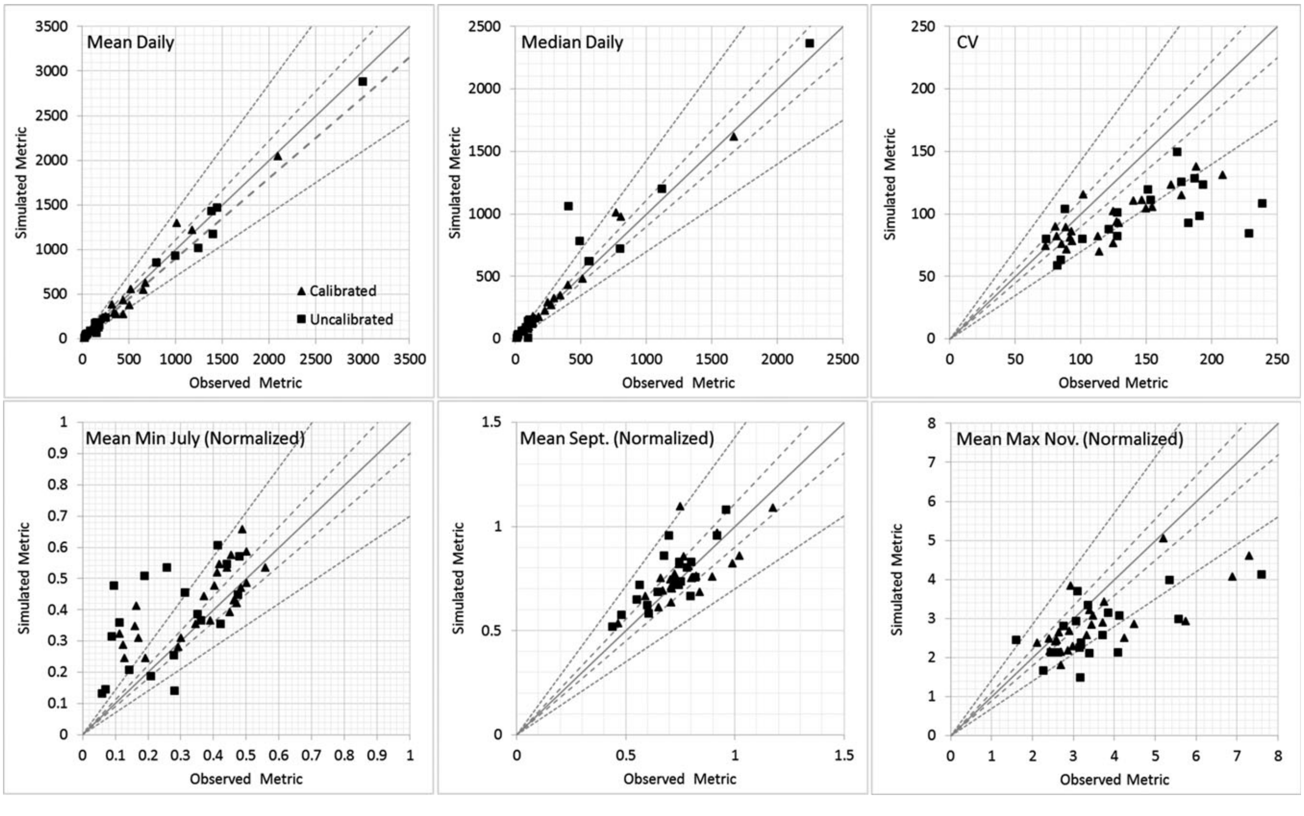


FIGURE 7. Model Performance Evaluation Using Hydrologic Metrics from the McManamay *et al.* (2012) Stream Classification System. Dotted lines on each graph represent ±10 and ±30% bounds of uncertainty as points of reference.

