ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE

10.1029/2021WR031751

Large Scale Evaluation of Relationships Between Hydrologic Signatures and Processes

Hilary K. McMillan[1], Sebastian J. Gnann[2], and Ryoko Araki[1]

[1]Department of Geography, San Diego State University, San Diego, CA, USA, [2]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, Germany

Key Points:

- Large sample signature distributions enabled us to put signature values into context as high or low, but differ by country
- Most signatures agreed with the processes they are supposed to represent, except for infiltration and saturation excess signatures
- We provide a table of signatures with recommendations on their reliability and use for process interpretation

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

H. K. McMillan,
hmcmillan@sdsu.edu

Citation:

McMillan, H. K., Gnann, S. J., & Araki, R. (2022). Large scale evaluation of relationships between hydrologic signatures and processes. *Water Resources Research*, *58*, e2021WR031751. https://doi.org/10.1029/2021WR031751

Received 3 DEC 2021
Accepted 11 MAY 2022

Author Contributions:

Conceptualization: Hilary K. McMillan
Data curation: Ryoko Araki
Funding acquisition: Hilary K. McMillan
Methodology: Hilary K. McMillan, Sebastian J. Gnann, Ryoko Araki
Software: Hilary K. McMillan, Sebastian J. Gnann, Ryoko Araki
Visualization: Sebastian J. Gnann
Writing – original draft: Hilary K. McMillan, Sebastian J. Gnann
Writing – review & editing: Hilary K. McMillan, Ryoko Araki

Abstract Dominant processes in a watershed are those that most strongly control hydrologic function and response. Estimating dominant processes enables hydrologists to design physically realistic streamflow generation models, design management interventions, and understand how climate and landscape features control hydrologic function. A recent approach to estimating dominant processes is through their link to hydrologic signatures, which are metrics that characterize the streamflow timeseries. Previous authors have used results from experimental watersheds to link signature values to underlying processes, but these links have not been tested on large scales. This paper fills that gap by testing signatures in large sample data sets from the U.S., Great Britain, Australia, and Brazil, and in Critical Zone Observatory (CZO) watersheds. We found that most inter-signature correlations are consistent with process interpretations, that is, signatures that are supposed to represent the same process are correlated, and most signature values are consistent with process knowledge in CZO watersheds. Some exceptions occurred, such as infiltration and saturation excess processes that were often misidentified by signatures. Signature distributions vary by country, emphasizing the importance of regional context in understanding signature-process links and in classifying signature values as “high” or “low.” Not all signatures were easily transferable from single, small watersheds to large sample studies, showing that visual or process-based assessment of signatures is important before large-scale use. We provide a summary table with information on the reliability of each signature for process identification. Overall, our results provide a reference for future studies that seek to use signatures to identify hydrological processes.

1. Introduction

1.1. Hydrologic Function and Dominant Processes

Within the hydrologic cycle, watersheds transport water from the land surface to its release as river flow, evapotranspiration (ET), or groundwater. This role is referred to as “watershed function” and can be divided into key categories, such as partitioning, storage, and release of water (Black, 1997; McDonnell & Woods, 2004; Wagener et al., 2007). For example, partitioning includes interception, infiltration, percolation, runoff, and return flow processes. Storage includes snow, unsaturated or saturated zone storage, perched or deeper aquifers, and lakes. Release of water includes ET, channel flow, and groundwater flow out of the watershed. Inherent in watershed process descriptions is the idea of “dominant processes.” Although watersheds might include a wide variety of processes under certain conditions, dominant processes are those most influential in controlling the hydrologic function and response (Grayson & Blöschl, 2001). For example, infiltration and saturation excess processes may both occur in a watershed, but the dominant process is the one that most strongly controls the magnitude and shape of the hydrograph.

There are many reasons to estimate the dominant processes in a watershed. Identifying the processes is a first step to developing models that provide physically realistic simulations (Grayson & Blöschl, 2001; Gupta et al., 2014). This is important given a new generation of hydrologic models with flexible structures that can simulate spatially variable processes, but may lack the corresponding spatial process knowledge (Clark et al., 2015). Watershed managers can apply process knowledge when designing interventions to intercept floodwater or prevent polluted runoff. More fundamentally, hydrologists seek to explain how climate, landscape, and critical zone features control watershed processes and runoff generation (Dunne, 1978; Fan et al., 2019; Sivapalan, 2006). To achieve these goals, accurate estimation of dominant processes is essential.

1.2. Estimating Dominant Processes Using Regionalization and Modeling

Two main approaches have been used to estimate dominant processes: regionalization and modeling. In the regionalization approach, knowledge is used from other similar watersheds. For example, Peschke et al. (1999) propose a regionalization method based on their experience in two experimental watersheds. After a literature review of conditions that favor the process of interest, they examined hydrograph shapes in the target basin and compared these with nearby basins. They examined potential runoff contributions from different land covers based on water balance estimates. The approach culminated in a rules-based assessment of which processes are possible given hydrologic and landscape characteristics. This method was automated and applied in a mesoscale basin by Hellie et al. (2002) to create process-oriented subdivisions. A similar, decision-tree approach was created by Scherrer and Naef (2003) to identify processes in highly instrumented plots, providing a structured method to translate hydrologic observations into dominant process identification.

In the modeling approach, dominant processes are those which show the greatest sensitivity and improvement when incorporated into a model. Each candidate process can be added into the model in turn, and the model tested for improved performance (Sivakumar, 2008). At the same time, the dimensionality of the system can be analyzed to estimate how many processes are needed, although this approach cannot identify specific processes. Where a flexible modeling framework is used, alternative process representations can be switched in and out, and the dominant processes inferred as those with the highest posterior probabilities in a Bayesian analysis (Prieto et al., 2021, 2022). Alternatively, a model which already incorporates all the candidate processes can be used in a sensitivity analysis. Markstrom et al. (2016) undertook a U.S.-wide assessment of parameter sensitivity for the PRMS model. Model parameters were grouped by process, and for each hydrological response unit, dominant processes were those with the highest sensitivity scores in their related parameters. This method was used to produce US-wide maps of process importance. Both regionalization and modeling approaches may suffer the same challenges as more conventional regionalization to select model parameter values based on landscape metrics. This method typically has weak results (Oudin et al., 2008), because processes can be highly variable even in superficially similar watersheds, and because model parameters may not represent the process they are supposed to.

1.3. Hydrologic Signatures Link to Processes

A promising approach to estimating dominant processes is through their link to hydrologic signatures. Signatures are quantitative metrics that describe statistical or dynamical features of streamflow timeseries, and are often used to assess model ability to simulate streamflow dynamics. Examples include annual flood, baseflow index, slope of the flow duration curve, and descriptors of recession shapes. Signatures are widely used in ecohydrology to summarize the flow regime, rate habitat suitability and assess hydrologic alteration (Olden & Poff, 2003; Yarnell et al., 2020). In hydrologic modeling applications, signatures can be used as a performance measure in calibration (Gupta et al., 2008; Kavetski et al., 2018) and to evaluate model structure (Hrachowitz et al., 2014). In ungauged basins, models can be calibrated against regionalized hydrologic signatures (Hrachowitz et al., 2013; Prieto et al., 2019). A recent review of hydrologic signatures and their applications is provided by (McMillan, 2021).

Some hydrologic signatures have well-understood links to processes in the upstream watershed, such as hydrograph recession shapes that can be derived from watershed storage-discharge behavior (Tallaksen, 1995). Work by Dunne (1978) and later, Kirkby (1988), discuss how climate, topography and soils control runoff generation processes, and how these processes lead to characteristic patterns of lag times and peak flows. Ecohydrology studies further demonstrate the link between watershed attributes and signatures values (Jowett & Duncan, 1990; Poff & Ward, 1989). Therefore, signature values can be used to understand the upstream watershed and assess dominant processes. Beighley et al. (2005) made a qualitative assessment of dominant processes based on total runoff ratio, hydrograph recession rate and flashiness, and change in event runoff ratio with season. They checked that the inferred processes were plausible given soil depth and impervious area, and added spatial data to understand process distribution within the watershed. The processes were then included in a watershed model. Recently, Wu et al. (2021) inferred patterns of runoff generation processes in the U.S. by using six signatures to cluster watersheds into eight classes. Signature values in each class were used to identify the dominant processes as infiltration excess, saturation excess/subsurface stormflow, lateral preferential flow, or baseflow. The classes were linked to physical watershed characteristics using random forest modeling. Both these studies rely on proposing

10.1029/2021WR031751

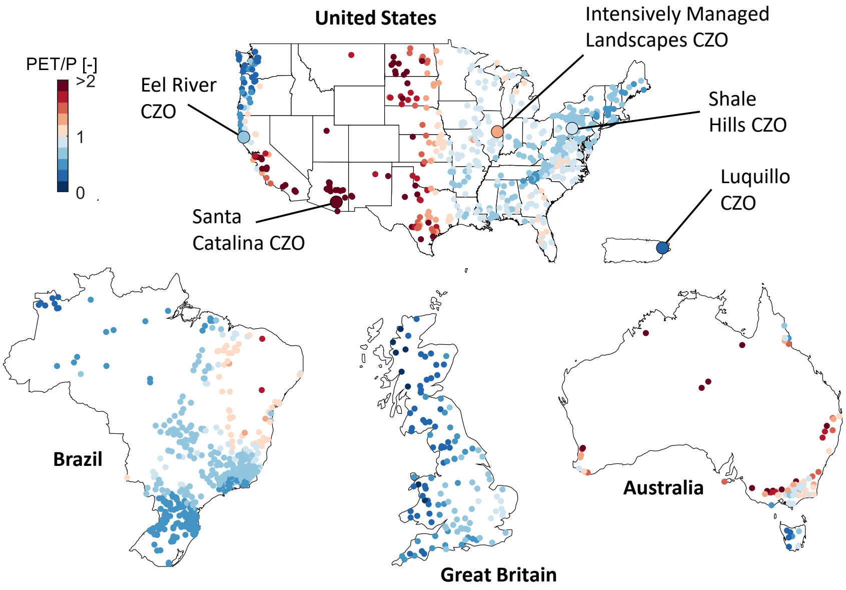

Figure 1. Map of CAMELS watersheds colored according to their aridity index (PET/P) and locations of Critical Zone Observatorys (CZOs). Note that the maps of the countries are not to the same scale. Aridity index for CAMELS watersheds is taken from Addor et al. (2017) who used Daymet data from 1989 to 2009 and the Priestly Taylor method for PET; aridity index for CZO watersheds is taken from Wlostowski et al. (2020) who used NLDAS data from 2000 to 2015 and the Penman-Monteith method for PET. Index values for CAMELS and CZO watersheds are therefore not directly comparable.

links between signatures and processes, which then enable a regionalization-type approach to estimate dominant processes.

A catalog of signature-process links was created by McMillan (2020) who collected streamflow signatures and independent information on dominant processes in papers from 45 experimental watersheds around the globe. However, these links might be specific to particular climates or hydrologic regimes. To provide a sound basis for large-scale estimation of processes based on streamflow data, we must be sure that these links are consistent across watersheds. Some evidence for consistency exists, for example, storage-related signatures (baseflow index and watershed sensitivity to runoff) were consistently linked to regolith development (weathering and creation of clay lenses) across the U.S. Critical Zone Observatory (CZO) network (Wlostowski et al., 2020). However, further evidence is required to establish consistency of interpretation for a wide range of signatures and watershed characteristics.

1.4. Aims of the Paper

The aim of this paper is to determine whether links between streamflow generation processes and streamflow signatures are consistent across a large sample of watersheds, or to determine for which signatures and processes the links hold. We focus our analysis on signatures relating to baseflow/groundwater processes and overland flow. We test the signature-process links using two types of data. Using large-sample data sets from four countries, we test whether inter-signature correlations and climate-signature correlations conform with process knowledge. Using in-depth data from five experimental watersheds in the U.S., we test whether processes inferred from signatures match with process knowledge from the literature. Understanding if and where the signature-process relationship is consistent will enable us to choose robust and reliable signatures to estimate dominant processes from large databases of streamflow data.

2. Data

We used two sources of hydrologic data: large sample data from four countries and in-depth data from CZO watersheds. Watershed locations are illustrated in Figure 1.

2.1. CAMELS Data Sets

We used several CAMELS data sets to test whether large-scale correlations and patterns in signature values conform with process knowledge. CAMELS data sets are national or continental-scale data sets of daily streamflow and forcing climate variables for watersheds, mostly with low influence from human impacts. We used CAMELS data from the U.S. (Addor et al., 2017; Newman et al., 2015), Great Britain (Coxon et al., 2020), Australia (Fowler et al., 2021) and Brazil (Chagas et al., 2020), and hourly CAMELS U.S. rainfall data from Gauch et al. (2020, 2021). The catchments used range from a size of 4.4–4,720,020 km^2, with a median size of 528.4 km^2.

