Metabolic rhythms in flowing waters: An approach for classifying river productivity regimes

Philip Savoy 1,*, Alison P. Appling 2, James B. Heffernan 3, Edward G. Stets 4, Jordan S. Read 5, Judson W. Harvey 6, Emily S. Bernhardt 1

1Department of Biology, Duke University, Durham, North Carolina
2U.S. Geological Survey, Integrated Information Dissemination Division, University Park, Pennsylvania
3Nicholas School of the Environment, Duke University, Durham, North Carolina
4U.S. Geological Survey, Earth System Processes Division, Boulder, Colorado
5U.S. Geological Survey, Integrated Information Dissemination Division, Middleton, Wisconsin
6U.S. Geological Survey, Earth System Processes Division, Reston, Virginia

Abstract

Although seasonal patterns of ecosystem productivity have been extensively described and analyzed with respect to their primary forcings in terrestrial and marine systems, comparatively little is known about these same processes in rivers. However, it is now possible to perform a large-scale synthesis on the patterns and drivers of river productivity regimes because of the recent sensor advances allowing for near-continuous estimates of river productivity. Here, we explore a dataset of 47 U.S. rivers to examine whether there are characteristic river productivity regimes. We use classification approaches to develop a typology of productivity regimes and then use these regimes to examine differences with respect to potential controls of productivity. We identified two distinct metabolic regimes, which we named Summer Peak and Spring Peak Rivers, within our dataset. These regimes meaningfully differed in both the timing and magnitude of productivity and were robust to different approaches to classification. We also found that several variables, including watershed area and characteristics of water temperature or discharge, were able to predict the class membership of these regimes with modest accuracy. Our results support the presence of characteristic metabolic regimes and suggests that these regimes may have common sets of environmental controls. We present classification as one approach to begin exploring the productivity regimes of rivers. The strength of our approach is that it fully leverages these newly available high-frequency productivity estimates to create classes that can be used to draw inferences about how the controls of river productivity differ between or within systems.

The seasonality of photosynthesis is fundamental to the seasonal oscillations of many ecosystem processes, including the cycling of carbon (C), water, and nutrients. Autotrophs respond to seasonal variations in climatic conditions; therefore, the seasonality of photosynthesis and its associated ecosystem processes can both be considered inherently phenological (Gu et al. 2003; Noormets 2009). Phenological responses to changes in broad-scale climate have been observed across a wide variety of taxa and locations (Parmesan 2006) including climate-change-induced shifts in the timing of the growing season in forests (Menzel and Fabian 1999; Menzel et al. 2006; Schwartz et al. 2006; Keenan et al. 2014), altered phytoplankton phenology from large-scale climatic oscillations in the oceans (Sasaoka et al. 2011; Racault et al. 2012), and advancement in the timing of spring plankton blooms in lakes (Parmesan and Yohe 2003; Peeters et al. 2007). At longer timescales, changing climatic conditions have altered patterns of primary production (Boisvenue and Running 2006; Chavez et al. 2011; Keenan et al. 2014). Given the likelihood of novel climate conditions in the future (Williams et al. 2007; Williams and Jackson 2007), understanding the relationship between environmental conditions and ecosystem phenology is critical for predicting shifts in C cycling associated with climate.

Across terrestrial and marine systems, productivity is fundamentally limited by a combination of light, nutrients, temperature, or water (Field et al. 1998). In rivers, the phenology of terrestrial vegetation and the annual patterns of flow may be

*Correspondence: prs15@duke.edu

Correction added on April 2nd, after first online publication: the first digit was accidentally duplicated for each entry in the GPP column of Table 1 during production.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

1835
more important proximal controls on productivity than are climate drivers. For example, forest canopies shade many river channels. In such cases, the phenology of adjacent terrestrial ecosystems controls the light regimes of rivers and the seasonality of gross primary production (GPP) (Hill and Dimick 2002; Roberts et al. 2007). The flow regime also exerts strong control on river GPP through the removal of autotrophic biomass during high flow events (Biggs and Thomsen 1995; Uehlinger 2000; Uehlinger et al. 2003), attenuation of light in the water column from suspended sediments (O’Connor et al. 2012; Hall et al. 2015), or by periods of intermittent flow (Acuña et al. 2005). Thermal regimes of rivers may also become decoupled from patterns of ambient air temperature because of heat exchange with groundwater and riparian shading (Poole and Berman 2001). Predicted increases in precipitation variability could lead to increased disturbance from floods or droughts (Milly et al. 2005), and climate-change-induced alterations in the phenology of adjacent riparian areas may shift the seasonality of the light or thermal regimes of rivers. These changes are likely to affect the magnitude, timing, and variability of productivity. Characterizing productivity regimes will be essential to document and understand stream ecosystem response to altered environmental drivers (Bernhardt et al. 2017).

In terrestrial, marine, and lentic systems, the ability to resolve estimates of GPP from remotely sensed and eddy covariance data has produced a rich literature base that examines the spatiotemporal patterns and drivers of GPP from regional to global scales. In rivers, despite considerable attention to thermal (Olden and Naiman 2010; Demars et al. 2011) and hydrologic regimes (Poff et al. 1997), little is known about how temporal patterns of GPP emerge in response to environmental controls (Bernhardt et al. 2017). This is partly because the lack of affordable and durable sensors has stymied efforts to capture the variation of GPP over the full course of the year in all but a handful of sites (Bernhardt et al. 2017). A recent effort to estimate continuous metabolism for a large number of U.S. Geological Survey (USGS) gauging stations, where dissolved oxygen (DO) is collected with in situ sensors, has led to a rapid expansion in the number of rivers with long-term high-frequency productivity estimates (Appling et al. 2018c). This effort has now generated and compiled continuous metabolism records for 356 sampling locations in rivers throughout the continental United States. In total, this compilation includes 490,907 individual days of continuous DO data that have been converted into estimates of metabolism for 2190 site years. The availability of a large dataset of near-continuous estimates of stream GPP provides us with the opportunity to perform the first large-scale synthesis of the seasonality of river productivity.

