Brian C. Brown[1], Aimee H. Fulerton[2], Darin Kopp[3], Flavia Tromboni[4,5], Arial J. Shogren[6], J. Angus Webb[7], Claire Ruffing[8], Matthew Heaton[9], Lenka Kuglerová[10], Daniel C. Allen[11], Lillian McGill[12], Jay P. Zarnetske[13], Matt R. Whiles[14], Jeremy B. Jones Jr.[15], Benjamin W. Abbott[1]

[1]Department of Plant and Wildilfe Sciences, Brigham Young University, Provo, Utah, USA.

[2]Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmosphereic Administration, Seatle, Washington, USA.

[3]Oak Ridge Institute for Science and Education (ORISE), Corvallis, Oregon, USA.

[4]Global Water Center and Department of Biology, University of Nevada, Reno, Nevada, USA.

[5]Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany.

[6]Earth and Environmental Sciences Department, Michigan State University, East Lansing, Michigan, USA.

[7]Environment and Agriculture Program, University of Melbourne, Victoria, Australia.

[8]The Nature Conservancy in Oregon, Portland, Oregon, USA.

[9]Department of Statistics, Brigham Young University, Provo, Utah, USA.

[10]Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, Umeå, Sweden.

[11]Department of Biology, University of Oklahoma, Norman, Oklahoma, USA.

[12]Center for Quantitative Science, University of Washington, Seattle, Washington, USA.

[13]Department of Earth and Environmental Sciences, Michigan State University, East Lansing, Michigan, USA.

[14]Soil and Water Sciences Department, University of Florida, Gainesville, Florida, USA.

[15]Institute for Arctic Biology, University of Alaska Fairbanks, Fairbanks, Alaska, USA.

Corresponding author: Brian Brown (`bcbrown365@gmail.com`)

Key Points:

- Math often used to describe music (frequency analysis) offers an elegant solution for describing complex timeseries such as streamflow.

- A global frequency analysis of streamflow reveals that flow variability at short timescales is linked with variability at long timescales.
- Machine learning analysis suggests that catchment size and climate mediate linkages in flow regime timescales, and a mechanism is proposed.

Abstract

River flow changes on timescales ranging from minutes to millennia. These variations influence fundamental functions of ecosystems, including biogeochemical fluxes, aquatic habitat, and human society. Efforts to describe temporal variation in river flow—i.e. flow regime—have resulted in hundreds of unique descriptors, complicating interpretation and identification of global drivers of overall flow regime. In this study, we used three analytical approaches to investigate three related questions: 1. how interrelated are flow regime metrics, 2. what catchment characteristics are most associated with flow regime at different timescales globally, and 3. what hydrological processes could explain these associations? To answer these questions, we analyzed a new global database of river discharge from 3,685 stations with coverage from 1987 to 2016. We calculated and condensed 189 traditional flow metrics via principal components analysis (PCA). We then used wavelet analysis to perform a frequency decomposition of each time series, allowing comparison with the flow metrics and characterization of variation in flow at different timescales across sites. Finally, we used three machine learning algorithms to relate flow regime to catchment properties, including climate, land-use, and ecosystem characteristics. For both the PCA and wavelet analysis, just a few catchment properties (catchment size, precipitation, and temperature) were sufficient to predict most aspects of flow regime across sites. The wavelet analysis revealed that variability in flow at short timescales was negatively correlated with variability at long timescales. We propose a hydrological framework that integrates these dynamics across daily to decadal timescales, which we call the Budyko-Darcy hypothesis.

1 Introduction

River flow sculpts landscapes on geological timescales and drives the structure and function of aquatic ecosystems on sub-daily to decadal timescales (Fisher et al., 1998; Pinay et al., 2018; Tucker & Hancock, 2010). For humans, variability in river flow regulates access to freshwater, with extreme flow events such as floods and droughts imposing immense personal and societal costs (Abbott, Bishop, Zarnetske, Minaudo, et al., 2019; Van Loon et al., 2016; Vörösmarty et al., 2010). For ecosystems, changes in water flow through soils, aquifers, and surface-water networks mediate aquatic and riparian biodiversity (Bochet et al., 2020; Hain et al., 2018; N. LeRoy Poff et al., 1997; N. Leroy Poff & Zimmerman, 2010). Additionally, the direction, amount, and timing of flow define terrestrial-aquatic connectivity, controlling the delivery of pollutants to aquatic and marine ecosystems, including human pathogens, excess nutrients, and novel entities (Raymond et al., 2016; Bernhardt et al., 2017; Moatar et al., 2017; Zarnetske et al., 2018; Frei et al., 2020; Gorski & Zimmer, 2021).

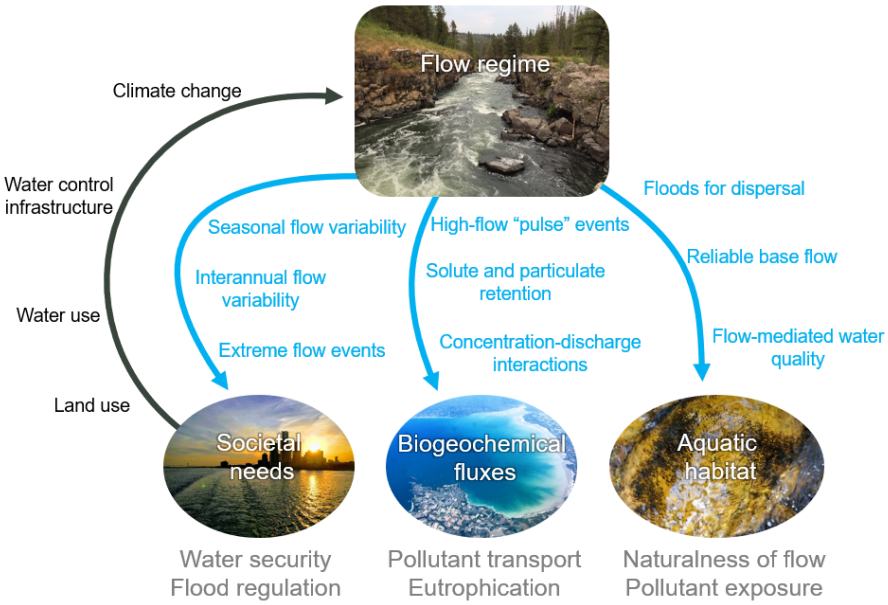

Figure 1. Conceptual diagram representing the societal, biogeochemical, and ecological importance of river flow regime. The relevant dimensions of flow regime are represented in blue, the consequences of flow regime are in gray, and the human influences on flow regime are in black.

In the Anthropocene, human interference with climate, land, and water is threatening aquatic ecosystems, human water security, and biogeochemical cycles at truly global scales (Abbott, Bishop, Zarnetske, Hannah, et al., 2019; Fig. 1). This creates an urgent challenge and opportunity to identify how climate and catchment parameters influence flow regime, and in turn shape the hydrological resilience of socioecological communities (Abbott et al., 2018; Berghuijs et al., 2019; Bunn & Arthington, 2002; Díaz et al., 2019; Harrison et al., 2018; Teixeira et al., 2019). As human modification of landscapes, water, and climate increases (Ascott et al., 2021; Minaudo et al., 2017; Pascolini-Campbell et al., 2021; Zhou et al., 2015), understanding how to describe and predict river flow is urgently needed (Fig. 1).

Though global datasets of river flow are now available (Gerten et al., 2008; Hannah et al., 2011; Masaki et al., 2017), a unified framework for describing and interpreting river flow has not been widely adopted (McMillan, 2021). Because of its importance to society and the environment, over 600 metrics describing flow regime have been proposed (George et al., 2021; Gnann et al., 2021; Jones et al., 2014; N. LeRoy Poff et al., 1997). Many of these metrics are specifically designed to describe key features of flow relevant to society and ecosystems, such

as interannual variability of low flows and the seasonal timing of flooding (Archfield et al., 2014; Carlisle et al., 2010; McMillan, 2021). While these metrics are useful within individual studies, their sheer range and redundancy creates a problem of comparability at regional to global scales (Olden & Poff, 2003). In addition, the strictly hydrological literature has relied heavily on the spectral properties of hydrographs obtained via wavelet decompositions (Carey et al., 2013; Coulibaly & Burn, 2004; Labat, 2005, 2008, 2010; Lafrenière & Sharp, 2003, p. 2003; Larsen et al., 2009; Rossi et al., 2009; Sabo & Post, 2008; Sang, 2013; Sang et al., 2009; Smith et al., 1998; White et al., 2005), which have the advantage of describing variability at multiple timescales simultaneously in a single analysis. However, wavelet decompositions and the suite of other metrics are rarely used in concert, and similarities and differences between the two approaches still need to be quantified. Regardless of the method, understanding variation and similarity in flow regime across biomes and ecoregions could reveal drivers of aquatic ecology and explain differences in success of water management and ecosystem protection in different conditions (Berghuijs et al., 2019; Bunn & Arthington, 2002; Zhou et al., 2015).

One of the problems posed by the diversity of flow regime descriptors is that the primary physical and biological factors controlling river flow remain unclear. Even when constraining the discussion to specific timescales or metrics such as annual flow or runoff ratios during storm events, the physical, biological, and human controls on flow remain debated (Reaver et al., 2020; Zhou et al., 2015). Climatic, surface, and subsurface parameters have been proposed as fundamental controls on the timing and magnitude of river flow across sites, including the amount of soil and aquifer water storage, the relative availability of energy and water, the configuration and size of the surface water network, and the extent and type of vegetation (Carlisle et al., 2010; Lane et al., 2017; Oldfield, 2016; Ryo et al., 2015; Sanborn & Bledsoe, 2006; Zhou et al., 2015).