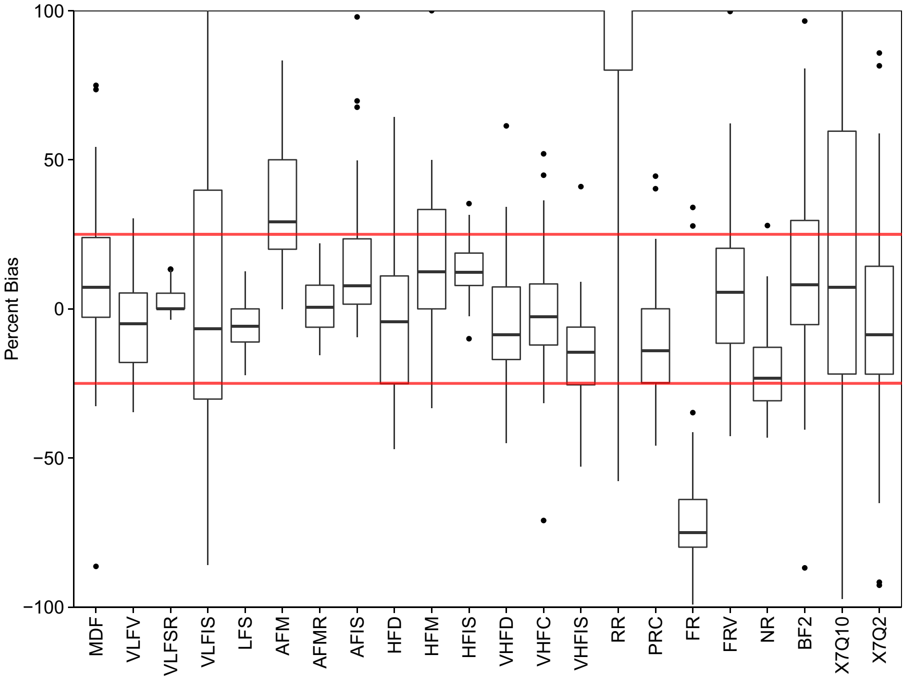


FIGURE 8. Percent Difference for 22 Hydrologic Metrics Used by the Hendricksen and Heasley North Carolina Stream Classification System across Calibration and Validation Sites. For metric definitions see Hendriksen and Heasley (2010).

(Caldwell *et al.*, 2015). This comparison was conducted at five sites within the Apalachicola-Chattahoochee-Flint Basin within Georgia and Alabama and examined classic hydrologic performance measures as well as percent difference (*i.e.*, bias) in 14 hydrologic metrics. Models used in the study were parameterized independently with WaterFALL set up to simulate unaltered conditions using an earlier land cover dataset than other models as well as a climate dataset which predicted the lowest amount of precipitation among the model inputs. Each model application also employed a different level of calibration, which also varied by site for some models. For example, WaterFALL used a combination of calibrated gages and applied calibration parameters for uncalibrated gages as explained for the North Carolina application in this study, while PRMS was specifically calibrated for each site with adjustments to climate inputs during calibration. HSPF and SWAT were calibrated similar to WaterFALL with a combination of calibrated and uncalibrated sites.

For the classic hydrologic performance metrics median bias among the sites range from −15% with WaterFALL to 1.3% with PRMS-SERAP. For NSE, monthly median values across sites ranged from 0.64 with SWAT to 0.87 with PRMS-SERAP (WaterFALL: 0.83), whereas daily median values ranged from 0.37 with HSPF to 0.80 with PRMS-SERAP (WaterFALL: 0.38). Finally, among the five sites and the 14 hydrologic metrics, the median absolute percent bias by model ranged from 18.7% for PRMS-DAYMET to 31.9% for SWAT and the number of metrics predicted outside of the ±30% uncertainty bounds (of 70 possible) ranged from 22 for HSPF to 38 for SWAT. WaterFALL had a median absolute bias across sites and metrics of 24.1% and predicted 31 metrics outside of the uncertainty bounds. Overall, WaterFALL performance was comparable to the other models, in particular the SWAT model for this application relied on the SCS-CN for runoff processes and was therefore similar to WaterFALL.

Potential Model Improvements

WaterFALL will continue to evolve as new applications and locations are tested. Presently, model enhancements in consideration for investigation of improved or varied methods include use of a more advanced representation of the evapotranspiration rates and inclusion of alternative methods of runoff simulation to supplement the use of the SCS-CN method. An evaluation of the single-stage calibration procedure is in process.

Model Enhancements. The Hamon (1961) evapotranspiration function used within WaterFALL relies solely on temperature and length of the day to calculate potential evapotranspiration, which are parameters included in the available model inputs.

Other methods to consider within WaterFALL include the Penman-Monteith method (Monteith, 1965; Jensen *et al.*, 1990), which requires air temperature, humidity, radiation, and wind speed, and is used within the SWAT model and the Jensen-Haise method (Jensen and Haise, 1963), which requires temperature, radiation, and latent heat inputs, and is used within PRMS. Modification of the evapotranspiration function from Hamon to either of these two initial suggestions would require either additional model input data or additional algorithms to estimate and calibrate the needed inputs from available climate data. When moving forward the benefit derived from using a more detailed evapotranspiration method must be balanced against the added complexity in the WaterFALL system required by these methods considering WaterFALL has purposely been designed to use the least intensive methods necessary to arrive at streamflow estimates.

