Journal of Hydrology 494 (2013) 20–31

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Applying hillslope-storage models to improve low flow estimates with limited streamflow data at a watershed scale

A.H. Matonse [*,1], C.N. Kroll

Environmental Resources Engineering, State University of New York (SUNY), College of Environmental Science and Forestry (ESF), 1 Forestry Dr., Syracuse, NY 13210, United States

ARTICLE INFO

Article history:
Received 29 November 2012
Received in revised form 19 March 2013
Accepted 13 April 2013
Available online 29 April 2013
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief

Keywords:
Hillslope-storage baseflow models
Sacramento model
Simulation of low-flow series and statistics
Model calibration with limited streamflow data

SUMMARY

This study presents a framework to evaluate the performance of rainfall-runoff models for the estimation of low flow at sites with limited streamflow data. Estimates of low flow statistics are important for water supply, waste-load allocation, irrigation, hydropower, and ecological and habitat assessment. Paradoxically most rainfall-runoff models focus on flood simulations and use oversimplified representations of baseflow processes resulting in poor performance simulating low flow statistics. Such baseflow models cannot account for variations in topography and hydrogeology that impact baseflow processes and have limited applicability to evaluate land use and climate change impacts on low flow. Both a hillslope-storage Boussinesq model (*hsB*) and a kinematic wave hillslope-storage model (*kw*) have shown good results in simulating baseflow in synthetic hillslopes; one major challenge is how to apply these models in real watersheds. In this study *hsB* and *kw* are coupled to the Sacramento Soil Moisture Accounting (SAC-SMA) model and tested at two similarly sized watersheds in North Carolina with different watershed slopes. The partitioned *kw* and *hsB* models are also compared to the original SAC-SMA model (*Sac*) and SAC-SMA applied to a partitioned watershed (*Sacm*). Both 5 years and 1 year of full and reduced ranges of streamflow data are employed for model calibration. All partitioned models improved their estimation of low flow when calibrated to a lower range of streamflows but with *kw* and *hsB* performing slightly better at the steeper sloped watershed. The performance of the coupled models with limited streamflow data is encouraging and can potentially improve the estimation of low flow statistics at sites with limited streamflow data.

1. Introduction

Estimates of low flow discharges and statistics are needed for water quality management, watershed ecosystem restoration and water supply planning under both present and climate change conditions (Pushpalatha et al., 2011; Ouyang, 2012). This includes determining the necessary dilution of waste-load allocations, allowable surface water and groundwater withdrawals for municipal and industrial water uses, the design of reservoir storage capacity, and minimum downstream releases for the aquatic ecosystem (Riggs, 1980; Metcalf and Eddy, 1991; Mosley and McKerchar, 1993; Vogel and Fennessey, 1995; Kroll and Vogel, 2002; Karim et al., 1995; Caruso, 2002). When a historic record of sufficient length is available at the streamflow site of interest, analytical methods have been developed to assess low flow characteristics. This includes estimation of annual minimum d-day, T-year flow statistics (Riggs, 1961), low flow quantiles from flow duration curves (Smakhtin, 2001), and low flow duration and severity (Salas et al., 2005). Often a period of 20–30 years of historic record is considered sufficient to estimate such statistics (Hisdal et al., 2004). Of particular concern is how best to estimate low flow characteristics when limited or no streamflow data is available. In addition, even when a historic record is available, it is often unclear how to predict changes in low flow characteristics when a watershed undergoes climatic or anthropogenic changes. While statistical methods, such as regional regression (Vogel and Kroll, 1996; Kroll et al., 2004), baseflow correlation (Stedinger and Thomas, 1985; Zhang and Kroll, 2007), index low flows (Clausen and Pearson, 1995; Madsen and Rosbjerg, 1998), and geostatistics (Skøien et al., 2006; Laaha et al., 2007), have been proposed to address the problem of no or limited streamflow data, these methods are not equipped to handle low flow prediction in watersheds undergoing change. One solution to this problem is to employ a physics-based rainfall runoff (RR) model to estimate low flow characteristics at the site of interest.

* Corresponding author. Present address: CUNY Institute for Sustainable Cities (CISC), Hunter College, City University of New York (CUNY), 695 Park Avenue, NY 10065, United States. Tel.: +1 (845) 340 7793 (O); fax: +1 (845) 338 1032.

E-mail addresses: amatonse@hunter.cuny.edu (A.H. Matonse), cnkroll@esf.edu (C.N. Kroll).

[1] Office location: WQ Modeling Group, NYC DEP, 71 Smith Avenue, Kingston, NY 12401, United States.

0022-1694/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhydrol.2013.04.032

While the hydrologic literature is inundated with RR models that focus on the estimation of peak discharges, there have been limited studies regarding the use of RR models for low flow prediction. During low flow events, streamflow consists primarily of groundwater discharge (Brutsaert and Nieber, 1977), which is commonly called baseflow, where this discharge recedes over time when there is no recharge to groundwater. Baseflow discharge models employed within RR models vary from a single linear reservoir with a lumped recession parameter, to multiple reservoirs with varying parameters, to more complex distributed models, such as MODFLOW, that is based on the 3-dimensional Richard's equation. Van Lanen et al. (1997) compared three different RR models of varying complexity for estimating hydrological drought. They found that even more complex models had difficulties in simulating drought details. Engeland and Hisdal (2009) compared regional regression to the semi-distributed HBV model for estimating low flow indices at ungauged watersheds. They found regional regression produced better low flow estimators than the HBV model, which employs a linear reservoir based model for baseflow. The difference in performance of the two models was larger for smaller values of the low flow index. Samuel et al. (2011) found that using a non-linear discharge-storage relationship in deeper soils to replace the linear discharge-storage in McMaster University-Hydrlogiska Byraens Vattenbalansavdelning (MAC-HBV) model combined with increasing the range for each parameter values and inclusion of low flow criteria in model calibration improved MAC-HBV ability to simulate baseflow time series at ungauged sites in northern Ontario, Canada. Querner and van Lanen (2001) employed the transient model SIMGRO to simulate the impact of urbanization strategies on hydrologic drought. They found that urban stormwater management has an impact on the length and severity of low flow periods. Wagener et al. (2009) analyzed parameter sensitivity for the upper and lower soil zones in the National Weather Service's (NWS) Sacramento Soil Moisture Accounting (SAC-SMA) model and how it relates to high and low flows. In their study they found that low flows are sensitive to both upper and lower zone parameters as opposed to only upper zone parameters which primarily impact high streamflow. Staudinger et al. (2011) studied the impact of model structure on model ability to simulate low flow and recession behavior for the Narsjo catchment in Norway. They found that different structural combination of conceptual models for a lower layer, subsurface flow and percolation have a significant impact on summer and winter model performance but with the lower layer and subsurface flow having more influence on winter performance. Son and Sivapalan (2007) describe a downward approach for model structure development including the use of auxiliary data (deuterium composition and groundwater level dynamics) and multiple wetting front water movement in the unsaturated zone to reduce model predictive uncertainty and improve the physical realism and streamflow simulation.

Of interest in this study is the development of models which account for variations in the topography, geometry and hydrogeology that impact baseflow processes (Tucker and Bras, 1998; Bogaart and Troch, 2004), yet are not overly complex such as distributed groundwater models. One potential solution is integrating hillslope discharge models into a RR model. Studies have shown how topography and geometry are two dominant factors that influence hillslope discharge (Troch et al., 2002). Fan and Bras (1998) proposed a methodology to simplify the 3-dimensional hillslope runoff processes into a 1-dimensional approach that maintains some of the 3-dimensional characteristics of baseflow. Extending this approach, Troch et al. (2003) applied two solutions to the Boussinesq Equation (Boussinesq, 1877), the kinematic wave (*kw*) and the hillslope-storage Boussinesq (*hsB*) model, a more general form that accounts for both gravity and diffusion flow generation mechanisms, to a series of synthetic hillslopes. Studies on synthetic hillslopes have shown that the relatively simple 1-dimensional *hsB* model can in general simulate synthetic hillslope flow dynamics as well as a model based on the 3-dimensional Richards equation (Paniconi et al., 2003). Hilberts et al. (2007) showed that coupling *hsB* with a 1-dimensional Richard's equation model to simulate vertical unsaturated flow can simulate water table levels and outflow in synthetic hillslopes with varying geometry better than the original *hsB* model, which does not account for an unsaturated zone component. Broda et al. (2011a) combined the *hsB* model for shallower baseflow with an analytical element (AE) model to account for deep regional groundwater flow using a connected leakage element based on Darcy's law. The resulting *hsB*/AE approach showed good results in simulating heads, hydrographs and exchange fluxes on shallow (0.2% and 5%) uniform and divergent synthetic hillslopes, and good simulation of heads in confined and unconfined aquifers of an open-book catchment when compared to a 3-dimensional Richards equation.

A major challenge is how to apply the *hsB* model concept in real watersheds. A model application requires the development of a model framework to account for other processes in the hydrologic cycle at a watershed scale. Fan and Bras (1998) applied the *kw* model in a watershed in the White Mountain National Forest, New Hampshire. For their application, they considered hillslopes as areas between two streamlines at the ends of a channel link or "from a channel head to the ridgelines" (Fan and Bras, 1998). The size of the hillslopes was extended for adjacent areas as long as the plan and profile curvatures were similar. Their model framework was based on the following assumptions: (i) the kinematic approximation of the hsB model applies; (ii) no evapotranspiration, interception or infiltration, and thus all rainfall becomes recharge directly; (iii) saturation excess overland flow is possible but not infiltration excess (Hortonian) runoff; and (iv) the hillslope plan profile is described by a second-order polynomial. Matonse and Kroll (2009) examined the use of the *kw* and *hsB* hillslope-storage models to improve the prediction of baseflow and low flow statistics. Using a relatively simple model framework based on a mass balance to estimate discharge, and a variable range of streamflow calibration, they showed that by partitioning the watershed into multiple hillslopes with variable parameters that account for changes in topography, geometry and hydrogeology can improve model performance. Though models calibrated with a lower range of streamflow data can better describe the corresponding lower envelope of streamflow data, there was no evidence of a significant improvement from using different calibration ranges in simulating the 7-day (Q_7) and 30-day (Q_{30}) low flow statistics. Their results also indicated that the *kw* and *hsB* models perform similarly for a steep sloped watershed (such as the 34° Maimai M8), where gravity processes dominate over diffusion processes. However, their results were constrained by the limited availability of input data, the small size of the study site, and the steepness of the watershed.

