Available online at www.sciencedirect.com
ScienceDirect
Advances in Water Resources 30 (2007) 1756–1774
Advances in Water Resources
www.elsevier.com/locate/advwatres

Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins

Maitreya Yadav [a,1], Thorsten Wagener [a,*], Hoshin Gupta [b]

[a] Department of Civil and Environmental Engineering, The Pennsylvania State University, Sackett Building, University Park, PA 16802, USA
[b] Department of Hydrology and Water Resources, The University of Arizona, Harshbarger Building, Tucson, AZ 85721, USA

Received 22 August 2006; received in revised form 21 January 2007; accepted 28 January 2007
Available online 11 February 2007

Abstract

Approaches to modeling the continuous hydrologic response of ungauged basins use observable physical characteristics of watersheds to either directly infer values for the parameters of hydrologic models, or to establish regression relationships between watershed structure and model parameters. Both these approaches still have widely discussed limitations, including impacts of model structural uncertainty. In this paper we introduce an alternative, model independent, approach to streamflow prediction in ungauged basins based on empirical evidence of relationships between watershed structure, climate and watershed response behavior. Instead of directly estimating values for model parameters, different hydrologic response behaviors of the watershed, quantified through model independent streamflow indices, are estimated and subsequently regionalized in an uncertainty framework. This results in expected ranges of streamflow indices in ungauged watersheds. A pilot study using 30 UK watersheds shows how this regionalized information can be used to constrain ensemble predictions of any model at ungauged sites. Dominant controlling characteristics were found to be climate (wetness index), watershed topography (slope), and hydrogeology. Main streamflow indices were high pulse count, runoff ratio, and the slope of the flow duration curve. This new approach provided sharp and reliable predictions of continuous streamflow at the ungauged sites tested.

Keywords: Rainfall-runoff modeling; Predictions in ungauged basins; Regionalization; Streamflow indices; Catchment classification; Uncertainty

1. Introduction

Rainfall-runoff models are standard tools for hydrologic analysis. These models are used for applications such as water resources studies and flood forecasting, or in support of ecological studies. Available watershed models range from parsimonious-lumped to complex distributed physically based representations [38,47]. A problem common to all such models is that they all require some degree of parameter calibration to achieve reliable predictions (e.g. [5,40]), in which process the model parameters are adjusted (manually or automatically) until the observed and simulated watershed responses match as closely as possible (e.g. [14,19,35,43,44,37]). Even physically based models usually require some degree of calibration since it is difficult to estimate values for all of the parameters through field measurements. This occurs because the scale of measurement is usually smaller than the effective scale at which the model parameters are applied [5].

Problems in hydrologic modeling are accentuated further when it comes to prediction in ungauged or altered (e.g. land use) basins, where sufficiently long streamflow time series for parameter estimation via calibration are typically not available. Two common approaches to overcome this problem in ungauged situations are

(a) use of physically based models, and
(b) regionalization of model parameters using physical characteristics of watersheds.

The hope for physically based models is that that their parameters can somehow be strongly related to observable

* Corresponding author. Tel.: +1 8148655673; fax: +1 8148637304.
E-mail address: thorsten@engr.psu.edu (T. Wagener).
[1] Now at Parsons Brinkerhoff, Baltimore, USA.

0309-1708/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2007.01.005

physical properties of the watershed. However, differences in scale, over-parameterization and model structural error, have so-far prevented this objective from being achieved and some degree of calibration is usually still required [4,5]. Such models are used extensively wherever a high level of spatial detail is required from the prediction, as for example in the prediction of inundation areas. When simpler predictions such as streamflow response are required, less complex conceptual lumped models have been shown to be equally reliable and are often preferred.

To apply conceptual-type models to ungauged watersheds, a regionalization approach is commonly used [46,20,28,30,1,31,15,23,43,44,22,41]. In a regionalization approach, a parsimonious hydrologic model structure is selected, and calibrated to observable watershed responses for a large number of gauged watersheds. Regression equations are then developed in an attempt to explain the calibrated values of the model parameters from observable physical watershed characteristics. While some of the model parameters may be found to exhibit strong correlations with physical watershed characteristics, it is common that little or no significant correlation is found for many of the parameters [43,44]. Wagener and Wheater [41] also point out that this approach typically suffers from model identification difficulties, model structure errors, and difficulties in finding an appropriate calibration strategy that appropriately preserves the physical meaning of the model parameters. At the same time there is general agreement that model uncertainty is inherent and unavoidable, and therefore the best practice would be to employ an ensemble of feasible models to provide a corresponding ensemble of predictions for informed decision making [22]. Unresolved, therefore, is the question of what might be a way forward to achieve reliable hydrologic ensemble predictions in ungauged and altered basins?

The approach introduced in this paper provides ensemble predictions in ungauged basins by regionalizing model independent dynamic hydrologic response characteristics (or indices) to physical characteristics of watersheds in an uncertainty framework as constraints on ensemble predictions. Our objective is to achieve a progressive reduction in predictive uncertainty by constraining the expected watershed behavior at ungauged locations, while maintaining reliable predictions, leading to an increased understanding of the relationship between watershed structure and watershed response behavior. The next section reviews the past use of indices of streamflow behavior. A case study using 30 watersheds within the UK is utilized to develop and demonstrate the new approach.

2. Indices of dynamic response behavior

Dynamic response characteristics (hydrologic response behavior indices or signatures) of a watershed can be derived from output or input–output time series measured within the watershed, including precipitation, evapotranspiration (or temperature) and streamflow (or other response variables) time series. Such response characteristics are often indicative for a specific watershed and how its response differs from others; examples include common descriptors of hydrograph shape such as runoff ratio and time to peak flow [33].

A wide range of biologically relevant streamflow indicators of this type have been widely used by the ecological community for the evaluation of flow regimes (e.g. [27,29,9,16,18]). Olden and Poff [27], for example, compiled 171 (mainly statistical) hydrologic indices from a number of published papers and examined their suitability for describing various aspects of streamflow. They used 36 years of daily streamflow data from 420 watersheds throughout the US to carry out their case study, in which the indices were categorized into five classes based on their function and form. Principal Component Analysis (PCA) was carried out to examine redundancy in hydrologic indices. They created a statistical framework to select the relevant hydrologic indices that explain most of the variability in streamflow regimes. Olden and Poff [27] concluded that their approach can be helpful for hydro-ecological studies by selecting high-information and non-redundant hydrologic indices for a particular region in which a stream of interest is located. A similar study by Harris et al. [17] investigated four large watersheds in the UK to classify their streamflow and temperature regimes by analyzing the timing and magnitude of flow and temperature variation. They classified the annual temperature regime into three shape classes (early, intermediate and late minimum) and four temperature classes (cold winter and cool summer; moderate winter and cool summer; moderate winter and hot summer; and warm winter and hot summer). Similarly, the annual streamflow regimes were classified into five shape categories and four magnitude categories. An overall classification of streams was then carried out by combining the shape and magnitude classes. Their study followed the work by Hannah et al. [16] for classifying diurnal discharge hydrographs from glacier basins. Hydrologic indices have also been used to study seasonal changes in evapotranspiration [12]. In their work the authors' studied three streamflow indices (long-term averaged precipitation minus runoff; streamflow recession time constant; and diurnal streamflow amplitude) as indicators of spring onset and leaf emergence for fairly small watersheds (area <200 km^2) located on the US east coast. They also tracked the springtime changes in the three indices along the east coast of the US. More importantly, they found that the second and third indices depended on two physical characteristics of watersheds, namely, riparian area and hydraulic conductivity.

A recent study by Chinnayakanahalli et al. [8] examined possible links between various hydrologic indices from Olden and Poff [27] and physical characteristics of watersheds to predict hydrologic flow regimes for biological assessment in ungauged basins. They used data from 491 watersheds located in the western US to develop various linear regression models and estimate hydrologic indices

from physical characteristics of watershed. They found that some of the indices of flow magnitude (e.g. mean discharge, seven day minimum and maximum flows) could be estimated from physical characteristics (with R^2 values > 0.4). In a similar study, Spate et al. [36] used a rule based approach to regionalize a hydrologic index (runoff ratio = long-term annual runoff/long-term annual precipitation) to eight watershed characteristics by developing a set of hypothesized rules. Their work, being in a very preliminary phase did not yet provide significant conclusions, but concentrated on the robustness of the rule based algorithm. In another study by Berger and Entekhabi [3], two hydrologic indices of physiographic and climatic variability (runoff ratio and evaporation efficiency) were related to six physical watershed characteristics to explain basin to basin differences in modeled hydrologic response for 10 watersheds throughout the US. They found that the variability in modeled hydrologic response could be attributed to terrain and climatic variables within their dataset and that this information can be extracted from topographic, soil texture and rainstorm statistics data.

Dynamic response characteristics of watersheds have more recently been introduced in the context of hydrologic model calibration (e.g. [51,25,32,33]). Shamir et al. [32] derived two hydrologic indices based on the concept of peak density developed by Morin et al. [25], and investigated their usefulness in improving the identification of hydrologic model parameters. Shamir et al. [32] found that these indices described some important channel and hill slope routing processes and that using their rising limb density index improved the reliability of model predictions. In a subsequent study, Shamir et al. [33] investigated the usability of additional hydrologic indices at different time scales to constrain ensemble model predictions from parameters obtained using these indices. They developed a scheme to estimate model parameters based on extracting hydrologic indices at different timescales and compared the resulting parameter space with four other existing parameter estimation schemes. They reported that their scheme showed better results. Furthermore, they concluded that there is a need to study relationships between various hydrologic indices and physical characteristics of watersheds.

Detenbeck et al. [13] found that differences in seasonal flow regimes – identified using different streamflow indices – could be related to differences in hydrogeomorphology, watershed storage and land use. They suggest that the relationship between land use and flow response found might provide a way forward to investigate the impacts of land use change, even in ungauged watersheds.

3. Example study of constraint regionalization

In this section we relate dynamic response characteristics to observable physical watershed characteristics by means of regressive relationships in a new approach to predictions in ungauged basins (Fig. 1). The idea of regionalizing such indices stems from the observation that the amount of uncertainty involved in regionalizing hydrologic model parameters can be large, particularly because it is difficult to account for the effects of model structural error during model calibration and because the calibration problem is often ill-posed [41]. Watershed response characteristics on the other hand are not model-specific. Therefore uncertainties and confounding influences that might arise from the process of model identification are eliminated (or at least significantly reduced). Once regionalized, the behavioral information summarized by the response characteristics can be used as constraints on the model predictions, by facilitating (for example) a separation into behavioral and non-behavioral model sets using a binary classification approach.