In this context, we analyzed a global dataset of river flow to compare methods for characterizing flow regime and to identify flow relationships with climatic and catchment factors. We were motivated by the observation that from the various viewpoints of human society, biogeochemical fluxes, and aquatic habitat, no one timescale stands out as singularly important regarding flow regime (Fig. 1). Consequently, we combined traditional flow metrics with a continuous mathematical tool (i.e., wavelet analysis) to describe the time series of river flow. While the broad range of flow behaviors across timescales are rarely analyzed in concert (McMillan, 2021; Olden & Poff, 2003), we hypothesized that variability in flow at different timescales acts as an interacting set of variables. Therefore, considering potential interactions between timescales could shed new light on theories of flow regime because the same climatic and catchment attributes influence flow across timescales. For example, because the relative abundance of energy and water influence vegetation and soil development, hot and dry catchments could simultaneously exhibit high seasonal variability in flow and greater extractive human water use. Likewise, because larger catchments integrate heterogenous subcatchments over larger and longer spatiotemporal scales (Chezik

et al., 2017), we predict they will show less short-term variability but greater sensitivity to long-term changes in water balance.

2 Materials and Methods

2.1 River flow and catchment characteristics data

We obtained daily river discharge time series from the Global Runoff Data Centre (GRDC; https://www.bafg.de/GRDC). We used several criteria to select from the 6544 stations with discharge data from a recent 30-year period of interest (1987-2016). Because continuous time series are required for the calculation of many flow metrics, we first removed stations that had <10 complete water years over the period of interest. This left us with 4762 candidate stations (2399 without any gaps and 2363 with some gaps). For all stations, we removed records for partial water years, i.e., those before the first complete water year or after the last complete water year. For the time series with gaps, we computed the number of days in each missing period and the total number of missing periods. We summarized the number of missing days (e.g., minimum, mean, maximum, and percentiles), and calculated the proportion of days in the record for which data were available. We filled gaps via linear interpolation for stations that met the following criteria: $\leq 25\%$ missing data, the longest data gap was <2 years, and the 75^{th} percentile of consecutive days of missing data was <3 months. For stations that passed this test (1163 of the 2363), we visually inspected the result of interpolation to ensure that obvious peaks or troughs in each station's data record were not omitted. We discarded 104 stations that showed anomalous effects during interpolation, leaving 1059 stations. For the stations with gaps that did not meet our criteria, 509 were located > 1 km from an included station, and many were in data-sparse regions with relatively few observations. Despite their gaps, some of these stations had long data records within the 30-y period of interest. Therefore, we determined which stations had sufficiently long ($\geq$10 y) intact stretches that could be extracted from the longer time series. We were able to salvage an additional 227 stations using an automated approach followed by visual inspection. Therefore, our final set of stations included those with complete records (2399), those with interpolation that met our inclusion criteria (1059), and additional salvaged stations (227), for a total of 3685 stations—56% of the original GRDC stations.

The GRDC streamflow dataset reports the upstream catchment area associated with each station but does not directly reference them to the hydrography we used in this study. As such, differences in data sources may create mismatches between the location of a GRDC station and the upstream catchment we delineated from the integrated Shuttle Radar Topography Mission (SRTM) digital elevation model and the GTOPO30 Digital Elevation Model (DEM, http://files.ntsg.umt.edu/data/DRT). Following Barbarossa et al. (2018), we geo-referenced each station to the pixel that was most similar in catchment area and within 5 km from its original location. We designated stations as high, medium, or low quality if the difference in catchment area was <5%, 5% to 10%, or 10% to 50%, respectively (Barbarossa et al., 2018).

After delineating each watershed, we intersected the shapefiles with 117 variables obtained from a variety of geospatial data sources (supplemental table S1). These variables capture the stream network structure, climate, landcover (including lakes and soils), and anthropogenic impacts (including population density and reservoirs) upstream of each GRDC location. Depending on the parameter, we calculated cumulative values (e.g., total precipitation) or catchment means (e.g., mean annual temperature). Because the configuration and density of stream networks can influence propagation of water and solutes (Godsey & Kirchner, 2014; Helton et al., 2011), we quantified stream network structure using TauDem (Terrain Analysis Using Digital Elevation Models, https://hydrology.usu.edu/taudem/taudem5/). This open source software implements highly parallelized algorithms that can efficiently process large datasets (Barbarossa et al., 2018). We used the AreaD8 function to calculate the number of pixels upslope from a station (i.e., the flow accumulation grid) and the GridNet function to calculate stream network attributes (e.g., stream order and total network length). In addition to comparing catchment attributes with flow regime metrics, we calculated pairwise correlations between catchment characteristics to test for collinearity.

2.2 Characterizing flow regime – conceptual introductions

Frequency decompositions rely on the fact that timeseries are fundamentally related to waves. Waves are phenomena that repeat through time, and can occur in any number of dimensions, though in this scope we consider one-dimensional waves that represent a single variable changing through time. Waves can be described with five fundamental descriptors: (1) through time, the variable may increase or decrease a certain amount away from the mean, or equilibrium point; this amount is known as *amplitude*. (2) Two variables with otherwise identical wave behavior may be out of sync with each other; the degree to which they are in sync is known as the *phase*. (3) Two waves can be perfectly in sync and identical in amplitude but never touch if one is shifted above the other; this is known as a *vertical translation*. (4) A variable can follow any pattern that repeats through time, following a typical sinusoid curve, or a more unusual shape such as a square, triangle, or saw-tooth shape; this is known as *waveform*. Finally, (5) the variable moves up and down either quickly or slowly; this speed is known as the *frequency*. Together, amplitude, phase, vertical translation, waveform, and frequency describe essentially any difference between any two waves (figure 2).

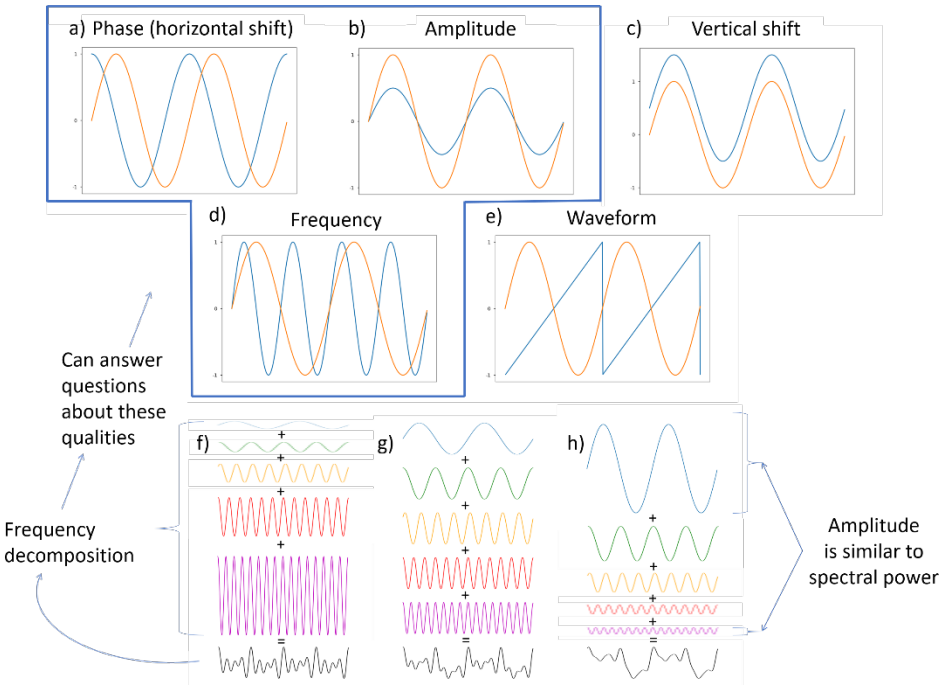

Figure 2. Five fundamental components of flow regime (or any time series): a) amplitude, or the height of the peaks in a wave b) phase, the degree to which a wave is shifted left or right, c) vertical translation, or the distance the wave is raised or lowered from the x-axis, d) waveform, the pattern the wave traces out through time, and e) frequency, the number of times a wave repeats itself over a given timescale. Many of the behaviors in streamflow timeseries relate back to these five fundamental principles. The lower portion of the figure represents frequency decompositions of three timeseries: f) a timeseries dominated by high-frequency variability, g) a timeseries with equal variability across all timescales, and h) a timeseries dominated by low-frequency variability. In each example, adding together the five colored waves produces the complex curve shown in black at the bottom.

The terms amplitude, phase, vertical translation, waveform, and frequency have familiar analogues in hydrology. Consider an imaginary catchment with a hydrograph that follows a perfect sinusoidal curve that goes up and down over the course of a year. The amplitude of this wave plus any vertical shift relates closely to the familiar concept of peak annual flow, and vertical shift minus the amplitude relates to baseflow, or minimum flow. In this catchment, the frequency of one cycle per *year* relates to the timescale containing the most variance. The phase of the wave indicates the time of year snowmelt or monsoon rains occur, and would be opposite for a northern vs southern hemisphere catchment. The

waveform relates to the rate of rise or fall of the year-long increase and decrease in flow. Now imagine a second catchment whose flow follows another perfect sinusoidal wave, but which oscillates at one cycle per two weeks. This "flashy" catchment neither accrues nor loses long-term storage, and might hypothetically occur in a warm climate with no snow and identical rain storms every two weeks. Both catchments exhibit variability in flow, but in the first, the variance is maximized at the timescale of one year, and in the second at two weeks.

