Signatures of Hydrologic Function Across the Critical Zone Observatory Network

Adam N. Wlostowski1,2, Noah Molotch2,3, Suzanne P. Anderson2,4, Susan L. Brantley5, Jon Chorover6, David Dralle7,8, Praveen Kumar9, Li Li10, Kathleen A. Lohse11, John M. Mallard12, Jennifer C. McIntosh13, Sheila F. Murphy13, Eric Parrish14, Mohammad Safeeq14,15, Mark Seyfried16, Yuning Shi17, and Ciaran Harman18,19

1Lynker Technologies, LLC, Boulder, CO, USA, 2Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, CO, USA, 3Department of Geography, University of Colorado Boulder, Boulder, CO, USA, 4Department of Geological Sciences, University of Colorado Boulder, Boulder, CO, USA, 5Earth and Environmental Systems Institute and Department of Geosciences, Pennsylvania State University, University Park, PA, USA, 6Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA, 7Department of Earth and Planetary Science, University of California Berkeley, Berkeley, CA, USA, 8Pacific Southwest Research Station, US Forest Service, Davis, CA, USA, 9Department of Civil and Environmental Engineering, University of Illinois, Urbana, IL, USA, 10Department of Civil and Environmental Engineering, Pennsylvania State University, University Park, PA, USA, 11Department of Geosciences, Idaho State University, Pocatello, ID, USA, 12Nicholas School of the Environment, Duke University, Durham, NC, USA, 13U.S. Geological Survey, Boulder, CO, USA, 14Sierra Nevada Research Institute, University of California, Merced, CA, USA, 15Pacific Southwest Research Station, United States Forest Service, Fresno, CA, USA, 16Northwest Watershed Research Center, Agricultural Research Service, United States Department of Agriculture, Boise, ID, USA, 17Department of Ecosystem Science and Management, Pennsylvania State University, University Park, PA, USA, 18Department of Environmental Health and Engineering, Johns Hopkins University, Baltimore, MD, USA, 19Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, MD, USA

Abstract Despite a multitude of small catchment studies, we lack a deep understanding of how variations in critical zone architecture lead to variations in hydrologic states and fluxes. This study characterizes hydrologic dynamics of 15 catchments of the U.S. Critical Zone Observatory (CZO) network where we hypothesized that our understanding of subsurface structure would illuminate patterns of hydrologic partitioning. The CZOs collect data sets that characterize the physical, chemical, and biological architecture of the subsurface, while also monitoring hydrologic fluxes such as streamflow, precipitation, and evapotranspiration. For the first time, we collate time series of hydrologic variables across the CZO network and begin the process of examining hydrologic signatures across sites. We find that catchments with low baseflow indices and high runoff sensitivity to storage receive most of their precipitation as rain and contain clay-rich regolith profiles, prominent argillic horizons, and/or anthropogenic modifications. In contrast, sites with high baseflow indices and low runoff sensitivity to storage receive the majority of precipitation as snow and have more permeable regolith profiles. The seasonal variability of water balance components is a key control on the dynamic range of hydraulically connected water in the critical zone. These findings lead us to posit that water balance partitioning and streamflow hydraulics are linked through the coevolution of critical zone architecture but that much work remains to parse these controls out quantitatively.

1. Introduction

The U.S. Critical Zone Observatories (CZOs) were created to advance understanding of the thin layer of the Earth’s surface extending from unweathered bedrock to the top of the vegetation canopy (Brantley et al., 2007; White et al., 2015). CZOs have focused the efforts of teams of scientists working across disciplines by collecting data sets that characterize the physical, chemical, and biological architecture of the critical zone and its fluxes (e.g., Brantley et al., 2016). CZOs are seeking to advance fundamental understanding of how processes interact within this thin epithelium of the Earth, particularly how these processes sustain life and react to changing climate, tectonic, and land use drivers.

The U.S. CZOs were not initially conceived as a network—each was initiated with its own set of science questions and research approaches. However, they offer a unique opportunity to develop a cross-site
synthesis of understanding gained in the first decade of CZO research. Scientific synthesis can promote the
development of broader questions and transferable hypotheses about the interactions between water, rock,
energy, and biota in the critical zone (Brooks et al., 2015). One such question at the core of this pursuit is the
following: how do landscapes store and partition water? The CZOs provide observations that allow investi-
gators to interrogate the relationship between critical zone structure and hydrologic response across a suite
of geologically and climatically diverse sites. Using data from 15 CZO sites, this paper compares catchments
based on hydrologic function (Black, 1997; Wagener et al., 2007)—the partitioning, storage, and release
water. In doing so, this work documents initial patterns of hydrologic variability across CZO catchments,
laying a foundation for future cross-site investigations.

The partitioning, storage, and transmission of water are central landscape functions with numerous direct
interconnections to other critical zone processes, including those that support human societies
(Black, 1997; Loomis et al., 2000; Poff et al., 1997; Wagener et al., 2008). For example, the quantity of water
available for ecological, industrial, and municipal use depends on the partitioning of precipitation between
evapotranspiration, runoff, and groundwater recharge (Bales et al., 2006, 2018). Similarly, the timing and
quantity of runoff depend on the orientation of flow paths responsible for runoff generation (Jencso
et al., 2010; Payn et al., 2012). The chemical quality of surface water and groundwater depends on the resi-
dence time and composition of hydrologic flow paths (Maher, 2011). Plant ecosystem structure and product-
vity depend on soil water storage (Eagleson, 1982; Graham et al., 2010). Hydrologically mediated landscape
functions such as these are relevant over short timescales (days to decades) and largely depend on the inter-
play of climate and landscape architecture (i.e., soils, regolith, and topographic properties) (Field et al., 2015;
Tague & Grant, 2009).

Over longer timescales (decades to millennia), water partitioning, storage, and transmission paces the evo-
lution of landscapes. Water is a primary control on the spatial distribution and properties of soil and regolith
(Jenny, 1941). The conversion of fresh bedrock to soil requires meteoric water to chemically and physically
break down the rock (Brantley & Lebedeva, 2011; Maher, 2010; Rempe & Dietrich, 2014). Conversely, fluvial
incision, itself affected by water transmission, affects the distribution of soil and regolith and the morphol-
ogy of hillslopes (Deal et al., 2018). Hydrologic processes also modify soil and regolith drainage characteris-
tics (Jefferson et al., 2010). For example, the development of subsurface clay horizons affects the orientation
of flow paths by promoting the development of shallow, transient, and perched water tables (Lohse &
Dietrich, 2005; Zimmer & McGlynn, 2017). Further, the development of reaction fronts in the subsurface
may have strong control on the orientation of deeper flow paths contributing to baseflow (Brantley
et al., 2017).

An underlying but implicit hypothesis of the CZOs was that intense measurements of the entire CZ at a few
locations would illuminate connections between landscape structure and hydrologic function (Brantley
et al., 2017; Jefferson et al., 2014; Kumar et al., 2018; Lohse & Dietrich, 2005; Troch et al., 2015; Yoshida
& Troch, 2016). And yet, putting together the disparate measurements at all CZOs has remained elusive
at least partly because of the many differences in sites. Here we make a first foray toward understanding
and simulating the processes underpinning landscape-hydrological interactions in the hopes that our effort
will ultimately foster new approaches for managing water resources (Kumar, 2011; Sivapalan &
Blöschl, 2017). Cross-site synthesis is needed to unpack the interactions between landscape structure and
hydrologic response, but cross-site comparisons are inherently difficult because of differences in definitions
and measurements of important landscape and hydrologic features. Large-sample synthesis studies have
sought to build understanding by linking hydrologic signatures to catchment properties (Addor et al., 2017,
2018; Berghuijs et al., 2014, 2016; Sawicz et al., 2011; Sivapalan et al., 2011). We build on this work by focusing
on a smaller set of sites with a more detailed documentation of their respective subsurface architectures.

The goals of this paper are to make three key initial steps toward understanding the interactions between
critical zone structure and hydrologic function. First, we completed a large effort to compile a common
set of hydrologic data for 15 catchments from across the U.S. CZO network for synthesis here and, hopefully,
for others in the future. Second, we extracted quantitative signatures of hydrologic function to characterize
how they partition, store, and release water. Finally, we examined variability across the sites in the network
seeking to highlight patterns. By summarizing cross-network variability in hydrologic behavior, the data sets
and analyses presented here contribute to a first small step toward the development of generalizable theories
for understanding the partitioning, storage, and transmission of water in the critical zone.
2. The U.S. Critical Zone Observatories

This study considers 15 catchments located across the U.S. CZO network where previous critical zone investigations have been conducted (Figure 1 and Table 1). The collection of catchments is geomorphically, ecologically, and geologically diverse (Table 1). Catchment drainage area varies by over 4 orders of magnitude and median catchment elevation spans nearly 3,000 m. Basin relief ranges from 35 to 966 m and median slope ranges from 1% to 24%. Catchment bedrock composition includes metamorphic, igneous, and sedimentary rock types ranging in age from Precambrian to Quaternary. Land cover, dominant vegetation species, and land use history exhibit similar variation across catchments. For a more detailed review of observatory landscape properties and characteristics, see White et al. (2015).

3. Catchment Storage as Synthesis

Catchment storage dynamics are a logical basis for synthesizing hydrologic function across a set of catchments (McNamara et al., 2011; Staudinger et al., 2017). The filling and spilling of catchment storage space are both a function of contemporary landscape architecture and a characteristic that drives landscape evolution (Beven et al., 1988). However, storage is also largely invisible. The general lack of subsurface characterization available in many catchments hampers our ability to draw meaningful linkages between catchment structure and hydrologic behavior. Catchment water storage can typically only be observed indirectly through hydrologic analysis.

Figure 1. National map of the conterminous United States showing the location of each Critical Zone Observatory (CZO) and subset digital elevation models depicting individual catchments at each CZO considered in this study. Contour lines on catchment maps are at 20-m intervals. Throughout the text, figures, and tables, catchments are referred to by the abbreviations adjacent to CZO and catchment names on this map. For example, Gordon Gulch at the Boulder Creek CZO is referred to as BC (observatory) GG (catchment), and Johnston Draw at Reynolds Creek CZO is referred to as RC (observatory) JD (catchment).
Table 1  
Topographic, Vegetation, and Geologic Characteristics of 15 Catchments Across the Critical Zone Observatory Network

