Check for updates

Zhang Liang (Orcid ID: 0000-0002-6111-5640)
Moges Edom (Orcid ID: 0000-0003-4559-0799)
Wymore Adam S. (Orcid ID: 0000-0002-6725-916X)

Accepted Article

CHOSEN: A synthesis of hydrometeorological data from intensively monitored catchments and comparative analysis of hydrologic extremes

Liang Zhang[1], Edom Moges[1], James Kirchner[2], Elizabeth Coda[1], Tianchi Liu[1], Adam S. Wymore[3], Zexuan Xu[4], Laurel G. Larsen[1]

[1] University of California, Berkeley, CA, USA
[2] ETH Zürich, Zurich, Switzerland
[3] University of New Hampshire, Department of Natural Resources and the Environment, Durham, NH, USA
[4] Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Keywords: hydrologic extremes, trends of extreme events, intensively monitored catchments, comprehensive hydrologic dataset, comparative hydrology, climate change

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/hyp.14429

This article is protected by copyright. All rights reserved.

Accepted Article

Abstract

Comparative hydrology has been hampered by limited availability of geographically extensive, intercompatible monitoring data on comprehensive water balance stores and fluxes. These limitations have, for example, restricted comprehensive assessment of multiple dimensions of wetting and drying related to climate change and hampered understanding of why widespread changes in precipitation extremes are uncorrelated with changes in streamflow extremes. Here we address this knowledge gap and underlying data gap by developing a new data synthesis product and using that product to detect trends in the frequencies and magnitudes of a comprehensive set of hydroclimatic and hydrologic extremes. CHOSEN (Comprehensive Hydrologic Observatory Sensor Network) is a database of streamflow, soil moisture, and other hydroclimatic and hydrologic variables from 30 study areas across the United States. An accompanying data pipeline provides a reproducible, semi-automated approach for assimilating data from multiple sources, performing quality assurance and control, gap-filling, and writing to a standard format. Based on the analysis of extreme events in the CHOSEN dataset, we detected hotspots, characterized by unusually large proportions of monitored variables exhibiting trends, in the Pacific Northwest, New England, Florida, and Alaska. Extreme streamflow wetting and drying trends exhibited regional coherence. Drying trends in the Pacific Northwest and Southeast were often associated with trends in soil moisture and precipitation (Pacific Northwest) and evapotranspiration-related variables (Southeast). In contrast, wetting trends in the upper Midwest and the Rocky Mountains showed few univariate associations with other hydroclimatic extremes, but their latitudes and elevations suggested the importance of changing snowmelt characteristics. On the whole, observed trends are incompatible with a “drying-in-dry, wetting-in-wet” paradigm for climate-induced hydrologic changes over land. Our analysis underscores the need for more extensive, longer-term observational data for soil moisture, snow, and evapotranspiration.

Accepted Article

Acknowledgments

This work is supported by the US Geological Survey Powell Center for Analysis and Synthesis, a Gordon and Betty Moore Foundation Data-Driven Discovery Investigator grant to LL, and the Jupyter Meets the Earth project, funded by NSF grant number (UC Berkeley: 1928406, NCAR: 1928374). Partial support for ASW is provided by the National Science Foundation and the Experimental Program to Stimulate Competitive Research (EPSCoR: EPS-1929148; Canary in the Watershed). Much of the data used in this study were obtained from the US. Long-Term Ecological Research Network (LTER), Critical Zone Observatory program (CZO), Lawrence Berkeley National Laboratory, Dry Creek Experimental Watershed (DCEW); we would like to acknowledge all the staff from these institutions for collecting and publishing these data. We thank Dr. Adrian Harpold for providing the data from the Sagehen catchment. We would especially like to thank the Powell Center Working Group on Watershed Storage and Controls for their contributions to this project. We also thank Dr. Lindsey Heagy and Dr. Fernando Pérez for their suggestions on data publication and future development. This manuscript greatly benefited from constructive comments from five anonymous reviewers and Associate Editor Steve Sebestyen.

Accepted Article

1. Introduction

Climatic and hydrologic extremes pose severe risks to human society and infrastructures and trigger irreversible transitions in ecosystems (AghaKouchak et al., 2020; Ainsworth et al., 2020; Hughes et al., 2019; McClymont et al., 2020). The magnitude and frequency of these extremes are increasing as a result of climate change (e.g., Ahn & Palmer, 2016; Pagán et al., 2016; Swain et al., 2018; Wentz et al., 2007), which results from basic physical principles. In accordance with Clausius-Clapeyron scaling, warmer air holds more moisture, which is associated with projected increases in rainfall intensity (Sillmann et al., 2013), the intensity and frequency of tropical cyclones (Marsooli et al., 2019), and the amount of water conveyed in atmospheric rivers (Gao et al., 2015; Payne et al., 2020). Warmer temperatures also increase potential evapotranspiration and are linked to increasing drought severity (Cook et al., 2015; Diffenbaugh et al., 2015). The balance between processes that promote catchment drying (e.g., enhanced evapotranspiration) and those that promote wetting (e.g., increased precipitation extremes) varies among catchments. Therefore, it can be difficult to generalize outcomes of increasing precipitation and temperature extremes for hydrological processes.

The difficulty in predicting how increased climatic extremes will impact hydrologic extremes is particularly apparent in the discrepancy between the projected and observed association between precipitation and discharge extremes. While climate models predict a strong correlation between extreme precipitation and extreme flood magnitude (e.g., Pall et al., 2011), observations show low correlation spatially and temporally (e.g., Archfield et al., 2016; Berghuijs et al., 2016; Blöschl et al., 2017; Do et al., 2020), except for rare floods with recurrence intervals longer than 10 years (Wasko & Nathan, 2019). Specifically, flood trends are not changing in accordance with climate model predictions (Sharma et al., 2018). The need to understand the link between changing precipitation and changing flooding has been argued to be one of the grand challenges in hydrology (Sharma et al., 2018).

Measurements of soil moisture and other variables indicative of water balance stores and fluxes may provide clues critical to reconciling Sharma et al. 's (2018) grand challenge, and, more broadly, understanding how shifting climate translates into a range of hydrological outcomes. Results of modeling and observational studies that derive (Berghuijs et al., 2016; Byun et al., 2019; Heidari et al., 2020; Ivancic & Shaw, 2015) or account for measured soil moisture (Wasko & Nathan, 2019) or changes in subsurface storage (Slater & Villarini, 2016) suggest that changes in hydrologic extremes are attributable to simultaneous shifts in several hydrologic variables, with soil moisture or subsurface storage of critical importance. One gap in these analyses is that, with the exception of Wasko and Nathan's (2019) study of Australian catchments, they rely on simple models or proxies for soil moisture rather than actual measurements. Meanwhile, the role of soil moisture, snow storage, and actual evapotranspiration in governing low-flow extremes remains underexplored. Exploration of causes of hydrologic extremes requires hydrologic databases that synthesize variables beyond the precipitation, temperature, and streamflow measurements that are more typically available.

Long-term observational records play an important role in understanding and projecting the impact of climate change on hydrological systems. They provide important ground truth for hydroclimatic models, highlighting uncertainties in their representation of certain processes (e.g., rainfall-runoff processes). Trends detected in the observational record are also commonly reliable indicators of future hydroclimatic change (Batibeniz et al., 2020). Despite their potential importance, long-term and spatially extensive databases that contain a range of hydrologic variables relevant to water-balance partitioning (e.g., soil moisture, snow data, vapor pressure deficit) are virtually nonexistent. One reason for limited spatial coverage is that extensive measurements of soil moisture and snow-water content are impractical to measure with gauging stations and uncertain when inferred from current remote sensing techniques, with estimates characterized by limited volumetric representativeness and high uncertainty (Ford & Quiring, 2019). Further, hydrologically comprehensive datasets are available at only a limited, albeit growing, number of catchments, often referred to as hydrologic observatories. Synthesis across these observatories has been hindered by a lack of standardization in variable naming conventions, file formats, time steps, metadata, and data processing procedures, which in turn has slowed the development of the subfield of comparative hydrology (Gupta et al., 2014).

Here we respond to the dearth of long-term, regionally extensive, hydrologically comprehensive databases by presenting CHOSEN (DOI: 10.5281/zenodo.4060384), the Comprehensive Hydrologic Observatory SEnsor Network database, a compilation of publicly available hydrometeorological and hydrological measurements from 30 LTER (Long-Term Ecological Research observatories; Servilla & Brunt, 2011), CZO (Critical Zone Observatories; Zaslavsky et al., 2011), and university field stations in the United State (Kakalia et al., 2021; McNamara, 2017; R. S. Petersky & Harpold, 2018). We developed CHOSEN using a novel operational pipeline that overcomes the challenges associated with a lack of standardization across observatories. The data pipeline ensures accessibility and reproducibility of the data cleaning procedures including quality control, gap-filling, and file formatting, thereby facilitating the expansion of CHOSEN to additional times and catchments. An open-source Jupyter Notebook tutorial with a user interface facilitates the modification of this pipeline to suit the needs of other investigators. Reproducible data analysis pipelines such as this one are an essential part of a modern practice of environmental science that requires rapid data assimilation capabilities to enable rapid response (Fer et al., 2021).

Although CHOSEN was developed to facilitate a range of comparative hydrology studies, we demonstrate another application here in evaluating associations between observed trends in streamflow extremes (both wet and dry) and a wide range of climatic and other hydrologic extremes from a water-balance perspective. Given the limited number of hydrologic observatories and the well-known difficulty in performing attribution analysis on trends in the observational record (Sillmann et al., 2013), this phenomenological analysis represents early progress toward resolving the challenge of understanding the relationship between hydrological and climatic extremes. The primary contributions of this work are to establish a baseline trend assessment for extreme values (high and low, for both magnitude and frequency of the extreme events) and to provide ground-truthing for extreme event detection and attribution analyses that rely on modeled/derived water-balance quantities.

We use CHOSEN to ground-truth four main predictions. First, both low extremes and high extremes in discharge and associated hydroclimatic variables are increasing in magnitude and frequency over a broad spectrum of study areas, with significant trends in frequency more common than trends in magnitude, as has been observed in streamflow records (e.g., Archfield et al., 2016; Hirsch & Archfield, 2015; Mallakpour & Villarini, 2015).

Accepted Article

Second, with respect to "hotspots" of hydrologic and hydroclimatic extremes, we expect that northern latitudes and high-elevation study areas will exhibit the largest proportion of monitored variables with trends in magnitude, given the expectation that climatic forcing at these locations will exceed the envelope of historical variability earlier (Batibeniz et al., 2020). Extreme event frequency trends will reflect climate model projections and previously reported hydrologic observations, with many significant trends concentrated within the eastern, southern, and upper-Midwest portions of the US (Archfield et al., 2016; Batibeniz et al., 2020; Mallakpour & Villarini, 2015). Because climate change forcing may alter water-balance partitioning in competing directions (e.g., enhancing rainfall while also enhancing evapotranspiration), regional hotspots for trends in discharge extremes will not necessarily coincide with regional hotspots for trends in other hydroclimatic extremes.