However, it is normal for a catchment to exhibit both some flashiness and some seasonal variability. Adding together the perfect sine wave from the first catchment with the perfect sine wave from the second catchment would produce a complex curve that can no longer be described with amplitude, phase, vertical translation, waveform, and frequency, but which more closely resembles a real-world catchment. This process of adding new catchments that epitomize behavior on different timescales could be repeated infinitely many times, producing an ever more complex, and hence more realistic hydrograph, but which could always be decomposed back into a collection of simple waves that can individually be described by the same few, succinct variables. Mathematical tools exist to run this process *backwards*—decomposing a timeseries into a set of perfect sinusoids that together recreate the original timeseries. These are known as *frequency decompositions,* and can be thought of as functioning similarly to a prism, which decomposes white light into a rainbow of colors ranging from high to low frequency, or to a computer program that might take in the sound recording of a symphony and output a musical score. No matter the timeseries, the amplitudes of the resultant decomposed waves at different frequencies relate to the amount of variability in the data that occurs on those timescales, reported in a characteristic known as *spectral power*. Spectral power thus provides a unit for describing variability in streamflow across every timescale present in a hydrograph.

A *metric space* is a space where the distance between points can be described. When we use the term metric space here, we mean an abstract space where a catchment's location describes its streamflow regime, and where the distance between two catchments describes their similarity in streamflow regime. Each axis, then, highlights some unique aspect of streamflow regime. Metric spaces can be composed of many axes or just a few, and some of these axes tend to be more useful than others, depending on how much variance occurs along the axis.

Decision trees are a primal machine learning model that are foundational to many more modern models, such as random forests and gradient boosting forests. Conceptually, decision trees take in an array of prediction features and step-by-step combine multiple points of data along the feature array. Using relatively simple logic, they distill information further and further until a single prediction is made (Myles et al., 2004). Decision trees are generally known to have high bias (generally viewed as negative) with low variance, though they are still occasionally used because of their inherent interpretability.

Random forests are called "forests" because they comprise many individual decision trees, usually of significant depth, whose collective predictions are averaged to produce an output that is generally less biased and more accurate than individual decision tree regressors (Biau & Scornet, 2016). The "random" aspect comes from an innovation in 2001 where successive trees are trained on independent random samples with replacement from the larger dataset (Breiman, 2001).

Gradient boosting regressors are similar to random forest regressors, but they differ in that new trees are added in a way that minimizes error in a targeted, rather than a random fashion. This targeted approach is achieved by adding new trees according to the gradient of a user-defined loss function, which is simply a function which characterizes the error of the model (Elith et al., 2008).

Principal Components Analysis, or *PCA,* projects high dimensional data onto a lower dimensional space where each axis is a linear combination of the original axes in the high dimensional space, and where the number of dimensions projected onto is the user's choice. As an intuitive example, imagine a "high-dimensional" dataset with two dimensions, x and y. If, for every step in the x direction, data tend to take two steps in the y direction, the two axes are redundant and linearly related; a linear regression might draw a line through the two axes with a slope of 2. PCA on these two axes would project data points onto that regression line. That is, instead of listing data points by their x and y coordinates, the PCA projection would list data points by their location on a new axis, z, which is two parts y, and one part x. The "two parts" and "one part" that describe how much each original axis contribute to the new projected axis are referred to as the *loadings matrix.* The loadings matrix effectively describes how correlated (positive or negative) each of the original axes in the high-dimensional space is with the low-dimensional axes PCA projects the data onto. Examining the loadings matrix is one of the best methods for adding interpretability to the abstract axes that result from a PCA projection.

2.3 Streamflow analysis

2.3.1 Similarities between streamflow metrics and frequency decomposition

A correlation analysis was run between each frequency and each of the 189 flow metrics. This was done by calculating the Spearman correlation between a given frequency and a given flow metric across all catchments in the dataset. To account for possible non-monotonic relationships between flow metrics and the frequency decomposition, we also trained machine learning models to predict each of the streamflow metrics using the frequency decompositions as inputs. To account for variability between models and divisions of data, 18 models were trained on each of the 189 streamflow metrics. For each metric, 9 were decision tree regressors and 9 were gradient boosting regressors, and data were divided with an 80:20 training to testing ratio, though the divisions were done randomly and independently for each model. Models were then validated on the 20% portion reserved for testing and an r-squared was calculated by taking a

linear regression between the model's output and the actual values for the given streamflow metric on the 20% testing data. Finally, the "feature importances" were extracted from each model to determine which input features were most important in the models' decision making processes. Models were implemented in Python using the Sci-kit Learn library and feature importances were extracted using the "feature_importance_" method (Pedregosa et al., 2011). To better understand the basic structure of the data, a pairwise spearman rank correlation was calculated between the spectral powers for each frequency.

In addition, we performed PCA on the suite of 189 flow metrics to compress the flow metrics down to a more manageable number of axes. To connect these PCA axes to frequency analyses, we ran a correlation analyses between each of these "PCA metrics" and the spectral power of each frequency. Similar to each of the 189 flow metrics, we also trained 360 machine learning models, with an even split between gradient boosting regressor and random forest regressor models, to predict each PCA metric using the frequency domain, again with an 80:20 split between training and testing data. PCA was implemented in R using the FRK package and the NNGP method (Zammit-Mangion & Cressie, 2017) and machine-learning models were implemented in Python using the Sci-kit Learn library (Pedregosa et al., 2011).

2.3.2 Identifying controls on streamflow regimes with PCA metrics

We ran several analyses to connect flow regime to catchment characteristics. First, we trained three separate machine learning models, a decision tree regressor, a random forest regressor, and a gradient boosting regressor to predict each of the PCA metrics using the 117 catchment characteristics as input features, for a total of 21 models. These models were trained and validated using 10 folds of the data, where models were trained on nine-tenths of the data and validated on the remaining one-tenth, 10 times for a total of 10 validations spanning the entire dataset. During each validation step, r-squared values were taken by running a linear regression between the model's output and the validation data, and the average r-squared for each model was recorded. Prior to training, 470 catchments were removed because they contained insufficient catchement characteristics data. As before, model "feature importances" were extracted to understand which input features (i.e. catchment characteristics) were most important in determining flow regime. A correlation analysis was also performed in which the values for many of the 117 streamflow metrics were correlated against the spectral power for each frequency using Spearman correlation. All correlation analyses were implemented using the Scipy library in python (Virtanen et al., 2020).

3 Results

3.1 Similarities between streamflow metrics and frequency decompositions

Spearman correlation between the 189 flow metrics and each frequency in the frequency decompositions revealed the average maximum coefficient of correlation at any point along the frequency domain was 0.46 (supplemental figures

S1 and S2), suggesting a substantial though not universal relationship between phenomena described by the wavelet decomposition and the 189 flow metrics. On a similar vein, the average r-squared for machine learning models trained to predict the 189 flow metrics exclusively using the frequency decomposition was 0.27 (supplemental figures S3 and S4). The average r-squared for the machine learning models that trained to predict the 7 PCA flow metrics exclusively using the frequency decompostion was 0.42. This may indicate that frequency decompositions such as the wavelet decomposition describe between 30-45 % the same information, or, it may mean that the two methods describe separate phenomena that are correlated.

3.2 Linkages between timescales

In each of the correlations between the 189 flow metrics and frequency decompositions, coefficients of correlation were coherent across a local range of frequencies (supplemental figure S1). However, metrics that were negatively correlated with low frequencies tended to be positively correlated with high frequencies. Because most of the flow metrics were designed to describe one phenomenon or one timescale, these relationships suggested that timescales themselves are inherently linked with each other. Seeking to isolate this phenomenon, we calculated the pairwise correlation between spectral powers for each frequency in each catchment (figure 3). This showed that high frequencies were indeed negatively correlated with low frequencies, meaning that a mass-balance relationship exists between changes in flow that occur over several days and changes in flow that occur over several months or years. More generally, we also found that on average, four distinct timescales emerge globally on which most variability in flow occurs (figure 4). These are multi-day variations, multi-month variations, annual variations, and multi-annual variations. Annual variation was the strongest, followed by multi-month variation and multi-day.

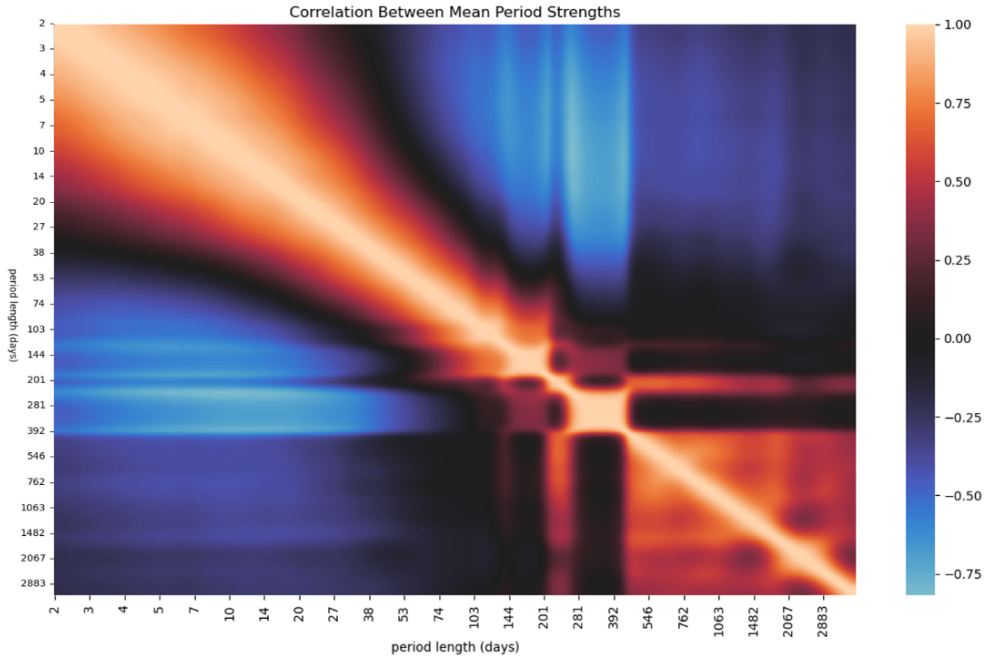

Figure 3. Pairwise correlations between spectral power for each frequency. The coefficient of correlation from the spearman correlation is represented as color, with brighter orange representing a stronger positive monotonic relationship and brighter blue representing a stronger negative monotonic relationship.