As previously noted, the SCS-CN method has its critics, as well as its required uses in the regulated community. Because the method was originally developed for small land areas, application of the method over NHDPlus catchment units rather than full watersheds is more true to the original formulation and appears to eliminate the need for more complex runoff estimation methods or modifications to the area over which the SCS-CN is applied. Evaluation of the method to represent a range of storm peaks and smaller runoff events throughout the varied geographies and physical conditions simulated with WaterFALL model runs will continue with each new application. If necessary, different or modified runoff methods can be incorporated into the model as enabled by the infrastructure of the WaterFALL modeling system.

Calibration Processes. Within the North Carolina application there are no trends seen in performance metrics due to basin size or location (Tables 4 and 5) indicating that model performance is likely influenced by the decisions made during calibration by personnel who set initial values and bounds during automated calibration and determine any manual calibration adjustments. This finding is supported by the conclusion of Caldwell *et al.* (2015) that "differences among model predictions for specific fit statistics or [hydrologic metrics] are as likely to be related to differences in model calibration strategy as they are related to differences in model structure." The first step to examining the impact of the calibration process used to date in WaterFALL applications is to vary the objective function used to determine optimum model parameters. Potential functions used by other similar models in the Caldwell *et al.* (2015) comparison include minimize the normalized root mean square error at the daily or monthly time step and maximize daily or monthly NSE. A second step to improving the calibration process is to use a multistage calibration that uses either a combination of objective functions or specific order of calibration of model parameters. A formal model sensitivity analysis will be conducted to aid in the specification of refined calibration processes.

Although it would further add new calibration parameters, thereby adding complexity to the calibration process and model formulation, there is a need to allow for the calibration of curve numbers applied to the range of land uses. The curve numbers used within the model to date were assigned based on designated land use categories and soil hydrologic group using average values from USDA-NRCS (1986). Expanding the range of values may improve runoff and therefore peak flow estimates.

Finally, continuing on the path to enhanced calibration processes, there is a need to explore formal approaches to extrapolation of calibrated parameters to uncalibrated (*i.e.*, ungaged) locations. Hrachowitz *et al.* (2013) provide a summary of multiple techniques that can be explored using formal analysis. Although the use of multipliers on the AWC and RCoeff parameters within calibration procedures preserves the underlying spatial variation in soils and land use properties throughout both gaged and ungaged areas of a watershed, the formal extrapolation of these multipliers through such methods as transfer functions or statistical relationships based on physiographic properties may prove valuable in terms of model performance improvements.

Strengths and Limitations of WaterFALL

WaterFALL has been demonstrated as a valid model to represent uncalibrated locations (as surrogates for ungaged locations). Although performance across the North Carolina application varies from site to site, a similar issue is seen in other watershed-based models (Caldwell *et al.,* 2015). Performance issues for some sites indicate deficiencies in estimating peak flows; however, these deficiencies are likely related to the calibration process used and there are investigations planned to minimize these impacts for future applications. The simpler formulation of the routed rainfall-runoff-based hydrologic model within WaterFALL provides a comparable level of streamflow simulation in terms of classic hydrologic performance metrics and hydrologic metrics as other widely used watershed models while also providing physically based streamflow estimates at individual stream reaches between the calibrated gaged locations. The spatial granularity of daily streamflow results available from WaterFALL allows for the potential application to a range of management studies such as ecological flow analysis (Patterson *et al.*, 2017; Phelan *et al.*, 2017). The server-based data management system houses all model input and output for WaterFALL and allows for easy tracking and analysis of model simulations. Currently, WaterFALL is best configured for applications within the eastern U.S. due to the upstream-to-downstream accounting of water uses, as opposed to prior appropriation needs within the western portions of the country.

CONCLUSIONS

The modeling system described in this article provides a highly scalable, distributed, and readily parameterized hydrologic model. The foundations of WaterFALL are built upon an extensively published and demonstrated hydrologic network and rainfall-runoff estimation method, as well as a centralized data management system. The main contributions of WaterFALL can be summarized as follows: (1) a readily scalable model that can provide daily streamflow estimates from a single stream reach up to any number of larger hydrologic units; (2) a portable model that can be applied across ecoregions and other geographic conditions within the eastern U.S.; and (3) a flexible system that can be adapted to accommodate any range of modeling scenarios, including naturalized flowing systems to highly altered and regulated urban stream corridors.