3.1. Watershed climate and hydrology data

A set of 30 small to medium sized watersheds (~50–1100 km^2) located in the UK was used for the present study (Fig. 2). The watersheds are located throughout England and Wales and cover a wide range of soil types, topography and land use. Soil types range from predominantly clay at Mimram@Panshanger Park, over mostly alluvium at Fal@Tregony, to the permeable chalk watershed Test@Broadlands, and the very low permeable Taw@Umberleigh. Almost all watersheds are rural with very little urbanization except for the Isebrook@Harrowdown Old Mill, which is considered moderately urbanized. There is a good

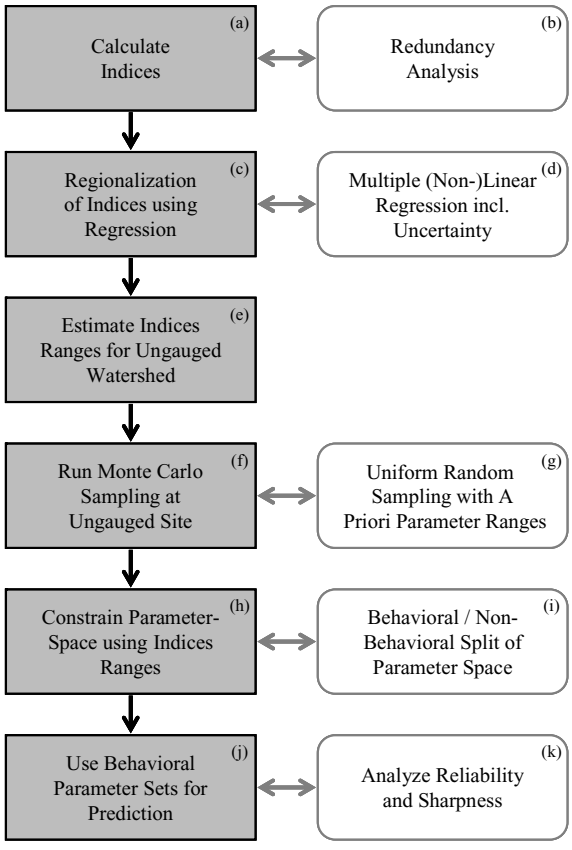
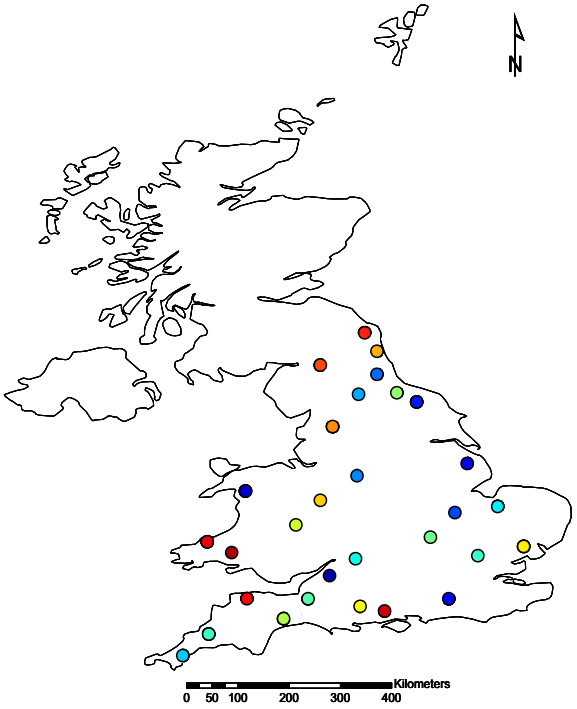


Fig. 1. Flowchart of procedure steps. These steps are explained in detail in subsequent subsections.



	Station Number	River	Gauging Station
	53017	Boyd	Bitton
	67018	Dee	New Inn
	29003	Lud	Louth
	39029	Tillingbourne	Shalford
	27042	Dove	Kirkby Mills
	31010	Chater	Fosters Bridge
	24004	Bedburn Beck	Bedburn
	28046	Dove	Isaac Walton
	25006	Greta	Rutherford Bridge
	48003	Fal	Tregony
	33029	Stringside	Whitebridge
	39020	Coln	Bibury
	47008	Thrushel	Tinhay
	38003	Mimram	Panshanger Park
	52010	Brue	Lovington
	32004	Ise Brook	Harrowdown Mill
	25005	Leven	Leven Bridge
	45005	Otter	Dotton
	55014	Lugg	Byton
	43006	Nadder	Wilton
	37005	Colne	Lexden
	54016	Roden	Rodington
	22006	Blyth	Hartford Bridge
	27035	Aire	Kildwick Bridge
	23006	South Tyne	Featherstone
	22001	Coquet	Morwick
	50001	Taw	Umberleigh
	62001	Teifi	Glan Teifi
	42004	Test	Broadlands
	60010	Tywi	Nantgaredig

Fig. 2. Map of UK showing the location of the 30 watersheds used in this study. Table shows the station numbers, name of rivers and locations of their gauging stations. Watersheds are ranked from smallest (top and blue) to largest (bottom and red) drainage area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mix of topographic gradient within the watersheds with Tywi@Nantgaredig having an altitude change of about 795 m and the Stringside@Whitebridge having a change in altitude of only 75 m.

None of the watershed streamflow used here is significantly affected by abstractions or other alterations, though a few watersheds have naturalized flow. The rainfall-runoff data for the rest are only affected by factors which are thought to be of minor impact and have therefore been ignored in this study. The quality of precipitation and runoff data was the criterion deemed most important in the selection of the 30 watersheds. A second criterion was to achieve a reasonable distribution of three watershed characteristics that are considered most relevant to pool hydrologically similar watersheds within the UK [6,22]: watershed area (AREA), a base-flow index derived from the Hydrology of Soil Types classification (BFIHOST) and standard annual average precipitation for the period of 1961–1990 (SAAR). Potential evapotranspiration was calculated from temperature data using Hargreaves equation [34]. Eleven consecutive years (1980–1990) of data were available for 29 watersheds. Time series data for the Test@Broadlands was available from 1983 to 1996. The period of time series used for analysis was from 1-1-1983 to 12-31-1990.

The long-term average monthly values of precipitation, streamflow and potential evapotranspiration are plotted in Fig. 3a–c, showing the similarity in regime between the watersheds. A flow duration curve showing the cumulative frequency of the normalized flow values is also shown (Fig. 3d). The flows are normalized by their mean flow values to facilitate comparison. A steep slope in the flow duration curve indicates flashiness of the streamflow response to precipitation inputs whereas a flatter curve indicates a relatively damped response and higher storage. Fig. 3d shows the diversity in watersheds with respect to their hydrologic response. Fig. 3e shows the climate gradient within the data set, and already indicates that a strong relationship exists between runoff ratio (R/P) versus the ratio of precipitation and potential evapotranspiration (P/EP).

3.2. Watershed physical characteristics

A comprehensive list of physical characteristics for all the watersheds was compiled from the National River Flow Archive (http://www.nwl.ac.uk/ih/nrfa) and the data CD of the Flood Estimation Handbook (FEH). The main features of these watersheds are presented in a parallel coordinates plot in Fig. 4 (each line represents one watershed). The plot shows that most of the watersheds tend to have small area, small woodland and built up areas, and large RMED-1H values (see Table 1).

Other available watershed characteristics include station elevation, hydrogeology, climate and land use. The elevation data is derived from 50 m grid cells having 0.1 m vertical resolution. Hydrogeology information consists of data

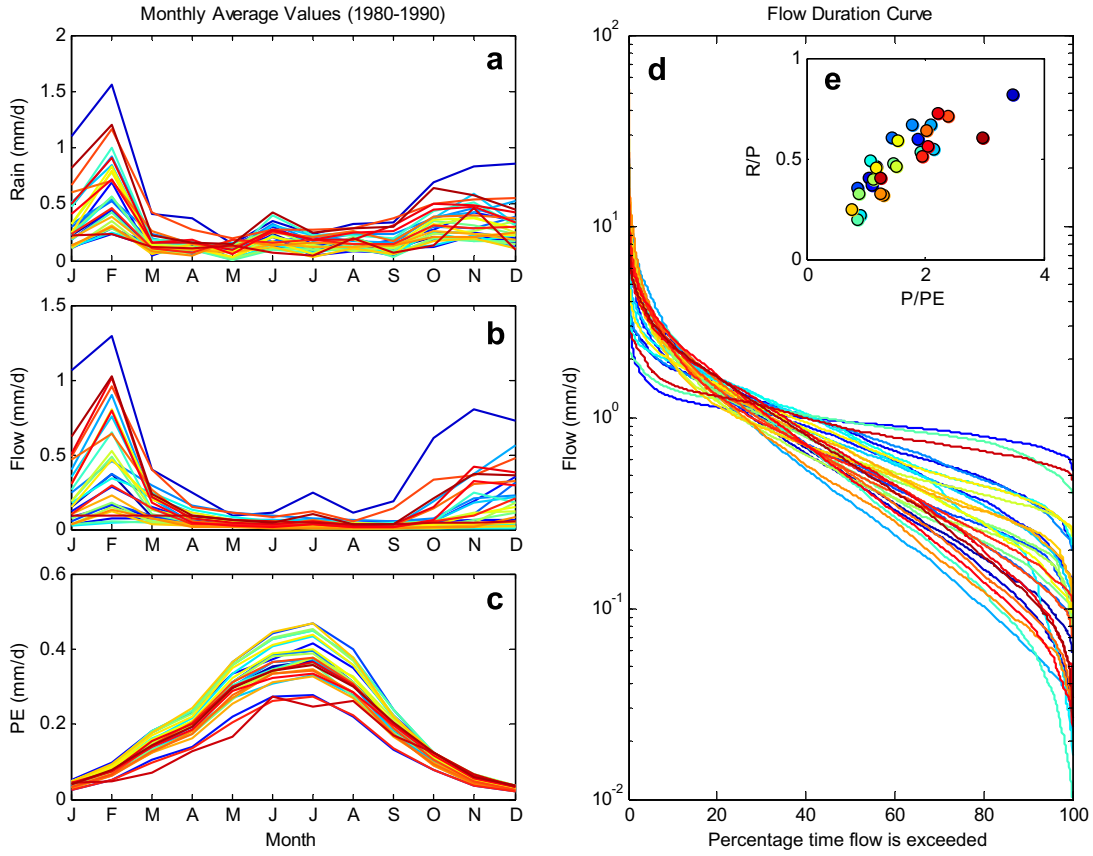


Fig. 3. (a–c) Average monthly values of precipitation, streamflow and potential evapotranspiration for the 30 UK watersheds. (d) Normalized flow duration curve. The flows are normalized by the mean flows. (e) Runoff/precipitation versus precipitation/potential evapotranspiration. The color-coding relates to watershed size as indicated in Fig. 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

representing permeability of watersheds and of another dataset for superficial deposits or drifts. This data is important because it is known to influence streamflow regimes (http://www.nwl.ac.uk/ih/nrfa). Land use has 27 land cover regions which have been broadly clustered into seven categories. Climatic characteristics include the variability in precipitation (PVAR) and the wetness index (P/PE), which is the ratio of long-term averaged precipitation to long-term averaged potential evapotranspiration.