<table>
<thead>
<tr>
<th>Observatory</th>
<th>Catchment name</th>
<th>Area (km²)</th>
<th>Median elevation (m)</th>
<th>Relief (m)</th>
<th>Median slope (%)</th>
<th>Depth to fresh bedrock (m)</th>
<th>Dominant vegetation</th>
<th>Geologic parent material</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boulder Creek (BC)</td>
<td>Gordon Gulch (GG)</td>
<td>2.77</td>
<td>2,638</td>
<td>270</td>
<td>14</td>
<td>10–14d</td>
<td>Mixed conifer forest, predominantly lodgepole and ponderosa pine</td>
<td>Metamorphic (Precambrian gneiss)</td>
</tr>
<tr>
<td>Shale Hills (SH)</td>
<td>Shale Hills (SH)</td>
<td>0.08</td>
<td>283</td>
<td>30</td>
<td>30</td>
<td>16–44a</td>
<td>Oak, hickory, and pine forest</td>
<td>Sedimentary (Silurian shale)</td>
</tr>
<tr>
<td>Eel River (ER)</td>
<td>Elder Creek (EC)</td>
<td>16.85</td>
<td>828</td>
<td>867</td>
<td>29</td>
<td>4.5–30c</td>
<td>Douglas-fir, tan oak, and Pacific madrone mixed forest</td>
<td>Sedimentary (Mesozoic argillite)</td>
</tr>
<tr>
<td>Jemez River Basin (JRB)</td>
<td>Upper Jaramillo (UJ)</td>
<td>3.05</td>
<td>2,919</td>
<td>601</td>
<td>12</td>
<td>27–34d</td>
<td>Spruce-fir forest and woodland</td>
<td>Extrusive igneous (Quaternary rhylotic tuff)</td>
</tr>
<tr>
<td>Santa Catalina Mountains (SCM)</td>
<td>La Jara (LJ)</td>
<td>3.68</td>
<td>3,086</td>
<td>727</td>
<td>15</td>
<td>42–49d</td>
<td>Mixed conifer and deciduous (oak and maple) forest</td>
<td>Intrusive igneous and metamorphic (Tertiary granite and schist)</td>
</tr>
<tr>
<td></td>
<td>Oracle Ridge (OR)</td>
<td>0.86</td>
<td>2,265</td>
<td>403</td>
<td>21</td>
<td>N/A</td>
<td>Ponderosa pine transitioning to pinyon-juniper-oak forest from high to low elevation</td>
<td>Intrusive igneous (Tertiary quartz monzonite)</td>
</tr>
<tr>
<td>Luquillo (LUQ)</td>
<td>Rio Mameyes (MAM)</td>
<td>17.80</td>
<td>508</td>
<td>967</td>
<td>22</td>
<td>27.7–37.2e</td>
<td>Tabunoco, Palm, and Colorado forest</td>
<td>Sedimentary (Cretaceous volcanioclastic)</td>
</tr>
<tr>
<td></td>
<td>Rio Icacos (IC)</td>
<td>3.26</td>
<td>686</td>
<td>212</td>
<td>14</td>
<td>N/A</td>
<td>Colorado and Palm forest</td>
<td>Intrusive igneous (Tertiary quartz diorite)</td>
</tr>
<tr>
<td>Reynolds Creek (RC)</td>
<td>Johnston Draw (JD)</td>
<td>1.83</td>
<td>1,747</td>
<td>374</td>
<td>14</td>
<td>13–31f</td>
<td>Sagebrush grassland with intermittent aspen and juniper stands</td>
<td>Intrusive igneous (Cretaceous granite)</td>
</tr>
<tr>
<td>Southern Sierra (SS)</td>
<td>P301</td>
<td>0.99</td>
<td>1,990</td>
<td>318</td>
<td>12</td>
<td>10–35g</td>
<td>Mixed conifer forest</td>
<td>Intrusive igneous (Cretaceous granodiorite)</td>
</tr>
<tr>
<td></td>
<td>P303</td>
<td>1.26</td>
<td>1,906</td>
<td>292</td>
<td>14</td>
<td>N/A</td>
<td>Corn and soybean agricultural rotation</td>
<td>Sedimentary (Silurian to Carboniferous dolomite, shale, limestone, and sandstone beneath thick glacial drift deposits)</td>
</tr>
<tr>
<td></td>
<td>P304</td>
<td>0.49</td>
<td>1,909</td>
<td>213</td>
<td>14</td>
<td>N/A</td>
<td>Pine-hardwood forest</td>
<td>Intrusive igneous (Silurian metagranite)</td>
</tr>
<tr>
<td>IML (IML)</td>
<td>Upper Sangamon (US)</td>
<td>615.73</td>
<td>232</td>
<td>88</td>
<td>1</td>
<td>100h</td>
<td>Tabonuco, Palm, and Colorado</td>
<td>Intrusive igneous (Silurian metagranite)</td>
</tr>
</tbody>
</table>

*Depth of fresh bedrock observed in deep boreholes situated along ridges and valleys of the Shale Hills catchment (Sullivan et al., 2016 and unpublished observations).  
Seismic refraction surveys along four north-south transects in Gordon Gulch, where the weathered-fresh bedrock boundary was inferred to reside at the 3,500 m/s P wave velocity contour (Befus et al., 2011).  
Depth to fresh bedrock observed in boreholes throughout catchment, deepest depth to fresh bedrock is beneath ridgetops and shallowest depth to bedrock beneath valley bottoms (Hahm et al., 2019; Rempe & Dietrich, 2018; Salve et al., 2012).  
Seismic refraction surveys along four north-south oriented transects, where the weathered-fresh bedrock boundary is inferred to reside at the 2.95 km/s P wave velocity contour (Olyphant et al., 2016).  
Depth to what is inferred to be “relatively unweathered bedrock” in two boreholes (Buss et al., 2013).  
Seismic refraction surveys along four north-south transects in Las Jara and Upper Jaramillo catchments, where the regolith to unweathered bedrock transition was inferred to occur at the 2.95 km/s P wave velocity contour (Olyphant et al., 2016).  
Depth to bedrock is deepest beneath ridges and shallowest beneath valley bottoms.

Dynamic storage refers to indirectly estimated water storage capacity that is actively filled and drained over timescales relevant to annual input and output fluxes (McNamara et al., 2011). Dynamic storage is inferred by analyzing water fluxes (precipitation, runoff, and evapotranspiration) under a water balance framework. We propose two perspectives on catchment-scale dynamic water storage to serve as a framework for the synthesis study: terrestrial and hydraulic storage (Carrer et al., 2018; Dralle et al., 2018).

### 3.1. Terrestrial Storage

Terrestrial storage, $s_{\text{terr}}$ [L], is the volume of water stored beneath the land surface as inferred by evaluating a catchment-wide water balance over some timescale. Changes in $s_{\text{terr}}$ are controlled by rates of surface water input ($SWI$ [L/T]), runoff ($Q$ [L/T]), evapotranspiration ($ET$ [L/T]), and net groundwater flows ($GW$ [L/T]).

$$\frac{ds_{\text{terr}}}{dt} = SWI - Q - ET - GW.$$  

10.1029/2019WR026635

WLOSTOWSKI ET AL.
SWI is the input of liquid water from aboveground to terrestrial storage. In this conceptual model, we assume that snowpack is an aboveground storage reservoir distinct from \( s_{\text{terr}} \). Other aboveground storage reservoirs, such as canopy interception and depression storage, are assumed to be negligible compared to the catchment’s total dynamic storage. While this is a major conceptual simplification, it is justified at 11 of 15 CZO catchments due to a substantial influence of snow (e.g., Tennant et al., 2017). In cold regions, the snowpack decouples the timing and intensity of SWI from precipitation \((P \ [L/T])\) because accumulated snow is not immediately available for infiltration but is retained until the melt period.

We define SWI as the sum of snow/ice melt, \( M \ [L/T] \), rainfall on snow, \( P_S \ [L/T] \), and direct rainfall, \( P_D \ [L/T] \), which falls directly onto the land surface and is not retained by the snowpack:

\[
SWI = M + P_S + P_D
\]  

(2)

At any time, total precipitation, \( P \ [L/T] \), is the sum of surface accumulated precipitation as ice/snow, \( P_{\text{acc}} \ [L/T] \), \( P_S \), and \( P_D \).

\[
P = P_D + P_{\text{acc}}
\]  

(3)

In catchments with no snow/ice accumulation (i.e., \( P_{\text{acc}} = 0 \) and \( P_S = 0 \)), \( SWI = P_D = P \).

For convenience, we assume that net groundwater exchange \((GW)\) is a negligible term in Equation 1. This assumption is unlikely to hold for most of the catchments considered in this study—small headwater catchments nested in much larger regional groundwater systems (Fan, 2019; Tóth, 1999). To accurately quantify net groundwater flows in CZO catchments, a more detailed synthesis of hydrogeologic data is needed. In absence of such a synthesis, we move forward with this simplifying assumption and discuss the implications for study results.

Finally, \( s_{\text{terr}} \) must be defined relative to some initial (and unknown) storage state, \( s_0 \), and is evaluated as a function of time by integrating 1 from some initial time \((t = 0)\) to time \(t\).

\[
s_{\text{terr}} - s_0 = \int_{0}^{t} (SWI - Q - ET) \, dt
\]  

(4)

3.2. Hydraulic Storage

Hydraulic storage, \( s_h \ [L] \), is defined as the storage that drives streamflow generation. Specifically, changes in \( s_h \) are hydraulically linked to streamflow through an assumed 1-to-1 relationship. Hydraulic storage is determined from the recession behavior of the streamflow hydrograph using methods developed by Kirchner (2009) and Wittenberg (1999). Dralle et al. (2018) refer to this as “direct storage,” as it is directly affiliated with streamflow generation. Hydraulic storage is a subset of \( s_{\text{terr}} \) because not all stored water is hydraulically linked to flow generation. For example, changes in unsaturated vadose zone storage may have no influence on streamflow generation.

Here we present a formulation of \( s_h \) slightly distinct from previous works (e.g., Kirchner, 2009) in the way the role of evapotranspiration is handled. We will assume that the rate of change in \( s_h \) with respect to time is controlled by water fluxes into and out of \( s_h \) zones:

\[
\frac{ds_h}{dt} = R - Q - \alpha ET.
\]  

(5)

Hydraulic storage zones are filled by recharge, \( R \ [L/T] \), and drained by \( Q \) and \( ET \). Recharge is conceptually different from SWI, because not all SWI fills hydraulically active storage zones (e.g., the filling of unsaturated zones). Further, not all \( ET \) affects hydraulic storage; rather, only a fraction of total catchment \( ET \), \( \alpha \ [\%] \), is drawn from hydraulic storage. The influence of evapotranspiration on runoff dynamics is apparent in diel streamflow oscillations and inflated streamflow recession rates as shown by Bart and Hope (2014), Wang and Cai (2010), and Weisman (1977).

Unlike \( s_{\text{terr}} \), it is difficult to directly resolve \( s_h \) with a water balance approach because \( R \) is difficult to observe. Instead, recession curve analysis is used to estimate \( s_h \) (Kirchner, 2009; Krier et al., 2012; Teuling et al., 2010). This approach hinges on the assumption that runoff is a single-valued and invertible function of storage.
\[ Q = f(s_h) \]  

Using the chain rule to differentiate 6 with respect to time and substituting 5 gives an expression for the time rate of change in streamflow:

\[ \frac{dQ}{dt} = \frac{dQ}{ds_h} \frac{ds_h}{dt} = g(Q)(R - Q - \alpha ET), \]

where \( g(Q) = \frac{dQ}{ds_h} [1/\text{day}] \) is the catchment sensitivity function, which quantifies the sensitivity of \( Q \) to changes in \( s_h \). Rearranging 7 to solve for \( g(Q) \) yields

\[ g(Q) = \frac{dQ/dt}{R - Q - \alpha ET}. \]

During periods of extended baseflow recession, it is reasonable to assume that \( R << Q \) following the redistribution of water from the vadose zone to the water table. During these periods, \( g(Q) \) may be directly evaluated from discharge observations using 8 if estimates of \( \alpha ET \) can be obtained. Once \( g(Q) \) is known, the hydraulic storage-discharge relationship, \( s_h(Q) \), is estimated by integrating the reciprocal of \( g(Q) \) with respect to \( Q \).

\[ s_h - s_0 = \int_{Q_0}^{Q} \frac{1}{g(Q)} dQ \]

### 4. Data and Methods

#### 4.1. Data Sets Used

A standard set of hydrologic data and estimates (daily precipitation, discharge, potential and actual evapotranspiration, snow water equivalent [SWE], and melt rate) was compiled for the 15 study catchments across the CZO network listed in Table 1. Those data are available for download through the Consortium of Universities for the Advancement of Hydrologic Science, Inc (CUAHSI) HydroShare (Wlostowski et al., 2021). Here we describe how those data were compiled.