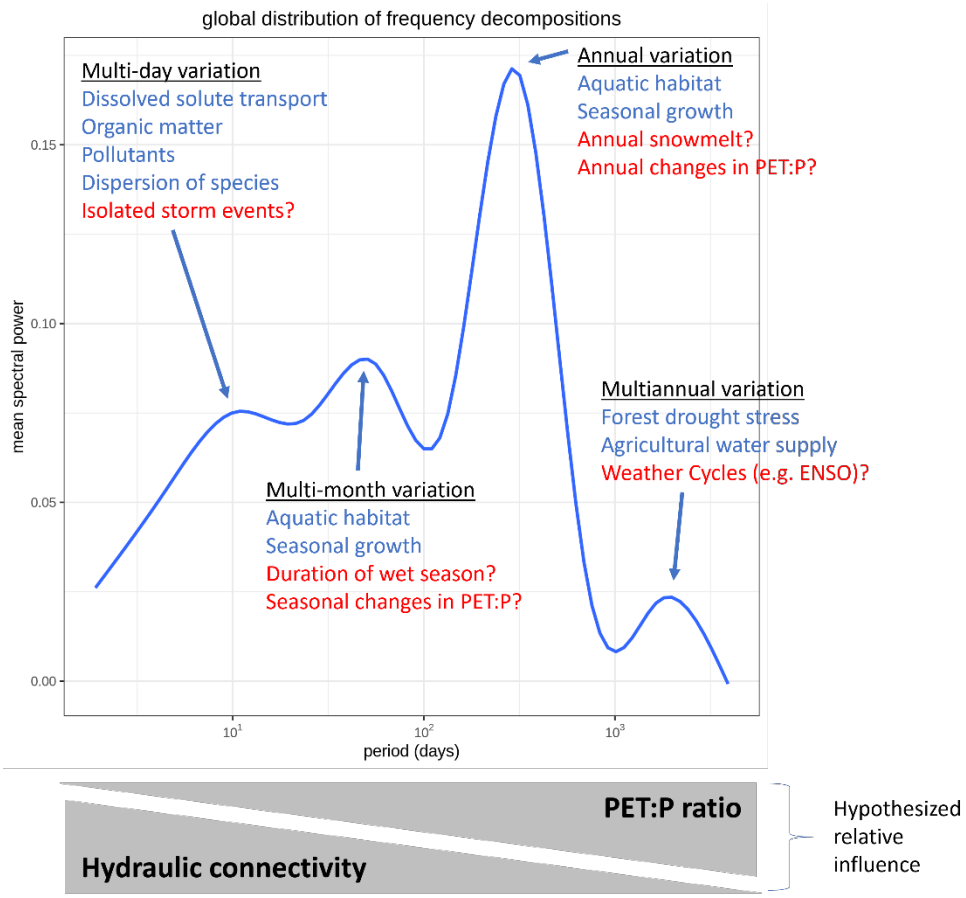

Figure 4. Mean spectral powers across timescales ranging from two days to ten years from our global dataset of streamflow timeseries.

We also found that 70% of the variance in the original 189 flow metrics could be explained in 7 PCA axes, each capturing increasingly less variability in the data (supplemental figure S5). A summary of the loadings matrices of each metric are found in table 1, and more extensive descriptions of the loadings matrices are given in supplemental tables S2-S8. Spatial distributions of the metrics across the world are plotted in supplemental figure S6. When we correlated the PCA metrics to the frequency domain, we found that PCA metrics that explained more variance in the original 189 metrics tended to relate more strongly to the frequency domain (e.g. metrics 1-4), while those that explained less variance in the original metrics tended to relate less strongly to the frequency domain (e.g. metrics 5-7) (supplemental figure S7).

Table 1. List of top seven principal components derived from 250 flow metrics calculated for 3,685 river flow

PCA

Table 1. List of top seven principal components derived from 250 flow metrics calculated for 3,685 river flow

1	
2	
3	
4	
5	
6	
7	

3.3 Identifying controls on streamflow regime with PCA metrics

Consistent with the relatively simple relationship between high and low-frequency variability in river flow, the three machine learning models trained to predict each of the 7 PCA flow metrics using the 117 catchment characteristics suggested just a few dominant controls of flow regime (figure 5). These included dominant contributions of cumulative precipitation for PCA metrics 1, 5, and 7, catchment area for metric 2, and climate variables for metrics 3 and 4, and land cover for metrics 3, 5, and 6. In addition, the length of the timeseries was an important feature for several metrics. We note that the r-squared values across the different models decreased from the higher variance-explaining metrics to the lower variance explaining metrics, and specifically that model accuracy decreased from a maximum of ~0.85 for metric 1 to a maximum of ~0.45 for metric 7 (supplemental figure S8).

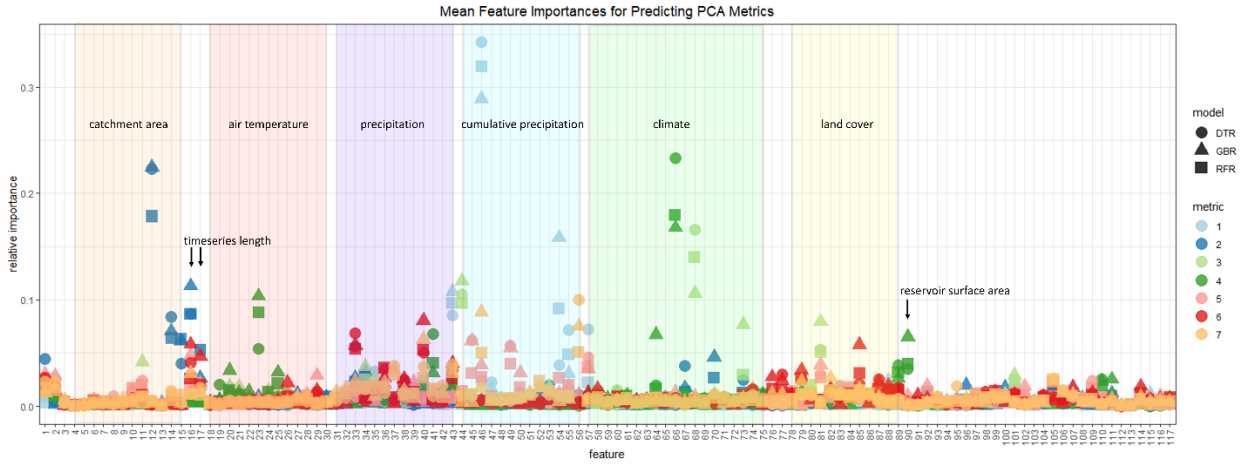

Figure 5. Feature importance from three different machine learning models trained to predict the 7 PCA flow metrics. Model type "DTR" stands for Decision Tree Regressor, while "GBR" stands for Gradient Boosting Regressor, and "RFR" stands for Random Forest Regressor. The 117 input features are listed on the x-axis. Some have been labeled for convenience, and the soft background color on

others describes the general category. For all monthly characteristics, values proceed from January to December, left to right. A full description of the catchment characteristic corresponding to each number can be found in supplemental table S1.

To visualize the relationship between flow metrics and catchment characteristics derived as important via the machine learning analysis, as well as catchment characteristics suggested to be important by hydrological theory, we plotted the relationships between the PCA metrics and several catchment characteristics (figures 6 and 7). The dominant role of catchment size was readily seen, including non-linear relationships between metrics 3, 5, and 7 and catchment size, in which the largest streams tended to behave similar to the smallest streams. Relationships between biome were surprisingly ambiguous given that biome integrates temperature and precipitation data. However, when split, temperature and precipitation did show relationships to the flow metrics, while the effect imposed by land use such as forest cover and net human alterations were less visible (figure 7). A more comprehensive set of visualizations across a broader range of catchment characteristics can be found in supplemental figures S9-S12.