Base-flow index (BFIHOST) in this study, is considered a physical watershed characteristic rather than a hydrologic index. Throughout the UK, it can be estimated from

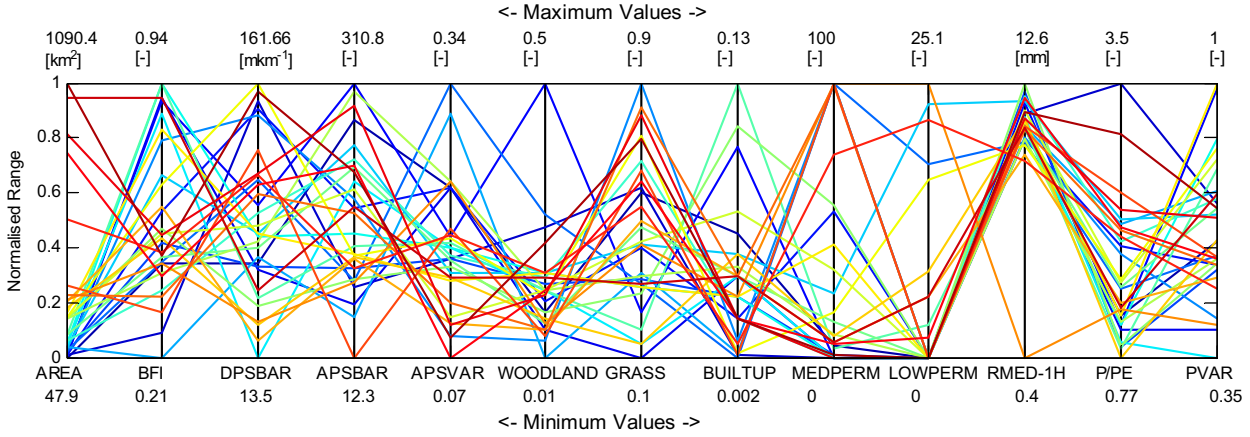


Fig. 4. Physical characteristics of watersheds shown in parallel coordinates plot. The color-coding relates to watershed size as indicated in Fig. 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Description of watershed characteristics

Characteristic	Unit	Description
AREA	km^2	Watershed drainage area
BFIHOST	[–]	Base-flow index derived using HOST classification
LDP	km	Longest drainage path
DPSBAR	$m\ km^{-1}$	Index of watershed steepness
DPLBAR	km	Index describing watershed size and drainage path configuration
APSBAR	[–]	Index representing the dominant aspect of watershed slopes
APSVAR	[–]	Index representing the invariability of aspect of watershed slopes
URBEXT	[–]	FEH index of fractional urban extent for 1990
LEVEL ST	mod[a]	Elevation of gauging station above ordinance datum
MAX ALT	mod	Elevation of point with maximum altitude above ordinance datum
ALT95	mod	95 Percentile elevation (95% points in watershed above this elevation)
ALT5	mod	5 Percentile elevation (5% points in watershed above this elevation)
WT. ALT	mod	Weighted averaged altitude of watershed
WOODLAND	[–]	Percentage woodland within watershed
ARABLE	[–]	Percentage arable land within watershed
GRASS	[–]	Percentage grasslands within watershed
MOUNTAIN	[–]	Percentage mountains within watershed
BUILTUP	[–]	Percentage built-up land within watershed
HIGH PERM	[–]	Percentage soil within watershed with high permeability
MED PERM	[–]	Percentage soil within watershed with medium/mixed permeability
LOW PERM	[–]	Percentage soil within watershed with low permeability
SAAR	mm	1961–1990 Standard period average annual rainfall
RMED-1D	mm	Median annual maximum 1-day precipitation
RMED-2D	mm	Median annual maximum 2-day precipitation
RMED-1H	mm	Median annual maximum 1-hour precipitation
P/PE	[–]	Ratio of average annual precipitation and average annual evapotranspiration
PVAR	[–]	Coefficient of variation in precipitation

[a] mod: meters above ordinance datum.

an available regression equation where BFIHOST is the dependent variable and the HOST classifications are the independent variables [6].

These physical characteristics of watersheds were divided into six categories depending on their function, thus extending the work by Wagener et al. [43,44]. This classification has been used here as well and verified by analyzing the correlation between the watershed characteristics. It was found that almost all of the characteristics within each category were highly correlated. This result was used to reduce the number of physical watershed characteristics for this study, e.g. among the landform descriptors (AREA, LDP and DPLBAR), only AREA was considered for this study and the others were rejected since they were unlikely to contain any additional information. A list of all characteristics and their grouping is shown in Table 2. Only the underlined physical characteristics were used for this study.

Not all characteristics in a group were found to be correlated. The ones showing significant correlation (linear correlation > 0.75) are shown by numbered brackets preceding them. For example, BFIHOST, MEDPERM and LOWPERM were under the same group, *Soils*, but were not strongly correlated. A final choice of characteristics was made by selecting one physical characteristic from each correlated group, along with the ones that were not strongly correlated with any other physical characteristic. Additionally, ARABLE and URBEXT were the only characteristics that show negative correlation. Among the group *Others*, MOUNTAINS was not kept because it was significantly correlated (correlation coefficient $= 0.7$) with the physical characteristics in the group *Topography*.

Table 2
Classification of physical characteristics of watersheds

Function	Physical characteristics of watershed
Landform	[1]LDP, [1]AREA, [1]DPLBAR
Topography	[2]DPSBAR, [2]LEVEL OF STATION, [2]MAX ALT, [2]ALT95, [2]ALT5, [2]WTAVALT
Land use	[3]BUILTUP, [2]ARABLE, [3]URBEXT
Soil	[4]BFIHOST, [4]HIGHPERM, MEDPERM, LOWPERM
Climate	[5]SAAR, [5]RMED-1D, [5]RMED-2D, RMED-1H, [5]P/PE, PVAR
Others	MOUNTAINS, GRASS, APSVAR, APSBAR, WOODLAND

Groupings are based on linear correlation coefficient (>0.75).
Underlined response characteristics were kept for regression analysis and designated by the letters within brackets.

3.3. Methodology

The earlier introduced Fig. 1 shows the stepwise procedure of the approach implemented in the subsequent sections. The procedure starts by calculating the response characteristics (hydrologic indices) for the 30 watersheds.

3.3.1. Dynamic response characteristics (Fig. 4a)

Thirty-nine dynamic response characteristics were derived from streamflow, precipitation and evapotranspiration time series at various timescales. To analyze the different aspects of streamflow, the response characteristics were divided into seven categories (extending the work by Olden and Poff [27]: magnitude of high flows, magnitude of low flows, magnitude of average flows, duration of flows, frequency, rate of change in flows, and timing of flow events. A list of these indices is shown in Table 3 (see also [50,2,11]).

3.3.2. Redundancy in response characteristics (Fig. 4b)

Similar to the manner in which the number of physical characteristics was pruned, the response characteristics must also be checked for redundancy. In the present study the number of response characteristics was 39 and an assessment of relationships via correlation coefficients was considered appropriate 43,44,3]. Linear and Spearman rank correlation coefficients were used to find patterns of variability among the response characteristics. The grouping of response characteristics is shown in Table 4. Square brackets indicate characteristics with high linear correlation (>0.75) and curly brackets indicate characteristics with high Spearman rank correlation (>0.75 and linear correla-

Table 3
Description and grouping of watershed response indices (D – Daily, W – Weekly, M – Monthly, A – Annual/yearly)

S. No.	Name	Units	Scale	Description	Ref.
Magnitude of high flow events					
MH2	Maximum February flow	mm	M	Maximum monthly flow for February across all years	Shamir et al. [33]
MH3	Maximum August flow	mm	M	Maximum monthly flow for August across all years	Shamir et al. [33]
MH4	Maximum November flow	mm	M	Maximum monthly flow for November across all years	Shamir et al. [33]
MH5	Maximum negative flow change	mm	M	Maximum difference in magnitude of flows in a rising limb	Shamir et al. [33]
MH6	Maximum positive flow change	mm	M	Maximum difference in magnitude of flows in a falling limb	Shamir et al. [33]
MH7	High flow discharge (1st percentile)	[–]	A	Mean of 1st percentile from flow duration curve divided by median Flow	Clausen et al. [10]
MH8	High flow discharge (10th percentile)	[–]	A	Mean of 10th percentile from flow duration curve divided by median flow	Clausen et al. [10]
Magnitude of low flow events					
ML1	Recession coefficient (slope)	mm	D	Slope of the slow recession limb	
Magnitude of average flow events					
MA1	Mean of daily flows	mm	D	Mean daily flow	Clausen and Biggs [9]
MA2	Median of daily flows	mm	D	Median daily flow	Clausen and Biggs [9]
MA3	Skewness in daily flows	[–]	D	Mean daily flows divided by median daily flows	Clausen and Biggs [9]
MA4	Streamflow variability	[–]	D	Coefficient of variation in Streamflow	Clausen and Biggs [9]
MA6	Annual specific runoff	mm	A	Average annual runoff divided by watershed area	
MA7	Runoff ratio	[–]	A	Average annual runoff divided by average annual precipitation	
MA8	Inter annual range	mm	A	Wettest yearly flow minus driest yearly flow	Shamir et al. [33]
MA9	Time series total flow	mm	11 yrs	Sum of daily flows for the record period of flow	Shamir et al. [33]
Rate of change in flow events					
R2	Slope of flow duration curve (mid)	mm	D	Slope of part of curve between the 33% and 66% flow exceedance values of streamflow normalized by their means	–
Frequency of flow events					
F2	Declining limb density	week^{-1}	W	Number of peaks divided by cumulative duration of declining limbs	Shamir et al. [33]
F3	High pulse count (3 times median)	year^{-1}	A	Number of annual occurrences during which flow remains above 3 times median daily flow divided by period of flow	Clausen and Biggs [9]

Table 4
Clustering of response characteristics

Function	Response characteristics
MAGNITUDE (H)	MH1, [1]MH2, [1]MH3, [1]MH4, [1]MH5, [1]MH6, [2]MH7, [2]MH8, [2]MH9
MAGNITUDE (L)	[1]ML1, ML2, [1]ML3, ML4[3]
MAGNITUDE (A)	[1]MA1, [1]MA2, [2]MA3, [2]MA4, MA5, [1]MA6, [1]MA7, [1]MA8, [1]MA9
DURATION	[2]D1, -{1}D2, [2]D3, -[3]D4
RATE OF CHANGE	R1, [2]R2, [3]R3, -R4
FREQUENCY	F1, F2, [2]F3, [2]F4, {1}F5
TIMING	[4]T1, -[4]T2
CLIMATE	-[1]C1, C2

[–]: groupings based on linear correlation (>0.75). {–}: groupings based on Spearman rank correlation (>0.75) and linear correlation <0.4. Underlined response characteristics are kept for subsequent regression analysis. Negative correlation is represented by a hyphen preceding the response characteristic.

tion < 0.6). Negative correlation is represented by a hyphen preceding the response characteristic (one variable without and one with hyphen would have a negative relationship, two variables with hyphen would have a positive relationship, etc.). The 13 underlined response characteristics were retained for use in regionalization based upon their grouping, correlation and individual importance (Fig. 5).

3.3.3. Regionalization of response characteristics (Fig. 4c–e)

Based on the analysis described above, 28 watershed response characteristics and 13 physical watershed characteristics were retained for further analysis. The next step in our approach was to derive relationships between the indices and the physical characteristics. This could be achieved in a number of ways. The simplest of the methods is linear regression which has been used in many regionalization studies e.g. [30,43,44,3]. Other methods include nonlinear regression, multivariate analysis, rule based classification, cluster analysis, etc. (e.g. [27,36]). In this study, stepwise linear regression was used to establish relationships between multiple physical characteristics and each individual response characteristic.