Volumetric streamflow data \([L^3/T]\) was obtained from stream gages within each catchment. Detailed information on the location, period of record, drainage area, instrumentation, and data source for each gage is provided in Table S1. Volumetric streamflow is converted to daily runoff \([L/T]\) by normalizing data by catchment area and aggregating subdaily data to a daily time step. At four catchments (LUQ IC, LUQ MAM, EEL EC, and IML US), streamflow is monitored by gages maintained by the U.S Geological Survey (USGS). At these sites, annual runoff is computed by converting USGS estimates of mean annual discharge by normalizing to catchment area. At the remaining sites, annual runoff is calculated by numerical integration of daily runoff data. As such, any gaps in the daily runoff record will lead to underestimates in annual runoff. The variable \( Q \) is used to denote daily runoff with units of millimeters per day.

Precipitation data were collated from 27 gages and two spatially distributed precipitation data products (Kormos et al., 2017; Murphy et al., 2017). The location, period of record, and data source for each gage (or data product) are specified in Table S2. By combining available precipitation observations within (or near) each catchment, catchment average daily precipitation intensity is approximated for each site. Table S2 articulates the methodological approach for estimating catchment average daily precipitation intensity at each site. Hereinafter, \( P \) is used to denote daily precipitation intensity in units of millimeters per day. The standard deviation of annual precipitation was calculated from the 8-km PRISM product (1895–2019) (PRISM Climate Group, 2019), except at LUQ IC and LUQ MAM, where the available period of record presented in Murphy et al. (2017) was used.

Due to high spatial rainfall variability in the Luquillo Mountains (Murphy et al., 2017) and the lack of long-term daily precipitation gages within the two small Luquillo CZO catchments, a different method of estimating annual and daily precipitation totals was needed. Mean annual precipitation (MAP) was
obtained from Murphy et al. (2017); for LUQ MAM, MAP was estimated by a spatial interpolation of all existing rainfall data, but due to probable underestimation of rainfall in LUQ IC, MAP was estimated from a linear model of annual runoff versus annual precipitation for many tropical forests across the world (Murphy et al., 2017). For LUQ IC and LUQ MAM, daily precipitation at a nearby rain gage (El Verde, <8 km away) was summed to obtain MAP for the period 1995–2015 and scaled to the modeled MAP for each watershed to estimate annual precipitation totals. To approximate daily rainfall intensities for each catchment, daily observations at the El Verde rain gage were scaled by the ratio of annual precipitation in the Rio Mameyes or Rio Icacos to annual precipitation at El Verde. Short-term spatial variability in daily precipitation might cause overprediction/underprediction of daily rainfall intensity in each catchment.

Data from two gridded meteorological products are used in this study. North America Land Data Assimilation System (NLDAS-2) longwave and shortwave radiation, surface pressure, specific humidity, 2-m air temperature, and 10-m wind speed data are used for all catchments located within the continental United States (Mitchell, 2009). NLDAS-2 forcing data have a 1/8° spatial resolution and hourly temporal resolution. Data were acquired for all grid cells covering the extent of each catchment. If a catchment extent spans more than one grid cell, the average value of each variable across the total number of grid cells is used. For catchments outside the continental United States (i.e., those in Puerto Rico), Modern-Era Retrospective analysis for Research Applications (MERRA2) data are used. Longwave and shortwave radiation (Global Modeling and Assimilation Office [GMAO], 2015b), air temperature, wind speed, pressure, and specific humidity data (GMAO, 2015a) were obtained at an hourly time step for the single 0.5° × 0.625° grid cell containing both of the Puerto Rico catchments, LUQ IC and LUQ MAM. We opted to use NLDAS-2 and MERRA2 data products in lieu of on-the-ground meteorological observations to enhance methodological consistency and temporal coverage. However, NLDAS-2 and MERRA2 data were bias corrected using available meteorological observations at each site.

Seasonal SWE accumulation/ablation and SWI rates are simulated with the SNOW17 model (Anderson, 2006). SNOW17 is a spatially lumped (i.e., assumes uniform spatial distribution of SWE across a catchment) snowpack model that simulates melt rates with a seasonally variable temperature index approach (Hock, 2003). The model is calibrated to match SWE observations at or near each catchment (Table S3). A shuffled-complex evolution algorithm (Duan et al., 1993) is used to exhaustively search the model parameter space and identify near-optimal parameter sets that minimize the root mean squared error between simulated and observed SWE. Where SWE observations are not available (SH SH and IML US), a standard parameterization is used. Model parameter values are presented in Table S3. Once parameter values are established, forward simulations estimate catchment average daily SWE and melt fluxes for each catchment. This spatially lumped approach to SWE modeling is admittedly simplistic, especially at sites where snow cover is spatially and temporally intermittent (e.g., RC and BC). Nevertheless, this approach offers a first-order approximation of basin-average SWE accumulation/ablation dynamics.

MODIS leaf area index (LAI) data (Myneni et al., 2015), which have a 500 × 500 m spatial resolution and 8-day temporal resolution, were obtained for grid cells covering each catchment. Mean LAI of all grid cells covering a catchment is calculated for every 8-day observation and linearly interpolated to a daily time step. Snow depth and/or SWE data were collated from five CZO sites. Data set names, locations, periods of record, and data sources are provided in Table S4. Digital elevation models are used to characterize the land surface topography of each catchment. Details and data sources of surface topography data sets can be found in Table S4. Soil horizon, depth, and texture data were obtained for each catchment from the Digital General Soil Map of the United States (STATSGO2) (Soil Survey Staff, Natural Resources Conservation Service, U. S. D. of A, 2018).

4.2. Estimating Actual Evapotranspiration

Actual evapotranspiration, $ET_a$ (mm/day), is estimated by scaling surface-dependent potential evapotranspiration, $PET_{surf}$ (mm/day), by the ratio of the average deficit between annual precipitation, $P_{ann}$ (mm), and annual runoff, $Q_{ann}$ (mm), to average annual $PET_{surf}$. 
Hydraulic storage is calculated from the catchment sensitivity function, \( \frac{P_{ann} - Q_{ann}}{PET_{surf,ann}} \) (10), where \( PET_{surf,ann} \) is total annual surface-dependent potential evapotranspiration (mm). \( PET_{surf} \) is calculated using a surface-dependent Penman-Monteith model (Monteith, 1973), forced with daily average meteorological data (section 5.1). Aerodynamic and canopy resistance terms in the model are dependent on land surface characteristics. Specifically, aerodynamic resistance is a function of vegetation height (Monteith, 1965), and canopy resistance is assumed to be a function of LAI (Shuttleworth & Wallace, 1985). Following Shuttleworth and Wallace (1985), mean stomatal conductance was set to 400 s/m at all sites for simplicity. Differences in vegetation between sites likely render a spatially constant value of stomatal conductance inappropriate, though we proceed for convenience’s sake.

This approach for estimating \( ET_a \) assumes that actual evapotranspiration is approximately equal to the long-term mean deficit between \( P \) and \( Q \) (Budyko, 1974). Furthermore, this approach assumes that the seasonality of \( ET_a \) is identical to the seasonality of \( PET_{surf} \). These assumptions are least plausible for catchments with substantial groundwater gains/losses, short periods of record focused on anomalously wet or dry periods, and/or catchments with significant intra-annual variations in water/energy limitation.

### 4.3. Hydroclimatic Indices

The aridity index (\( AI \)) is calculated as the ratio of annual reference land cover potential evapotranspiration, \( PET_{ref,ann} \) (mm), to \( P_{ann} \) (Budyko, 1974).

\[
AI = \frac{PET_{ref,ann}}{P_{ann}}
\]

(11)

\( PET_{ref} \) (mm/day) is calculated from a Penman model with a “reference crop” run at a daily time step (Penman, 1947), forced by daily average meteorological data. Unlike \( PET_{surf} \), between-site variations in \( PET_{ref} \) are only due to climatic differences because \( PET_{ref} \) assumes that land cover is homogeneous across sites (reference crop), making \( AI \) a purely climatic metric.

The seasonality index, \( \phi \), summarizes the relative timing and magnitude of terrestrial water supply (\( SWI \)) and evaporative demand (\( PET \)) (Woods, 2009):

\[
\phi = A_{SWI} \text{sgn}(A_{PET}) \frac{\cos(2\pi(\theta_{SWI} - \theta_{PET}))}{365},
\]

(12)

where \( A_{SWI} \) and \( \theta_{SWI} \) are the amplitude and time offset of mean annual \( SWI \) variation and \( A_{PET} \) and \( \theta_{PET} \) are the amplitude and time offset of mean annual \( PET \) variation. \( A_{SWI}, \theta_{SWI}, A_{PET}, \theta_{PET} \) are determined by fitting sinusoidal models through time series of mean \( SWI \) and \( PET \) for each day of the water year (1 October through 30 September). Sinusoidal models of annual \( SWI \) and \( PET \) variation are denoted as \( SWI' \) and \( PET_{ref}' \), respectively:

\[
SWI'(t) = SWI[1 + A_{SWI}\sin(2\pi(t - \theta_{SWI})/365)]
\]

(13)

\[
PET_{ref}'(t) = PET_{ref}[1 + A_{PET,ref}\sin(2\pi(t - \theta_{PET,ref})/365)].
\]

(14)

The seasonality index, \( \phi \), is strongly sensitive to the phase lag between water supply (\( SWI \)) and demand (\( PET_{ref} \)). When supply and demand are out of phase, \( \phi \) is negative. Conversely, when supply and demand are in-phase, \( \phi \) is positive. Additionally, the magnitude of \( \phi \) is sensitive to \( A_{SWI} \). Catchments with no seasonal \( SWI \) variation (i.e., same \( SWI \) intensity every day of the year) have \( A_{SWI} = 0 \) and \( \phi \) equal to zero.

Automated baseflow separation is used to partition hydrographs into baseflow and storm flow. The baseflow index (\( BFI \)) is the ratio of baseflow to total streamflow over the available period of record. We partition baseflow from observed hydrographs using a single-pass digital filter technique (Arnold & Allen, 1999).

### 4.4. Hydraulic Storage Calculation

Hydraulic storage is calculated from the catchment sensitivity function, \( g(Q) \), which is derived from streamflow recession observations using 8. Other analyses of this type (Kirchner, 2009) use recession observations...
where $ET_a << Q$, and therefore, $ds_h/dt \approx -Q$. Further, $g(Q)$ is estimated as a continuous function of $Q$ by fitting an empirical model through recession plot observations ($-dQ/dt$ v. $Q$) and dividing the model by $Q$. However, in arid and semi-arid catchments, $ET_a$ is rarely much less than $Q$ leaving little (or no) observations remaining from which to derive $g(Q)$. In 10 of the 15 catchments considered in this work, $ET_a$ is much greater than $Q$ for over 90% of all $Q$ observations. At these catchments, it is critically necessary to account for the influence of $ET$ on $g(Q)$.