We used CAMELS streamflow (Q), precipitation (P), and potential evapotranspiration (PET) data for water years 1989–2009 and only kept watersheds with at least 99% complete records. Water years are defined as starting from 1 October for the U.S. and Great Britain, 1 April for Australia, and 1 September for Brazil. A few watersheds had very small negative PET values, and those were set to zero. We removed watersheds with more than 30% of precipitation falling as snow, because event-based signatures in particular are unreliable under high snowfall conditions. Future options for including landscapes with significant snow could be to exclude snowmelt periods from the analysis (although current methods for determining spring snowmelt onset do not perform well in rivers with winter rains (Lundquist et al., 2004)); or to run a snow model to simulate soil water input, although this might lead to unwanted signature dependence on model characteristics. We also removed watersheds with significant flow regulation or diversions, based on the following criteria, noting that this had a very small impact on the results. The number of watersheds used is shown in brackets:

1. *CAMELS U.S.*: we kept all watersheds as they are near-natural (546 watersheds).
2. *CAMELS Great Britain*: we only used benchmark watersheds from the UK Benchmark Network (Harrigan et al., 2018), a subset of near-natural watersheds (120 watersheds).
3. *CAMELS Australia*: we removed watersheds with a river disturbance index >0.2 (87 watersheds).
4. *CAMELS Brazil*: we removed watersheds with consumptive_use_perc >5% and watersheds with regulation_degree >10% (486 watersheds).

2.2. CZO Data Sets

Critical zone observatories are highly instrumented watersheds that are used to study interconnected hydrological, physical, biological, and chemical processes at the Earth's surface. These observatories offer precipitation, climate and streamflow data, and extensive literature describing hydrological processes, which can be compared with processes inferred from signature values. We conducted a signature analysis at five CZO sites with a total of eight streamflow gauges: Eel River, California (Elder and Dry Creeks); Shale Hills, Pennsylvania (Shale Hills Creek); Luquillo, Puerto Rico (Rio Mameyes and Rio Icaros); Intensively Managed Landscapes (IML), Illinois/Iowa (Upper Sangamon River); Santa Catalina, Arizona (Marshall Gulch and Oracle Ridge streams).

These observatories encompass a wide range of hydrological and climatological conditions, from arid, mountainous landscapes in Arizona, to tropical forest in Puerto Rico, to a humid, steep watershed in Pennsylvania. The CZOs with paired sites offer the opportunity to compare signature values in contrasting sites under similar climate conditions. In particular, Elder and Dry Creeks, and Rios Mameyes and Icaros differ significantly in underlying geology. These five observatories were selected from the CZO network as those with less than 30% of precipitation falling as snow, matching the criterion used for the CAMELS watersheds.

For each site, raw data were processed into precipitation, streamflow, and PET time series with consistent hourly and daily timesteps, with the exception of Luquillo for which only daily data was available. Data were obtained from the Level 1 streamflow, precipitation, and meteorological data sets provided by Wlostowski et al. (2020), which comprise re-formatted versions of raw data. Data time periods varied by site, but comprised between 7 and 20 years of data during the period 1995–2017. For Luquillo, we used Level 2 precipitation data sets that had been corrected based on annual totals (Wlostowski et al., 2020). For Eel River, additional streamflow and precipitation data for the neighboring Dry Creek were provided by D. Dralle (personal commication). Where multiple precipitation gauges were available, we calculated areal averages following the site-specific methods described by Wlostowski et al. (2020). Where necessary, we used daily streamflow values from the United States Geological Survey (USGS) gauges, disaggregated using linear interpolation, to infill missing hourly data. PET values were calculated from meteorological variables using the algorithm described by Zotarelli et al. (2010).

10.1029/2021WR031751

3. Hydrologic Signatures

The signatures tested in this paper relate to baseflow/groundwater processes and overland flow (saturation and infiltration excess) and are taken from the McMillan (2020) catalog. MATLAB codes to calculate these signatures were implemented as part of the Toolbox for Streamflow Signatures in Hydrology (TOSSH; Gnann et al., 2021a). The TOSSH toolbox provides standardized methods for hydrologic signature calculations, including recommended parameter values. Minor changes from the original catalog were made to revise or remove three signatures for ease of interpretation (see Table S1 in Supporting Information S1). Full descriptions of the signatures calculated for this paper are given in Table 1 (17 signatures for groundwater/baseflow processes and 9 signatures for overland flow processes), and their MATLAB code can be found at https://tosshtoolbox.github.io/TOSSH/p2_signatures.html#process-based-signature-sets. Some of these signatures rely on common functionality, as follows. Several baseflow signatures use a function to identify recession periods; these were found as periods of decreasing streamflow (or allow increases of up to a small tolerance value), with a minimum length specified. Several overland flow signatures use a function to separate rainfall and flow series into individual events. Events are defined based on the rainfall time series, and occur when more than 2 mm/hr or 10 mm/day of precipitation fell, are separated into distinct events when 12 dry hours occur, and are deemed to end 5 days after rainfall end. Where overland flow signatures require values such as “flow volume” or “total precipitation,” these are calculated for the period of each event.

Several of the signatures require watershed-specific parameters, and their values are described in Table S2 in Supporting Information S1. We inspected TOSSH warnings to check for problems, typically due to default parameters being unsuitable for the watershed, or to data errors. Most TOSSH signatures offer a “plot_results” parameter for diagnostic graphical display of signature calculation and values. We used this option for CZOs and for 5 randomly selected CAMELS watersheds per country to visually check the signature calculations (e.g., the fitted recessions). For the CZO watersheds, we made visual checks of the baseflow separation function, that baseflow was adequately separated from quickflow during events. For the event identification function, we checked that the event periods cover major rainfall periods, and those event recession periods include flow peaks occurring immediately after rainfall. We checked that the recession identification function selected all major recession periods and adjusted recession selected parameters if required. For plots where threshold functions were fitted (e.g., to a plot of quickflow against antecedent condition metrics), we checked whether the fit was influenced by a few large or unusual rainstorms. For CZO watersheds, parameters were manually adjusted if found unsuitable, and for CAMELS watersheds—the large sample case where parameters were not individually adjusted—we retained the default parameters but discuss issues found in Section 6.2.

4. Signature Analysis

We used two approaches to test whether the links between signatures and processes described in McMillan (2020) hold true across multiple watersheds. The first approach used a large sample analysis of signature values across the CAMELS watersheds. For these data, we tested whether signatures related to the same process are correlated, and how signature values are related to climate aridity. The second approach used detailed analyses of signature values in CZO watersheds, to test whether processes inferred from signature values agree with information from watershed-specific literature.

4.1. Large-Scale Signature and Process Patterns in CAMELS Watersheds

4.1.1. Distribution of Signature Values

We applied the overland flow and groundwater signature sets across the CAMELS data sets to determine the distributions of values for each signature in each country. This enabled us to classify signature values in terms of their quantile values, that is, high or low compared to the median for their country, or quasi-globally. This information is valuable for interpretation of the signatures in new watersheds, for example, to say whether a recession constant should be considered fast or slow.

4.1.2. Correlation Between Signature Values

Several of the signatures target the same or similar processes, for example, multiple signatures indicate high water storage in the watershed (Table 1). If these signatures represent the same process, we should find correlations

Table 1
Groundwater and Overland Flow Signatures Used in This Paper

Signature	Unit	Description	Technical definition	Related process
Groundwater and baseflow signatures				
TotalRR	–	Total runoff ratio	Mean streamflow divided by mean precipitation	Evapotranspiration or other flow bypassing gauge (Safeeq & Hunsaker, 2016)
EventRR	–	Event runoff ratio (average over all events)	Average of runoff ratios (streamflow divided by precipitation) calculated for all identified events; events are defined based on the precipitation time series (detailed explanation in text)	Low ratios show rapid vertical drainage of water to groundwater (Noguchi et al., 1997)
RR_Seasonality	–	Runoff ratio seasonality (summer total RR/winter total RR)	Runoff ratio (streamflow divided by precipitation) calculated for summer months divided by runoff ratio calculated for winter months; winter is defined as the 6 months following the start of the water year	Low ratios show high bedrock permeability (Pfister et al., 2017)
StorageFraction	–	Ratio between active and total storage	Active storage divided by total storage (storage terms are explained below)	Low ratios show permeable bedrock and high total storage (Pfister et al., 2017)
ActiveStorage	mm	Active storage defined as maximum storage deficit	Storage deficit corresponding to the 99th percentile of the observed flow duration curve; storage deficit is calculated using a simple water balance model	Active watershed storage (Pfister et al., 2017)
TotalStorage	mm	Total storage calculated by extrapolation to find storage deficit at near-zero flow	Fits an envelope line (tangent to the hysteretic loop between daily values of discharge and storage deficit) and extrapolates it to nearly zero-flow conditions (0.001 mm/day)	Total watershed storage (Pfister et al., 2017)
Recession_a_Seasonality	–	Seasonal variations in recession "a" parameter, related to recession timescale	Assumes that recessions have a slope of 2 when dQ/dt is plotted against Q in log-log space, then calculates the y-intercept for all individual recession events, and returns the difference between the maximum and minimum monthly median y-intercept; recessions are defined as periods of decreasing streamflow (detailed explanation in text)	Impact of evapotranspiration on watershed storage (Shaw & Riha, 2012)
AverageStorage	mm	Average storage derived from average baseflow and storage-discharge relationship	Uses a simple water balance model to calculate changes in storage, then finds the relationship between storage and discharge, and then estimates average storage from average baseflow	Average magnitude of watershed storage (Peters & Aulenbach, 2011)
RecessionParameters_b	–	Recession analysis parameters (T0, b) approximate storage-discharge relationship. b is a shape parameter	Fits a line to the dQ/dt versus Q point cloud in log-log space for each individual recession and returns the median slope; recessions are defined as periods of decreasing streamflow (detailed explanation in text)	Storage-discharge relationship (Tallaksen, 1995)
RecessionParameters_T0	d	Characteristic timescale of recessions, at median flow	Fits a line to the dQ/dt versus Q point cloud in log-log space for each individual recession, with Q scaled by median Q. T0 is the median value of -1/intercept	Typical watershed response timescale (McMillan et al., 2014; Tallaksen, 1995)
MRC_num_segments	–	Number of different segments in nonparametric master recession curve (MRC)	Fits successively more linear segments (maximum 3) to a log-transformed MRC until the RMSE is reduced by <25% for an extra segment. MRC is derived using a new matrix-solution implementation of the adaptive matching strip method	Presence of multiple reservoirs contributing to flow (Clark et al., 2009)

Table 1
Continued

Signature	Unit	Description	Technical definition	Related process
BFI	–	Baseflow index, that is, fraction of flow classified as baseflow	Mean baseflow divided by mean streamflow; baseflow is estimated using the UKIH smoothed minima method (UKIH, 1980)	Baseflow proportion and baseflow residence time (Bulygina et al., 2009; Yilmaz et al., 2008)
BaseflowRecessionK	1/d	Exponential recession constant fitted to master recession curve (MRC)	Fits exponential function to MRC and returns the time constant; MRC is derived using the adaptive matching strip method	Low values show greater groundwater influence and longer subsurface flow paths (Safeeq et al., 2013)
First_Recession_Slope	1/d	Steep section of MRC, related to storage that is quickly depleted	Fits a straight line to the first segment of the log-transformed MRC (steep part; segments as defined by MRC_num_segments) and returns its slope	Storage near the soil surface that is quickly depleted (Estrany et al., 2010)
Mid_Recession_Slope	1/d	Mid-section of MRC, related to water retention capacity of the watershed	Fits a straight line to the second segment of the log-transformed MRC (segments as defined by MRC_num_segments) and returns its slope	Water retention capacity of the watershed (Estrany et al., 2010)
EventRR_TotalRR_ratio	–	Ratio between event and total runoff ratio	EventRR divided by TotalRR	Low event runoff coefficients and high yearly runoff coefficients show large storage capacity (Blume et al., 2008)
VariabilityIndex	–	Variability index of flow	Standard deviation of log-transformed discharge values determined at 10% intervals from 10% to 90% of the cumulative frequency distribution (flow duration curve)	Low variability index shows higher water storage (Estrany et al., 2010)
Overland flow signatures				
IE_effect	–	Infiltration excess importance	Average of the standardized z-score coefficients for mean and maximum event intensity in regression equations to predict event peak magnitude and quickflow volume	Infiltration excess occurrence and relative importance compared to saturation excess (Estrany et al., 2010)
SE_effect	–	Saturation excess importance	Average of the standardized *z*-score coefficients for total event precipitation and 3 and 7-day antecedent precipitation totals in regression equations to predict event peak magnitude and quickflow volume	Saturation excess occurrence and relative importance compared to infiltration excess (Estrany et al., 2010)
IE_thresh_signif	–	Infiltration excess threshold significance	*P*-value for the significance of a non-zero change in slope above and below a threshold in a plot of event quickflow volume versus event maximum intensity. Slopes are calculated using a “broken stick” fit to minimize squared errors	Significant values (<0.05) imply infiltration excess occurs (Ali et al., 2013)
SE_thresh_signif	–	Saturation excess threshold significance	*P*-value for the significance of a non-zero change in slope above and below a threshold in a plot of event quickflow volume versus event total precipitation. Slopes are calculated using a “broken stick” fit to minimize squared errors	Significant values (<0.05) imply saturation excess occurs (Ali et al., 2013; McGrath et al., 2007)
IE_thresh	mm/time-step	Infiltration excess threshold depth (intensity of precipitation needed to produce quickflow)	Value (location) of the threshold identified in the IE_thresh_signif signature	Rainfall intensity required to generate infiltration excess (Ali et al., 2013)
SE_thresh	mm	Saturation excess threshold location (depth of precipitation needed to produce quickflow)	Value (location) of the threshold identified in the SE_thresh_signif signature	Event precipitation depth required to generate saturation excess (Ali et al., 2013; McGrath et al., 2007)