To test the robustness of the regression models, a sixfold cross validation approach was utilized, where 30 watersheds were divided into 6 groups of 5 watersheds each. Each group was treated as ungauged, in turn, for the purpose of validation. The remaining 25 watersheds were treated as gauged to develop regression relationships. In this manner, the proposed approach was applied to all 30 watersheds (5 validation watersheds per group) with the results of regression developed from the remaining 25 watersheds per group. Regression was carried out for each of the 6 groups individually and the combinations of physical characteristics for predicting response characteristics were observed.

Linear regression equations were developed between individual response characteristics and physical characteristics, for the 25 watersheds in each group, based on an equation of the following form [21]:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_{p-1} x_{p-1} + \epsilon \quad (1)$$

where Y is the response characteristic of interest, $x_1, x_2, \ldots, x_{p-1}$ are $p-1$ physical characteristics, $\beta_0, \beta_1, \ldots, \beta_{p-1}$ are the p regression coefficients, and ε is an error term. Fig. 6 shows one of the regression relationships in watershed group 1 between P/PE and Maximum November flow (MH4) including estimates of uncertainty in the regression. The coefficient of determination (r^2) of this regression was 0.88. The regression includes the estimation of prediction (dark grey band) and confidence intervals (light grey band);

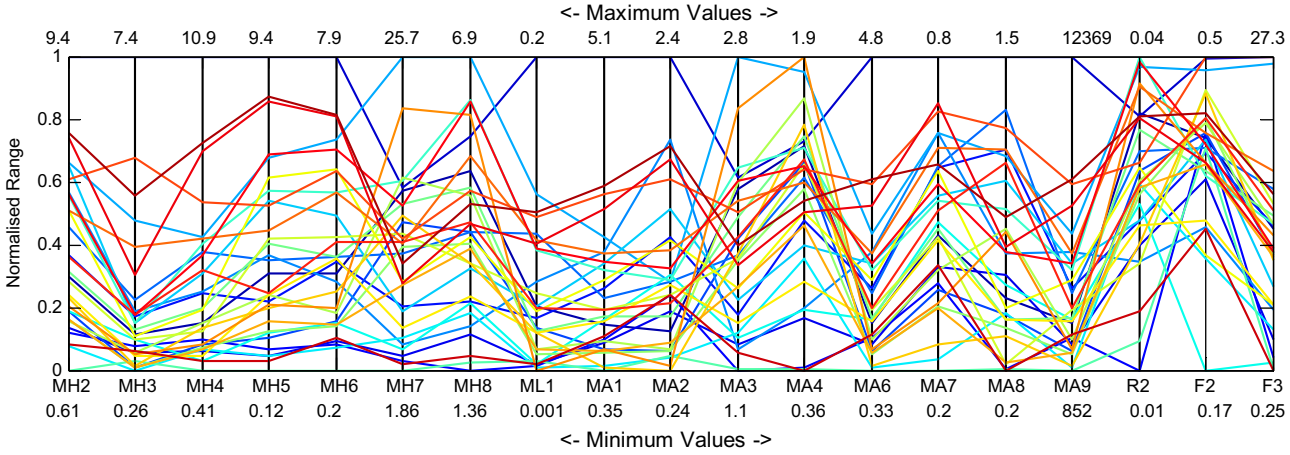


Fig. 5. Response characteristics of watersheds shown in parallel coordinates plot. Numbers on the upper axis show maximum values of response characteristics. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

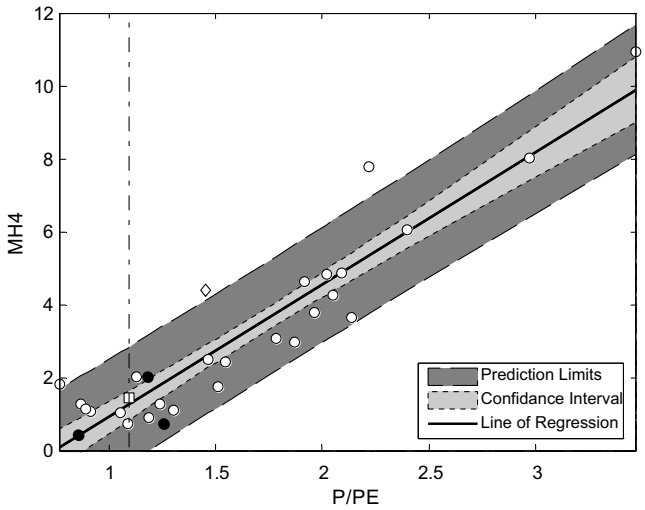


Fig. 6. Example of single index regression for Maximum November flow (MH4) and climatic gradient P/PE. The white circles show 25 watersheds used to develop regression relationship. *R*-squared statistic for this regression was 0.88. Three black circles, one white square and one white diamond show 5 watersheds left out for validation. One watershed (represented by a white diamond) lies outside the prediction limits while the ones represented by the three black circles and a white square lie inside. The result of MH4 as constraint on hydrologic model prediction for the watersheds represented by white square and diamond is analyzed in detail in the results and discussions section.

the confidence interval is a measure of the certainty (or uncertainty) of predicting the true (expected or mean) value of the variable while the prediction interval is a measure of the certainty of predicting some future (possible) value of that variable. Since the uncertainty in prediction intervals includes the uncertainty in the regression parameters $(\beta_0, \beta_1, \ldots, \beta_{p-1})$ and any new measurement (Y), this interval is wider than the confidence interval, which considers uncertainty in regression parameters only, while the measurements themselves are assumed to be random variables. In Fig. 6, the 25 watersheds indicated by white circles were used to develop a linear regression relationship and the five watersheds indicated by black circles, white square and diamond were not included in the derivation of the regression relationship. The known value of P/PE for the ungauged watershed (e.g. black vertical line in Fig. 6 for one of the ungauged watersheds) can then be used to calculate the confidence and prediction limits of MH4 from the regression equations for the validation watershed. It can be seen that three black circles and a white square lie within the prediction interval while one does not. This means that if MH4 is used as constraint on model prediction on the white diamond lying outside the prediction interval, the regionalized constraint might be too narrow. On the other hand, the watersheds represented by three black circles and a white square lying close to the line of regression are expected to yield better results. A detailed analysis of this issue is carried out in results and conclusions section.

The technique of forward stepwise regression automatically eliminates the variables that do not contribute significantly towards the prediction of the response characteristics. This was implemented using the *p*-values of individual physical characteristics for each regression relationship. The critical value of p was set to 0.05 as suggested by Kottegoda and Rosso [21]. If this value is greater than 0.05 for a certain physical characteristic, the characteristic is eliminated from the regression. For example, Fig. 6 shows the regression between response characteristic MH4 and physical watershed characteristic P/PE for group 1. This shows that in group 1, when stepwise regression was carried out on MH4 and 13 physical watershed characteristics, only P/PE showed a *p*-value less than 0.05. That is why this particular relationship became a simple univariate linear regression model. However, when stepwise regression results of MH4 in the other five groups were examined, BFIHOST and APSVAR also figured in the regression equation along with P/PE. Stepwise regression results for the remaining response characteristics also presented a similar situation, where different patterns of relationships were found across six groups for individual single response characteristics.

The results of the stepwise regressions are shown in Table 5. Here the number of times each physical characteristic appears in the stepwise regression is presented, with 6 being the maximum, since for each response characteristic, separate regression relationships are developed for 6 groups of watersheds for cross validation. The shades of gray also indicate the same; with black indicating that the physical characteristic appears all 6 times in the equation for a response characteristic. A blank cell indicates that the physical characteristic does not appear at all in the relationship for the corresponding response characteristic. The last column in the table shows the range of r^2 statistics for all six regression equations. It was noticed that some of the physical characteristics (e.g. BFIHOST and P/PE) figured prominently in most of the regression equations, while others (e.g. AREA and WOODLAND) featured only a few times. This suggests that BFIHOST and P/PE were the most important physical (climatic) characteristics that described most of the response characteristics.

The confidence limits and prediction limits were calculated for each response characteristic for 5 watersheds in each group treated as ungauged. The bracketed response characteristics in Table 5 were rejected for further analysis. It was noticed that some of the response characteristics did not show very good R^2 statistics (e.g. MH1, MH9, MA5, D4, R4, F1 and T1 with R^2 value less than 0.6) and these were subsequently removed for further analysis. It was also observed that confidence and prediction limits for some of the response characteristics included negative values. For example, although response characteristics D1 and F4 in Table 5 showed good R^2 statistic for the overall regression, the confidence and prediction limits for the validation watersheds came out to be negative and hence these response characteristics were also removed from further analysis. An exception

Table 5
Stepwise regression analysis details

	AREA	BFI HOST	DPS BAR	APS BAR	APS VAR	WOOD LAND	GRASS	BUILT UP	MED PERM	LOW PERM	RMED -1H	P/PE	PVAR	r^2 Range	F Statistic Range (10^{-10})
[MH1]		6		1	5									0.59–0.74	21171–564170
MH2		3					2					6		0.84–0.92	0.06–9.5
MH3				6	4					5		6		0.84–0.94	0.13–14.44
MH4		1			1							6		0.83–0.92	0.003–2.5
MH5		4		1			1		1			6	2	0.76–0.89	3.9–181.1
MH6	1	6						1				6	4	0.85–0.93	0.31–13.92
MH7	2	5	3		1				1		5	3		0.87–0.98	0.002–20.61
MH8		6									1			0.75–0.89	0.44 –212.8
[MH9]		6												0.55–0.7	355.38–224150
ML1		6		1	1	1				1	3	6		0.82–0.92	2.55–62.93
MA1	1		1							4		6		0.91–0.94	0.0002–0.011
MA2		6	6									6		0.87–0.94	0.01–23.5
MA3		6	5						1		5		1	0.88–0.95	0.015–9.42
MA4		6				1					5	6		0.90–0.94	0.02 – 0.53
[MA5]		5							1			3		0–0.57	991810–27041000
MA6			1							4		6		0.91–0.93	0.0001–0.024
MA7			6					5	2	1	1	6	1	0.82–0.93	1.47–546.93
MA8			1		6				2			6	5	0.82–0.90	1.65–476.78
MA9			1							4		6		0.91–0.93	0.0001–0.024
[D1]		6		1		2	4	1	1	1	1		2	0.70–0.91	19.44–10761
[D4]		6		1		2		1					4	0.54–0.65	97336–6243100
R2		6	1					2						0.46–0.76	6367.4–2025300
[R4]							5	5				4	1	0.28–0.69	56053–653226000
[F1]														0–0	0–0
F2		6					1	5				2		0.60–0.78	20831–386450
F3	3	6		5			2		2	2		5		0.89–0.98	0.001–1.79
[F4]		6			2					1			1	0.87–0.92	0.013–0.114
[T1]			5								5			0.19–0.76	166.18–286540000

Cell values in the table indicate number of times a physical characteristic appears in the regression for 6 bins or groups of watersheds. This number can be visually interpreted through the color of cells also with darker shades of grey indicating a higher value. The last two columns show the ranges of R^2 and F statistic.

to this rule is the response characteristic R2, which was retained for further application even though the lowest R^2 value, was below the chosen threshold. R2 was thought to be an important descriptor of the hydrologic response and this is corroborated in the results later.