To account for the influence of $ET$ on $g(Q)$, we compute an $ET$-adjusted recession rate, $\varepsilon$ (mm/day$^2$):

$$\varepsilon = \frac{-dQ/dt}{Q + \alpha ET},$$

(15)

where $\alpha$ is the fraction $ET_a$ drawn from $s_h$. If $\alpha ET_a$ is negligible relative to $Q$, $\varepsilon$ reduces to $-dQ/dt$. Physically, we might expect this to be the case when water table levels are lower than the rooting depth of catchment vegetation. However, if $\alpha ET_a$ is comparable in magnitude to $Q$, then the $\varepsilon$ is less than $-dQ/dt$. We might expect this to be the case when water table levels intersect the rooting zone, and plant water use is drawn from the same storage compartments involved in runoff generation. In this work, we arbitrarily set $\alpha = 0.25$ and test the sensitivity of $s_h$ to $\alpha$ by setting $\alpha$ to end-member values, 0 and 1.

We define recession periods as times when $SWI = 0$, simulated $SWE = 0$, and the recession duration is at least 4 days (Figure 2a). For all daily recession period observations, $-dQ/dt$ is estimated with a 3-day sliding window regression and $\varepsilon$ is calculated using $\alpha$ values are plotted against daily $Q$ observations to create $ET$-adjusted recession plots (Figure 2b). Following Kirchner (2009), daily $\varepsilon$ values are grouped into 100 logarithmically distributed $Q$ bins. Bins are merged with adjacent bins until each bin has at least three $\varepsilon$ estimates and the standard error of $\varepsilon$ in any bin is less than half of the mean of $\varepsilon$ in that bin. Binned $\varepsilon$ values ($\varepsilon_{fit}$) are plotted as red markers in Figure 2b.

A novel dynamic power law function, $\varepsilon_{fit}$, is fit through the binned $\varepsilon$ estimates to empirically describe the variation of $\varepsilon$ as function of $Q$ (Figure 2b; Equations 16 and 17).

$$\varepsilon_{fit} = a \left( \frac{Q}{Q_{min}} \right)^{b(Q)},$$

(16)

where

$$b(Q) = b_L + (b_U + b_L) \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{\ln(Q) - \ln(Q_{min})}{\ln(\sigma)/\sqrt{2}} \right) \right].$$

(17)

The power law exponent, $b$, in 17 is a sigmoidal function of $Q$ (Equation 17), ranging from $b_L$ and $b_U$ following a lognormal cumulative distribution of $Q$ with mean and standard deviation $\ln(\sigma)$ (Figure 3a). When $b_L = b_U$, $\ln(\varepsilon_{fit})$ v. $\ln(Q)$ plots as a straight line. When $b_U > b_L$, the slope of $\ln(\varepsilon_{fit})$ v. $\ln(Q)$ is variable, steeper at high flows, and more gradual at low flows (Figure 3b). The variability of $b$ makes 16 and 17 sufficiently flexible to fit nonlinear trends observed commonly in log-log recession plots.

The catchment sensitivity function is finally derived as

$$g(Q) = \varepsilon_{fit}.$$  

(18)

The relationship between hydraulic storage and $Q$ is derived by integrating the reciprocal of $g(Q)$ from the minimum nonzero $Q (Q_{min})$ to the 99th percentile flow rate ($Q_{99}$) and appointing an arbitrary initial storage state ($s_h$) (Figure 2c).

$$s_h(Q) = \int_{Q_{min}}^{Q_{99}} \frac{1}{g(Q)} dQ + s_h(Q_{min}).$$

(19)

The absolute value of $s_h$ and $s_a$ cannot be known, but $s_h$ was selected such that the median of $s_h$ was zero. The 99th percentile flow rate is used as the upper integration boundary instead of the peak flow, because
Figure 2. (a) Streamflow recession periods are extracted from the hydrograph by finding times when there is no snow on the ground, SWI = 0, and streamflow continuously recedes for at least 3 days. (b) A dynamic power law function describes the relationship between ln(ε) and ln(Q) by fitting ln(ε<sub>bin</sub>) and ln(Q<sub>bin</sub>). (c) The hydraulic storage-discharge relationship is calculated by integrating the reciprocal of g(Q) from Q<sub>min</sub> > 0 to Q<sub>99</sub>, informed and uniformed parts of the relationship are shown. (d) Hydraulic storage-discharge relationships are used to transform Q(t) to sh(t). Hydraulic storage estimates are not made for Q > Q<sub>99</sub>. The data shown in this figure are from a single water year at Boulder Creek’s Gordon Gulch catchment.

extremely high flows are usually associated quick-flow processes that bypass storage zones (e.g., overland flow). Equation 19 is used to map Q(t) to sh(t) (Figure 2d). It is important to note that sh(Q) is only informed by recession Q data, a small subset of all the Q data indicated by a red line in Figure 2c. Estimating sh beyond the range of recession Q requires extrapolation of ε<sub>fit</sub>.

Figure 3. (a) Power law exponent, b, is a lognormal cumulative distribution function of ln(Q). In the example shown, b<sub>L</sub> < b<sub>U</sub> resulting in an increasing cumulative distribution from low to high Q. (b) Changes in b as a function of ln(Q) lead to nonlinear ln(ε) v. ln(Q) shapes. Dashed lines show ln(ε) v. ln(Q) shapes resulting if b = b<sub>U</sub> and b = b<sub>L</sub>.
5. Results

5.1. Annual Water Partitioning and Hydroclimatic Seasonality

MAP ($P_{ann}$), runoff ($Q_{ann}$), reference-surface potential evapotranspiration ($PET_{ref, ann}$), and actual evapotranspiration ($ET_{a, ann}$) vary strongly among network catchments (Figure 4 and Table 2). MAP (Table 2 and Figure 4a) varies by an order of magnitude from 546 mm/yr at SCM OR (see Table 1 for catchment abbreviations) to 5,313 mm/yr at LUQ IC. Mean annual runoff (Table 2 and Figure 4b) varies from 56 mm/yr at JRB LJ to 4,151 mm/yr at LUQ IC. Mean annual reference-surface potential evapotranspiration (Table 2 and Figure 4c) ranges from 1,002 mm/yr at SH SH to 1,889 mm/yr at SC MG and 1,854 mm/yr at SC OR. Mean annual actual evapotranspiration, as approximated by an interannual water balance (Table 2), spans from 376 mm/yr at SCM MG to 1,162 mm/yr at LUQ IC.

It is important to note that annual $ET_a$ estimates are solely based on the long-term difference between $P$ and $Q$ (Equation 10, Table 2) and will therefore be biased by groundwater gains/losses or multiyear changes in net storage. The consequence of this assumption becomes apparent by comparing annual $ET_a$ estimates from this work to independently derived estimates of actual evapotranspiration synthesized from the literature (Table 2, $ET_a$-literature). For example, flux tower derived annual $ET_a$ at SS range from 500 to 800 mm/yr (Bales et al., 2018), which are much greater than $ET_a$ estimated by 10. Such a discrepancy may be caused by shallow or deep groundwater losses, which are not explicitly accounted for in the water balance approach used in this study.

The aridity index ($AI$) is a first-order control on long-term mean annual water partitioning, as illustrated by Figure 5. Network catchments occupy a wide range of $AI$, from 0.22 (LUQ IC) to 3.63 (SCM MG). Generally, as $AI$ increases, a smaller fraction of $P_{ann}$ is partitioned to $Q_{ann}$. Accordingly, the long-term mean annual water balance (larger markers in Figure 5b) coarsely follows Budyko’s function ($\pm 20\%$; Li et al., 2014), an empirical function describing a nearly ubiquitous relationship between mean annual water balance and $AI$ (Budyko, 1974). However, three catchments (SH SH, EEL EC, and SCM MG) deviate from Budyko’s function, plotting below the 20% error boundary. Deviations below the Budyko function indicate an “overproduction” of runoff relative to Budyko’s hypothesis.

Temporal patterns of annual SWI and $PET_{surf}$ are diverse among network sites (Figure 6). Four CZOs (EEL, RC, SS, and CH) exhibit Mediterranean-type hydroclimatic regimes, where SWI is highest during the winter while $PET_{surf}$ is lowest. At BC, JRB, and SCM, summer monsoon storms act to coalign the timing of SWI and $PET_{surf}$. At these sites, annual precipitation is bimodal and poorly fit by a sinusoidal model (Figure 6). Similarly, at IML and SH, the lag time between SWI and $PET_{surf}$ is reduced by spring and summer rains. In contrast to other network sites, there is relatively little variation of SWI and $PET_{ref}$ at LUQ, yet SWI and $PET_{surf}$ are approximately in phase. Seasonal water delivery and demand dynamics are summarized by the seasonality index ($\phi$, Equation 12), which ranges from $-1.06$ at EEL (strongly out of phase) to 0.46 at SCM (moderately in phase) (Table 2).

5.2. Terrestrial Storage Distributions

Terrestrial storage ($s_{terr}$) is evaluated at a daily time step for each catchment using 4. Boxplots of median-adjusted $s_{terr}$ distributions are plotted in Figure 7 as a way of comparing $s_{terr}$ states among catchments. The range of $s_{terr}$ distributions (vertical lines in Figure 7) conveys the total dynamic $s_{terr}$ capacity of each catchment. In Figure 7, boxplot ranges include negative values because $s_{terr}$ is a relative, not absolute, term. The interquartile range (IQR) of each distribution, indicated by the vertical extent of the boxes in...
et al. (2017). cAnnual actual evapotranspiration estimates from various literature sources. dStandard deviation of annual precipitation at LUQ IC and LUQ MAM was calculated from the available precipitation record synthesized by Murphy et al. (2012). eLower than average annual precipitation across California was recorded from 2012 to 2015. As a result, the average annual precipitation and stream observations are available.