Table 1
Continued

Signature	Unit	Description	Technical definition	Related process
SE_Slope	mm/mm	Above-threshold slope in a plot of quickflow volume versus total precipitation	The above-threshold slope of the broken-stick fit in the SE_thresh_signif signature	Rate at which saturated areas expand with additional rainfall (Tani, 1997)
Storage_thresh_signif	–	Storage threshold significance	*P*-value for the significance of a non-zero change in slope above and below a threshold in a plot of event quickflow volume versus event total precipitation + antecedent precipitation index. Slopes are calculated using a "broken stick" fit to minimize squared errors	Significant values (<0.05) imply saturation excess occurs (Ali et al., 2013; McGrath et al., 2007)
Storage_thresh	mm	Storage threshold location (storage depth needed to produce quickflow)	Value (location) of the threshold identified in the Storage_thresh_signif signature	Storage depth proxy (API + event precipitation depth) required to generate saturation excess (Ali et al., 2013; McGrath et al., 2007)

Note. Technical information on how signatures are calculated is drawn from Gnann et al. (2021a). Information on the processes related to each signature, and reference(s) establishing those interpretations, are drawn from McMillan (2020).

between their values. We therefore created a correlation matrix showing rank correlations between each pair of signatures. Spearman rank correlation was used as a nonparametric correlation measure, as relationships between signatures may not be linear. We assessed whether signatures that represent the same or similar processes have high correlations (where we used a subjective threshold of 0.7 or greater to indicate high correlation).

4.1.3. Hydro-Climate Relationship to Signature Patterns

Climate is a strong control on many signatures (Knoben et al., 2018). In particular, the aridity index (PET/P) has shown strong (empirical) links to many signatures (Addor et al., 2018). We therefore investigated to what extent aridity explains observed signature patterns and whether these patterns are consistent across different countries. We calculated rank correlations and plotted signature values against the aridity index, separated by country. The results will serve as a first assessment of similarity in signature controls across countries, therefore showing how transferable our results might be.

4.2. Signature-Process Links at Critical Zone Observatories

We based our analysis of signature-process links in CZOs on the summary findings of McMillan (2020; their Table 1), with an overview in Table 1, this paper. We reorganized their findings into a series of questions about processes in the watershed that could potentially be answered from literature descriptions, and matched the signature values that relate to each question, for example, "*Do riparian zones contribute to flow?*" a positive answer implies that there is no rainfall depth threshold before flow occurs, that is, *SE_thresh* is close to 0 and/or *SE_thresh_signif* >0.05 (see Results in Section 5.2 for the full list of questions, corresponding signature values, and answers). Restructuring the analysis in this way allowed for multiple signatures relating to one process. For each observatory, we collected journal articles describing the watershed processes, and used these to answer the questions. In several cases, the observatories included contrasting sub-watersheds, and process information was collected about each one.

We calculated signature values for all signatures described in Section 3. Where hourly data were available (i.e., all CZOs except Luquillo), we calculated signatures at both hourly and daily timesteps. We compared the impact of timestep choice and noted cases where signature values depend strongly on timestep. We used the distributions of signature values described in Section 4.1.1 to assign a percentile to each value, within the distribution of values across the CAMELS U.S. data set. We chose to use only the U.S. data set to quantify percentiles, as we assume that descriptions of processes as being of high or low importance are most likely to implicitly imply a comparison against other U.S. watersheds.

We then placed each question-signature pair into one of three categories: good agreement, mixed agreement, or poor agreement between signature and described process. For CZOs with multiple watersheds with contrasting properties, we assessed whether differences in signatures between these watersheds correspond to known contrasts in hydrologic processes.

5. Results

5.1. Large-Scale Patterns and Distributions

We calculated signature values across each CAMELS data set. Maps for the four regions and for eight representative signatures are shown in Figures S1–S8 in Supporting Information S1.

5.1.1. Distribution of Signature Values

The distributions of each signature for each CAMELS data set are shown in Figure 2. These distributions calculated using large samples can be used to assess empirically whether signature values should be considered high or low. The percentiles of each signature are given in Tables S3 and S4 in Supporting Information S1. Some signatures have clearly defined limits (e.g., *BFI*), while others such as watershed storage (*AverageStorage*) have no upper limit. We found that distributions can vary substantially between countries, such that a high signature value in one country might not be considered high in another country. For example, low values of the *BFI* in Brazil (BFI of 0.59 falls on the 25th percentile) would be considered average in the U.S. (BFI of 0.58 falls on the 50th percentile), Great Britain (BFI of 0.63 falls on the 50th percentile), and Australia (BFI of 0.53 falls on the 50th percentile).

5.1.2. Correlation Between Signature Values

To test whether signature correlations aligned with proposed physical interpretations of the signatures, we looked for examples where multiple signatures related to the same feature of a flux or store: that is, its magnitude, spatial variation, temporal variation, or response time. See Table 1 for a list of signatures and their corresponding processes, with references establishing those interpretations. The signature correlations across all four CAMELS data sets are shown in Figure 3.

Watersheds with high baseflow magnitude and long baseflow response time are supposed to be characterized by high *BFI* and low *BaseflowRecessionK*. This is correctly represented by a strong correlation (−0.88) between the two signatures. We note a strong correlation between *BFI* and *VariabilityIndex* (−0.80), suggesting that *VariabilityIndex* is related to baseflow magnitude more than storage as originally proposed (Estrany et al., 2010). There are further strong correlations between *BaseflowRecessionK*, *RecessionParameters_T0* (−0.74) and *Mid_Recession_Slope* (0.75), all of which relate to the response time of the watershed. These high correlations between *BFI* and multiple other signatures provide evidence for the common use of BFI as an overarching measure of baseflow importance (Figure 3a). However, the high correlation of baseflow magnitude and response time signatures means that the signatures do not provide a robust method to separate these two aspects of baseflow. Using multiple BFI signatures with different time windows would help to resolve this issue (Gnann, McMillan, et al., 2021).

Several signatures are supposed to be related to the magnitude of groundwater storage (Table 1), including *AverageStorage*, *ActiveStorage*, *TotalStorage*, and *RR_Seasonality*. High positive correlations occur between *ActiveStorage*, *AverageStorage* (0.77) and *TotalStorage* (0.86), confirming this interpretation. A lower correlation with *RR_Seasonality* (0.49) suggests that this signature is also influenced by processes other than storage. Small values of event runoff ratio (*EventRR*) and its fraction of total runoff ratio (*EventRR_Total_RR_ratio*) are supposed to signify fast drainage to groundwater and high storage capacity of the watershed (e.g., permeable bedrock), but this is not supported by the data. Instead, we found that these signatures are most highly correlated with total runoff ratio (correlations of 0.96 and 0.42), suggesting that they are controlled by losses to ET or deep groundwater as part of the overall water balance. The *StorageFraction* signature should relate to storage magnitude but was found to be unreliable, often giving unrealistic values and a poor fit when plotted against the underlying data. This signature was originally developed for a set of 16 watersheds in Luxembourg (Pfister et al., 2017), but modification or generalization of the signature would be needed for it to translate well to other watersheds.

The signature *MRC_num_segments* (number of segments in the master recession curve) has notably low correlations with all other signatures, including *RecessionParameters_b* (nonlinearity in the shape of recession curve) as

10.1029/2021WR031751

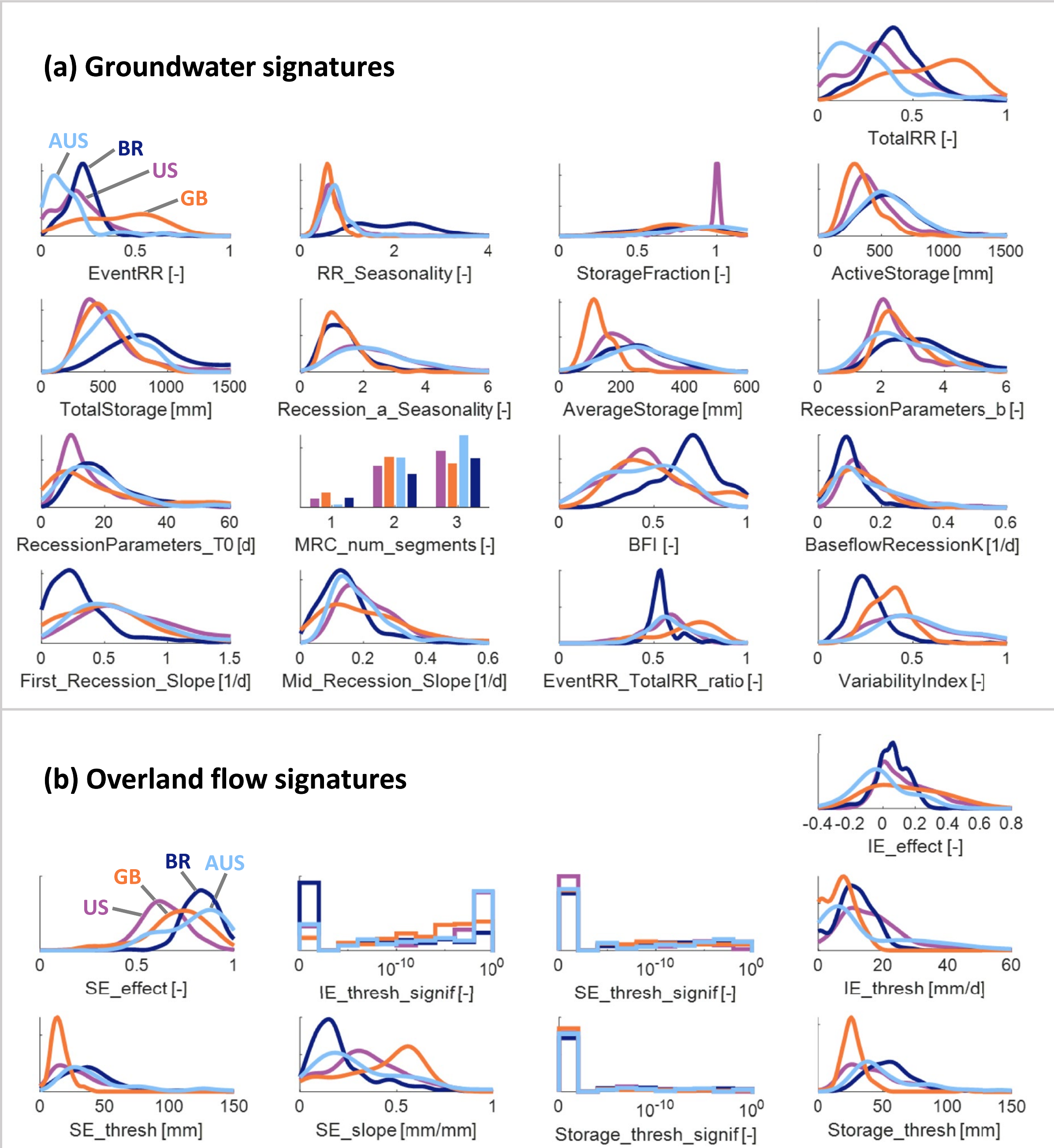

Figure 2. Distributions of (a) groundwater signatures and (b) overland flow signatures, smoothed using a kernel density estimation. Exceptions are MRC_num_segments, which can only take three values and is thus shown as a bar plot, and IE/SE/Storage_thresh_signif whose values cluster at 0 and 1, and are shown as histograms with a value at zero and other values on a log-axis. Plot ranges are adjusted for better visibility.

shown in Figure 3a. Its highest correlation is 0.27 with *Mid_Recession_Slope*. This is likely due to *MRC_num_segments* being an ordinal signature that can only take values of 1, 2, or 3, and therefore provides less information about the relative values of different watersheds.