3.3.4. Hydrologic model application and Monte Carlo analysis (Fig. 4f–j)

While the regionalized index ranges derived above provide constraints on the expected behavior of the ungauged watersheds, the approach does not by itself provide any predictive capability of the rainfall-runoff response. A hydrologic model is required to gain this capability. Note that *any* model structure for which the required input data are available can be implemented in the framework developed here. For illustration of the method, we have selected a simple and commonly used parsimonious model. The lumped hydrologic model (e.g. [7,42,39]) chosen for this study (Fig. 7) has 5 adjustable parameters, H_{UZ}, b, α, K_q and K_s (Table 6). It consists of a probability-distributed

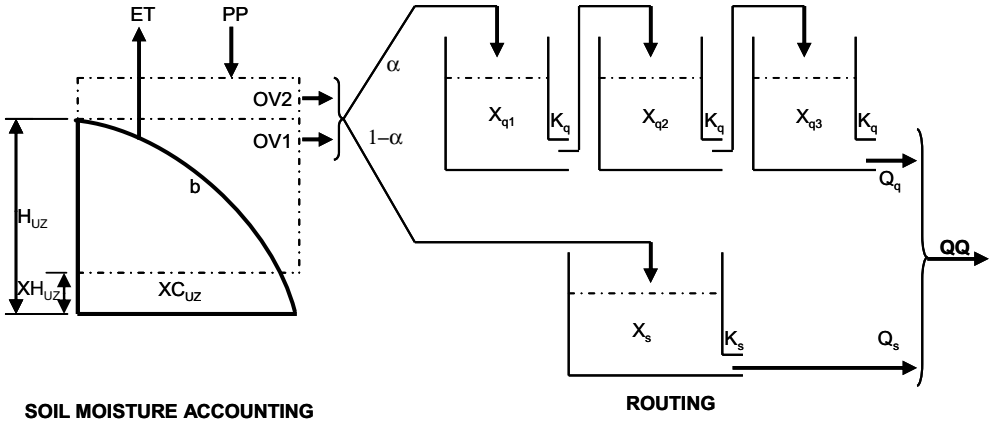


Fig. 7. Lumped 5-parameter model structure. ET and PP are potential evapotranspiration and precipitation respectively [mm]. OV1 and OV2 are model simulated effective precipitation components [mm]. X_i is the state of individual buckets of the routing model. QQ is model simulated streamflow [mm]. XH_{UZ} and XC_{UZ} are Soil moisture accounting tank state contents [mm].

store that describes the watershed storage as a Pareto distribution of buckets of varying depth as the soil moisture accounting component. Excess precipitation is produced through overflow of the buckets and routed through a parallel combination of a 3-reservoir Nash Cascade for quick flow and a single reservoir slow flow routing component. XH_{UZ} and XC_{UZ} are state variables describing the soil moisture accounting content. E is potential evapotranspiration, PP is precipitation, OV1 and OV2 are excess precipitation to the routing module generated from overflow of the soil moisture accounting component. See Moore [24] for a detailed description of the soil moisture accounting model. X_{q1}, X_{q2}, X_{q3} and X_s are the states of the individual reservoirs of the routing module. Q_q and Q_s are the flow values generated from the quick and slow reservoirs respectively.

The model was run within a Monte Carlo framework by randomly sampling 10,000 parameter sets drawn from a uniform distribution covering the predefined feasible parameter space. For plotting purposes, the Nash Sutcliffe Efficiency measure (NSE) [26, Eq. (3.12)] and the Root Mean Squared Error measure ($RMSE^{ast}$) (Eq. (3.13)) were calculated; the latter is computed using a Box-Cox transformation applied to the observed and simulated flows,

$$NSE = 1 - \frac{\sum_{i=1}^{N}(q_i - \hat{q}_i)^2}{\sum_{i=1}^{N}\left(q_i - \frac{1}{N}\sum_{i=1}^{N} q_i\right)^2} \quad (2)$$

$$RMSE^* = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(q_i^* - \hat{q}_i^*)^2} \quad (3)$$

$$q_i^* = (q_i^{\lambda} - 1)/\lambda \quad (4)$$

where q_i is mean daily streamflow (mm), $\hat{q}_i$ is the model simulated daily streamflow (mm), N is the length of time series of flow, q_i^* and $\hat{q}_i^*$ are Box-Cox transformed values of daily observed and simulated stream flows respectively (Eq. (4)), with a value of 0.3 for the transformation parameter λ, since this value provides a good fit to low flow periods in our experience. The two measures were selected because of their emphasis on fitting different parts of the time series of flow (i.e. peak flows and low flows respectively).

It is important to stress that these two model performance measures were *not* used for selecting parameter sets, since by definition a comparison to observed flows would not be possible at ungauged locations. Instead, the regionalized ranges of indices only (based on confidence and prediction limits) were used for this purpose. Simulations that produced indices that fall within the ranges were considered behavioral, while those that fell outside were considered non-behavioral. This step was first applied for each of the regression equations separately, and then using combinations of response characteristics. The maximum and minimum simulated flows generated by the behavioral parameter sets (i.e. lying within the confidence and prediction intervals) were determined for each time step, and ensemble streamflow ranges were plotted.

Table 6
Description of model parameters

Parameter	Description	Unit	Min	Max
H_{UZ}	Maximum storage capacity of watershed	mm	1	500
b	Index describing spatial soil moisture distribution	–	0.1	2
α	Flow distribution coefficient	–	0.1	0.99
K_q	Residence time of quick flow reservoir	s^{-1}	0.1	0.99
K_s	Residence time of slow flow reservoir	s^{-1}	0	0.1

3.3.5. Analysis of reliability and sharpness of ensemble predictions (Fig. 4k)

Evaluation of the performance of the approach was done in terms of Reliability and Sharpness (precision) of the predictive ranges obtained from the behavioral simulations. *Reliability* is defined here as a measure of the fraction of time the observed streamflow is within the prediction band of the model. Reliability values are calculated by counting the number of times the observed streamflow falls within the prediction band, divided by the length of the time series. How far the observations fall outside the prediction bands, is not currently considered. *Sharpness* is a measure of the ensemble spread. A single line would be a sharpness of 100%, while no reduction from the original ensemble range produced by the a priori feasible parameter range represents a sharpness of 0%. A small value of sharpness represents a small reduction in uncertainty (i.e. a wide ensemble spread). Both measures are expressed in percent and an optimal value for both, reliability and sharpness, is at 100%.

To investigate the effect of response characteristics on different parts of the hydrograph, the streamflow time series was divided into 10 levels based on flow exceedance percentiles of the flow duration curve. Reliability and sharpness values of these 10 different levels for individual response characteristics for each ungauged watershed were noted. To visualize the large amount of data available for analysis average values of reliability and sharpness for all the watersheds were also noted as discussed below.

4. Results and discussion

The method described above was tested using each watershed in turn as a 'validation' watershed, as stated before. For reasons of brevity, the results from the first group are presented in detail followed by the overall result for single response characteristics. A detailed discussion on the combined effect of multiple response characteristics is presented later. The results obtained using the regression equation (MH4 [Maximum November Flow] versus P/PE; example of regression demonstrated in previous sections; see also Fig. 6), for the watershed (shown by white square marker in Fig. 6) from group 1 are shown in Fig. 8a–c. The observed value of the response characteristic

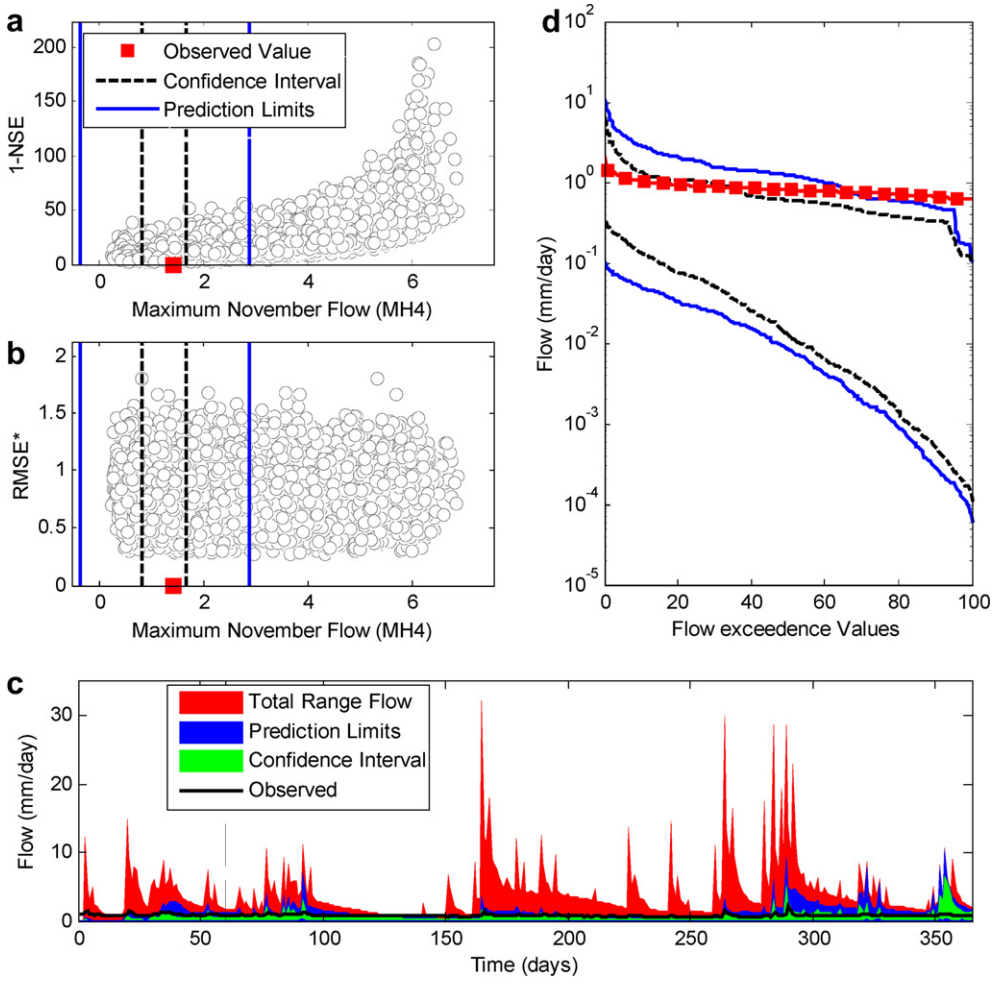


Fig. 8. Dotty plots of simulated MH4 versus (a) 1-NSE and (b) RMSE*. (c) Observed and (constraint) predicted streamflow ranges. (d) Flow duration curves for observed and (constraint) predicted streamflow ranges for the watershed W11 represented by white square marker in Fig. 6. This watershed lies inside the prediction limits of regression equation. A warm-up period of 60 days is shown by the dashed line.

