Flow dynamics at the continental scale: streamflow correlation and hydrological similarity

Andrea Betterle^{1,2,3,*}, Mario Schirmer^{1,2}, Gianluca Botter³

1.EAWAG (Switzerland), 2.University of Neuchâtel (Switzerland), 3.University of Padova (Italy)

Abstract

Streamflow variability in space and time critically affects anthropic water uses and ecosystem services. Unfortunately, spatiotemporal patterns of flow regimes are often unknown, as discharge measurements are usually recorded at a limited number of hydrometric stations unevenly distributed along river networks. Advances in understanding the physical processes that control the spatial patterns of river flows are therefore necessary to predict water availability at ungauged locations or to extrapolate pointwise streamflow observations. This work explores the use of the spatial correlation of river flows as a metric to quantify the similarity between hydrological responses of two catchments. Following a stochastic framework, 340,000 cross-correlations between pairs of daily streamflows time series are predicted at a seasonal timescale across the contiguous United States using 413 catchments of the MOPEX dataset. Model predictions of streamflow correlation obtained in absence of runoff information are successfully used to identify catchment outlets sharing similar discharge dynamics and flow regimes across a broad range of geomorphoclimatic conditions, without relying on calibration. The selection of reference streamgauges based on predicted streamflow correlation generally outperforms the selection based on spatial proximity, especially as the density of available gauged sections decreases. Interestingly, correlated outlets share a broad spectrum of hydrological signatures (mean discharge, flow variability, recession properties), suggesting that catchments forced by analogous frequency and intensity of effective rainfall events might exhibit common geomorphoecological traits leading to similar hydrological responses. The proposed framework provides a physical basis to assist the regionalization of flow dynamics, and to interpret the spatial variability of flow regimes along stream networks.

Keywords: Streamflow Correlation, Hydrological Similarity, Stochastic Hydrology, Ungauged Catchments, PUB, Regionalization, Flow Patterns

*Preprint submitted to Hydrological Processes; *:corresponding author*

October 5, 2018

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', <https://doi.org/10.1002/hyp.13350>

Published version available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.13350>

Accepted for publication in *Hydrological Processes*. Copyright (2019) Wiley. Further reproduction or electronic distribution is not permitted.

1. Introduction

Streamflow dynamics and their spatial patterns along hydrographic networks critically affect anthropogenic water uses and ecosystem services worldwide [Postel and Richter, 2003; Ziv et al., 2012; Hurford and Harou, 2014; Lazzaro et al., 2017]. Understanding the physical processes that modulate the spatiotemporal patterns of river flows is crucial for infrastructure design, water resources management, hydropower production and to face extreme events like floods and droughts [Sabo et al., 2010; Widder et al., 2014]. Moreover, the environmental function of riverine ecosystems is strongly affected by streamflow dynamics, which in turn control the potential of rivers to sustain life and the biogeochemical turnover of nutrients and pollutants [Boano et al., 2014; Ceola et al., 2014]. Thus, the study of catchment-scale hydrological processes that drive spatial patterns of flow regimes has important implications to understand geomorphoclimatic legacies on physical and biogeochemical functioning of rivers.

Spatial and temporal patterns of flow regimes can hardly be obtained from direct measurements. A limited number of sites are provided with streamgauges, and flow dynamics are generally unknown at most locations along river networks [Sivapalan et al., 2003; Razavi and Coulibaly, 2013]. Especially where economical restrictions constrain the monitoring of hydrological variables, the lack of direct discharge information poses serious limitations to the optimal management of water resources and to the development of floods and droughts mitigation strategies [Hrachowitz et al., 2013].

Different approaches have been developed in the literature to cope with the need for flow regimes estimates in ungauged areas [Blöschl et al., 2013]. The concept of regionalization consists in defining suitable regions of the landscape (or group of catchments) that are expected to share similar hydrological features (e.g. discharge time series, flow statistics, recession rates, seasonality). These regions are assumed to be homogeneous in terms of the geomorphoclimatic characteristics that are deemed critical for the considered hydrological signature. In such approaches, gauged sites (i.e. locations where discharge is recorded) are necessary to identify the relationship between hydrological response and catchment attributes. This is usually done by defining empirical correlations between calibrated model parameters (or runoff signatures) and physiographic characteristics at gauged sites [Kiang et al., 2013; Merz and Blöschl, 2014; Pugliese et al., 2016]. Such relationships are then extended to catchments where the physical attributes are known but corresponding flow records are not available. Statistical correlations emerging from regression-based approaches overlook the causal link between geomorphoclimatic attributes and catchment response, thus they might be difficult to interpret in terms of hydrological functioning. Weak correlations that are frequently observed between catchment characteristics and runoff responses also suggest that critical descriptors are often missing [Oudin et al., 2007; Merz and Blöschl, 2014]. Moreover, colinearities and spurious correlations emerging from the complex nature of hydrological processes can hinder mechanistic

dependencies between the covariates, and challenge the physical interpretation of hydrological systems. In fact, regression-based methods do not allow to causally disentangle climatic from landscape contributions to the hydrological response of river basins and are problematic in case of non-stationary conditions [Müller and Thompson, 2016].

As an alternative to regression methods, the concept of “proximity” has been widely used in regionalization practices. The approach relies on the definition of distance metrics to quantify differences between geomorphoclimatic attributes of catchments. Catchments that are close to each other in the attributes space are assumed to be similar, and thus they are assigned to hydrologically homogeneous regions [Wagener et al., 2007; He et al., 2011]. The hypothesis is that similar catchments in terms of the selected set of attributes will also be similar in terms of hydrological functioning. A common approach to group catchments into homogeneous clusters (i.e. regions) is by maximizing the inter-cluster variance while minimizing the intra-cluster variability [Rao and Srinivas, 2006; Ganora et al. 2009; Rubio-Alvarez and McPhee, 2010]. Principal components analysis can be used to define new orthogonal combinations of catchment attributes, which can be employed to better describe inter-catchment heterogeneity and facilitate catchment classification [Chiang et al., 2002]. Cluster and principal components analysis, despite being powerful statistical tools, are subjected to some degree of arbitrariness and their application could be limited by practical constraints. The threshold above which inter-catchment distances in the attribute space become “too large” (thereby demanding a new region to be added) can not be objectively derived on mechanistic principles and it is often arbitrarily set. Moreover, the choice and the number of physiographic attributes that are assumed to explain the inter-catchment variability of hydrological response might be biased or simply limited by data availability [Oudin et al., 2007; Arsenault and Brisette 2016]. Regression and proximity-based methods usually provide better performances in regions where catchments are homogeneous in terms of their hydrological functioning. At larger scales, spatiotemporal changes in the dominant hydrological processes driving runoff dynamics might require the use of alternative sets of physiographic descriptors for specific catchments (or during certain seasons). In these cases, the challenge of identifying consistent and representative attributes valid for the entire set of study catchments makes the application of the method problematic [Oudin et al., 2010; Arsenault and Brisette, 2016].

Regionalization methods based on geographical distance can be considered as a special case of proximity-based methods. These methods probably represent the oldest and yet the most widely used procedure to quickly and inexpensively identify hydrologically similar locations [Vandewiele et al., 1991; Blöschl, 2005; Mohamoud, 2010]. In this vein, daily streamflow time series at ungauged locations can be estimated by importing normalized streamflow records from the closest gauged catchment [Hirsch, 1979; Smakhtin, 1999]. In general, functional similarity is not necessarily entailed by spatial proximity [Ali et al., 2012] and in some cir-

cumstances nearby sites can display significant differences in terms of streamflow dynamics. Nevertheless, spatial proximity can be efficiently used as a proxy of hydrological similarity in a wide range of geomorphoclimatic conditions, as it often outperforms regionalization based on physiographic similarity or regressions [Merz and Blöschl, 2003; Oudin, 2010]. Indeed streamflows are controlled by processes that are strongly autocorrelated in space, and geographical distance can implicitly account for the (often unknown) smooth spatial variability of hydrological features and climatic forcing [Blöschl, 2005].

Autocorrelation of geomorphoclimatic variables is the foundation of geostatistical methods in hydrology. Geostatistics aim to reproduce the spatial variability of geophysical variables by accounting for their autocorrelation structure through empirical variograms [De Marsily, 1986; Dowd, 1991; Bourges et al., 2011; Ly et al., 2012]. From the first attempts to spatially extend pointwise observations using interpolation techniques with different weighting schemes (e.g. linear interpolation, inverse distances etc.), to more sophisticated unbiased methods (e.g. kriging), geostatistical techniques have provided valuable support to spatial analysis in hydrology. To better estimate the spatial distribution of flow-related variables, geostatistical techniques were additionally developed that explicitly account for the topological arrangement of catchments along a drainage network [Skøjen et al., 2006; Archfield et al., 2013; Müller and Thompson, 2015]. These methods account for the shape and structure of river systems and thus represent valuable tools for spatial analyses in hydrology.

Despite statistical and geostatistical techniques generally succeed in densely monitored settings, their use in sparsely gauged areas remains problematic due to data requirements [Archfield and Vogel, 2010; Parajka et al., 2005]. Similarly to statistical approaches, geostatistical methods are sensitive to the quality of the available data and they are extremely data-intensive [Blöschl et al., 2013]. Computational requirements can further limit large-scale applications of more sophisticated geostatistical techniques [Müller and Thompson, 2015].

Physically-based classification frameworks based on similarity of hydrological functions are a key means to assess the dominant controls of water movement across the landscape [e.g. McDonnell and Woods, 2004; Doulatyari et al., 2017]. In particular, classification can help understanding catchment hydrological functioning by linking similarities of catchment responses to specific geomorphoclimatic attributes [Botter et al., 2013, Berghuijs et al., 2014]. Despite this, currently available classification frameworks fall short of providing a comprehensive picture of hydrological response patterns in relation to the physical similarities between catchment-scale processes [Hrachowitz et al., 2013].

Statistically-based prediction of streamflow correlation show potential for the identification of index streamgauges in poorly gauged locations [Yuang, 2010; Archfield and Vogel, 2013]. However, the few studies that consider correlation as an index of hydrological similarity focus on specific regions. Additionally, the statisti-

cal nature of these methods prevents a mechanistic interpretation of the processes underlying spatial patterns of streamflow dynamics. On the other hand, mechanistic approaches to catchment response can reduce the gap in understanding the link between river dynamics and hydrological forcings [Duan *et al.*, 2006; Doulatyari *et al.*, 2015]. Therefore, the development of physically based methods can be useful to improve model predictions, especially in data scarce settings or in cases where strong geomorphoclimatic gradients are observed.