APPENDIX A: ADDITIONAL MODEL COMPONENT DETAILS

Point Withdrawals/Discharges

Alterations are cumulative for a catchment in WaterFALL, which means that all withdrawals (W) within a catchment are totaled and discharges (D) within a catchment are totaled and applied at the downstream end of the catchment after considering inflow from any upstream catchment(s) and the locally generated flow (QL) (Figure 1; Equation 5). WaterFALL accounts for withdrawals first and then adds any discharges from human operations. Withdrawals are subtracted from streamflow and only a volume up to the corresponding streamflow can be withdrawn (*i.e.*, the cumulative streamflow [QC] will never be <0). Any discharges/returns are then added to the remaining cumulative streamflow. Because WaterFALL runs on a daily time step, it requires that daily alteration rates be included in streamflow calculations; however, daily variations in withdrawal/discharge rates are not typically tracked in a readily available format or reported to an overseeing agency. For this reason, WaterFALL is currently configured to consider changes in alterations on a monthly basis, meaning that the same daily alteration rate will be withdrawn or discharged to the stream channel every day of the same month (*i.e.*, one flow value is considered by month for each alteration).

Reservoirs

The first option to include a reservoir within a WaterFALL simulation is to designate a catchment as a control structure within the WaterFALL database. This designation requires that a complete time series of streamflows that accounts for the managed flow be included within the WaterFALL database for this catchment. For example, records of reservoir releases from a hydroelectric dam can replace the WaterFALL-predicted streamflow in the catchment that includes the dam. The dates of the replacement time series for the control structure limit the time period for any WaterFALL simulation considering the regulated flows. Multiple control structures can be considered within a single watershed. During routing, the input time series for this catchment is used in place of the calculation of cumulative streamflow (QC). When using a control structure with input time series the mass balance typically maintained during routing is not confirmed, so this option should be used with caution.

When physically representing a reservoir within a WaterFALL simulation, the reservoir is defined in the WaterFALL database by the catchments containing the actual reservoir (that may provide direct runoff), providing inflow from a tributary, and acting as the outflow point (Figure 2). Simulation of the outflow from the reservoir using mass balance can proceed using one of three formulations, set prior to the simulation for each reservoir in the watershed: (1) uncontrolled reservoir with average annual release rate; (2) measured daily outflow with preservation of the mass balance; or (3) static monthly reservoir volume targets. Any reservoir simulated is parameterized in the WaterFALL database by an initial volume, shape parameters to establish a volume-surface area relation, cover coefficient, hydraulic conductivity, and additional values based on the outflow calculation method chosen. Each day the volume of the reservoir (V_t) is calculated through the mass balance in Equation (A1):

$$V_t = V_{t-1} - \text{Outflow} + V_{\text{prcp}} - V_{\text{evap}} - V_{\text{seep}} \qquad (A1)$$

The precipitation volume (V_{prcp}) is calculated as the product of the rainfall depth and the area of the reservoir. The evaporation volume (V_{evap}) is calculated based on the same temperature-based potential evaporation calculation, but multiplied by reservoir surface area and a reservoir-specific cover coefficient. Any volume loss due to seepage (V_{seep}) from the reservoir is a linear function of a reservoir-specific hydrologic conductivity and the reservoir surface area. The reservoir outflow volume each day (Outflow_t) depends on the formulation set for the reservoir in the database.

For Option 1, an uncontrolled reservoir, outflow is dependent on the daily reservoir volume compared to both the principal reservoir volume (V_{prin}) and flood storage volume (V_{fld}). When the daily volume is between the principal and flood storage volumes, Equation (A2) governs the outflow calculation. When the volume is greater than flood storage volume, Equation (A3) is used to calculate the outflow for the day. If the volume is less than the principal volume then outflow is set to zero for the day. This option requires the specification of V_{prin}, V_{fld}, and $\text{Outflow}_{\text{avg}}$, the average daily outflow, within the WaterFALL database.

$$\text{Outflow}_t = \begin{cases} V_t - V_{\text{prin}}, & V - V_{\text{prin}} < \text{Outflow}_{\text{avg}} \\ \text{Outflow}_{\text{avg}}, & V - V_{\text{prin}} > \text{Outflow}_{\text{avg}} \end{cases} \qquad (A2)$$

$$\text{Outflow}_t = \begin{cases} V_t - V_{\text{prin}}, V_{\text{fld}} - V_{\text{prin}} < \text{Outflow}_{\text{avg}} \\ V_t - V_{\text{fld}} + \text{Outflow}_{\text{avg}}, \\ \qquad V_{\text{fld}} - V_{\text{prin}} > \text{Outflow}_{\text{avg}} \end{cases} \qquad (A3)$$

Option 2 utilizes a time series of measured daily outflows like the control structure option. However, unlike the control structure option, the daily reservoir volume is calculated based on the specified outflow and no release will be made if sufficient volume is not available thereby preserving the mass balance. Like the control structure option, the time series of daily outflows must correspond to the dates of the model simulation.