MH4 for this watershed (denoted by W11; the first validation watershed in the first cross validation group) lies inside the prediction limits of MH4 obtained from the regression of MH4 with physical characteristic P/PE for 25 watersheds shown by 25 white circles in Fig. 6. The confidence and prediction intervals derived from the regression analysis have clearly constrained the parameter space in terms of the performance evaluation criterion used. Fig. 8c shows the maximum and minimum simulated flows for these intervals and for the complete range of simulations. Fig. 8d also shows the flow duration curves for the predictive ranges of flow. The 60 day period before the dashed vertical line was used as a model warm up period. The observed streamflow is seen to lie almost fully inside the prediction intervals after the warm up period. The number of behavioral simulations when flow is constrained by prediction limits of MH4 ratio was 6002 (60%), and the corresponding number for flow constrained by confidence limits of the MH4 was 2022 (20%).

Fig. 9a–c present similar plots for the watershed that is indicated by a white diamond marker in Fig. 6. The observed value of the response characteristic MH4 for this watershed (denoted by W15) lies outside the prediction limits of MH4 obtained from the regression of MH4 with physical characteristic P/PE for 25 watersheds shown by 25 white circles in Fig. 6. The observed streamflow is seen to lie almost fully within the prediction intervals after the warm up period for this case too. The number of behavioral simulations, when flow is constrained by prediction limits of the runoff ratio, was 8639 (86%), and the corresponding number for flow constrained by confidence limits of the MH4 was 1633 (16%). Although, the number of behavioral simulations in both the cases is similar (confidence intervals), the predictive uncertainties in simulated flow for the latter case are higher than the former case. This is evident from a visual inspection of difference in magnitudes of maximum simulated flows and observed flows from Figs. 8c and 9c.

It was seen that upper flow limits are higher for W11 than W15 during high flow periods. For low flow periods, W11 shows more constrained flows in terms of the upper limit. It can be seen that for W11, reliability values are high

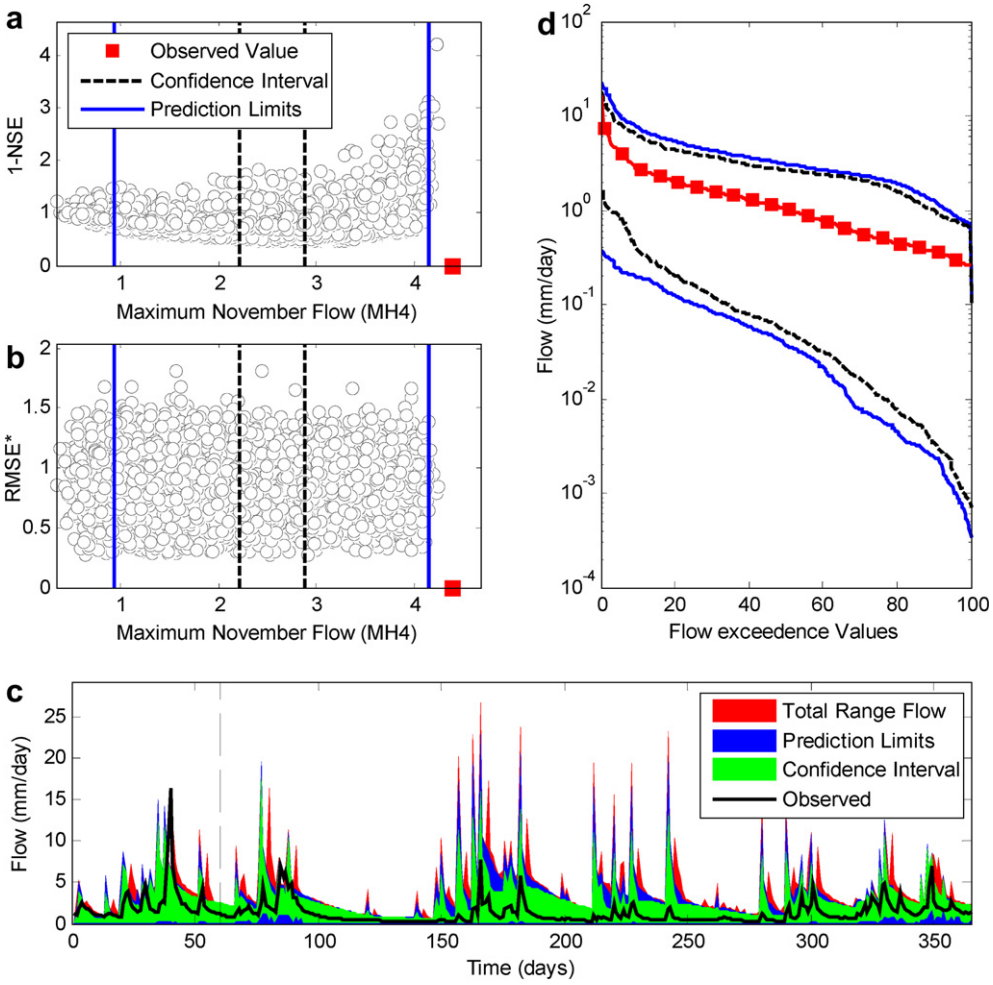


Fig. 9. Dotty plots of simulated MH4 versus (a) 1-NSE and (b) RMSE*. (c) Observed and (constraint) predicted streamflow ranges. (d) Flow duration curves for observed and (constraint) predicted streamflow ranges for the watershed W15 represented by white diamond marker in Fig. 6. This watershed lies outside the prediction limits of regression equation. A warm-up period of 60 days is shown by the dashed line.

for most of the flow percentiles though predicted ranges are too narrow for low flows. Sharpness values of high flows for W11 are smaller than that for W15 indicating increased uncertainty for the case of W11.

On the other hand it was also noticed that the difference in both reliability and sharpness for the two watersheds is small indicating that the approach might be robust for most cases. To see if this is a conclusive statement, a detailed study of all 19 response characteristics as constraints for all 30 watersheds was done. For this purpose, the reliability and sharpness values for independent response characteristics for all 30 watersheds were aggregated to analyze their impact on different parts of streamflow time series. It was observed that while reliability values for all the flow percentiles were mainly above 90%, the sharpness values varied extensively. This suggests that there is a trade-off between these values since higher values of both reliability and sharpness are wanted for a favorable result. It was not possible to identify the role of individual response characteristics with respect to reliability and sharpness values from these figures. Thus another set of figures was developed that shows the impact of individual response characteristics on different percentiles of flow exceedance values (Fig. 10a–d).

Fig. 10a–d show the average reliability and sharpness values of each response characteristic on different flow

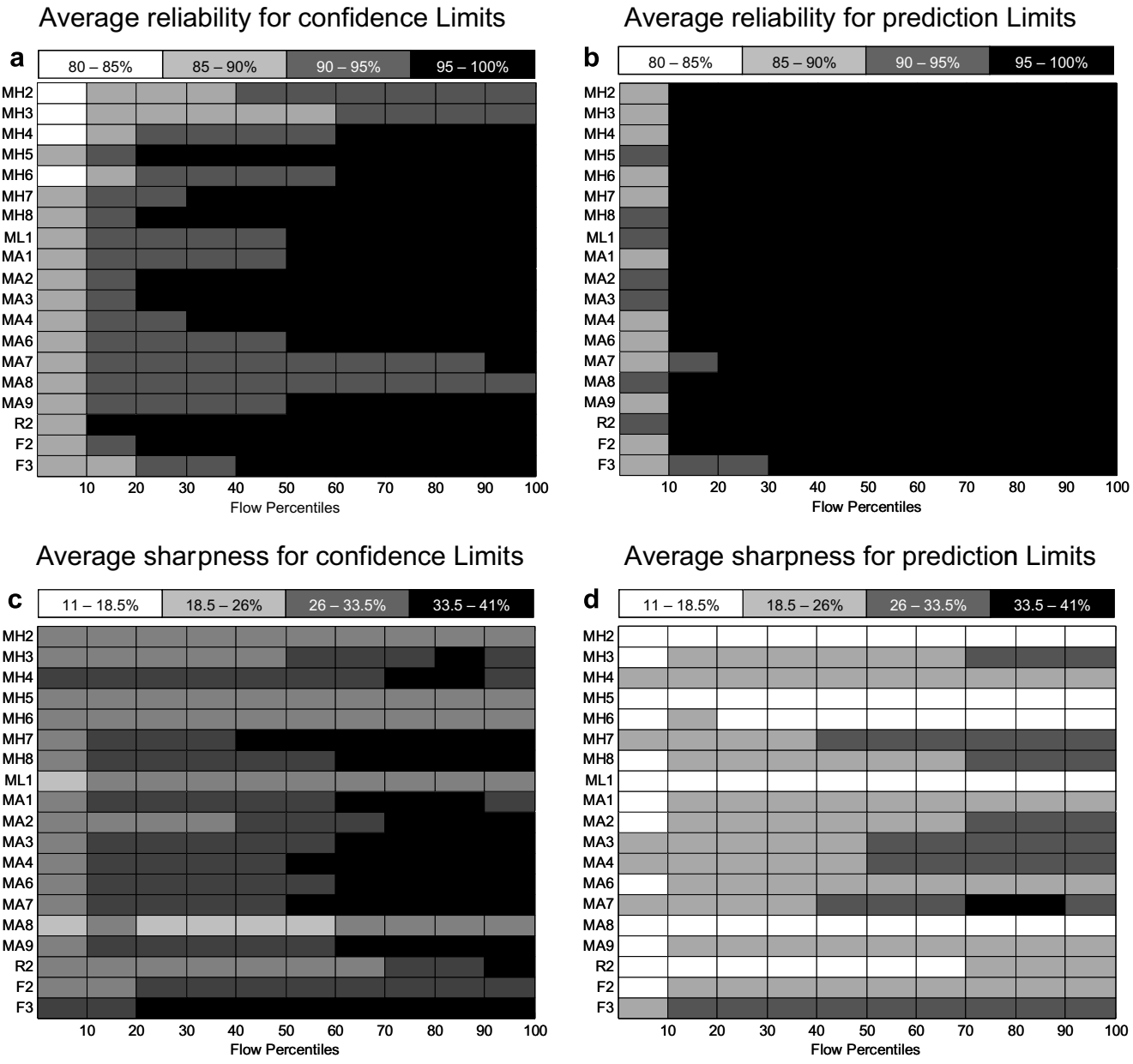


Fig. 10. Average reliability and sharpness values of response characteristics calculated when parameter space is constrained by confidence intervals and prediction intervals of response characteristics. Values are shown for different percentile flow exceedance values and are averaged for all the watersheds. Darker shades of grey show higher values. The legend above each figure shows the range of values for different shades of gray.

exceedance percentiles for all watersheds. The shades of gray represent the relative values from white (minimum) to black (maximum). It was observed from Fig. 10a that almost all response characteristics showed high reliability values for low flows or when the percentile flow exceedance value is higher. A few of the response characteristics (e.g. MH5, MH8, MA2, MA3, R2 and F2) also have high reliability values for high flows. This shows that