This study explores on a large scale the use of cross-correlation between pairs of streamflow time series as a metric of hydrological similarity. A recently developed physically-based model that accounts for the probabilistic nature of joint streamflow dynamics at two river sections [Betterle *et al.*, 2017a; Betterle *et al.*, 2017b] is used to estimate 340,000 inter-catchment seasonal streamflow correlations at 413 catchments of the MOPEX dataset. The model relies on a limited number of parameters and on a set of simple mechanistic hypothesis concerning catchment responses to link the spatial variability of streamflow dynamics to the underlying heterogeneity of catchment-scale hydrological and climatic drivers. Specifically, the model application presented in this paper is used to address the following research questions:

- Can model predictions of streamflow correlation across strong physiographic gradients help us understanding catchment hydrological functioning?
- Are intra-annual dynamics of streamflow correlation relevant? Can they be explicitly accounted for by considering seasonally variable meteorological inputs?
- Can the information embedded in the correlation between river flows be efficiently used to identify catchments with similar hydrographs? What are the emerging properties shared by catchments that experience analogous hydroclimatic forcings that make them hydrologically similar?
- How does correlation perform with respect to spatial proximity in the identification of reference streamgauges (i.e. stations where discharge data are available that can be associated to ungauged locations)? What is the effect of streamgauges density on the performance of distance-based and correlation-based methods in the selection of reference streamgauges?

The reminder of this paper is organized as follows. Section 2 and 3 introduce the analytical model and two alternative methods to estimate model parameters. The study sites are presented in section 4. The performance metrics introduced in Section 5 will be used in Section 6 to evaluate the model predictions, and in Section 7 to explore the relationship between streamflow correlation and hydrological similarity. Section 8 shows how model estimates of streamflow correlation compare to spatial proximity in the identification of sites having analogous streamflow dynamics. Section 9 analyses streamgauge density as a critical constraint to identify

catchments sharing similar hydrologic responses. Discussion and conclusion close the paper.

2. Methods

This study takes advantage of a parsimonious probabilistic description of joint streamflow dynamics at arbitrary pairs of catchment outlets. The method, which is designed to characterize statistically the spatial variability of river flows, relies on a stochastic representation of specific (i.e. *mm/day*) streamflow dynamics [Botter *et al.*, 2007; Doulatyari *et al.*, 2017]. The Pearson correlation between synchronous daily streamflow time series at two river sites is expressed as a function of a limited number of hydroclimatic parameters [Betterle *et al.*, 2017a, 2017b]. The method is suited to describe the hydrological response of catchments whose streamflow dynamics are directly driven by intermittent precipitation and where the effects of water storages (snow, lakes), anthropogenic activities or inter-seasonal carryover effects can be neglected. The model also assumes that streamflow propagation time in the channel network is negligible compared to the hillslope response time, an assumption best suited for catchments with sizes up to about $10,000 \text{ km}^2$ [Robinson and Sivapalan, 1997; Botter and Rinaldo, 2003]. Additional assumptions are poissonian (i.e. non-autocorrelated) rainfall and linear catchment response.

The seasonal steady-state correlation between daily streamflow time series is analytically quantified as a function of a synthetic set of catchment-scale parameters describing the physical processes responsible for the joint streamflow dynamics at two arbitrary stations within a given region. Seasonal streamflow dynamics at a station are modeled as a function of the following parameters: i) the average catchment-scale frequency of effective rainfall events (i.e. streamflow-generating rainfall events); ii) the average catchment-scale intensity of effective rainfall events, and iii) the characteristic response time of the upstream contributing catchment. Specifically, at a catchment outlet streamflow dynamics are represented as a sequence of streamflow jumps $\Delta q(t)$ — triggered by effective rainfall events $h(t)$ — followed by exponential recessions with rate k ($dq(t)/dt = -kq(t)$). Streamflow jumps are proportional to the effective rainfall intensities and are modeled according to a Poisson process of frequency λ_t . The intensities of effective rainfall events are described by an exponentially distributed random variable of mean α_t [Laio *et al.*, 2001]. This formulation implicitly includes the interplay of stochastic rainfall and catchment-scale soil moisture dynamics [Porporato *et al.*, 2004; Botter *et al.*, 2007].

The analytical characterization of streamflow spatial correlation requires the joint streamflow dynamics at two catchment outlets to be specified. When a generic pair of outlets is concerned, effective rainfall events can be divided into two independent classes: joint and disjoint [Betterle *et al.*, 2017a, 2017b]. Joint events produce a simultaneous streamflow increment in the hydrographs of the two sites. On the

206 contrary, disjoint events produce a daily flow increment in only one of the two out-
 207 lets. At each outlet, the complete sequence of effective rainfall events can thus be
 208 decomposed in two sequences, including either joint or disjoint events. Joint and
 209 disjoint effective rainfall events are described as independent Poisson processes of
 210 frequencies λ_{12} and λ_i ($i = 1, 2$), whereas their intensities are described by expo-
 211 nentially distributed rainfall depths with means α_i^{12} and α_i , respectively. It follows
 212 that:

$$\lambda_{it} = \lambda_{12} + \lambda_i \quad (1)$$

$$\langle q_i \rangle = \alpha_{it} \lambda_{it} = \alpha_i \lambda_i + \alpha_i^{12} \lambda_{12} \quad (2)$$

213 where the subscript i identifies one of the two sites, 12 refers to joint events, t
 214 denotes the total sequence of events (joint and disjoint) and $\langle q_i \rangle$ (mm/day) is the
 215 mean daily specific discharge during the considered season. The joint streamflow
 216 process at the two sites is described in probabilistic terms by the Master Equation
 217 for the joint probability density function (PDF) of q_1 and q_2 . From the steady-
 218 state solution of the Master Equation the following analytical expressions for the
 219 streamflow correlation are obtained [Betterle *et al.*, 2017a,]:

$$\rho_{model}^{(1)} = \frac{\lambda_{12}}{\sqrt{\lambda_{1t}\lambda_{2t}}} \frac{1}{2} (1 + r_\alpha) \frac{2\sqrt{k_1 k_2}}{k_1 + k_2} \quad (3)$$

$$\rho_{model}^{(2)} = \frac{\lambda_{12} \alpha_1^{12} \alpha_2^{12}}{\sqrt{[\lambda_1 (\alpha_1)^2 + \lambda_{12} (\alpha_1^{12})^2] [\lambda_2 (\alpha_2)^2 + \lambda_{12} (\alpha_2^{12})^2]}} \frac{1}{2} (1 + r_\alpha) \frac{2\sqrt{k_1 k_2}}{k_1 + k_2} \quad (4)$$

220 In equations (3) and (4), r_α is the correlation between the intensities of joint
 221 effective rainfall events and k_1 and k_2 are the recession rates in the two catchments
 222 (for a comprehensive summary of the model parameters the reader is referred to
 223 Table 1 in Betterle *et al.*, 2017b). Equation (3) and (4) differ in the probability
 224 distributions assigned to the intensities of effective rainfall events. Equation (3)
 225 assumes that joint and disjoint events are characterized by the same exponential
 226 distribution of intensities, whereas Equation (4) assumes two different exponential
 227 distributions for the intensities of joint and disjoint events (with means α_i and α_i^{12} ,
 228 respectively). The simpler structure of equation (3) makes the formulation effective
 229 in cases where limited information is available on the considered catchments, while
 230 the greater flexibility and larger number of parameters of equation (4) makes this
 231 version of the model preferable in more controlled settings, where parameters can
 232 be effectively constrained by data. The expressions given by equations (3) and (4)
 233 quantify the cross correlation between streamflow time series as a simple function
 234 of the frequency and intensity of effective rainfall, together with the response rates

of the corresponding draining areas. The equations state that daily streamflow correlation increases as the two catchments share a relatively high number of joint effective rainfall events (relative to the total number of effective rainfall events). Furthermore, the larger and more correlated the intensities of joint events are with respect to the intensities of the disjoint events, the more correlated the streamflow time series will be. Finally, equations (3) and (4) show that flow correlation is higher when the recession rates of the two catchments are relatively homogeneous. For a comprehensive analysis of the analytical solutions and of the sensitivity of streamflow correlation to the physical processes represented by the different model parameters the reader is directed to *Betterle et al., 2017a*.

3. Estimation of model parameters

In this paper, two alternative methods have been adopted to estimate model parameters: the first is based on rainfall data only, while the second relies on discharge records. The rainfall-based method requires robust estimates of synchronous daily rainfall records in the contributing catchments. Therefore, only daily rainfall fields and catchment boundaries are required to estimate model parameters in this case. Catchment boundaries can be extracted from widespread available digital terrain models, whereas daily rainfall fields can be obtained (or computed) based on pointwise rain measurements, ground radar or satellite sensors. In the rainfall-based estimation of model parameters, the relative frequency and intensity of joint rainfall events (with respect to the total frequency and intensity of rainfall) is assumed to be the same as the relative frequency and intensity of effective rainfall (i.e. the streamflow-producing rain events). The two terms:

$$F_{\lambda}^{(1)} = \frac{\lambda_{12}}{\sqrt{\lambda_{1t}\lambda_{2t}}} \quad (5)$$

$$F_{\lambda\alpha}^{(2)} = \frac{\lambda_{12}\alpha_1^{12}\alpha_2^{12}}{\sqrt{\left[\lambda_1(\alpha_1)^2 + \lambda_{12}(\alpha_1^{12})^2\right]\left[\lambda_2(\alpha_2)^2 + \lambda_{12}(\alpha_2^{12})^2\right]}} \quad (6)$$

in equations (3) and (4) are therefore estimated based on the corresponding frequency and intensities of rainfall events. Consequently, λ_{it} , λ_{12} , λ_i are in this case the average seasonal frequency of total, joint and disjoint rainfall events (computed as a ratio between the number of rainy days and the length of the available rainfall records). Analogously, α_{it} , α_i^{12} , α_i are the mean rainfall depths of total, joint and disjoint rainfall events respectively, and r_{α} is the Pearson correlation coefficient between the joint rainfall intensities. In this study daily rainfall intensities are calculated as the exceedance of a threshold (1 mm) representing canopy interception [*Lai and Katul, 2000; Laio et al., 2001; Doulatyari et al., 2017*].

