Hydrol. Earth Syst. Sci. Discuss., 8, 9323–9355, 2011
www.hydrol-earth-syst-sci-discuss.net/8/9323/2011/
doi:10.5194/hessd-8-9323-2011

Hydrology and Earth System Sciences Discussions

This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Controls on hydrologic similarity: role of nearby gauged catchments for prediction at an ungauged catchment

S. Patil[1] **and M. Stieglitz**[1,2]

[1]School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA
[2] School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA

Received: 28 September 2011 – Accepted: 11 October 2011 – Published: 19 October 2011

Correspondence to: S. Patil (sopan.patil@gatech.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Prediction of streamflows at ungauged catchments requires transfer of hydrologic information (e.g., model parameters, hydrologic indices, streamflow values) from gauged (donor) to ungauged (receiver) catchments. One of the most reliable metrics for selection of ideal donor catchments is the spatial proximity between donor and receiver catchments. However, it is not clear whether information transfer among nearby catchments is suitable across a wide range of climatic and geographic regions. We examine this issue using the data from 756 catchments within the continental United States. Each catchment is considered ungauged in turn and daily streamflow is simulated through distance-based interpolation of streamflows from neighboring catchments. Results show that distinct geographic regions exist in US where transfer of streamflow values from nearby catchments is useful for retrospective prediction of daily streamflow at ungauged catchments. Specifically, the high predictability catchments (Nash-Sutcliffe efficiency NS > 0.7) are confined to the Appalachian Mountains in eastern US, the Rocky Mountains, and the Cascade Mountains in the Pacific Northwest. Low predictability catchments (NS < 0.3) are located mostly in the drier regions west of Mississippi river, which demonstrates the limited utility of gauged catchments in those regions for predicting at ungauged basins. The results suggest that high streamflow similarity among nearby catchments (and therefore, good predictability at ungauged catchments) is more likely in humid runoff-dominated regions than in dry evapotranspiration-dominated regions. We further find that higher density and/or closer distance of gauged catchments near an ungauged catchment does not necessarily guarantee good predictability at an ungauged catchment.

1 Introduction

Long-term measurements of river streamflow are essential for a number of applications in water resources, such as, planning of water supply and irrigation projects (Dunne and Leopold, 1978; Jain and Singh, 2003), delineation of river floodplains (Merwade et al., 2008; Tate et al., 2002), day-to-day management of dams and canals (Hirsch and Costa, 2004), to name a few. Streamflow measurements are also important for characterizing the hydrologic behavior of river basins within modeling frameworks, so that future assessments of hydrologic behavior in response to climate and/or land-use change can be obtained. However, in many parts of the world, developed as well as developing, rivers are not gauged for continuous monitoring. Developing strategies for prediction at ungauged basins (PUB; Sivapalan et al., 2003) is therefore required not only for the above practical applications, but also for advancing our process understanding of the controls on regional variability in landscape hydrologic response (Patil and Stieglitz, 2011; Wagener et al., 2004).

Prediction of streamflows at ungauged basins typically requires transfer of hydrologic information (e.g., model parameters, hydrologic indices, streamflow values) from gauged to ungauged catchments. Transfer of hydrologic model parameters is perhaps the most common procedure (Merz and Blöschl, 2004). Hydrologic models that have been used extensively for this purpose include: HBV (Götzinger and Bárdossy, 2007; Merz and Blöschl, 2004; Seibert, 1999), IHACRES (Kokkonen et al., 2003; Post and Jakeman, 1996; Schreider et al., 1996), PDM (Kay et al., 2007; Lamb and Kay, 2004). Alternatively, model-independent methods have been developed that incorporate the spatial variability of streamflow within a region (Archfield and Vogel, 2010; Skøien and Blöschl, 2007; Smakhtin, 1999). These model-independent methods have the advantage that they can circumvent the deficiencies associated with any particular hydrologic model structure. Smakhtin et al. (1997) developed regionalized flow duration curves for catchments in South Africa and estimated streamflows at ungauged catchments through transfer of daily streamflow data from nearby gauged catchments using the interpolation technique described in Hughes and Smakhtin (1996). Archfield and Vogel (2010) developed the map-correlation method to identify the donor gauged catchment that is likely to have high correlation to the ungauged catchment for direct transfer of daily streamflow time series. Skøien and Blöschl (2007) used the concept of topological kriging, or top-kriging (Skøien et al., 2006), on 376 catchments in Austria to predict hourly and daily streamflow in ungauged catchments.

A common challenge for all information transfer procedures is the search for ideal donor (gauged) catchments from which hydrologic information can be successfully transferred. Recent studies have shown that choosing the donor catchments based on spatial proximity to the ungauged catchment alone is by far the most reliable approach. For example, Merz and Blöschl (2004) compared eight parameter transfer methods using the HBV model for 308 catchments in Austria and concluded that methods based on spatial proximity alone performed better than those based on catchment attributes. Oudin et al. (2008) compared three different regionalization approaches on 913 French catchments using two hydrologic models and found that the spatial proximity approach outperformed other approaches. Zhang and Chiew (2009) obtained similar results using data from 210 catchments in Australia. While these studies aptly demonstrate the advantage of spatial proximity based approach, what is not clear is whether the information transfer from nearby gauged to ungauged catchments is suitable across a wide range of hydro-climatic settings.

In this study, we characterize the transferability of hydrologic information among nearby catchments. Our objectives are to: (1) determine if distinct geographic regions exist where nearby catchments tend to have similar streamflows, so that information can be easily transferred between gauged and ungauged catchments, and (2) identify the physiographic and hydro-climatic conditions that favor streamflow similarity among nearby catchments within a region. We use the data from 756 gauged catchments across the continental United States to simulate daily streamflow through inverse distance weighted (IDW) interpolation of streamflow from neighboring gauged catchments. The prediction efficiency at ungauged catchments is then compared against physiographic and hydro-climatic properties to identify the conditions that favor high streamflow similarity within a region. In Sect. 2, we describe the 756 catchments chosen for this study and the associated environmental data used for our analysis. In Sect. 3, we outline the distance-based interpolation method used for streamflow simulation, the goodness-of-fit measures used to assess predictability, and metric used to analyze the relationship between predictability and catchment properties. In Sect. 4, we present results of our analysis. Sections 5 and 6 provide the discussion of our results and the conclusions of this study, respectively.

2 Data

We use the streamflow records of catchments from US Geological Survey's Hydro-Climate Data Network (HCDN) (Slack et al., 1993). The HCDN database consists of data of 1659 catchments located within the United States that are not severely affected by human activity. The streamflow records in HCDN span from 1874 to 1988. A majority of the catchments have consistent and continuous records from water year 1970 onwards. As such, we consider only those catchments that have a continuous daily streamflow record from water year 1970 to 1988 (i.e., 1 October 1969 to 30 September 1988), which reduces the number of acceptable catchments to 756 (see Fig. 1). The drainage area of the catchments ranges from 23 km^2 to 5000 km^2, and the median drainage area is 715 km^2.

Monthly time-series of precipitation and potential evapotranspiration in each of the 756 catchments are obtained from the climate dataset developed by Vogel and Sankarasubramanian (2005). Estimates of monthly precipitation in this dataset are obtained from the PRISM (Daly et al., 1994) climate analysis system as described in Vogel et al. (1999), whereas the monthly potential evapotranspiration estimates are obtained using the formula suggested by Hargreaves and Samani (1982). From the streamflow and precipitation data, we derive five hydrologic indices (or signatures) for each of the 756 catchments. These hydrologic signatures are: baseflow index, runoff ratio, baseflow runoff ratio (baseflow/rainfall), slope of flow duration curve, and inter-annual streamflow elasticity. We also use the data of three physiographic attributes from Vogel and Sankarasubramanian (2005) dataset, viz., channel slope, soil permeability, and soil water holding capacity. Details of the methods used for deriving the hydrologic signatures are provided in Appendix A.