Table 2

<table>
<thead>
<tr>
<th>Catchment</th>
<th>Record years</th>
<th>$P_{ann}$ (mm)</th>
<th>$s(P_{ann})$ (mm)</th>
<th>$Q_{ann}$ (mm)</th>
<th>$PET_{ref ann}$ (mm)</th>
<th>$ET_{a ann}$ literature (mm)</th>
<th>Aridity index</th>
<th>Seasonality index, $\phi$</th>
<th>Annual snow fraction* (%)</th>
<th>mean (std)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LUQ IC</td>
<td>1995 – 2014</td>
<td>5,313</td>
<td>1,021</td>
<td>4,151</td>
<td>1,135</td>
<td>1,162</td>
<td>306 – 1904d</td>
<td>0.22</td>
<td>0.17</td>
<td>0</td>
</tr>
<tr>
<td>LUQ MAM</td>
<td>1995 – 2014</td>
<td>3,953</td>
<td>760</td>
<td>2,883</td>
<td>1,135</td>
<td>1,052</td>
<td>0.30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EEL EC</td>
<td>2009 – 2016a</td>
<td>1,636</td>
<td>556</td>
<td>1,012</td>
<td>1,254</td>
<td>624</td>
<td>339 – 514c</td>
<td>0.81</td>
<td>−1.06</td>
<td>0</td>
</tr>
<tr>
<td>SH SH</td>
<td>2006 – 2012</td>
<td>1,036</td>
<td>144</td>
<td>520</td>
<td>1,002</td>
<td>516</td>
<td>391 – 566d</td>
<td>0.98</td>
<td>0.24</td>
<td>25 (5)</td>
</tr>
<tr>
<td>SS P301</td>
<td>2004 – 2010</td>
<td>1,336</td>
<td>310</td>
<td>399</td>
<td>1,267</td>
<td>937</td>
<td>500 – 800b</td>
<td>1.06</td>
<td>−1.04</td>
<td>72 (14)</td>
</tr>
<tr>
<td>SS P303</td>
<td>2004 – 2010</td>
<td>1,336</td>
<td>310</td>
<td>276</td>
<td>1,267</td>
<td>1,060</td>
<td>1.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS P304</td>
<td>2004 – 2010</td>
<td>1,336</td>
<td>310</td>
<td>410</td>
<td>1,267</td>
<td>926</td>
<td>1.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IML US</td>
<td>2000 – 2015</td>
<td>926</td>
<td>147</td>
<td>317</td>
<td>1,263</td>
<td>609</td>
<td>416 – 618h</td>
<td>1.40</td>
<td>0.36</td>
<td>16 (4)</td>
</tr>
<tr>
<td>CH WS4</td>
<td>2015 – 2016</td>
<td>1,164</td>
<td>197</td>
<td>369</td>
<td>1,782</td>
<td>795</td>
<td>N/A</td>
<td>1.61</td>
<td>−0.22</td>
<td>0</td>
</tr>
<tr>
<td>RC JD</td>
<td>2003 – 2015</td>
<td>674</td>
<td>139</td>
<td>96</td>
<td>1,105</td>
<td>578</td>
<td>150 – 491d</td>
<td>1.67</td>
<td>−0.47</td>
<td>62 (9)</td>
</tr>
<tr>
<td>BC GG</td>
<td>2010 – 2013</td>
<td>557</td>
<td>93</td>
<td>86</td>
<td>1,023</td>
<td>470</td>
<td>N/A</td>
<td>1.85</td>
<td>0.41</td>
<td>47 (8)</td>
</tr>
<tr>
<td>JRB UJ</td>
<td>2008 – 2015</td>
<td>704</td>
<td>143</td>
<td>97</td>
<td>1,447</td>
<td>607</td>
<td>433 – 465d</td>
<td>2.13</td>
<td>0.45</td>
<td>40 (12)</td>
</tr>
<tr>
<td>JRB LJ</td>
<td>2008 – 2012</td>
<td>742</td>
<td>125</td>
<td>56</td>
<td>1,461</td>
<td>686</td>
<td>2.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCM OR</td>
<td>2012 – 2015</td>
<td>546</td>
<td>235</td>
<td>103</td>
<td>1,854</td>
<td>443</td>
<td>N/A</td>
<td>3.46</td>
<td>0.46</td>
<td>29 (5)</td>
</tr>
<tr>
<td>SCM MG</td>
<td>2008 – 2015</td>
<td>594</td>
<td>239</td>
<td>218</td>
<td>1,889</td>
<td>376</td>
<td>3.49</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Mean annual runoff, $Q_{ann}$, precipitation $P_{ann}$, and potential evapotranspiration, $PET_{ann}$, are calculated over consecutive record years when all three observations are available. *Lower than average annual precipitation across California was recorded from 2012 to 2015. As a result, the average annual precipitation and streamflow totals reported for the period of record at EEL EC are lower than the longer-term average reported by Hahm et al. (2019). †Standard deviation of annual precipitation calculated from PRISM Climate Group, 2019 at all sites except for LUQ IC and LUQ MAM, which are outside of the PRISM domain. Standard deviation of annual precipitation at LUQ IC and LUQ MAM was calculated from the available precipitation record synthesized by Murphy et al. (2017). ‡Annual actual evapotranspiration estimates from various literature sources. §Synthesized by Murphy and Stallard (2012) and Murphy et al. (2017). ‖Water balance-derived estimates from Salve et al. (2012). ‡Model-derived estimates from Shi et al. (2013) and flux tower-derived estimates using Davis and Shi (2019). ¶Flux tower-derived estimates from Bales et al. (2018). ‡Flux tower-derived estimates using Meyers (2019). ‡Flux tower-derived estimates using Litvak (2019). Flux tower record includes a pre- and post-fire period. *Annual snow fraction is the percentage of total annual precipitation falling as snow. This is calculated based on a temperature threshold approach in the SNOW17 model and may differ from other independent estimates of snow fraction.

5.3. Recession Plots and Hydraulic Storage-Discharge Relationships

Recession plots, the log of ET-adjusted recession rates (Equation 15) plotted as a function of ln(Q), are unique signatures of drainage from hydraulic storage (Figure 8). Recession data shown in Figure 8 assume $\alpha = 0.25$ at all times for all catchments. A dynamic power law model (Equations 16 and 17) quantifies the general shape of recession plots (Figure 3, black lines). The average of $b_U$ and $b_L$ model parameters ($b_{av}$, Table 3) describes the steepness of recession plots. Higher $b_{av}$ values correspond to more steeply sloping recession plots than lower $b_{av}$ values. The most steeply sloping recession plots are observed in four catchments (LUQ MAM, LUQ IC, SCM OR, and JRB UJ) that have $b_{av}$ values greater than 2 (Table 3). The nonlinearity of a recession plot in logarithmic space is summarized by the difference between $b_U$ and $b_L$ ($b_{curve}$, Table 3). Five catchments (SS P301, SS P303, SS P304, EEL EC, and IML US) have nearly log-linear recession plot shapes, as indicated by $b_{curve}$ values below 0.2. Another five catchments (JRB LJ, JRB UJ, SCM OR, LUQ IC, and LUQ MAM) have recession plots that are concave downward, as indicated by negative $b_{curve}$ values. All other catchments exhibit concave upward recession plot shapes (Figure 8).

The coefficient of determination ($R^2$) is used to assess the goodness of fit between dynamic power law models and observed recession data. Coefficients of determination are calculated for comparing models against binned recession observations ($R^2_{bin}$, Table 3) and comparing models against discrete recession observations ($R^2_i$, Table 3). Models produce satisfactory fits to $Q_{50a}$ at all catchments as indicated by $R^2_{bin}$ greater than or equal to 0.6. However, the ability for models to fit clouds of discrete ET-adjusted recession...
Figure 5. Catchment evaporative partitioning (approximated as $1 - Q_{ann}/P_{ann}$) as a function of the aridity index, also known as a Budyko plot. Smaller markers indicate individual water years, and larger markers denote long-term interannual mean values. Most of the interannual data fall within the theoretical bounds of water and energy limitation, indicated by diagonal and horizontal dashed lines, respectively. Budyko’s curve is plotted as a solid black line, with gray lines denoting 20% uncertainty bounds about the feasible range. Ten of 15 catchments fall within the 20% uncertainty bounds of Budyko’s curve.

Observations is variable between catchments. Models explain less than 50% of the ln($\varepsilon$) variance at three catchments (JRB LJ, SCM OR, and LUQ IC). At one of these (JRB LJ), $R^2$ $\varepsilon$ is less than 0, indicating that the model is no better than the mean. Conversely, models produce satisfactory fits to ln($\varepsilon$) at five catchments (BC GG, SS P301, SS P303, RC JD, and IML US), where $R^2$ $\varepsilon$ is greater than 0.8. The discrepancy of $R^2$ $\varepsilon$ between catchments calls into question whether recession analysis is appropriate for all catchments.

Hydraulic storage-discharge relationships shown in Figure 9 highlight differing hydraulic storage dynamics among catchments. The total dynamic range of $s_h$ is highly variable between catchments. For example, while RC JD and BC GG span similar ranges of discharge, the dynamic range of $s_h$ is much narrower at BC GG, implying that runoff is more sensitive to changes in $s_h$. The narrowest dynamic ranges of $s_h$ are observed at SH SH and CH WS4, where runoff is extremely responsive to $s_h$. Except for JRB LJ, all $s_h$-Q relationships exhibit increasing concave upward shapes. While the shape of $s_h$ at JRB LJ is unique, the validity of the relationship is questionable because the dynamic power law model poorly fits recession plot data at this site (Table 3). The nonlinear increasing concave upward shape of $s_h$-Q relationships implies that discharge becomes more sensitive to changes in $s_h$ as $s_h$ increases.

5.4. Hydraulic Storage and Catchment Sensitivity Distributions

Hydraulic storage-discharge relationships (Figure 9) display the total range of $s_h$ states occupied by each catchment. As storage is filled and drained, catchments dynamically navigate the $s_h$ state space. Distributions of $s_h$ encountered by each catchment are summarized with boxplots in Figure 10. The widest IQR of $s_h$ is observed at four catchments (EEL EC, SS P301, P303, and P304), which also exhibit the widest ranges of $s_{terr}$ (Figure 7). These catchments exhibit the widest dynamic storage ranges. The narrowest IQR of $s_h$ is observed at two catchments (CH WS4 and SH SH). As expected, the range of $s_h$ variation at any catchment is less than the range of $s_{terr}$ variation (Figure 7).

Catchment sensitivity $g(Q)$ (day$^{-1}$) (Equation 18) describes the sensitivity of runoff to changes in hydraulic storage (i.e., the derivative of $s_h$-Q curves). Curvilinear $s_h$-Q shapes indicate that catchment sensitivity depends on hydraulic storage state (or Q). Using 18, $g(Q)$ is evaluated for all Q observations at each site.
Figure 6. Thirty-day moving window average annual terrestrial water input (SWI, solid black lines) and surface-dependent potential evapotranspiration (PET_{surf}, solid green lines) with best-fitting sinusoidal models (dashed lines) for each observatory.

Figure 7. Median-adjusted boxplots summarizing terrestrial storage distributions at 15 CZO catchments. Boxes span the interquartile range and vertical lines show the distribution extent. Sites are arranged from left to right along the x axis in order of increasing seasonality index magnitude. Mediterranean sites are on the left side of the x axis, whereas monsoonal sites are on the right side of the x axis.

Sensitivity distributions are summarized by boxplots in Figure 11. Median catchment sensitivity is highest at LUQ, SH, and CH. Conversely, median catchment sensitivity is lowest at JRB LJ, RC JD, SS P303, and SS P304. RC JD, EEL EC, and IML US exhibit the widest IQRs of catchment sensitivity. The logarithm of median catchment sensitivity is negatively related to baseflow index (BFI) (Figure 12, \( p < 0.01, R^2 = 0.55 \)). The highest BFI is observed at SS and JRB catchments and these catchments also exhibit low median \( g(Q) \). Conversely, the lowest BFI is observed at SH SH and CH WS4, which also exhibit high median \( g(Q) \).

5.5. Relationship Between Dynamic Storage Range and Hydroclimatic Seasonality

The functional ranges of \( s_{terr} \) and \( s_h \) are widely variable among catchments (Figures 7 and 10). The IQRs of \( s_{terr} \) and \( s_h \) distributions are significantly correlated with \( \phi \) (Figure 13). Where \( \phi \) is lower (more negative, meaning that surface water inputs and vegetation water demand are seasonally out of phase), wider ranges of \( s_{terr} \) and \( s_h \) are observed. Where \( \phi \) is greater (more positive, meaning surface water inputs and vegetation water demand are seasonally in phase), narrower ranges of \( s_{terr} \) and \( s_h \) are observed. This result is somewhat intuitive. Under Mediterranean-type climates, storage is continuously filled during wet winters and then...
drained by evapotranspiration throughout dry summers when vegetation water use is highest, causing wide variations in dynamic storage (e.g., EEL EC; Dralle et al., 2018; Hahm et al., 2019). Conversely, the filling and draining of storage zones occurs coincidentally under humid and monsoonal hydroclimatic regimes, when surface water inputs are coincident with periods of high vegetative water use, leading to relatively small ranges of dynamic storage.