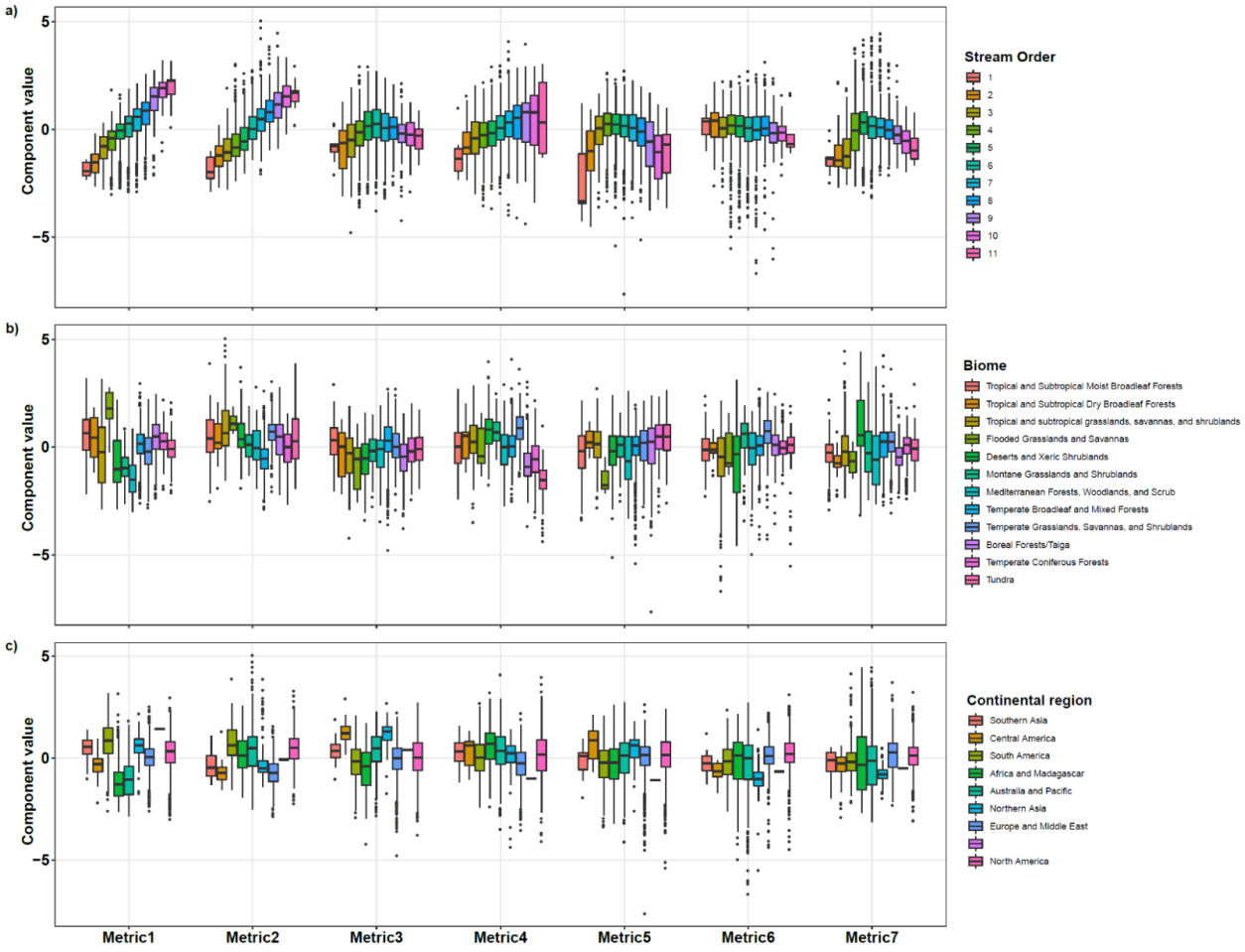

Figure 6. The 7 PCA flow metrics divided according to a) stream order, b) biome, and c) continental region.

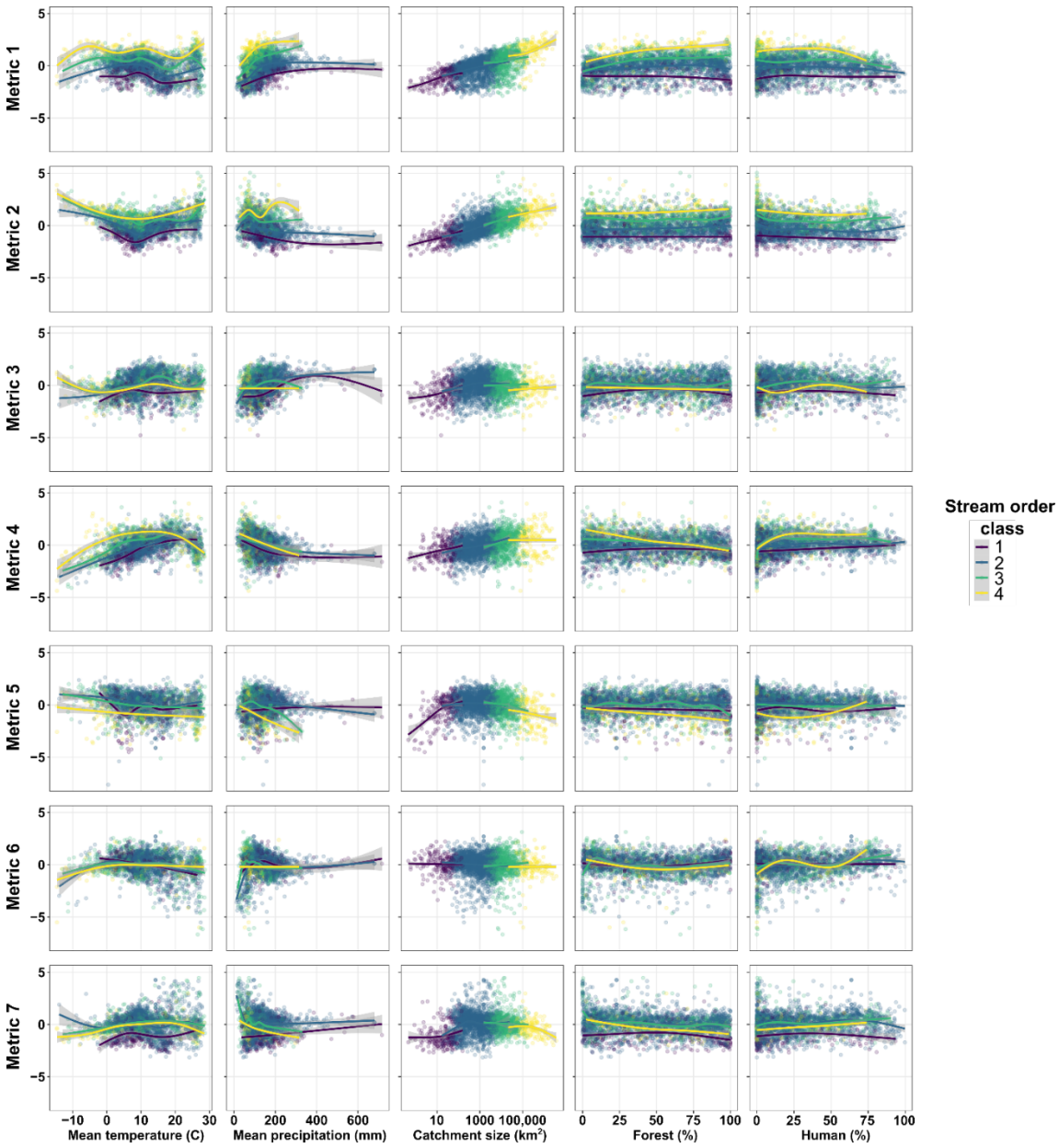

Figure 7. Continuous relationships between 7 PCA flow metrics and catchment properties of mean annual temperature, mean annual precipitation (normalized for catchment size), catchment size, percent forest cover, and percent human influence.

3.4 Identifying controls on streamflow regime with Frequency Decompositions

A correlation analysis between each catchment characteristic and the spectral power for each catchment at each frequency revealed that many streamflow metrics showed similar correlation patterns to those of the correlations between the 189 flow metrics and PCA metrics and the frequency domain (figure 8). For example, metrics of catchment size were strongly negatively correlated with high frequency (short-term) phenomena but were positively correlated with low frequency (long-term) phenomena. Temperature followed a more complex curve where high winter temperatures were positively correlated with multi-day phenomena and negatively correlated with multi-month to year-long phenomena, and where summer temperatures were most strongly correlated with multi-year phenomena. Many land-use characteristics followed similar trends (supplemental figure S13).

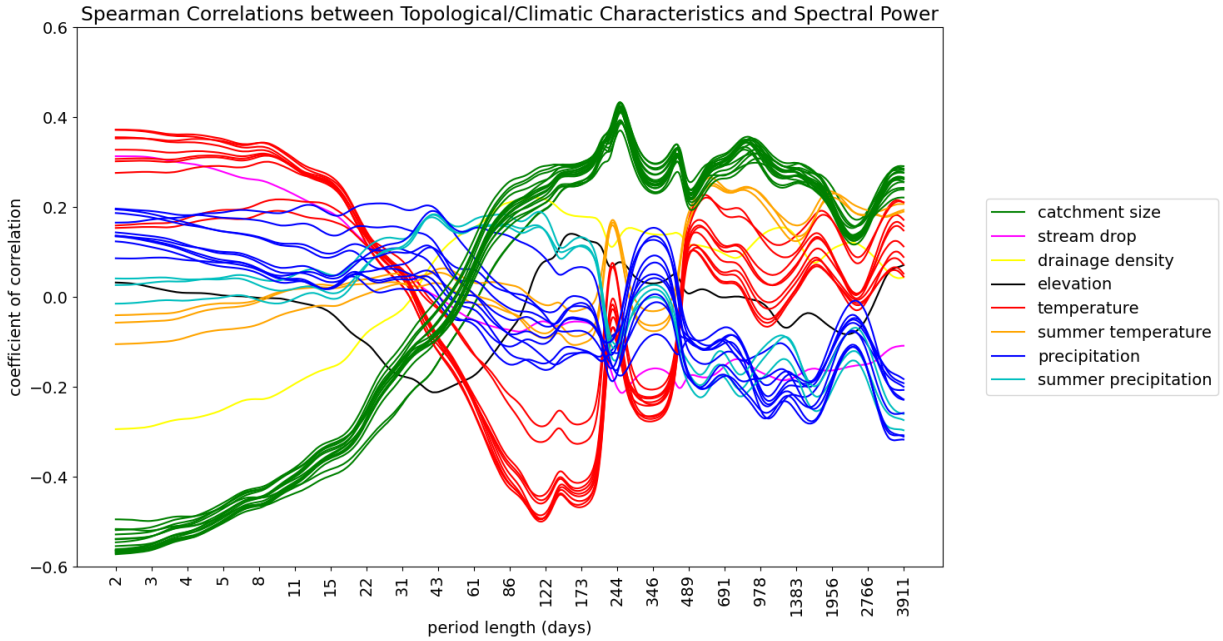

Figure 8. Correlations between catchment characteristics and spectral power across period lengths ranging from two days to almost ten years.