HESSD
8, 9323–9355, 2011

As illustrated in Fig. 1, the number of stream gauges is higher in the eastern half of the country than in the western half. Since the regionalized predictions in this study are based on proximity of gauged and ungauged catchments, this variable gauge density may bias our analysis. Thus, in Sect. 4.4, we evaluate whether gauge density has an appreciable effect on streamflow predictions at ungauged catchments.

3 Methods

In this section, we first outline the distance based interpolation method used for simulating daily streamflows. We then describe the goodness-of-fit measures used for assessing the prediction efficiency, followed by a brief explanation of the metric used for assessing the relationships between prediction efficiency and catchment properties.

The inverse distance weighted (IDW) interpolation is one of the simplest methods to determine whether streamflow values among spatially proximate catchments are similar. Nonetheless, as will be shown in results, this method is highly effective in characterizing the regionalized predictability patterns over the scale of continental US. Comparison of different interpolation methods is beyond the scope of this study. The mathematical expression of the IDW interpolation scheme is as follows:

$$q(x) = \sum_{k=1}^{N} \frac{w_k(x)}{\Sigma_{k=1}^{N} w_k(x)} \times q(x_k) \tag{1}$$

And,

$$w_k(x) = \frac{1}{d(x,x_k)^p} \tag{2}$$

where, $q(x)$ is the discharge value at the ungauged catchment that is located at point x in the region, $q(x_k)$ is the discharge value at neighboring donor catchment k located at point x_k in the region, and N is the total number of neighboring donor catchments

HESSD

8, 9323–9355, 2011

considered for the interpolation. The daily streamflow values are normalized by catchment drainage area and have the units of mm day^{-1}. Distance between the gauged and ungauged catchment d is calculated individually for each of the N neighboring catchments. d is the distance between stream gauges of the catchments. The interpolation weights w are calculated for all the donor catchments using Eq. (2). The exponent p in Eq. (2) is a positive real number, called as power parameter.

Each of the 756 catchments is considered ungauged in turn (jack-knife procedure), and daily streamflows are simulated using Eqs. (1) and (2). We use power parameter $p = 2$ (i.e., the inverse square distance weighted method) and vary the number of neighboring donor catchments N from 1 to 50.

Goodness of fit for predicted hydrograph is measured using two metrics: Nash-Sutcliffe efficiency (NS) and water balance error (WBE). These two measures convey different information about prediction performance. The WBE verifies whether long-term differences in the magnitudes of observed and simulated streamflows are within an acceptable range, whereas the NS verifies whether the fluctuations in daily hydrograph are appropriately captured. These metrics have been extensively used in the hydrology literature to determine the simulation efficiency of daily hydrographs. The Nash-Sutcliffe efficiency criterion (Nash and Sutcliffe, 1970) is defined as follows:

$$\mathrm{NS} = 1 - \frac{\Sigma_{i=1}^{n}(Q_{\mathrm{obs},i} - Q_{\mathrm{pred},i})^2}{\Sigma_{i=1}^{n}(Q_{\mathrm{obs},i} - \overline{Q}_{\mathrm{obs}})^2} \tag{3}$$

where, $Q_{\mathrm{pred},i}$ and $Q_{\mathrm{obs},i}$ are the predicted and the observed stream discharge values on the i-th day respectively, $\overline{Q}_{\mathrm{obs}}$ is the mean of all the observed discharge values and n is the total number of days in the record.

The water balance error is defined as follows:

$$\mathrm{WBE} = 100 \times \left| \frac{\sum_{i=l}^{n} Q_{\mathrm{pred},i} - \sum_{i=l}^{n} Q_{\mathrm{obs},i}}{\sum_{i=l}^{n} Q_{\mathrm{obs},i}} \right| \tag{4}$$

We analyze the relationship of prediction efficiency (NS values) with numerous catchment properties. These relationships are analyzed to identify the factors that favor streamflow similarity among nearby catchments. To this end, we use the Spearman's rank correlation (Spearman, 1904), which quantifies the increasing/decreasing trend in a relationship. The formula for Spearman's correlation (ρ) is as follows:

$$\rho = 1 - \frac{6 \times \sum d^2}{M \times (M^2 - 1)} \tag{5}$$

Where, d is the difference between the ranks of each observation on the two variables under consideration, and M is the total number of observation points ($M = 756$ in our case). Spearman's ρ varies from -1 to $+1$, with -1 being a perfect monotonically decreasing relationship and $+1$ being perfect monotonically increasing.

4 Results

4.1 Choosing the optimal number of donor gauged catchments

To find the optimal number of donor catchments required for a good streamflow estimate we vary the number of nearest donor catchments from 1 to 50 and calculate the associated NS. This approach for choosing the optimal number of donors has been used previously (Oudin et al., 2008; Zhang and Chiew, 2009). Figure 2a shows the relationship between the number of donor gauged catchments used for simulating daily streamflow and the median NS from each simulation run for all the 756 catchments.

HESSD
8, 9323–9355, 2011

The median NS increases sharply from 0.49 to 0.61 as the number of donor catchments increase from 1 to 4 followed by small increases in median NS for subsequent increases in the number of donor catchments. The median NS reaches its highest value of 0.615 at 15 donor catchments. Beyond 15 donor catchments there is decline in simulation efficiency that can be attributed to the relative reduction in influence of the nearby catchments. For subsequent analysis, we therefore limit the number of donors to five nearest gauged catchments and perform the distance based interpolation to simulate daily streamflows. For simulations with five donor catchments, the maximum NS is 0.97, the median value is 0.61, and the 25th percentile value is 0.29. Figure 2b shows the 25th and 75th percentile NS values along with median NS against the number of donor catchments. Similar to the median, other percentile values also show that increasing the number of donors far beyond 4 or 5 does not cause an increase in prediction performance. Figure 2c shows the distribution of NS values of simulated streamflows using five donor catchments.

4.2 Geographic patterns of streamflow predictions

Distinct geographic patterns are observed in the NS and WBE values of catchment streamflows using the IDW interpolation method (Figs. 3 and 4). For better identification of these geographic patterns, we partition the catchments into three groups: Group 1 for NS greater than 0.7, Group 2 for NS between 0.3 and 0.7, and Group 3 for NS less than 0.3. Figure 3a shows the location of all the Group 1 catchments, 288 in total (~40 %), which have the highest predictability of daily streamflow. The majority of the Group 1 catchments are located in three geographic regions: (1) the Appalachian mountain ranges in the eastern US, (2) the Rocky Mountains, and (3) the Pacific Northwest region to the west of Cascade Mountain range. The remaining Group 1 catchments are located across the eastern half of continental US, especially in the states of Indiana and Illinois (Fig. 3a). The Group 2 catchments (Fig. 3b, a total of 277 catchments; ~35 %), are located across the eastern part of the United States. The poorest performers, Group 3 catchments are located in the western half of continental US,

HESSD
8, 9323–9355, 2011

especially to the west of Mississippi river (Fig. 3c). There are 191 catchments (~25 %) that belong in Group 3 and these are considered as practically unpredictable using the spatial proximity based regionalization method.

For water balance errors (WBE), we group catchments as follows: Group A for WBE < 20 %, Group B for 20 % < WBE < 50 %, and Group C for WBE >50 %. Figure 4 shows the geographic patterns of WBE for these three catchment groups. Figure 4a shows the Group A catchments, which have the smallest water balance errors. A total of 473 (~63 %) catchments belong to Group A, and majority of them are located in the eastern part of US and along the west coast. The Group B catchments (152 in total; ~20 %, see Fig. 4b) are spread throughout the US and do not have any preferential geographic patterns. Figure 4c shows the Group C catchments, 131 in total (~17 %), which have the highest water balance errors and are located mostly in drier western part of the US. The geographic patterns of Group C catchments are similar to the catchments with lowest NS values (Group 3, Fig. 3c).