Several existing conceptual frameworks of stream ecosystems could inform such a synthesis. Most notable is the representation of rivers as a continuous gradient of physical drivers and responses (Vannote et al. 1980). Others have suggested that these continuous gradients are punctuated by discontinuities (Ward and Stanford 1983) or that geomorphic control of some processes are better explained when described as distinct units (Montgomery 1999). A discontinuum view of lotic systems does not preclude the possibility that a continuum exists and these ideas can be reconciled, for example, by considering rivers as hierarchical systems that may or may not arrange themselves approximately as a continuum (Poole 2002). Seasonal patterns of GPP reflect a highly dimensional and multiscale set of physical, chemical, and biological controls that can make it difficult to draw comparisons across sites or determine the relative importance of factors driving these patterns. Assessing rivers from a hierarchical classification framework can aid our understanding by describing how classes differ by focusing on only the most relevant sets of variables at each level of the hierarchy (Frissell et al. 1986). Classification has a long history and widespread use in ecology (Whittaker 1962) and can be a useful organizational tool to facilitate comparisons between classes or to understand the importance of processes within classes. Classifications of temporal regimes of river flow (Poff 1996; Olden et al. 2012; Archfield et al. 2014), water temperature (Maheu et al. 2015) or terrestrial phenology (Lloyd 1990; White et al. 2005) have been used for the purpose of monitoring, hypothesis testing, and identifying common drivers of observed patterns.

In this study, we used classification approaches to develop a typology of productivity regimes and we used the resulting typology to explore the following questions: (1) Are there characteristic regimes of river GPP? And (2) do these regimes reflect differences in site characteristics or environmental drivers that may mediate productivity?

Data

Daily metabolism data

Daily estimates of metabolism were produced by the USGS Powell Center working group on continental patterns of stream metabolism (Appling et al. 2018c) and are freely available to download (Appling et al. 2018d). Modeled estimates of GPP were generated using the streamMetabolizer R package (Appling et al. 2018b), which uses inverse modeling to produce daily estimates of GPP, ecosystem respiration (ER), and gas exchange rate coefficients \( K_{600} \) based on subdaily time series of DO, water temperature, discharge, stream depth, and photosynthetically active radiation (PAR) (Appling et al. 2018a). The specific model variant used in Appling et al. (2018c) was based on a standard mass-balance representation of DO as a function of light-dependent inputs from GPP, constant ER, and temperature- and concentration-dependent gas exchange between the water and air. This model variant included Bayesian partial pooling of \( K_{600} \) where estimates of \( K_{600} \) on individual days were pooled toward a multiday piecewise-linear relationship between \( K_{600} \) and daily mean discharge, which has been shown to reduce unrealistic model estimates (Appling et al. 2018a).

This nation-wide application of streamMetabolizer used stream monitoring data (DO, water temperature, and discharge)
from the USGS’s National Water Information System (NWIS) (USGS 2017) and environmental data (barometric pressure and shortwave radiation) from the North American Land Data Assimilation System (NLDAS) (Mitchell et al. 2004; Xia et al. 2012) as described in Appling et al. (2018c). Stream depth was estimated from discharge using regionalized hydraulic geometry coefficients from the NEXSS database (Gomez-Velez et al. 2015). Linear interpolation was used to fill gaps of 3 h or less for most variables, and a theory-based light model was used to interpolate PAR. The data were combined and passed to the selected model variant (Appling et al. 2018c). It is possible for this model variant to produce negative GPP estimates, which are not biologically possible, so all negative GPP values were set to not available (NA) before use in our analysis of productivity regimes.

For this analysis, we selected a subset of these 356 sites based on a combination of data quality, coverage within years, and length of the record. First, all sites that had a dam or canal within the upstream footprint of the oxygen probe, estimated as the 95% turnover distance of DO in the water column based on water velocity and the gas exchange coefficient (e.g., Chapra and Di Toro 1991), were excluded because of the potential for major discontinuities in slope or flow to violate the assumptions of metabolism models. Second, individual site-years of data were screened to only include years that had modeled daily GPP for at least 50% of the year and for which no single data gap was greater than 90 d during the calendar year. We chose these constraints to retain as much data as possible while still considering that a minimum level of data support was required for the subsequent analysis. Finally, only sites that had data for the entire period of 2013–2016 were selected. This selection was done to ensure that all sites had equal representation in the dataset (n = 4 yr) and that all datasets were compared within the same period of broad-scale climatology. The specific dates were chosen by examining the filtered dataset to provide the best trade-off between the number of sites and total number of site years included in the analysis. The final filtered dataset consisted of 47 sites and 188 site years of data (Table 1).

### Site characteristics and environmental drivers

Site characteristics and time series of relevant environmental data over the study period were compiled for use in post hoc comparisons of how regimes differed with respect to their drivers. Because annual mean rates of river GPP have been shown to be positively related to watershed area (Lamberti and Steinman 1997; Finlay 2011), we included watershed area information compiled from several sources (Appling et al. 2018c). Watershed area encapsulates other covariates related to GPP such as changes in discharge or channel width. Variation in light availability has also been shown to be related to watershed area, and light is an important control on river GPP both across (Mulholland et al. 2001) and within sites (Roberts et al. 2007). Although data on light reaching the water surface was not available at our sites, we included several surrogates of light availability. Estimates of river width from regionalized hydraulic geometry coefficients (Gomez-Velez et al. 2015) were used to more directly represent changing channel characteristics on available light. To capture variation in the structure of riparian vegetation, we used canopy cover in the 100 m riparian buffer for the accumulated watershed derived from the National Land Cover Database 2011 (Homer et al. 2015) and tree heights from the global lidar-derived estimates of Simard et al. (2011).

In addition to site-level characteristics, we also included time series of several environmental drivers. These time series were used to summarize the average environmental conditions at sites as well to compare the seasonality of GPP and environmental drivers. Subdaily discharge (m$^{-3}$ s$^{-1}$) and water temperature (°C) were acquired from a set of data compiled by the Powell Center working group on stream metabolism from publicly available NWIS data (USGS 2017). Because data on measured light at the river surface were unavailable, hourly surface downwelling shortwave radiation (W m$^{-2}$) from the North NLDAS (Xia et al. 2012) were downloaded via the National Aeronautics and Space Administration Goddard Earth Sciences Data and Information Center Data Rods Explorer (Teng et al. 2016). This was converted to PAR (μmol m$^{-2}$ s$^{-1}$) using the conversion from Britton and Dodd (1976). Although these data may not reflect the light regimes of PAR actually received by the river, it does include information about distal controls of GPP in the form of geographic location or cloudines.