4 Discussion

River networks connect and unite us all by providing critical ecosystem and societal services (Figure 1). Like an ecological heartbeat, river flow rises and falls across myriad timescales, sculpting aquatic habitat, driving biogeochemical flux, and quenching human water needs. In an increasingly human-dominated world of dams, agricultural water use, and changing climate, it is critical to understand which hydrological processes drive variability in streamflow at different timescales, and which climate, land cover, and water-use factors in turn drive

those hydrological processes. One of the necessary milestones needed to achieve this understanding has been the development of a metric space for describing streamflow regime that is both concise enough to favor meaningful insight, yet broad enough to capture the wide range of behaviors seen in streams around the world. Therefore, our primary purpose in this paper was to explore possible metric spaces for describing streamflow regime, and then to use those metric spaces to gain insight into the global distribution of streamflow regimes in order to identify patterns in and drivers of that distribution. Below, we discuss our findings in light of current ecological challenges and hydrological theory, with particular emphasis on the importance of understanding timescales as interacting units.

4.1 Are streamflow metrics or frequency decompositions better?

Streamflow metrics and frequency decompositions such as wavelet analyses facilitate different, albeit related insights into streamflow regime. Frequency decompositions offer a continuous view of variability across timescales ranging from days to decades. Similarly, streamflow metrics often describe variability in flow, but they can also capture other wave-like behaviors of streamflow timeseries such as amplitude (magnitude) and waveform (the temporal dynamics relating to rise rate and fall rate). In addition, the frequency decomposition method as used here ignored phase, or timing of flow, and also averaged spectral power across time, obfuscating any information about seasonality of variability at all temporal scales, whereas many flow metrics describe seasonal variability such as the coefficient of variation for a specific month of flow data. Thus, the dexterity of the 189 flow metrics to describe temporal, waveform, and amplitude patterns in flow may explain the roughly 55-70% difference between the frequency decompositions and original flow metrics measured here. We nonetheless suggest that the fundamental mathematics of waves remains the most intuitive method for categorizing streamflow metrics.

We also note that limitations in the ability of frequency decompositions to describe asymmetrical waveforms such as possible dramatic differences in rise-rates and fall-rates across any timescale could potentially be improved by developing a frequency decomposition which produces a two-dimensional *waveform-frequency* output, where a frequency decomposition is performed using a series of waveforms that interpolate between rapid rise rates and low fall rates, even rise rates and fall rates, and low rise rates and rapid fall rates. We also note that analyses using non-time-averaged spectral power from wavelet decompositions can more readily answer questions about the seasonality of different kinds of flow variability.

While admittedly less broad in scope as the larger suite of flow metrics, frequency decompositions benefit from a sort of meta-scope visible only by arranging streamflow phenomena in a chronosequential way. This allowed us to find the unexpected result that a mass balance relationship exists between short and long-term variability, with the fulcrum of the relationship residing somewhere near the 60-70 day time period. This inverse relationship between multi-day

and multi-month variability had artifacts in the correlations between the 189 flow metrics and the frequency domain as well as the correlations between the PCA metrics and the frequency domain. It may also help explain the inherent compressibility of the 189 flow metrics down to just a few dimensions. That is, many flow metrics may not at first glance describe the same phenomenon, but because of some previously unidentified, causal hydrological phenomenon, many metrics may be inherently (anti)correlated with each other though they describe very different aspects of streamflow regime.

4.2 Streamflow Metrics Can be Predicted with Climate, Precipitation, and Catchment Size

The correlation analysis between catchment characteristics and spectral power of multiple frequencies revealed that the positive and negative correlations between many catchment characteristics roughly paralleled the correlations between high and low frequency hydrological phenomena. Furthermore, the machine learning analysis suggested that catchment size, temperature, and precipitation were the most important predictors of flow regime. Given the large number of possible controls of flow regime, the identification of just a few fundamental drivers of flow regime, in tandem with the relatively simple structural relationship between short-term and long-term variability, suggests that a single or few fundamental mechanisms comprehensively drive flow variability at multiple timescales.

4.3 The Budyko-Darcy Hypothesis

The simplicity of the structural linkages in flow regime qualities as well as the small number of catchment characteristics needed to predict flow regime invites a simple explanation of one or a few broad-reaching underlying hydrological phenomena controlling most aspects of streamflow regime. Here we propose a marrying of two previously separate famous frameworks in hydrology: connectivity-driven Darcian flow, and climate-driven Budyko descriptions of water balance. We hypothesize that both frameworks are constantly active in mediating variability of flow through natural systems, though they tend to be most influential at opposite temporal and spatial scales. Furthermore, we hypothesize that flow of water through soils and across the land surface, as described by Darcy's law, as well as evapotranspiration, as described by Budyko's equation, are the primary forces that control variability in streamflow regime across all spatial and temporal scales (figure 9).

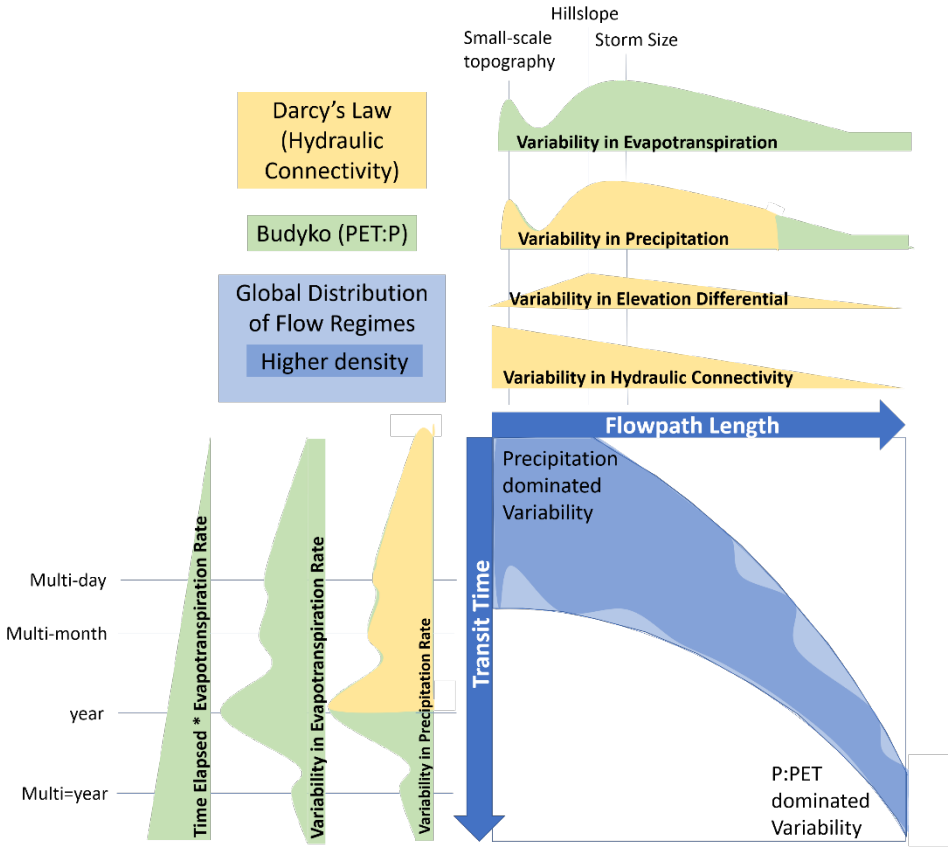

Figure 9. Conceptual diagram of the Budyko-Darcy hypothesis. Streamflow is a phenomenon that occurs across space and time which we hypothesize to be mediated by two fundamental forces: piston pressure forces pushing water through porous soils along hydrologically connected pressure gradients (Darcy's law), and evapotranspiration forces which subtract subsurface water away from precipitation-delivered water stores that might eventually enter surface flow (Budyko's Equation), reducing hydraulic connectivity in the process.

Because flow represents the movement of water across space over time, we note that time and space are inherently connected in hydrology, according to a distribution whose variance across either time or space is mediated by variance in Darcian and Budyko forces. Furthermore, we hypothesize that variance in Darcian and Budyko forces is driven by variance in the variables involved in their respective equations. For Darcy's equation, these include the pressure differential between two hydrologically connected bodies of water, the viscosity of water, and porosity of the medium that water is traveling through. In natural

systems, Darcian flow is also implicitly activated by the presence or absence of hydrological connections, which occur through soil structures and are ultimately mediated by precipitation events. For Budyko's equation, this includes potential evapotranspiration (PET) and precipitation (P).

We note that the factors mediating variability in Darcian flow vary mostly along spatial scales. For example, on the smallest scales, hydrological connections occur through small pores in the soil created by insects, tree roots, and rock fissures. As spatial scales increase, the likelihood of a high-porosity hydrological connection spanning that entire distance decreases to zero relatively quickly. This forces subsurface water to eventually flow throw less porous mediums as it travels between pressure gradients along hydrologically connected networks, causing the speed at which subsurface flow occurs to become increasingly homogeneous as spatial scales increase, ultimately resulting in a decrease in variance contributed to streamflow via Darcian flow. A similar pattern likely occurs with interruptions in hydrological connectivity via less porous paths, i.e. through different swaths of soil types, though probably at a larger scale. As spatial scales continue to increase, the next factor to reach a peak in variability is the elevation gradient within networks, likely at the spatial scale corresponding to the average horizontal length of the steepest hillslopes. Finally, differences in precipitation across space can contribute to a difference in pressure differential within a hydrological network. We hypothesize that globally, variance introduced via this phenomenon is maximized at the spatial scale corresponding to the average spatial extent of precipitation events. We further hypothesize that beyond this scale, variability in the factors contributing to Darcian flow only shrinks, thereby reducing the variability in flow contributed via Darcian flow. We note that Darcian flow is still active at larger spatial scales, and that only the variance it contributes is diminished.