4.3 Impact of catchment proximity on predictability at ungauged catchments

Figure 5a shows the relationship of prediction efficiency (NS) with the average distance of donor catchments from the ungauged catchment, while Fig. 5b shows its relationship with the distance of nearest donor catchment. As expected, the observed trend is that high NS catchments have donor catchments in closer proximity, i.e., smaller distances. The Spearman rank correlation (ρ) for the relationships of NS with the average and minimum distance is −0.44 and −0.41 respectively (p-value < 0.01 in both cases). However, at any given NS value, there is a surprisingly wide scatter of distances between donor and receiver catchments (Fig. 5a and b). This suggests that the donor-receiver catchment proximity alone cannot fully explain the prediction performance at a given location. Among catchments with NS > 0, the R^2 value of relationship between NS and average donor distance is 0.12, i.e., the average distance from donor catchments explains only 12 % of the spatial variability in NS.

HESSD
8, 9323–9355, 2011

4.4 Impact of gauge density on predictability at ungauged catchments

Gauge density is possibly an important factor that can influence the transfer of information to ungauged catchments. If more gauged catchments are present in the vicinity of an ungauged catchment, we can intuitively expect that catchment to have better predictability. Therefore, we test quantitatively whether disparity in gauge density across different regions of the US influences predictability at an ungauged catchment. Gauge density around a catchment is defined as the number of gauged catchments within the 200 km radius of its location. We tested the gauge density metric by varying the search radius from 100 km to 500 km and found that the relationship between NS and gauge density is not affected by the choice of the search radius (result not shown). Figure 6 shows the relationship between NS and gauge density near the ungauged catchment. Contrary to our a priori expectation, high gauge density around an ungauged catchment does not guarantee good predictability. Moreover, there are numerous catchments that have low gauge density in their vicinity and still have high NS values. No significant trend is observed in the relationship between NS and gauge density. Among catchments with NS > 0, the R^2 value of relationship between NS and gauge density is 0.06, i.e., the density of gauged catchments surrounding within a region explains only 6 % of the spatial variability in NS.

4.5 Impact of climate on predictability at ungauged catchments

We analyze the high and low predictability catchments using the Budyko curve (Budyko, 1974). A Budyko curve characterizes the relationship between aridity index (PET/P) and evaporation index (ET/P) of the catchments. Figure 7 shows the Group 1 catchments (NS > 0.7, blue squares) and Group 3 catchments (NS < 0.3, red squares) on the Budyko curve. Majority of the high predictability catchments (Group 1) have low values of evaporation and aridity indices and are located in the lower portion of the curve. This suggests that the water balance in these high predictability catchments is controlled by energy limitation, i.e., more water is present than can be evaporated. On the other hand, low predictability catchments (Group 3) have higher values of evaporation and aridity indices and are located in the higher portion of the curve. About 48 % of the Group 3 catchments have aridity index >1, suggesting that their water balance is controlled by water limitation, i.e., less water is present than can be evaporated. Thus, the Budyko curve shows that the predictability is higher in regions where the ET of catchments is demand limited (i.e., humid) and low where the ET is supply limited (i.e., arid).

4.6 Physical conditions favoring good predictions at ungauged catchments

To identify the physical conditions that favor high streamflow similarity (and therefore good predictability), we explore the relationships between NS and catchment attributes. Eight catchment properties are considered: three physiographic properties (channel slope, soil permeability, and soil water holding capacity) that are obtained for each catchment from the Vogel and Sankarasubramanian (2005) dataset; and five hydrologic signatures (baseflow index, runoff ratio, baseflow runoff ratio, slope of flow duration curve, and inter-annual streamflow elasticity) that are derived from the streamflow and precipitation data (see Appendix A for details).

Figure 8 shows the relationship between prediction efficiency (NS) and each of the three physiographic properties. While none of these properties have a distinct relationship with NS, a majority of the catchments with higher channel slope (>1 %) have high NS value (Fig. 8a). This trend is consistent with the observation that a majority of high NS catchments are located along the three large mountain ranges of the US (Fig. 3a). However, high NS values are not exclusive to catchments with high channel slope. Of the three physiographic properties, only channel slope shows a statistically significant trend in its relationship with NS (Spearman $\rho = 0.21$; see Table 1). No distinct trend is observed in soil permeability except that the preference of higher permeability catchments is towards high NS values (Fig. 8b). No trend whatsoever is observed in the relationship between NS and soil water holding capacity (Fig. 8c).

Figure 9 shows each of the five hydrologic signatures plotted against NS. High scatter is observed in all the five relationships, similar to the observations of physiographic attributes (Fig. 8). Nonetheless an increasing trend with respect to NS is observed in the relationships of runoff ratio (Spearman $\rho = 0.51$), baseflow runoff ratio (Spearman $\rho = 0.46$), and slope of FDC (Spearman $\rho = 0.31$) (Fig. 9a, c, and d, respectively). Although many high NS catchments are clustered towards high values of baseflow index (Fig. 9b), it does not have a significant trend in its relationship with NS. No particular trend (increasing or decreasing) is observed in the relationship between NS and streamflow elasticity (Fig. 9e).

5 Discussion

Distinct geographic regions exist where transfer of streamflow information from nearby gauged catchments results in good streamflow prediction at an ungauged catchment. High streamflow predictability is obtained in humid mountainous regions, whereas the low predictability catchments are predominantly located in the drier regions (Figs. 3 and 4). To our knowledge, the geographic patterns of streamflow similarity (and predictability at ungauged catchments) shown here not been shown before within the continental US, specifically at a daily time-scale and using an information transfer method. The Budyko curve (Fig. 7) illustrates the preference of high predictability catchments in humid regions. Our previous work (Patil and Stieglitz, 2011) characterized streamflow similarity among nearby catchments across multiple flow conditions. Patil and Stieglitz (2011) suggested that the competing influences of precipitation and evaporative demand determine the conditions at which streamflow similarity is manifested. Consistent with their suggestion, the results presented here show that streamflow similarity is more likely to occur in regions where annual precipitation exceeds evaporative demand (i.e., low energy environments). The preference for humid environment is further evident from the tendency of high predictability catchments to be located in regions of high forest density. Figure 10 shows all the 756 catchments mapped with the forest

HESSD
8, 9323–9355, 2011

cover within the US. The forest cover map is obtained from the USGS Global Land Cover Characteristics (GLCC) project (Loveland et al., 1991). With the exception of catchments in the mid-West, almost all the high predictability catchments (Group 1) are located in regions with high amount of forest cover.

While humid climate is certainly favorable for similarity among nearby catchments, climate alone is not sufficient for identifying regions of high streamflow similarity. The clustering of Group 1 catchments along the mountain ranges suggests that topography is also an important factor in determining streamflow similarity (and predictability). For instance, the catchments in southeastern states of Louisiana, Mississippi, Alabama and Florida have humid climate, but a flatter terrain (and most are Group 2 catchments). Due to the strong connection of predictability with geographic features, we had an a priori expectation that the catchments with high (or low) predictability will have distinct physiographic and hydrologic signatures associated with them. However, the relationship of NS with individual catchment properties is weak. Of the eight catchment properties considered, statistically significant positive trends with respect to NS are observed in only four properties: channel slope, runoff ratio, baseflow runoff ratio, and the slope of FDC (see Table 1). These weak relationships are indicative of the difficulties faced by hydrologists in achieving a universally acceptable hydrologic classification of catchments (McDonnell et al., 2007; Wagener et al., 2007).