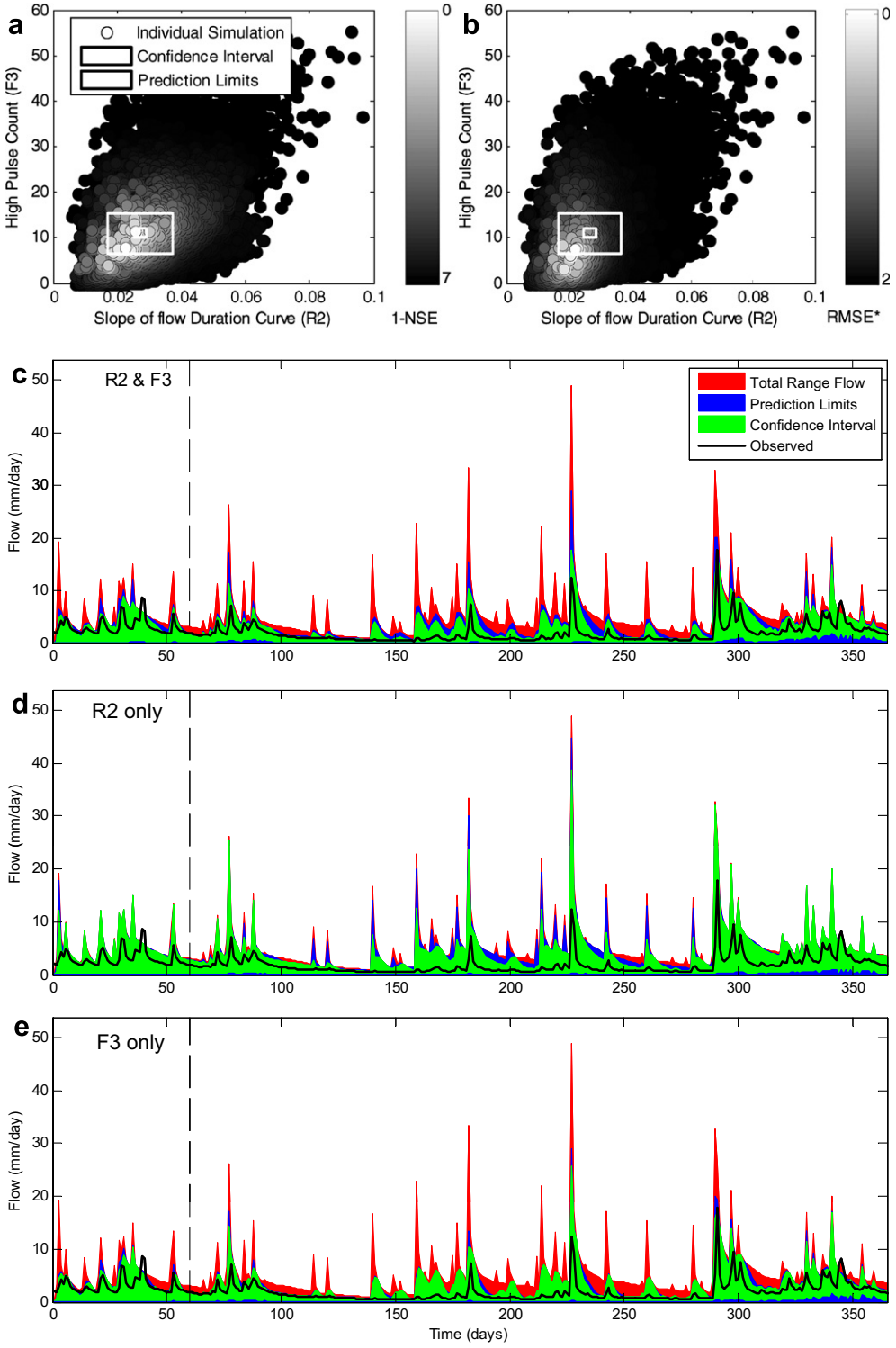


Fig. 11. Dotty plots of R2 and F3 showing their impact as constraints on hydrologic prediction. The circles in (a) and (b) show a single simulation value of R2 and F3 with shades of gray representing the value of the performance measure (1-NSE for (a) and RMSE* for (b)). (c) shows the prediction ranges as a result of R2 and F3 as constraints. The bottom figures show the prediction bands when R2 (slope of flow duration curve; d) and F3 (high pulse count; e) alone are used to constrain simulated stream flows.

observed low flows were almost always within the prediction band obtained from the regression relationship and that the high flows were either not characterized by most of the response characteristics or were not simulated well by the hydrologic model used in this study. The trade-off between reliability and sharpness is evident from the comparison of using the wider prediction limits as constraints (Fig. 10b and d) versus the narrower confidence limits (Fig. 10a and c).

The next logical step is then to test how combinations of constraints can be applied. As an example, Fig. 11 shows the effect of response characteristics R2 and F3 on watershed W53, where both response characteristics were used to determine whether a simulation was behavioral (i.e. if the simulated values of both the response characteristics for 3rd watershed in 5th cross validation group were within their prediction limits the simulation was considered as behavioral).

It was noticed that the number of behavioral simulations decreased when the constraints imposed by the two regression equations were applied simultaneously (see Fig. 11c): resulting in a further narrowing of the confidence and prediction bands (compare with Fig. 11d and e). Again, the observed flow and best simulations lie within

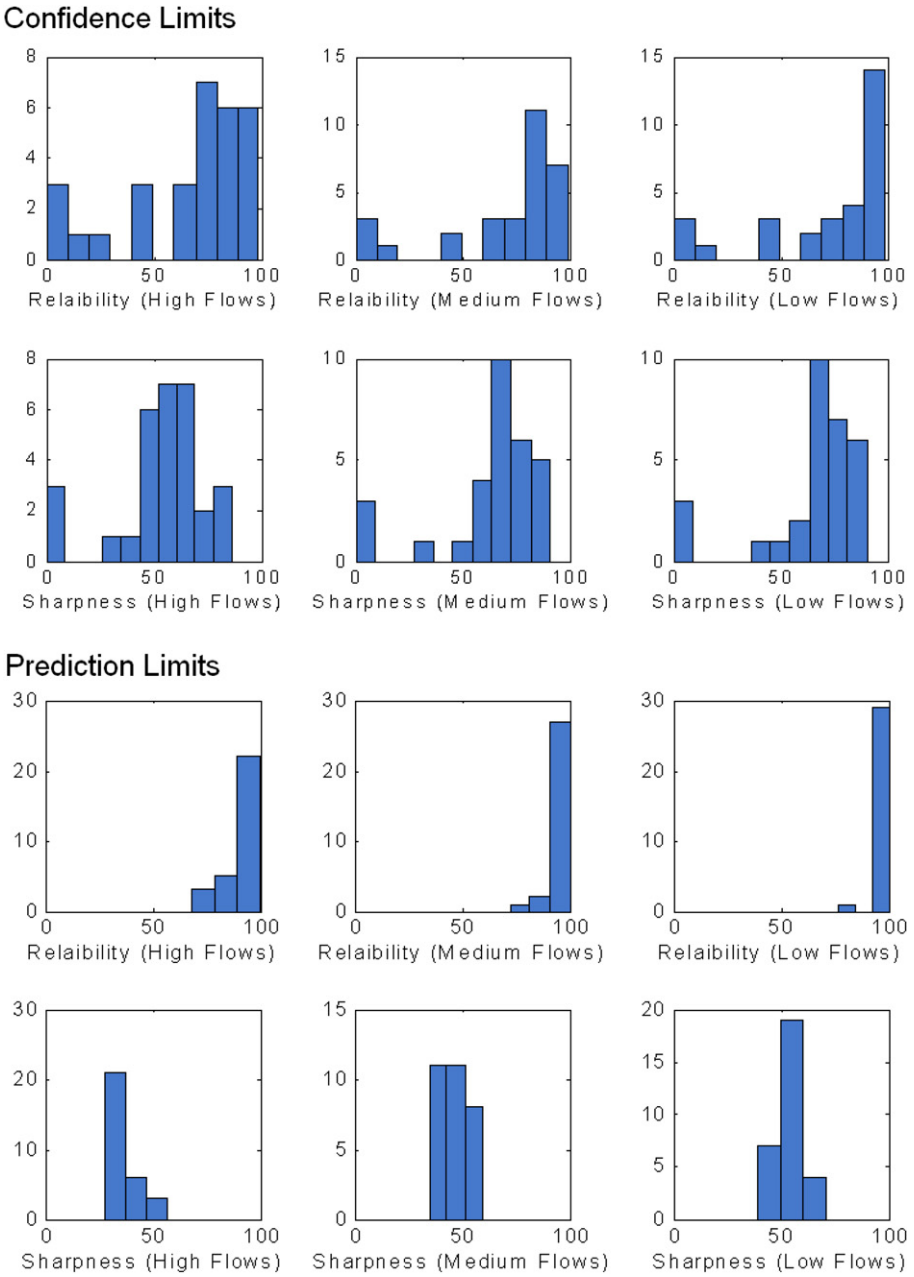


Fig. 12. Histogram of reliability and sharpness values for high, medium and low flows when confidence limits of high pulse count, runoff ratio and slope of flow duration curve are used to constrain ensemble predictions (a). (b) shows the histogram of reliability and sharpness values when prediction limits of these response characteristics are used to constrain flows.

the predicted range. The confidence and prediction limits for this case are shown in Fig. 11a and b. When the flow was constrained by the wider ranges (prediction limits) of both response characteristics, then the number of behavioral simulations was 4421 (44%), reducing to only 642 (6%) when the flow was constrained by the narrower confidence limits. The reliability values obtained from the analysis range from 81% for high flows to 100% for low flows and the sharpness values range from 33% for high flows to 53% for low flows. This shows that even after the streamflow is constrained by more than one response characteristic simultaneously, good reliability and sharpness values can be obtained. However, when all 19 response characteristics were used simultaneously to constrain the streamflow prediction limits, it was observed that 12 out of 30 validation watersheds did not yield any behavioral simulations.

Since some of the response characteristics better constrain the low flows while others better constrain the high flows, a sensible combination of these different response characteristics should constrain the predictive uncertainty even further. This was implemented by grouping the response characteristics based on the specific portions of streamflow time series they impact. For example, using the prediction intervals as the constraining factor, the sharpness and reliability values can be used to classify the response characteristics based on their impact on either high, medium or low flows. The following classification was found:

(a) High Flows – MH7, MA3, F3
(b) Medium Flows – MH4, MA1, MA4, MA7
(c) Low Flows – MH8, MA2, MA3, MA4, MA6, MA7, MA9, ML1, R2

To maintain a balance between the reliability and sharpness values, three response characteristics were selected from the above groups to constrain high, medium and low flows respectively. These response characteristics were: high pulse count (high flows), runoff ratio (medium flows) and slope of the flow duration curve (low flows). Fig. 12 shows the histogram of reliability and sharpness values for high, medium and low flows when these response characteristics were used to simultaneously constrain the ensemble flows for all 30 watersheds. The reliability and sharpness values are seen to be high across the range of flows.

Finally, we note that three physical characteristics appear to explain most of the observed streamflow behavior, i.e. the dominant independent variables in the regression analysis for high pulse count, runoff ratio and slope of flow duration curve. They are the wetness index (P/PE) describing climate, the average surface slope (DPSBAR) describing topography, and the base-flow index (BFIHOST), an integrated measure of subsurface characteristics. This empirical result corroborates the watershed classification suggested by Winter [48]; see also Wolock et al. [49] who suggests that watersheds should be grouped by similarity in climate, topography and geology.