### Summarizing time series of productivity and environmental drivers

Several measures were derived from the raw data to characterize different scales of temporal variability. Archfield et al. (2014) described a suite of seven ecologically relevant streamflow statistics (ERSS) that capture the five components of the natural flow regime as defined by Poff et al. (1997). The first four L-moments of the distribution of daily data (mean, coefficient of variation [L-C.V.], skewness [L-skew], and kurtosis [L-kurt]) are used to summarize daily data. The autoregressive lag-one (AR(1)) correlation coefficient captures temporal autocorrelation of the preceding day and is used as a proxy for quantifying the rate of change of hydrologic conditions. Finally, amplitude and phase describe the variation in the timing and magnitude of the time series (Archfield et al. 2014). We calculated these ERS from mean daily discharge data for the entire study period to describe variation in the flow regime that might influence patterns of productivity. Additionally, we adapted these metrics to describe the temporal characteristics of GPP, water temperature, and incoming PAR. These metrics were calculated as per Archfield et al. (2014). Annual GPP (g O$_2$ m$^{-2}$ yr$^{-1}$) was also calculated for each site.

### Methods

#### Approach to analysis

Uncovering patterns in time series datasets without established groups is commonly done through clustering (Liao
Table 1. Site locations, watershed area, and average annual GPP for sites used in this analysis.

| USGS site                  | Lat  | Lon   | Watershed area (km²) | GPP (g O₂ m⁻² yr⁻¹) |
|----------------------------|------|-------|----------------------|----------------------|
| 01480617 West Branch Brandywine Creek at Modena | 39.96 | −75.80 | 143                  | 1216                 |
| 01480870 East Branch Brandywine Creek near Cowningtown | 39.97 | −75.67 | 233                  | 1287                 |
| 01493112 Chesterville Branch near Crumpton | 39.26 | −75.94 | 18                   | 533                  |
| 01548303 Straight Run (site 1) at Marsh Creek | 41.78 | −77.40 | 19                   | 150                  |
| 01608500 South Branch Potomac River near Springfield | 39.45 | −78.65 | 19                   | 1347                 |
| 01611500 Cacapon River near Great Cacapon | 39.58 | −78.31 | 1749                 | 1268                 |
| 01632900 Smith Creek near New Market | 38.69 | −78.64 | 246                  | 643                  |
| 01645704 Difficult Run above Fox lake near Fairfax | 38.88 | −77.33 | 15                   | 164                  |
| 01646305 Dead Run at Whann ave. near Mclean | 38.96 | −77.18 | 7                    | 253                  |
| 01648010 Rock Creek at Joyce rd | 38.96 | −77.04 | 171                  | 370                  |
| 01649190 Paint Branch near College Park | 39.03 | −76.96 | 34                   | 260                  |
| 01650800 Sligo Creek near Takoma park | 38.99 | −77.00 | 19                   | 306                  |
| 01656903 Flatlick Branch above Frog Branch at Chantilly | 38.88 | −77.43 | 10                   | 310                  |
| 01657300 Pamunkey River near Hanover | 37.77 | −77.33 | 2794                 | 163                  |
| 02101726 Rocky river at US 64 near Siler City | 35.74 | −79.42 | 180                  | 293                  |
| 02110500 Waccamaw River near Longs | 33.91 | −81.55 | 1969                 | 461                  |
| 02110700 Enoree River at Whitmire | 34.51 | −81.60 | 1153                 | 324                  |
| 02203655 South River at Forrest Park rd. | 33.68 | −84.36 | 60                   | 129                  |
| 02217643 Parks Creek at Lyle Field rd. near Jefferson | 34.16 | −83.53 | 8                    | 84                   |
| 02217644 Whittenhouse Creek near Vineland | 28.39 | −81.62 | 36                   | 123                  |
| 02226300 Reedy Creek near Vineland | 28.33 | −81.58 | 202                  | 138                  |
| 02336152 South Fork Peachtree at Casa dr. near Clarkston | 33.81 | −84.25 | 23                   | 134                  |
| 023362095 Burnt Fork Creek at Millwood Way near Clarkston | 33.82 | −84.27 | 13                   | 87                   |
| 02336300 Peachtree Creek at Atlanta | 33.82 | −84.41 | 225                  | 476                  |
| 02336313 Woodall Creek at Defoors Ferry rd. | 33.82 | −84.44 | 7                    | 214                  |
| 02336410 Nancy Creek at West Wesley rd. | 33.84 | −84.44 | 97                   | 318                  |
| 02336526 Proctor Creek at Jackson Parkway | 33.79 | −84.47 | 41                   | 534                  |
| 03067510 Shavers Fork near Cheat Bridge | 34.62 | −83.87 | 160                  | 213                  |
| 04121944 Little Muskegon River near Oak Grove | 43.43 | −85.60 | 966                  | 1255                 |
| 04121970 Muskegon River near Croton | 43.43 | −85.67 | 6073                 | 683                  |
| 04125460 Pine River at High School Bridge near Hoxeyville | 44.19 | −85.77 | 652                  | 675                  |
| 04136900 Au Sable River near McKinley | 44.61 | −83.84 | 3725                 | 2236                 |
| 04137005 Au Sable River near Curtisville | 44.56 | −83.80 | 3950                 | 299                  |
| 04166500 River Rouge at Detroit | 42.37 | −83.25 | 479                  | 469                  |
| 04176500 River Raisin near Monroe | 41.96 | −83.53 | 2680                 | 2573                 |
| 04213500 Cattaraugus Creek at Gowanda | 42.46 | −78.93 | 1129                 | 1024                 |
| 05435943 Badger Mill Creek at Verona | 42.98 | −89.54 | 55                   | 1248                 |
| 07072520 South Fork of Little Red River (Gulf Mt. Wilderness) | 35.55 | −92.71 | 121                  | 94                   |
| 07072570 South Fork of Little Red River near Scotland | 35.57 | −92.62 | 194                  | 72                   |
| 07143672 Little Arkansas River at HWY 50 near Halstead | 38.03 | −97.54 | 1997                 | 616                  |
| 07177500 Bird Creek near Sperry | 36.28 | −95.95 | 2362                 | 313                  |
| 08052745 Doe Branch at US HWY 380 near Prosper | 33.22 | −96.89 | 112                  | 264                  |
| 10133980 East Canyon Creek at East Canyon Resort | 40.87 | −111.59 | 286              | 1575                 |
| 11044000 Santa Margarita River near Temecula | 33.47 | −117.14 | 1495              | 1160                 |
| 11273400 San Joaquin River at Merced River near Newman | 37.35 | −120.98 | 17,551             | 3750                 |
| 14206950 Fanno Creek at Durham | 45.40 | −122.75 | 82              | 161                  |
2005; Aghabozorgi et al. 2015), but there is no single best set of methods for analysis. Clustering aims to identify groups based on a measure of dissimilarity, and there are two key decision points that must be considered: (1) how to quantify the similarity between time series, and (2) how to use this similarity to construct a classification (if one exists). The selection of a similarity measure is critical for clustering analysis and the definition of similarity measures for time series is non-trivial due to their high dimensionality (Fu 2011). Measures of similarity generally fall into one of several categories: (1) raw-data-based approaches that utilize the full time series, (2) feature-based approaches that reduce data dimensionality prior to computing dissimilarity measures, and (3) model-based approaches that define dissimilarity based on differences in model parameters (Aghabozorgi et al. 2015). There are also several common approaches used for time series clustering, which can generally be classified as partitional, hierarchical, or model-based (Liao 2005). Generally, the objective of partitional clustering is to create a set of nonoverlapping groups by minimizing the difference between cluster members, whereas hierarchical clustering groups data into a nested tree (Aghabozorgi et al. 2015).