Third, trends toward wetter conditions will predominantly occur in humid locations, whereas trends toward drier conditions will predominantly occur in more arid locations. This prediction originates from the "drier-in-dry, wetter-in-wet" (DIDWIW) hypothesis from climate models (Feng & Zhang, 2015), which replaces the wet-gets-wetter, dry-gets-drier paradigm (Held & Soden, 2006; Knutson & Manabe, 1995; Wentz et al., 2007) commonly applied to oceans but now thought inapplicable to the terrestrial setting (Byrne & O'Gorman, 2015; Hu et al., 2018).

Fourth, based on findings that discharge extremes result from interactive processes (Byun et al., 2019), changes in the magnitude and frequency of discharge extremes will be associated with changes in the magnitude and frequency of extremes in other hydroclimatic variables in a regionally coherent manner that reflects their contribution to water-balance processes (Table 1). Given that climate-induced changes in water balance stores and fluxes may have opposing effects, associations among trends that accord with the signs in Table 1 will be indicative of dominant water-balance processes triggering changes in discharge extremes. We expect that extremes in antecedent moisture, as represented through soil moisture or snow variables, will exhibit associations with both high and low discharge extremes at many study areas.

Table 1. Hypothesized sign* of correlation between trends in extreme discharge frequency and magnitude and trends in extremes of associated hydroclimatic variables, based on analysis of seasonal anomalies.

2. Data pipeline

The data synthesis followed the workflow (Figure 1) of data cleaning (downloading, quality control, data aggregation, naming standardization), gap-filling (section 2.2), and compilation (section 2.3). We implemented this workflow by using a set of Jupyter Notebooks as a pipeline on data from each study area (e.g., Harris et al., 2020). To make the pipeline reproducible, we provided an interactive Jupyter Notebook as a tutorial for data gap-filling which allows users to interactively tune parameters in the gap-filling functions and graphically view the result. The data products and Jupyter Notebooks are available on the Zenodo (DOI: 10.5281/zenodo.4060384) and GitLab platforms.

Accepted Article

Accepted Article

2.1 Data cleaning

First, any available time series of streamflow, precipitation, air temperature, solar radiation, relative humidity, wind direction, wind speed, SWE, snow depth, vapor pressure, soil moisture, soil temperature, and water isotopes were downloaded for each study area. Subsequent quality control consisted of exclusion of erroneous values (i.e., unrealistic values such as negative precipitation or relative humidity greater than 100%, obvious typos or errors due to equipment malfunction), and cross-checking with pre-flagged entries in the downloaded product. Next, we aggregated time series data to daily time steps if the original time series were on a sub-daily scale: cumulative variables were summed for the day, and rate variables were averaged for the day. Finally, variable names were standardized using the format suggested by Addor et al. (2020) for large sample hydrology datasets.

2.2 Gap-filling methods

Gaps in the cleaned and aggregated daily data (excluding isotope data) were filled using one of three techniques, depending on the length of the gap and availability of complementary data. We applied the three techniques sequentially, meaning gaps not filled by the first technique would undergo the second method, etc (Figure 1). First, for gaps of less than seven days, linear interpolation was applied. Though using linear interpolation may be improper for variables like precipitation, we made this operational decision for reasons of internal consistency, noting that our data processing pipeline gives researchers the necessary information to implement alternative processing conventions.

To fill gaps longer than seven days, we applied spatial regression for study areas that have multiple adjacent stations, and then applied temporal regressions for study areas that have long records (Pappas et al., 2014). To implement spatial regression, we first evaluated the correlation coefficients between the station with missing values and all the other stations in the same study area. We then used the data from the station with the highest correlation coefficient to estimate the linear regression parameters. If the highest correlation coefficient was less than 0.7, or if no data were available from other stations contemporarily, the missing values were reconstructed by the climate catalog technique. In the climate catalog (i.e., temporal regression) method, we filled gaps using data from the most highly correlated year at the same site, selected from among years with at least 9 months of data and a correlation coefficient greater than 0.7 to the missing-data year. Gaussian random noise was added to the resulting regression-based estimate, scaled by the standard deviation of the record of each date in the gap across all years, in order to maintain the variation statistics of the original time series. However, this technique may not be useful for reconstructing non-random variations in time series that are large-scale (i.e. wet and dry years) or small-scale (i.e. before and after a storm).

To assure the quality of the gap-filled data, we deleted any values that exceeded the maximum or fell below the minimum of the original time series. Finally, flags were generated to differentiate between raw, missing, and filled data, indicating the technique used to create each reconstructed data point.

Accepted Article

Figure 1 Data pipeline and visualizations of cleaning methods: a) interpolation, b) spatial regression and c) climate catalog (i.e., temporal regression).

2.3 NetCDF data product

We stored and published the processed data in NetCDF (Network Common Data Form) format. NetCDF data have hierarchical structures and are self-explanatory, which means the descriptions of the attributes of the data tables are accessible from the file by different programming interfaces, for example, C++, Java, Python, and MATLAB. NetCDF is emerging as the data standard for large-sample hydrology, as well as for other large-sample products across the geosciences, particularly climate science and remote sensing (Liu et al., 2016; Signell et al., 2008). The NetCDF library is designed to read and write multi-dimensional scientific data in a well-structured manner. The library enables writing data in several coordinate dimensions, accommodating multiple measurement stations.

We stored the data and metadata from each study area in one NetCDF file. In the NetCDF files, the hydrometeorological variable data and associated data flags are two-dimensional arrays (i.e., by time and location). There is a timestamp variable for conveniently checking the first starting date and last ending date for data in this study area. The grid variable contains information about monitoring stations, providing the names, latitudes, and longitudes and elevations if available.

3. Dataset description

We synthesized data from 30 intensively monitored study areas across the United States (Figure 2). Sixteen of the 30 study areas are from the LTER network (Servilla & Brunt, 2011), 11 from the CZO network (Zaslavsky et al., 2011), and the remaining three are East River, Dry Creek, and Sagehen Creek (Kakalia et al., 2021; McNamara, 2017; Petersky & Harpold, 2018). Table S1 includes additional information about the study areas in the CHOSEN dataset such as data source links, geographical information, and climate conditions.

Figure 2 Geographical distribution of the study areas. “CZO” represents Critical Zone Observatories; “LTER” represents Long-term Ecological Research Stations; “Other” represents observatories managed by other entities.

The availability of different variables in CHOSEN varies by site. The H.J. Andrews and Bonanza LTER datasets contain all 13 variables, with most other datasets having around 10 variables. Discharge record lengths range from three years at Calhoun to 78 years at the San Diego River (California Current Ecosystem LTER), with a median of 19 years.

Figure 3 The span of time series availability and duration across study areas

Discharge and precipitation time series are available in all CHOSEN study areas, and seven catchments have soil moisture and snow measurements with records exceeding five years. Although publicly available water isotope data are limited, we identified six study areas with water isotope time series longer than one year (Figure 4). The measured isotopes include ^{18}O and deuterium in streamflow, precipitation, and snowpack. Note that, unlike other variables, the resolution of isotope data is sparse, usually weekly or biweekly.

Figure 4 Distributions of record spans for a selection of variables in CHOSEN

Accepted Article

4. Extreme events analysis with CHOSEN data

4.1 Methods

Extreme events are occurrences above or below certain thresholds of exceedance over a period of time. In this paper, we evaluated extreme events based on seasonal anomalies, in which we first removed the seasonality calculated by a moving average of 30 days (Figure S1). Then we picked out local minima/maxima in the time series as independent events and identified the high or low extremes as the independent events above the value at the percentile ranking of 95% or below the value at the percentile ranking of 5%. We studied the extreme events of each hydro-meteorological variable with a record longer than 10 years for each study area in CHOSEN, with the exception of water isotopes. If the study area had multiple measurement records for a single variable, we chose the longest.

We used the Mann-Kendall trend test (M-K test) to identify the significance (with a p-value less than 0.05) and sign (increasing or decreasing) of monotonic trends in extreme event magnitudes and frequencies over time (Kendall, 1975; Mann, 1945). For convenience, we refer to significant test results as trends, though we recognize that the M-K test is specific only to the monotonicity, rather than to the magnitude of the trend. The M-K tests were performed on two kinds of statistics: annual counts and the annual median of the extreme event magnitudes, to detect trends in frequency and magnitude, respectively. The analyses were conducted for both high and low extremes and implemented using the python package scikit-learn (Pedregosa et al., 2021). It is worth noting that autocorrelated time series remain a challenge in the M-K test. Autocorrelated time series may artificially inflate test statistics, resulting in false positives in the trend detection (Storch & Navarra, 1999; Yue et al., 2002). Our usage of the annual interval for the statistics decreases the likelihood of within-water-year autocorrelation that would arise from using shorter intervals.

Following the M-K trend analyses, the percentage of extreme-value time series available at each study area (for all variables, excluding isotopes, with record length longer than 10 years) with significant trends was computed as a first step in identifying locations that are “hotspots” for change across multiple hydrologic and hydroclimatic variables. For example, this value would be 25% for a study area with sufficient precipitation and discharge record lengths that exhibited a trend only in high-flow extremes (because one of the four possible extremes -- high flow, low flow, high precipitation, and low precipitation -- exhibited a trend). Hotspots were operationally defined as study areas that exhibited trends in over two-thirds (66.67%) of the available extreme-value time series.

Next, based on significant trends in the magnitude and frequency of extreme discharge, study areas were classified as “wetting” or “drying” with respect to discharge. Specifically, increases in the magnitude of extreme high or low-discharge, decreases in the frequency of extreme low-discharge, and increases in the frequency of extreme high-discharge were all classified as "wetting" trends, and vice-versa. We caution readers that these labels are not intended to apply to total water availability within the study area and that they are not necessarily representative of water availability outside of extreme flow events.

Last, we evaluated whether wetting or drying trends with respect to discharge were associated with trends indicative of wetting or drying in other water-balance stores and fluxes in a manner consistent with a simple water-balance explanation (i.e., Table 1). Namely, we evaluated correlations between significant trends in extreme discharge and significant trends in other monitored hydroclimatic variables. A positive correlation means that both variables trended in the same direction; a negative correlation means they trended in opposite directions. We compared these correlations with our predictions in Table 1 and counted how many correlations matched the predictions. Meanwhile, we identified counterfactuals to the predictions. Here, a counterfactual is an observed trend in the extremes of an associated hydroclimatic variable that has a sign opposite that predicted in Table 1 and, for sites where a trend in high or low-discharge extremes was also detected, is likewise inconsistent with the high or low-flow predictions. This complex definition accounts for the fact that trends in extremes contain no information about within-year timing, and that high-discharge and low-discharge extremes may be sensitive to different hydroclimatic extremes that occur at different times of the year. For example, increasing frequency of low-SWE events associated with an increasing frequency of high-discharge events is a counterfactual if there is no significant trend in low-discharge. However, it is not a counterfactual if that same catchment also shows an increasing frequency of low-discharge events; indeed, low flows may be most sensitive to wintertime delivery of snow, whereas high-flow events may be most sensitive to warm-season rainfall.