Although the streamflow predictions in this study are obtained through distance-based interpolation, results show that the distance between donor and receiver catchments cannot fully explain the prediction patterns. It could have been argued that the high NS catchments are preferentially located in humid regions because of the higher gauge density in those regions. However, no clear relationship is found between NS and gauge density (Fig. 6). This suggests that factors other than the spatial proximity among catchments and gauge density play an important role in regional similarity of streamflows. The higher predictability in humid environments is likely to be due to similarity in climatic inputs over larger spatial scales. However, low predictability an ungauged catchment can be due to either one of the three primary causes: (1) the

HESSD
8, 9323–9355, 2011

ungauged catchment is too far from the donor catchments, or (2) the spatial variability in climatic inputs is high in the region surrounding the ungauged catchment, or (3) the hydrologic behavior of the ungauged catchment is idiosyncratic (and therefore, non-representative of the region surrounding it) either due to contributions from deep groundwater sources, loss of water to regional aquifers, or other complex geologic factors.

6 Summary and conclusion

This study examined whether information transfer from nearby gauged to ungauged catchments is suitable across multiple environments. Distinct geographic patterns of daily streamflow predictability at ungauged catchments were observed within the continental US. Specifically, high predictability catchments are located along the Appalachian Mountains in eastern US, the Rocky Mountains, and the Cascade Mountains in Pacific Northwest, whereas the low predictability catchments are located in the drier regions west of Mississippi river. Identification of these patterns provides essential information regarding the usefulness of gauged catchments within a region for predicting streamflow at a nearby ungauged catchment. While the direct transfer of streamflows is useful for retrospective prediction, future forecasts of streamflows will still require implementation of hydrologic models. Our results suggest that streamflow similarity in the high predictability regions increases the likelihood that gauged and ungauged catchments in those regions will have similar model parameters. However, we suspect that model regionalization studies will need to additionally consider whether their chosen model structure is suitable for characterizing the hydrologic response within their region of interest.

Comparison of catchments using the Budyko curve suggests that climate has a dominant control over the regional extent of similarity in hydrologic response. Nonetheless, among the humid regions, high predictability catchments are still preferentially clustered among the mountainous environments. This suggests that the topography of

HESSD
8, 9323–9355, 2011

the region also has the ability to influence similarity in catchment streamflows. However, analysis of individual catchment attributes provides, at best, a weak picture of the physiographic and hydro-climatic conditions that favor high streamflow similarity (and predictability at ungauged catchments). More importantly, our results show that the spatial proximity between gauged and ungauged catchments alone cannot fully explain the prediction performance at a given location. This suggests that a combined influence of spatial proximity, regional climate variability and geologic settings contributes towards meaningful information transfer between the gauged and ungauged catchments.

Appendix A

Deriving the hydrologic signatures of a catchment

Five hydrologic indices (or signatures) are derived individually for each of the 756 catchments. These hydrologic signatures are: baseflow index, runoff ratio, baseflow runoff ratio, slope of flow duration curve, and inter-annual streamflow elasticity. Sawicz et al. (2011) used four of the above signatures (baseflow index, runoff ratio, slope of flow duration curve, and inter-annual streamflow elasticity) in their catchment classification study and showed that each individual hydrologic signature explains a different aspect of the hydrologic response of a catchment.

The baseflow index (BFI) is defined as the ratio of baseflow to total streamflow of a catchment. We use the one parameter single-pass digital filter method (Arnold and Allen, 1999; Eckhardt, 2008) to calculate the BFI. The baseflow filter is applied on daily streamflow time-series through the following equation:

$$B_k = \alpha \times B_{k-1} + \frac{1-\alpha}{2} \times (Q_k + Q_{k-1}) \tag{A1}$$

where, B is the baseflow and Q is the total streamflow. The values of filter parameter $\alpha = 0.925$. Equation (4) is applicable provided that $B_k \le Q_k$ (or else $B_k = Q_k$). After

HESSD
8, 9323–9355, 2011

applying the above filter, the baseflow index is calculated as:

$$\mathrm{BFI} = \sum_{k=1}^{N} \frac{B_k}{Q_k} \tag{A2}$$

A high value of BFI suggests that the influence of subsurface flow on the overall flow output from a catchment is higher. On the other hand, a low BFI value suggests that the catchment is fast responding.

The runoff ratio (RR) is defined as the ratio of average annual streamflow (Q) to average annual precipitation (P). We consider the annual average values of Q and P over the entire period of WY 1970-1988 to calculate the RR values. The runoff ratio is a metric for partitioning the incoming precipitation input into the fraction that exits the catchment as runoff and the fraction that exits the catchment as evapotranspiration (Sankarasubramanian et al., 2001; Yadav et al., 2007). Catchments with high RR value are considered to be streamflow dominated, while those with low RR values are evapotranspiration dominated.

The baseflow runoff ratio is the ratio of average annual baseflow and precipitation. It is a similar metric to runoff ratio, but gives a direct estimate of the proportion of incoming rainfall that reaches the catchment outlet through slower subsurface paths. The baseflow runoff ratio is calculated as the product of baseflow index and runoff ratio of a catchment.

The flow duration curve (FDC) of a catchment is a graphical illustration of the amount of time (expressed as a percentage) a specific streamflow value is equaled or exceeded in a catchment within a specified period of hydrologic record (Searcy, 1959; Smakhtin, 2001). The slope of flow duration curve (S_{FDC}) is defined as the slope of the middle section of the FDC (between 33rd and 66th percentile flows) when the curve can be considered as approximately linear (Sawicz et al., 2011; Yadav et al., 2007). S_{FDC} is calculated using the following formula:

HESSD
8, 9323–9355, 2011

$$S_{\mathrm{FDC}} = \frac{\ln(Q_{66}) - \ln(Q_{33})}{0.66 - 0.33} \tag{A3}$$

A high value of S_{FDC} indicates that the catchment is subject to high flow variability, while a low S_{FDC} values is typical of catchments with damped response behavior and stable flows.

The inter-annual streamflow elasticity (E_{QP}) is defined as the ratio of percentage change in annual streamflow and the percentage change in annual precipitation. E_{QP} is an indicator of the sensitivity of streamflow to relative changes in precipitation inputs (Sankarasubramanian et al., 2001; Sawicz et al., 2011). We calculate the E_{QP} using the following formula:

$$E_{QP} = \text{median}\left(\frac{dQ}{dP} \times \frac{P}{Q}\right) \tag{A4}$$

An E_{QP} value of 1 suggests that the relationship between precipitation change and streamflow change is linear. $E_{QP} > 1$ indicates that the catchment is elastic (or more sensitive) to precipitation change, while $E_{QP} < 1$ indicates that the catchment is inelastic.

Acknowledgements. The authors would like to thank Alex Abdelnour and Yiwei Cheng for providing comments on the early versions of the manuscript.

References

Archfield, S. A. and Vogel, R. M.: Map correlation method: Selection of a reference streamgage to estimate daily streamflow at ungaged catchments, Water Resour. Res., 46, W10513, doi:10.1029/2009wr008481, 2010.

Arnold, J. G. and Allen, P. M.: Automated Methods for Estimating Baseflow and Ground Water Recharge from Streamflow Records, J. Am. Water Resour. As., 35, 411–424, doi:10.1111/j.1752-1688.1999.tb03599.x, 1999.