The frequency of rainfall is generally higher than the frequency of effective rainfall since a fraction of the incoming rainfall is normally buffered by the soil moisture

dynamics in the root zones and does not appear as streamflow. However, the assumption behind the rainfall-based estimation of the model parameters is rooted in the idea that the relative frequency and intensity of simultaneous runoff events is mainly controlled by the joint rainfall dynamics in the two catchments [Betterle *et al.*, 2017a, 2017b]. Additionally, as reliable estimates of catchment drainage rates in absence of direct streamflow measurements are problematic [Biswal and Marani, 2014; Doulatyary *et al.*, 2015], and considering that moderate inter-catchment heterogeneity in recession properties bear a limited impact on the spatial correlation of streamflows [Bettele *et al.*, 2017a, 2017b], as a first approximation it is assumed that $2\sqrt{k_1 k_2}/(k_1 + k_2) = 1$ in equations (3) and (4). The assumption implies the recession rates to be the same in the two catchments (i.e. $k_1 = k_2$). In this case, no additional assumptions are required on recession properties because the recession-related factor disappears from equations (3) and (4) (it equals 1). Since reliable estimates of daily rainfall records are often available in most regions of the world, the model with rainfall-estimated parameters can be used to predict streamflow correlation between arbitrary pairs of sites along river networks in most settings. The method is computationally inexpensive and does not need calibration over observed discharge data.

When streamflow records are available, the frequency and intensities of effective rainfall events can be inferred from the frequency and magnitude of the flow increments observed in the hydrographs. Streamflow increments are, according to the adopted formulation, the catchment response to an effective rainfall event. The frequency of total, joint and disjoint effective rainfall events (λ_{it} , λ_{12} , λ_i) can therefore be computed based on the number of total, joint and disjoint jumps observed in synchronous daily streamflow records at the two outlets. Similarly, the average effective rainfall intensities can be evaluated from the magnitude of the corresponding streamflow increments $\Delta q(t)$. Since the model assumes exponential recessions, the depth of each effective rainfall pulse $h(t)$ can be evaluated as $h(t) = \Delta q(t)/k$ (see e.g. equation (1) in Betterle *et al.* [2017a] and equation (4) in Botter *et al.* [2007a]). Consequently, $\alpha = \langle h(t) \rangle$ can be computed as: $\langle \Delta q(t) \rangle / k$, where $\langle \cdot \rangle$ denotes the ensemble mean. The analysis is performed separately for each class of streamflow-producing events (namely total, joint and disjoint) in order to estimate the corresponding mean depths (α_{it} , α_i^{12} , α_i). Additionally, r_α is computed as the Pearson correlation coefficient between the joint effective rainfall depths in the two catchments. Finally, recession analysis is used to evaluate k_1 and k_2 . The drainage rates are estimated by fitting linear regressions between values of $dq_i(t)/dt$ and $q_i(t)$ extracted from the descending limbs of the observed hydrograph at each outlet i [Ceola *et al.*, 2010; Basso *et al.*, 2015; Dralle *et al.*, 2015].

4. Study sites and hydrological data

A detailed model testing on a statistically significant number of geomorphoclimatically heterogeneous basins is paramount for model benchmarking and for an

improved process understanding [Duan *et al.*, 2005]. In large-scale applications, insights on region-specific hydrological features are shifted to general catchment functioning, thereby allowing the identification of fundamental processes [Andreassian *et al.* 2006]. In this study, the analyses have been performed on the study catchments included in the MOPEX dataset ([Schaake *et al.*, 2006], http://www.nws.noaa.gov/ohd/mopex/mo_datasets.htm). The dataset includes 438 catchments slightly affected by anthropogenic activities and major impoundments. From the original MOPEX dataset, 23 sites have been removed due to large gaps in their streamflow and/or precipitation records. Other two sites were additionally excluded from the analyses because they were found to include large artificial reservoirs that were constructed during the period of record. The remaining 413 study sites feature up to 56 years of synchronous precipitation and streamflow records (period: 1948-2003). Continuous daily precipitation and streamflow time series are available for all stations during the entire study period, with a few minor gaps in the flow data (quartiles of the percentage of missing flow data are [2.2 2.3 12.2]%). The daily streamflow records included in the MOPEX are provided by the national USGS streamflow gauging network (<https://waterdata.usgs.gov/nwis/rt>), whereas spatially averaged daily precipitation measurements are obtained combining observations from the National Climate Data Center (<http://www.ncdc.noaa.gov/>) and from the Natural Resources Conservation Service SNOTEL network (<https://www.wcc.nrcs.usda.gov/snow/>) with physiographic monthly precipitation fields derived from the PRISM model [Daly, 2008]. The quality of precipitation data is critical for parameter estimation in most hydrological models [Duan *et al.*, 2005]. The MOPEX sites fulfill an empirical criterion prescribing a minimum density of rainfall gauges within each catchment, which ensures reliable estimates of spatially averaged daily precipitations [Schaake *et al.*, 2000].

The maps in Figure 1 show the distribution of the 413 study sites, which span across the entire contiguous USA and feature a wide variety of sizes, morphologies and geomorphoclimatic conditions. Across the study area, complex climatic patterns emerge from the spatial variability of precipitation and evapotranspiration conditions, which are in turn affected by the strong elevation gradients. Inter-seasonal differences of wetness conditions in combination with heterogeneous land covers are additionally expected to bear a significant contribution to the variability of the hydrological response of the study catchments. The boxplots in Figure 1 further highlight how the main physical characteristics of the study sites range across several orders of magnitude.

5. Performance metrics and indexes of similarity

Model performances are evaluated by comparing equations (3) and (4) with the observed correlation, ρ_{meas} . The measured Pearson correlation between the

discharge records at each couple of outlet is computed as:

$$\rho_{meas} = \frac{\sum_{j=1}^n [(q_1(j) - \langle q_1 \rangle) (q_2(j) - \langle q_2 \rangle)]}{\sqrt{\sum_{j=1}^n [q_1(j) - \langle q_1 \rangle]^2 \sum_{j=1}^n [q_2(j) - \langle q_2 \rangle]^2}} \quad (7)$$

where $q_i(j)$ is the specific discharge (mm/day) during the j -th day, and $\langle q_i \rangle$ denotes the mean seasonal discharge in catchment i .

Additionally, in order to estimate the similarity of flow regimes between pairs of outlets, two similarity indexes are used. The first index quantifies the difference between normalized flow duration curves. Flow duration curves (FDCs) represent the exceedance probability of a discharge value during a reference time period. The exceedance probability is obtained empirically by summing the number of days having discharge larger than q , divided by the total number of days considered. Seasonal FDCs effectively represent the variability of streamflow in the frequency domain and the consequent water availability during a season [Vogel and Fennessey, 1995]. Given two river sites, it is possible to quantify the differences between their seasonal flow statistics by defining a measure of distance between their observed FDCs [Ganora et al., 2009; Pugliese et al., 2014]. At a given streamflow exceedance probability P , the distance δ_{12} between two FDCs can be expressed as the difference between the corresponding normalized discharge ($q(P)/\langle q \rangle$) as:

$$\delta_{12}(P) = \left| \frac{q_1(P)}{\langle q_1 \rangle} - \frac{q_2(P)}{\langle q_2 \rangle} \right| \quad (8)$$

The smaller δ_{12} , the more similar the flows observed at the two river locations for the considered exceedance probability.

The second index, instead, is based on the concept of Nash-Sutcliffe Efficiency (NSE). The NSE is a synthetic indicator of similarity between discharge time series [Nash and Sutcliffe, 1970]. In case where normalized discharges are concerned, NSE can be expressed as:

$$NSE(q_1, q_2) = 1 - \frac{\sum_{j=1}^n \left(\frac{q_1(j)}{\langle q_1 \rangle} - \frac{q_2(j)}{\langle q_2 \rangle} \right)^2}{\sum_{j=1}^n \left(\frac{q_1(j)}{\langle q_1 \rangle} - 1 \right)^2} \quad (9)$$

Without loss of generality, in equation (9), the site 1 is chosen as reference. NSE decreases from 1 to $-\infty$ as the similarity between the streamflow time series decreases. The NSE is often used to compare predictions of rainfall-runoff models with observed streamflow records at a river section. In this study NSE is used to quantify the error associated to exporting streamflow time series from one site to another.

6. Continental-scale prediction of seasonal streamflow correlation

The analytical model is applied to reproduce the observed correlation between daily streamflow time series at the outlets of the 413 MOPEX sites. The analysis is performed for each possible pair of study sites at seasonal times-scale, resulting in more than 340,000 pairs of synchronous streamflow time series. Seasons are defined based on fixed calendar dates as follows: spring: March, April, May; summer: June, July, August; autumn: September, October, November; winter: December, January, February.

6.1. Model performance

Figure 2 compares the measured streamflow correlation versus the model predictions (equations (3) and (4)) for all the possible combinations of study sites in the 4 seasons. Case 1 (panel a)) refers to equation (3) with model parameters estimated based on rainfall records, whereas case 2 (panel b)) refers to equation (4) when model parameters are estimated based on streamflow data.

Both versions of the method capture reasonably well the observed large-scale variability of streamflow spatial correlation. As expected, the model performs better when streamflow records are used to estimate model parameters. Streamflow dynamics are the byproduct of catchment-scale soil moisture dynamics and transport processes. Therefore, estimating model parameters from discharge data allows the relevant hydrological processes responsible for runoff formation and drainage to be accounted for, and model parameters to be better constrained. On the other hand, rainfall time series provide more limited information on hydrological dynamics as they represent only the primary input of the water cycle. The additional assumptions concerning the rainfall-based estimation procedure of model parameters (see section 3) might also be responsible for the higher scattering and lower model performances in Fig 2a. In particular, overestimations of the relative frequencies and correlations of joint effective rainfall events might cause the overestimation of streamflow correlation when all parameters are estimated just relying on rainfall records [Betterle *et al.*, 2017 b]. Nevertheless, in Figure 2a the observed variability of streamflow spatial correlation is reasonably captured by the model, especially for high correlations. As discussed in the following sections, the identification of strongly correlated catchments outlets is critical for process understanding and engineering purposes. A proper characterization of common features in the hydrological responses of ungauged catchments is indeed important to identify functional similarities useful for regionalization methods [Blöschl *et al.*, 2013; Sivapalan *et al.*, 2003; Hrachowitz *et al.*, 2013]. Therefore, the reduced performance of the model to capture low or negative correlations is of minor concern for practical applications, which are typically focused on similarities of streamflow dynamics. As the paper mainly focuses on the prediction in ungauged sections, the reminder of this study will only consider equation 3 with rainfall-estimated model parameters, unless otherwise specified.