To address our broader research questions, we performed two separate sets of analyses using different clustering approaches. The first analysis used the full time series of GPP to classify productivity regimes with hierarchical clustering. We then examined whether these clusters differed with respect to site characteristics or environmental drivers through a series of post hoc comparisons. The second analysis classified sites based on differences in their site or environmental conditions and examined how these compared to classifications based on productivity. To allow inclusion of both site characteristics and time series of environmental drivers, we used a feature-based partitional clustering approach to perform this analysis. The specific details of each analysis are presented in the following sections.

Analysis I—identifying productivity regimes with data-based hierarchical clustering

Our first analysis created a typology by leveraging the full data record to directly compare time series of productivity. We used dynamic time warping (DTW) (Sakoe and Chiba 1978; Berndt and Clifford 1994) to compare the shapes of time series. The primary advantage of DTW is that it is elastic with respect to time, meaning that it is less sensitive to small temporal shifts, noisy data, or outliers than one-to-one measures of similarity like Euclidean distance. Comparing a pair of time series with DTW involves constructing a warping matrix that compares a reference and query time series where points along the diagonal represent a one-to-one match of the time series and deviations from this represent warping (Fig. 1). DTW has been used to quantify similarity in remotely sensed vegetation phenology across regions (Xue et al. 2014; Zhang and Hepner 2017), providing both precedent and an opportunity to compare aquatic and terrestrial ecosystem phenology using similar approaches.

Although the strength of DTW lies in its elasticity, it is necessary to constrain the time period over which the algorithm searches, both for efficiency and accuracy (Sakoe and Chiba 1978; Berndt and Clifford 1994). We used the Sakoe-Chiba band (Sakoe and Chiba 1978) to constrain the warping path so only points up to 2 wk apart could be considered. This window width was chosen by first calculating the standard deviation (SD) of the phase of GPP at each site during the study.

---

**Fig. 1.** An example of comparing two time series (a) with DTW. A cumulative distance matrix (M) is constructed (b) and the resulting optimal alignment is found between the two series (c), resulting in the final alignment (d).
period to represent interannual variability in the timing of seasonality. The average SD of phase was 16 d, which is roughly equivalent to the 2 wk window that has previously been used when assessing phenological variation in remotely sensed data for terrestrial phenology (Zhang and Hepner 2017). The DTW distance was calculated, using the symmetric2 step pattern, in the dtw package (Giorgino 2009) in R version 3.3.3 (R Core Team 2017).

Several preprocessing steps were required prior to performing DTW. At each site, gaps in the GPP time series were filled by first fitting a generalized additive model (GAM) with both seasonal trend components using the mgcv package (Wood 2006) in R and then interpolating missing values from the GAM model fit. Gap-filling generated time series of equal lengths, which could be assessed for both temporal variation and calculating annual sums of GPP. Although DTW is often cited for its ability to compare series of different lengths, Ratnamahatana and Keogh (2005) demonstrated that there was no significant difference between using series of unequal length or reinterpolating series to have equal length. As our comparisons focused on differences between regimes and sites rather than interannual variability, we calculated a representative time series of GPP at each site by taking the mean GPP for each day of the year across all 4 yr of data. This time series was then z-normalized prior to computing the DTW distance. Processing raw time series by z-normalization is a critical step for accurate classification (Mueen and Keogh 2016), because calculating distances on un-normalized data can exaggerate differences in the dissimilarity between time series (Keogh and Kasety 2003).

Hierarchical agglomerative clustering, which is widely used for hydrologic classification (Olden et al. 2012) or for identifying stream thermal regimes (Maheu et al. 2015), was used to identify emergent productivity regimes within our dataset based on a DTW distance matrix. Hierarchical agglomerative clustering begins by treating each observation as a cluster and successively merges these observations into larger clusters until all observations belong to a single cluster. The method for successively merging clusters together and the final number of clusters must be determined by users based on the objectives of clustering. We used the complete linkage method, which determines when two clusters merge during agglomeration based on the most distant members of each cluster (Sørensen 1948), to merge clusters. Complete linkage is useful in ecological applications because it reduces chaining and generates distinct clusters (Legendre and Legendre 2012). Clustering was done using the hclust function in R version 3.3.3 (R Core Team 2017).