The contrasting ranges of $s_{sh}$ and $s_{terr}$ is emphasized by Figure 13c. The IQR of $s_{sh}$ varies between 132 and 9 mm. The IQR of $s_{terr}$, varies between 560 and 93 mm. At all catchments, IQR ($s_{sh}$) is less than IQR ($s_{terr}$). The ratio of IQR ($s_{sh}$) to IQR ($s_{terr}$) is greatest at EEL EC (0.30) and smallest at CH WS4 (0.05). The hydraulically active storage range accounts for a minority of the total dynamic storage range at each catchment, suggesting that vegetation water use is a sizeable draw on catchment storage reserves.

6. Discussion

The striking spatial variations this analysis has revealed show that the CZOs represent a wide range of hydroclimatic variability. Here we will discuss potential linkages between this variety in functional hydrologic

![Figure 8](image)

**Figure 8.** ET-adjusted recession observations ($\varepsilon$ in Equation 15, colored dots) and binned recession values (red markers) plotted as a function of $\ln(Q)$. Best-fitting regression models are shown as black lines. Axis extents are different on each subplot to accentuate differences in recession plot shape.

<table>
<thead>
<tr>
<th>Catchment ID</th>
<th>$b_U$</th>
<th>$b_L$</th>
<th>$a$</th>
<th>$b_{curve}$</th>
<th>$b_{avg}$</th>
<th>$R^2_{bin}$</th>
<th>$R^2_{\varepsilon}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RC JD</td>
<td>1.77</td>
<td>1.71</td>
<td>4633.60</td>
<td>0.06</td>
<td>1.74</td>
<td>0.99</td>
<td>0.90</td>
</tr>
<tr>
<td>SS P301</td>
<td>1.54</td>
<td>1.47</td>
<td>837.89</td>
<td>0.07</td>
<td>1.51</td>
<td>0.99</td>
<td>0.86</td>
</tr>
<tr>
<td>SS P303</td>
<td>1.80</td>
<td>1.64</td>
<td>1253.28</td>
<td>0.16</td>
<td>1.72</td>
<td>0.99</td>
<td>0.86</td>
</tr>
<tr>
<td>IML US</td>
<td>2.00</td>
<td>1.82</td>
<td>552.06</td>
<td>0.18</td>
<td>1.91</td>
<td>0.99</td>
<td>0.86</td>
</tr>
<tr>
<td>BC GG</td>
<td>2.22</td>
<td>1.10</td>
<td>580.98</td>
<td>1.12</td>
<td>1.66</td>
<td>0.98</td>
<td>0.80</td>
</tr>
<tr>
<td>SH SH</td>
<td>2.19</td>
<td>1.47</td>
<td>12.62</td>
<td>0.71</td>
<td>1.83</td>
<td>0.98</td>
<td>0.78</td>
</tr>
<tr>
<td>EEL EC</td>
<td>1.95</td>
<td>1.80</td>
<td>108.16</td>
<td>0.15</td>
<td>1.88</td>
<td>1.00</td>
<td>0.78</td>
</tr>
<tr>
<td>CH WS4</td>
<td>2.22</td>
<td>1.08</td>
<td>55.17</td>
<td>1.13</td>
<td>1.65</td>
<td>0.99</td>
<td>0.62</td>
</tr>
<tr>
<td>JRB UJ</td>
<td>2.07</td>
<td>2.99</td>
<td>757.84</td>
<td>−0.93</td>
<td>2.53</td>
<td>0.96</td>
<td>0.59</td>
</tr>
<tr>
<td>SCM MG</td>
<td>2.12</td>
<td>0.63</td>
<td>208.72</td>
<td>1.48</td>
<td>1.37</td>
<td>0.86</td>
<td>0.57</td>
</tr>
<tr>
<td>SS P304</td>
<td>1.73</td>
<td>1.70</td>
<td>208.29</td>
<td>0.02</td>
<td>1.71</td>
<td>0.97</td>
<td>0.56</td>
</tr>
<tr>
<td>LUQ MAM</td>
<td>1.80</td>
<td>4.03</td>
<td>2.87</td>
<td>−2.23</td>
<td>2.92</td>
<td>0.98</td>
<td>0.53</td>
</tr>
<tr>
<td>LUQ IC</td>
<td>1.77</td>
<td>3.09</td>
<td>2.02</td>
<td>−1.32</td>
<td>2.43</td>
<td>0.96</td>
<td>0.41</td>
</tr>
<tr>
<td>SCM OR</td>
<td>2.09</td>
<td>2.94</td>
<td>1216.58</td>
<td>−0.85</td>
<td>2.52</td>
<td>0.89</td>
<td>0.33</td>
</tr>
<tr>
<td>JRB LJ</td>
<td>−1.39</td>
<td>2.46</td>
<td>1257.99</td>
<td>−3.85</td>
<td>0.54</td>
<td>0.64</td>
<td>−0.30</td>
</tr>
</tbody>
</table>

**Table 3**

Dynamic Power Law Regression Model Best-Fit Parameters, Shape Metrics, and Goodness of Fit Statistics

Note. Table rows are organized from best to worst $R^2$ statistic for the model fit against discrete observations.
behavior and variations in critical zone architecture across the sites. Continuing the theme developed throughout the paper, this discussion focuses on dimensions of hydrologic storage as a key functional property of the critical zone.

We can examine the results through two lenses. The first is to consider the annual water balance in terms of the Budyko curve. This well-studied empirical relationship is believed to be driven by the capacity of the landscape to store the water intermittently supplied by rainfall and snowmelt such that it is available for transpiration by vegetation at other times. The second is to consider the catchment sensitivity function, which expresses the capacity of the landscape to store and later release water as streamflow. This relationship is also well studied and has been hypothesized to arise from the coevolved architecture of the critical zone. How that coevolution might occur is considerably less well understood relative to the Budyko case, though some mechanisms have been suggested for basaltic landscapes (Jefferson et al., 2010). Finally, we will examine an intriguing point of overlap between these two perspectives: the relationship between the seasonality of the water balance drivers and the range of hydraulic storage revealed by the catchment sensitivity function.

6.1. Annual Water Balance Partitioning

Annual water balance dynamics can be examined through the Budyko framework, where observed relationships between climatic aridity, $AI$, are plotted against evaporative partitioning (approximated here as $1 - Q/P$) (Figure 5). The Budyko relationship has been used to describe the annual water balance of many catchments throughout the world (Budyko, 1974; Gentine et al., 2012). The relationship is believed to arise from a coevolution of climate and critical zone properties, particularly vegetation and soil characteristics (Troch et al., 2013). That is, catchment vegetation evolves with climate to utilize the largest sustainable fraction of soil moisture supplied by infiltration (Eagleson, 1982; Gao et al., 2014; Horton, 1933).

Most of the CZO catchments fall close to the Budyko curve (12 of 15 are within 20%), but others do not. Mean annual water balances at three
Figure 11. Catchment sensitivity, \( g(Q) \), describes the sensitivity of runoff to changes in hydraulic storage state. Boxplots summarize the operational distributions of catchment sensitivity at 15 CZO catchments. Horizontal lines indicate median \( g(Q) \) values and sites are ranked left to right by median \( g(Q) \). Boxes indicate the interquartile range and vertical lines indicate distribution extents.

The Shale Hills catchment (SH SH) also exhibits annual and interannual departures from Budyko’s hypothesis. This may be attributed to the dominance of shallow subsurface flow paths in streamflow generation (Jin et al., 2011; Li et al., 2017; Lin, 2006). Shallow subsurface flow, or interflow, occurs through a highly fractured zone in the upper 5–8 m of regolith (Sullivan et al., 2016). This fractured zone overlies less fractured sites (SH SH, EEL EC, and SCM MG) plot below Budyko’s curve in Figure 5. The most arid catchment, SCM MG, exhibits annual and interannual water balances that fall below Budyko’s curve. In June 2003, nearly half of the SCM MG catchment burned during the Aspen Fire (Canfield et al., 2005). Since the fire, which predated observations, vegetation has been undergoing successive phases of recovery (P. Troch and J. Chorover, personal communication, February 8, 2018). The apparent overproduction of runoff relative to Budyko’s hypothesis may be due to a disturbance-induced long-term disequilibrium between climate and vegetation.

Elder Creek at the Eel River CZO also exhibits an overproduction of runoff relative to Budyko’s hypothesis. The climate at EEL EC is strongly Mediterranean \((\phi = -1.06, \text{ Table 2})\), where surface water input and evaporative demand are out of phase. Several studies have attributed Budyko anomalies to climatic seasonality (e.g., Berghuijs et al., 2014; Sankarasubramanian & Vogel, 2003), which may also explain the annual inefficiency of plant water use at EEL EC. Because 100% of the precipitation falls as rain, the subsurface is the only storage reservoir capable of seasonally retaining water between the wet winter and dry summer seasons. As shown by Rempe and Dietrich (2018), subsurface storage is filled during the wet season and then vegetation draws from storage throughout the relatively dry growing season to sustain productivity. Perhaps subsurface storage deficits in the dry season impose a seasonal water limitation on vegetation, resulting in relatively inefficient water use compared to other catchments of similar aridity (e.g., SH or SS), where growing season storage does not limit plant water use.

Figure 12. Median catchment sensitivity is negatively related to baseflow index \((p < 0.01, R^2 = 0.55)\). The size of each marker is scaled by the average annual fraction of precipitation falling as snow at each site (Table 2).
Figure 13. Terrestrial (a) and hydraulic (b) storage interquartile range (IQR) is negatively related to seasonality index ($p << 0.01$). Mediterranean-type climates ($\phi < 0$) exhibit a wider range of dynamic storage states than continental, tropical, or monsoonal-type climates ($\phi > 0$). Vertical error bars of hydraulic storage IQR markers indicate the uncertainty associated with ET fluxes from hydraulic storage zones. The upper extent of error bars assumes that all ET is drawn from hydraulic storage zones (i.e., $\alpha = 1$) and the lower extent of error bars assume that no ET is drawn from hydraulic storage zones (i.e., $\alpha = 0$). (c) Terrestrial and hydraulic storage interquartile ranges are positively related ($p << 0.01$).

rock, upon which an annually persistent perched water table develops and drives interflow. The rapid transmission of meteoric water to the stream via shallow interflow poses a possible explanation for the Budyko anomaly at Shale Hills. Supporting this notion, Troch et al. (2013) showed that short subsurface residence times tend to be associated with lower values of the evaporative index than would be expected from Budyko’s hypothesis.

The Calhoun CZO site, CH WS4, does not depart significantly from Budyko’s hypothesis in Figure 5. A prominent Bt horizon at CH WS4 (and other catchments throughout the Carolina Piedmont) promotes lateral interflow through shallow transiently perched aquifers. Additionally, the agricultural legacy along the South Carolina Piedmont has led to substantial erosion and a top-down decapitation of the A horizon. The reduced A horizon storage capacity, and associated brief residence times in the perched aquifer, would seem to favor runoff partitioning and non-Budyko behavior (Troch et al., 2013), yet this is not necessarily observed. Currently, our ability to elucidate interannual and intra-annual hydrologic partitioning at the Calhoun CZO is limited by a relatively short discharge record of two water years. As the record continues to accrue and ongoing research matures, our hydrologic picture of this system will sharpen.