4.2 Results

Among 26 study areas with records longer than 10 years, trends in the magnitude and frequency of extreme hydro-climatological and hydrological events were common. All variables in CHOSEN exhibited significant trends in magnitude and in frequency for at least one study area. These trends were distributed among 23 unique sites, with 22 sites exhibiting trends in frequency and 22 sites exhibiting trends in magnitude (Figure 5). On the whole, 81 trends in frequency and 101 trends in magnitude were observed.

Observed trends were indicative of changes to the full suite of water-balance stores and fluxes considered (evapotranspiration, snow storage, soil moisture storage, precipitation, discharge), with "hotspots" of change (defined here as areas with significant trends in over two-thirds of the observed variables) in the southeast (Florida Coastal Everglades), northeast (Hubbard Brook), Pacific Northwest (H.J. Andrews), and Alaska (Bonanza; Figure 5). These hotspots were geographically consistent across magnitude and frequency trends, except for Bonanza, which fell just short of the hotspot threshold for magnitude.

Within sites, trends in frequency and magnitude of extremes generally provided similar information about changes in water balance processes (i.e., Figure 5A compared to Figure 5B). Across sites, trends indicating changes in evapotranspiration were most common (19 study

Accepted Article

areas), followed by changes in precipitation (15 study areas), discharge (11 study areas), snow storage (two study areas), and soil moisture storage (two study areas). Most trends commonly associated with controls on evapotranspiration suggested increases, though at many sites, increasing high-relative-humidity events that accompanied increasing high-temperature events (Figure S2a) exerted competing influences. Trends indicative of changes in extreme runoff, snow storage, and soil moisture storage showed more geographic and temporal heterogeneity (e.g., increases in "high" extremes coupled with decreases in "low" extremes) in the direction of the change compared with trends related to the evapotranspiration. We have repeated this experiment excluding the climate-catalog data and found consistent results (Figure S2b).

Figure 5 Distribution of significant magnitude (A) and frequency (B) trends among study sites and related hydrological stores and fluxes. The size of the bubbles indicates the percentage of the extreme-value time series available at a study area (for variables, excluding isotopes, with record lengths longer than 10 years) that exhibit significant trends in frequency or magnitude. Variables were classified as relevant to evapotranspiration (soil and air temperature, relative humidity, solar radiation), snow (SWE, snow depth), soil moisture, precipitation, or discharge, and those stores and fluxes exhibiting significant trends are indicated by color; note that size of the colored regions within each bubble has no relation to the proportion of trends relevant to that process. Arrows denote whether the trends suggest an increase or decrease in the time-averaged behavior of the associated store or flux (e.g., more frequent low-precipitation extremes would suggest a decrease in precipitation and be denoted with a down arrow in B). When multiple variables associated with the store or flux (e.g., temperature and relative humidity for evapotranspiration) or trends in high and low extremes for the same variable suggest opposing behavior (e.g., increasing low-precipitation and high-precipitation extremes), both arrows are depicted, though a larger arrow indicates the direction implied by a majority of trends.

Similar to the whole suite of variables (Figure 5), extreme discharge exhibited more trends in magnitude (17) than frequency (14), though more study areas (10) exhibited trends in frequency than in magnitude (nine; Figure S2a). Study areas often exhibited significant trends in both low and high discharge events that indicated either consistent wetting or drying (e.g., increasing magnitude of both low- and high-discharge events, or increasing frequency of high-discharge events coupled with decreasing frequency of low-discharge events), with the exception of Hubbard Brook, which exhibited increasing frequency of both high- and low-discharge extremes. Similarly, for study areas showing trends in both frequency and magnitude, the trends pointed consistently toward wetting or drying (Figure 6). Namely, study areas that had trends toward drier conditions with respect to discharge were clustered in the Southeast (Georgia and Florida) and Northwest (Oregon, Idaho, and central California). Meanwhile, study areas that exhibited trends toward wetter or predominantly wetter conditions in terms of discharge were located in the high-elevation West (Colorado, New Mexico), coastal Southwest (southern California), upper Midwest (Michigan), and Northeast (New Hampshire). With respect to the drying observed in the southeast and wetting observed in the montane west, this geographic pattern diverged from the DIDWIW prediction.

Figure 6 Trends in low-flow (A) and high-flow (B) extremes and associated trends in hydroclimatic stores and fluxes consistent with water-balance explanations of how those stores and fluxes impact streamflow (Table 1). Shades of blue suggest wetting trends (with respect to

Accepted Article

Accepted Article

the particular discharge extreme plotted, based on trends in frequency and/or magnitude), while shades of red suggest drying trends. Purple represents a combination of a wetting trend (i.e., increased low-flow magnitude) and drying trend (i.e., increased low-flow frequency). Colored outlines show one or more associated significant trends in other hydroclimatic extremes that are consistent with a simple water-balance explanation (i.e., Table 1). Circles without outlines imply that the study area has no univariate associations between extreme discharge and other hydroclimatic extremes consistent with a simple water-balance explanation.

Five out of nine study areas exhibiting significant trends in low-flow events and three of the 10 study areas exhibiting trends in high-flow events showed associated trends in other hydroclimatic variables consistent with the predictions in Table 1. In the Southeast, drying trends in discharge extremes were associated with trends indicative of increased evapotranspiration, while the drying trends in discharge extremes observed in the northwest had more diverse associations: decreased precipitation, decreased soil moisture, and increased evapotranspiration (Figure 6; Tables 2 and 3). In the northeast, more frequent low-discharge extremes were associated with trends indicative of increased evapotranspiration and more frequent low-precipitation extremes. Meanwhile, wetter discharge extremes had almost no associations with trends in hydroclimatic variables, with the exception of Hubbard Brook, where wetter high-flow extremes were associated with more frequent and higher-magnitude precipitation extremes.

Overall, observed associations between discharge extremes and extremes in other hydroclimatic variables partially upheld our fourth prediction. Specifically, in many locations, trending extremes in discharge could be associated with trending extremes in one or more water balance processes. As expected, interactions among these processes were complex and often confounding; study areas with associations consistent with changing hydroclimatic inputs also commonly exhibited counterfactuals (Tables 2 and 3). Just two study areas exhibited associations that were only counterfactual to the water-balance expectations; both California Current Ecosystem and Jemez exhibited wetter low- and high-flow extremes based on trends in discharge frequency and magnitude (Figure 6), despite trends indicative of higher evapotranspiration. The remainder of the study areas with significant trends in discharge extremes exhibited no other trends in hydroclimatic variables. In contrast to our third prediction, widespread associations between variables indicative of antecedent moisture (i.e., soil moisture, snow depth, SWE) and discharge extremes were not observed. Only at H.J. Andrews was an association between soil moisture and discharge extremes observed.

Table 2. Hypothesis testing of correlation between trends in extreme discharge frequency and trends in the frequency of extremes of associated hydroclimatic variables. Only study areas with significant trends in discharge are included. Counterfactuals are compiled across the low-flow and high-flow analyses and represent correlations between significant trends in discharge and significant trends in other monitored variables that have signs opposite those depicted in Table 1.

Table 3. Hypothesis testing of correlation between trends in extreme discharge magnitude and trends in the magnitude of extremes of associated hydroclimatic variables. Only study areas with significant trends in discharge are included. Counterfactuals are compiled across the low-flow and high-flow analyses and represent correlations between significant trends in discharge and significant trends in other monitored variables that have signs opposite those depicted in Table 1.

4.3 Discussion

CHOSEN contains an uncommon breadth of variables that allows for analysis of trends in multiple extremes, which is typically beyond the scope of observational extreme events studies. One advantage of analyzing multiple types of extremes simultaneously is the potential to evaluate multiple types of wetting or drying processes that affect different hydrological stores and fluxes. Such an analysis addresses the critique that the pronouncement of "wetting" or "drying" based on trends in a single variable (e.g., discharge, soil moisture, evapotranspiration flux) may be misleading (Roth et al., 2021). Indeed, our overall portrait of trends in hydrologic and hydroclimatic extremes (Figure 5) confirms that processes typically assigned the label "drying" or "wetting" may coexist within single locations (e.g., co-occurrences of "up" arrows for precipitation and "down" arrows for discharge or soil moisture). Further, with respect to single variables within single locations, trends in extremes often indicated both "wetting" and "drying" by exhibiting an increase in the magnitude of high extremes coupled to a decrease in the magnitude of low extremes. With respect to discharge, however, trends in low and high extremes tended to point toward consistent wetting or drying within individual study areas (i.e., Figure 6A compared to 6B), evidencing a shift in the whole distribution of streamflow, as has also been overwhelmingly observed at the global scale (Gudmundsson et al., 2019).

Though most observational studies have been limited to one type of extreme, climate modelers have used a multivariate Climate Extremes Index (Gleason et al., 2008) to identify likely "hotspots" of combined wet, dry, hot, and cold extremes from downscaled global climate models (Batibeniz et al., 2020), which our observations largely corroborate. Consistent with our finding of multivariate extreme "hotspots" in south Florida, Oregon, and New Hampshire, study indicated that by 2050, Florida, New England, and the Pacific Northwest are likely to develop the most extreme conditions across a suite of variables (Batibeniz et al., 2020; Figure S3). Note that they did not consider Alaska, our fourth hotspot, but they did find that the extreme conditions would extend into the Rocky Mountain west, where our observations indicated a less comprehensive set of trends to date. Furthermore, the study found that these patterns were primarily driven by warming and drying conditions, as the majority of areas did not exceed the historical envelope of variability for intense precipitation events until 2050. Namely, the Florida hotspot primarily arose from extreme warm conditions, consistent with the decreased discharge/increased evapotranspiration associations that we observed. Meanwhile, the Pacific Northwest and New England hotspots predominantly arose from extremely dry conditions, consistent with our observed decreased soil moisture and increased evapotranspiration trends at H.J. Andrews and increased evapotranspiration trend at Hubbard Brook, together with an increased frequency of low-discharge extremes.

Although our observations generally upheld climate model-based projections of extreme event hotspots, they deviated from projections and previous observations in a few ways. First, our analysis resolved no trends in extremes for any of the five sites in the Mid-Atlantic region (Figure 5), in contrast to projected drying trends in streamflow extremes (Naz et al., 2016), observed wetting trends in high-streamflow extremes (Archfield et al., 2016), projected increases in hurricane-related flood hazards (Marsooli et al., 2019), and observed increasing trends in the climate extremes index for the 1981-2005 period, encompassing both drought and intense wet events (Batibeniz et al., 2020). Our lack of trends in the Mid-Atlantic region was likely strongly

Accepted Article

driven by the limited data record length (among the shortest of all sites for variables other than discharge) for most of the Mid-Atlantic observatories (Figure 3). To test whether short record length had impeded our ability to detect trends, we carried out two-sample t-tests. Results showed that the time series with identified trends for both frequency and magnitude of extreme events were significantly longer (p<0.01) than those with no trends. For most of the study areas, the record lengths for discharge, precipitation, and air temperature were sufficient, whereas, for other hydroclimatic variables, the scarcity of long records substantially restricted the trend analysis.