For overland flow, we expected negative correlations between saturation excess importance (*SE_effect*) and the significance *P*-value of its threshold (*SE_thresh_signif*) and similarly for infiltration excess importance (*IE_effect*) and the significance *P*-value of its threshold (*IE_thresh_signif*). In other words, for watersheds with saturation excess process (where regression coefficients related to event precipitation depth most strongly predict quickflow volume and peak), there will be a threshold in a plot of event precipitation depth and quickflow volume. For

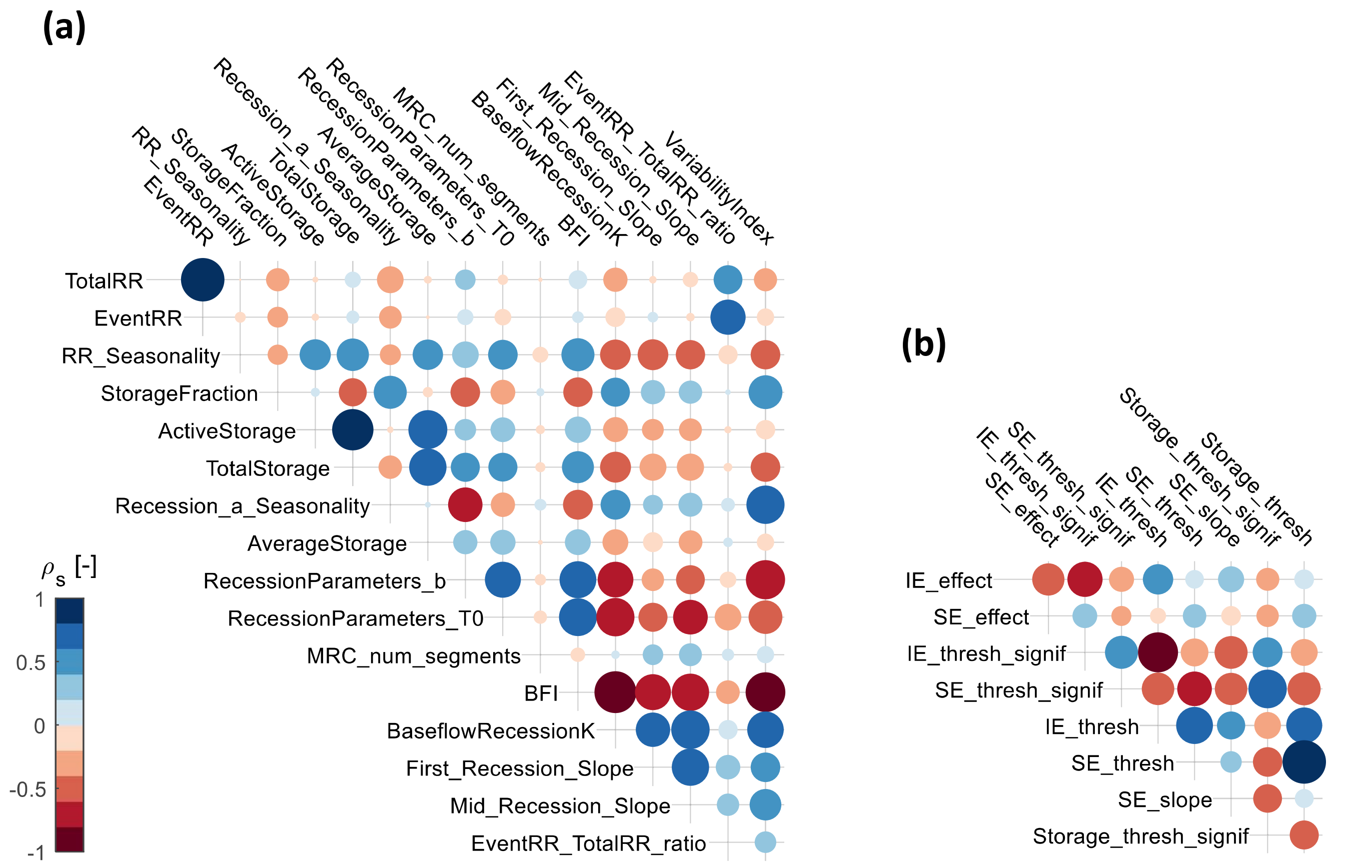

Figure 3. Correlations between (a) groundwater signatures and (b) overland flow signatures for all four CAMELS data sets.

watersheds with infiltration excess process (where regression coefficients related to event precipitation intensity most strongly predict quickflow volume and peak), there will be a threshold in a plot of event precipitation intensity and quickflow volume. This holds true for *IE_effect* (correlation −0.62). For *SE_effect*, the correlation is weak (−0.2), suggesting that event precipitation depth can control flow peak and volume but without a threshold in the relationship (Figure 3b). The data also show that threshold size and significance are negatively correlated (*IE_thresh_signif* has a correlation of −0.81 with *IE_thresh* and *SE_thresh_signif* has a correlation of −0.61 with *SE_thresh*), showing correctly that the signature algorithm will not identify a large threshold if it is not significant. Although not predicted in advance, we found that all the threshold sizes (*IE_thresh*, *SE_thresh*, and *Storage_thresh*) are strongly positively correlated. High values identify watersheds that require a lot of water to start producing flow, whether this be via infiltration or saturation excess mechanisms.

Overall, we find that expected relationships in baseflow and overland flow signatures hold in general, with a few exceptions as noted above.

5.1.3. Hydro-Climate Relationship to Signature Patterns

Aridity is correlated only with a few signatures (e.g., runoff ratios, see Figure 4a) when all CAMELS watersheds are lumped together (see Figures S9 and S10 and Tables S5 and S6 in Supporting Information S1 for all signatures). This changes when countries are investigated separately. For instance, the *IE_effect* signature has a strong correlation with aridity in Great Britain (rank correlation −0.92), while overall it shows only a very weak correlation (rank correlation −0.24), see Figure 4b. Sometimes, the relationships even have opposite signs, as is the case for the BFI in Great Britain and Australia, see Figure 4c. As we expected, these results show that signatures are not solely controlled by climate, but also by other watershed characteristics (e.g., soils and geology). The results may further point to climate-signature relationships from individual countries giving small windows into an underlying distribution; in this case, BFI might have a nonlinear relationship with aridity, maximized for intermediate values. Our results demonstrate that relationships between climate characteristics and signatures from a

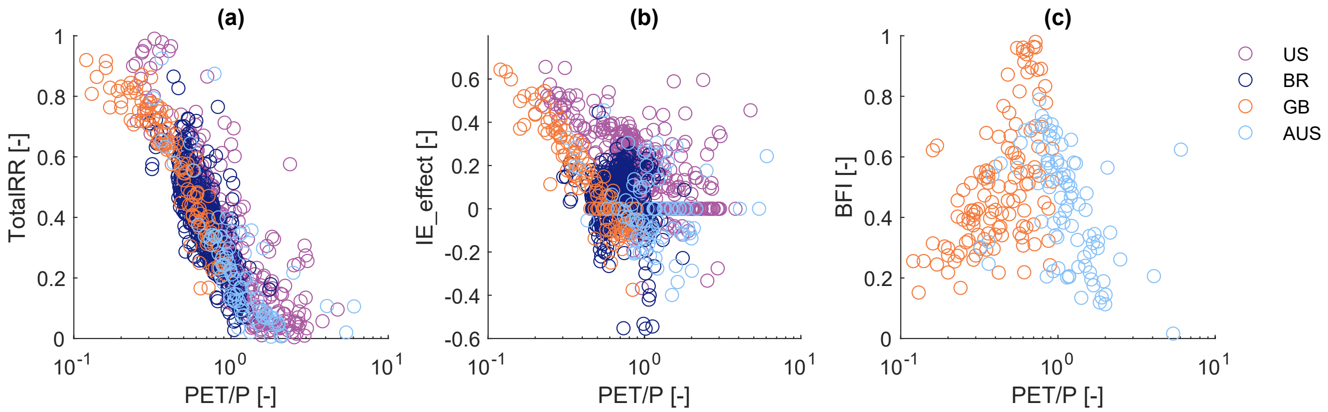

Figure 4. Relationships between (a) TotalRR and aridity (PET/P) for all CAMELS countries, (b) between IE_effect and aridity for all CAMELS countries, and (c) between BFI and aridity for Great Britain and Australia.

single country should not be assumed to hold in other countries. For example, the relationships of signatures to climate and catchment attributes found in the CAMELS U.S. data set (Addor et al., 2018) might not hold for other countries. This result shows the benefits of using multi-continent data sets to understand overarching drivers of signature patterns.

5.2. Signature-Process Links at Critical Zone Observatories

As described in Section 4.2, we collected literature from each CZO to answer each process question, and calculated the values and percentiles of each signature matched to that question. This information allowed us to describe the signature-process agreement in text, and ascribe good, partial, or poor agreement between each pair. For these well-studied watersheds, most of the process questions could be answered by searching the literature, but there were some gaps (white cells in the table). A full spreadsheet showing all signature values (with data at hourly and daily timestep), percentiles, descriptions of each process, and key references for each CZO watershed is given in the GitHub repository (see Data Availability Statement section). A summary showing signature-process agreement is shown here (Figure 5).

In general, there was good agreement between the processes interpreted from literature and the corresponding signatures (blue-green in Figure 5). Overall, groundwater signatures (71% agreement, 20% partial agreement, 8% poor agreement) were more reliable than overland flow signatures (46% agreement, 28% partial agreement, 26% poor agreement).

Good matches between signatures and processes occur in Eel River, Shale Hills, and Luquillo CZOs. Eel River in Northern California has two contrasting sub-watersheds, Elder Creek with high groundwater storage, and Dry Creek with low groundwater storage and frequent saturation excess flow (Lovill et al., 2018; Oshun et al., 2016). Groundwater signatures accurately represent the contrasts in storage and seasonality. However, SE flow in Dry Creek is incorrectly identified as IE at daily timescale, and BFI is moderate (not high as expected) in Elder Creek, representing a compromise between the Mediterranean climate that favors low baseflow and high storage that favors high baseflow. Shale Hills in Pennsylvania lies on sedimentary geology in the Appalachian Mountains (Brantley et al., 2018). In wet conditions, rising water tables generate interflow with a transmissivity-feedback mechanism (Scaife et al., 2020). Deep groundwater and surface flow are smaller components (Brantley et al., 2018; Li et al., 2018), although saturated riparian areas generate flow (Lin & Zhou, 2008; Takagi & Lin, 2012). Signatures and processes agree across almost all overland flow and groundwater processes, predicting low storage and BFI, fast recessions, a storage threshold for flow, seasonal ET influence, and low/moderate surface flow. Luquillo watersheds in Puerto Rico comprise tropical, montane forest, with Rio Icaros (granitoid) and Rio Mameyes (volcaniclastic) on contrasting geologies (McDowell et al., 1996). Fast/shallow processes dominate despite deep soils, with event water flowing as a perched water table (Schellekens et al., 2004; Shanley et al., 2011). Signatures and processes mostly agree, predicting fast processes (high event runoff coefficient and steep initial recessions), a storage threshold for flow, low seasonality, and no clear IE or SE dominance. Granitoid

10.1029/2021WR031751

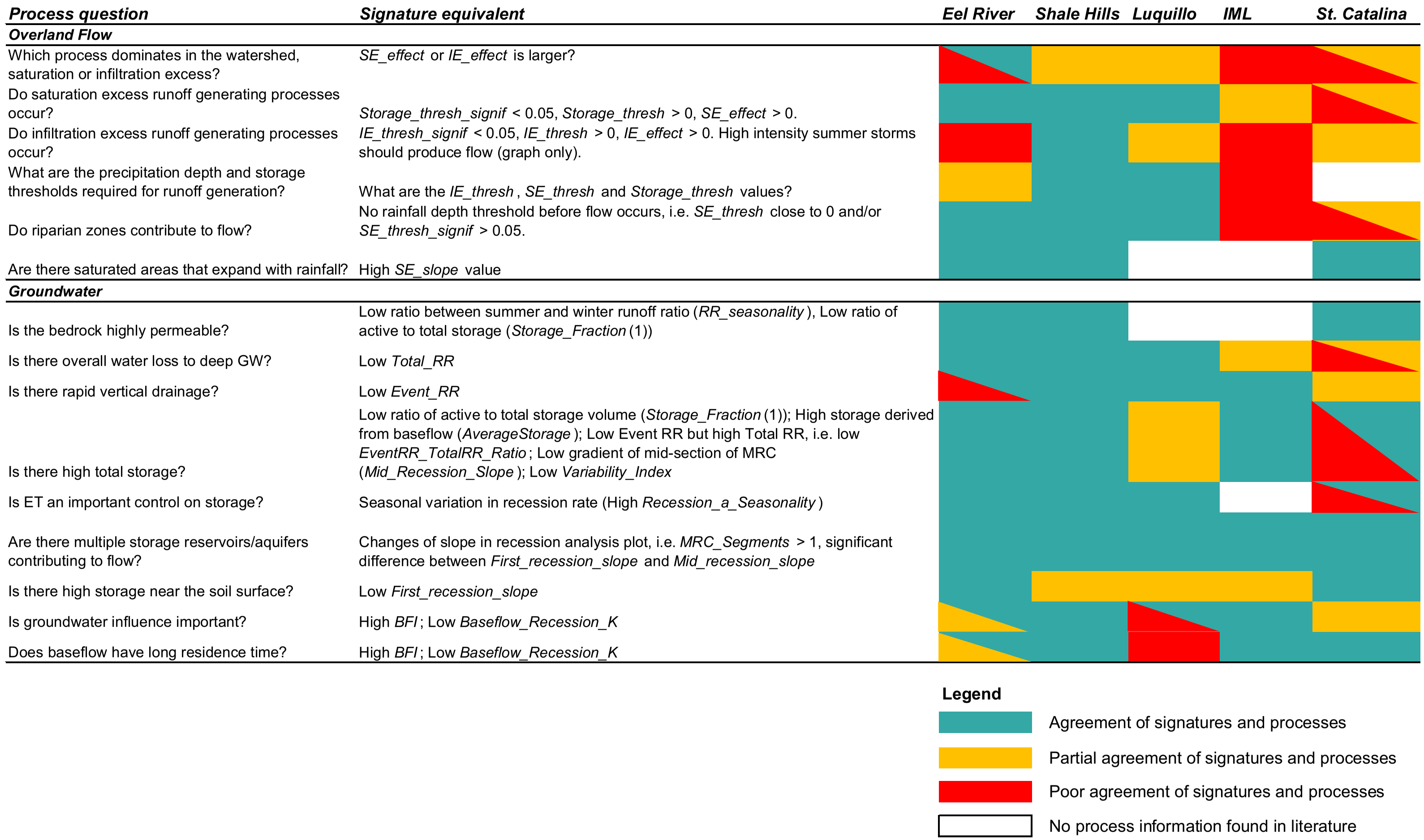

Figure 5. Questions derived from McMillan (2020), used to compare signature values and process knowledge. For each Critical Zone Observatory watershed, cells are color-coded for process knowledge and signature values agreement (blue-green), partial agreement (yellow), or poor agreement (red). Where results differ substantially between signatures and/or subwatersheds, split colors are used. White cells imply that we found insufficient literature to answer the question.