In this analysis, we develop a modeling framework by coupling the *kw* and *hsB* models to the SAC-SMA model and apply these models to two larger watersheds with similar drainage areas but different watershed slopes. The watersheds are partitioned into multiple hillslopes, and the *kw* and *hsB* models are compared to the original SAC-SMA model, which represents baseflow as two homogeneous linear reservoirs, and a partitioned SAC-SMA where baseflow from each hillslope is individually parameterized. Using this application, calibrating these models with either 5 years or 1 year of data, and examining the models ability to predict low flow discharges and statistics, the following is investigated:

(1) If improvements in the baseflow component in SAC-SMA can lead to improved low flow prediction.

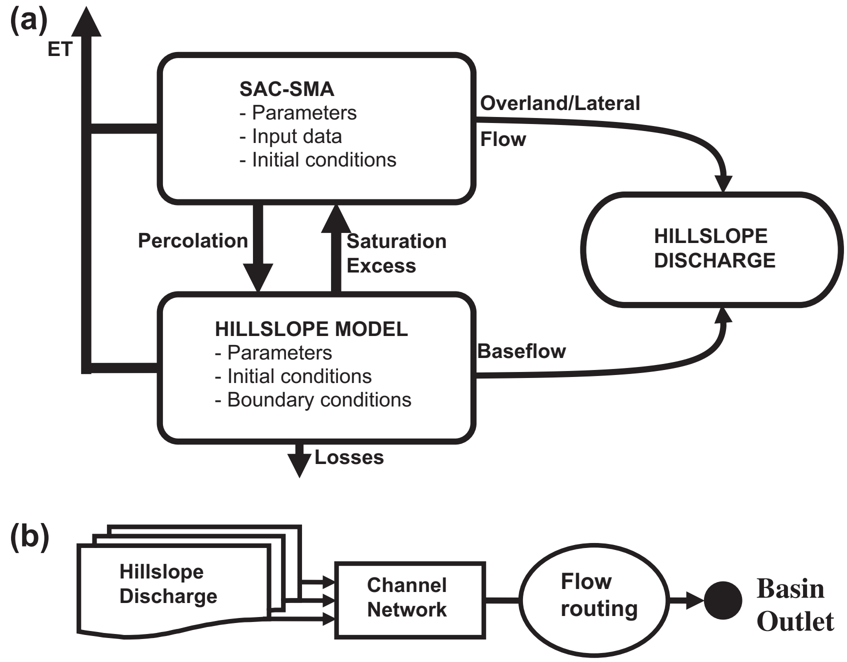

Fig. 1. Schematic representation of the coupling between (a) the SAC-SMA model and hillslope models, and (b) the routing of each hillslope discharge to the basin outlet. (IC – initial conditions; BC – boundary conditions; ET – evapotranspiration.)

(2) Whether applying the *hsB* model to a watershed with a lower mean average slope (or shallower sloped watershed), where hillslope diffusion processes should be present, will improve low flow prediction.
(3) If calibrating models using
  a. A range of streamflow data below the 20th or 50th percentile.
  b. A variable calibration length, i.e. 1-year versus 5-year calibration period.
  c. Data representing a particular hydrologic regime, i.e. data from a dry, normal, wet, or mixed period, will improve low flow prediction.

2. Model framework development

To account for most elements of the hydrologic cycle and preserve the overall watershed mass balance, we adopted a model framework based on the coupling of the *kw* and *hsB* hillslope-storage models to the SAC-SMA model, a well-accepted model within the hydrologic community. SAC-SMA was originally developed for flood forecasting. It is a lumped-parameter continuous simulation RR model that incorporates land cover, evapotranspiration, infiltration, lateral flow (interflow), baseflow and overland flow to estimate runoff at the watershed outlet (Burnash, 1995). The model represents subsurface processes as two layers: one active root zone layer and one lower soil layer. In each layer current tension water and free water content are estimated. The free water in the upper layer is available for percolation to the lower zone and for lateral flow. In the lower layer two reservoirs of free water coexist. Water discharges from these two linear reservoirs for baseflow at two different rates of decline. In this manuscript we use the abbreviation "*Sac*" to refer to lumped SAC-SMA watershed application, and "*Sacm*" when the model is applied to multiple hillslopes.

Here we compare four models for estimating low flow series and statistics: the original *Sac* model, *Sac* applied to multiple hillslopes (*Sacm*), the *Sac* model with the lower soil layer replaced by the *kw* hillslope-storage model (*Sackw* or simply *kw*), and the *Sac* model with the lower soil layer replaced by the *hsB* hillslope-storage model (*SachsB* or *hsB*). Our objective is to improve baseflow simulation by accounting for physical variations in hydrogeology, hillslope geometry and topography that are important for groundwater flow at a hillslope scale. Fig. 1a illustrates the coupling between *Sac* and the two hillslope-storage models. The coupling is performed at a hillslope basis meaning that both *Sac* and hillslope-storage models are connected and run for each individual hillslope. As in the original *Sac* structure, the saturation excess moisture from the lower zone is sent back to the upper layer and is subject to evaporation, becoming lateral flow, or re-percolating to the lower layer during the following time step. Because of the large size of the watersheds, contributions from individual hillslopes are routed through the channel network to the basin outlet (as illustrated in Fig. 1b) using the looped-rated Muskingum–Cunge method by Ponce and Lugo (2001). As with the original *Sac* the models are run at a 6 h time step.

3. Hillslope-storage models

Two hillslope-storage models are applied in this study, the kinematic wave (*kw*) and the hillslope-storage Boussinesq (*hsB*) models. The formulation of the *hsB* model results from combining Darcy's law:

$$Q = -\frac{kS(x,t)}{f}\left[\cos i \frac{\partial}{\partial x}\left(\frac{S(x,t)}{fw(x)}\right) + \sin i\right] \tag{1}$$

where Q [L^3T^{-1}] is volumetric groundwater flow, k [LT^{-1}] is the soil (effective) saturated hydraulic conductivity, S [L^2] is the storage at location x at time t, f [−] is the drainable porosity, and i is the hillslope slope angle, with the continuity equation:

$$\frac{\partial S}{\partial t} = \frac{\partial Q}{\partial x} + Nw \tag{2}$$

where N represents recharge and w a width function. The final *hsB* formulation has the form:

$$f\frac{\partial S}{\partial t}=\frac{k\cos i}{f}\frac{\partial}{\partial x}\left[\frac{S(x,t)}{w(x)}\left(\frac{\partial S(x,t)}{\partial x}-\frac{S(x,t)}{w(x)}\frac{\partial w(x)}{\partial x}\right)\right]+k\sin i\frac{\partial S(x,t)}{\partial x}+f\ N(t)w(x) \quad (3)$$

Fan and Bras (1998) introduced the storage capacity function:

$$S_c(x)=w(x)\,\bar{d}_m(x)f \quad (4)$$

where $w(x)$ [L] is the hillslope width function, and $\bar{d}_m(x)$ is the average hillslope depth, to reduce the 3-dimensional soil mantle to 1-dimension. Eq. (3) accounts for both diffuse and gravity drainage and is applicable to 3-dimensional flow in hillslopes with different plan and profile curvature (Troch et al., 2002, 2003). For this application the *hsB* model in Eq. (3) is solved numerically by discretizing in space using finite differences and applying a multistep ordinary differential equation (ODE) solver. This solution can accommodate different boundary conditions, as well as temporal and spatial variability of recharge, hydraulic parameters, and slope angles. In addition, the hydraulic conductivity k was determined for each time step and for each hillslope according to a power function (Rupp and Selker, 2006; Matonse and Kroll, 2009)

$$k(d)=k_S d^{\mu} \quad (5)$$

where k_s is the effective saturated hydraulic conductivity at the storage capacity (calibrated for each hillslope) and d [−] represents a ratio between the volumetric storage at the beginning of the current time step divided by the volumetric storage at full capacity. The use of effective hydraulic conductivity is justified because (i) most traditional small-scale k_s measurements do not capture the effect of macropores at a hillslope scale (Brooks et al., 2004), (ii) given the size of hillslopes employed in this study, topographic gradient (as opposed to water table gradient) is more likely to be the dominant factor, and (iii) different (localized) k_s values associated with heterogeneous porous medium across individual hillslopes are difficult to parameterize (Harman and Sivapalan, 2009).

The *hsB* model becomes a *kw* approximation under relatively steep impermeable bed slopes where it is assumed that the rate of groundwater flow is relatively high. As a consequence the second-order diffusive term in Eq. (1) can be dropped (Paniconi et al., 2003). For a more detailed description and derivation of the *kw* and *hsB* models see Fan and Bras (1998), Paniconi et al. (2003), Troch et al. (2002, 2003), and Matonse and Kroll (2009).