The final mass balance reservoir option uses a monthly target volume ($V_{\text{target},m}$ where $m = 1$ to 12) to determine the outflow from the reservoir (Equation A4). These target volumes are set within the WaterFALL database by reservoir. In addition, for this option, a target number of days needed to meet the monthly target ($\text{Days}_{\text{target}}$) must be specified within the database.

$$\text{Outflow}_t = \begin{cases} \frac{V_t - V_{\text{target},m}}{\text{Days}_{\text{target}}}, & V > V_{\text{target},m} \\ 0, & V \leq V_{\text{target},m} \end{cases} \qquad (A4)$$

ACKNOWLEDGMENTS

We wish to thank Yu Zhang (OHD/NOAA) for providing the data related to the SAC-SMA parameters and Jan Curtis (USDA Portland) and Mauro DiLuzio (Texas A&M) for providing the national 4-km resolution climate data. We acknowledge the following people for their work and support in conducting the initial modeling projects with WaterFALL: Jessie Allen, Lauren Patterson, Jennifer Phelan, Catherine Burns, Kimberley Meitzen, Rebecca Benner, Eloise Kendy, Rua Mordecai, Bryan Piazza, and Sam Pearsall. WaterFALL development was funded internally by RTI International. WaterFALL is currently used by RTI International as a project tool.

LITERATURE CITED

Anderson, R.M., V.I. Koren, and S.M. Reed, 2006. Using SSURGO Data to Improve Sacramento Model *a priori* Parameter Estimates. Journal of Hydrology 320(1):103-116.

Arnold, J.G., D.N. Moriasi, P.W. Gassman, K.C. Abbaspour, M.J. White, R. Srinivasan, C. Santhi, R.D. Harmel, A. van Griensven, M.W. Van Liew, N. Kannan, and M.K. Jha, 2012. SWAT: Model Use, Calibration, and Validation. Transactions of the ASABE 55(4):1491-1508.

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

Bicknell, B.R., J.C. Imhoff, J.L. Kittle, Jr., A.S. Donigian, Jr., and R.C. Johanson, 1997. Hydrological Simulation Program—Fortran, User's Manual for Version 11: U.S. Environmental Protection Agency, National Exposure Research Laboratory, Athens, Georgia, EPA/600/R-97/080, 755 pp.

Bingner, R.L. and F.D. Theurer, 2001. AGNPS 98: A Suite of Water Quality Models for Watershed Use. Proceedings of the Sediment: Monitoring, Modeling, and Managing, 7th Federal Interagency Sedimentation Conference, Reno, Nevada, March 25-29.

Blöschl, G. and M. Sivapalan, 1995. Scale Issues in Hydrological Modelling: A Review. Hydrological Processes 9:251-290.

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

Brakebill, J.W., D.M. Wolock, and S.E. Terziotti, 2011. Digital Hydrologic Networks Supporting Applications Related to Spatially Referenced Regression Modeling. Journal of the American Water Resources Association 45(5):916-932.

Brodie, R.S. and S. Hostetler, 2005. A Review of Techniques for Analysing Baseflow from Stream Hydrographs, Proceedings of the NZHS-IAH-NZSSS 2005 Conference, Auckland, New Zealand, Vol. 28.

Bunn, S.E. and A.H. Arthington, 2002. Basic Principles and Ecological Consequences of Altered Flow Regimes for Aquatic Biodiversity. Environmental Management 30(4):492-507.

Burnash, R.J.C., 1995. The NWS River Forecast System—Catchment Modeling. *In*: Computer Models of Watershed Hydrology, V.P. Singh (Editor). Water Resources Publications, Littleton, Colorado, pp. 311-366.

Caldwell, P.V., J.G. Kennen, G. Sun, J.E. Kiang, J.B. Butcher, M.C. Eddy, L.E. Hay, J.H. LaFontaine, E.F. Hain, S.A.C. Nelson, and S.G. McNulty, 2015. A Comparison of Hydrologic Models for Ecological Flows and Water Supply. Ecohydrology 8 (8):1525-1546, DOI: 10.1002/eco.1602.

Dai, T., J. Zhen, M. Chokshi, S. Paul, and L. Shoemaker, 2005. New Enhancements and TMDL Tools for Generalized Watershed Loading Functions (GWLF) Model. *In*: Proceedings of the Water Environment Federation, National TMDL Science and Policy, Water Environment Federation, Alexandria, Virginia, pp. 955-976, ISSN 1938-6478.

DiLuzio, M., G.L. Johnson, C. Daly, J.K. Eischeid, and J.G. Arnold, 2008. Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States. Journal of Applied Meteorology and Climatology 47:475-497.