Icaros correctly has higher BFI and lower Recession K than Mameyes, but for both rivers the recession parameters incorrectly merge fast event processes with a small but sustained baseflow component.

Poorer matches between signatures and processes occur in IML and Santa Catalina CZOs. The IML Upper Sangamon watershed in Illinois is in row-crop agriculture with tile drains (Kumar et al., 2018; Schilling et al., 2018). SE and IE flow are both reported in the literature, but with IE dominant (Abban et al., 2014; Davis et al., 2014; Wilson et al., 2012). Groundwater rises quickly after events and runs off via tile drains (Kim et al., 2020; Schilling & Helmers, 2008; Wilson et al., 2018). Groundwater signatures and processes mostly agree, with low event runoff ratio suggesting drainage to groundwater, and low storage signatures suggesting that this groundwater drains quickly to the stream. However, overland flow processes are incorrectly identified as dominated by SE, with IE found not significant although it is known to occur. The Santa Catalina watersheds in Northern Arizona are arid, mid to high-elevation sites. Heavy summer monsoon storms produce overland and near-surface flows dominated by event and soil water (Desilets et al., 2008; Dwivedi et al., 2019). IE dominates, with some near-stream SE (Lyon et al., 2008). Streams gain some water from deep/regional groundwater (Dwivedi et al., 2019). Signatures and processes often disagree; signatures incorrectly suggest that saturation and storage processes dominate, and very low runoff ratios make interpretation difficult. However, moderate BFI and recession K agree with the limited but important groundwater contribution.

6. Discussion

In this paper, we applied hydrological signatures and assessed their process interpretations in a diverse set of basins, in many cases well outside of the hydroclimatic regimes for which the signatures were designed. Therefore, we gained many useful insights into signature use.

6.1. Benefits of Calculating Signature Distributions

To interpret signature values as "high" or "low," it was essential to know the distributions of signature values (Figure 2). For example, *SE_effect* values (saturation excess importance) are consistently higher than *IE_effect* values (infiltration excess importance), so these values are interpreted differently. It was therefore useful to present the signature value as a percentile of the national distribution rather than an absolute value. We recommend a comparison with regional rather than global distributions when transferring knowledge about links between signatures and processes (e.g., where a high signature value indicates a dominant process), as we hypothesize that authors implicitly compare processes within the same region. For example, a low *BFI* in Brazil might be considered an average *BFI* in the U.S. Understanding signature distributions and their spread (e.g., variance) is important when using signatures to assess watershed similarity, as it enables us to calculate how similar two signature values are, for example, within the same decile of the distribution.

Analyzing signature values on a national scale was also useful when understanding correlations of signatures with aridity. We found examples where a correlation that occurs for the combined set of four CAMELS data sets does not hold for individual countries. Strong but diverging correlations (e.g., *BFI*, see Figure 4c) might point at relationships that are not causal. For instance, the most productive aquifers in Great Britain happen to be in the least humid places, so a correlation between aridity and *BFI* here might be a coincidence, and this could be the reason why it does not hold for other countries. Understanding such regional differences in relationships is essential for prediction of signatures in ungauged basins, which is used in applications that regionalize signatures to ungauged basins, and then use those signature values as a performance measure to calibrate a rainfall-runoff model. One approach to this is to group sites by region or climate before analyzing the physiographic drivers of signature values within each group, as used in recent studies of non-perennial rivers and river flashiness (Gannon et al., 2022; Hammond et al., 2021).

6.2. Signature Robustness

Large sample signature calculations pose several challenges. Some signatures are straightforward to calculate (e.g., *TotalRR*), or have been widely used (e.g., *BFI*) and are relatively robust. Some watershed types may prove more susceptible to uncertainties or difficulties in signature calculation, such as leaky watersheds that cause errors in estimated ET (Wlostowski et al., 2020). Signatures that have only been used in a few small scale studies,

10.1029/2021WR031751

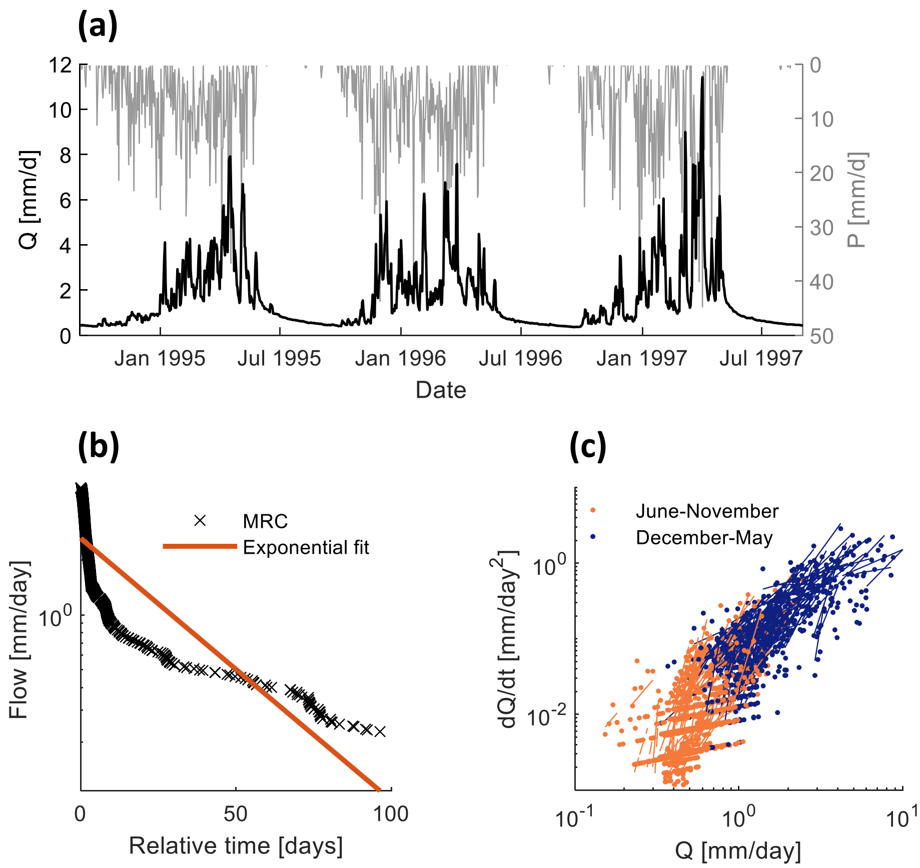

Figure 6. (a) Hydrograph of a watershed in Brazil that shows very seasonal (monsoonal) precipitation and thus streamflow. (b) The corresponding master recession curve flattens out for late recessions, indicating the transition from wet season recessions to dry season recessions. (c) The same can be seen from the dQ/dt plot where dry season recessions are systematically steeper.

or are sensitive to parameter selections (e.g., recessions, see Dralle et al., 2017; Stoelzle et al., 2013) are less robust for large samples. Their results might not be reliable, even though the values might be within a realistic range. For example, the monsoonal climate in some parts of Brazil leads to a distinct seasonal flow regime, with many short recessions during the wet season and a long recession during the dry season (Figure 6). This is not picked up by the recession signatures we used, as they return a single (average) storage-discharge relationship. It is thus important to check visually whether the signature results are reasonable, and to test signatures when transferring them to other scales or other places. It is then possible to tailor the signatures to certain regions, for example, by dividing the time series according to season (Euser et al., 2013). Quantifying the sources and magnitudes of uncertainty that impact signature values is necessary for many applications (Westerberg & McMillan, 2015). For example, uncertainty information is needed when using signatures as metrics for model calibration (where uncertainty information is used in the likelihood function), regionalizing signatures and their uncertainties for flow prediction in ungauged basins (Prieto et al., 2019; Westerberg et al., 2016), or using signature changes through time to assess hydrologic alteration (Vigiak et al., 2018).

Although we expected hourly data to provide more accurate estimates of event characteristics and recession dynamics in small watersheds, we found that working with hourly data required a hands-on approach to prevent errors. This included changing recession selection tolerance due to diurnal flow fluctuations, and filling gaps in timeseries; for example, USGS flow data is usually infilled at the daily timescale, but not for hourly data. Such interventions cause hourly signature values to be more uncertain, trading off accuracy and data processing time. We suggest comparing hourly and daily signature values, and identifying reasons for significant differences. In the arid Santa Catalina CZO, hourly data gave poor results that did not match literature information on processes. We do however recommend hourly data for identifying overland flow processes, as it produced results that better matched field observations, and were often substantially different from daily results. For example, the IE_effect and SE_effect signatures represent the importance of infiltration and saturation excess processes. At Dry Creek in Eel River CZO, saturation excess is known to dominate (Lovill et al., 2018). With daily data, IE_effect is 0.46 (96% quantile), and SE_effect is 0.47 (9% quantile); while with hourly data, IE_effect is 0.34 (87% quantile) and SE_effect is 0.65 (58% quantile). Therefore hourly data more close represents the SE dominance. At Shale Hills, some surface flow is known to occur, although infiltration versus saturation excess processes are not described (Brantley et al., 2018). If daily data is used, neither IE nor SE flow are indicated (IE_effect = 0 (9% quantile), SE_effect = 0.42 (6% quantile)), while with hourly data, moderate SE_flow is indicated (IE_effect = −0.09 (3% quantile), SE_effect = 0.61 (43% quantile)). Therefore, hourly data provides a better match to observations. This finding confirms that of Wu et al. (2021) who found that hourly precipitation data was preferred when calculating signatures for infiltration excess. Some large scale data sets have hourly data available, for example, CAMELS in the U.S. (Gauch et al., 2021) and LamaH-CE in Central Europe (Klingler et al., 2021), and we therefore recommend using the hourly versions if investigating infiltration excess processes.

6.3. Lessons From Comparisons of Signatures and Processes for CZOs

We found some challenges in the CZO watersheds when comparing signature values to process descriptions. It could be difficult to obtain standardized process data, such as depth to bedrock which was sometimes quantified differently by different authors, even in the same watershed. We sometimes found conflicting information, such as in the IMLs CZO, where saturation excess flow was said to occur, but the water table was said to be low due to tile drains. Such conflicts could be due to differences in exact location, or in wetness conditions at the time of observation, and illustrate the difficulties in summarizing complex understanding of the landscape. Similarly, some signature values failed to capture the full insights of field studies, such as in Shale Hills CZO that has a known discharge of old (20–30 years) water to the stream, but which comprises only a small percentage of flow (Li et al., 2018). Although previous studies have linked high *BFI* and low *BaseflowRecessionK* to groundwater residence time (Bulygina et al., 2009; Safeeq et al., 2013), in this case the BFI signature value is low (due to the small volume of old water) and so does not indicate the old water discharge.

Some processes were less reliable in their match to signature values across multiple CZOs. In particular, IE and SE processes (first three rows in Figure 5) were not well differentiated. For example, SE is known to dominate in Dry Creek in the Eel River CZO, but signatures show IE; whereas IE is known to dominate in IML CZO, but signatures show SE. Additionally, watersheds differed in their reliability, for example, Shale Hills showed high reliability across all signatures with no disagreements. In the arid but high-elevation Santa Catalina CZO, the reliability of signatures based on events and recession periods was reduced, because only a small number of storms produced flow, and some of these were impacted by snow. The IML CZO, where the hydrology reflects significant human impacts in cropped areas, showed low reliability for overland flow signatures. The low reliability could be due to high variability in processes between impacted and non-impacted areas of the watershed, as overland flow signatures face a scale conflict between location-specific observations of flow, and signature values that reflect integrated watershed response. Low reliability could also be due to human impacts such as tile drains. The tile drain response could mimic saturation excess processes if the water table rises high enough to intersect the tile drain layer, providing a fast pathway from groundwater into the channel. Signatures are known to be modified by human activities, such as baseflow index being affected by groundwater abstraction and effluent discharges to rivers (Bloomfield et al., 2021), but were not originally designed for use in human-impacted watersheds. The gaps found here in signature availability and accuracy suggest opportunities to develop signatures targeted at deep groundwater contributions for ecohydrology studies, or targeted at human impacts to assess drivers of hydrologic alteration.