4. Description of the study sites

Two watersheds were modeled in this study: the Linville River watershed near Nebo, NC (LRN) (USGS 02138500), and the Indian Creek watershed near Laboratory, NC (ICL) (USGS 02143500). Table 1 presents general topographical, meteorological, and hydrological characteristics of these sites. While the drainage areas are similar, the average channel and watershed slopes vary greatly, with LRN being the steeper of the watersheds. In addition, LRN is at a higher elevation and has an average annual precipitation approximately 30% more than ICL. While discharges are larger at LRN, the coefficient of variation (CV) (standard deviation divided by the average) of the discharge at these two sites is similar. Note that in table 1, and throughout this manuscript, streamflow [L^3T^{-1}] values were converted into specific discharge units [L T^{-1}] (i.e. streamflow divided by the basin drainage area).

Both LRN and ICL are part of the Catawba River Basin, and their outlets are approximately 100 km apart. LRN is located at the eastern edge of the Blue Ridge physiographic province, while ICL is located at the western edge of the Piedmont physiographic province. At LRN the first 10 cm of soil are predominantly loam and sandy loam, while at ICL they are predominantly sandy clay loam (Daniel et al., 1997). Most of the Piedmont and Blue Ridge physiographic provinces are underlain by dense, almost impermeable bedrock that yields water primarily from secondary porosity and permeability provided by fractures. These bedrock fractures provide a network of channels for water movement, but provide little storage volume with porosity typically less than 1%. Most of the water is stored in the unconsolidated materials overlying the bedrock, where the porosity is 20–40%. The thickness of the regolith overlying the bedrock can vary between 0 to more than 150 ft in the region (Daniel and Sharpless, 1983; Daniel and Dahlen, 2002). While the hydrogeology of the region varies and is challenging to model, it is important to note that the hillslope-storage models employ "effective" hydraulic conductivities, where the overall drainage characteristics of the watershed hillslopes are aggregated and characterized. In the work of Matonse and Kroll (2009) and that presented here, it was assumed that hydraulic conductivity decreases with depth (Eq. (5)), as is commonly observed in practice.

Table 1
Watershed characteristics (CV: coefficient of variation).

Watershed characteristic	Watershed name	
	LRN	ICL
Area (km^2)	173	179
Average slope (°)	14.8	4.6
Main channel length (km)	56	30
Channel average slope (m/km)	14	2.5
Mean elevation (m)	1009	287
Minimum elevation (m)	366	224
Maximum elevation (m)	1782	410
Average annual precipitation [1951–2002] (mm/year)	1619	1207
Average annual discharge [1951–2002] (mm/year)	795	444
CV – annual discharge	0.30	0.29
20th Percentile of streamflow [1951–2002] (mm)	0.76	0.42
50th Percentile of streamflow [1951–2002] (mm)	1.41	0.75
CV – lower 20% streamflow	0.26	0.27
CV – lower 50% streamflow	0.37	0.37
CV of all streamflow	1.78	1.77

5. Data and methods

5.1. Input data and hillslope derivation

Data for both study sites are available via the Internet through the Model Parameter Estimation Experiment (MOPEX) project, US MOPEX Data Set at (http://www.nws.noaa.gov/oh/mopex/mo_datasets.htm) (Hogue et al., 2004; Schaake et al., 2006). The input data include: observed daily streamflow, daily potential evaporation (PE) with monthly PE adjustment factors, 6-h rainfall, and eleven a priori (Duan et al., 2001) *Sac* parameters listed in Table 2.

GIS pre-processing was employed to partition the watersheds into hillslopes. Each watershed was partitioned into nine hillslopes following an empirical approach similar to Fan and Bras (1998) using Arc Hydro tools and the "Editor" function in ArcGIS. Hillslopes were classified as straight, convergent, or divergent, or a combination of these shapes, based on the general flow line patterns (Fan and Bras, 1998; Troch et al., 2002; Paniconi et al., 2003). Areas with similar hillslope characteristics were aggregated to form larger hillslopes, and width function parameters were chosen to preserve the surface area of the hillslopes. Fig. 2 presents both watersheds partitioned into nine hillslopes.

Bogaart and Troch (2006) showed how multiple hillslope from a basin can be represented by a single hillslope folded around the channel network and how such a hillslope exhibit divergent characteristics. Matonse and Kroll (2009) found that partitioning a small headwater catchment into multiple hillslope with varying parameters resulted in improved model simulation of low statistics than either a single hillslope representation or using multiple hill-

Table 2
List and values of Sac a priori parameters and monthly PE adjustment factors used for LRN and ICL watersheds.

Parameter name	Parameter value	
	LRN	ICL
UZTWM – upper zone tension water capacity (mm)	57.7	30.9
UZFWM – upper zone free water capacity (mm)	51.4	27.7
UZK – fractional daily upper zone free water withdrawal rate	0.47	0.46
ZPERC – maximum percolation rate coefficient	75.02	30.02
REXP – percolation equation exponent	1.879	2.626
LZTWM – lower zone tension water capacity (mm)	192.2	196.5
LZFSM – lower zone supplemental free water capacity (mm)	21.4	45.8
LZFPM – lower zone primary free water capacity (mm)	142.8	84.3
PFREE – fraction of percolated water going directly to lower zone free water storage	0.14	0.36
PE_Adj (1) – January PE adjustment factor	0.65	0.61
PE_Adj (2) – February PE adjustment factor	0.69	0.67
PE_Adj (3) – March PE adjustment factor	0.72	0.73
PE_Adj (4) – April PE adjustment factor	0.87	0.96
PE_Adj (5) – May PE adjustment factor	1.22	1.09
PE_Adj (6) – June PE adjustment factor	1.33	1.06
PE_Adj (7) – July PE adjustment factor	1.31	1.01
PE_Adj (8) – August PE adjustment factor	1.32	1.03
PE_Adj (9) – September PE adjustment factor	1.28	1.08
PE_Adj (10) – October PE adjustment factor	0.98	0.95
PE_Adj (11) – November PE adjustment factor	0.66	0.69
PE_Adj (12) – December PE adjustment factor	0.60	0.61

slope with uniform parameters; however, the difference in model performance between 3 and 10 hillslopes was minor. While the tradeoff between number of hillslope partitions, model complexity and level of improvement in model performance is important, this issue is not investigated in this study. For these simulations all hillslopes were discretized using 100 m Δx intervals. The length of hillslopes ranged between 1900 and 10,400 m for ICL and 1300–7300 m for LRN.

5.2. Model calibration and verification

Because models are a simplified representation of real world systems, models are associated with errors and uncertainties (Gupta et al., 2005). For most RR models the accuracy of the model output depends, among others, on the following key aspects: the model's ability to correctly represent the hydrologic processes involved, the quality of the input data, the calibration of model parameters, and the quality of observed output data to evaluate model performance (DeVries and Hromadka, 1993; Gupta et al., 2005). Assuming the model structure is correct; calibration is a process by which model parameters are estimated to minimize model error so that the model best represents the behavior of the system under consideration (Gupta et al., 1998, 2005). Often optimization techniques are applied to: (i) estimate unknown physical parameters, (ii) adjust for effective parameter values when available parameter measurements and model spatial domain are different (for example hillslope hydraulic conductivity as discussed in Brooks et al. (2004)) or the parameter has a high spatial variability within the model domain, and (iii) introduce correction factors (multipliers or additive constants) to account for causal variables known to be part of the physical system, but not entirely described by the model (Morton and Suarez, 2001).

One important aspect for a successful calibration process is the selection of calibration data. A previous study by Sorooshian and Gupta (1995) found that both the quality and length of calibration data can have an impact on model performance. As a rule of thumb it is suggested that the minimum calibration data length should be 20 times the number of calibration parameters, and the data should be representative of the variability in watershed behavior (Sorooshian and Gupta, 1995). An initial analysis of observed rainfall and streamflow data in the study area revealed that the period between 1964 and 1969 included a combination of wet, normal and dry years for both the LRN and ICL watersheds. The attribution of a yearly hydrologic regime (dry, normal, and wet) is relative to the entire period of record available. Though based on Sorooshian and Gupta's (1995) recommendation two years of data would have been sufficient, the selection of five water years was justified in order to maintain the same period of calibration in which at least one normal, one wet and one dry year were included for both watersheds. In addition, by comparing results from five and one year calibration periods we initiate an analysis of tradeoffs between calibration data length and hydrologic regime characteristics (wet, normal, or dry regime), and hillslope-storage model

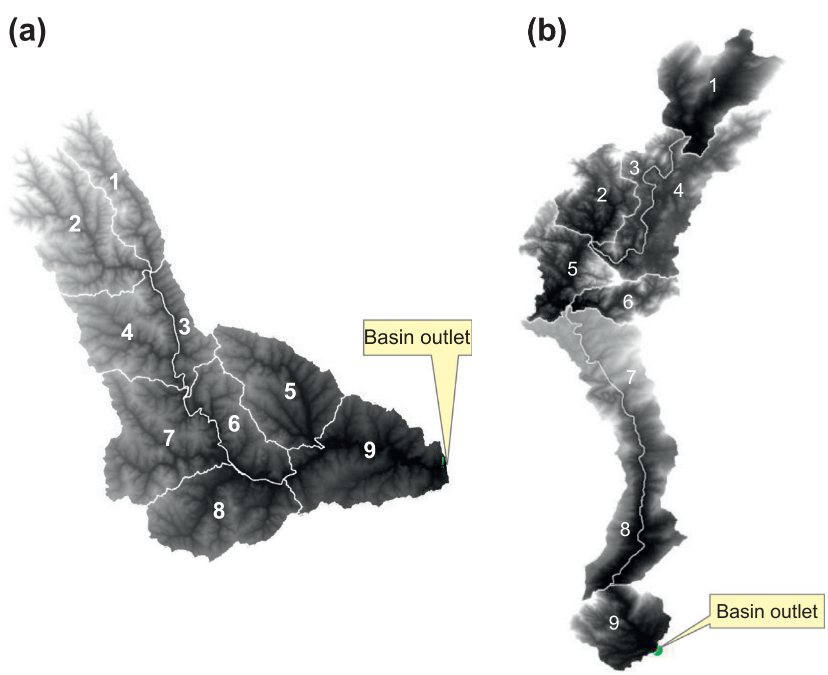

Fig. 2. The ICL (a) and LRN (b) watersheds with 9 hillslope partitions.

performance when estimating low flow statistics. Tables 3 and 4 present selected annual streamflow statistics for the calibration years at LRN and ICL, respectively. In general and as expected, streamflow statistics at the shallower sloped ICL watershed are of lower magnitude than at LRN given the differences in slope and average annual precipitation (Table 1).