Doherty, J., 2010. PEST: Model-Independent Parameter Estimation User Manual (Fifth Edition). Watermark Numerical Computing, Brisbane, Australia. http://www.pesthomepage.org/Downloads.php, *accessed* January 8, 2015.

Doherty, J. and J.M. Johnston, 2003. Methodologies for Calibration and Predictive Analysis of a Watershed Model. Journal of the American Water Resources Association 39(2):251-265.

Duda, P.B., P.R. Hummel, A.S. Donigian, Jr., and J.C. Imhoff, 2012. BASINS/HSPF Model Use, Calibration and Validation. Transactions of the ASABE 55(4):1523-1547.

Falcone, J.A., D.M. Carlisle, D.M. Wolock, and M.R. Meador, 2010. GAGES: A Stream Gage Database for Evaluating Natural and Altered Flow Conditions in the Conterminous United States: Ecological Archives E091-045. Ecology 91(2):621.

Feldman, A.D., 1995. HEC-1 Flood Hydrograph Package. *In*: Computer Models of Watershed Hydrology, V.P. Singh (Editor). Water Resources Publications, Highlands Ranch, Colorado, pp. 119-150.

Forsythe, W.C., E.J. Rykiel, Jr., R.S. Stahl, H.I. Wu, and R.M. Schoolfield, 1995. A Model Comparison for Daylength as a Function of Latitude and Day of Year. Ecological Modelling 80(1):87-95.

Fry, J., G. Xian, S. Jin, J. Dewitz, C. Homer, L. Yang, C. Barnes, N. Herold, and J. Wickham, 2011. Completion of the 2006 National Land Cover Database for the Conterminous United States. Photogrammetric Engineering & Remote Sensing (PE&RS) 77(9):858-864.

Garen, D.C. and D.S. Moore, 2005. Curve Number Hydrology in Water Quality Modeling: Uses, Abuses, and Future Directions. Journal of the American Water Resources Association 41(2):377-388.

Gentine, P., T.J. Troy, B.R. Lintner, and K.L. Findell, 2012. Scaling in Surface Hydrology: Progress and Challenges. Journal of Contemporary Water Research & Education 147(1):28-40.

Georgas, N., S. Rangarajan, K.J. Farley, and S.C.K. Jagupilla, 2009. AVGWLF-Based Estimation of Nonpoint Source Nitrogen Loads Generated within Long Island Sound Subwatersheds. Journal of the American Water Resources Association 45(3):715-733.

Haith, D.A., R. Mandel, and R.S. Wu, 1992. GWLF, Generalized Watershed Loading Functions, Version 2.0, User's Manual. Department of Agricultural & Biological Engineering, Cornell University, Ithaca, New York.

Haith, D.A. and L.L. Shoemaker, 1987. Generalized Watershed Loading Functions for Stream Flow Nutrients. Journal of the American Water Resources Association 23(3):471-478.

Hamon, W.R., 1961. Estimating Potential Evapotranspiration. Proceedings of the American Society of Civil Engineers. Journal of the Hydraulics Division 87(HY3):107-120.

Hendriksen, J. and J. Heasley, 2010. Hydroecological Stream Classification and StreamFlow© Software Development for North Carolina. Report prepared for Environmental Defense Fund and North Carolina Division of Water Resources, 46 pp., Fort Collins, Colorado.

Howarth, R.W., J.R. Fruci, and D. Sherman, 1991. Inputs of Sediment and Carbon to an Estuarine Ecosystem: Influence of Land Use. Ecological Applications 1(1):27-39.

Hrachowitz, M., H.H.G. Savenije, G. Blöschl, J.J. McDonnell, M. Sivapalan, J.W. Pomeroy, B. Arheimer, T. Blume, M.P. Clark, U. Ehret, F. Fenicia, J.E. Freer, A. Gelfan, H.V. Gupta, D.A. Hughes, R.W. Hut, A. Montanari, S. Pande, D. Tetzlaff, P.A. Troch, S. Uhlenbrook, T. Wagener, H.C. Winsemius, R.A. Woods, E. Zehe, and C. Cudennec, 2013. A Decade of Predictions in Ungauged Basins (PUB)—A Review. Hydrological Sciences Journal 58(6):1198-1255, DOI: 10.1080/02626667.2013.803183.

Jackson, R.B., S.R. Carpenter, C.N. Dahm, D.M. McKnight, R.J. Naiman, S.L. Postel, and S.W. Running, 2001. Water in a Changing World. Ecological Applications 11:1027-1045.

Jakeman, A.J. and G.M. Hornberger, 1993. How Much Complexity Is Warranted in a Rainfall-Runoff Model? Water Resources Research 29(8):2637-2649.