6.4. Comparison With Previous Studies

It is useful to compare our results with previous studies that related process descriptions to signature values. For overland flow processes, Wu et al. (2021) identified infiltration and saturation excess using Spearman correlations between event runoff ratios and rainfall intensity, rainfall volume and rainfall storage. They found few watersheds with dominant infiltration excess, agreeing with previous findings that IE flow is rare in the U.S. (Buchanan et al., 2018; Wolock, 2003). However, there are substantial differences in spatial patterns of IE between our study and these previous studies, and among the previous studies. One explanation for high uncertainty is that overland flow signatures are sensitive to calculation methods, particularly whether hourly or daily rainfall intensity is used, and require multiple choices including baseflow separation, event definition, and storage calculation methods. Our *IE_effect* and *SE_effect* signatures are based on a study by Estrany et al. (2010) in a Mediterranean watershed, but may not function correctly in other climates, as also evidenced by the unexpected positive correlation between wetness and *IE_effect* in daily CAMELS-GB data. Our *IE_thresh* and *SE_thresh* signatures were more consistent with previous studies and process knowledge, particularly when using hourly rainfall intensity. They showed positive infiltration excess thresholds in arid and Southeastern U.S., where infiltration excess is expected to occur, and positive saturation excess thresholds in most of the U.S. except the arid West and in the North East where antecedent conditions may outweigh event volume. In summary, *IE_effect* and *SE_effect* signatures are not reliable, and future work is needed to design and test signatures that better differentiate these processes.

For groundwater processes, we can compare our results with those of Wlostowski et al. (2020), who studied how critical zone architecture controls signature values. Our results agree with theirs in finding that baseflow and storage signatures are controlled not by depth to bedrock but rather by properties and structures of the soil. For example, expert observations of whether shallow interflow and return flow processes occur were more likely to match signature values than a simple depth to bedrock value. An example occurs in the Luquillo CZO, where depth to bedrock at the Rio Mameyes site is 30–40m, but streamflow dynamics are dominated by rapid delivery of event water to the stream by fast, shallow runoff processes including lateral macropore flow. Wlostowski et al. (2020) further agreed with our findings in noting a clear influence of tile drains in signatures for the IML CZO.

6.5. Recommendations for Signature Choice

In this section, we record signature-specific conclusions from our study, in particular whether signatures related to the process interpretations as previously proposed in Table 1, whether signatures could be robustly calculated across large samples of watersheds, and whether signatures relied on any watershed-specific fitting parameters. Fitting parameters that affect multiple signatures, and would ideally be checked visually against the flow time-series, are those that control event selection and recession selection. These common parameters are noted in Table 1, although other signatures may also require parameters such as for baseflow separation. Our recommendations are based on analyses across the CAMELS and CZO watersheds. Signature robustness and correlations were tested for CAMELS watersheds, and therefore these findings hold across a wide variety of basin sizes, hydroclimates and geophysical attributes. The process interpretations were tested for five CZO sites with a total of eight streamflow gauges. While these include a wide range of hydroclimates and landscapes (arid to humid to tropical), the smaller number of locations means that more caution should be applied when transferring the results to other watersheds. The recommendations are summarized in Table 2.

7. Conclusions

This study tested whether relationships between signatures and processes developed from experimental watershed studies hold true when applied over large scales and diverse hydro-climates. The relationships were tested using two types of data: large sample CAMELS data sets from four countries, and detailed information from five CZO watersheds in the U.S. We note that when single signature values are used to summarize complex watershed responses, they might represent a compromise value between climate and process effects, or between multiple processes (e.g., fast and slow recession processes, or spatially variable overland flow). This compromise demonstrates the difficulty of summarizing processes using quantitative values, without losing some information.

We recommended a small set of preferred signatures that provide accurate and robust metrics for common process characteristics across diverse watersheds, as follows: *TotalRR* for water balance, *Recession_a_seasonality* for seasonal variability in storage and recessions, *AverageStorage* for storage magnitude, *BFI* for baseflow magnitude, *BaseflowRecessionK* for baseflow response time, *IE/SE/Storage_thresh* for importance of infiltration excess, saturation excess and pre-event storage on flow generation, and *SE_slope* for saturated area expansion. We identified a small number of signatures that were not reliable (*SE_effect*, *IE_effect*, and *StorageFraction*) or had different interpretations than expected (*EventRR*) and would therefore not be recommend for use to identify the related processes in other watersheds. We made recommendations for adapting some signatures to better differentiate between processes (using multiple BFI timescales to separate baseflow magnitude and response time, calculating *Storage_thresh* with longer memory to differentiate from *SE_thresh*). Overall, the results showed that

Table 2
Summary of Findings for Groundwater and Overland Flow Signatures

Signature	Process interpretation	Robustness/limitations	Event	Recession
Groundwater and baseflow signatures				
TotalRR	**Recommended signature for water balance.** Strongly related to aridity across all CAMELS data sets (Figure 4a). At large scales describes climate more than hydrology, modified by storage dynamics at small scales (e.g. compare Dry and Elder watersheds at Eel CZO)	Easy and robust to calculate		
EventRR	High (0.96) correlation to Total RR, therefore relates to total water balance rather than to watershed storage as previously claimed		x	
RR_Seasonality	Our results in CZO watersheds confirm previous interpretations that this signature relates to bedrock permeability and watershed storage size. Correlates well to other storage magnitude signatures			x
StorageFraction (incl. ActiveStorage and TotalStorage)		Unreliable signature for large samples as seen in unrealistic values and poor fit in plotting		
Recession_a_Seasonality	**Recommended signature for seasonal variability in storage and recessions.** Our results in CZO watersheds confirm previous interpretations that this signature relates to ET influence on storage. Moderate to strong correlations with *RecessionParameters_b* (−0.60) and *VariabilityIndex* (0.74)			x
AverageStorage	**Recommended signature for storage magnitude.** Our results in CZO watersheds confirm previous interpretations that this signature relates to watershed storage	More reliable than the StorageFraction above, recommended when estimates of storage are needed		
RecessionParameters (*b*, T0)	Established signatures with theoretical link to watershed storage-discharge relationship, highly correlated to other signatures of baseflow magnitude and response time	Does not distinguish short event recessions from longer dry season recessions		x
MRC_num_segments	Useful signature to identify complexity of recession shapes	Robust across a wide range of recession characteristics. For recessions with multiple segments of different slopes, *BaseflowRecessionK* may be unreliable		x
BFI	**Recommended signature for baseflow magnitude.** Reliable signature that correlates strongly with most other baseflow and storage signatures	BFI integrates multiple aspects of baseflow (volume of baseflow, response time, multiple baseflow sources or pathways), which cannot be distinguished based on a single BFI value		
BaseflowRecessionK	**Recommended overall signature for baseflow response time.** Our results confirm relation to baseflow magnitude and residence time for most CZOs. Strong correlation with BFI (−0.88), and with other indicators of residence or response time, *RecessionParameters_T0* (−0.74) and *Mid_Recession_Slope* (0.75)	Usually reliable, however, if using high BFI and low Baseflow_Recession_K to indicate long GW residence times, may not differentiate between low baseflow with long residence time or moderate baseflow with moderate residence times. We advise a visual check of fit when MRC_num_segments is >1		x
First_Recession_Slope	A low first slope is supposed to indicate high storage near the soil surface, that is, the fastest flow path is delayed due to this storage. However, it is not easy to determine whether this matches with soil profile observations that typically record soil texture and/or importance of shallow flow processes	MRC fitting was robust across varied recession shapes		x

Table 2
Continued

Signature	Process interpretation	Robustness/limitations	Event	Recession
Mid_Recession_Slope	Correlates strongly (0.75) with *BaseflowRecessionK* for response time, confirms previous description that it relates to retention capacity of watershed	MRC fitting was robust across varied recession shapes		x
EventRR_TotalRR_ratio	Showed reasonable match to process interpretation related to storage capacity in most CZOs, except for arid Santa Catalina watershed	Event and Total RR are highly correlated, making this ratio more uncertain	x	
VariabilityIndex	Showed a moderate fit to storage information in CZO watersheds and easier to calculate than AverageStorage. Alternative to Recession_a_Seasonality for quantifying seasonal variation in storage and recessions	Easy and robust to calculate. AverageStorage signature is preferred for storage estimates		
Overland flow signatures				
IE_effect	Watersheds where IE or SE dominates can be incorrectly identified by IE_effect and SE_effect values	IE_effect is not a reliable signature. Hourly data produces a closer match to process knowledge	x	
SE_effect	Not strongly related to the threshold signatures; may show control of rainfall depth on flow, independent of the existence of a threshold	SE_effect is not a reliable signature. SE_effect is above 0.5 for most watersheds, so need the percentile to quantify high/low values. Hourly data produces a closer match to process knowledge	x	
IE_thresh, SE_thresh, Storage_thresh	**Recommended signatures for importance of infiltration excess, saturation excess and pre-event storage on flow generation.** Large, significant thresholds suggest that these processes are important. All the thresholds (IE, SE, storage thresh) are strongly positively correlated. That is, this identifies watersheds that require a lot of water to start producing flow	More reliable than IE/SE_effect. The very strong correlation between Storage and SE thresholds shows a difficulty in separating the impacts of pre-event storage and event depth. To achieve this, a longer storage memory than Mosley (1979) 30-day API used in the Storage_thresh signature may be required. Hourly rainfall data is preferred for IE_thresh	x	
IE_thresh_signif, SE_thresh_signif, Storage_thresh_signif	Useful to confirm threshold existence. For Storage_thresh_signif see comments on Storage_thresh above	Strongly negatively correlated to IE_thresh (for IE_thresh_signif) or SE_thresh (for SE_thresh_signif and Storage_thresh_signif), that is, the code will generally not identify a threshold if it is not significant	x	
SE_slope	**Recommended signature for saturated area expansion.** Our results in CZO watersheds confirm previous interpretations that this signature relates to expansion of saturated areas with event precipitation		x	

Note. Column "Event" shows where a signature depends on prior event separation, column "Recession" shows where a signature depends on prior identification of recession periods; both these cases require additional parameters that may be watershed-dependent. Rows shaded in gray identify a smaller number of preferred signatures that provide accurate and robust metrics for common process characteristics, these are described in bold.

most signature patterns agreed with process interpretations, with groundwater and baseflow signatures being more reliable than overland flow signatures. Based on the CZO watershed results, signature-process relationships were most reliable in humid and Mediterranean-climate watersheds, and less reliable in arid and human-impacted watersheds. This difference reflects the history of signature development which has been concentrated in natural, humid basins, and points to scope for future signature development in a wider range of watersheds.

Data Availability Statement

The CAMELS U.S. data set is available at https://dx.doi.org/10.5065/D6MW2F4D (Addor et al., 2017; Newman et al., 2015). The hourly rainfall data set corresponding to the CAMELS U.S. locations is available at https://doi.org/10.5281/zenodo.4072700 (Gauch et al., 2020, 2021). The CAMELS Great Britain data set is available at https://doi.org/10.5285/8344e4f3-d2ea-44f5-8afa-86d2987543a9 (Coxon et al., 2020). The CAMELS Australia data set is available at https://doi.pangaea.de/10.1594/PANGAEA.921850 (Fowler et al., 2021). The CAMELS Brazil data set is available at https://zenodo.org/record/3964745 (Chagas et al., 2020). Critical Zone Observatory (CZO) data products are available at https://doi.org/10.4211/hs.29e2ec85770b42c881ef0750696463e5 (Wlostowski et al., 2021). The Toolbox for Streamflow Signatures in Hydrology toolbox (Gnann et al., 2021a) used to calculate hydrologic signatures is available at https://github.com/TOSSHtoolbox/TOSSH and archived at https://zenodo.org/record/6462813 (Gnann et al., 2021b). The code used to load the CAMELS data into MATLAB is available at https://github.com/SebastianGnann/CAMELS_Matlab and archived at https://zenodo.org/record/6462821 (Gnann, 2022). The code to reproduce our analysis, and a full spreadsheet showing all signature values (with data at hourly and daily timestep), percentiles, descriptions of each process and key references for each CZO watershed, are available at https://github.com/SebastianGnann/LargeScaleSigs and archived at https://zenodo.org/record/6462823 (McMillan et al., 2022).

Acknowledgments
Support to H. McMillan and R. Araki was provided by the NSF Hydrologic Sciences Program, Division of Earth Sciences, Grant/Award Number 2124923. Partial support to S. J. Gnann was provided by the Water Informatics Science and Engineering Centre for Doctoral Training (WISE CDT) under a grant from the Engineering and Physical Sciences Research Council (EPSRC), Grant Number EP/L016214/1 and the Alexander von Humboldt Foundation in the framework of the Alexander von Humboldt Professorship endowed by the German Federal Ministry of Education and Research.