For the present analysis, model calibration was performed using the Shuffled Complex Evolution (SCE) algorithm (Duan et al., 1992, 1993). The SCE algorithm has been shown to be a robust, effective, and efficient method for finding the global optimal set of parameters for many hydrologic models (Duan et al., 1992, 1993; Gan and Biftu, 1996; Gupta et al., 1998). The following 8 parameters were calibrated to be constant for all hillslopes in each watershed: minimum impervious area (PCTIM), additional impervious area (ADIMP), ratio of deep recharge to channel baseflow (SIDE), fraction of lower zone free water not transferable to lower zone tension water (RSERV), beta and alpha parameters for the gamma-function in the unit hydrograph calculations, a parameter for channel routing (only for partitioned models; does not apply to *Sac*), and a factor setting the amount of saturation excess water that is returned to the upper layer. For the *Sac* model, the 2 baseflow recession parameters in the lower soil layer were calibrated as being constant for the entire watershed; in the *Sacm* model application these parameters were calibrated for each hillslope. For the *kw* and *hsB* models, saturated hydraulic conductivity (k_S) and soil porosity (f) were calibrated for each of the 9 individual hillslopes, while the constant μ for the hydraulic conductivity power function decrease in Eq. (5) was assumed constant for all hillslopes.

Model calibration was performed at a daily time step by minimizing the Root Mean Square Error (*RMSE*) (Lettenmaier and Wood, 1993; Wagener et al., 2004):

$$RMSE = \left[\frac{\sum_{i=1}^{n}(O_i - S_i)^2}{n}\right]^{\frac{1}{2}} \tag{6}$$

where O_i [L] and S_i [L] are the observed and simulated streamflows at day i, respectively, and n is the number of streamflow days. The lower the *RMSE*, the better the model fit to measured streamflow. As a calibration stopping criteria we applied the same method of function convergence described by Sorooshian and Gupta (1995)

$$(f_{i-1} - f_i)/f_i \leqslant \varepsilon_f \tag{7}$$

where f_{i-1} and f_i are the best function values from the previous and actual iteration step, respectively, and ε_f is the convergence criterion, here set equal to 10^{-4}.

For model verification a thirty year period (1971–2000) was selected. A Scaled Root Mean Square Difference (*SRMSE)* was applied to evaluate model performance at the watershed outlet for the lower 20% and 50% of streamflow (i.e. comparing simulated versus observed streamflow equal or smaller than the 20th or 50th percentiles of the observed streamflow distribution, respectively), and the full range of streamflow. In addition we compared the average Q_7 and Q_{30}, as well as the 7-day, 10-year ($Q_{7,10}$) and 30-day, 2-year ($Q_{30,2}$) low flow statistics. The scaled Root Mean Square Difference is calculated as

$$SRMSE = \left[\frac{(n-1)}{n}\frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}\right]^{\frac{1}{2}} \tag{8}$$

where $\bar{O}$ is the mean of the observed values. *SRMSE* [−] represents the *RMSE* divided by the standard deviation of the observed streamflow over the validation range. While the *RMSE* provides a measure of the differences between model and measured data and is easy to interpret when evaluating the models fit to the same range of streamflow, it is scale dependent. We apply the *SRMSE* to indicate the relative magnitude of model errors to the standard deviation of the measured streamflow in the verification range. The *SRMSE* is useful when comparing model performance across different ranges of data.

Table 3
Annual statistics for the streamflow discharge for each of the 5 years applied during model calibration in LRN watershed.

Statistic (mm/day)	Water year				
	1964–65	1965–66	1966–67	1967–68	1968–69
Average discharge	2.66	1.68	1.91	1.71	1.87
Minimum discharge	0.41	0.38	0.42	0.38	0.27
Maximum discharge	32.3	56.8	14.5	19.0	39.3
20th Percentile	0.93	0.68	1.01	0.83	0.73
Q7	0.49	0.42	0.51	0.41	0.28
Q30	0.65	0.59	0.88	0.63	0.37

Table 4
Annual statistics for the streamflow discharge for each of the 5 years applied during model calibration at ICL watershed.

Statistic (mm/day)	Water year				
	1964–65	1965–66	1966–67	1967–68	1968–69
Average discharge	1.93	1.23	0.70	1.30	0.93
Minimum discharge	0.45	0.45	0.18	0.35	0.22
Maximum discharge	35.0	19.5	4.62	11.9	12.1
20th Percentile	0.75	0.56	0.41	0.53	0.36
Q7	0.47	0.47	0.21	0.40	0.23
Q30	0.62	0.54	0.32	0.54	0.27

6. Results and discussion

6.1. Model calibration using five years of streamflow data

6.1.1. LRN

Initial results are presented for LRN, the steeper and wetter watershed. Fig. 3 shows the *RMSE* values for the four models run over a five year calibration period for LRN.

The four models are the *Sac* lumped on the entire watershed with 11 a priori and 9 calibrated parameters, *Sacm* run on a hillslope scale with 11 a priori, 8 calibrated constants, and 2 varying parameters per hillslope, *Sackw* (*kw*) using 11 a priori, 8 calibrated constant *Sac* model parameters, and 2 calibrated *kw* parameters

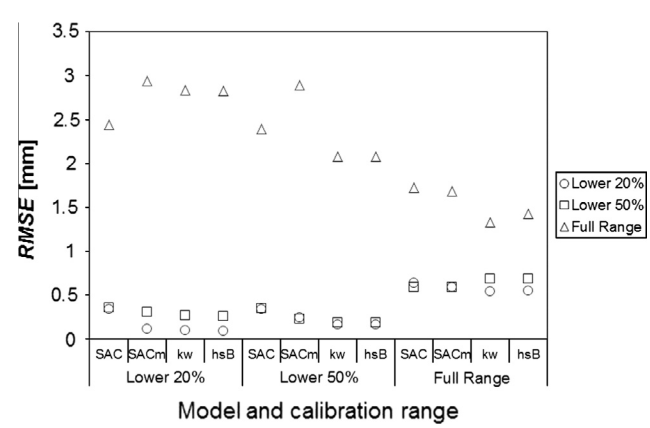

Fig. 3. Calibration *RMSE* for the LRN watershed with the *Sac*, *Sacm*, *kw*, and *hsB* models calibrated using varying ranges of streamflow (shown in the *x*-axis) at the watershed outlet over a 5 year calibration period.

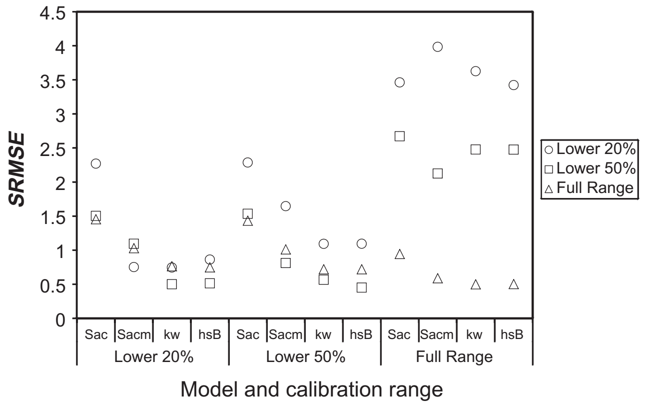

Fig. 4. *SRMSE* over a 30-year verification period for the LRN watershed with the *Sac*, *Sacm*, *kw*, and *hsB* models calibrated using varying ranges of streamflow over the 5-year period.

per hillslope, and *SachsB* (*hsB*) with 11 a priori, 8 calibrated constant *Sac* model parameters and 2 calibrated *hsB* model parameters per hillslope. The four models were calibrated using the full range (all data), and the lower 20% and 50% of streamflow (*x*-axis in Fig. 3) and were evaluated over the same range as the calibration. As expected, models calibrated using the lower 20% of data show the best *RMSE* for streamflow data below the 20th percentile and models calibrated with the lower 50% of data exhibit a better performance than all other models when describing streamflow under the 50th percentile. All models calibrated with the full range exhibit the best performance to describe the full streamflow hydrograph, but they perform poorly in describing data at lower ranges. These results are similar to those of Matonse and Kroll (2009), showing that as the range of data used during calibration increases (particularly above the 50th percentile), the difficulty for the model in describing data at lower ranges also increases. For all evaluated streamflow ranges models calibrated at a hillslope scale performed better than *Sac* (with lumped parameters across the entire watershed). Though *Sacm* performs better than *Sac* across all calibration ranges, the *RMSE* associated with this model was always slightly higher than for *kw* or *hsB*. At the LRN watershed the *kw* and *hsB* models exhibit similar performance at all ranges of streamflow.

Fig. 4 shows the *SRMSE* results over the thirty year (1971–2000) verification period. A *SRMSE* value less than 1 indicates that the *RMSE* is less than the standard deviation of the observed streamflow in the corresponding range. The *Sac* model shows overall a poor performance at all ranges of streamflow. Interestingly the partitioned models calibrated using data below the 50th percentile perform well at all ranges, with the hillslope-storage models showing better results when calibrated using the lower 50% of streamflow. Models calibrated using the full range of streamflow continue to show poor performance at predicting the lower range of streamflow.