Jensen, M.E., R.D. Burman, and R.G. Allen, 1990. Evapotranspiration and Irrigation Water Requirements. ASCE, New York, New York.

Jensen, M.E. and H.R. Haise, 1963. Estimating Evapotranspiration from Solar Radiation. Journal of Irrigation and Drainage Division 89:15-41.

Jobson, H.E., 1996. Prediction of Traveltime and Longitudinal Dispersion in Rivers and Streams. U.S. Geological Survey, Reston, Virginia.

Kircher, J.W., 2009. Catchments as Simple Dynamical Systems: Catchment Characterization, Rainfall-Runoff Modeling, and Doing Hydrology Backward. Water Resources Research 45: W02429.

Kling, H. and H. Gupta, 2009. On the Development of Regionalization Relationships for Lumped Watershed Models: The Impact of Ignoring Sub-Basin Scale Variability. Journal of Hydrology 373(3):337-351.

Knight, R.R., J.C. Murphy, W.J. Wolfe, C.F. Saylor, and A.K. Wales, 2013. Ecological Limit Functions Relating Fish Community Response to Hydrologic Departures of the Ecological Flow Regime in the Tennessee River Basin, United States. Ecohydrology 7(5):1262-1280, DOI: 10.1002/eco.1460.

Krause, P., D.P. Boyle, and F. Bäse, 2005. Comparison of Different Efficiency Criteria for Hydrological Model Assessment. Advances in Geosciences 5:89-97.

Li, X., D.E. Weller, and T.E. Jordan, 2010. Watershed Model Calibration Using Multi-Objective Optimization and Multi-Site Averaging. Journal of Hydrology 380(3):277-288.

Liang, X., D.P. Lettenmaier, E.F. Wood, and S.J. Burges, 1994. A Simple Hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs. Journal of Geophysical Research 99 (D7):14415-14428.

Linsley, R., M. Kohler, and J. Paulhus, 1975. Hydrology for Engineers. McGraw Hill, New York, New York, 512 pp.

Magilligan, F.J. and K.H. Nislow, 2005. Changes in Hydrologic Regime by Dams. Geomorphology 71(1):61-78.

Markstrom, S.L., R.S. Regan, L.E. Hay, R.J. Viger, R.M.T. Webb, R.A. Payn, and J.H. LaFontaine, 2015. PRMS-IV, the Precipitation-Runoff Modeling System, Version 4, U.S. Geological Survey Techniques and Methods, book 6, chap. B7, 158 pp., DOI: 10.3133/tm6B7.

McKay, L., T. Bondelid, A. Rea, C. Johnston, and R. Moore, 2013. NHDPlus Version 2: User Guide. ftp://ftp.horizon-systems.com/NHDPlus/NHDPlusV21/Documentation/NHDPlusV2_User_Guide.pdf, *accessed* January 2014.

McManamay, R.A., D.J. Orth, C.A. Dolloff, and E.A. Frimpong, 2012. A Regional Classification of Unregulated Stream Flows: Spatial Resolution and Hierarchical Frameworks. River Research and Applications 28(7):1019-1033.

Merz, R. and G. Blöschl, 2004. Regionalisation of Catchment Model Parameters. Journal of Hydrology 287(1):95-123.

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

Moriasi, D.N., J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, and T.L. Veith, 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the ASABE 50(3):885-900.

Murphy, J.C., R.R. Knight, W.J. Wolfe, and W.S. Gain, 2013. Predicting Ecological Flow Regime at Ungaged Sites: A Comparison of Methods. River Research and Applications 29(5):660-669.

Nash, J. and J.V. Sutcliffe, 1970. River Flow Forecasting through Conceptual Models Part I—A Discussion of Principles. Journal of Hydrology 10(3):282-290.

NCDC (National Climatic Data Center), 2002. Climate Atlas of the United States. http://geodata.lib.ncsu.edu/fedgov/noaa/climate/DOCS/datadoc.html, *accessed* November 2016.

NCDWR (North Carolina Division of Water Resources), 2013. Central Coastal Plain Capacity Use Area. North Carolina Department of Environment and Natural Resources. http://www.ncwater.org/Permits_and_Registration/Capacity_Use/Central_Coastal_Plain/, *accessed* September 2013.

NRCS (Natural Resources Conservation Service), 2009. Small Watershed Hydrology WinTR-55 User Guide. http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1042897.pdf, *accessed* January 2013.

NRCS (National Resources Conservation Service), 2015. WinTR-20 User Guide Version 3.10. http://www.wcc.nrcs.usda.gov/ftpref/wntsc/H&H/WinTR20/WinTR20UserGuideVer310Mar2015.pdf, *accessed* June 2015.