Our quantification and interpretation of annual water balance is obscured in part by uncertainties in $P$ and $Q$. We did not conduct an uncertainty analysis on the $Q$ or $P$ data reported by each CZO. Uncertainties in streamflow arise from uncertainties in stage discharge rating curves, discharge measurement error, and stage measurement error. Uncertainties in $P$ arise from gauge under catch and inaccurate spatial extrapolation of precipitation quantities between gauges in a catchment. In the context of the Budyko plotting space, overestimates in discharge would cause underestimates in the evaporative index and vice versa. Conversely, overestimates in precipitation would cause underestimates in the evaporative index and vice versa. Most of the discharge data provided by the CZOs do not include error estimates, yielding another synthesis opportunity for future cross-CZO investigations. Similarly, the relatively simple scheme used to approximate catchment-wide effective precipitation rates from discrete gauges leaves much room for improvement.

The small size of most CZO catchments relative to those studied by Budyko also obscures our interpretation of the annual water balance. In very large catchments, net groundwater inflows and outflows are likely to be a smaller portion of the long-term annual water balance (Fan, 2019). However, in geologically complex small catchments, like most of the catchments considered here, net groundwater flows likely play a larger role. We ignored net groundwater flows from the annual water balance for convenience and must consider the consequence of this simplification on our understanding of catchment water balances.
Ignoring groundwater import/export biases water balance-derived estimates of annual evapotranspiration. If a catchment is leaky (net groundwater exporter), ET will be overestimated by a simplified water balance framework. Many of the catchments studied here are likely to be leaky, which would cause on overestimation of actual ET and 1 − Q/P. Indeed, the water balance-derived ET estimates are consistently greater than literature-derived ET values, most of which are based on surface energy balance investigations (Table 2). Annual average ET estimates derived from water balances exceeded literature-derived estimates by as much as 80% and 35% on average (Table 2). This discrepancy may be driven by either groundwater leakage or groundwater storage that is active on longer timescales than can be detected from the relatively short records at each site. Additional synthesis of interbasin groundwater flows and storage is necessary to reduce the uncertainty of annual water balance partitioning in these catchments.

Indeed, ignoring groundwater leakage obscures the plotting positions of each site in Figure 5, such that the plotting distance below Budyko’s hypothesis is likely underestimated in our investigation. So, while ignoring groundwater exports causes a high bias in our annual ET estimates, it does not discount the observation that SH SH, EEL EC, and SCM MG seem to overproduce runoff, relative to Budyko’s hypothesis.

The water balance analysis presented here is further limited by the relatively short record lengths available at the CZOs. Other studies utilizing the Budyko framework typically consider catchments with at least 50 years of data because over sufficiently long periods, changes in catchment storage are assumed to be negligible, and ET ≈ P − Q (assuming groundwater flows are negligible). To better understand similarities and differences in annual partitioning behavior across the CZO network, it is necessary to (1) develop longer records by continuing the collection of runoff and precipitation data and/or (2) leverage ancillary data sets, such as soil moisture or groundwater depth, to approximate changes in catchment storage and groundwater import/export.

### 6.2. Hydraulic Response and Critical Zone Architecture

We observed a significant negative relationship between catchment sensitivity and baseflow index (Figure 12). According to this result, CZO catchments may be conceptually organized between two contrasting end-members: (1) baseflow dominated with low runoff sensitivity to storage and (2) storm flow dominated with high runoff sensitivity to storage.

Prior studies have attributed catchment hydraulic response to critical zone architecture and regolith development. Notably, Jefferson et al. (2010), Lohse and Dietrich (2005), and Yoshida and Troch (2016) conducted controlled experiments across chronosequences of basaltic catchments. Results showed that baseflow is less dominant in catchments with more developed regolith because extensive primary mineral weathering facilitates the development of clay lenses, favoring shallow interflow (storm flow) over deeper baseflow. Along these same lines, Brantley et al. (2017) used a coupled hydrologic and reactive transport model to illustrate how saturated lateral flow paths can develop at geochemical reaction fronts in a way that allows reaction fronts of minerals of different solubilities to coexist in a steady-state critical zone profile (Harman & Cosans, 2019). Their model suggested that baseflow emanates from deep water tables developed at reaction fronts delineating the interface between unweathered and weathered bedrock. Similarly, interflow emanates from perched water tables atop shallower secondary reaction fronts, usually associated with argillic horizons or the interface between soil and saprolite.

The hydraulic response of CZO catchments observed in this study does appear to be linked to the architecture of the critical zone. Catchments with clay-rich mobile regolith that might divert water laterally prior to reaching the deeper bedrock exhibit less baseflow and higher g(Qmed) than catchments with more permeable regoliths that allow for recharge to the bedrock. This is not to say that the architecture of the deeper critical zone is not relevant for understanding the hydraulics, only that a first-order control on baseflow variability is whether the properties of the surficial regolith permit temporally variable recharge to that deeper zone.

To better articulate this hypothesis, regolith profiles for each U.S. CZO site were assembled using a range of available data, including the deep local site knowledge available at each CZO (Figure 14). These profiles are necessarily gross simplifications of the true complexity of each site, but they capture some essential differences between the sites. Profiles are organized from left to right according to baseflow index and g(Qmed).
Figure 14. Soil depth and texture profiles for each CZO site. Initial profiles were created following data obtained from the NRCS Web Soil Survey for all sites except Santa Catalina, where no information was available. Each profile was modified based on additional soil information provided by investigators at each site and peer-reviewed literature. The soil texture and depth profile at Santa Catalina were synthesized from information in Holleran et al. (2015). Results at Elder Creek are best explained by storage dynamics in deeper weathered bedrock not pictured above, as detailed in Hahm et al. (2019) and Rempe and Dietrich (2018). Figure design by E. Parrish.

Sites that exhibit the low baseflow indices and high \( g(Q)_{med} \) also have structural attributes that favor rapid lateral water transmission as interflow and precipitation regime dominated by rainfall. Calhoun and Luquillo CZOs are both deeply weathered with well-developed regolith and prominent clay horizons (Buss et al., 2013; Orlando et al., 2016; St. Clair et al., 2015). The clay-rich Bt horizon (Calabrese et al., 2018) at CH WS4 may facilitate perched water table development and interflow, though this has not been directly observed. However, Zimmer and McGlynn (2017) observed interflow from perched water tables developed atop clay horizons at another site in the North Carolina Piedmont with similar soil characteristics. At LUQ, hydraulic conductivity decreases sharply at the interface of the surficial organic horizon and clay-rich A horizon in the Mameyes watershed (Schellekens et al., 2004). This permeability contrast cues the development of perched water tables and interflow through shallow organic soils (Schellekens et al., 2004). Both CH and LUQ receive nearly 100% of precipitation as rainfall. At SH, interflow responsible for the majority of runoff generation occurs through highly fractured shale at depths of 5–8 m (Sullivan et al., 2016). Additionally, flow through macropores in channery shale soils at SH is also understood to contribute to runoff generation (Lin, 2006). At SH, between 63% and 82% of the precipitation falls as rain.

The Upper Sangamon basin at the Intensively Managed Landscape (IML) CZO also exhibits low baseflow and high \( g(Q)_{med} \); however, unique from other CZOs, anthropogenic manipulation of the critical zone likely plays an important role in controlling basin-wide hydraulic response. European settlement across the Midwestern United States throughout the 19th and 20th centuries prompted drastic changes in landscape...
morphology to alleviate poor drainage conditions on fertile agricultural lands. In particular, streams were straightened and widened, and channel networks were extended further into the headwaters (Rhoads et al., 2016). Additionally, tile drains were installed in the shallow subsurface to lower water tables, drain the rooting zone after storms, and reduce surface ponding in low-lying areas (Hewes & Frandson, 1952; Schilling & Helmers, 2008). These human modifications promote the rapid transport of water through and across the landscape by artificially modifying drainage regimes and accelerating transit times (Kumar et al., 2018). However, it remains difficult to establish a causal linkage between anthropogenic impacts and hydraulic response in this study, because the CZO network lacks an un-impacted Midwestern catchment for comparison. Similar to other catchments characterized by low baseflow and high \( g(Q_{med}) \) (SH, LUQ, and CH), 74–90% of annual precipitation falls as rain (Table S3).

Two CZO sites, SS and JRB, stand out as being baseflow dominated (baseflow index > 60%) with relatively low \( g(Q_{med}) \), compared to other CZO sites (Figure 12). At JRB UJ and LJ, stream flow is primarily fed by groundwater emanating from fractured bedrock (Olshansky et al., 2018; Zapata-Rios et al., 2016), and ~41–63% of annual precipitation falls as snow (McIntosh et al., 2017). Stable isotope data indicate that snowmelt is the dominant source of water to groundwater reservoirs, which are the dominant source of streamflow generation (Liu et al., 2008). During the relatively brief wet season following snowmelt at JRB, interflow through perched aquifers atop clay-rich lenses can also contribute to stream flow (McIntosh et al., 2017); however, annual contributions are dominated by flow from fractured bedrock 15–18 m beneath the surface. Similarly, at SS, groundwater is a key source of runoff, accounting for 61–69% of total streamflow in the Providence Creek subcatchments (P301, P303, and P304) according to the results of a geochemical mixing analysis (Li et al., 2013). Snow is a critical water subsidy to Providence Creek, accounting for 35–60% of annual precipitation (Hunsaker et al., 2012). In these catchments, both snowmelt and rainfall rapidly infiltrate vertically through well-drained sandy regolith to an aquifer oriented atop fresh bedrock, which is responsible for generating the majority of runoff (Bales et al., 2011; Liu et al., 2013). Moreover, regolith depth exhibits strong lateral variability over short spatial scales throughout Providence Creek (Holbrook et al., 2014), invoking the possible control of flow and spill dynamics on runoff generation (Tromp-van Meerveld & McDonnell, 2006).

Four CZOs (EEL, BC, SCM, and RC) cluster in the middle of the baseflow index and \( g(Q_{med}) \) space (Figure 12). Three of these four sites (BC, SCM, and RC) are architecturally and climatically similar. These sites have metamorphic and igneous parent materials (Table 1) and are thinly weathered, with mobile regolith thickness less than 2.4 m (Foster et al., 2015; Holleran et al., 2015; Patton et al., 2018; Pelletier & Rasmussen, 2009). Moreover, these sites receive approximately similar annual quantities of precipitation as snow and rain, with snowfall accounting for 23% to 68% of annual precipitation across sites. End-member mixing analysis at BC GG and SCM MG reveals that the majority of runoff is geochemically similar to precipitation, with precipitation contributing 69% of runoff at BC GG (Cowie et al., 2017) and 51–60% of runoff at SCM MG (Dwivedi et al., 2018). Consistent with this, end-member mixing analysis in a headwater catchment at RC indicates that groundwater averages 25% of annual streamflow, snowmelt 50%, and water traveling along the saprolite/bedrock boundary 25%. These results indicate that runoff may be predominately generated via shallow subsurface flow paths through saprolite above weathered bedrock (Radke et al., 2019). Indeed, hillslope observations at SCM MG confirm that runoff generation is tightly coupled to the development of a perched water table atop shallow fractured bedrock (Heidbüchel et al., 2013). Limited data are available at BC GG and RC JD for identifying specific runoff generating pathways, an opportunity for future research.