Budyko, M. I.: Climate and life, Academic press, 1974.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀| |▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Daly, C., Neilson, R. P., and Phillips, D. L.: A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain, J. Appl. Meteorol., 33, 140–158, doi:10.1175/1520-0450(1994)033¡0140:ASTMFM¿2.0.CO;2, 1994.
Dunne, T. and Leopold, L. B.: Water in environmental planning, WH Freeman & Co., 1978.
Eckhardt, K.: A comparison of baseflow indices, which were calculated with seven different baseflow separation methods, J. Hydrol., 352, 168–173, doi:10.1016/j.jhydrol.2008.01.005, 2008.
Götzinger, J. and Bárdossy, A.: Comparison of four regionalisation methods for a distributed hydrological model, J. Hydrol., 333, 374–384, doi:10.1016/j.jhydrol.2006.09.008, 2007.
Hargreaves, G. H. and Samani, Z. A.: Estimating potential evapotranspiration, J. Irr. Drain. Div.-ASCE, 108, 225–230, 1982.
Hirsch, R. M. and Costa, J. E.: U.S. stream flow measurement and data dissemination improve, Eos Trans. AGU, 85, doi:10.1029/2004eo200002, 2004.
Hughes, D. A. and Smakhtin, V.: Daily flow time series patching or extension: a spatial interpolation approach based on flow duration curves, Hydrol. Sci. J., 41, 851–871, doi:10.1080/02626669609491555, 1996.
Jain, S. K. and Singh, V. P.: Water resources systems planning and management, Elsevier Science, 2003.
Kay, A. L., Jones, D. A., Crooks, S. M., Kjeldsen, T. R., and Fung, C. F.: An investigation of site-similarity approaches to generalisation of a rainfallrunoff model, Hydrol. Earth Syst. Sci., 11, 500–515, doi:10.5194/hess-11-500-2007, 2007.
Kokkonen, T. S., Jakeman, A. J., Young, P. C., and Koivusalo, H. J.: Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina, Hydrol. Process., 17, 2219–2238, doi:10.1002/hyp.1329, 2003.
Lamb, R. and Kay, A. L.: Confidence intervals for a spatially generalized, continuous simulation flood frequency model for Great Britain, Water Resour. Res., 40, W07501, doi:10.1029/2003wr002428, 2004.
Loveland, T. R., Merchant, J. W., Ohlen, D. O., and Brown, J. F.: Development of a land-cover characteristics database for the conterminous US, Photogr. Eng. Remote Sens., 57, 1453–1463, 1991.
McDonnell, J. J., Sivapalan, M., Vaché, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M. L., Selker, J., and Weiler, M.: Moving beyond heterogeneity and

HESSD
8, 9323–9355, 2011

Controls on hydrologic similarity

S. Patil and M. Stieglitz

process complexity: A new vision for watershed hydrology, Water Resour. Res., 43, W07301, doi:10.1029/2006wr005467, 2007.
Merwade, V., Olivera, F., Arabi, M., and Edleman, S.: Uncertainty in Flood Inundation Mapping: Current Issues and Future Directions, J. Hydrol. Eng., 13, 608, doi:10.1061/(ASCE)1084-0699(2008)13:7(608), 2008.
Merz, R. and Blöschl, G.: Regionalisation of catchment model parameters, J. Hydrol., 287, 95–123, doi:10.1016/j.jhydrol.2003.09.028, 2004.
Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, J. Hydrol., 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.
Oudin, L., Andréassian, V., Perrin, C., Michel, C., and Le Moine, N.: Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments, Water Resour. Res., 44, W03413, doi:10.1029/2007wr006240, 2008.
Patil, S. and Stieglitz, M.: Hydrologic similarity among catchments under variable flow conditions, Hydrol. Earth Syst. Sci., 15, 989–997, doi:10.5194/hess-15-989-2011, 2011.
Post, D. A. and Jakeman, A. J.: Relationships between catchment attributes and hydrological response characteristics in small Australian mountain ash catchments, Hydrol. Process., 10, 877–892, doi:10.1002/(SICI)1099-1085(199606)10:6¡877::AID-HYP377¿3.0.CO;2-T, 1996.
Sankarasubramanian, A., Vogel, R. M., and Limbrunner, J. F.: Climate elasticity of streamflow in the United States, Water Resour. Res., 37, 1771–1781, doi:10.1029/2000wr900330, 2001.
Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A., and Carrillo, G.: Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA, Hydrol. Earth Syst. Sci., 15, 2895–2911, doi:10.5194/hess-15-2895-2011, 2011.
Schreider, S. Y., Jakeman, A. J., and Pittock, A. B.: Modelling rainfall-runoff from large catchment to basin scale: The Goulburn Valley, Victoria, Hydrol. Process., 10, 863–876, doi:10.1002/(SICI)1099-1085(199606)10:6¡863::AID-HYP376¿3.0.CO;2-8, 1996.
Searcy, J. K.: Flow-duration curves: Water Supply Paper 1542-A, US Geological Survey, Reston, VA, 1959.
Seibert, J.: Regionalisation of parameters for a conceptual rainfall-runoff model, Agr. Forest Meteorol., 98–99, 279–293, doi:10.1016/s0168-1923(99)00105-7, 1999.
Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., and Zehe, E.: IAHS Decade on Predictions in Ungauged Basins (PUB),

2003–2012: Shaping an exciting future for the hydrological sciences, Hydrol. Sci. J., 48, 857–880, doi:10.1623/hysj.48.6.857.51421, 2003.
Skøien, J. O. and Blöschl, G.: Spatiotemporal topological kriging of runoff time series, Water Resour. Res., 43, W09419, doi:10.1029/2006wr005760, 2007.
Skøien, J. O., Merz, R., and Blöschl, G.: Top-kriging – geostatistics on stream networks, Hydrol. Earth Syst. Sci., 10, 277–287, doi:10.5194/hess-10-277-2006, 2006.
Slack, J. R., Lumb, A., and Landwehr, J. M.: Hydro-Climatic Data Network (HCDN) Streamflow Data Set, 1874–1988: USGS Water-Resources Investigations Report 93-4076, US Geological Survey, Reston, VA, 1993.
Smakhtin, V. U.: Low flow hydrology: a review, J. Hydrol., 240, 147–186, doi:10.1016/s0022-1694(00)00340-1, 2001.
Smakhtin, V. Y.: Generation of natural daily flow time-series in regulated rivers using a non-linear spatial interpolation technique, Regulated Rivers: Research & Management, 15, 311–323, doi:10.1002/(SICI)1099-1646(199907/08)15:4¡311::AID-RRR544¿3.0.CO;2-W, 1999.
Smakhtin, V. Y., Hughes, D. A., and Creuse-Naudin, E.: Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa, Hydrol. Sci. J., 42, 919–936, doi:10.1080/02626669709492088, 1997.
Spearman, C.: The proof and measurement of association between two things, Am. J. Psychol., 15, 72–101, 1904.
Tate, E., Maidment, D., Olivera, F., and Anderson, D.: Creating a Terrain Model for Floodplain Mapping, J. Hydrol. Eng., 7, 100–108 doi:10.1061/(ASCE)1084-0699(2002)7:2(100), 2002.
Vogel, R. M. and Sankarasubramanian, A.: Monthly Climate Data for Selected USGS HCDN Sites, 1951–1990, edited, Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, 2005.
Vogel, R. M., Wilson, I., and Daly, C.: Regional Regression Models of Annual Streamflow for the United States, J. Irr. Drain. Eng., 125, 148–157, doi:10.1061/(ASCE)0733-9437(1999)125:3(148), 1999.
Wagener, T., Sivapalan, M., McDonnell, J., Hooper, R., Lakshmi, V., Liang, X., and Kumar, P.: Predictions in ungauged basins as a catalyst for multidisciplinary hydrology, Eos Trans. AGU, 85, doi:10.1029/2004eo440003, 2004.
Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment Classification and Hydrologic Similarity, Geogr. Compass, 1, 901–931, doi:10.1111/j.1749-8198.2007.00039.x, 2007.
Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed

response behavior for improved predictions in ungauged basins, Adv. Water Res., 30, 1756–1774, doi:10.1016/j.advwatres.2007.01.005, 2007.
Zhang, Y. and Chiew, F. H. S.: Relative merits of different methods for runoff predictions in ungauged catchments, Water Resour. Res., 45, W07412, doi:10.1029/2008wr007504, 2009.