Fig. 5 summarizes modeled and observed low flow statistics for the thirty year verification period, including average Q_7 and Q_{30}, as well as $Q_{7,10}$ and $Q_{30,2}$.

As one would expect given the difficulty in fitting low flow data, the *Sac* model performs poorly in simulating low flow statistics. *Sacm* and the coupled models calibrated with data below the 20th percentile can predict $Q_{7,10}$ relatively well, with *Sacm* and *hsB* model showing slightly better performance than the *kw* model; both *kw* and *hsB* perform much better than the *Sac* model. It appears from these results that partitioned models calibrated with the lower 20 and 50% of streamflow exhibit a better ability to estimate the average Q_7 and Q_{30} as well as $Q_{7,10}$ and $Q_{30,2}$ than the *Sac* model. At a type I error $\alpha = 0.05$, one would reject the null hypothesis (Ho) that observed and modeled average Q_7 and Q_{30} are equal except for the average Q_{30} of the partitioned models calibrated with the lower 20% of data. These hypothesis tests were performed with results from the verification period. In general all partitioned (*Sacm, kw* and *hsB*) models simulate both low flow statistics better when calibrated using streamflows below the 20th percentile, although there is a trend of slightly overestimating $Q_{7,10}$ and underestimating $Q_{30,2}$. Coupled models calibrated with the lower 50% of streamflow can also simulate $Q_{30,2}$ relatively well (within a 1 mm/day range difference), but with the hillslope coupled models showing a slightly better performance. Models calibrated using the full range of streamflow data systematically underestimate both low flow statistics, suggesting a difficulty of these models to describe the lower range data while trying to fit to high streamflow values (Matonse and Kroll, 2009). When calibrated with the full range the partitioned models are as bad as the *Sac* in estimating low flow statistics.

6.1.2. ICL

The following results are for ICL, the shallower sloped and drier watershed. Based on the *RMSE* values from the five year calibration period presented in Fig. 6, the partitioned models continue to perform similarly and better than a lumped *Sac* model when simulating the lower ranges of streamflow. The *kw* and *hsB* models

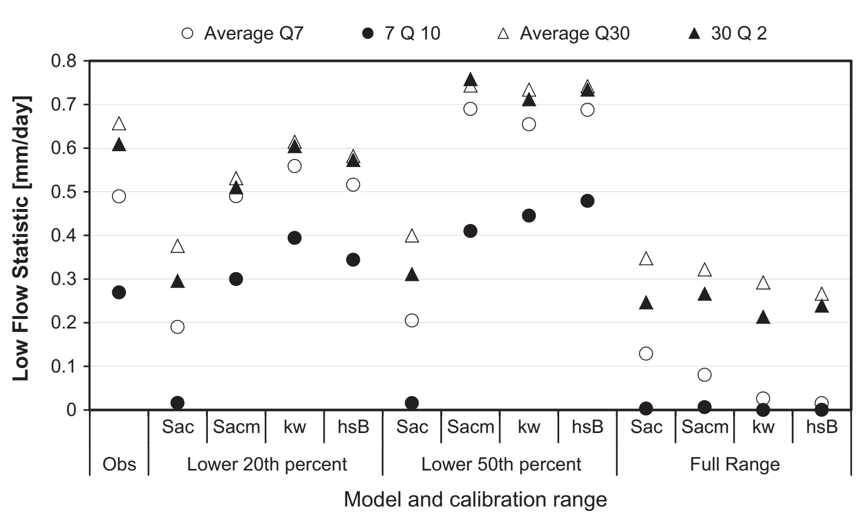

Fig. 5. Low flow statistics for a relatively steep slope LRN watershed over the 30 year verification period using models calibrated with 5 years of streamflow data and varying streamflow ranges.

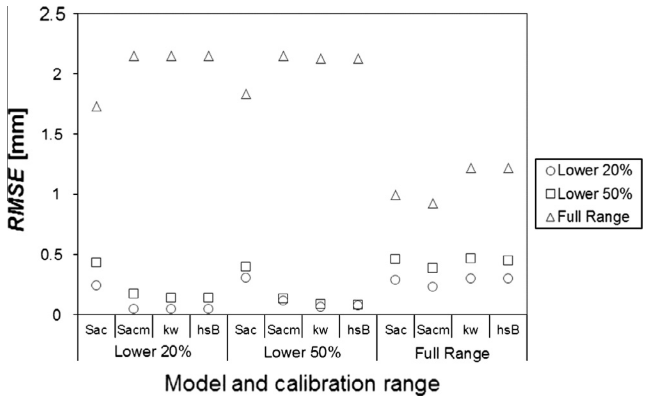

Fig. 6. *RMSE* for the ICL watershed with the *Sac*, *Sacm*, *kw*, and *hsB* models calibrated using varying ranges of streamflow at the watershed outlet over the 5 year calibration period.

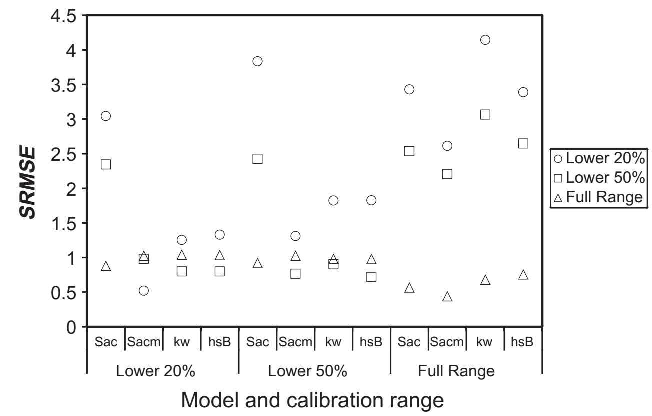

Fig. 7. *SRMSE* over a 30-year verification period for the shallower sloped ICL watershed with the *Sacm*, *kw*, and *hsB* models calibrated using varying ranges of streamflow over a 5-year period.

perform slightly better than *Sacm* for the lower 50% of streamflow but *Sac* and *Sacm* appears to better fit the full range of streamflows when calibrated using the full range of streamflows. Overall, these results suggest that calibrating *Sac* using a lower percentage of streamflows does not improve the model's ability to simulate the lower range of streamflows.

Fig. 7 compares the performance of all four models over the 30-year verification period. As in LRN, here again the *Sac* model has a poor performance at the lower percent streamflow. All three partitioned models calibrated using the lower 20% of streamflow show a *SRMSE* between 0.5 and 1.5 for the lower 20% of streamflow but with *Sacm* exhibiting the lowest SRMSE, a sign that the calibration process allows this model to have enough flexibility to fit streamflow at this range relatively well. The *hsB* model calibrated with the lower 50% of streamflow shows the best performance for a similar streamflow range over the verification period, but all models had a *RMSE* that is smaller than a standard deviation of the observed streamflow. All models calibrated using the full range of streamflow perform similarly over the full range of streamflow but poorly for the streamflow below the 50th percentile.

The performance of coupled models in predicting low flow statistics is shown in Fig. 8. As with the steeper sloped LRN watershed, partitioned models calibrated using the lower 20% of streamflow simulate ICL low flow statistics that are closer to the observed values. However, the overestimation of $Q_{7,10}$ is larger for the shallower sloped ICL compared to LRN. Interestingly, the *Sacm* model estimates are slightly higher than *kw* and *hsB*, suggesting that a small SRMSE with the lower 20% of streamflow in Fig. 7 is a result of a better fit to the higher values in the range. These results may be an indication that to better simulate $Q_{7,10}$ we may need to calibrate our models using a threshold lower than the 20th percentile of streamflows or, possibly, include Q_7 as one of the calibration metrics. Another possible reason for the overestimation of $Q_{7,10}$ may be that during extreme low flow events, these rivers may be losing water to the groundwater system, a situation that is not accounted for in the presented model formulations. During the same verification period partitioned models calibrated using the lower 20 and 50% of streamflow show better $Q_{30,2}$ predictions with models calibrated using the lower 20% showing the best statistics. While the *kw* and *hsB* models calibrated with the lower 50% of streamflow show close Q_{30} and $Q_{30,2}$ results, the values from the *hsB* model appear to be slightly larger than from the *kw* model (although slightly smaller than *Sacm*), indicating the possible effect of the diffusive term in the *hsB* model at this relatively shallow sloped watershed.

Hypothesis tests comparing mean Q_7 and Q_{30} values from ICL resulted in rejecting the hypothesis that observed and modeled averages are equal for Q_7, except for the partitioned models calibrated with the lower 20% of streamflow. For Q_{30} the test failed

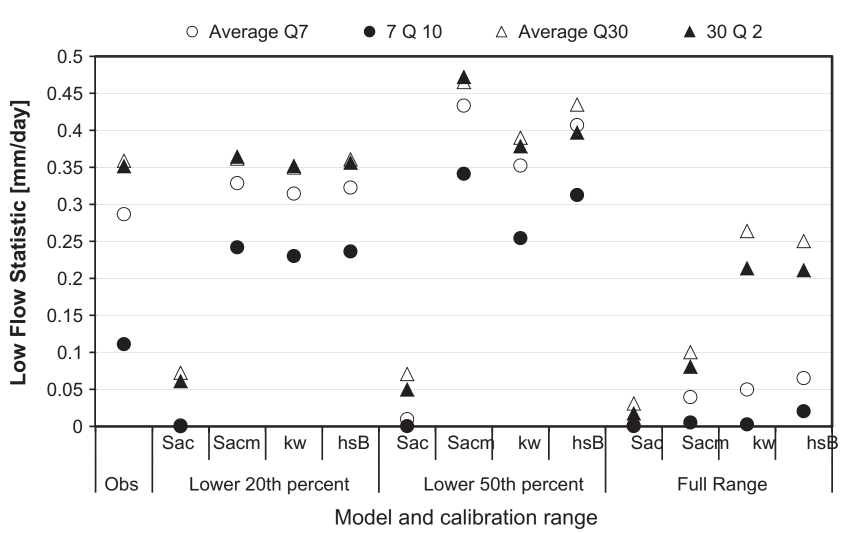

Fig. 8. Low flow statistics for the ICL watershed over the 30 year verification period using models calibrated with 5 years of streamflow data and varying streamflow ranges.

to reject the null hypothesis that simulated averages are equal to observed averages for all partitioned models calibrated with the lower 20% of streamflow and for the *kw* model calibrated using the lower 50% of streamflow. All models calibrated using the full range and *Sac* using any range of streamflow performed poorly in simulating ICL low flow statistics.