References

Abban, B., Papanicolaou, A. N., Cowles, M. K., & Wilson, C. G. (2014). Examining seasonal trends in sediment source contributions in an intensely cultivated midwestern sub-watershed using Bayesian unmixing. In *World Environmental and Water Resources Congress 2014* (pp. 1453–1463).

Addor, N., Nearing, G., Prieto, C., Newman, A. J., Le Vine, N., & Clark, M. P. (2018). A ranking of hydrological signatures based on their predictability in space. *Water Resources Research*, *54*(11), 8792–8812. https://doi.org/10.1029/2018WR022606

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, *21*(10), 5293–5313. https://doi.org/10.5194/hess-21-5293-2017

Ali, G., Oswald, C. J., Spence, C., Cammeraat, E. L., McGuire, K. J., Meixner, T., & Reaney, S. M. (2013). Towards a unified threshold-based hydrological theory: Necessary components and recurring challenges. *Hydrological Processes*, *27*, 313–318. https://doi.org/10.1002/hyp.9560

Beighley, R. E., Dunne, T., & Melack, J. M. (2005). Understanding and modeling basin hydrology: Interpreting the hydrogeological signature. *Hydrological Processes*, *19*(7), 1333–1353. https://doi.org/10.1002/hyp.5567

Black, P. E. (1997). Watershed functions. *Journal of the American Water Resources Association*, *33*(1), 1–11. https://doi.org/10.1111/j.1752-1688.1997.tb04077.x

Bloomfield, J. P., Gong, M., Marchant, B. P., Coxon, G., & Addor, N. (2021). How is Baseflow Index (BFI) impacted by water resource management practices? *Hydrology and Earth System Sciences*, *25*(10), 5355–5379. https://doi.org/10.5194/hess-25-5355-2021

Blume, T., Zehe, E., Reusser, D. E., Iroumé, A., & Bronstert, A. (2008). Investigation of runoff generation in a pristine, poorly gauged catchment in the Chilean Andes I: A multi-method experimental study. *Hydrological Processes: International Journal*, *22*, 3661–3675.

Brantley, S. L., White, T., West, N., Williams, J. Z., Forsythe, B., Shapich, D., et al. (2018). Susquehanna Shale Hills critical zone observatory: Shale Hills in the context of Shaver's Creek watershed. *Vadose Zone Journal*, *17*, 1–19. https://doi.org/10.2136/vzj2018.04.0092

Buchanan, B., Auerbach, D. A., Knighton, J., Evensen, D., Fuka, D. R., Easton, Z., et al. (2018). Estimating dominant runoff modes across the conterminous United States. *Hydrological Processes*, *32*(26), 3881–3890. https://doi.org/10.1002/hyp.13296

Bulygina, N., McIntyre, N., & Wheater, H. (2009). Conditioning rainfall-runoff model parameters for ungauged catchments and land management impacts analysis. *Hydrology and Earth System Sciences*, *13*, 893–904. https://doi.org/10.5194/hess-13-893-2009

Chagas, V. B., Chaffe, P. L., Addor, N., Fan, F. M., Fleischmann, A. S., Paiva, R. C., & Siqueira, V. A. (2020). CAMELS-BR: Hydrometeorological time series and landscape attributes for 897 catchments in Brazil. *Earth System Science Data*, *12*(3), 2075–2096. https://doi.org/10.5194/essd-12-2075-2020

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D. E., Woods, R. A., et al. (2015). A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resources Research*, *51*(4), 2515–2542. https://doi.org/10.1002/2015WR017200

Clark, M. P., Rupp, D. E., Woods, R. A., Tromp-van Meerveld, H. J., Peters, N. E., & Freer, J. E. (2009). Consistency between hydrological models and field observations: Linking processes at the hillslope scale to hydrological responses at the watershed scale. *Hydrological Processes: International Journal*, *23*, 311–319. https://doi.org/10.1002/hyp.7154

Coxon, G., Addor, N., Bloomfield, J. P., Freer, J., Fry, M., Hannaford, J., et al. (2020). CAMELS-GB: Hydrometeorological time series and landscape attributes for 671 catchments in Great Britain. *Earth System Science Data*, *12*(4), 2459–2483. https://doi.org/10.5194/essd-12-2459-2020

Davis, C. A., Ward, A. S., Burgin, A. J., Loecke, T. D., Riveros-Iregui, D. A., Schnoebelen, D. J., et al. (2014). Antecedent moisture controls on stream nitrate flux in an agricultural watershed. *Journal of Environmental Quality*, *43*, 1494–1503. https://doi.org/10.2134/jeq2013.11.0438

Desilets, S. L., Ferré, T. P., & Ekwurzel, B. (2008). Flash flood dynamics and composition in a semiarid mountain watershed. *Water Resources Research*, *44*, W12436. https://doi.org/10.1029/2007WR006159

Dralle, D. N., Karst, N. J., Charalampous, K., Veenstra, A., & Thompson, S. E. (2017). Event-scale power law recession analysis: Quantifying methodological uncertainty. *Hydrology and Earth System Sciences*, *21*(1), 65. https://doi.org/10.5194/hess-2016-341

Dunne, T. (1978). Field studies of hillslope flow processes. In M. J.Kirkby (Ed.), *Hillslope hydrology* (pp. 227–293). John Wiley & Sons.

Dwivedi, R., Meixner, T., McIntosh, J. C., Ferré, P. A. T., Eastoe, C. J., Niu, G.-Y., et al. (2019). Hydrologic functioning of the deep critical zone and contributions to streamflow in a high-elevation catchment: Testing of multiple conceptual models. *Hydrological Processes*, *33*, 476–494. https://doi.org/10.1002/hyp.13363

Estrany, J., Garcia, C., & Batalla, R. J. (2010). Hydrological response of a small Mediterranean agricultural catchment. *Journal of Hydrology*, *380*(1–2), 180–190. https://doi.org/10.1016/j.jhydrol.2009.10.035

Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., & Savenije, H. H. G. (2013). A framework to assess the realism of model structures using hydrological signatures. *Hydrology and Earth System Sciences*, *17*(5), 1893–1912. https://doi.org/10.5194/hess-17-1893-2013

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., et al. (2019). Hillslope hydrology in global change research and Earth system modeling. *Water Resources Research*, *55*(2), 1737–1772. https://doi.org/10.1029/2018wr023903

Fowler, K. J., Acharya, S. C., Addor, N., Chou, C., & Peel, M. C. (2021). CAMELS-AUS: Hydrometeorological time series and landscape attributes for 222 catchments in Australia. *Earth System Science Data Discussions*, 1–30. https://doi.org/10.5194/essd-13-3847-2021

Gannon, J. P., Kelleher, C., & Zimmer, M. (2022). Controls on watershed flashiness across the continental US. *Journal of Hydrology*, *609*, 127713. https://doi.org/10.1016/j.jhydrol.2022.127713

Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., & Hochreiter, S. (2020). Models and predictions for "Rainfall-Runoff Prediction at Multiple Timescales With a Single Long Short-Term Memory Network" [Data set]. Zenodo. https://doi.org/10.5281/zenodo.4095485

Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., & Hochreiter, S. (2021). Rainfall–runoff prediction at multiple timescales with a single Long Short-Term Memory network. *Hydrology and Earth System Sciences*, *25*(4), 2045–2062. https://doi.org/10.5194/hess-25-2045-2021

Gnann, S. J. (2022). *SebastianGnann/CAMELS_Matlab: First release (1.0.0)*. Zenodo. https://doi.org/10.5281/zenodo.6462821

Gnann, S. J., Coxon, G., Woods, R. A., Howden, N. J., & McMillan, H. K. (2021a). TOSSH: A Toolbox for Streamflow Signatures in Hydrology. *Environmental Modelling & Software*, *138*, 104983. https://doi.org/10.1016/j.envsoft.2021.104983

Gnann, S. J., Coxon, G., Woods, R. A., Howden, N. J., & McMillan, H. K. (2021b). *TOSSHtoolbox/TOSSH: First release – Minor update (1.0.1)*. Zenodo. https://doi.org/10.5281/zenodo.6462813

Gnann, S. J., McMillan, H. K., Woods, R. A., & Howden, N. J. K. (2021). Including regional knowledge improves baseflow signature predictions in large sample hydrology. *Water Resources Research*, *57*(2), e2020WR028354. https://doi.org/10.1029/2020WR028354

Grayson, R., & Blöschl, G. (2001). Summary of pattern comparison and concluding remarks. *Spatial Patterns in Catchment Hydrology: Observations and Modelling*, 355–396.

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., & Andréassian, V. (2014). Large-sample hydrology: A need to balance depth with breadth. *Hydrology and Earth System Sciences*, *18*(2), 463–477. https://doi.org/10.5194/hess-18-463-2014

Gupta, H. V., Wagener, T., & Liu, Y. (2008). Reconciling theory with observations: Elements of a diagnostic approach to model evaluation. *Hydrological Processes: International Journal*, *22*(18), 3802–3813. https://doi.org/10.1002/hyp.6989

Hammond, J. C., Zimmer, M., Shanafield, M., Kaiser, K., Godsey, S. E., Mims, M. C., et al. (2021). Spatial patterns and drivers of nonperennial flow regimes in the contiguous United States. *Geophysical Research Letters*, *48*(2), e2020GL090794. https://doi.org/10.1029/2020gl090794

Harrigan, S., Hannaford, J., Muchan, K., & Marsh, T. J. (2018). Designation and trend analysis of the updated UK Benchmark Network of river flow stations: The UKBN2 dataset. *Hydrology Research*, *49*(2), 552–567. https://doi.org/10.2166/nh.2017.058

Hellie, F., Peschke, G., Seidler, C., & Niedel, D. (2002). Process-oriented subdivision of basins to improve the preprocessing of distributed precipitation-runoff-models. *Interdisciplinary Approaches in Small Catchment Hydrology: Monitoring and Research*, *137*.

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., et al. (2014). Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. *Water Resources Research*, *50*(9), 7445–7469. https://doi.org/10.1002/2014wr015484

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., et al. (2013). A decade of predictions in ungauged basins (PUB)—A review. *Hydrological Sciences Journal*, *58*, 1198–1255. https://doi.org/10.1080/02626667.2013.803183

Jowett, I. G., & Duncan, M. J. (1990). Flow variability in New Zealand rivers and its relationship to in-stream habitat and biota. *New Zealand Journal of Marine & Freshwater Research*, *24*(3), 305–317. https://doi.org/10.1080/00288330.1990.9516427

Kavetski, D., Fenicia, F., Reichert, P., & Albert, C. (2018). Signature-domain calibration of hydrological models using approximate Bayesian computation: Theory and comparison to existing applications. *Water Resources Research*, *54*(6), 4059–4083. https://doi.org/10.1002/2017wr020528

Kim, J., Blair, N. E., Ward, A. S., & Goff, K. (2020). Storm-induced dynamics of particulate organic carbon in clear Creek, Iowa: An intensively managed landscape critical zone observatory story. *Frontiers in Water*, *2*, 37. https://doi.org/10.3389/frwa.2020.578261

Kirkby, M. (1988). Hillslope runoff processes and models. *Journal of Hydrology*, *100*(1–3), 315–339. https://doi.org/10.1016/0022-1694(88)90190-4

Klingler, C., Schulz, K., & Herrnegger, M. (2021). LamaH-CE: LArge-SaMple DAta for hydrology and environmental sciences for central Europe. *Earth System Science Data*, *13*(9), 4529–4565. https://doi.org/10.5194/essd-13-4529-2021

Knoben, W. J. M., Woods, R. A., & Freer, J. E. (2018). A quantitative hydrological climate classification evaluated with independent streamflow data. *Water Resources Research*, *54*(7), 5088–5109. https://doi.org/10.1029/2018WR022913

Kumar, P., Le, P. V., Papanicolaou, A. T., Rhoads, B. L., Anders, A. M., Stumpf, A., et al. (2018). Critical transition in critical zone of intensively managed landscapes. *Anthropocene*, *22*, 10–19. https://doi.org/10.1016/j.ancene.2018.04.002

Li, L., DiBiase, R. A., Del Vecchio, J., Marcon, V., Hoagland, B., Xiao, D., et al. (2018). The effect of lithology and agriculture at the Susquehanna Shale Hills critical zone observatory. *Vadose Zone Journal*, *17*, 1–15. https://doi.org/10.2136/vzj2018.03.0063

Lin, H., & Zhou, X. (2008). Evidence of subsurface preferential flow using soil hydrologic monitoring in the Shale Hills catchment. *European Journal of Soil Science*, *59*(1), 34–49. https://doi.org/10.1111/j.1365-2389.2007.00988.x

Lovill, S. M., Hahm, W. J., & Dietrich, W. E. (2018). Drainage from the critical zone: Lithologic controls on the persistence and spatial extent of wetted channels during the summer dry season. *Water Resources Research*, *54*(8), 5702–5726. https://doi.org/10.1029/2017WR021903

Lundquist, J. D., Cayan, D. R., & Dettinger, M. D. (2004). Spring onset in the Sierra Nevada: When is snowmelt independent of elevation? *Journal of Hydrometeorology*, *5*(2), 327–342. https://doi.org/10.1175/1525-7541(2004)005%3C0327:SOITSN%3E2.0.CO;2