While factors mediating Darcian flow tend to vary across space, the factors that mediate PET:P ratios vary most across time. We also note that precipitation events deliver water inputs at a rapid rate for a short period of time in comparison to the much slower-acting but longer-lasting process of evapotranspiration. Thus, at small spatiotemporal scales, the residence time of water in a given catchment may not be long enough for evapotranspiration losses to become significant relative to losses due to Darcian flow. However, as temporal scales increase, which inherently follow increases in spatial scale, the amount of time elapsed multiplied by the average evapotranspiration rate (i.e. net evapotranspiration) during that time becomes significant relative to the size of precipitation inputs. Thus, variability in evapotranspiration rate vs precipitation rate can have a significant impact on variability in streamflow at long timescales, and by necessity, at large spatial scales, just as variability contributed from Darcian flow collapses. Variability in PET:P ratios occurs on diurnal, to weather event, to seasonal, to annual, to multiannual timescales. But we hypothesize that the effective variance contributed does not become significant relative to Darcian flow losses until at least storm-event to seasonal timescales.

5 Conclusions

Finally, given the emergent importance of timescale we observed from our frequency decomposition and PCA analysis of traditional flow metrics, and from the surprising discovery that variance at multiple timescales is linked via catchment size, precipitation, and temperature, we propose the term *chronohydrology* to describe the comparison of hydrological phenomena such as variability, timing, and waveform, at multiple timescales. We believe a frequency-based approach to understanding flow regimes provides a greatly improved metric space upon which to consider changes in flow by allowing hydrological phenomena to be organized in a chronosequential format. While we demonstrate here that chronohydrology offers insights into fundamental governing principles of streamflow regime, we have not considered the role that chronohydrology might play in explaining trickle-down effects of streamflow regime on biogeochemical cycles, aquatic habitat, and human societal needs. Future work may benefit immensely from examining the relationships between timescales in hydrological phenomena, and we suggest that such analyses are most easily facilitated by streamflow metrics derived from the mathematics of waves.

Acknowledgments

This project was funded by the U.S. National Science Foundation (grant numbers DEB-1354867, EAR-2011439, EAR-2012123) and the Utah Division of Natural Resources Watershed Restoration Initiative. We thank the Stream Resiliency Research Coordination Network for initiating and coordinating this collaboration.

Data Availability

The data used in this study are available on researchgate.net at https://doi.org/10.13140/RG.2.2.24985.95842 and https://doi.org/10.13140/RG.2.2.31696.84487 under CC BY 4.0 licenses.

References

Abbott, B. W., Gruau, G., Zarnetske, J. P., Moatar, F., Barbe, L., Thomas, Z., et al. (2018). Unexpected spatial stability of water chemistry in headwater stream networks. *Ecology Letters*, *21*(2), 296–308. https://doi.org/10.1111/ele.12897Abbott, B. W., Bishop, K., Zarnetske, J. P., Hannah, D. M., Frei, R. J., Minaudo, C., et al. (2019). A water cycle for the Anthropocene. *Hydrological Processes*, *33*(23), 3046–3052. https://doi.org/10.1002/hyp.13544Abbott, B. W., Bishop, K., Zarnetske, J. P., Minaudo, C., Chapin, F. S., Krause, S., et al. (2019). Human domination of the global water cycle absent from depictions and perceptions. *Nature Geoscience*, *12*(7), 533–540. https://doi.org/10.1038/s41561-019-0374-yArchfield, S. A., Kennen, J. G., Carlisle, D. M., & Wolock, D. M. (2014). An objective and parsimonious approach for classifying natural flow regimes at a continental scale. *River Research and Applications*, *30*(9), 1166–1183. https://doi.org/10.1002/rra.2710Ascott, M. J., Gooddy, D. C., Fenton, O., Vero, S., Ward, R. S., Basu, N. B., et al. (2021).

The need to integrate legacy nitrogen storage dynamics and time lags into policy and practice. *Science of The Total Environment*, *781*, 146698. https://doi.org/10.1016/j.scitotenv.2021.146698Barbarossa, V., Huijbregts, M. A. J., Beusen, A. H. W., Beck, H. E., King, H., & Schipper, A. M. (2018). FLO1K, global maps of mean, maximum and minimum annual streamflow at 1 km resolution from 1960 through 2015. *Scientific Data*, *5*(1), 180052. https://doi.org/10.1038/sdata.2018.52Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. W. (2019). The Relative Importance of Different Flood-Generating Mechanisms Across Europe. *Water Resources Research*, *0*(0). https://doi.org/10.1029/2019WR024841Bernhardt, E. S., Rosi, E. J., & Gessner, M. O. (2017). Synthetic chemicals as agents of global change. *Frontiers in Ecology and the Environment*, *15*(2), 84–90. https://doi.org/10.1002/fee.1450Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, *25*(2), 197–227. https://doi.org/10.1007/s11749-016-0481-7Bochet, O., Bethencourt, L., Dufresne, A., Farasin, J., Pédrot, M., Labasque, T., et al. (2020). Iron-oxidizer hotspots formed by intermittent oxic–anoxic fluid mixing in fractured rocks. *Nature Geoscience*, 1–7. https://doi.org/10.1038/s41561-019-0509-1Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324Bunn, S. E., & Arthington, A. H. (2002). Basic Principles and Ecological Consequences of Altered Flow Regimes for Aquatic Biodiversity. *Environmental Management*, *30*(4), 492–507. https://doi.org/10.1007/s00267-002-2737-0Carey, S. K., Tetzlaff, D., Buttle, J., Laudon, H., McDonnell, J., McGuire, K., et al. (2013). Use of color maps and wavelet coherence to discern seasonal and interannual climate influences on streamflow variability in northern catchments. *Water Resources Research*, *49*(10), 6194–6207. https://doi.org/10.1002/wrcr.20469Carlisle, D. M., Falcone, J., Wolock, D. M., Meador, M. R., & Norris, R. H. (2010). Predicting the natural flow regime: models for assessing hydrological alteration in streams. *River Research and Applications*, *26*(2), 118–136. https://doi.org/10.1002/rra.1247Chezik, K. A., Anderson, S. C., & Moore, J. W. (2017). River networks dampen long-term hydrological signals of climate change. *Geophysical Research Letters*, *44*(14), 7256–7264. https://doi.org/10.1002/2017GL074376Coulibaly, P., & Burn, D. H. (2004). Wavelet analysis of variability in annual Canadian streamflows. *Water Resources Research*, *40*(3). https://doi.org/10.1029/2003WR002667Díaz, S., Settele, J., Brondízio, E. S., Ngo, H. T., Agard, J., Arneth, A., et al. (2019). Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science*, *366*(6471). https://doi.org/10.1126/science.aax3100Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, *77*(4), 802–813. https://doi.org/10.1111/j.1365-2656.2008.01390.xFisher, S. G., Grimm, N. B., Martí, E., Holmes, R. M., & Jones Jr, J. B. (1998). Material spiraling in stream corridors: a telescoping ecosystem model. *Ecosystems*, *1*(1), 19–34.Frei, R. J., Abbott, B. W., Dupas, R., Gu, S., Gruau, G., Thomas, Z., et al. (2020). Predicting Nutrient Incontinence in the Anthropocene at Watershed Scales. *Frontiers in Envi-*

ronmental Science, *7*. https://doi.org/10.3389/fenvs.2019.00200George, R., McManamay, R., Perry, D., Sabo, J., & Ruddell, B. L. (2021). Indicators of hydro-ecological alteration for the rivers of the United States. *Ecological Indicators*, *120*, 106908. https://doi.org/10.1016/j.ecolind.2020.106908Gerten, D., Rost, S., von Bloh, W., & Lucht, W. (2008). Causes of change in 20th century global river discharge. *Geophysical Research Letters*, *35*(20), L20405. https://doi.org/10.1029/2008GL035258Gnann, S. J., Coxon, G., Woods, R. A., Howden, N. J. K., & McMillan, H. K. (2021). TOSSH: A Toolbox for Streamflow Signatures in Hydrology. *Environmental Modelling and Software*, *138*, 104983. https://doi.org/10.1016/j.envsoft.2021.104983Godsey, S. E., & Kirchner, J. W. (2014). Dynamic, discontinuous stream networks: hydrologically driven variations in active drainage density, flowing channels and stream order. *Hydrological Processes*, *28*(23), 5791–5803. https://doi.org/10.1002/hyp.10310Gorski, G., & Zimmer, M. A. (2021). Hydrologic regimes drive nitrate export behavior in human-impacted watersheds. *Hydrology and Earth System Sciences*, *25*(3), 1333–1345. https://doi.org/10.5194/hess-25-1333-2021Hain, E. F., Kennen, J. G., Caldwell, P. V., Nelson, S. A. C., Sun, G., & McNulty, S. G. (2018). Using regional scale flow–ecology modeling to identify catchments where fish assemblages are most vulnerable to changes in water availability. *Freshwater Biology*, *63*(8), 928–945. https://doi.org/10.1111/fwb.13048Hannah, D. M., Demuth, S., Lanen, H. A. J. van, Looser, U., Prudhomme, C., Rees, G., et al. (2011). Large-scale river flow archives: importance, current status and future needs. *Hydrological Processes*, *25*(7), 1191–1200. https://doi.org/10.1002/hyp.7794Harrison, I., Abell, R., Darwall, W., Thieme, M. L., Tickner, D., & Timboe, I. (2018). The freshwater biodiversity crisis. *Science*, *362*(6421), 1369–1369. https://doi.org/10.1126/science.aav9242Helton, A. M., Poole, G. C., Meyer, J. L., Wollheim, W. M., Peterson, B. J., Mulholland, P. J., et al. (2011). Thinking outside the channel: modeling nitrogen cycling in networked river ecosystems. *Frontiers in Ecology and the Environment*, *9*(4), 229–238. https://doi.org/10.1890/080211Jones, N. E., Schmidt, B. J., & Melles, S. J. (2014). Characteristics and distribution of natural flow regimes in Canada: a habitat template approach. *Canadian Journal of Fisheries and Aquatic Sciences*, *71*(11), 1616–1624. https://doi.org/10.1139/cjfas-2014-0040Labat, D. (2005). Recent advances in wavelet analyses: Part 1. A review of concepts. *Journal of Hydrology*, *314*(1), 275–288. https://doi.org/10.1016/j.jhydrol.2005.04.003Labat, D. (2008). Wavelet analysis of the annual discharge records of the world's largest rivers. *Advances in Water Resources*, *31*(1), 109–117. https://doi.org/10.1016/j.advwatres.2007.07.004Labat, D. (2010). Cross wavelet analyses of annual continental freshwater discharge and selected climate indices. *Journal of Hydrology*, *385*(1), 269–278. https://doi.org/10.1016/j.jhydrol.2010.02.029Lafrenière, M., & Sharp, M. (2003). Wavelet analysis of inter-annual variability in the runoff regimes of glacial and nival stream catchments, Bow Lake, Alberta. *Hydrological Processes*, *17*(6), 1093–1118. https://doi.org/10.1002/hyp.1187Lane, B. A., Dahlke, H. E., Pasternack, G. B., & Sandoval-Solis, S. (2017). Revealing the Diversity of Natural Hydrologic Regimes in California with Relevance