6.2. Physical drivers of streamflow correlation

The effect on streamflow correlation played by each physical process responsible for flow dynamics can be assessed taking advantage of the analytical model. In fact, the model can help disentangling landscape and climate contributions to the ensuing correlation of river flows. The analytical expression for streamflow correlation (equation (3)) is the product of three factors ($\rho_{model} = F_\lambda \cdot F_\alpha \cdot F_k$, see inset of Figure 3) that account for the frequency and intensity of effective rainfall events and for the catchment recession rates respectively [Betterle *et al.*, 2017b]. By independently considering the frequency distribution of each single factor across the study sites, it is possible to evaluate how the corresponding process affects streamflow correlation in the study region. In particular, the more and more often a factor approaches 0, the more the corresponding process is responsible for a loss of streamflow correlation.

Figure 3 shows that a clear hierarchy exists among the physical controls on streamflow correlation: $\langle F_\lambda \rangle < \langle F_\alpha \rangle < \langle F_k \rangle$ (mean values of F_λ , F_α and F_k are 0.40, 0.53 and 0.91, respectively). In most cases, significant drops of correlation are induced by the lack of synchronicity in streamflow-producing rain events (i.e. small frequencies of joint effective rainfall events compared to the overall frequency of effective rainfall events in the two sites). Inter-catchment heterogeneities in the intensities of joint events generally provide the second largest contribution to streamflow correlation losses.

Although the timing of runoff events is generally dominant over their intensities, F_λ displays a large variability compared to F_α . This can be related to the wide range of joint frequencies that are possible when considering catchments spanning such a large study area. In fact, once the correlation range of rainfall is exceeded, rainfall events are completely uncorrelated and the variability in their occurrence is maximized. Hence, very different frequencies of synchronous events can arise. On the other hand, rainfall depths can be highly correlated over larger distances, thereby reducing the spatial variability of F_α among pairs of sites.

Interestingly, despite the strong geomorphological differences between the study sites – and the consequent enhanced inter-catchment heterogeneities of recession rates – decreases of flow correlation due to recession characteristics are mostly limited. The skewness of the corresponding distribution towards lower values of F_k shows that just in a limited number of cases substantial correlation drops are directly caused by inter-catchment differences in drainage properties. The low sensitivity of F_k to small to moderate differences between k_1 and k_2 that emerges from Figure 3 explains the reasonable model performances when inter-catchment heterogeneity of recession rates is neglected (i.e. $F_k = 1$, Figure 2a).

6.3. Seasonal dynamics of streamflow correlation

The scatterplots in Figure 4 explore how seasonality of climate and hydrology modulates streamflow correlation and model performances. Model overestimations

of the observed streamflow spatial correlation are more visible in summer and autumn. The lower runoff coefficients that are generally observed in these seasons (see Figure 4) hint to a strong role of evapotranspirative fluxes, which possibly hinder the direct link between catchment responses and rainfall inputs. High evapotranspiration rates can also enhance the effect of inter-catchment heterogeneities of landscape features (e.g. vegetation, morphology, geology), which control the capacity of catchments to buffer the incoming rainfall and ultimately control the spatial patterns of hydrologic responses. On the other hand, when and where runoff coefficients are higher, the tighter dependence between rainfall and streamflow dynamics makes more reliable the estimation of model parameters based exclusively on rainfall data. The arch of points deviating from the 1 : 1 line in Figure 4 corresponds to pairs of sites where the model significantly underestimates the observed streamflow correlation. This is likely the case of pairs of hydrographs that are impacted by melting of snow stored during colder seasons. In those cases, regardless of the nature of precipitation inputs, streamflows can be strongly correlated as a result of simultaneous melting triggered by large-scale seasonal climatic patterns (e.g. temperature, solar radiation etc.). Interestingly, the deviation is especially evident during summer, suggesting the effect of late snow melting in some cases. Anticorrelated streamflow time series on the other hand, are possibly related to snow melting affecting only one of the two catchments. In these cases, long recessions corresponding to dry conditions at one catchment are associated with increasing melting-driven discharges at the other outlet. Additionally, Figure 4 suggests that higher runoff coefficients increase streamflow correlation, as correlation is larger during wet seasons (i.e. winter and spring).

6.4. Streamflow correlation and geographical distance

Figure 5 shows that streamflow correlation decreases as inter-catchment distance increases. As expected, ρ reflects the autocorrelation structure of the hydrological variables underlying streamflow dynamics (e.g. climate, morphology, land cover), which vary smoothly in space. However, spatial proximity does not explicitly account for the physical process controlling the hydrological cycle and the resulting streamflow dynamics. The inset in Figure 5a shows that, despite streamflow correlation decreases – on average – as inter-catchment distance increases, pairs of sites at the same distance can span a wide range of correlations, even if they are relatively close to each other (inter-catchment distance < 50 km).

If the effect of seasonality is included in the relation between catchment distance and streamflow correlation, it can be noted how a consistent contribution to the variability observed in Figure 5a is due to unsteady correlations between river flows over the year. Significant seasonal dynamics of streamflow correlation patterns can in fact be observed in Figure 5 b) c) d) and e), where the extension and retreat of the correlation range is highlighted across seasons.

Neighboring catchments may thus be characterized by significantly heterogeneous streamflow dynamics, thereby implying that distance might not always be

a reliable index of hydrological similarity. On the other hand, physically-based methods of the type proposed in this study can be extremely useful to interpret the spatiotemporal variability of hydrological responses, especially when strong geomorphoclimatic gradients or large spatial scales are involved.

7. Streamflow correlation: a metric for hydrological similarity

Unlike inter-catchment distance, correlation is a normalized index with a limited range of variability (say $[-1, 1]$). Physically based estimates of streamflow correlation can thus provide valuable insights about streamflow similarity at different locations, and they can be used to classify and compare catchments based on their hydrological response. In this section, the relationship between streamflow correlation and inter-catchment hydrological responses are explored. In particular we aim at assessing the relationship between streamflow correlation and: i) seasonal flow duration curves ii) mean seasonal flows iii) flow variability iv) catchment response rates, and v) hydrographs. In particular, it is shown how model predictions of ρ (section 6.1) can be used, in absence of discharge data, to identify river sites characterized by similar streamflow characteristics across a wide range of geomorphoclimatic conditions.

7.1. Relationship between streamflow correlation and flow statistics

Figure 6 summarizes the distribution of the observed differences δ_{12} between all possible pairs of observed seasonal FDCs at the study sites (see equation 8). The distributions are aggregated based on decreasing values of modeled correlation. Figure 6 demonstrates that, without using discharge information, model prediction of streamflow correlation can efficiently group catchments based on similarities of their streamflow statistics. Highly correlated outlets ($\rho > 0.9$) are characterized by analogous exceedance probabilities across the entire range of flow magnitudes (i.e. small differences between the corresponding FDCs). On the other hand, pairs of outlet that are poorly correlated display significant inter-catchment heterogeneity in terms of FDCs. The histograms in Figure 6 show three examples of the typical differences between the frequency distribution of normalized streamflows for pairs of sites with decreasing values of ρ_{model} .

The seasonal FDCs considered in Figure 6 effectively represent the streamflow variability at one river site regardless of the corresponding average flow (flows are normalized by means of their seasonal average). The relationship between average seasonal flows and streamflow correlation is explored in Figure 7. Panels a) and b) compare the average seasonal discharge at all the possible pairs of study sites for different degrees of measured and modeled streamflow correlations. The Figure shows that the mean seasonal flows at two catchment outlets progressively approach as the correlation between their streamflow time series increase. As a consequence, highly correlated outlets share similar seasonal flows, and the model is able to couple

sites characterized by similar mean discharge across several orders of magnitudes ($10^{-3} < \langle q \rangle < 10^1 \text{ mm/day}$). Mathematically speaking, correlation is insensitive to shifts in the means of the covariates. The fact that highly correlated outlets have similar seasonal streamflows is far from being trivial from a purely statistical viewpoint and suggests a clear link between flow dynamics and catchment-scale water balance.

The coefficient of variation of daily flows ($CV(q) = \sqrt{\sigma^2(q)} / \langle q \rangle$) is an important statistical descriptor of the flow regime at-a-station. Based on $CV(q)$, two categories of flow regimes can be identified: erratic and persistent [Botter et al., 2013]. Erratic flow regimes characterize river reaches that often run dry, and whose streamflow dynamics are highly variable ($CV(q) > 1$). On the other hand, persistent flow regimes are typical of river sites where flows are more stable and are characterized by lower coefficients of variation ($CV(q) < 1$). Erratic flow regimes are normally associated to fast-responding catchments forced by sporadic effective rainfall, whereas catchments that slowly release consistent amounts of water stored during frequent effective rainfall events are typically persistent. Scatterplots in Figure 8 show that model estimates of streamflow correlation can be used to identify pairs of sites having analogous flow regimes (in terms of relative streamflow variability). Increasing values of ρ_{model} correspond to outlets sharing the same type of flow regime across a wide range of CVs.

Additionally, Figure 9 displays the recession rates of catchments with a different degree of correlation. High values of streamflow correlation correspond to pairs of sites sharing similar characteristic response times, k^{-1} . On the other hand, as streamflow correlation decreases, catchment response times tend to diverge.

7.2. Relationship between streamflow correlation and flow dynamics

In what follows, we investigate to what extent streamflow correlation bears significant information on the discharge dynamics of two catchments. Although highly correlated signals are not necessarily similar, the results reported in Figures 6, 7, 8 and 9 suggest that analogies in the hydrological response of catchments is intimately related to the correlation between their flows.