HESSD
8, 9323–9355, 2011

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Table 1. Correlation of catchment properties with Nash-Sutcliffe (NS) efficiency of simulation. Bold values indicate statistically significant value ($p < 0.01$).

Type	Property	Spearman rank correlation (ρ)	p-value
Physiographic	Channel slope	**0.21**	0.00
	Soil permeability	0.08	0.03
	SWHC	0.04	0.26
Hydrologic	Runoff ratio	**0.51**	0.00
	Baseflow index	0.03	0.46
	Baseflow runoff ratio	**0.46**	0.00
	Slope of FDC	**0.31**	0.00
	Streamflow elasticity	0.01	0.83

HESSD
8, 9323–9355, 2011

Controls on hydrologic similarity

S. Patil and M. Stieglitz

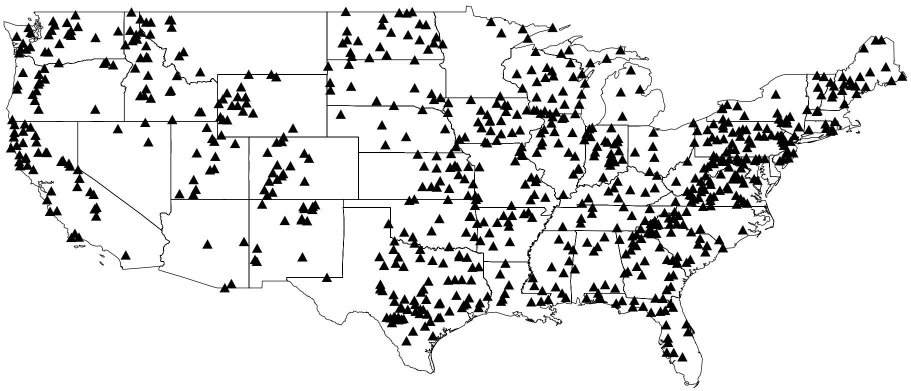

Fig. 1. Location of all the 756 catchments (black triangles) within the continental United States.

HESSD

8, 9323–9355, 2011

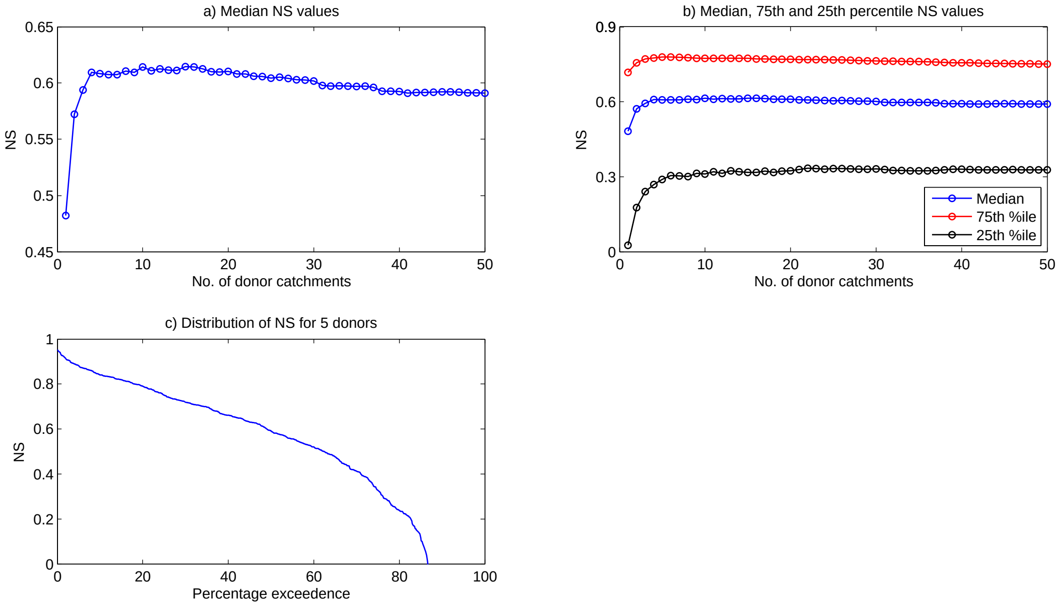

Fig. 2. (a) Relationship between number of donor catchments and median NS values, **(b)** relationship between number of donor catchments and median, 75th percentile, and 25th percentile NS values, and **(c)** distribution of NS values of simulated streamflows for the configuration of five donor catchments.

HESSD
8, 9323–9355, 2011

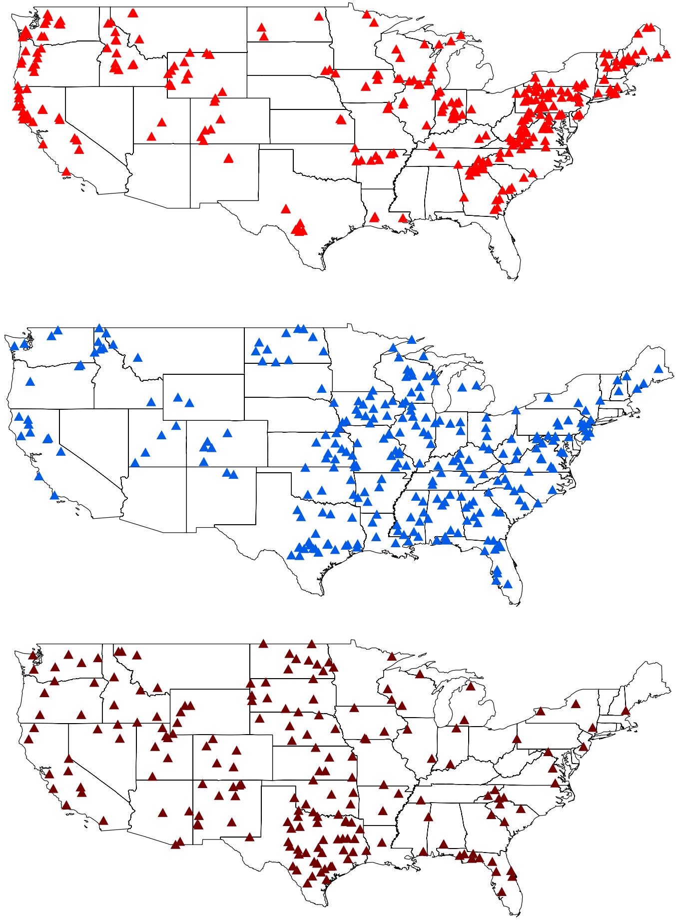

Fig. 3. **(a)** Group 1 catchments with NS > 0.7 (Red triangle), **(b)** Group 2 catchments with 0.3 < NS < 0.7 (Blue triangle), and **(c)** Group 3 catchments with NS < 0.3 (Brown triangle).