Parajka, J., R. Merz, and G. Blöschl, 2005. A Comparison of Regionalisation Methods for Catchment Model Parameters. Hydrology and Earth System Sciences Discussions, European Geosciences Union 9(3):157-171.

Patterson, L., J. Phelan, C. Goudreau, and R. Dykes, 2017. Flow-Biology Relationships Based on Fish Habitat Guilds in North Carolina. Journal of the American Water Resources Association 53(1):56-66 DOI: 10.1111/1752-1688.12495.

Phelan, J., T. Cuffney, L. Patterson, M. Eddy, R. Dykes, S. Pearsall, C. Goudreau, J. Meade, and F. Tarver, 2017. Fish and Invertebvrate Flow-Biology Relationships to Support the Determination of Ecological Flows for North Carolina. Journal of the American Water Resources Association 53(1):42-55 DOI: 10.1111/1752-1688.12497.

Poff, N.L., B.D. Richter, A.H. Arthington, S.E. Bunn, R.J. Naiman, E. Kendy, M. Acreman, C. Apse, B.P. Bledsoe, M.C. Freeman, J. Henriksen, R.B. Jacobson, J.G. Kennen, D.M. Merritt, J.H. O'Keeffe, J.D. Olden, K. Rogers, R.E. Tharme, and A. Warner, 2010. The Ecological Limits of Hydrologic Alteration (ELOHA): A New Framework for Developing Regional Environmental Flow Standards. Freshwater Biology 55:147-170.

Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreda, D.J. Seo, and D.M.I.P. Participants, 2004. Overall Distributed Model Intercomparison Project Results. Journal of Hydrology 298(1):27-60.

Richter, B.D., J.V. Baumgartner, R. Wigington, and D.P. Braun, 1997. How Much Water Does a River Need? Freshwater Biology 37:231-249, DOI: 10.1046/j.1365-2427.1997.00153.x.

Schneiderman, E.M., D.C. Pierson, D.G. Lounsbury, and M.S. Zion, 2002. Modeling the Hydrochemistry of the Cannonsville Watershed with Generalized Watershed Loading Functions (GWLF). Journal of the American Water Resources Association 38 (5):1323-1347.

Schneiderman, E.M., T.S. Steenhuis, D.J. Thongs, Z.M. Easton, M.S. Zion, A.L. Neal, G.F. Mendoza, and M.T. Walter, 2007. Incorporating Variable Source Area Hydrology into a Curve-Number-Based Watershed Model. Hydrological Processes 21:3420-3430.

Simley, J.D. and W.J. Carswell, Jr., 2009. The National Map—Hydrography: U.S. Geological Survey Fact Sheet 2009-3054, 4 pp.

USACE (U.S. Army Corps of Engineers), 2014. Hydrologic Engineering Center. HEC-HMS (Hydrologic Modeling System). http://www.hec.usace.army.mil/software/hec-hms/, *accessed* November 2016.

USDA-NRCS (U.S. Department of Agriculture-Natural Resources Conservation Service), 1986. Urban Hydrology for Small Watersheds. Conservation Engineering, Technical Release 55. USDA Natural Resources Conservation Service, Washington, D.C.

USDA-NRCS (U.S. Department of Agriculture-Natural Resources Conservation Service), 2014. Soil Survey Geographic (SSURGO) Database. USDA Natural Resources Conservation Service. http://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm.

Van de Griend, A.A., J.J. De Vries, and E. Seyhan, 2002. Groundwater Discharge from Areas with a Variable Specific Drainage Resistance. Journal of Hydrology 259:203-220.

Wagener, T., H.S. Wheater, and H.V. Gupta, 2004. Rainfall-Runoff Modelling in Gauged and Ungauged Catchments. Imperial College Press, London.

Walter, M.T. and S.B. Shaw, 2005. Discussion: "Curve Number Hydrology in Water Quality Modeling: Uses, Abuses, and Future Directions," by David C. Garen and Daniel S. Moore. Journal of the American Water Resources Association 41 (6):1491-1492.

Young, R.A., C.A. Onstad, D.D. Bosch, and W.P. Anderson, 1989. AGNPS: A Nonpoint-Source Pollution Model for Evaluating Agricultural Watersheds. Journal of Soil and Water Conservation 44(2):168-173.

Zhang, Y., Z. Zhang, S. Reed, and V. Koren, 2011. An Enhanced and Automated Approach for Deriving *a priori* SAC-SMA Parameters from the Soil Survey Geographic Database. Computers & Geosciences 37(2):219-231.