Figure 10 displays the performances – as quantified in terms of NSE (equation (9)) – associated to exporting seasonal streamflow time series from one site to another, plotted against the measured inter-catchment streamflow correlation. The figure shows the relationship between the two metrics, with the median NSE that linearly increases (with slope close to 2) as correlation increases. Given the good performance of the analytical model in identifying highly correlated sites (Figure 2), the model appears suited to identify pairs of outlets having similar streamflow dynamics (if $\rho_{meas} > 0.9$, then $\langle NSE \rangle > 0.8$). In particular, the function $NSE(\rho)$ has an upper parabolic limit, which highlights the intrinsic relationship between the two metrics [Gupta et al., 2009; McCuen et al., 2006; Weglarczyk, 1998]. In fact, if the residuals ($q_1 - q_2$) have zero average (non-systematically biased) and are uncorrelated with q_1 and q_2 (not conditionally biased), ρ is equivalent to $\sqrt{NSE}$. The

two metrics become strongly related as ρ increases because of the lack of systematic and conditional bias that results from the similarity of $\langle q \rangle$ and $CV(q)$ at highly correlated sites (Figures 7 and 8).

8. Prediction of the hydrologic response using correlation and distance

Reference streamgauges are typically used to export hydrological information to target ungauged locations. The identification of reference streamgauges is therefore a key step for the estimation of flow characteristics where hydrologic information is required. In this section, the model predictions of streamflow correlation obtained in absence of discharge data are used to identify highly correlated outlets that are eligible as reference streamgauges for all the MOPEX sites selected in this study. The method is compared to spatial proximity. Both methods are applied to the 413 MOPEX sites, assuming each catchment in turn to be ungauged. In our exercise, at the target site, streamflow data are assumed to be unavailable and only daily rainfall depths (spatially averaged over the upstream contributing area) are assumed to be known. Streamflow and precipitation time series are assumed to be known across the remaining sites. A leave-one-out cross-validations procedure is then performed. The model with rainfall-estimated parameters (see Section 6), is recursively used to predict the seasonal streamflow correlation between each single target site and all the remaining 412 outlets. The outlet having the highest modeled streamflow correlation with the target site is elected as the most hydrologically similar reference streamgauge. Measured correlations obtained from observed streamflow time series are also considered for the selection of reference streamgauges. Although this method would not be applicable to ungauged sites, it represents an upper limit in terms of model performance for the identification of hydrologically similar outlets. The use of measured flow correlations can help clarifying the reasons of possible erroneous identifications of optimal donor sites when using model estimates of streamflow correlation. In particular, the use of observed correlation can help assessing if poor performances of the proposed framework are related to the underlying model hypothesis and estimates of the parameters, or if they are rather a consequence of the inability of streamflow correlation to be a sound proxy for hydrological similarity. To evaluate how similar the streamflow dynamics are in the donor and target site, the corresponding normalized streamflow time series $(q_1(t)/\langle q_1 \rangle, q_2(t)/\langle q_2 \rangle)$ measured during individual seasons are compared.

8.1. Performance of correlation-based regionalization of flow dynamics

The boxplot in Figure 11 shows the NSE between the measured streamflow time series at the donor and target outlets for the three methods adopted to identify the optimal donor (i.e. maximum modeled correlation, minimum inter-catchment distance and maximum measured correlation). The NSE distribution in case of pairs of catchments featuring the highest NSE between their observed streamflow records

is additionally shown in transparency (gray quartiles). This represents the upper limit achievable in terms of regionalization performances when exporting normalized streamflow records from a gauged to an ungauged site, given the set of catchments available in the MOPEX database.

The comparison of the median values of NSE obtained with different selection criteria shows that, during all seasons, the observed streamflow dynamics in all target sites are satisfactory reproduced by the normalized streamflow time series at the most correlated outlet. The analytical model, despite its simplicity and parsimony in terms of data requirements, seems to capture the main processes controlling spatial patterns of flow dynamics across all the USA. The distance-based criterion also provides acceptable performances. Nevertheless, during all seasons performances are systematically lower than those provided by the correlation-based analytical approach. This is especially true in the wettest periods (spring and winter), whereas in summer (the driest season) the two methods provide comparable performances. As noted in section (6.3), the performances of the analytical model during summer are likely impacted by snowmelt. Snow melting processes are not included in the analytical formulation, and they might be strongly autocorrelated in space. Thus, similarity of flow regimes impacted by snow melting could be better captured by spatial proximity. Figure 11 also highlights the more skewed distribution of NSE towards lower values of performances when distance-based selection criterion is used, thereby implying that in a relevant number of cases geographical distance is outperformed by the correlation-based method.

The seasonal dynamics of streamflow correlation can be analyzed looking at the fluctuations of the NSE (shaded grey quantiles in the four boxes in Figure 11). In particular, streamflow dynamics are spatially more heterogeneous (lower median NSE) with a higher inter-catchment variability in the dry seasons (autumn and summer). During spring and winter, on the other hand, the inter catchment variability of NSE is lower, and the distribution shifts upward indicating more similar flow dynamics across space. The boxplot in Figure 11 finally shows that, when reference streamgauges are identified by maximizing the observed correlation with a target site, the resulting distribution of NSE approaches its upper limit. This suggests that the most correlated sites in the MOPEX database are also the most similar in terms of NSE. Therefore, streamflow correlation represents a sound indicator of similarity for streamflow dynamics.

8.2. When, where and how much does correlation outperforms spatial proximity for streamflow regionalization?

Figure 12 compares the correlation-based versus the distance-based methods for the selection of a reference stream gauges at each individual site. Green and red marks correspond to target sites whose reference streamgauge selected with the two methods differs. A green dot is assigned to sites having the higher NSE with the most correlated donor outlet. On the other hand, a red dot is assigned to sites having

the higher NSE with the closest outlet. The size of the mark additionally informs on the relative improvement in terms of NSE that the best performing selection criteria provides compared to the other. In the remaining cases (gray dots), the two criteria either identify the same donor site, or they are both unable to identify a suitable donor ($\text{NSE} < 0$). The analysis is performed at seasonal scale and three approaches to identify the reference streamgauge are compared: 1.) using the analytical model in absence of discharge data with parameters described as in Section 3; 2.) using the analytical model in absence of discharge data but relaxing the hypothesis of homogeneous recession rates (recession rates are assumed to be known from recession analysis (i.e. equation (3) in the full form is used); 3) using measured streamflow correlations. Figure 12 shows that, in general, the analytical model outperforms spatial proximity in identifying donor sites in the study area. In some cases, the selection of the most correlated reference streamgauge can dramatically improve the estimates of daily streamflow when compared to the selection of the nearest site (an example is shown in Figure 13). Instead, in cases where proximity outperforms modeled correlation, the increase of performance is generally limited.

9. Effect of streamgauge density on the selection of reference sites

The performances of regionalization procedures strongly depend on the density of sites where streamflow records and catchments attributes are available [Oudin *et al.*, 2008]. As expected, hydrological prediction in ungauged sites located within densely gauged areas are more robust [Blöschl *et al.*, 2013]. In these circumstances, spatial patterns of hydrological forcing can be captured more accurately and the presence of catchment sharing similar attributes is more likely. Despite the relatively large number of sites in the MOPEX dataset, catchment density is rather variable across the study area: the average distance to the closest catchment is $49 \pm 41 \text{ km}$ (median: 35, max: 218 km).

Figure 14 explores the effect of streamgauge density on the performances of the correlation-based and of the distance-based methods for the identification of reference streamgauges across the MOPEX dataset. The figure shows the distribution of the NSE between pairs of hydrographs as a function of inter-catchment distance (i.e. streamgauge density). The effect of streamgauge density on the identification of the optimal donor site is analysed by progressively excluding the closest sites to each target outlet. The distribution of NSEs obtained between pairs of outlets identified with the two criteria is plotted against their corresponding median distance while potential neighbouring donors are iteratively excluded. As expected, it is increasingly more difficult to find hydrologically similar outlets when the density of potential donor outlets decreases. However, the selection based on the model estimates of streamflow correlation outperforms spatial proximity as a criterion to identify sites sharing similar streamflow dynamics across the entire range of streamgauge density. Moreover, the gap between the performances of the two methods increases as the streamgauge density decreases.

The distance based approach relies on the hypothesis of smooth spatial variability of hydrological forcing which might become inappropriate as the inter catchment distance exceeds the integral scale of the relevant hydrological drivers. On the contrary, when explicitly considering the inter-catchment variability of fundamental hydrological drivers (e.g. rainfall), a suitable reference streamgauge can be identified even for relatively high inter-catchment distances. For example, using estimates of streamflow correlation, a 212 *km* apart pair of catchments is identified having NSE=0.84.

10. Discussion

Our results indicate that the proposed framework provides robust predictions of seasonal streamflow correlation at arbitrary locations along river networks, in absence of discharge data and without requiring calibrations (Figure 2). The only requirement is the availability of synchronous average daily rainfall records on the upstream drainage areas. The analysis shows that a parsimonious mechanistic description of basic physical processes responsible for streamflow dynamics can be used to identify similarities of river flows across a vast area spanning a wide range of physiographic conditions.

10.1. Streamflow correlation and hydrological similarity

Figure 7, 8 and 9 indicate that pairs of catchments with analogous seasonal flows, flow regimes and response characteristics, can be identified by means of model prediction of streamflow correlation. In statistical terms, correlation quantifies the synchronicity of two signals, but it does not necessarily inform about the means and variances of the corresponding variables (i.e. two variables with different means and/or variances can be perfectly correlated). However, our results indicate that, when applied to catchment functioning, the correlation between discharge time series embeds fundamental information on the similarity of the hydrological response of river basins in a broad sense. In particular, the evidence that streamflow spatial correlation can be used to identify long-term similarities in catchments response hints to the intimate relationship between hydrological processes in catchments with correlated flows.