6.2. Model calibration using one year of streamflow data in the ICL watershed

To analyze the impact of reducing the amount of data and the data regime characteristics for calibration on model results, the *kw* and *hsB* models were calibrated at the ICL watershed using three different 1 year periods representing the following hydrologic regimes: a dry year (1968), a normal year (1967), and a wet year (1964). For this analysis we focus on coupled hillslope-storage models only. Note that a low flow hydrologic year was employed (April 1 – March 31). In these simulations, all models were calibrated using only the lower 20% of streamflow, which appears warranted given the results from calibrating using 5 years of data presented in Section 6.1.

The model's *SRMSE* over the 30 year verification period is shown in Fig. 9. In this figure, results for 1 year calibrations are also compared to those with 5 year calibrations. For the normal year, models with 1 year of calibration data have similar performance to models calibrated using 5 years of streamflow. However, for the lower 20% range of streamflow models calibrated using the wet year show overall the best (lowest and <1) *SRMSE* values, with this performance decreasing from wet to normal and normal to dry years. These results suggest that because the five year period includes wet, dry and normal years its combined behavior is similar to a normal year, whereas the verification data is more representative of a wet period. Also, the results in Fig. 9 indicate that model performance at the lower 20% of streamflow is less sensitive to calibration data length compared to regime characteristic. However, for larger streamflows the model performance did not change at all except with dry year data for the *kw* model which shows a slightly poorer performance for the lower 50% streamflow.

Fig. 10 presents the predicted average Q_7 and Q_{30}, and the $Q_{7,10}$ and $Q_{30,2}$ for the 30 year verification period. These results indicate that all models overestimate $Q_{7,10}$ over the 30 year period. Overall, for the 30 year verification period the *kw* model calibrated using streamflow from a dry year and both models calibrated using the wet year data exhibit the best Q_7 and $Q_{7,10}$ estimates, while models calibrated with five years of streamflow show the best Q_{30} and $Q_{30,2}$ estimates.

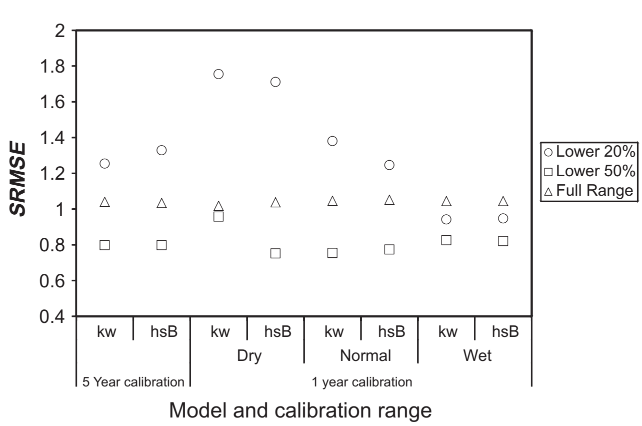

Fig. 9. *SRMSE* for ICL watershed over the 30 year verification period with the *kw* and *hsB* models calibrated using the lower 20th percentage of streamflow from 5 and 1 year periods.

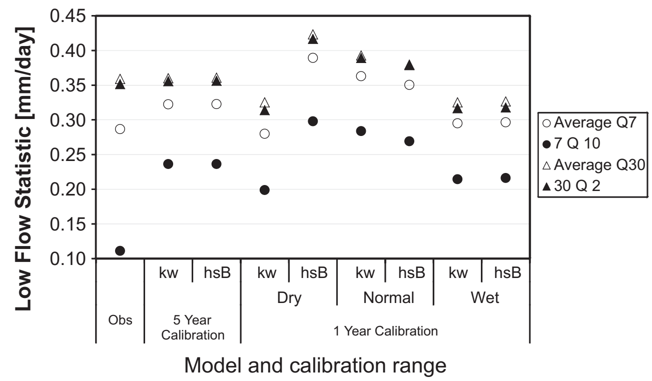

Fig. 10. Low flow statistics for the ICL watershed over the 30 year verification period with the *kw* and *hsB* models calibrated using the lower 20th percentage of streamflow from 5 and 1 year periods.

Though these results represent only the particular case of the ICL watershed and the selected validation period, the overall performance of the *kw* model calibrated with one year of streamflow data for a relatively dry year and the *kw* and *hsB* models for a relatively wet year show similar performance as the models calibrated with five years of streamflow data. The *hsB* model, though, is overestimating streamflow for the verification period. This may be due to the diffusion term in this model. In general Q_{30} and $Q_{30,2}$ model estimates from 5-year and wet and normal 1-year calibrations are relatively good, differing by less than 0.05 mm/day from the observed values. These results raise important questions about the length and characteristics of streamflow data necessary to calibrate partitioned models intended for low flow prediction. An important question is whether hydrologic regime characterization based on yearly average is the right approach for low flow modeling. Further investigation is necessary to test whether these results are due to this particular simulation or represents a more general behavior that is likely to occur at other watersheds with varying hydrogeological characteristics.

6.3. Hydraulic conductivity

In Fig. 11 we compare the box plots of calibrated hillslope saturated hydraulic conductivities for the *kw* and *hsB* model simulations at ICL and LRN. These are effective saturated conductivities that include the effect of macropore flow at a hillslope scale

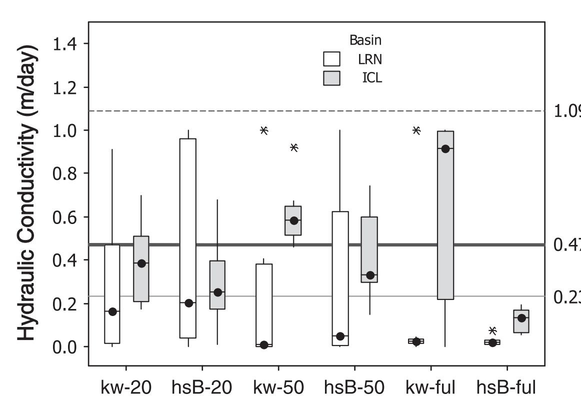

Fig. 11. Box plots showing the median (horizontal line in the boxes), mean (dark dots), interquartile, whiskers and outliers of hydraulic conductivity (k) for the nine hillslopes for the LRN and ICL watersheds with the *kw* and *hsB* models calibrated with 5 years of streamflow data and varying streamflow ranges. Also shown in the graph are three lines representing average k values from drilled wells at different depths for valleys and draws (dashed line), slopes (middle dark line), and hills and ridges (lower gray line) across the study region (Daniel et al., 1997).

(Brooks et al., 2004). In general the average saturated hydraulic conductivity at ICL is higher than at LRN with higher rainfall and steeper slopes. The variability of saturated hydraulic conductivity for the different calibration ranges is larger for LRN than for ICL, except when calibrating at the full range. Unlike at LRN, for the shallower sloped ICL watershed *hsB* models show lower mean hydraulic conductivity than the *kw* models. This indicates the calibration technique is allowing the *kw* model to compensate for the absence of the diffusion term by converging to higher *k* values in the shallower sloped watershed. Though for ICL the *kw* and *hsB* models showed relatively close *SRMSE* values (Fig. 7), Fig. 11 shows how the *kw* calibrated with the lower 50% and the full range of streamflow converged on average to hydraulic conductivity values that are higher than average values reported in previous studies for drilled wells in slope areas at this watershed (Daniel et al., 1997). Higher hillslope saturated hydraulic conductivity than small-scale k_s measurements is not surprising and it is extensively discussed in Brooks et al. (2004). Though, our calibrated saturated hydraulic conductivity values are, in general, within the range of measured values at various locations across this region as reported in Daniel et al. (1997), this analysis is limited by the fact that drainable porosity f and the power constant μ, which have an impact on hillslope-storage baseflow modeling, were both allowed to change during the calibration process. While, this procedure provides more flexibility in capturing heterogeneous properties across a watershed (Harman and Sivapalan, 2009), it makes capturing model sensitivity to a single parameter (hydraulic conductivity) more difficult.

A guidance criteria for the kinematic wave validity (Beven, 1981; Fan and Bras, 1998) is the λ index, a ratio of rainfall accumulation to soil water drainage (Henderson and Wooding, 1964)

$$\lambda = \frac{4N\cos i}{k\sin^2 i} \tag{9}$$

where N is the recharge, k is the saturated hydraulic conductivity, and i is the slope of the impermeable bed, should be less than 0.75. Based on an average total annual precipitation of 1207 mm with an estimate of 80 1-day events per year and a 4.6° average slope, for λ to be less than or equal to 0.75 the hydraulic conductivity must be at least 12 m/day. Though this value is very uncertain and is sensitive to the assumed number of rainfall events per year and recharge amounts (in our calculations we assumed all precipitation becomes recharge), this hydraulic conductivity value is very high given the characteristics of the soils at ICL. The results in Fig. 11 indicate values that are less than the minimum k value required for the validity of the kinematic model, based on the λ index. Regardless of these results, in our simulations the *kw* and *hsB* models perform similarly for watersheds with different slopes. One explanation is that the calibration process is flexible enough to allow the *kw* models to converge to parameter values to adequately model low flow in shallower sloped watersheds.