HESSD
8, 9323–9355, 2011

Controls on hydrologic similarity

S. Patil and M. Stieglitz

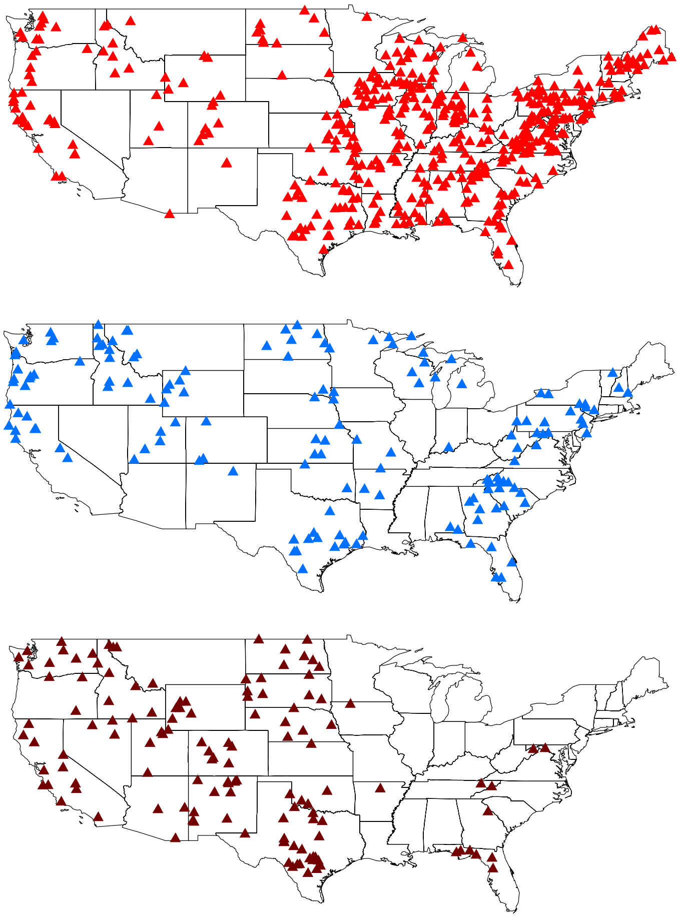

Fig. 4. **(a)** Group A catchments with WBE < 20 % (Red triangle), **(b)** Group B catchments with 20 % < WBE < 50 % (Blue triangle), and **(c)** Group C catchments with WBE > 50 % (Brown triangle).

HESSD

8, 9323–9355, 2011

Controls on hydrologic similarity

S. Patil and M. Stieglitz

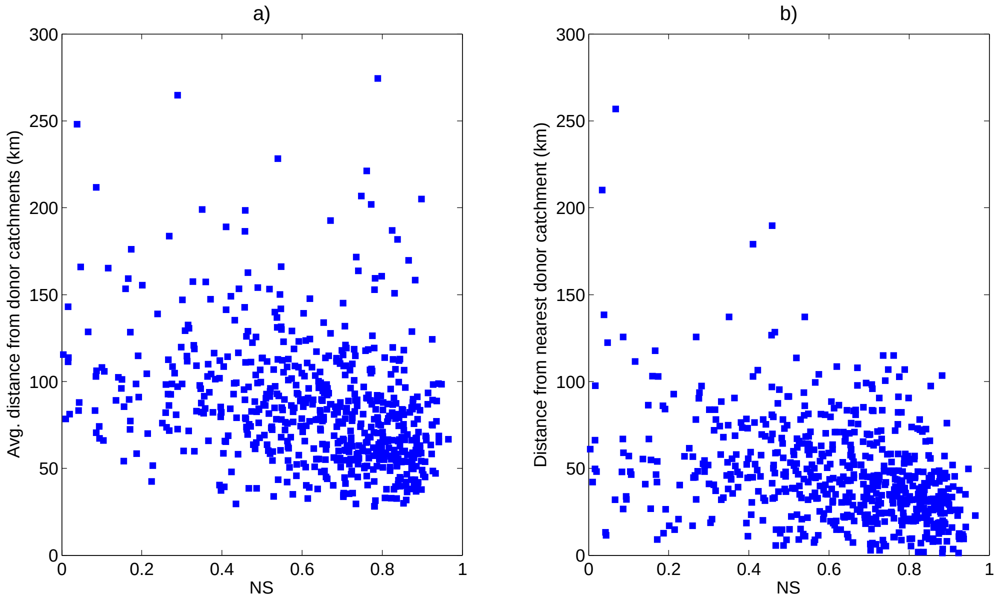

Fig. 5. Relationship of Nash Sutcliffe efficiency (NS) with **(a)** average distance from donor catchments, and **(b)** distance from nearest donor catchment.

HESSD
8, 9323–9355, 2011

Controls on hydrologic similarity

S. Patil and M. Stieglitz

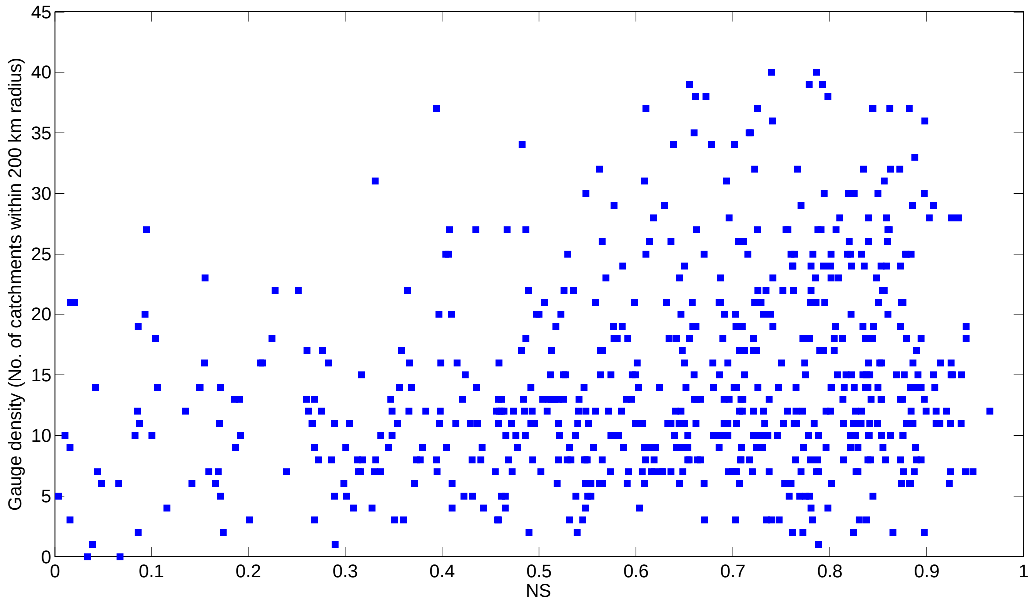

Fig. 6. Relationship of Nash Sutcliffe efficiency (NS) with gauge density around an ungauged catchment.

HESSD
8, 9323–9355, 2011

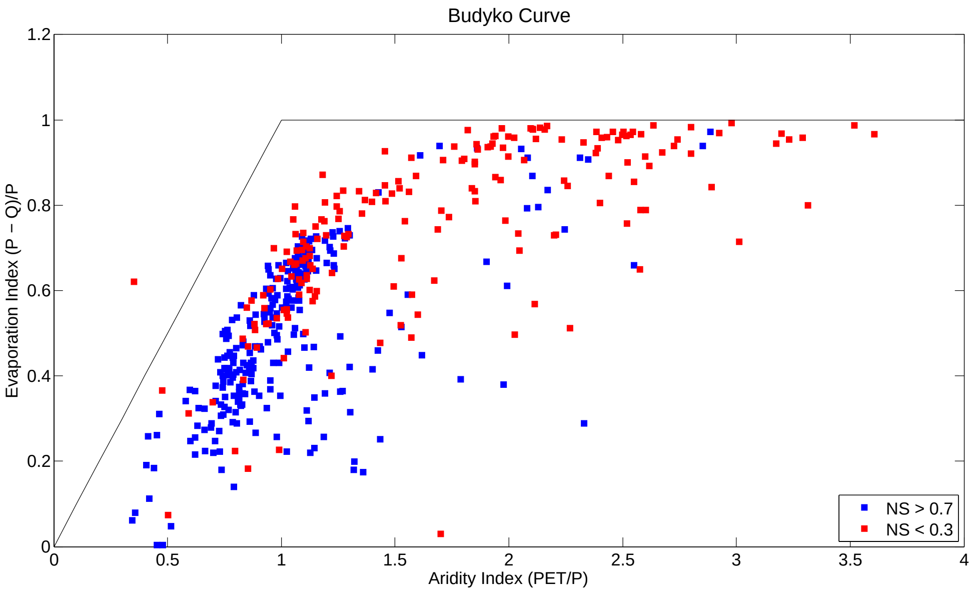

Fig. 7. Budyko diagram showing the high predictability ($NS > 0.7$) and low predictability ($NS < 0.3$) catchments.