The hydrological similarity displayed by highly correlated catchments can be interpreted based on two complementary arguments. The first refers to catchments coevolution. Catchments experiencing correlated streamflow dynamics triggered by similar hydroclimatic inputs (e.g. rainfall), might coevolve to develop analogies in a wide range of features of their hydrological response. Despite the scattering in the relationship between inter-catchment distance and streamflow correlation (Figure 5), highly correlated catchments tend to be close to each other as a consequence of the spatial autocorrelation of landscape and climatic characteristic. This might

contribute to further enhance the similarities across different time-scales of the hydrological response of catchments that additionally experience analogous rainfall forcing.

The second argument refers to potentially critical hydrological processes/forcing that jointly affect a wide set of streamflow signatures, including streamflow correlation. In fact, if correlation and other catchment response properties depend on the same processes (i.e. variables), they are likely to be intimately related. For example, frequency and intensities of effective rainfall events are a first-order driver for the catchment-scale water balance (see equation (2)) as well as for streamflow correlation (Figure 3). Catchments sharing frequent joint effective rainfall events with correlated intensities are therefore prone to experience similar seasonal flows as well as high values of streamflow correlation. Moreover, catchment recession rates are likely to depend on the interarrivals between effective precipitations [Dralle *et al.*, 2017]. Since streamflow correlation strongly depends on the frequency of effective rainfall events, relationships between streamflow dynamics and recession rates might be observed at highly correlated sites.

10.2. Streamflow correlation as a means to classify flow regimes and regionalize hydrological signatures

Because highly correlated catchments share similar hydrological responses, streamflow time series and other hydrological signatures observed at one gauged location can be exported to ungauged outlets. For example, by identifying highly correlated outlets it is possible to identify locations sharing similar flow regimes and ecohydrological features (Figure 8). River reaches characterized by persistent/erratic flow regime have peculiar ecosystem functioning as well as different hydromorphological behaviour (e.g.: erratic regimes are characterized by larger flooding potential, sediment transport capacity and biogeochemical activity [Botter *et al.*, 2008; Basso *et al.*, 2015, 2016]). More generally, the possibility to infer flow regimes, FDCs, catchment response rates, mean flows and hydrographs along river networks based on point discharge records represents a valuable opportunity for optimal infrastructure design, water resources management and ecological studies.

The increase of similarity between hydrological signatures at two locations that is observed as streamflow correlation increases (Figures 6, 7, 8, 9, 10) shows how correlation can provide a normalized classification metric for catchment functioning. Alternative metrics (e.g. distance) are dimensional, lack of a reference scale and have a potentially unlimited range of variability. Correlation can thus help quantifying streamflow similarity and provide a means to evaluate the reliability of streamflow estimation procedures at ungauged locations. In particular, the relation between streamflow correlation and NSE (Figure 10), which is especially strong for high correlations, proves that the analytical model represents a good alternative for selecting reference streamgauges in the regionalization of daily flows (Figures 11, and 12).

10.3. Inter-seasonal dynamics of flow correlation

In all seasons, modeled correlation outperforms distance-based selection criterion for donor sites. However, improvement in performances are especially visible during wet seasons (spring, winter). The model in fact takes advantage of the direct link between rainfall dynamics and streamflow response ensured by high runoff coefficients during spring and winter by explicitly accounting for the inter-catchment variability of rainfall. For example, strong precipitation gradients have been observed during spring in eastern regions of the USA [Messinger and Paybins, 2014]. In cases where relevant spatial variability in rainfall patterns are expected, the smoothness of geomorphoclimatic forcing implicitly assumed by spatial proximity can represent a strong limitation for distance-based approaches. Geographic distance additionally overlooks the variability of runoff dynamics triggered by seasonally switching precipitation mechanisms from frontal to convective. In these cases, heterogeneous catchment responses can be better captured by accounting for the spatial variability of rainfall dynamics. On the other hand, procedures based on rainfall data might be less robust in seasons where the rainfall signature on streamflow dynamics is beclouded by high soil water deficits and evapotranspiration rates. The reduced similarity of streamflow dynamics during autumn and summer (grey quantiles in Figure 11) hints to enhanced spatial heterogeneity of runoff dynamics caused by arid conditions. Lower runoff coefficient can enhance the effects of inter-catchment heterogeneities in land use/cover, and they can be responsible for the reduced improvement in the performances of the analytical model with rainfall-estimated parameters compared to spatial proximity during summer and autumn. Figure 11 additionally highlights how the most suited donor site in terms of discharge time series (i.e. the site having the highest NSE with the target outlet) can be accurately identified using measured streamflow correlations. Although the method would not be applicable in ungauged settings, it emphasizes that poorer performances of the analytical model are related to a combination of the following factors: i) the effect of snow dynamics, which violate some fundamental hypothesis of the model; ii) biased rainfall-based estimates of some model parameters, especially recession rates ($F_k \neq 1$).

The seasonal variability in the performance of the different classification methods tested in this paper (Figures 11, 12 and 14) encourages the development of dynamical approaches to regionalization problems. In particular, the choice of a static reference streamgauge overlooks seasonally varying runoff dynamics that can be crucial in some catchments during specific seasons (Figure 5). The reference streamgauge for a target location generally changes across seasons and does not necessarily coincide with the closest gauged outlet. Explicitly accounting for the intra-annual dynamics of hydrological processes is therefore paramount for process understanding and catchment classification.

10.4. Final remarks and pivotal findings

The good performances provided by the analytical model in reproducing the observed spatial correlation among the 413 MOPEX sites shows that the fundamental model assumptions hold across a wide range of geomorphoclimatic characteristics. However, in cases where consistent snowfall is expected, the model should be reformulated to explicitly account for snow accumulation and melting. Additionally, caution should be used when applying the methods presented in this study to settings where significant delays in catchment responses are expected due to relatively large time-scales of the flood wave propagation along the river channel (say, more than one day). Time lags caused by inter-catchment differences in channel response properties is not explicitly accounted for by the model and might lead to overestimated correlations between the synchronous flows. In these cases, similarities in streamflow dynamics might still exist, yet they cannot be expressed in terms of synchronous flow correlation.

Additionally, attention should be paid in defining a minimum depth for daily rainfalls in large catchments. Especially when spatially averaged precipitation fields are used to estimate catchment-scale rainfalls, large number of days having negligible rainfall depths might arise. In case of nested catchments, for example, larger frequencies of rainfall events can lead to overestimated frequencies of joint events and, ultimately, to overestimated streamflow correlations. Nevertheless, setting a threshold on rainfall depth to accounts for canopy interception significantly mitigates this effect.

In summary, the initial research questions can be addressed as follows:

- A clear hierarchy in the physical control of streamflow dynamics emerges from this study. Frequency and intensity of effective rainfall events are the main drivers of streamflow correlation, whereas only substantial inter-catchment differences in drainage rates bear a significant contribution to the spatial correlation of river flows.
- Correlation strength changes significantly among different neighboring catchments and varies across seasons. The analytical model, by explicitly considering first-order hydrological processes, allows an improved interpretation of similarities in streamflow dynamics and analogies of catchment functioning.
- Correlation of river flows embeds information on a broad spectrum of hydrological signatures: highly correlated catchments share similarities in terms of flow statistics (mean discharge and relative flow variability), catchment response rates and flow dynamics.
- Rainfall-based model predictions of streamflow correlation can be used to identify reference streamgauges more efficiently than spatial proximity. Performances further increase as the density of the available gauged sites decreases.

The good model performances and the insights obtained across a broad variety of physiographic, ecological and climatic conditions, encourage the application of the methods presented in this work to multiple settings. An improved understanding of the spatial patterns of river flows can help a more conscious management of water resources, possibly within an ecologically aware perspective.

11. Conclusions

This study provided a large-scale benchmark for the use of spatial correlation as an index of hydrological similarity. Streamflow correlation is shown to be a synthetic and effective indicator quantifying analogies between the hydrological response of two catchments in a broad sense. Correlated outlets in fact share similar hydrological signatures across a wide range of geomorphoclimatic conditions, suggesting the emergence of common hydrological responses in catchments forced by synchronous and intense joint effective rainfall events. Additionally, a framework is developed which uses model predictions of streamflow correlation – derived in absence of discharge data – to identify hydrologically similar locations. The physically-based stochastic model adopted succeeds in reproducing the observed steady-state seasonal cross-correlation between synchronous daily streamflow time series at arbitrary pairs of catchment outlets across the USA.

The possibility offered by the model to predict streamflow correlation without requiring discharge data and/or calibration is appealing for the estimate of the hydrological response in ungauged locations. In particular, the analytical model can be used as a regionalization tool in poorly gauged areas where limited geomorphoclimatic information is available. The advantage of the method is that it explicitly describes how the spatiotemporal variability of basic hydrological drivers affects flow dynamics. This is especially evident in sparsely gauged areas, where model prediction of streamflow correlation significantly outperforms spatial proximity in identifying hydrologically similar sites.

The model is computationally inexpensive due to its analytical nature. Thus, it can be applied pointwise along river networks at large spatial scales to identify gaps and redundancies in streamflow gauging networks. Overall, the model can help understanding the climatic and geomorphoecological controls on spatial patterns of flow dynamics and their contribution to the biogeochemical functioning of river networks.

Acknowledgements

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 641939. This study was also funded by the Swiss National Science Foundation (SNF, Project No.200021-149126). Additional support was provided by the Competence Center Environment and Sustainability (CCES) of the ETH domain in

899 the framework of the RECORD Catchment project. All the hydrologic data used
900 in this study are taken from the MOPEX dataset and are freely available online at:
901 <http://www.nws.noaa.gov/ohd/mopex/>. The authors thank Marc Müller for the
902 insightful comments provided to an earlier version of the manuscript.