Lyon, S. W., Desilets, S. L., & Troch, P. A. (2008). Characterizing the response of a catchment to an extreme rainfall event using hydrometric and isotopic data. *Water Resources Research*, *44*(6). https://doi.org/10.1029/2007wr006259

Markstrom, S. L., Hay, L. E., & Clark, M. P. (2016). Towards simplification of hydrologic modeling: Identification of dominant processes. *Hydrology and Earth System Sciences*, *20*(11), 4655–4671. https://doi.org/10.5194/hess-20-4655-2016

McDonnell, J. J., & Woods, R. (2004). On the need for catchment classification. *Journal of Hydrology*, *299*(1), 2–3. https://doi.org/10.1016/j.jhydrol.2004.09.003

McDowell, W. H., McSwiney, C. P., & Bowden, W. B. (1996). Effects of hurricane disturbance on groundwater chemistry and riparian function in a tropical rain forest. *Biotropica*, *28*(4), 577–584. https://doi.org/10.2307/2389098

McGrath, G. S., Hinz, C., & Sivapalan, M. (2007). Temporal dynamics of hydrological threshold events. *Hydrology and Earth System Sciences*, *11*(2), 923–938. https://doi.org/10.5194/hess-11-923-2007

McMillan, H. (2020). Linking hydrologic signatures to hydrologic processes: A review. *Hydrological Processes*, *34*(6), 1393–1409. https://doi.org/10.1002/hyp.13632

McMillan, H. (2021). A review of hydrologic signatures and their applications. *Wiley Interdisciplinary Reviews: Water*, *8*(1), e1499. https://doi.org/10.1002/wat2.1499

McMillan, H., Gnann, S. J., & Araki, R. (2022). *SebastianGnann/LargeScaleSigs: First release (1.0.0)*. Zenodo. https://doi.org/10.5281/zenodo.6462823

McMillan, H., Gueguen, M., Grimon, E., Woods, R., Clark, M., & Rupp, D. E. (2014). Spatial variability of hydrological processes and model structure diagnostics in a 50 km^2 catchment. *Hydrological Processes*, *28*(18), 4896–4913. https://doi.org/10.1002/hyp.9988

Mosley, M. P. (1979). Streamflow generation in a forested watershed, New Zealand. *Water Resources Research*, *15*(4), 795–806. https://doi.org/10.1029/wr015i004p00795

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., et al. (2015). Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance. *Hydrology and Earth System Sciences*, *19*(1), 209–223. https://doi.org/10.5194/hess-19-209-2015

Noguchi, S., Nik, A. R., Yusop, Z., Tani, M., & Sammori, T. (1997). Rainfall-runoff responses and roles of soil moisture variations to the response in tropical rain forest, Bukit Tarek, Peninsular Malaysia. *Journal of Forest Research*, *2*(3), 125–132. https://doi.org/10.1007/bf02348209

Olden, J. D., & Poff, N. L. (2003). Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River Research and Applications*, *19*(2), 101–121. https://doi.org/10.1002/rra.700

Oshun, J., Dietrich, W. E., Dawson, T. E., & Fung, I. (2016). Dynamic, structured heterogeneity of water isotopes inside hillslopes. *Water Resources Research*, *52*(1), 164–189. https://doi.org/10.1002/2015WR017485

Oudin, L., Andréassian, V., Perrin, C., Michel, C., & Le Moine, N. (2008). Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. *Water Resources Research*, *44*(3). https://doi.org/10.1029/2007wr006240

Peschke, G., Etzenberg, C., Töpfer, J., Zimmermann, S., & Müller, G. (1999). *Runoff generation regionalization: Analysis and a possible approach to a solution* (pp. 147–156). IAHS-AISH Publication.

Peters, N. E., & Aulenbach, B. T. (2011). Water storage at the Panola Mountain research watershed, Georgia, USA. *Hydrological Processes*, *25*, 3878–3889. https://doi.org/10.1002/hyp.8334

Pfister, L., Martínez-Carreras, N., Hissler, C., Klaus, J., Carrer, G. E., Stewart, M. K., & McDonnell, J. J. (2017). Bedrock geology controls on catchment storage, mixing, and release: A comparative analysis of 16 nested catchments. *Hydrological Processes*, *31*(10), 1828–1845. https://doi.org/10.1002/hyp.11134

Poff, N. L., & Ward, J. V. (1989). Implications of streamflow variability and predictability for lotic community structure: A regional analysis of streamflow patterns. *Canadian Journal of Fisheries and Aquatic Sciences*, *46*(10), 1805–1818. https://doi.org/10.1139/f89-228

Prieto, C., Kavetski, D., Le Vine, N., Álvarez, C., & Medina, R. (2021). Identification of dominant hydrological mechanisms using Bayesian inference, multiple statistical hypothesis testing, and flexible models. *Water Resources Research*, *57*(8), e2020WR028338. https://doi.org/10.1029/2020wr028338

Prieto, C., Le Vine, N., Kavetski, D., Fenicia, F., Scheidegger, A., & Vitolo, C. (2022). An exploration of Bayesian identification of dominant hydrological mechanisms in ungauged catchments. *Water Resources Research*, *58*(3), e2021WR030705. https://doi.org/10.1029/2021WR030705

Prieto, C., Vine, N. L., Kavetski, D., García, E., & Medina, R. (2019). Flow prediction in ungauged catchments using probabilistic random forests regionalization and new statistical adequacy tests. *Water Resources Research*, *55*(5), 4364–4392. https://doi.org/10.1029/2018WR023254

Safeeq, M., Grant, G. E., Lewis, S. L., & Tague, C. L. (2013). Coupling snowpack and groundwater dynamics to interpret historical streamflow trends in the western United States. *Hydrological Processes*, *27*(5), 655–668. https://doi.org/10.1002/hyp.9628

Safeeq, M., & Hunsaker, C. T. (2016). Characterizing runoff and water yield for headwater catchments in the Southern Sierra Nevada. *JAWRA Journal of the American Water Resources Association*, *52*(6), 1327–1346. https://doi.org/10.1111/1752-1688.12457

Scaife, C. I., Singh, N. K., Emanuel, R. E., Miniat, C. F., & Band, L. E. (2020). Non-linear quickflow response as indicators of runoff generation mechanisms. *Hydrological Processes*, *34*(13), 2949–2964. https://doi.org/10.1002/hyp.13780

Schellekens, J., Scatena, F. N., Bruijnzeel, L. A., Van Dijk, A., Groen, M. M. A., & Van Hogezand, R. J. P. (2004). Stormflow generation in a small rainforest catchment in the Luquillo Experimental Forest, Puerto Rico. *Hydrological Processes*, *18*(3), 505–530. https://doi.org/10.1002/hyp.1335

Scherrer, S., & Naef, F. (2003). A decision scheme to indicate dominant hydrological flow processes on temperate grassland. *Hydrological Processes*, *17*(2), 391–401. https://doi.org/10.1002/hyp.1131

Schilling, K. E., & Helmers, M. (2008). Effects of subsurface drainage tiles on streamflow in Iowa agricultural watersheds: Exploratory hydrograph analysis. *Hydrological Processes*, *22*(23), 4497–4506. https://doi.org/10.1002/hyp.7052

Schilling, K. E., Streeter, M. T., Bettis, E. A., Wilson, C. G., & Papanicolaou, A. N. (2018). Groundwater monitoring at the watershed scale: An evaluation of recharge and nonpoint source pollutant loading in the Clear Creek Watershed, Iowa. *Hydrological Processes*, *32*(4), 562–575. https://doi.org/10.1002/hyp.11440

Shanley, J. B., McDowell, W. H., & Stallard, R. F. (2011). Long-term patterns and short-term dynamics of stream solutes and suspended sediment in a rapidly weathering tropical watershed. *Water Resources Research*, *47*(7). https://doi.org/10.1029/2010wr009788

Shaw, S. B., & Riha, S. J. (2012). Examining individual recession events instead of a data cloud: Using a modified interpretation of dQ/dt–Q streamflow recession in glaciated watersheds to better inform models of low flow. *Journal of Hydrology*, *434*, 46–54. https://doi.org/10.1016/j.jhydrol.2012.02.034

Sivakumar, B. (2008). Dominant processes concept, model simplification and classification framework in catchment hydrology. *Stochastic Environmental Research and Risk Assessment*, *22*(6), 737–748. https://doi.org/10.1007/s00477-007-0183-5

Sivapalan, M. (2006). Pattern, process and function: Elements of a unified theory of hydrology at the catchment scale. In *Encyclopedia of Hydrological Sciences*. John Wiley & Sons. https://doi.org/10.1002/0470848944.hsa012

Stoelzle, M., Stahl, K., & Weiler, M. (2013). Are streamflow recession characteristics really characteristic? *Hydrology and Earth System Sciences*, *17*(2), 817–828. https://doi.org/10.5194/hess-17-817-2013

Takagi, K., & Lin, H. S. (2012). Changing controls of soil moisture spatial organization in the Shale Hills Catchment. *Geoderma*, *173*, 289–302. https://doi.org/10.1016/j.geoderma.2011.11.003

Tallaksen, L. M. (1995). A review of baseflow recession analysis. *Journal of Hydrology*, *165*(1–4), 349–370. https://doi.org/10.1016/0022-1694(94)02540-R

Tani, M. (1997). Runoff generation processes estimated from hydrological observations on a steep forested hillslope with a thin soil layer. *Journal of Hydrology*, *200*(1–4), 84–109. https://doi.org/10.1016/s0022-1694(97)00018-8

UKIH. (1980). *UK Institute of Hydrology (Great Britain), Low Flow Studies Reports*. Institute of Hydrology.

Vigiak, O., Lutz, S., Mentzafou, A., Chiogna, G., Tuo, Y., Majone, B., et al. (2018). Uncertainty of modelled flow regime for flow-ecological assessment in Southern Europe. *The Science of the Total Environment*, *615*, 1028–1047. https://doi.org/10.1016/j.scitotenv.2017.09.295

Wagener, T., Sivapalan, M., Troch, P., & Woods, R. (2007). Catchment classification and hydrologic similarity. *Geography Compass*, *1*(4), 901–931. https://doi.org/10.1111/j.1749-8198.2007.00039.x

Westerberg, I. K., & McMillan, H. K. (2015). Uncertainty in hydrological signatures. *Hydrology and Earth System Sciences*, *19*(9), 3951–3968. https://doi.org/10.5194/hess-19-3951-2015

Westerberg, I. K., Wagener, T., Coxon, G., McMillan, H. K., Castellarin, A., Montanari, A., & Freer, J. (2016). Uncertainty in hydrological signatures for gauged and ungauged catchments. *Water Resources Research*, *52*(3), 1847–1865. https://doi.org/10.1002/2015wr017635

Wilson, C. G., Abban, B., Keefer, L. L., Wacha, K., Dermisis, D., Giannopoulos, C., et al. (2018). The intensively managed landscape critical zone observatory: A scientific testbed for understanding critical zone processes in agroecosystems. *Vadose Zone Journal*, *17*(1), 180088. https://doi.org/10.2136/vzj2018.04.0088

Wilson, C. G., Papanicolaou, A. N. T., & Denn, K. D. (2012). Partitioning fine sediment loads in a headwater system with intensive agriculture. *Journal of Soils and Sediments*, *12*(6), 966–981. https://doi.org/10.1007/s11368-012-0504-2

Wlostowski, A. N., Molotch, N., Anderson, S. P., Brantley, S. L., Chorover, J., Dralle, D., et al. (2020). Signatures of hydrologic function across the critical zone observatory network. *Water Resources Research*, *57*(3). https://doi.org/10.1029/2019wr026635

Wlostowski, A. N., Molotch, N., Anderson, S. P., Brantley, S. L., Chorover, J., Dralle, D., et al. (2021). Signatures of hydrologic function across the critical zone observatory network. *Water Resources Research*, *57*(3), e2019WR026635. https://doi.org/10.1029/2019wr026635

Wolock, D. M. (2003). *Infiltration-excess overland flow estimated by TOPMODEL for the conterminous United States*. US Geological Survey. Retrieved from https://pubs.er.usgs.gov/publication/ofr03310

Wu, S., Zhao, J., Wang, H., & Sivapalan, M. (2021). Regional patterns and physical controls of streamflow generation across the conterminous United States. *Water Resources Research*, *57*(6). https://doi.org/10.1029/2020wr028086

Yarnell, S. M., Stein, E. D., Webb, J. A., Grantham, T., Lusardi, R. A., Zimmerman, J., et al. (2020). A functional flows approach to selecting ecologically relevant flow metrics for environmental flow applications. *River Research and Applications*, *36*(2), 318–324. https://doi.org/10.1002/rra.3575

Yilmaz, K. K., Gupta, H. V., & Wagener, T. (2008). A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resources Research*, *44*(9). https://doi.org/10.1029/2007wr006716

Zotarelli, L., Dukes, M. D., Romero, C. C., Migliaccio, K. W., & Morgan, K. T. (2010). *Step by step calculation of the Penman-Monteith evapotranspiration (FAO-56 method)*. Institute of Food and Agricultural Sciences, University of Florida.