Because no a priori optimal number of clusters exists, clustering was used to define between 2 and 10 clusters, and two separate methods were used to select the final number of clusters. When no existing reference groups are known, clusters can be validated through internal cluster validity indices (CVIs). There are a variety of CVIs, most of which assess a combination of within cluster cohesion or between cluster separation, but there is no universally best CVI (Arbelaitz et al. 2013). Each CVI determines an optimal number of clusters based on the assumptions of the index. A suite of six CVIs were calculated using the dtwclust package (Sarda-Espinosa 2017), and then a final number of clusters was determined based on a majority rules approach based on the optimal number of clusters suggested by each index.

Next, at each level of clustering, the stability of clusters was assessed by a bootstrapping analysis where sites were resampled with replacement and then a hierarchical agglomerative clustering was performed on the resampled data. The similarity between clusters was quantified by the Jaccard index ($J$), which is the ratio of the number of elements in the union and the number of elements in the intersection within two clusters. The Jaccard index ranges from 0 to 1, indicating that two clusters share no or all members, respectively. For each bootstrapped cluster, the most similar original cluster was identified and the Jaccard index value was then recorded. The processes of resampling with replacement, performing clustering, and assessing the similarity of the bootstrapped clusters to the original clusters were repeated 100 times at each level of clustering. Clusters with a mean Jaccard coefficient of less than 0.5 were considered to be unstable and may not reflect a true pattern in the data.

We explored potential differences among the productivity regimes derived from this analysis through several post hoc analyses. Pearson’s correlation coefficient ($r$) was used to assess the correlation between site or environmental drivers and annual rates of GPP ($g O_2 m^{-2} yr^{-1}$). Correlations were performed both across the entire dataset and within clusters to explore the relative importance of drivers within regimes. The seasonal dynamics of GPP with respect to environmental drivers were assessed by calculating the correlation of daily values of GPP ($g O_2 m^{-2} d^{-1}$) to water temperature, PAR, and discharge. Average correlations for each cluster were calculated by performing Fisher’s $Z$ transformation on the $r$ values, calculating the mean, and then back-transforming to $r$ values (Fisher 1915).

If clusters are treated as existing classes, then several common methods exist to infer which factors are useful for distinguishing between these classes including discriminant function analysis and logistic regression (Legendre and Legendre 2012). These methods differ from clustering in that they do not try to resolve a set of classes from data but rather they identify how to discriminate between existing classes. We used logistic regression to draw inferences about how site characteristics or environmental drivers may be able to distinguish between productivity regimes. Logistic regression was chosen because many of the site characteristics or environmental metrics in our dataset do not meet the assumption of multivariate normality required to perform discriminant function analysis. Rather than constructing a multivariate model to attempt to predict the classification, we took a more
exploratory approach to this analysis as our relatively small sample size ($n = 47$) limited our ability to consider multivariate models. A random sample without replacement was used to reserve 60% of sites for model training and 40% for validation. Next a univariate logistic regression model was constructed for each variable, and the accuracy of the model at predicting cluster membership in the validation dataset was assessed with a confusion matrix to calculate the overall classification accuracy. Some variables were indistinguishable between clusters, resulting in a model that classified all sites as the same. When this was the case, an NA was recorded, indicating that the variable could not distinguish any sites. The processes of random selection and model classification accuracy were bootstrapped 10,000 times, and the total number of times that a variable failed to discriminate sites as well as the average classification accuracy were recorded. We refer to the percentage of these 10,000 runs where a variable could distinguish between sites as the success rate.

Analysis II—identifying productivity regimes with feature-based partitional clustering

We performed a second clustering analysis of productivity regimes using different techniques to assess the robustness of our clustering results to methodological changes. The seven time series summary statistics we calculated from the GPP time series were used in a $k$-means partitional clustering (MacQueen 1967). Partitional clustering approaches, notably $k$-means, have been used for the classification of streamflow (Olden et al. 2012) as well as the phenology of phytoplankton (Sasaoka et al. 2011). Prior to clustering, each variable was z-normalized to minimize the influence of the relative magnitude of each variable on the clustering results and then a dissimilarity matrix based on Euclidean distance was computed. As before, clustering was used to define between 2 and 10 clusters. A suite of 26 CVIs was computed using the NbClust package (Charrad et al. 2014) in R (R Core Team 2017) and the final clustering solution was determined based on a majority rules approach. In the event of ties, the final clustering solution in NbClust is a conservative solution where the smallest number of clusters is chosen as optimal. It is worth noting that different numbers of CVIs are used between Analysis I and Analysis II, because the use of DTW distance in Analysis I precludes the use of many CVIs that rely on Euclidean distance. The similarity of cluster membership between the clusters derived from DTW hierarchical clustering and the metric-based partitional $k$-means clustering was compared using the Jaccard index.

Analysis III—classification based on site characteristics and environmental drivers

We also examined how clustering sites based on site characteristics or environmental drivers compared to those derived from the DTW hierarchical clustering from Analysis I. Our intention here was to see whether certain factors that potentially mediate productivity could predict similar productivity regimes to those derived directly from the time series of GPP. For this analysis, clustering was done using the same methods for $k$-means clustering presented in the preceding paragraph. Clustering was done on site characteristics and environmental drivers separately. Variables for clustering based on site characteristics included watershed area, river width, percent canopy, and riparian tree height. We did not include site location as a clustering input, because these influences are already encapsulated by some of the environmental drivers, for example, seasonal patterns of incoming PAR are influenced by latitude and longitude. Clusters were generated based on environmental drivers by using the seven time series metrics derived from water temperature, discharge, and PAR ($n = 21$ total metrics). The resulting clusters from both site characteristics and environmental drivers were compared to those derived from Analysis I using the Jaccard index. Finally, we examined the resulting productivity regimes from clustering based on site characteristics or environmental drivers.

Results

Analysis I—identifying productivity regimes with data-based hierarchical clustering

Classification of productivity regimes with hierarchical clustering resulted in a tie between 2 and 10 clusters, each selected as optimal by 3 CVIs. The bootstrapped cluster stability showed that only a single cluster with more than one member had a mean Jaccard index greater than 0.5, and so we selected two clusters as a conservative solution. We named these regimes *Summer Peak Rivers* ($n = 32$) and *Spring Peak Rivers* ($n = 15$) based on the difference in the timing of their peak productivity (Fig. 2).