References
References

- [1] Ali G., Tetzlaff D., Soulsby C., McDonnell J., Capell R., (2012), A comparison of similarity indices for catchment classification using a cross-regional dataset, *Advances in water resources*, doi:10.1016/j.advwatres.2012.01.008
- [2] Archfield, S., Vogel R., (2010), Map correlation method: Selection of a reference streamgage to estimate daily streamflow at ungaged catchments, *Water resources research*, doi:10.1029/2009WR008481
- [3] Archfield, S., Pugliese A., Castellarin A., Skøien J., Kiang J., (2013), Topological and canonical kriging for design flood prediction in ungauged catchments: an improvement over a traditional regional regression approach?, *Hydrol. Earth Syst. Sci.*, doi:10.5194/hess-17-1575-2013, 2013.
- [4] Arsenault R., Brisette F., (2016), Analysis of continuous streamflow regionalization methods within a virtual setting, *Hydrological Sciences Journal*, doi:10.1080/02626667.2016.1154557
- [5] Basso S., Frascati A., Marani M., Schirmer M., Botter G., (2015), Climatic and landscape controls on effective discharge, *Geophys. Res. Lett.*, 42, doi:10.1002/2015GL066014.
- [6] Basso S., Schirmer M., Botter G., (2016), A physically based analytical model of flood frequency curves, *Geophys. Res. Lett.*, doi:10.1002/2016GL069915
- [7] Berghuijs W., Sivapalan M., Woods R., Savenije H., (2014), Patterns of similarity of seasonal water balances: A window into streamflow variability over a range of time scales , *Water Resour. Res.*, 10.1002/2014WR015692
- [8] Betterle A., Schirmer M., Botter G., (2017a), Characterizing the spatial correlation of daily streamflows, *Water Resour. Res.*, 10.1002/2016WR019195
- [9] Betterle A., Radny D., Schirmer M., Botter G., (2017b), What do they have in common? Drivers of streamflow spatial correlation and prediction of flow regimes in ungauged Locations, *Water Resour. Res.*, 10.1002/2017WR021144
- [10] Biswal B., Marani M., (2014), ‘Universal’ recession curves and their geomorphological interpretation, *Advances in Water Resources*, doi:10.1016/j.advwatres.2014.01.004
- [11] Blöschl G., (2005), Rainfall-runoff modeling of ungauged catchments, *Encyclopedia of hydrological sciences*, doi:10.1002/0470848944.hsa140

- [12] Blöschl G., Sivapalan M., Wagener T., Viglione A., Savenije H., (2013), Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places and Scales, *Cambridge University Press*, ISBN: 978-1107028180
- [13] Boano F., Harvery J., Marion A., Packman A., Ravelli R., Ridolfi L., Wörman A., (2014), Hyporheic flow and transport processes: mechanisms, models, and biogeochemical implications, *Water resources research*, doi:10.1002/2012RG000417.
- [14] Bourges M., Mari J., Jeanne N., (2011), A practical review of geostatistical processing applied to geophysical data: methods and applications, *Geophysical Prospecting*, doi:10.1111/j.1365-2478.2011.00992.x
- [15] Botter G., Basso S., Rodriguez-Iturbe I., Rinaldo A., (2013), Resilience of river flow regimes, *Proc. Natl. Acad. Sci.*, doi:10.1073/pnas.1311920110
- [16] Botter G., Porporato A., Rodriguez-Iturbe I., Rinaldo A., (2007), Basin-scale soil moisture dynamics and the probabilistic characterization of carrier hydrologic flows: Slow, leaching-prone components of the hydrologic response, *Water Resources Research*, doi:10.1029/2006WR005043
- [17] Botter G., Rinaldo A., (2003), Scale effect on geomorphologic and kinematic dispersion, *Water Resources Research*, doi:10.1029/2003WR002154, 2003
- [18] Botter G., Daly E., Porporato A., Rodriguez-Iturbe I., Rinaldo A., (2008), Probabilistic dynamics of soil nitrate: Coupling of ecohydrological and biogeochemical processes, *Water Resources Research*, doi:10.1029/2007WR006108
- [19] Ceola S., Botter G., Bertuzzo E., Porporato A., Rodriguez-Iturbe I., Rinaldo A., (2010), Comparative study of ecohydrological streamflow probability distributions, *Water Resources Research*, doi:10.1029/2010WR009102
- [20] Ceola S., Bertuzzo E., Singer G., Battin T., Montanari A., Rinaldo A., (2014), Hydrologic controls on basin-scale distribution of benthic invertebrates, *Water Resources Research*, doi:10.1002/2013WR015112
- [21] Chiang S., Tsay T., Nix S., (2002), Hydrologic Regionalization of Watersheds. I: Methodology Development, *Journal of Water Resources Planning and Management*, doi:10.1061/(ASCE)0733-9496(2002)128:1(3)
- [22] Daly C., Halbleib M., Smith J., Gibson W., Doggett M., Taylor G., Curtis J., Pasteris P., (2008), Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, *International Journal of Climatology*, doi:10.1002/joc.1688.
- [23] Dowd P., (1991), A review of recent developments in geostatistics, *Computers and Geosciences*, doi:10.1016/0098-3004(91)90009-3

- [24] Doulatyari B., Betterle A., Basso S., Biswal B., Schirmer M., Botter G., (2015), Predicting streamflow distributions and flow duration curves from landscape and climate, *Advances in Water Resources*, doi:10.1016/j.advwatres.2015.06.013
- [25] Doulatyari B., Betterle A., Radny D., Alessi Celegon E., Fanton P., Schirmer M., Botter G., (2017), Patterns of streamflow regimes along the river network: The case of the Thur River, *Environmental Modelling and Software*, doi:10.1016/j.envsoft.2017.03.002
- [26] Dralle D., Karst N., Thompson S., (2015), a, b careful: The challenge of scale invariance for comparative analyses in power law models of the streamflow recession, *Geophysical Research Letters*, doi:10.1002/2015GL066007
- [27] Dralle D., Karst N., Charalampous K., Veenstra A., Thompson S., (2017), Event-scale power law recession analysis: quantifying methodological uncertainty, *Hydrol Earth Syst Sci*, doi:10.5194/hess-21-65-2017
- [28] Duan Q. et al., (2006), Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, *Journal of Hydrology*, doi:10.1016/j.jhydrol.2005.07.031
- [29] Ganora D., Claps P., Laio., Viglione A., (2009), An approach to estimate non-parametric flow duration curves in ungauged basins, *Water resources research*, doi:10.1029/2008WR007472
- [30] Gupta H., Kling H., Yilmaz K., Martinez G., (2009), Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *Journal of Hydrology*, doi:10.1016/j.jhydrol.2009.08.003
- [31] Hrachowitz M., Savenije H., Blöschl G., McDonnell J., Sivapalan M., Pomeroy J., Arheimer B., Blume T., Clark M., Ehret U., Fenicia F., Freer J., Gelfan A., Gupta H., Hughes D., Hut R., Montanari A., Pande S., Tetzlaff D., Troch P., Uhlenbrook S., Wagener T., Winsemius H., Woods R., Zehe E., Cudennec C., (2013), A decade of Predictions in Ungauged Basins (PUB) - a review, *Hydrological sciences journal*, doi:10.1080/02626667.2013.803183
- [32] He T., Bardossy A., Zehe E., (2011), A review of regionalisation for continuous streamflow simulation, *Hydrol Earth Syst Sci*, doi:10.5194/hess-15-3539-2011
- [33] Hirsch R., (1979), Evaluation of some record reconstruction techniques, *Water resources research*, 15(6), 1781–1790, ISSN 0043-1397.
- [34] Hurford A.P., Harou J.J., (2014), Balancing ecosystem services with energy and food security-assessing trade-offs from reservoir operation and irrigation investments in Kenya's Tana Basin, *Hydrol Earth Syst Sci*, doi:10.5194/hess-18-3259-2014, 2014.

- [35] Kiang, J. E., Stewart, D. W., Archfield, S. A., Osborne, E. B., Eng, K., (2013), A national streamflow network gap analysis, *US Geological Survey Scientific Investigations Report 2013-5013*, Reston, Virginia, USA
- [36] Lai, C., Katul G., Lai, (2000), The dynamic role of root-water uptake in coupling potential to actual transpiration, *Advances in Water Resources*, 10.1016/S0309-1708(99)00023-8
- [37] Laio F., Porporato A., Ridolfi L., Rodriguez-Iturbe I., (2001), Plants in water-controlled ecosystems: Active role in hydrologic processes and response to water stress, II. Probabilistic soil moisture dynamics, *Adv. Water Resour.*, doi:10.1016/S0309-1708(01)00005-7
- [38] Ly S., Charles C. Degre A., (2012), Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review, *Biotechnol. Agron. Soc. Environ*
- [39] Lazzaro G., Soulsby C., Tetzlaff D., Botter G., (2017), A probabilistic approach to quantifying hydrologic thresholds regulating migration of adult Atlantic salmon into spawning streams, *Water resources research*, doi:10.1002/2016WR019244
- [40] De Marsily, G., Quantitative hydrogeology, *Academic Press*, London, 1986.
- [41] Messinger T., Paybins K., (2014), Correlations of Daily Flows at Streamgages in and near West Virginia, 1930–2011, and Streamflow Characteristics Relevant to the Use of Index Streamgages , *U.S. Geological Survey*, Scientific Investigations Report 2014–5061
- [42] McCuen H., Knight Z., Cutter G., (2006), Evaluation of the Nash–Sutcliffe Efficiency Index, *Journal of Hydrologic Engineering* , doi:10.1061/(ASCE)1084-0699(2006)11:6(597)
- [43] McDonnell J., Woods R., (2004), On the need for catchment classification, *Journal of Hydrology*, doi: 10.1016/j.jhydrol.2004.09.003
- [44] Merz R., Blöschl G., (2004), Regionalisation of catchment model parameters, *Journal of hydrology*, doi:10.1016/j.jhydrol.2003.09.028
- [45] Mohamoud Y., (2010), Prediction of daily flow duration curves and streamflow for ungauged catchments using regional flow duration curves, *Hydrological sciences journal*, doi:10.1623/hysj.53.4.706
- [46] Müller M., Thompson S., (2015), TopREML: a topological restricted maximum likelihood approach to regionalize trended runoff signatures in stream networks, *Hydrol. Earth Syst. Sci.*, doi:10.5194/hess-19-2925-2015