In addition to insufficient record lengths for some variables and study areas, geographic undersampling may also explain discrepancies between our findings and the literature. In the Mid-Atlantic region, both high-flow (Archfield et al., 2016) and low-flow (Kam & Sheffield, 2016) trends exhibit strong variability in sign and significance, making it likely that observations from just a few sites would not be representative of the regional mean. Undersampling of the Midwest in CHOSEN might also explain why we observed just one study area with a significant change in the frequency of high flows in this region (i.e., Kellogg), despite the prevalence of increased flood frequency observed for the region in other observational studies (Ahn & Palmer, 2016; Hirsch & Archfield, 2015; Mallakpour & Villarini, 2015).

The geographic undersampling inherent in CHOSEN may additionally provide an explanation for why our second prediction--that we would observe more trends in extreme event frequency than magnitude, as observed in geographically extensive discharge records (Hirsch & Archfield, 2015)--was not upheld. In contrast to this prediction, we observed a comparable number in trends in magnitude as in frequency (Figure 5). Small-sample bias may have been exacerbated in CHOSEN by the preferential siting of many of the observatories in areas where rapid climate-driven change is expected. Furthermore, given that observed trends in extreme discharge are highly variable in sign and significance throughout the US (Ahn & Palmer, 2016; Archfield et al., 2016), it is not unexpected that the slight dominance of magnitude trends among our subset of sites would emerge from chance. A second potential explanation for the surprisingly large number of trends in magnitude is that many of these trends involved temperature or variables thought to be directly driven by temperature (Figure 5), and recent climate models (Batibeniz et al., 2020) suggest near-term (median: by 2025) emergence from the envelope of historic variability for temperature for most of the US.

Though undersampling provides a partial explanation for why aspects of our first and second predictions were not upheld, discrepancies from the DIDWIW prediction are likely not attributable to random sampling artifacts. Consistently across sites and variables, study areas in the arid Southwest showed trends toward wetter extremes, reflected in precipitation and discharge magnitude and frequency trends, while those in the humid Southeast showed trends toward drier extremes, reflected in discharge and evapotranspiration-related trends (Figure 5). This discrepancy underscores the importance of considering multiple variables in assessing wetting and drying trends (*sensu* Roth et al., 2021); the DIDWIW hypothesis was developed based on analysis of long-term, remotely sensed soil moisture changes between 1979 and 2013 (Feng & Zhang, 2015), whereas the increase in intense precipitation events forecasted for the Southwest (Batibeniz et al., 2020) may trigger high-flow extremes through Hortonian overland flow without a long-term increase in soil moisture, which would be consistent with our limited

Accepted Article

observations. Meanwhile, in humid environments like the Southeast, evapotranspiration may impact peak flow volumes while soils remain moist. Further, the soil moisture observations from 1979 to 2013 in Feng and Zhang (2015) may not have captured more recent changes in the Southwest present in CHOSEN. In fact, it is likely that the trends detected in this analysis are recent, as a 1981-2005 observational study of historical trends in intense precipitation events also shows no significant trends for the region (Batibeniz et al., 2020). Our results, taken together with model projections (e.g., Batibeniz et al., 2020), suggest that the DIDWIW paradigm will become less applicable as climate change advances.

Our ability to attribute observed trends in discharge to changes in dominant water balance processes was limited by the logical incongruity of correlative associations and causality and by a lack of long-term records of soil moisture and/or snow storage in most study areas. Nonetheless, the associations depicted in Figure 6 are generally consistent with previous studies that attribute changes in extreme discharge to underlying hydrological processes. In a statistical study based on precipitation and temperature measurements and modeled soil moisture and snowmelt, Berghuijs et al. (2016) found that increasing soil moisture storage is a strong predictor of extreme high-discharge throughout the Pacific Northwest, consistent with the soil moisture/high-flow association we found at H.J. Andrews (Figure 6B). Further, the association between precipitation extremes and high-flow extremes that we found at Hubbard Brook and Reynolds Creek (where snow data records were too short for trend analysis) may be indicative of the importance of extreme precipitation for the rain-on-snow events found to be the dominant factor explaining trends in high flow for these regions (Berghuijs et al., 2016). Meanwhile, climate-model based attribution of decreasing magnitude of low flows in the Southeast to warmer temperatures (Hayhoe et al., 2007) is consistent with our observations (Figure 6A), as is a statistically based attribution of decreased low flows in Idaho to decreased precipitation inputs (Kormos et al., 2016). However, in contrast to the Kormos et al. study, we found no association between precipitation and low-flow extremes in Oregon (H.J. Andrews). Instead, we found associations to soil moisture and evapotranspiration extremes, the former of which was not considered in their study.

Attributional studies in the literature suggest mechanisms that may explain observed trends in discharge extremes that were not associated with other trends in our study (Figure 6). Increasing frequency and/or magnitude of high-flow extremes observed at the Kellogg (Michigan), Boulder Creek (Colorado), and Jemez (high-elevation New Mexico) observatories may be attributable to increasingly rapid snowmelt events triggered by warmer temperatures or rain on snow (Mallakpour & Villarini, 2015). These mechanisms would not be captured by our data, which lacked long-term snow records for these sites, or our analysis, which did not consider multivariate interactions between temperature or precipitation and snow storage. Meanwhile, less snow storage over time as a result of precipitation falling increasingly as rain instead of snow may explain drying trends in both high- and low-flow extremes at the Providence observatory (McCabe & Wolock, 2009; Miller et al., 2003). Lastly, climate models suggest that the wetting trends projected for the Southwest (e.g., California Current Ecosystem) are attributable to increased total precipitation delivery (Heidari et al., 2020), which might not be reflected in precipitation extremes.

Accepted Article

Attributional studies typically assume that evapotranspiration plays no role in high-discharge extremes (e.g., Berghuijs et al., 2016 and Table 1 of this study). However, this assumption may not be valid for coastal and low-gradient parts of the Southeast, where watershed areas are large, flows are slow-moving, and the highest flows occur during the warmest part of the year and are not associated with snowmelt or frontal systems. At both the Georgia Coastal Ecosystem and Florida Coastal Everglades observatories, decreasing trends in the magnitude and/or frequency of high flow extremes are observed despite increasing (Florida) or no significant (Georgia) trends in high-precipitation extremes (Figure 5). Both of these areas, however, have exhibited increasing temperature trends (Figure S2a) that are among the strongest in the US (Batibeniz et al., 2020).

In summary, though our study was not attributional, it supports other attributional studies in suggesting that drying shifts in extreme streamflow in the Pacific Northwest and Southeast are likely linked to decreased precipitation inputs, decreased soil moisture, and increased evapotranspiration due primarily to warming. Wetting shifts in streamflow extremes are more challenging to explain via simple statistical analyses, as evidenced by a prevailing lack of associations to other hydroclimatic variables (Figure 6). Though our findings fall short of reconciling Sharma et al.'s grand challenge (2018) to attribute changing streamflow extremes to changes in hydroclimatic forcing, they suggest three hypotheses that are potentially addressable through more sophisticated statistical analyses or longer periods of record as CHOSEN continues to grow. First, the preferential location of wetting high-flow extremes in regions with snowpack suggests that these trends may be linked to increasingly rapid snowmelt, due to interactions between temperature or precipitation and snow storage. Second, higher rates of evapotranspiration may decrease high-flow extremes in locations without a snowmelt peak or dominantly frontal mechanisms of precipitation delivery. And finally, given the modeling results of Berghuijs et al. (2016) and the observed association at H.J. Andrews, changes in soil storage (Dymond et al., 2014) likely also drive changes in streamflow extremes in many regions.

5. Conclusions

To the best of our knowledge, the CHOSEN database is the largest open-source collection of comprehensive data from hydrological observatories, containing variables important to understanding water-balance partitioning that are not typically present in existing large-sample databases. It thus fulfills critical data needs for comparative hydrology. In particular, it lays a foundation for studies that establish hydrologic baselines, synthesize information on multiple aspects of "wetting" and "drying," ground-truth model projections of highly uncertain, derived hydrological quantities, and attempt to attribute observed changes to underlying hydrological processes.

Our simple synthesis of trends in hydroclimatic extremes generated generally consistent results with model projections and statistical studies that use derived quantities for soil moisture, instilling confidence in model projections. Consistency was strong in the identification of geographic hotspots for multivariate change in extremes and in the hydrologic stores and fluxes dominantly associated with those extremes. However, observations were less consistent with projections of discharge trends (Naz et al., 2016). Namely, many areas where we resolved drying trends in high-flow extremes (i.e., red points in Figure 6B) were projected to exhibit wetting trends by 2050, with the exception of south Florida, where both model projections and observed

Accepted Article

trends indicated drying. We propose that this inconsistency may reflect late emergence (i.e., around or after 2050) from the historic envelope of variability for wet extremes in most regions of the US (Batibeniz et al., 2020) rather than fundamental flaws of the model.

Impending emergence from the envelope of historical variability for both wet and dry extremes underscores the need for synthesis products from hydrologic observatories that can document baselines for wetting or drying across different components of the water balance. Our analysis, for example, suggests that in the Southwest, which is projected to show wetter extremes by 2050 (Batibeniz et al., 2020; Naz et al., 2016), a signature of wetting extremes in both precipitation and streamflow (Figure 5) is emerging. It further suggests that this emergence is recent, as these trends were absent in 1981-2005 observations (Batibeniz et al., 2020). The emergence of this wetting trend, together with drying in the Southeast with respect to discharge and evapotranspiration extremes, suggest that the WIWDID paradigm may be inadequate to describe ongoing climate-induced hydrological change across a suite of variables.

Lastly, though simple associations between hydroclimatic and hydrologic extremes were often consistent with a water-balance framework (Table 1) and prior attributional studies (Section 4.3), they were not sufficient to attribute most wetting trends in streamflow extremes to underlying mechanisms. These shortcomings underscore the need for analyses based on longer-term (i.e., >10 years), comprehensive, and openly available records of soil moisture and snow variables. The data record lengths in CHOSEN will continue to grow, and calls for more soil moisture data nationally are increasingly being heard (e.g., Sungmin & Orth, 2021; Petersky & Harpold, 2018; Wasko & Nathan, 2019). We echo that call and build upon it, highlighting that comprehensive observations related to changes in evapotranspiration (e.g., relative humidity, solar radiation, soil and air temperature, wind speed, and/or direct moisture flux data) may be relevant to explaining a wider range of hydrologic extremes than previously thought.