HESSD
8, 9323–9355, 2011

Controls on hydrologic similarity

S. Patil and M. Stieglitz

HESSD
8, 9323–9355, 2011

Controls on hydrologic similarity

S. Patil and M. Stieglitz

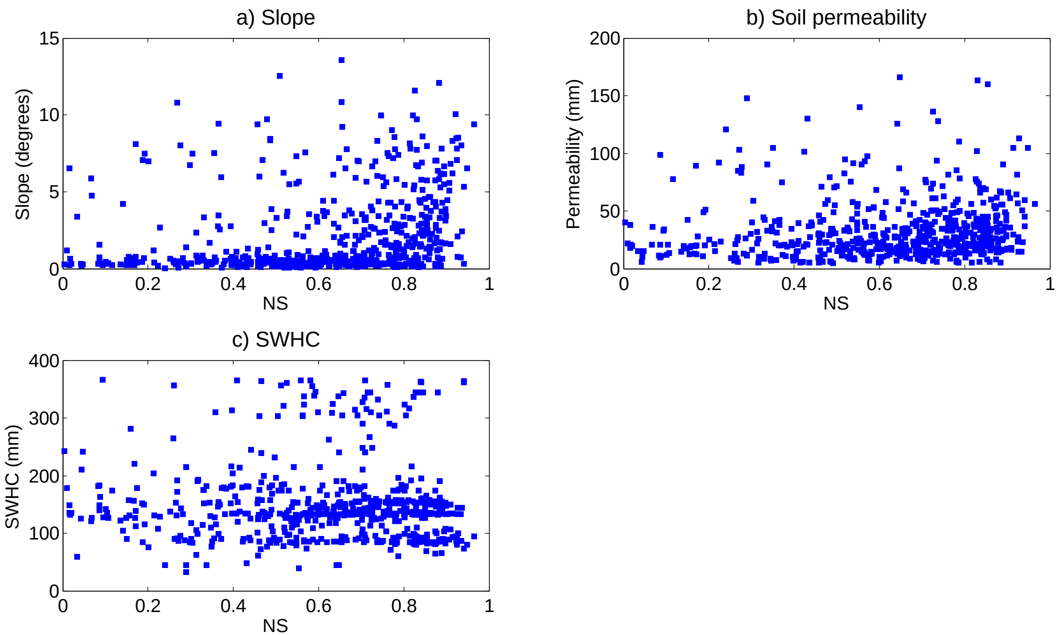

Fig. 8. Relationship between Nash Sutcliffe efficiency (NS) and **(a)** channel slope, **(b)** soil permeability, and **(c)** soil water holding capacity (SWHC).

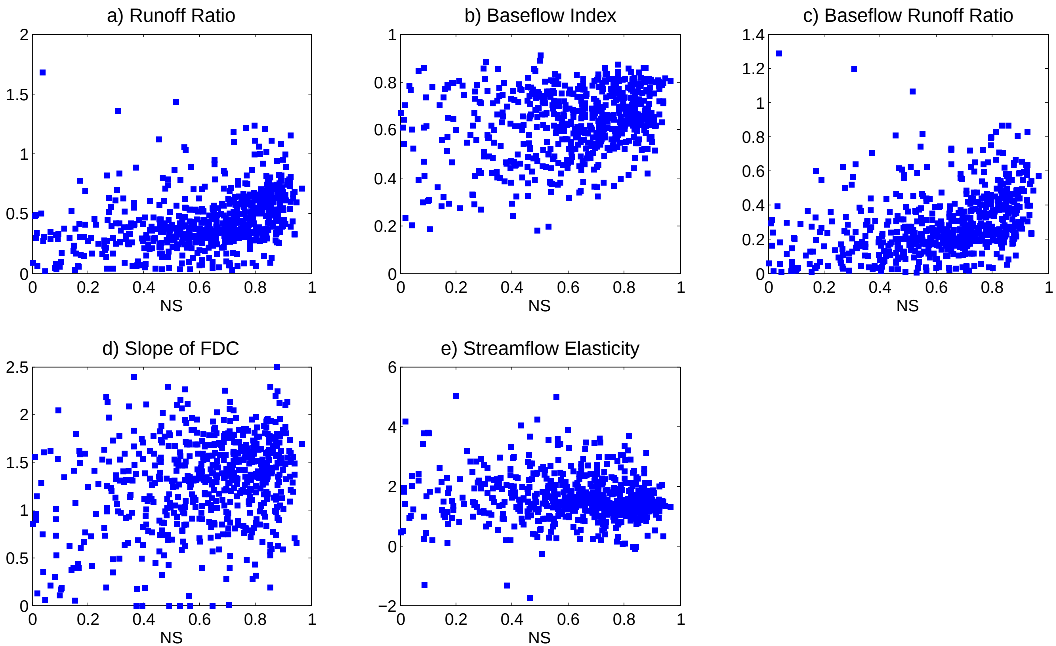

Fig. 9. Relationship between Nash Sutcliffe efficiency (NS) and **(a)** runoff ratio, **(b)** baseflow index, **(c)** baseflow runoff ratio, **(d)** slope of FDC, and e) streamflow elasticity.

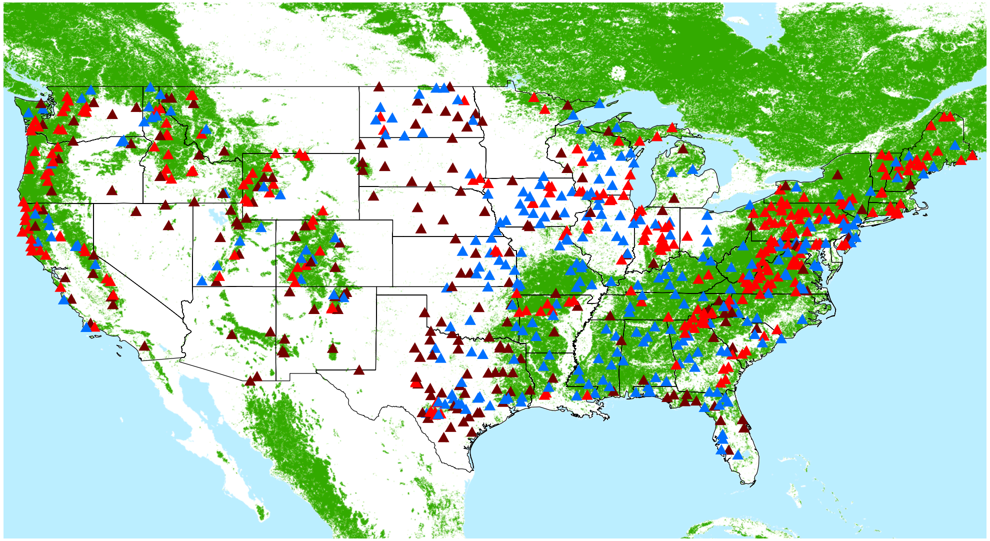

Fig. 10. All the 756 catchments mapped along with the forest cover within the United States: Group 1 (NS > 0.7; Red), Group 2 (0.7 > NS > 0.3; Blue), Group 3 (0.3 > NS; Brown).

HESSD

8, 9323–9355, 2011