Elder Creek at the Eel River CZO (EEL EC) stands out from other sites clustered in the middle of the baseflow index and \( g(Q_{med}) \) space (Figure 12). Unlike the other three catchments in this cohort, EEL EC is more deeply weathered (>20 m to fresh bedrock beneath ridgetops) and receives 100% of precipitation as rain (Hahm et al., 2019). Detailed observations on critical zone architecture and water storage reveal that runoff is primarily generated by deep groundwater flow (Kim et al., 2017; Rempe & Dietrich, 2018). The groundwater table in EEL EC resides within a fractured and weathered bedrock on top of a fresh bedrock boundary and exhibits seasonal vertical variations of 4–6 m (Kim et al., 2017; Rempe & Dietrich, 2018). Groundwater is seasonally recharged during wet winter months by infiltration through a deep vadose zone, once soil and rock moisture storages reach field capacity (Dralle et al., 2018). Given the clear dominance of deep
(>10 m) groundwater in generating streamflow, it is counterintuitive that basin-scale analysis reveals a moderate baseflow index and \( g(Q)_{med} \), relative to other CZO catchments. However, during the wet season (winter and spring), the water table is high and in a very transmissive layer of fractured and weathered rock (Salve et al., 2012). As such, shallow groundwater contributions manifest as a flashy streamflow response and may not be interpreted as baseflow by the simple digital filter approach used for baseflow separation.

Regolith depth (depth to bedrock) does not seem to control catchment hydraulic response dynamics. Catchments with similar regolith depths exhibit contrasting hydraulic response metrics (baseflow index and \( g(Q)_{med} \)). For example, JRB UJ and LJ, EEL EC, and CH WS4 may all be considered deeply weathered, with fresh bedrock residing more than 20 m beneath ridgetops (Table 1). Nevertheless, these sites exhibit contrasting hydraulic response metrics (Figure 12). This observation suggests that catchment storage states and more nuanced regolith properties, such as texture, permeability contrasts, or fracture properties, may exert a stronger control on catchment hydraulic regimes than regolith depth (depth to bedrock) alone. However, exceptions to this observation exist. Hahm et al. (2019) illustrate how regolith depth limits storage capacity into neighboring watersheds underlain by contrasting lithologies in coastal Northern California.

The fraction of precipitation falling as snow also seems to be related to the emergent hydrologic response dynamics of the watersheds considered in this study. Snow-dominated catchments exhibit lower runoff sensitivities to storage and higher baseflow indices, than catchments where the majority of precipitation falls as rain (Figure 12). This observation is in agreement with Barnhart et al. (2016), who observed that seasonal snow melt can bring soils to field capacity and facilitate deep critical zone recharge that drains to streams as baseflow. Moreover, this result is corroborated by Knoben et al. (2018), who argue that the fraction of precipitation falling as snow is one of three climate indices that control global variations in hydrologic responses.

Looking to the future, a shift in winter precipitation phase from snow to rain under future climate scenarios is predicted to occur at many of the U.S. CZOs, especially RC, BC, SS, and JRB (Klos et al., 2014). Results from this study suggest that changing precipitation phase may prompt changes in a catchment’s hydrologic signatures, such as the shape of recession curves, the dynamic range of storage, and the baseflow index.

The hypothesized linkage between regolith development and catchment hydraulic response is appealing because it implies a coevolutionary link between catchment form and function. That is, the transmission of water through soils and regolith is both essential to critical zone evolution and a manifestation of the unique historical pathway of such evolution. The hydraulic response metrics quantified in this manuscript (catchment sensitivity and baseflow index) provide an elusive critical zone pattern begging for a unifying explanation. The critical zone community seeks to use emergent patterns in hydrologic behavior among watersheds to infer historical pathways of critical zone evolution. Future work might explore this challenge by deducing the emergent hydrological properties associated with different pathways of critical zone evolution (Riebe et al., 2017).

6.3. Seasonality Explains Hydrologic Dynamic Storage Range

The results suggest that hydroclimatic seasonality partially explains the dynamic storage ranges of CZO catchments. The IQR of \( S_{terr} \) and \( S_{sh} \) is significantly related to seasonality index, \( \phi \) (Figure 13). At sites where SWI and PET are strongly out of phase, such as SS and EEL, \( S_{terr} \) and \( S_{sh} \) exhibit large dynamic ranges as storage is progressively filled during the wet season and gradually drained during the dry season. In contrast, at sites where SWI and PET are in phase, such as at JRB and SCM, dynamic ranges of \( S_{terr} \) and \( S_{sh} \) are small because storage filling and spilling occur at the same time.

We also observed that the dynamic range of \( S_{sh} \) was smaller than \( S_{terr} \) (Figure 13c), suggesting that hydraulically relevant storage is a subset of total storage. This result corroborates Dralle et al. (2018), who attribute the disparity between hydraulic (direct) and total storage to rock moisture at EEL EC. This result also aligns with the definitions of perceptual catchment storage put forth by Staudinger et al. (2017), who define “dynamic storage” as that which controls streamflow dynamics (what we consider to be \( S_{sh} \)) and “extended dynamic storage” as that which controls all catchment fluxes (what we consider to be \( S_{terr} \)). Staudinger et al.’s use of the prefix “extended” implies that \( S_{sh} \) is only a subset of \( S_{terr} \). Clearly, there remains to be an established textbook lexicon for discussing storage.
The relationship between hydroclimatic seasonality and dynamic storage ranges is somewhat unsurprising since dynamic storage accumulation (and depletion) should depend on the timing and magnitude of terrestrial water inputs and atmospheric water demands. However, the range of dynamic and hydraulic storage also depends on the rate of release of that accumulated storage as streamflow. If the rate of release is fast enough, the range of hydraulic storage will depend only on the ability of individual storms to recharge the hydraulically connected pore space, which will be largely depleted again before the next storm. This can be seen at the IML US where tile drainage likely limits the range of hydraulic storage. Where the release of water is slow, storage can accumulate over multiple years and the range of hydraulic storage ought to be better explained by interannual variability, rather than intra-annual.

Here, the seasonality is likely the primary control on $s_h$ largely because for most of the catchments, and for most of the time, the sensitivity function has a value between $g(Q) = 1/\text{week}$ and $g(Q) = 1/\text{year}$ (Figure 11). That is still a wide range of variability, but not as wide as it could be given that hydraulic conductivity (which idealized analytical models suggest ought to be a primary control on streamflow recession; Brutsaert & Nieber, 1977) varies over 14 orders of magnitude.

This raises two questions: (1) why do catchments tend to have sensitivity functions between 1/week and 1/year? And (2) what determines the variability between catchments within that range? From a mechanistic “Newtonian” point of view, the answer lies in the physical properties of the landscapes. But from a geological point of view, the answer could lie in the explanation for why those properties arise as emergent phenomena that modulate the climatic drivers with coevolutionary feedbacks between hydrologic dynamics and critical zone architecture.

7. Conclusions

This is the first meta-analysis of the commonly observed hydrologic data sets across the CZO network. The implicit hypothesis of this network was that studying the CZ in its entirety at each location would yield insights about form, function, and dynamics. Here, we started the process of extracting quantitative signatures of water storage, partitioning, and release and used these to compare CZO sites. We then drew on the deep accumulated understanding developed at these sites to examine how these patterns could be explained by the critical zone architecture.

The data compiled for this study should propel others forward in drawing on the depth of knowledge amassed at the CZOs. Although the water balance accounting models used here are simple, the time needed to collate them was significant, and we only made a first step in synthesis. More work is needed to synthesize these data and calculate signatures of hydrologic behavior that can be compared across the sites. Metrics of water balance partitioning and catchment hydraulics were chosen in part because they allowed us to focus on the critical role of hydrologic storage as a modulator of hydrologic variability and as a capacity of the landscape that arises from the evolved architecture of the critical zone.

While aridity is a first-order control on evaporative partitioning across the network, seasonality, disturbance, and other factors may drive sites away from the idealized Budyko relationship. Twelve of 15 catchments closely follow Budyko’s hypothesis, suggesting a coevolved equilibrium between climate, soils, and vegetative water use at these sites. The remaining catchments produce more runoff than is predicted by Budyko, perhaps suggesting ecohydrological disequilibrium. Catchment sensitivity functions (i.e., storage-discharge relationship), hydraulic storage, and baseflow indices were derived and used as a basis for hydraulic comparison. We observed a negative relationship between baseflow index and catchment sensitivity. Hydroclimatic seasonality and climatic variability appear to explain most of the variation in dynamic storage range between catchments, as well as a substantial portion of the variation in hydraulically connected storage range. Catchments with Mediterranean-type hydroclimates exhibit large dynamic storage ranges, whereas catchments with monsoonal-type hydroclimates exhibit small dynamic storage ranges.

While this pattern may seem to suggest the dominance of climate controls over landscape structure, we believe this is misleading—we propose that the resonance between catchment storage and seasonal dynamics is only possible because of the capacity of the landscape to store and shed water over similar timescales.
It seems to be whether this and other hydrologic properties of the landscape can be understood in terms of the coevolution of critical zone architecture, but patterns did emerge from the extensive comparison that suggest this is possible. Catchments with low baseflow indices and high catchment sensitivity receive the majority of precipitation as rain and contain clay-rich regolith profiles, prominent argillic horizons, and/or anthropogenic modifications. On the other hand, sites with high baseflow indices and low catchment sensitivity receive a large portion of precipitation as snow, have more permeable regolith profiles, and range from deeply to thinly weathered.

Data Availability Statement

Data products derived from these parent data sets have been made available through a CUAHSI HydroShare data resource (Wlostowski, A., N. Molotch, C. Harman (2021). CZO Network water budget data products, HydroShare, https://doi.org/10.4211/hs.29e2ec85770b42e881ef075069646).

Acknowledgments

Much of the data used in this study were provided by sites within the U.S. Critical Zone Observatory (CZO) network, each of which are funded by the U.S. National Science Foundation: Boulder Creek (grant number EAR-1331828), Shale Hills (grant number EAR-1311726), Eel River (grant number EAR-1331940). Reynolds Creek (grant number EAR-1331872), Southern Sierra (grant number EAR-1331939), Catalina-Jemez (grant number EAR-1331408), intensively managed landscapes (grant number EAR-1311906), Calhoun (grant number EAR-1331846), and Loquillo (grant number EAR-1331841). The Pacific Southwest Experimental Research Station, USDA Forest Service, which operates the Kings River Experimental Watersheds, and the Sierra National Forest, provided additional data and support at the Southern Sierra CZO. The USDA Reynolds Creek Experimental Watershed provided additional data and support at the Reynolds Creek CZO. Additional monetary support for this research was graciously provided by Timothy White, through National Science Foundation grant number ICER-1445246. Acknowledgement and citation for the parent data sets used in this study are provided within the main text and the supporting information.

References


Monteith, J. L. (1973). Principles of environmental physics. London: Edward Arnold. Retrieved from http://uclibraries.summon.serial-sslolutions.com/2.0.0/link/0edwhfcmXmyw2AwiNl0eUfOcnuZKUY61UnmGhkppkmSgkpQGb7qAPmMzxwNzi4HtaAAWtQElEwJINNgYDIJEMUJcRdRrUJWKOvpMO2D8wncMkMLnjmZmYAUWGbAnMm507J4wun0h8aA4ltxcNJTf6pMFTQHIZjGSoGdEcy2zkyFkusZNkIEF9P9All