References

Addor, N., Do, H. X., Alvarez-Garreton, C., Coxon, G., Fowler, K., & Mendoza, P. A. (2020). Large-sample hydrology: Recent progress, guidelines for new datasets and grand challenges. *Hydrological Sciences Journal*, *65*(5), 712–725. https://doi.org/10.1080/02626667.2019.1683182

AghaKouchak, A., Chiang, F., Huning, L. S., Love, C. A., Mallakpour, I., Mazdiyasni, O., Moftakhari, H., Papalexiou, S. M., Ragno, E., & Sadegh, M. (2020). Climate Extremes and Compound Hazards in a Warming World. *Annual Review of Earth and Planetary Sciences*, *48*(1), 519–548. https://doi.org/10.1146/annurev-earth-071719-055228

Ahn, K.-H., & Palmer, R. N. (2016). Trend and Variability in Observed Hydrological Extremes in the United States. *Journal of Hydrologic Engineering*, *21*(2), 04015061. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001286

Ainsworth, T. D., Hurd, C. L., Gates, R. D., & Boyd, P. W. (2020). How do we overcome abrupt degradation of marine ecosystems and meet the challenge of heat waves and climate extremes? *Global Change Biology*, *26*(2), 343–354. https://doi.org/10.1111/gcb.14901

Archfield, S. A., Hirsch, R. M., Viglione, A., & Blöschl, G. (2016). Fragmented patterns of flood change across the United States. *Geophysical Research Letters*, *43*(19), 10,232-10,239. https://doi.org/10.1002/2016GL070590

Accepted Article

Accepted Article

Batibeniz, F., Ashfaq, M., Diffenbaugh, N. S., Key, K., Evans, K. J., Turuncoglu, U. U., & Önol, B. (2020). Doubling of U.S. Population Exposure to Climate Extremes by 2050. *Earth's Future*, *8*(4), e2019EF001421. https://doi.org/10.1029/2019EF001421

Berghuijs, W. R., Woods, R. A., Hutton, C. J., & Sivapalan, M. (2016). Dominant flood generating mechanisms across the United States. *Geophysical Research Letters*, *43*(9), 4382–4390. https://doi.org/10.1002/2016GL068070

Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A. P., Merz, B., Arheimer, B., Aronica, G. T., Bilibashi, A., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B., Claps, P., Fiala, K., Frolova, N., Gorbachova, L., Gül, A., Hannaford, J., … Živković, N. (2017). Changing climate shifts timing of European floods. *Science*, *357*(6351), 588–590. https://doi.org/10.1126/science.aan2506

Byrne, M. P., & O'Gorman, P. A. (2015). The Response of Precipitation Minus Evapotranspiration to Climate Warming: Why the "Wet-Get-Wetter, Dry-Get-Drier" Scaling Does Not Hold over Land. *Journal of Climate*, *28*(20), 8078–8092. https://doi.org/10.1175/JCLI-D-15-0369.1

Byun, K., Chiu, C.-M., & Hamlet, A. F. (2019). Effects of 21st century climate change on seasonal flow regimes and hydrologic extremes over the Midwest and Great Lakes region of the US. *Science of The Total Environment*, *650*, 1261–1277. https://doi.org/10.1016/j.scitotenv.2018.09.063

Cook, B. I., Ault, T. R., & Smerdon, J. E. (2015). Unprecedented 21st century drought risk in the American Southwest and Central Plains. *Science Advances*, *1*(1), e1400082. https://doi.org/10.1126/sciadv.1400082

Diffenbaugh, N. S., Swain, D. L., & Touma, D. (2015). Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences*, *112*(13), 3931–3936. https://doi.org/10.1073/pnas.1422385112

Do, H. X., Mei, Y., & Gronewold, A. D. (2020). To What Extent Are Changes in Flood Magnitude Related to Changes in Precipitation Extremes? *Geophysical Research Letters*, *47*(18), e2020GL088684. https://doi.org/10.1029/2020GL088684

Dymond, Salli F.; Kolka, Randall K.; Bolstad, Paul V.; Sebestyen, Stephen D. 2014. Long-term soil moisture patterns in a northern Minnesota forest. Soil Science Society of America Journal, 78 (S1): S208-S216. https://doi.org/10.2136/sssaj2013.08.0322nafsc.

Feng, H., & Zhang, M. (2015). Global land moisture trends: Drier in dry and wetter in wet over land. *Scientific Reports*, *5*(1), 18018. https://doi.org/10.1038/srep18018

Fer, I., Gardella, A. K., Shiklomanov, A. N., Campbell, E. E., Cowdery, E. M., Kauwe, M. G. D., Desai, A., Duveneck, M. J., Fisher, J. B., Haynes, K. D., Hoffman, F. M., Johnston, M. R., Kooper, R., LeBauer, D. S., Mantooth, J., Parton, W. J., Poulter, B., Quaife, T., Raiho, A., … Dietze, M. C. (2021). Beyond ecosystem modeling: A roadmap to community cyberinfrastructure for ecological data-model integration. *Global Change Biology*, *27*(1), 13–26. https://doi.org/10.1111/gcb.15409

Ford, T. W., & Quiring, S. M. (2019). Comparison of Contemporary In Situ, Model, and Satellite Remote Sensing Soil Moisture With a Focus on Drought Monitoring. *Water Resources Research*, *55*(2), 1565–1582. https://doi.org/10.1029/2018WR024039

Gao, Y., Lu, J., Leung, L. R., Yang, Q., Hagos, S., & Qian, Y. (2015). Dynamical and thermodynamical modulations on future changes of landfalling atmospheric rivers over western North America. *Geophysical Research Letters*, *42*(17), 7179–7186. https://doi.org/10.1002/2015GL065435

Accepted Article

Gleason, K. L., Lawrimore, J. H., Levinson, D. H., Karl, T. R., & Karoly, D. J. (2008). A Revised U.S. Climate Extremes Index. *Journal of Climate*, *21*(10), 2124–2137. https://doi.org/10.1175/2007JCLI1883.1

Gudmundsson, L., Leonard, M., Do, H. X., Westra, S., & Seneviratne, S. I. (2019). Observed Trends in Global Indicators of Mean and Extreme Streamflow. *Geophysical Research Letters*, *46*(2), 756–766. https://doi.org/10.1029/2018GL079725

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

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., … Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hayhoe, K., Wake, C. P., Huntington, T. G., Luo, L., Schwartz, M. D., Sheffield, J., Wood, E., Anderson, B., Bradbury, J., DeGaetano, A., Troy, T. J., & Wolfe, D. (2007). Past and future changes in climate and hydrological indicators in the US Northeast. *Climate Dynamics*, *28*(4), 381–407. https://doi.org/10.1007/s00382-006-0187-8

Heidari, H., Arabi, M., Warziniack, T., & Kao, S.-C. (2020). Assessing Shifts in Regional Hydroclimatic Conditions of U.S. River Basins in Response to Climate Change over the 21st Century. *Earth's Future*, *8*(10), e2020EF001657. https://doi.org/10.1029/2020EF001657

Held, I. M., & Soden, B. J. (2006). Robust Responses of the Hydrological Cycle to Global Warming. *Journal of Climate*, *19*(21), 5686–5699. https://doi.org/10.1175/JCLI3990.1

Hirsch, R. M., & Archfield, S. A. (2015). Not higher but more often. *Nature Climate Change*, *5*(3), 198–199. https://doi.org/10.1038/nclimate2551

Hu, Z.-Y., Chen, X., Chen, D., Li, J., Wang, S., Zhou, Q., Gang, Y., & Guo, M. (2018). "Dry gets drier, wet gets wetter": A case study over the arid regions of Central Asia. *International Journal of Climatology*, *39*. https://doi.org/10.1002/joc.5863

Hughes, T. P., Kerry, J. T., Connolly, S. R., Baird, A. H., Eakin, C. M., Heron, S. F., Hoey, A. S., Hoogenboom, M. O., Jacobson, M., Liu, G., Pratchett, M. S., Skirving, W., & Torda, G. (2019). Ecological memory modifies the cumulative impact of recurrent climate extremes. *Nature Climate Change*, *9*(1), 40–43. https://doi.org/10.1038/s41558-018-0351-2

Ivancic, T., & Shaw, S. (2015). Examining why trends in very heavy precipitation should not be mistaken for trends in very high river discharge. *Climatic Change*, *133*(4), 681–693.

Kakalia, Z., Varadharajan, C., Alper, E., Brodie, E. L., Burrus, M., Carroll, R. W. H., Christianson, D. S., Dong, W., Hendrix, V. C., Henderson, M., Hubbard, S. S., Johnson, D., Versteeg, R., Williams, K. H., & Agarwal, D. A. (2021). The Colorado East River Community Observatory Data Collection. *Hydrological Processes*, *35*(6), e14243. https://doi.org/10.1002/hyp.14243

Kam, J., & Sheffield, J. (2016). Changes in the low flow regime over the eastern United States (1962–2011): Variability, trends, and attributions. *Climatic Change*, *135*(3), 639–653.

Kendall, M. G. (1975). *Rank correlation methods*. Griffin.

Knutson, T. R., & Manabe, S. (1995). Time-Mean Response over the Tropical Pacific to Increased C02 in a Coupled Ocean-Atmosphere Model. *Journal of Climate*, *8*(9), 2181–

Accepted Article

2199. https://doi.org/10.1175/1520-0442(1995)008<2181:TMROTT>2.0.CO;2
Kormos, P. R., Luce, C. H., Wenger, S. J., & Berghuijs, W. R. (2016). Trends and sensitivities of low streamflow extremes to discharge timing and magnitude in Pacific Northwest mountain streams. *Water Resources Research*, *52*(7), 4990–5007. https://doi.org/10.1002/2015WR018125

Liu, H., van Oosterom, P., Hu, C., & Wang, W. (2016). Managing Large Multidimensional Array Hydrologic Datasets: A Case Study Comparing NetCDF and SciDB. *Procedia Engineering*, *154*, 207–214. https://doi.org/10.1016/j.proeng.2016.07.449

Mallakpour, I., & Villarini, G. (2015). The changing nature of flooding across the central United States. *Nature Climate Change*, *5*(3), 250–254. https://doi.org/10.1038/nclimate2516

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, *13*(3), 245–259. https://doi.org/10.2307/1907187

Marsooli, R., Lin, N., Emanuel, K., & Feng, K. (2019). Climate change exacerbates hurricane flood hazards along US Atlantic and Gulf Coasts in spatially varying patterns. *Nature Communications*, *10*(1), 3785. https://doi.org/10.1038/s41467-019-11755-z

McCabe, G. J., & Wolock, D. M. (2009). Recent Declines in Western U.S. Snowpack in the Context of Twentieth-Century Climate Variability. *Earth Interactions*, *13*(12), 1–15. https://doi.org/10.1175/2009EI283.1

McClymont, K., Morrison, D., Beevers, L., & Carmen, E. (2020). Flood resilience: A systematic review. *Journal of Environmental Planning and Management*, *63*(7), 1151–1176. https://doi.org/10.1080/09640568.2019.1641474

McNamara, J. (2017). Long-Term, Continuous Stream Discharge Time Series from Measurement Sites in Dry Creek Experimental Watershed, Southwest Idaho. *Dry Creek Experimental Watershed Data*. https://doi.org/10.18122/B2VG6G

Miller, N., Bashford, K., & Strem, E. (2003). Potential Impact of Climate Change on California Hydrology. *JAWRA Journal of the American Water Resources Association*, *39*, 771–784. https://doi.org/10.1111/j.1752-1688.2003.tb04404.x

Naz, B. S., Kao, S.-C., Ashfaq, M., Rastogi, D., Mei, R., & Bowling, L. C. (2016). Regional hydrologic response to climate change in the conterminous United States using high-resolution hydroclimate simulations. *Global and Planetary Change*, *143*, 100–117. https://doi.org/10.1016/j.gloplacha.2016.06.003