- [47] Müller M., Thompson S., (2016), Comparing statistical and process-based flow duration curve models in ungauged basins and changing rain regimes, *Hydrol. Earth Syst. Sci.*, doi:10.5194/hess-20-669-2016
- [48] Nash J., Sutcliffe J., (1970), River flow forecasting through conceptual models part I – A discussion of principles, *Journal of Hydrology*, doi:10.1016/0022-1694(70)90255-6
- [49] Oudin L., Andreassian V., Perrin C., Michel C., Le Moine N., (2008), Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments, *Water resources research*, doi:10.1029/2007WR006240.
- [50] Oudin L., Kay A., Andreassian V., Perrin C., (2010), Are seemingly physically similar catchments truly hydrologically similar?, *Water resources research*, doi:10.1029/2009WR008887
- [51] Parajka J., Merz R., Blöschl G., (2005) , A comparison of regionalisation methods for catchment model parameters , *Hydrol. Earth Syst. Sci.*, doi:10.5194/hess-9-157-2005
- [52] Porporato A., Daly E., Rodriguez-Iturbe I., (2004), Soil Water Balance and Ecosystem Response to Climate Change, *The American Naturalist*, doi:10.1086/424970
- [53] Postel S., Richter B., (2003), Rivers for life: managing water for people and nature, Washington DC, *Island Press*, doi:10.1002/rra.820
- [54] Pugliese A., Farmer W., Castellarin A., Archfield S., Vogel R., (2016), Regional flow duration curves: Geostatistical techniques versus multivariate regression, *Advances in water resources*, doi:10.1016/j.advwatres.2016.06.008
- [55] Rao A., Srinivas V., (2006), Regionalization of watersheds by hybrid-cluster analysis, *Journal of hydrology*, doi:10.1016/j.jhydrol.2005.06.003
- [56] Razavi T., Coulibaly P., (2013), Streamflow prediction in ungauged basins: review of regionalization methods, *Journal of hydrologic engineering*, doi:10.1061/(ASCE)HE.1943-5584.0000690
- [57] Robinson J., Sivapalan M., (1997), Temporal scales and hydrological regimes: Implications for flood frequency scaling, *Water resources research*, doi:10.1029/97WR01964
- [58] Sabo, J.L., T. Sinha, L.C. Bowling, G.H.W. Schoups, W. W. Wallender, M.E. Campana, K.A., Cherkauer, P.L. Fuller, W.L. Graf, J.W. Hopmans, J.S. Kominsky, C. Taylor, S.W. Trimble, R.H., Webb, E.E. Wohl, (2010), Reclaiming sustainable watersheds in the Cadillac Desert, *Proc. Natl. Acad. Sci.*, doi:10.1073(50), 21263-21269.

- [59] Schaake J., Cong S., Duan Q., (2006), U.S. MOPEX Data Set, *IAHS Publication Series*
- [60] Schaake J., Duan Q., Smith M., Koren, V., (2000), Criteria to select basins for hydrologic model development and testing, *15th Conf. On Hydrology(Long Beach, California, USA, Am. Met. Soc., 10–14 January 2000)*, Paper P1.8.
- [61] Sivapalan M., Takeuchi K., Franks S., Gupta V., Karambiri H., Lakshimi V., Liang X., McDonnell J., Mendiondo E., O’Connel P., Oki T., Pomeroy J., Schertzer D., Uhlenbrook S., Zehe E., (2003), IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, *Hydrological sciences journal*, doi:10.1623/hysj.48.6.857.51421
- [62] Skøjen J.O., Merz R., Blöschl G., (2006), Top-kriging – geostatistics on stream networks, *Hydrol. Earth Syst. Sci.*, doi:10.5194/hess-10-277-2006.
- [63] Smakhtin V., (1999), Generation of natural daily flow time series in regulated rivers using a nonlinear spatial interpolation technique, *River research and applications*, doi:10.1002/(SICI)1099-1646(199907/08)15:4<311::AID-RRR544>3.0.CO;2-W
- [64] Vandewiele G., Xu C., Huybrechts W. (1991), Regionalisation of physically-based water balance models in Belgium. Application to ungauged catchments, *Water resources management*, doi: 10.1007/BF00421989
- [65] Vogel R.M., Fennessey N.M., (1995), Flow duration curves II. A review of application in water-resources planning, *Water Resour. Bull.*, 31(6), 1029 – 1039.
- [66] Wagener T., Sivapalan M., Troch P., Woods R., (2007), Catchment Classification and Hydrologic Similarity, *Geography Compass*, doi: 10.1111/j.1749-8198.2007.00039.x
- [67] Weglarczyk S., (1998), The interdependence and applicability of some statistical quality measures for hydrological models, *Journal of Hydrology*, doi:10.1016/S0022-1694(98)00094-8
- [68] Widder, S., Besemer K., Singer G., Ceola S., Bertuzzo E., Quince C., Sloan W., Rinaldo A., Battin T., (2014), Fluvial network organization imprints on microbial co-occurrence networks, *Proc. Natl. Acad. Sci.*, doi: 10.1073/pnas.1411723111.
- [69] Yuan L., (2013), Using correlation of daily flows to identify index gauges for ungauged streams, *Water resources research*, doi: 10.1002/wrcr.20070.

- 1115 [70] Ziv G., Baran E., Nam S., Rodriguez-Iturbe I., Levin S., (2012), Trading-off
1116 fish biodiversity, food security and hydropower in the Mekong river basin, *Proc*
1117 *Nat Acad Sci*, doi: 10.1073/pnas.1201423109

Figure 1: The 413 study catchments display a wide variety of geomorphoclimatic conditions. Inter catchment differences in terms of catchment sizes, shapes, topological arrangement (i.e. nested versus non-nested catchments) and morphology suggest enhanced differences in hydrological responses. The maps show: a) Distribution and delineation of the study catchments; b) Seasonality (defined as $\sum_{s=1}^4 |p_s/PET_s - p_a/PET_a| / (p_a/PET_a)$, where p_s (p_a) and PET_s (PET_a) are the average seasonal (and annual) precipitation and potential evapotranspiration depths; c) Elevation; d) Annual aridity (defined as PET_a/p_a); e) Land cover. The boxplots in panel f) summarizes the distribution of the main physical properties across the study sites.

Figure 2: Comparison between predicted and measured seasonal streamflow correlations at the outlets of each possible pair of the 413 study sites (340,312 pairs). Panel a): equation (3) with rainfall-estimated model parameters; panel b): equation (4) with discharge-estimate model parameters.

Figure 3: Distributions of the values assumed by the three factors that constitute the analytical expression for the seasonal streamflow correlation (equation (3)) across the study sites. Model parameters in F_λ and F_α are estimated based on rainfall data. Streamflow records are used in case of F_k to evaluate recession rates. Values close to 0 correspond to pairs of sites where inter-catchment heterogeneity in the physical processes represented by the corresponding factor F significantly contribute to a loss of streamflow correlation.

Figure 4: Intra-annual dynamics of runoff coefficients (showed in the boxplots) possibly affect the seasonal variability of streamflow correlation and model performances. Anticorrelated flows and underestimated correlations during summer might correspond to cases affected by relevant snow dynamics.

Figure 5: Streamflow correlation generally decreases as inter-catchment distance increases. Nevertheless, the complex geomorphoclimatic heterogeneity of the study sites and the spatial variability of hydroclimatic processes result in consistent scattering (even at a short scale). Seasonal dynamics of the correlation range between river flows introduces additional variability in the relation between distance and ρ . Distances are computed between the center of mass of the catchments.

Figure 6: Up: Distribution of the exceedence frequency of the differences between normalized streamflows at couples of sites. Sites are aggregated based on their expected streamflow correlation. As model estimates of streamflow correlation decreases, inter-catchment differences in observed streamflow distributions increase (i.e larger differences between the corresponding FDCs). Down: example of representative streamflow PDFs for decreasing values of modeled correlations. From higher to lower correlations, data correspond to the following pairs of sites (USGS id) during autumn: 3451500 vs 3448000; 01048000 vs 01197500; 03136000 vs 11497500.

Figure 7: High values of streamflow correlation characterize pairs of catchments having similar seasonal flows across several orders of magnitudes. As streamflow correlation decreases, inter-catchment differences in average flows progressively increase.

Figure 8: High values of streamflow correlation characterize pairs of catchments having similar coefficient of variation $CV(q)$ across several orders of magnitudes. (i.e. similar flow regimes).

Figure 9: High values of streamflow correlation characterize pairs of catchments having recession rates that are similar across several orders of magnitudes. As streamflow correlation decreases, inter-catchment differences in catchment response rates progressively increase.

Figure 10: Scatterplot between NSE and streamflow correlation for all possible pairs of seasonal streamflow time series obtained from the study sites. A clear relationship exists between streamflow correlation and NSE, which becomes increasingly stronger as correlation increases.

Figure 11: Seasonal comparison of the NSE between streamflows time series at target-donor sites. Each study outlet is in turn assumed as ungauged (target site) and a potential reference donor site is identified by means of different criteria: i) maximizing model prediction of streamflow correlation (green); ii) minimizing inter-catchment distance (red); iii) maximizing observed streamflow correlation (blue). When catchments are paired with the ones providing the maximum NSE, the ensuing distribution is represented by the gray shaded boxes. The labels at the bottom of the plot denote the limits of the lower whiskers. Model prediction of streamflow correlation systematically over performs geographic distance in the identification of reference streamgauges.

Figure 12: Each study site is in turn assumed as ungauged and a best reference donor streamgauge is selected. Green/red marks correspond to sites having highest NSE with the most correlated/near site. Differences in performances are evaluated in terms of the NSE between the streamflow time series at each pair of donor-target outlets. The size of the marks corresponds to the improvement in NSE provided by the specific selection criterion compared to the other. Grey marks refer to sites where, either the two methods identify the same reference streamgauge, or none of them is able to identify a suitable donor site ($NSE < 0$).

Figure 13: Comparison between the streamflow time series at a target outlet (usgs 01512500, Autumn 1957) and at two potential donor sites identified with the proximity and maximum modeled correlation criteria (usgs 01514000 and 01503000 respectively). Based on model estimates of streamflow correlation it is possible to identify a reference outlet having significantly more similar streamflow dynamics to the target outlet ($NSE_{12} = 0.62$, $NSE_{13} = -6.73$).

Figure 14: Distribution of the NSE (median, 45 and 55 quantiles) between the flows at pairs of catchments coupled based on maximum correlation (model prediction) and minimum geographic distance, as the average inter-catchment distance increases. It becomes more difficult to identify hydrologically similar locations when the density of a streamflow gauging network decreases (i.e. the average distance to the closest gauged site increases). However, model predictions of streamflow correlation systematically outperform spatial proximity in the identification of reference stream-gauge to be associated to an arbitrary ungauged location. Differences in the performances of the two methods increase as the density of the gauging network decreases.