The two clusters differ in both the timing of peak GPP (the phase of the time series) and in the duration and variability of these annual peaks. *Summer Peak Rivers* ($n = 32$) had the mean ($\pm$1 SD) date of peak productivity in midsummer (phase = day 180 $\pm$ 35) with an extended period of high GPP (Table 2). *Spring Peak Rivers* were characterized by a discrete spring GPP peak (phase = day 100 $\pm$ 14) followed by a very unproductive summer. Although the time series were classified only on the basis of their normalized shape, annual rates of GPP differed substantially between regimes (Table 2). *Summer Peak Rivers* were the most productive, with average annual rates of GPP more than double ($820 \pm 831$ g O$_2$ m$^{-2}$ yr$^{-1}$) that of *Spring Peak Rivers* ($279 \pm 169$ g O$_2$ m$^{-2}$ yr$^{-1}$).

In addition to differences in GPP, the regimes also differed with respect to site and environmental drivers. *Summer Peak Rivers* were generally wider (mean, $\pm$ 1 SD; 29.9 $\pm$ 28.6 m) than *Spring Peak Rivers* (8.7 $\pm$ 7.2 m) (Table 3). Similarly, *Summer Peak Rivers* had larger watershed area (1840 $\pm$ 3248 km$^2$) than did *Spring Peak Rivers* (92 $\pm$ 133 km$^2$). Differences in mean daily discharge between the regimes were consistent with differences in river size, with *Summer Peak* (12.8 $\pm$ 15.72 m$^{-3}$ s$^{-1}$)
having higher discharge than Spring Peak Rivers (0.94 ± 1.21 m$^{-3}$ s$^{-1}$). Daily rates of GPP (g O$_2$ m$^{-2}$ d$^{-1}$) were moderately correlated to environmental drivers in Summer Peak Rivers (water temperature \( r = 0.35 \); discharge \( r = -0.14 \); PAR \( r = 0.39 \)). By contrast, daily patterns of GPP in Spring Peak Rivers were only weakly correlated to water temperature \( r = -0.04 \), discharge \( r = 0.01 \), or PAR \( r = 0.17 \).

Only 15 of 27 variables examined were able to distinguish between the two clusters more than 50% of the time (Table 4). The highest classification accuracy was the AR(1) of discharge (76%), followed by mean discharge (74%) and watershed area (73%). All three of these variables had a high success rate at distinguishing the two regimes, and overall, there was a strong correlation between success rate and classification accuracy.

The amplitude and skewness of water temperature also had a success rate of at least 90%, resulting in a total of five variables with high classification accuracy. Overall, models had higher classification accuracies for Summer Peak Rivers than for Spring Peak Rivers, but this could potentially be related to larger representation of Summer Peak Rivers within the dataset resulting in a more accurate model for class membership.

### Analysis II—identifying productivity regimes with feature-based partitional clustering

The clustering results from the $k$-means and hierarchical clustering of GPP were compared to assess the robustness of classified productivity regimes despite methodological differences between the two analyses. Two distinct patterns of
productivity were also identified by the k-means clustering analysis that used derived metrics of GPP. It is pertinent to note that the majority rules approach identified two different numbers of clusters as optimal, with both two and four clusters suggested as optimal by six CVIs each. Thus, a conservative solution of two clusters was selected for the final number of clusters. The two patterns identified by k-means clustering reflect the same patterns of Summer Peak and Spring Peak Rivers determined in the hierarchical clustering from Analysis I. The corresponding clusters derived from k-means clustering had high agreement to Summer Peak Rivers from the hierarchical clustering analysis \((I = 0.83)\) as well as a similar number of

| Table 3. Site characteristics and derived metrics from time series of environmental drives for each of the two regimes. The mean (±1 SD) values for each regime were calculated from all of the sites included in a regime. |
|---|---|---|
| **Site characteristics** | **Summer peak** | **Spring peak** |
| Latitude | 38.57 ± 4.02 | 36.43 ± 4.16 |
| Longitude | −88.34 ± 2.95 | −80.03 ± 3.29 |
| Accumulated watershed area (km²) | 1840 ± 3248 | 92 ± 133 |
| River width (m) | 29.9 ± 28.6 | 8.7 ± 7.2 |
| Canopy cover (%) | 45 ± 25 | 39 ± 13 |
| Tree height (m) | 14.5 ± 10.4 | 16.8 ± 9.5 |
| Water temperature metrics* | | |
| Mean daily (°C) | 14.07 ± 3.39 | 15.16 ± 3.23 |
| Coefficient of variation (L-CV) | 0.32 ± 0.09 | 0.27 ± 0.09 |
| Skewness (L-skew) | −0.04 ± 0.03 | −0.02 ± 0.03 |
| Kurtosis (L-kurt) | −0.01 ± 0.04 | −0.06 ± 0.04 |
| AR(1) correlation coefficient | 0.86 ± 0.03 | 0.84 ± 0.02 |
| Amplitude | 1.36 ± 0.02 | 1.33 ± 0.03 |
| Phase (day of year) | 205 ± 6 | 204 ± 2 |
| Discharge metrics* | | |
| Mean daily (m³ s⁻¹) | 12.8 ± 15.72 | 0.94 ± 1.21 |
| Coefficient of variation (L-CV) | 0.53 ± 0.19 | 0.56 ± 0.09 |
| Skewness (L-skew) | 0.56 ± 0.14 | 0.59 ± 0.1 |
| Kurtosis (L-kurt) | 0.41 ± 0.15 | 0.44 ± 0.11 |
| AR(1) correlation coefficient | 0.66 ± 0.24 | 0.38 ± 0.24 |
| Amplitude | 0.42 ± 0.22 | 0.24 ± 0.15 |
| Phase (day of year) | 79 ± 44 | 110 ± 77 |
| PAR metrics* | | |
| Mean daily (μmol m⁻² s⁻¹) | 399.01 ± 42.75 | 408.92 ± 33.22 |
| Coefficient of variation (L-CV) | 0.27 ± 0.03 | 0.25 ± 0.03 |
| Skewness (L-skew) | 0 ± 0.02 | 0.02 ± 0.03 |
| Kurtosis (L-kurt) | 0.03 ± 0.03 | 0.02 ± 0.02 |
| AR(1) correlation coefficient | 0.3 ± 0.1 | 0.28 ± 0.09 |
| Amplitude | 1.2 ± 0.05 | 1.15 ± 0.04 |
| Phase (day of year) | 176 ± 3 | 175 ± 3 |

*Calculated following Archfield et al. (2014).

members \((n = 29)\). Similarly, the k-means cluster corresponding to Spring Peak Rivers also had high agreement of cluster membership \((I = 0.91)\).