O., S., & Orth, R. (2021). Global soil moisture data derived through machine learning trained with in-situ measurements. *Scientific Data*, *8*(1), 170. https://doi.org/10.1038/s41597-021-00964-1

Pagán, B. R., Ashfaq, M., Rastogi, D., Kendall, D. R., Kao, S.-C., Naz, B. S., Mei, R., & Pal, J. S. (2016). Extreme hydrological changes in the southwestern US drive reductions in water supply to Southern California by mid century. *Environmental Research Letters*, *11*(9), 094026. https://doi.org/10.1088/1748-9326/11/9/094026

Pall, P., Aina, T., Stone, D. A., Stott, P. A., Nozawa, T., Hilberts, A. G. J., Lohmann, D., & Allen, M. R. (2011). Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature*, *470*(7334), 382–385. https://doi.org/10.1038/nature09762

Pappas, C., Papalexiou, S. M., & Koutsoyiannis, D. (2014). A quick gap filling of missing hydrometeorological data. *Journal of Geophysical Research: Atmospheres*, *119*(15), 9290–9300. https://doi.org/10.1002/2014JD021633

Payne, A. E., Demory, M.-E., Leung, L. R., Ramos, A. M., Shields, C. A., Rutz, J. J., Siler, N., Villarini, G., Hall, A., & Ralph, F. M. (2020). Responses and impacts of atmospheric rivers to climate change. *Nature Reviews Earth & Environment*, *1*(3), 143–157. https://doi.org/10.1038/s43017-020-0030-5

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2021). Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*, 6.

Petersky, R., & Harpold, A. (2018). Now you see it, now you don't: A case study of ephemeral snowpacks and soil moisture response in the Great Basin, USA. *Hydrology and Earth System Sciences*, *22*(9), 4891–4906. https://doi.org/10.5194/hess-22-4891-2018

Petersky, R. S., & Harpold, A. A. (2018). *A Long-Term Micrometeorological and Hydrological Dataset Across an Elevation Gradient in Sagehen Creek, Sierra Nevada, California* [Data set]. University of Nevada Reno. https://doi.org/10.5281/zenodo.2590799

Roth, N., Jaramillo, F., Wang-Erlandsson, L., Zamora, D., Palomino-Ángel, S., & Cousins, S. A. O. (2021). A call for consistency with the terms 'wetter' and 'drier' in climate change studies. *Environmental Evidence*, *10*(1), 8. https://doi.org/10.1186/s13750-021-00224-0

Servilla, M., & Brunt, J. (2011). *The LTER Network Information System: Improving Data Quality and Synthesis through Community Collaboration*. *2011*, IN51C-1598.

Sharma, A., Wasko, C., & Lettenmaier, D. P. (2018). If Precipitation Extremes Are Increasing, Why Aren't Floods? *Water Resources Research*, *54*(11), 8545–8551. https://doi.org/10.1029/2018WR023749

Signell, R. P., Carniel, S., Chiggiato, J., Janekovic, I., Pullen, J., & Sherwood, C. R. (2008). Collaboration tools and techniques for large model datasets. *Journal of Marine Systems*, *69*(1), 154–161. https://doi.org/10.1016/j.jmarsys.2007.02.013

Sillmann, J., Kharin, V. V., Zwiers, F. W., Zhang, X., & Bronaugh, D. (2013). Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *Journal of Geophysical Research: Atmospheres*, *118*(6), 2473–2493. https://doi.org/10.1002/jgrd.50188

Slater, L. J., & Villarini, G. (2016). Recent trends in U.S. flood risk. *Geophysical Research Letters*, *43*(24), 12,428-12,436. https://doi.org/10.1002/2016GL071199

Storch, H. von, & Navarra, A. (Eds.). (1999). *Analysis of Climate Variability: Applications of Statistical Techniques Proceedings of an Autumn School Organized by the Commission of the European Community on Elba from October 30 to November 6, 1993* (2nd ed.). Springer-Verlag. https://doi.org/10.1007/978-3-662-03744-7

Swain, D. L., Langenbrunner, B., Neelin, J. D., & Hall, A. (2018). Increasing precipitation volatility in twenty-first-century California. *Nature Climate Change*, *8*(5), 427–433. https://doi.org/10.1038/s41558-018-0140-y

Wasko, C., & Nathan, R. (2019). Influence of changes in rainfall and soil moisture on trends in flooding. *Journal of Hydrology*, *575*, 432–441. https://doi.org/10.1016/j.jhydrol.2019.05.054

Wentz, F. J., Ricciardulli, L., Hilburn, K., & Mears, C. (2007). How Much More Rain Will Global Warming Bring? *Science*, *317*(5835), 233–235. https://doi.org/10.1126/science.1140746

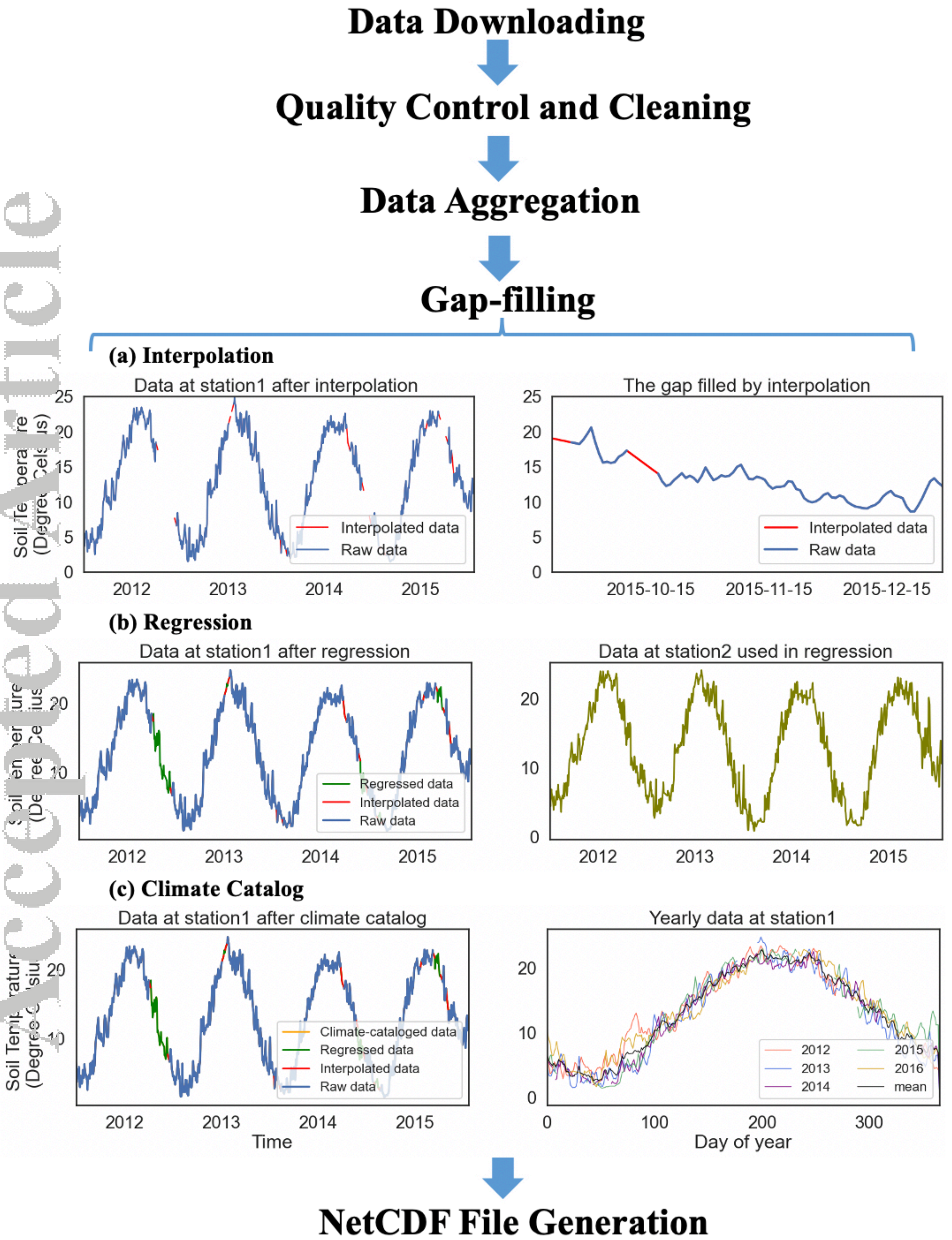
Yue, S., Pilon, P., & Cavadias, G. (2002). Power of the Mann–Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of Hydrology*, *259*(1), 254–271. https://doi.org/10.1016/S0022-1694(01)00594-7

Accepted Article

Zaslavsky, I., Whitenack, T., Williams, M., Tarboton, D., Schreuders, K., & Aufdenkampe, A. (2011). *The Initial Design of Data Sharing Infrastructure for the Critical Zone Observatory*. 6.