for Environmental Flows Applications. *JAWRA Journal of the American Water Resources Association*, *53*(2), 411–430. https://doi.org/10.1111/1752-1688.12504Larsen, I. J., MacDonald, L. H., Brown, E., Rough, D., Welsh, M. J., Pietraszek, J. H., et al. (2009). Causes of Post-Fire Runoff and Erosion: Water Repellency, Cover, or Soil Sealing? *Soil Science Society of America Journal*, *73*(4), 1393–1407. https://doi.org/10.2136/sssaj2007.0432Masaki, Y., Hanasaki, N., Biemans, H., Schmied, H. M., Tang, Q., Yoshihide Wada, et al. (2017). Intercomparison of global river discharge simulations focusing on dam operation—multiple models analysis in two case-study river basins, Missouri–Mississippi and Green–Colorado. *Environmental Research Letters*, *12*(5), 055002. https://doi.org/10.1088/1748-9326/aa57a8McMillan, H. (2021). A review of hydrologic signatures and their applications. *WIREs Water*. https://doi.org/10.1002/wat2.1499Minaudo, C., Dupas, R., Gascuel-Odoux, C., Fovet, O., Mellander, P.-E., Jordan, P., et al. (2017). Nonlinear empirical modeling to estimate phosphorus exports using continuous records of turbidity and discharge: P exports from turbidity and discharge. *Water Resources Research*, *53*(9), 7590–7606. https://doi.org/10.1002/2017WR020590Moatar, F., Abbott, B. W., Minaudo, C., Curie, F., & Pinay, G. (2017). Elemental properties, hydrology, and biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major ions. *Water Resources Research*, *53*(2), 1270–1287. https://doi.org/10.1002/2016WR019635Myles, A. J., Feudale, R. N., Liu, Y., Woody, N. A., & Brown, S. D. (2004). An introduction to decision tree modeling. *Journal of Chemometrics*, *18*(6), 275–285. https://doi.org/10.1002/cem.873Olden, J. D., & Poff, N. L. (2003). Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River Research and Applications*, *19*(2), 101–121. https://doi.org/10.1002/rra.700Oldfield, J. D. (2016). Mikhail Budyko's (1920–2001) contributions to Global Climate Science: from heat balances to climate change and global ecology. *Wiley Interdisciplinary Reviews: Climate Change*, *7*(5), 682–692. https://doi.org/10.1002/wcc.412Pascolini-Campbell, M., Reager, J. T., Chandanpurkar, H. A., & Rodell, M. (2021). A 10 per cent increase in global land evapotranspiration from 2003 to 2019. *Nature*, *593*(7860), 543–547. https://doi.org/10.1038/s41586-021-03503-5Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 6.Pinay, G., Bernal, S., Abbott, B. W., Lupon, A., Marti, E., Sabater, F., & Krause, S. (2018). Riparian Corridors: A New Conceptual Framework for Assessing Nitrogen Buffering Across Biomes. *Frontiers in Environmental Science*, *6*. https://doi.org/10.3389/fenvs.2018.00047Poff, N. Leroy, & Zimmerman, J. K. H. (2010). Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. *Freshwater Biology*, *55*(1), 194–205. https://doi.org/10.1111/j.1365-2427.2009.02272.xPoff, N. LeRoy, Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., et al. (1997). The Natural Flow Regime. *BioScience*, *47*(11), 769–784. https://doi.org/10.2307/1313099Raymond, P. A., Saiers, J. E., & Sobczak, W. V. (2016). Hydrological and biogeochemical controls on watershed dis-

solved organic matter transport: pulse-shunt concept. *Ecology*, *97*(1), 5–16. https://doi.org/10.1890/14-1684.1Reaver, N. G. F., Kaplan, D. A., Klammler, H., & Jawitz, J. W. (2020). Reinterpreting the Budyko Framework. *Hydrology and Earth System Sciences Discussions*, 1–31. https://doi.org/10.5194/hess-2020-584Rossi, A., Massei, N., Laignel, B., Sebag, D., & Copard, Y. (2009). The response of the Mississippi River to climate fluctuations and reservoir construction as indicated by wavelet analysis of streamflow and suspended-sediment load, 1950–1975. *Journal of Hydrology*, *377*(3), 237–244. https://doi.org/10.1016/j.jhydrol.2009.08.032Ryo, M., Iwasaki, Y., Yoshimura, C., & V, O. C. S. (2015). Evaluation of Spatial Pattern of Altered Flow Regimes on a River Network Using a Distributed Hydrological Model. *PLOS ONE*, *10*(7), e0133833. https://doi.org/10.1371/journal.pone.0133833Sabo, J. L., & Post, D. M. (2008). Quantifying periodic, stochastic, and catastrophic environmental variation. *Ecological Monographs*, *78*(1), 19–40. https://doi.org/10.1890/06-1340.1Sanborn, S. C., & Bledsoe, B. P. (2006). Predicting streamflow regime metrics for ungauged streamsin Colorado, Washington, and Oregon. *Journal of Hydrology*, *325*(1), 241–261. https://doi.org/10.1016/j.jhydrol.2005.10.018Sang, Y.-F. (2013). A review on the applications of wavelet transform in hydrology time series analysis. *Atmospheric Research*, *122*(Supplement C), 8–15. https://doi.org/10.1016/j.atmosres.2012.11.003Sang, Y.-F., Wang, D., Wu, J.-C., Zhu, Q.-P., & Wang, L. (2009). The relation between periods' identification and noises in hydrologic series data. *Journal of Hydrology*, *368*(1), 165–177. https://doi.org/10.1016/j.jhydrol.2009.01.042Smith, L. C., Turcotte, D. L., & Isacks, B. L. (1998). Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrological Processes*, *12*(2), 233–249. https://doi.org/10.1002/(SICI)1099-1085(199802)12:2<233::AID-HYP573>3.0.CO;2-3Teixeira, H., Lillebø, A. I., Culhane, F., Robinson, L., Trauner, D., Borgwardt, F., et al. (2019). Linking biodiversity to ecosystem services supply: Patterns across aquatic ecosystems. *Science of The Total Environment*, *657*, 517–534. https://doi.org/10.1016/j.scitotenv.2018.11.440Tucker, G. E., & Hancock, G. R. (2010). Modelling landscape evolution. *Earth Surface Processes and Landforms*, *35*(1), 28–50. https://doi.org/10.1002/esp.1952Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I. J. M., Stahl, K., Hannaford, J., et al. (2016). Drought in the Anthropocene. *Nature Geoscience*, *9*(2), 89–91. https://doi.org/10.1038/ngeo2646Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., et al. (2010). Global threats to human water security and river biodiversity. *Nature*, *467*(7315), 555–561. https://doi.org/10.1038/nature09440White, M. A., Schmidt, J. C., & Topping, D. J. (2005). Application of wavelet analysis for monitoring the hydrologic effects of dam operation: Glen Canyon Dam and the Colorado River at Lees Ferry, Arizona. *River Research and Applications*, *21*(5), 551–565. https://doi.org/10.1002/rra.827Zammit-Mangion, A., & Cressie, N. (2017). FRK: An R Package for Spatial and Spatio-Temporal

Prediction with Large Datasets.Zarnetske, J. P., Bouda, M., Abbott, B. W., Saiers, J., & Raymond, P. A. (2018). Generality of Hydrologic Transport Limitation of Watershed Organic Carbon Flux Across Ecoregions of the United States. *Geophysical Research Letters*, *45*(21), 11,702-11,711. https://doi.org/10.1029/2018GL080005Zhou, G., Wei, X., Chen, X., Zhou, P., Liu, X., Xiao, Y., et al. (2015). Global pattern for the effect of climate and land cover on water yield. *Nature Communications*, *6*(1), 5918. https://doi.org/10.1038/ncomms6918