7. Summary and conclusions

In this paper we present a coupling between the Sacramento Soil Moisture Accounting model (*Sac*) and the hillslope-based *kw* and *hsB* models. In addition we apply the *Sac* model on multiple hillslopes (*Sacm*). While SAC-SMA accounts for most basin characteristics and the overall mass balance, the hillslope sub-models replace the lower storages in the original *Sac* model to simulate baseflow discharge to a stream. *Sacm* and the new coupled models (*kw* and *hsB*) are run on a hillslope basis on a 6 h time step to generate daily streamflow discharge at the hillslope outlet. Contributions from each of nine individual hillslopes are routed along a channel network to the basin outlet. These coupled models were applied to two watersheds: the Linville River near Nebo (LRN) and Indian Creek near Laboratory (ICL), and compared to each other and the original *Sac* model as well as to *Sacm*. While the location and size of both watersheds is similar, LRN has an average slope of 14.8°, while ICL has an average slope of 4.6°. Data from a 5-year period that included wet, normal and dry hydrologic years were applied to calibrate the models. In addition, models were also calibrated at the shallower sloped ICL watershed using 1 year of streamflow data. Models were calibrated to minimize the Root Mean Square Difference (*RMSE*) using the full range of streamflow and the lower 20% and 50% of streamflow. Calibrated models were then applied to examine their performance over three ranges of streamflow using a scaled *RMSE* (*SRMSE*) and their ability to predict the lower 20%, 50% and full range of streamflow, and 7-day and 30-day low flow statistics during a 30 year verification period.

Our results indicate that:

1. The use of hillslope-storage baseflow models improves the SAC-SMA model's ability to simulate low flow at a watershed scale. Except at the full range for the shallower sloped and drier ICL watershed the lumped *Sac* model performs poorly for all ranges of streamflow when compared to partitioned models. This was expected as the partitioned models increase the number of model parameters to be calibrated. The performance of all partitioned models is similar, with the coupled models being slightly better in simulating the lower 50% of streamflow when calibrated with the same range of streamflow.
2. The change in calibration range has little or no effect on *Sac* performance at a lower range streamflow.
3. The coupled models perform better than *Sac* and *Sacm* when calibrated at a full range for the steeper LRN watershed. Conversely, *Sac* and *Sacm* perform better at the full range when applied to ICL watershed. However, it appears that the improvement on the full range of streamflow comes with a decrease in model performance for streamflow below the 50th percentile and low flow statistics. As suggested in Staudinger et al. (2011), this behavior may indicate the need to improve model structure in order to achieve a better and more balanced model performance.
4. A comparison between $Q_{7,10}$ and $Q_{30,2}$ low flow statistics calculated from observed and modeled streamflow indicates that partitioned models calibrated with the lower 20% of streamflow can better predict low flow statistics. Our results do not show any evidence of a difference in performance between *Sacm*, *kw* and *hsB* models. This conclusion was not reached by Matonse and Kroll (2009) due to the limited data available for their study. Based on these results it appears that model calibration with the lower 20% of streamflow leads to better $Q_{7,10}$ predictions. When estimating $Q_{30,2}$ the results are mixed, indicating that calibration using the lower 20% and 50% streamflow both have the potential for good prediction.
5. Results from the *kw* and *hsB* models were similar for both watersheds. This was surprising, as we expected at the shallower sloped ICL to observe differences in these models. Unless we have knowledge to restrict the range of effective hydraulic conductivity (k) values, the calibration process gives the *kw* model enough flexibility to compensate for the absence of the diffusive term by selecting higher k values, leading to results that are similar to the *hsB* model. Another important factor is that values for parameters such as the hydraulic conductivity are very sensitive to hillslope depth. It would be of interest to investigate the validity of the hillslope depth similarity assumption at these watersheds. Also, the impact of fractured bedrock and regional groundwater systems (Broda et al., 2011a, 2011b) were not investigated in this study. In theory one would expect that *Sacm* exhibit better results because the two calibrated

recession coefficients should account for both regional and local groundwater flow. Coupled models potentially have an advantage in representing water table levels at the hillslope, if necessary.
6. Our results from models calibrated with 1 year of data indicate the model performance is more sensitive to data characteristics than to data length. However, it remains unclear what would be the best indicator to characterize the hydrologic regime (wet, dry, normal, or a combination) for data intended for modeling low flow. The coupled *kw* and *hsB* models calibrated with 1-year of streamflow can perform relatively well in predicting low flow series and statistics. A traditional method to estimate low flow statistics at partially gauged sites, baseflow correlation, is based on having a nominal number of observations (measurements) of low flow at the ungauged site and from a concurrent gauged site (Reilly and Kroll, 2003; Zhang and Kroll, 2007). Previous studies have also indicated that the inclusion of baseflow indices improved predictions of low flow statistics using regression methods (Riggs, 1961; Vogel and Kroll, 1996; Kroll et al., 2004). The performance of coupled models with limited streamflow data is encouraging and can potentially improve the estimation of low flow statistics at partially gauged sites.

Of interest for further investigation is what length of calibration data is sufficient to produce "good" estimators of low flow statistics with our coupled models, and whether that data is better employed with other methods such as baseflow correlation or regional regression. Given our focus on low flow, of further interest is the regime characteristic of streamflow employed for model calibration and the best metrics to guide model calibration intended to simulate low flow statistics, while preserving good performance in the full streamflow hydrograph. Combining low flow statistics with other criteria such as recession (Staudinger et al., 2011) and soil moisture in future studies involving hillslope-storage models can be a valuable exercise to improve model structure and evaluate RR models performance in simulating the full streamflow hydrograph.

Acknowledgements

The authors want to thank Professor Richard Vogel (Tufts University, MA) and the anonymous reviewers for their comments and suggestions that helped improve this manuscript.