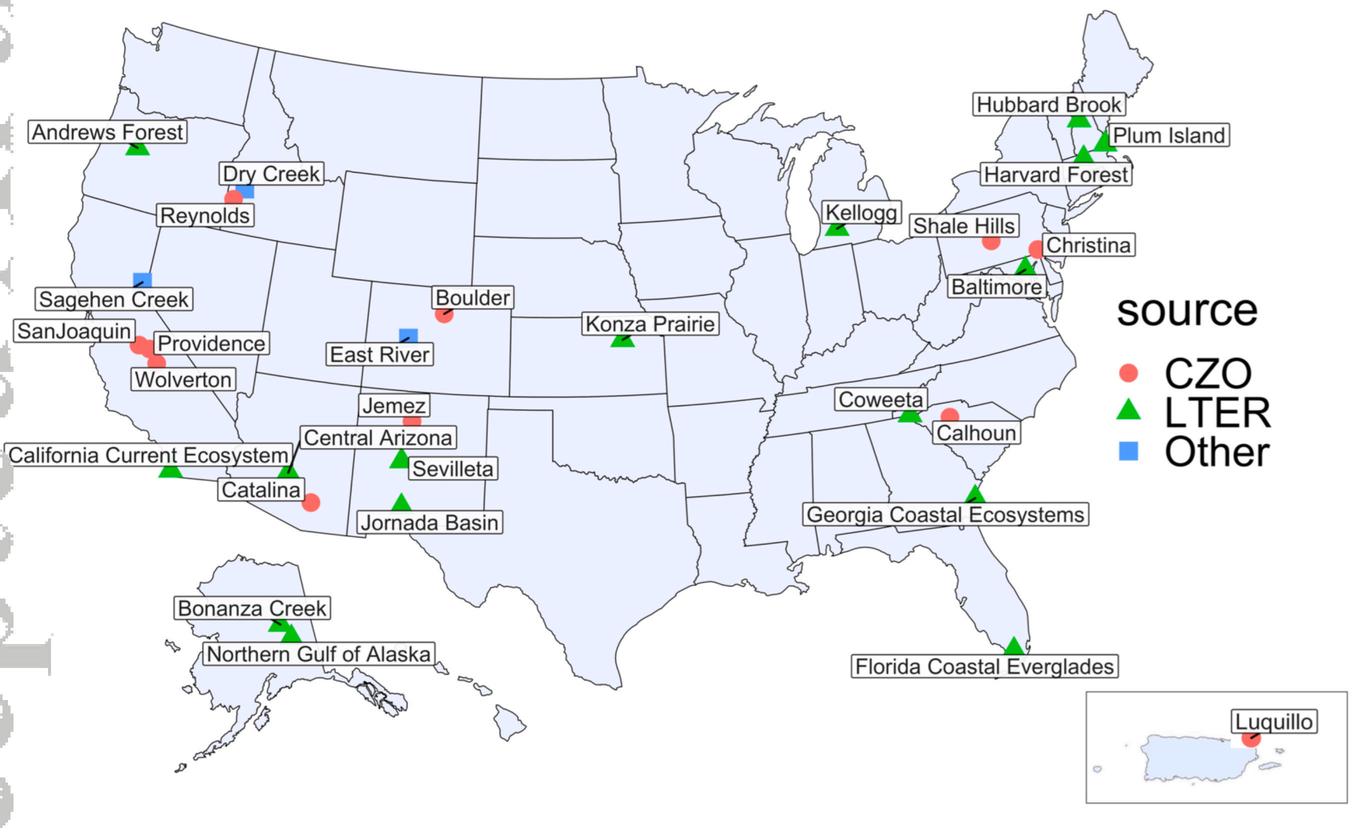
Accepted Article

Accepted Article



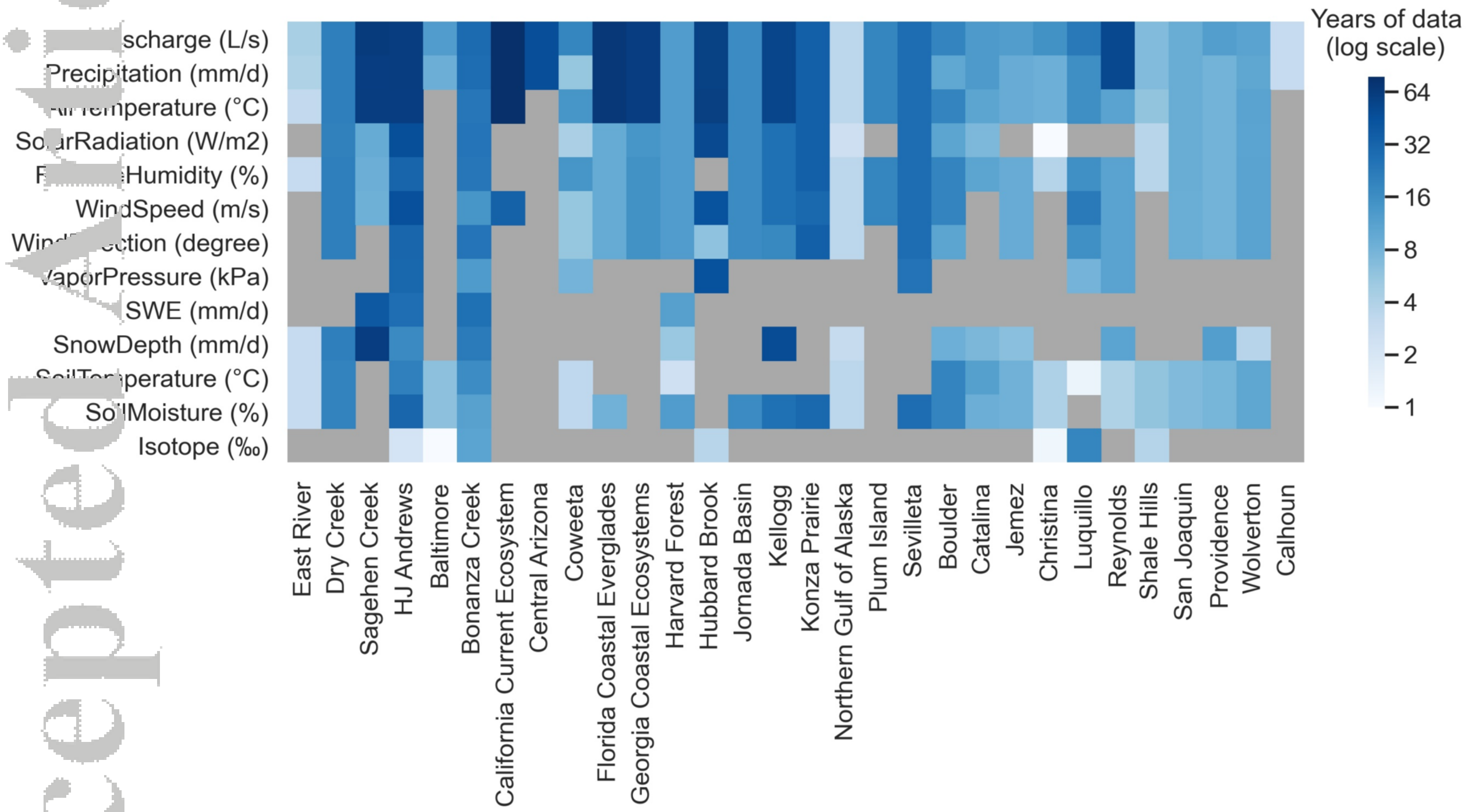
HYP_14429_Figure 1.Data pipeline and visualizations.tiff

Accepted Article



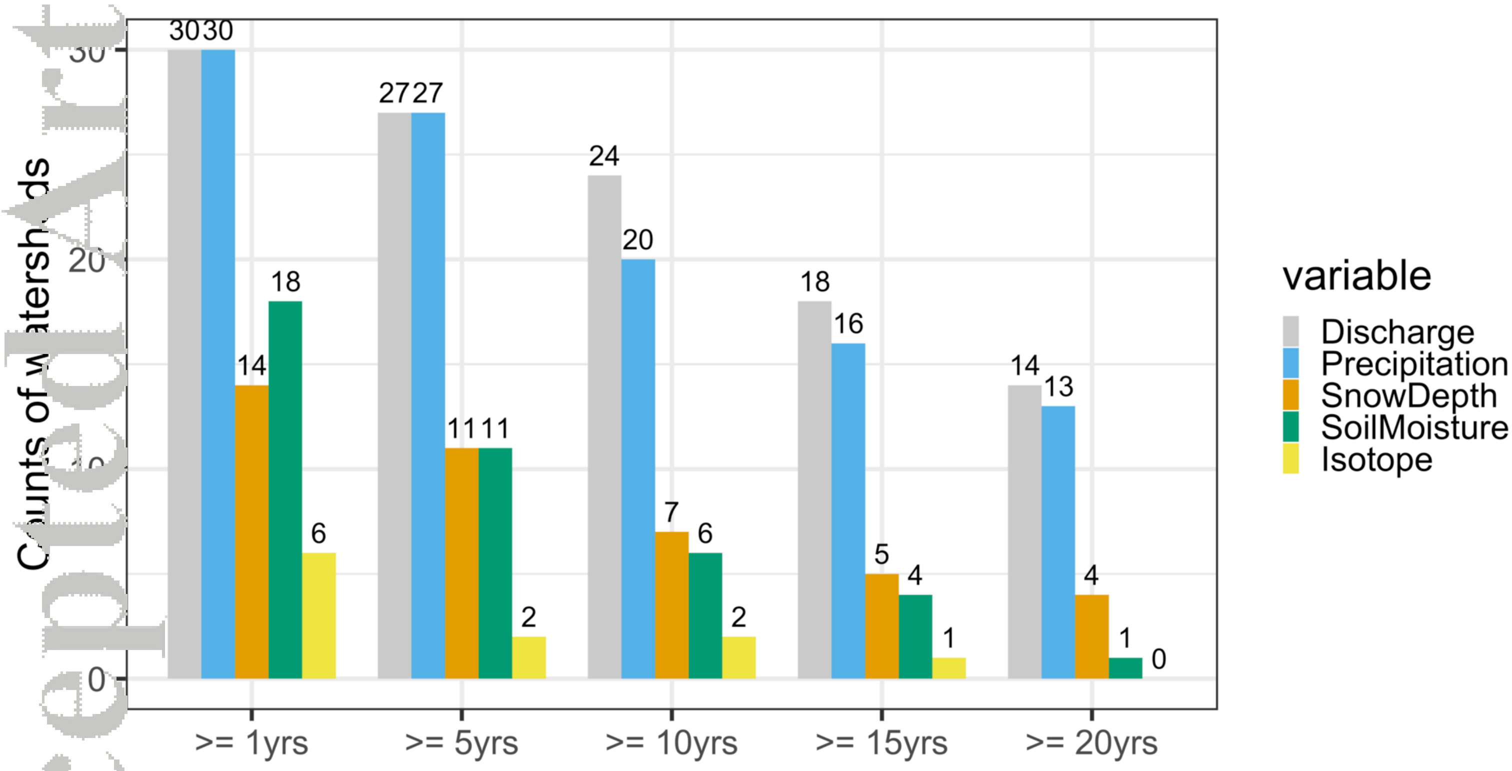
HYP_14429_Figure 2.Geographical distribution of the watersheds.tiff