5. Conclusions

This study presents a novel approach to hydrologic ensemble predictions in ungauged basins. The approach is based on identifying relationships between physical characteristics and dynamic response characteristics of watersheds. The relationships between physical and dynamic characteristics found were generally higher than those commonly reported for hydrologic model parameters and physical characteristics e.g. [30,43,44]. The classical approach for regionalizing model parameters requires choosing a particular model structure and calibration of the model parameters in the gauged watersheds – both elements are likely to have a negative impact on correlations between parameters and catchment characteristics since model structural error and uncertainty in finding hydrologically realistic parameter values are unavoidable [41]. The response characteristics determined by our approach are subsequently regionalized within an uncertainty framework to provide constraints on the model behavior at ungauged sites. The overall approach was tested on 30 watersheds distributed throughout England and Wales. It was found that high pulse count, runoff ratio and the slope of the flow duration curve provided strong regionalizable constraints while still allowing for reliable predictions, i.e. most of the observed flow was captured by the ensemble. The dominant physical characteristics are climate (wetness index, P/PE), watershed topography (slope, DPSBAR) and subsurface geology and soils (base-flow index, BFIHOST). Using too many dynamical characteristics simultaneously often resulted in a rejection of all models.

In general, the approach yielded very promising results and has several advantages compared to the common regionalization of models. The main advantages are that it (1) is applicable to *any* hydrologic model (lumped or distributed), (2) is not impacted by problems of parameter calibration or model structural error, (3) does not try to establish relationships between conceptual (effective) model parameters and watershed characteristics, and (4) yields increased understanding about the controls on watershed response behavior at the scale of interest, which could guide an improved approach to watershed classification [45].

Acknowledgements

Partial support for this work was provided by SAHRA under NSF-STC grant EAR-9876800, the National Weather Service Office of Hydrology under grant numbers NOAA/NA04NWS4620012, UCAR/NOAA/COMET/S0344674, NOAA/DG 133W-03-SE-0916, and the United States Geological Survey under USDI/USGS 432-41(69AR). We thank Ian Littlewood, the PUB Top-Down modeling Working Group, and the Centre for Ecology and

Hydrology for some of the datasets used in this study. We also thank The British Atmospheric Data Center for providing the temperature data (http://badc.nerc.ac.uk/home/index.html).

References

[1] Abdulla FA, Lettenmaier DP. Development of regional parameter estimation equations for a macro scale hydrologic model. J Hydrol 1997;197:230–57.
[2] Beighley RE, Moglen GE. Trend assessment in rainfall runoff behavior in urbanizing watersheds. J Hydrol Eng, ASCE 2002;7(1):27–34.
[3] Berger KP, Entekhabi D. Basin Hydrologic Response relations to distributed physiographic descriptors and climate. J Hydrol 2001;247:169–82.
[4] Beven KJ. Changing ideas in hydrology – the case of physically-based models. J Hydrol 1989;105:157–72.
[5] Beven KJ. Rainfall-runoff modeling – the primer. Chichester: Wiley; 2001.
[6] Boorman DB, Hollis JM, Lilly A. 1995. Hydrology of soil types: a hydrologically-based classification of the soils of the United Kingdom. Institute of Hydrology Report No. 126, Wallingford, UK.
[7] Boyle DP, Gupta HV, Sorooshian S. Toward improved calibration of hydrologic models: combining the strengths of manual and automatic approach. Water Resour Res 2000;36(12):3663–74.
[8] Chinnayakanahalli KJ, Tarboton DG, Hawkins CP. Predicting hydrologic flow regime for biological assessment at ungauged basins in the western United States. Eos Trans AGU 2005;86(52). Fall Meet. Suppl. 2005.
[9] Clausen B, Biggs BJF. Flow variables for ecological studies in temperate streams: grouping based on covariance. J Hydrol 2000;237:184–97.
[10] Clausen B, Iversen HL, Oversen NB. Ecological flow indices from Danish streams. In: Nilssen T, editor. Nordic hydrological conference 2000, Uppsala, Sweden, 2000; p. 3–10.
[11] Colwell RK. Predictability, constancy, and contingency of periodic phenomena. Ecology 1974;55:1148–53.
[12] Czikowsky MJ, Fitzjarrald DR. Effect of seasonal changes in evapotranspiration in eastern US hydrological records. J Hydrometeorol 2004;5(5):974–88.
[13] Detenbeck NE, Brady VJ, Taylor DL, Snarski VM, Batterman SL. Relationship of stream flow regime in the western Lake Superior basin to watershed type characteristics. J Hydrol 2005;309:258–76.
[14] Duan Q, Sorooshian S, Gupta HV. Shuffled complex evolution approach for effective and efficient global optimization of conceptual rainfall-runoff models. Water Resour Res 1992;28(4):1015–31.
[15] Fernandez W, Vogel RM, Sankarasubramanian S. Regional calibration of a watershed model. Hydrol Sci J 2000;45(5):689–707.
[16] Hannah DM, Smith BPG, Gurnell AM, McGregor GR. An approach to hydrograph classification. Hydrol Process 2000;14:317–38.
[17] Harris NM, Gurnell AM, Hannah DM, Petts GE. Classification of river regimes: a context for hydro ecology. Hydrol Process 2000;14:2831–48.
[18] Hughes JMR, James B. A hydrological regionalization of streams in Victoria, Australia, with implication for stream ecology. Austral J Marine Freshwater Res 1989;40:303–26.
[19] Gupta HV, Sorooshian S, Yapo PO. Toward improved calibration of hydrologic models: mutliple and noncommensurable measures of information. Water Resour Res 1998;34(4):751–63.
[20] Jakeman AJ, Hornberger GM, Littlewood IG, Whitehead PG, Harvey JW, Bencala KE. A systematic approach to modelling the dynamic linkage of climate, physical catchment descriptors and hydrological response components. Math Comput Simul 1992;33:359–66.
[21] Kottegoda NT, Rosso R. Statistics, probability, and reliability for civil and environmental engineers. New York, USA: McGraw-Hill; 1997.
[22] McIntyre N, Lee H, Wheater HS, Young A, Wagener T. Ensemble predictions of runoff in ungauged catchments. Water Resour Res 2005;41:W12434. doi:10.1029/2005WR004289.
[23] Merz B, Bloeschl G. Regionalization of catchment model parameters. J Hydrol 2004;287(1–4):95–123.
[24] Moore RJ. The probability-distributed principle and runoff production at point and basin scales. Hydrol Sci J 1985;30(2):273–97.
[25] Morin E, Georgakakos KP, Shamir U, Garti R, Enzel Y. Objective, observation-based, automatic estimation of the catchments response time scale. Water Resour Res 2002;38(10):1212–27.
[26] Nash JE, Sutcliffe JV. River flow forecasting through conceptual models – Part I: A discussion of principles. J Hydrol 1970;10(3):282–90.
[27] Olden JD, Poff NL. Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. River Res Appl 2003;19:101–21.
[28] Post DA, Jones JA, Grant GE. An improved methodology for predicting the daily hydrologic response of ungauged catchments. Environ Mod Software 1998;13:395–403.
[29] Richter BD, Baumgartner JV, Powell J, Braun DP. A method for assessing hydrologic alteration within ecosystems. Conservat Biol 1996;10(4):1163–74.
[30] Sefton CEM, Howarth SM. Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. J Hydrol 1998;211:1–16.
[31] Seibert J. Regionalization of parameters for a conceptual rainfall-runoff model. Agric Forest Meteor 1999;98–99:279–93.
[32] Shamir E, Imam B, Morin E, Gupta HV, Sorooshian S. The role of hydrograph indices in parameter estimation of rainfall-runoff models. Hydrol Proc 2004;19:2187–207.
[33] Shamir E, Imam B, Gupta HV, Sorooshian S. Application of temporal streamflow descriptors in hydrologic model parameter estimation. Water Resour Res 2005;41:W06021. doi:10.1029/2004WR003409.
[34] Shuttleworth JW. Evaporation. In: Maidment DR, editor. Handbook of hydrology. McGraw-Hill; 1992. p. 4.1–4.53.
[35] Sivapalan M, Takeuchi K, Franks SW, Gupta VK, Karambiri H, Lakshmi V, et al. IAHS decade on predictions in ungauged basins (PUB), 20 03-20 12: shaping an exciting future for the hydrological sciences. Hydrol Sci J 2003;48(6):857–80.
[36] Spate JM, Croke BFW, Norton JP. A proposed rule-discovery scheme for regionalization of rainfall-runoff characteristics in New South Wales, Australia. In: Transactions of the 2nd biennial meeting of the international environmental modelling and software society, iEMSs, Manno, Switzerland, 2004, ISBN 88-900787-1-5.
[37] Tang Y, Reed P, Wagener T. How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration? Hydrol Earth Syst Sci Discuss 2005;2:2465–520.
[38] Todini E. Rainfall-runoff modeling – past, present and future. J Hydrol 1988;100:341–52.
[39] Vrugt JA, Gupta HV, Bouten W, Sorooshian S. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resour Res 2003;39(8):1201. doi:10.1029/2002WR001642.
[40] Wagener T, Gupta HV. Model identification for hydrological forecasting under uncertainty. Stoch Environ Res Risk Assess 2005. doi:10.1007/s00477-005-0006-5.
[41] Wagener T, Wheater HS. Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. J Hydrol 2006;320:132–54.
[42] Wagener T, Boyle DP, Lees MJ, Wheater HS, Gupta HV, Sorooshian S. A framework for the development and application of hydrological models. Hydrol Earth System Sci 2001;5(1):13–26.

[43] Wagener T, Wheater HS, Gupta HV. Rainfall-runoff modelling in gauged and ungauged catchments. London, UK: Imperial College Press; 2004. p. 300.
[44] Wagener T, Sivapalan M, McDonnell JJ, Hooper R, Lakshmi V, Liang X, Kumar P. Predictions in ungauged basins (PUB) – a catalyst for multi-disciplinary hydrology. Eos Trans AGU 2004;85(44):451–2.
[45] Wagener T, Sivapalan M, Troch PA, Woods R. Catchment classification and hydrologic similarity. Geogr Compass, in press.
[46] Weeks WD, Ashkanasy NM. Regional parameters for the Sacramento models: a case study. In: Hydrology and water resources symposium, Hobart, Australia, 1984.
[47] Wheater HS, Jakeman AJ, Beven KJ. Progress and directions in rainfall-runoff modelling. In: Jakeman AJ et al., editors. Modelling change in environmental systems. John Wiley and Sons; 1993. p. 101–32.
[48] Winter TC. The concept of hydrologic landscapes. J Am Water Resour Assoc 2001;37:335–49.
[49] Wolock DM, Winter TC, McMahon G. Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses. Environ Manage 2004;34:S71–88.
[50] Yadav M, Wagener T, Gupta HV. Regionalization of dynamic watershed behavior. In: Andréassian V, Chahinian N, Hall A, Perrin C, Schaake J. editors. Large sample basin experiments for hydrological model parameterization Results of the Model Parameter Estimation Experiment (MOPEX) Paris (2004) and Foz de Iguaçu (2005) workshops, IAHS Redbook, Publ. no. 307, 2006.
[51] Yu PS, Yang TC. Using synthetic flow duration curves for rainfall-runoff model calibration at ungauged sites. Hydrol Process 2000;14(1):117–33.