References

Beven, K., 1981. Kinematic subsurface stormflow. Water Resour. Res. 17 (5), 1419–1424.
Bogaart, P.W., Torch, P.A., 2004. On the use of soil-landscape evolution modeling in understanding the hillslope hydrological response. Hydrology: Science and Practice for the 21st Century, vol. 1. Br. Hydrol. Soc., London, pp. 251–259.
Bogaart, P.W., Troch, P.A., 2006. Curvature distribution within hillslopes and catchments and its effect on the hydrological response. Hydrol. Earth Syst. Sci. 10, 925–936.
Boussinesq, J., 1877. Essai sur la theorie des eaux courantes. Mem. Acad. Sci. Inst. France 23 (1), 252–260.
Broda, S., Paniconi, C., Larocque, M., 2011a. Numerical investigation of leakage in sloping aquifers. J. Hydrol. 409, 49–61. http://dx.doi.org/10.1016/j.jhydrol.2011.07.035.
Broda, S., Larocque, M., Paniconi, C., Haitjema, H., 2011b. A low-dimentional hillslope-based catchment model for layered groundwater flow. Hydrol. Process.. http://dx.doi.org/10.1002/hyp.8319.
Brooks, E.S., Boll, L., McDaniel, P.A., 2004. A hillslope-scale experiment to measure lateral saturated hydraulic conductivity. Water Resour. Res. 40, W04208. http://dx.doi.org/10.1029/2003WR002858.
Brutsaert, W., Nieber, J.L., 1977. Regionalized drought flow hydrographs from a mature glaciated plateau. Water Resour. Res. 13, 637–643.
Burnash, R.J.C., 1995. The NWS river forecast system – catchment modeling. In: Singh, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publications, Colorado, USA, pp. 311–366.
Caruso, B.S., 2002. Temporal and spatial patterns of extreme low flows and effects on stream ecosystems in Otago, New Zealand. J. Hydrol. 257, 115–133.
Clausen, B., Pearson, C.P., 1995. Regional frequency analysis of annual maximum streamflow drought. J. Hydrol. 173 (1–4), 111–130. http://dx.doi.org/10.1016/0022- 1694(95)02713-Y.
Daniel, C.C., III, Dahlen, P.R., 2002. Preliminary hydrogeologic assessment and study plan for a regional ground-water resource investigation of the Blue Ridge and Piedmont Provinces of North Carolina: U.S. Geological Survey Water-Resources Investigations 02-4105, p. 60.
Daniel, C.C., III, Sharpless, N.B., 1983. Ground-water supply potential and procedures for well-site selection in the upper Cape Fear River Basin, North Carolina: North Carolina Department of Natural Resources and Community Development, p. 73.
Daniel, C.C., III, Smith, D.G., Eimers, J.L., 1997. Hydrogeology and simulation of groundwater flow in the thick regolith-fractured crystalline rock aquifer system of Indian Creek Basin, North Carolina, U.S. Geological Survey Water Supply Paper 2341, pp. 137. <http://pubs.usgs.gov/wsp/2341c/report.pdf> [I19.13:2341-C].
DeVries, J.J., Hromadka, T.V., 1993. Computer models for surface water. In: Maidment, D.R. (Ed.), Handbook of Hydrology. McGraw-Hill, Inc., USA (Chapter 21).
Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28 (4), 1015–1031.
Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optim. Theory Appl. 76 (3), 501–521.
Duan, Q.Y., Schaake, J., Koren, V., 2001. A priori estimation of land surface model parameters. In: Lakshmi, V., et al. (Eds.), Land Surface Hydrology, Meteorology, and Climate: Observations and Modeling, Water Science and Application 3, American Geophysical Union, Washington, DC, pp. 77–94.
Engeland, K., Hisdal, H., 2009. A comparison of low flow estimates in ungauged catchments using regional regression and the HBV-model. Water Resour. Manage. 23 (12), 2567–2586.
Fan, Y., Bras, R.L., 1998. Analytical solutions to hillslope subsurface storm flow and saturation overland flow. Water Resour. Res. 34 (4), 921–927.
Gan, T.Y., Biftu, G.F., 1996. Automatic calibration of conceptual rainfall-runoff models: optimization algorithms, catchment conditions, and model structure. Water Resour. Res. 32 (12), 3513–3524.
Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. Water Resour. Res. 34 (4), 751–763.
Gupta, H.V., Beven, K.J., Wagener, T., 2005. Model calibration and uncertainty estimation. In: Anderson, M. (Ed.), Encyclopedia of Hydrological Sciences. John Wiley & Sons. Lda., p. 142.
Harman, C., Sivapalan, M., 2009. Effects of hydraulic conductivity variability on hillslope-scale shallow subsurface flow response and storage-discharge relations. Water Resour. Res. 45, W01421. http://dx.doi.org/10.1029/2008WR007228.
Henderson, F.M., Wooding, R.A., 1964. Overland flow and groundwater flow from a steady rainfall of finite duration. J. Geophys. Res. 69 (8), 1531–1540.
Hilberts, A.G.J., Troch, P.A., Paniconi, C., Boll, J., 2007. Low-dimensional modeling of hillslope subsurface flow: relationship between rainfall, recharge, and unsaturated storage dynamics. Water Resour. Res. 43, W03445. http://dx.doi.org/10.1029/2006WR004964.
Hisdal, H., Tallaksen, L.M., Clausen, B.P., Gustard, A., 2004. Hydrological Drought Characteristics. In: Tallaksen, L.M., and van Lanen, H.A.J., (Eds), Hydrological Drought Processes and Estimation Methods for Streamflow and Groundwater. Developments in Water Sciences 48. Elsevier Science Publisher, the Netherlands, pp. 139–198.
Hogue, T., Wagener, T., Schaake, J., Duan, Q., Hall, A., Gupta, H.V., Leavesley, G., Andreassian, V., 2004. A new phase of the model parameter estimation experiment (MOPEX), *Eos*. Trans. AGU 85 (22), 217–218.
Karim, K., Gubbels, M.E., Goulter, I.C., 1995. Review of determination of instream flow requirements with special application to Australia. Water Resour. Bull. 31 (6), 1063–1077.
Kroll, C.N., Vogel, R.M., 2002. Probability distribution of low streamflow series in the United States. J. Hydrol. Eng. 7 (2), 137–146.
Kroll, C.N., Luz, J., Allen, B., Vogel, R.M., 2004. Developing a watershed characteristics database to improve low streamflow prediction. J. Hydrol. Eng. 9 (2), 116–125.
Laaha, G., Skøien, J., Blöschl, G., 2007. A comparison of Top-kriging and regional regression for low flow regionalisation. Geophys. Res. Abstr. 9 (07015), 1–2, ISSN 1029-7006.
Van Lanen, H.A.J., Tallaksen, L.M., Kašpárek, L., Querner, E.P., 1997. Hydrological drought analysis in the Hupsel basin using different physically-based models. In: Gustard, A., et al. (Eds.), FRIEND'97-Regional Hydrology: Concepts and Models for Sustainable Water Resource Management. IAHS Publ. No. 246, pp. 189–196.
Lettenmaier, D.P., Wood, E.F., 1993b. Hydrologic forecasting. In: Maidment, D.R. (Ed.), Handbook of Hydrology. McGraw-Hill, Inc., USA (Chapter 26).
Madsen, H., Rosbjerg, D., 1998. A regional Bayesian method for estimation of extreme streamflow droughts. In: Parent, E., Bobee, B., Hubert, P., Miquel, J., (Eds.), Statistical and Bayesian Methods in Hydrological Sciences. UNESCO, PHI Series, pp. 327–340.
Matonse, A.H., Kroll, C.N., 2009. Simulating low streamflows with hillslope storage models. Water Resour. Res. 45, W01407. http://dx.doi.org/10.1029/2007WR006529.

Metcalf & Eddy, Inc., 1991. Wastewater Engineering – Treatment, Disposal, and Reuse, third ed. Irwin/McGraw-Hill, New York, USA.
Morton, A., Suarez, M., 2001. Kinds of models. In: Anderson, M.G., Bates, P.D. (Eds.), Model Validation: Perspectives in Hydrological Science. John Wiley & Sons, Ltd., New York, USA.
Mosley P.M., and McKerchar A.I., 1993. Streamflow. In: Maidment D.R. (Ed.), Handbook of Hydrology, McGraw-Hill, New York, pp. 8.1–8.39.
Ouyang, Y., 2012. A potential approach for low flow selection in water resource supply and management. J. Hydrol. 454–455, 56–63.
Paniconi, C., Troch, P., van Loon, E.E., Hilberts, A.G.J., 2003. Hillslope-storage Boussinesq model for subsurface flow and variable source areas along complex hillslopes: 2. Intercomparison with a 3D Richards equation model. Water Resour. Res. 39 (11), 1317. http://dx.doi.org/10.1029/2002WR001730.
Ponce, V.M., Lugo, A., 2001. Modeling looped ratings in Muskingum-Cunge routing. J. Hydrol. Eng., ASCE 6 (2), 119–124.
Pushpalatha, R., Perrin, C., Le Moine, N., Mathevet, T., Andréassian, V., 2011. A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. J. Hydrol. 411, 66–76.
Querner, E.P., Van Lanen, H.A.J., 2001. Impact assessment of drought mitigation measures in two adjacent Dutch basins using simulation modelling. J. Hydrol. 252 (1–4), 51–64.
Reilly, C.F., Kroll, C.N., 2003. Estimation of 7-day, 10-year low-streamflow statistics using baseflow correlation. Water Resour. Res. 39 (9), 1236. http://dx.doi.org/10.1029/2002WR001740.
Riggs, H.C., 1961. Regional Low Flow Frequency Analysis, U.S. Geological Survey Professional Paper 424-B, B-21-B-23, Washington, DC.
Riggs, H.C., 1980. Characteristics of low flow. J. Hydraul. Eng. 106 (5), 717–731.
Rupp, D.E., Selker, J.S., 2006. On the use of the Boussinesq equation for interpreting recession hydrographs from sloping aquifers. Water Resour. Res. 42, W12421. http://dx.doi.org/10.1029/2006WR005080.
Salas, J.D., Fu, C., Cancelliere, A., Dustin, D., Bode, D., Pineda, A., Vincent, E., 2005. Characterizing the severity and risk of drought in the Poudre River, Colorado. ASCE J. Water Resour. Plan. Manage. 131 (5), 383–393.
Samuel, J., Coulibaly, P., Metcalfe, R., 2011. Identification of rainfall-runoff model for improved baseflow estimation in ungauged basins. Hydrol. Process.. http://dx.doi.org/10.1002/hyp.8133.
Schaake, J., Cong, S., Duan, Q., 2006. U.S. MOPEX data set, IAHS Publication Series, Lawrence Livermore National Laboratory, UCRL-JRNL-221228.
Skøien, J.O., Merz, R., Blöschl, G., 2006. Top-kriging – geostatistics on stream networks. Hydrol. Earth Syst. Sci. 10 (2), 277–287. http://dx.doi.org/10.5194/hess-10-277.
Smakhtin, V., 2001. Low flow hydrology: a review. J. Hydrol. 240 (3–4), 147–186.
Son, K., Sivapalan, M., 2007. Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data. Water Resour. Res. 43, W01415. http://dx.doi.org/10.1029/2006WR005032, 2007.
Sorooshian, S., Gupta, V.K., 1995. Model Calibration. In: Sing, V.P. (Ed.), Computer Models of Watershed Hydrology. Water Resources Publications, Highlands Ranch, Colorado, USA, pp. 23–68.
Staudinger, M., Stahl, K., Seibert, J., Clark, M.P., Tallaksen, L.M., 2011. Comparison of hydrological model structures based on recession and low flow simulations. Hydrol. Earth Syst. Sci. 15, 3447–3459. http://dx.doi.org/10.5194/hess-15-3447-2011.
Stedinger, J.R., Thomas Jr, W.O., 1985. Low-flow frequency estimation using base-flow measurements. U. S. Geol. Surv. Open File Rep., 85–95.
Troch, P., van Loon, A., Hilbert, A., 2002. Analytical solutions to a hillslope-storage kinematic wave equation for subsurface flow. Adv. Water Resour. 25 (6), 637–649.
Troch, P., Paniconi, C., van Loon, E.E., 2003. Hillslope-storage boussinesq model for subsurface flow and variable source areas along complex hillslopes: 1. Formulation and characteristic response. Water Resour. Res. 39 (11), 1316. http://dx.doi.org/10.1029/2002WR001728.
Tucker, G.E., Bras, R.L., 1998. Hillslope processes, drainage density, and landscape morphology. Water Resour. Res. 34 (10), 2751–2764, W021474.
Vogel, R.M., Fennessey, N.M., 1995. Flow duration curves II: a review of applications in water resources planning. Water Resour. Bull. 31 (6), 1029–1039.
Vogel, R.M., Kroll, C.N., 1996. Estimation of baseflow recession constants. Water Resour. Manage. 10, 303–320.
Wagener, T., Wheater, W.S., Gupta, H.V., 2004. Rainfall-runoff Modeling in Gauged and Ungauged Catchments. Imperial College Press, London, p. 109.
Wagener, T., van Werkhoven, K., Reed, P., Tang, Y., 2009. Multi-objective sensitivity analysis to understand the information content in streamflow observations for distributed watershed modeling. Water Resour. Res. 45, W02501. http://dx.doi.org/10.1029/2008WR007347.
Zhang, Z., Kroll, C.N., 2007. A closer look at baseflow correlation. J. Hydrol. Eng., ASCE 12 (2), 190–196.