Analysis III—classification based on site characteristics and environmental drivers

We compared the clustering solutions based on site characteristics and environmental drivers to the two regimes determined from the hierarchical clustering analysis. Clustering
The resulting productivity regimes from k-means clustering based on derived metrics of GPP (a, e), site conditions (b, f), and derived metrics from time series of water temperature, discharge, and PAR (c, g). The productivity regimes from k-means clustering on a selected combination of factors (watershed area, water temperature amplitude, L-skew of water temperature, the autoregressive lag-one correlation coefficient of discharge, and mean daily discharge) are also presented (d, h). The regimes are displayed as the median and interquartile range for all sites belonging to a regime for a given clustering.

Based on site characteristics and environmental drivers both resulted in two clusters; however, once again, the final number of clusters for environmental drivers resulted from a conservative selection between two and four clusters. There was moderately low agreement between site-based clusters and Summer Peak Rivers (f = 0.41) and Spring Peak Rivers (f = 0.47). The same was also true for environmental driver-based clusters (Summer Peak Rivers f = 0.42; Spring Peak Rivers f = 0.47). Additionally, clustering was performed using a combination of site characteristics and environmental drivers identified to regularly discriminate between Summer Peak and Spring Peak Rivers in the previous analysis (watershed area, water temperature amplitude, water temperature L-skew, AR(1) of discharge, and mean daily discharge). This final clustering based on a limited set of proximal controls of productivity had the highest agreement in cluster membership for both Summer Peak (f = 0.54) and Spring Peak Rivers (f = 0.64) and was the only set of clusters with greater than 50% membership to these regimes.

Despite classifying rivers based purely on site characteristics or environmental drivers, the resulting temporal patterns of GPP are remarkably similar to the previously identified Summer Peak and Spring Peak Rivers (Fig. 3). Although the overall temporal patterns are similar, the analogous Summer Peak pattern produced by clustering on site characteristics does have a higher amplitude than those derived from other approaches (Fig. 3b). Across the various sets of k-means clustering, the corresponding Spring Peak pattern had less variation among the resulting time series of GPP.

**Beyond two regimes?**

The presence of two distinct productivity regimes, Summer Peak Rivers and Spring Peak Rivers, manifested themselves from our collection of sites from different clustering methodologies and datasets. We have taken a conservative approach and presented the majority of results using these clusters, because they were most robust to changes in clustering inputs or methodologies. However, both of our clustering analyses also hint at the presence of more than two regimes. The hierarchical clustering analysis was unable to identify an optimum clustering solution with CVIs. Additionally, k-means clustering of both GPP metrics and a limited set of proximal controls on productivity only narrowly selected two over four clusters. Consequently, we examined the dendrogram from the hierarchical clustering and found that there were three stable clusters (Jaccard index > 0.5) and one additional cluster with weak temporal patterns and low absolute rates of productivity. Although this fourth cluster was not robust to the bootstrapping analysis (mean Jaccard coefficient = 0.39), we included it in the results below because this pattern may arise in frequently disturbed (Blaszczak et al. 2018) or blackwater rivers (Meyer and Edwards 1990).

Within these four clusters, the same temporal patterns of Summer Peak Rivers (n = 12) and Spring Peak Rivers (n = 15) are evident. In fact, the membership of Spring Peak Rivers is identical to those presented in the previous analysis. Two additional regimes are present, and we named these Aseasonal Rivers (n = 7) and Summer Decline Rivers (n = 13) following the naming conventions of the previously identified regimes (Fig. 4). Aseasonal Rivers had relatively invariant low rates of GPP year-round. Summer Decline Rivers (n = 13) had productivity peaks that occurred slightly earlier in the year (phase = day 146 ± 13), with rates of GPP declining gradually through the summer. Although annual rates of GPP were moderately positively correlated to channel width across all sites (r = 0.49), this relationship was tighter within Spring Peak Rivers (Fig. 5a). Outside the Aseasonal Rivers, the correlation between river width and annual GPP strengthened with decreasing river size.

Across all sites, annual GPP was most negatively correlated to L-C.V. of discharge (r = −0.37) but was strongly negatively correlated in Summer Decline Rivers (r = −0.85; Fig. 5b). Daily rates
of GPP (g O$_2$ m$^{-2}$ d$^{-1}$) were more strongly correlated to environmental conditions in Summer Peak Rivers (water temperature \( r = 0.55 \); discharge \( r = -0.29 \); PAR \( r = 0.51 \)) once Aseasonal and Summer Decline Rivers were no longer included.

**Discussion**

**Controls of productivity regimes**

Our analysis of the seasonality of river productivity supports the presence of characteristic river metabolic regimes. We identified two clearly distinct metabolic regimes, Summer Peak and Spring Peak Rivers, within our dataset of 47 U.S. streams and rivers. These regimes emerged repeatedly from our dataset regardless of the specific clustering analysis or whether the clusters were based on GPP or various site and environmental drivers. Although the exact membership of these clusters was not identical, the striking resemblance of their representative GPP regimes (Fig. 3) suggests that these temporal patterns are robust. These two GPP regimes are meaningfully different in both their timing and annual rates of GPP. This suggests that the timing and magnitude of GPP in streams and rivers are intimately linked. We also found that clustering based on a handful of easily derived site characteristics or environmental drivers (watershed area, amplitude of water temperature, L-skew of water temperature, AR(1) of discharge, and mean discharge) was able to predict more than half of the membership of Summer Peak Rivers and Spring Peak Rivers.