Accepted Article



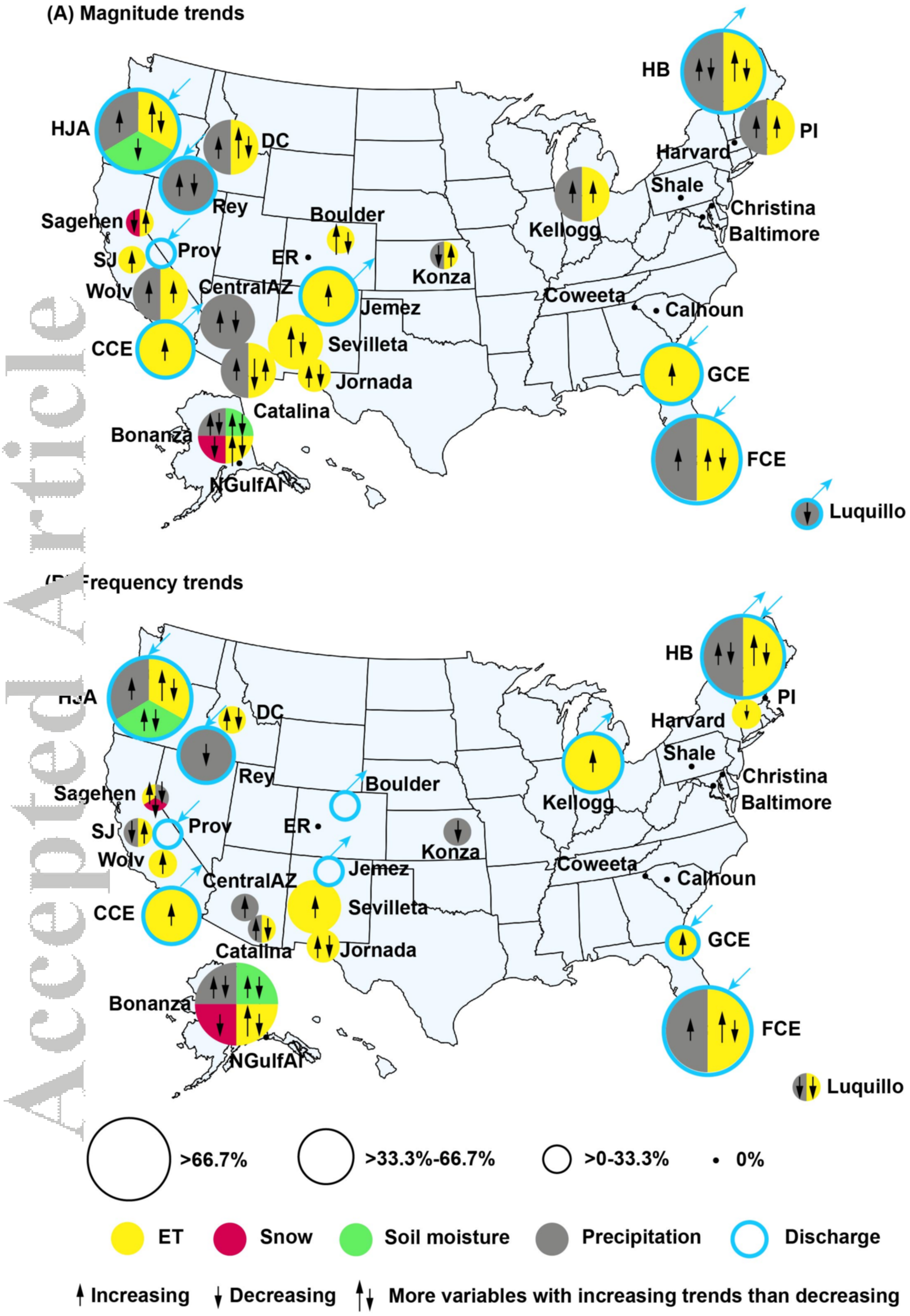
HYP_14429_Figure 3.Span of time series across study areas.tiff

Accepted Article

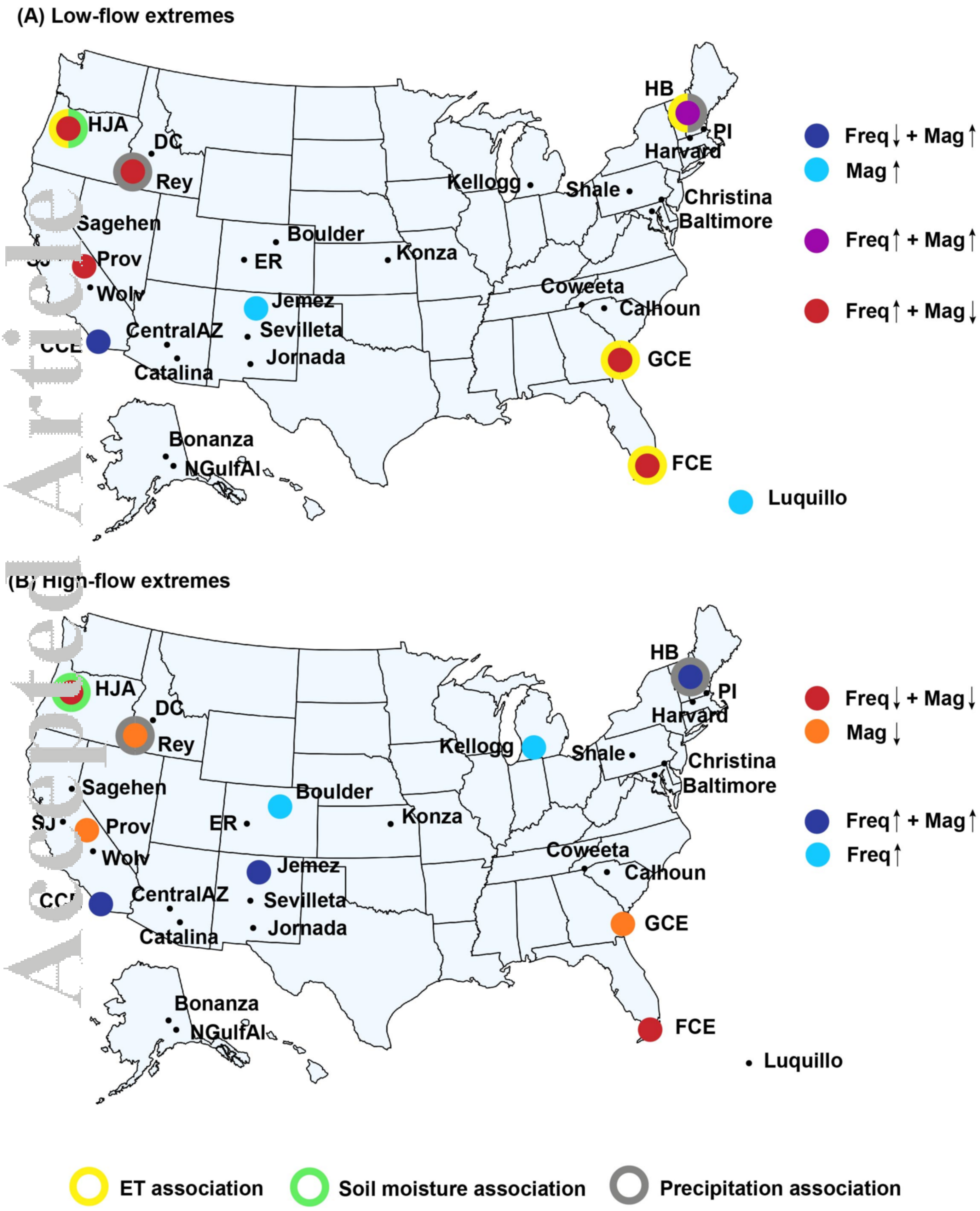


HYP_14429_Figure 4.Distributions of record spans for different variables.tiff

Accepted Article



HYP_14429_Figure 5.Extremes of frequency and magnitude map.tiff



HYP_14429_Figure 6.Low flow and high flow extremes map.tiff

Accepted Article

Table 1. Hypothesized sign* of correlation between trends in extreme discharge frequency and magnitude and trends in extremes of associated hydroclimatic variables, based on analysis of seasonal anomalies.

Correlated extreme	**Sign of correlation, frequency comparison**	**Sign of correlation, magnitude comparison**	**Associated hydrological process**
Expected correlates to low-flow extremes			
Low precipitation (unseasonably dry)	+	+	Precipitation
Low solar radiation (unseasonably cloudy)	-	-	Evapotranspiration
Low relative humidity (unseasonably dry air)	+	+	Evapotranspiration
Low SWE (low snow water content)	+	+	Snow storage
Low snow depth (low snowpack)	+	+	Snow storage
Low soil moisture (unseasonably dry soils)	+	+	Soil storage
High air temperature (unseasonably hot)	+	-	Evapotranspiration
High solar radiation (unseasonably sunny)	+	-	Evapotranspiration
High relative humidity (unseasonably humid)	-	+	Evapotranspiration
High SWE (high snow water content)	-	+	Snow storage
High snow depth (high snowpack)	-	+	Snow storage
High soil temperature (unseasonably hot soils)	+	-	Evapotranspiration
High soil moisture (unseasonably wet soils)	-	+	Soil storage
Expected correlates to high-flow extremes			
Low precipitation (unseasonably dry)	-	+	Precipitation
Low SWE (low snow water content)	-	+	Snow storage
Low snow depth (low snowpack)	-	+	Snow storage
Low soil moisture (unseasonably dry soils)	-	+	Soil storage
High precipitation (unseasonably wet)	+	+	Precipitation
High SWE (high snow water content)	+	+	Snow storage
High snow depth (high snowpack)	+	+	Snow storage
High soil moisture (unseasonably wet soils)	+	+	Soil storage

* The "+" sign of correlation for frequency comparison represents the same direction (both positive or negative) of significant trends (p-value≤0.05) in frequencies of two extremes. The "+" sign of correlation for magnitude comparison represents the positive Pearson correlation coefficient (>0.7) with significance (p-values≤0.05) of trends in magnitudes of two extremes.

Accepted Article

Table 2. Hypothesis testing of correlation between trends in extreme discharge frequency and trends in the frequency of extremes of associated hydroclimatic variables. Only study areas with significant trends in discharge are included. Counterfactuals are compiled across the low-flow and high-flow analyses and represent correlations between significant trends in discharge and significant trends in other monitored variables that have signs opposite those depicted in Table 1.

Study Areas	Discharge record (yrs)	# Total trends	Low-flow extremes frequency	# Consistent with low-flow hypotheses	Low-flow related processes	High-flow extremes frequency	# Consistent with high-flow hypotheses	High-flow related processes	# Counterfactuals	Counterfactuals related processes
H.J. Andrews	62	13	increasing	4	ET, Soil Storage	decreasing	1	Soil Storage	4	P, ET, Soil Storage
California Current Ecosystem	78	4	decreasing	0		increasing			1	ET
Florida Coastal Everglades	68	7	increasing	1	ET	decreasing	0		3	P, ET
Georgia Coastal Ecosystems	61	2	increasing	1	ET					
Hubbard Brook	59	7	increasing	2	P, ET	increasing	1	P	1	ET
Kellogg	55	5				increasing	0		0	
Reynolds Creek	52	3	increasing	1	P				0	
Boulder Creek	19	1				increasing	0		0	
Jemez	12	1				increasing	0		0	
Providence	12	1	increasing	0					0	

Accepted Article

Accepted Article

Table 3. Hypothesis testing of correlation between trends in extreme discharge magnitude and trends in the magnitude of extremes of associated hydroclimatic variables. Only study areas with significant trends in discharge are included. Counterfactuals are compiled across the low-flow and high-flow analyses and represent correlations between significant trends in discharge and significant trends in other monitored variables that have signs opposite those depicted in Table 1.

Study Area	Discharge record (yrs)	# Total trends	Low-flow extremes magnitude	# Consistent with low-flow hypotheses	Low-flow related processes	High-flow extremes magnitude	# Consistent with high-flow hypotheses	High-flow related processes	# Counterfactuals	Counterfactuals related processes
H.J. Andrews	62	13	decreasing	5	ET, Soil Storage	decreasing	2	Soil Storage	3	P, ET
California Current Ecosystem	78	4	increasing	0		increasing	0		1	ET
Florida Coastal Everglades	68	7	decreasing	1	ET	decreasing	0		3	P, ET
Georgia Coastal Ecosystems	61	4	decreasing	1	ET	decreasing	0		0	
Hubbard Brook	59	8	increasing	2	ET	increasing	1	P	2	P, ET
Jemez	12	3	increasing	0		increasing	0		1	ET
Luquillo	23	2	increasing	0					0	
Reynolds	52	4	decreasing	1	P	decreasing	1	P	1	P
Providence	12	2	decreasing	0		decreasing	0		0	