The prominence of these two regimes within our dataset affirms previously documented temporal patterns typical to large rivers vs. smaller headwater streams. Comparing these regimes based on the seasonality of productivity, the characteristics of rivers which belong to them, the correlates of productivity within each regime, and using illustrative examples from existing literature provide evidence for different primary constraints on productivity within each regime. The large watershed area and channel widths associated with Summer Peak Rivers are consistent with observed summer peak patterns in large rivers (Uehlinger 2006), although similar patterns could also occur in smaller streams with open canopies. This was the only regime where peak productivity was coincident
with the timing of peak incoming light. The coherence between GPP and climatic conditions suggests that patterns of GPP may be more closely related to distal controls of productivity, and additional proximal constraints imposed by canopy shading, hydrologic disturbance, or other potentially limiting factors are of secondary importance. *Summer Peak Rivers* effectively represent the upper envelope of productivity within our dataset.

The productivity regimes of *Spring Peak Rivers* were truncated and thus on average had cumulative annual GPP approximately one third of *Summer Peak Rivers*. The mismatch between their annual peak productivity with incoming PAR and water temperature peaks suggests canopy shading is an important constraint. (Bernhardt et al. 2017). *Spring Peak Rivers* exhibited a small period of productivity followed by a sharp decline similar to the pattern observed previously in Walker Branch, a small forested headwater stream in Tennessee (e.g., Roberts et al. 2007). However, it is worth noting that many *Spring Peak Rivers* in our dataset are large enough to be considered third- or fourth-order rivers (Downing et al. 2012) and thus could potentially represent a large fraction of river networks. In *Spring Peak Rivers*, productivity may be limited to a brief window prior to canopy closure when both light and temperature support higher rates of GPP. The relative importance of terrestrial shading on mediating productivity is expected to vary with river size (Vannote et al. 1980) and temporal patterns of productivity have been shown to be constrained by terrestrially mediated light availability (Roberts et al. 2007; Izagirre et al. 2008). The importance of these controls on seasonal patterns of productivity in *Spring Peak Rivers* is supported by their comparatively smaller width and catchment area relative to other clusters (Table 2) and the correlation between width and productivity within this cluster (Fig. 5). We expect that network position and terrestrial pheno-ology may represent the most important drivers of variation in productivity across *Spring Peak Rivers*.

We also suggest there may be two additional regimes that, although less robust to statistical measures, provide a glimpse into other potential patterns beyond large rivers and smaller streams. *Summer Decline Rivers* are characterized by an early spring peak followed by sustained, but reduced, productivity throughout the summer months. These rivers were generally similar to but slightly smaller than *Summer Peak Rivers*. Our observation that flow variability was strongly negatively correlated with annual productivity in these *Summer Decline Rivers* suggests that hydrologic disturbance plays an important role within this group of rivers. *Aseasonal Rivers* were also generally unproductive, suggesting that these rivers experience strong constraints on productivity throughout the year. A variety of mechanisms could interact to produce these patterns, including non-deciduous canopies, persistent turbidity or color, and frequent hydrologic disturbance. Aseasonal patterns have been observed in blackwater rivers where water clarity limits productivity throughout the year (Meyer and Edwards 1990) and in some urban streams where frequent disturbance limits biomass accumulation (Larsen and Harvey 2017).

**Classification**

The prevalence of classification in ecology (Whittaker 1962) is due in part to the fact that it provides a conceptual framework for organizing complex systems into manageable units for comparison between or within systems. We found

---

**Fig. 5.** The relationship between mean annual GPP and potential drivers of GPP. The relationship between river width and GPP was examined for all sites (gray points) and only for sites in the *Spring Peak* regime (black points) (a). Similarly, the relationship between the L-C.V. of daily discharge and average annual GPP was assessed for sites in the *Summer Decline* regime (white points) (b). Mean values of L-C.V. were calculated for each site based on yearly values from 2013–2016.
similar temporal regimes of productivity whether classification was performed directly on daily estimates of GPP, site characteristics, or environmental drivers. Furthermore, a relatively small number of proximal variables were able to predict the membership of sites to our two regimes with modest accuracy. This supports the notion that similar patterns of GPP should have a common set of environmental controls mediating these patterns. Therefore, although the primary proximal constraints on GPP (light, heat, disturbance, nutrient supply, grazing) are the same in all systems (Bernhardt et al. 2017), we expect that the relative importance would differ among regimes. The proximal controls on productivity have potentially complex relationships with broader-scale drivers such as climate, geology, topography, and land use and the gradients of channel size and shape that are embedded within them. More empirical data on key environmental patterns and drivers (e.g., light and turbidity) will be necessary to rigorously evaluate the network of causal relationships that determine the metabolic regimes of flowing waters.

We have proposed one possible approach to begin investigating both the patterns of and processes mediating stream metabolic regimes through classification. For several reasons, it is highly unlikely that this initial analysis has uncovered the totality of different annual metabolic regimes that occur in streams worldwide. The relatively small number of sites and their limited geographic range may have been magnified by the relatively strict criteria we used for inclusion of sites in this analysis. Our dataset also almost certainly over-represents larger channels owing to the distribution of USGS observation platforms. Other classification schemes might resolve different clusters of metabolic patterns. Therefore, our results are not meant to suggest a definitive set of river productivity regimes but to provide one framework for analysis of river productivity patterns. The strength of our classification approach is that it takes full advantage of near-continuous metabolism estimates while still ultimately distilling those records to a manageable, interpretable set of classes. There are of course numerous possibilities for performing similar synthetic studies. For example, the predictability and timing of flooding events has been treated as a continuum for assessing impacts on biological communities (Jardine et al. 2015), and wavelet analysis has been used to assess the relevant timescales of drivers on observed patterns of net ecosystem exchange on land (Stoy et al. 2005). Such methods could complement our exploration into the seasonality of river productivity.

**Implications and future directions**

Characterizing and classifying metabolic regimes will help illuminate links between productivity and the biomass or structure of aquatic communities. The preferential selection of high-quality C from autotrophic activity by stream consumers means that even small fluxes of autotrophic C production can substantially impact aquatic food webs (Minshall 1978; Marcarelli et al. 2011). Changes in the magnitude of this flux may affect the total energy available to support secondary consumers, whereas changes in the timing of autochthonous production may select for different dominant consumers. Furthermore, as new insights are gained about the factors mediating algal community composition (Lange et al. 2016), it may be possible to incorporate relevant drivers of these changes into classification schemes. As we better understand the variety of metabolic regimes and their resilience or sensitivity to climate change and hydrologic alterations, we are likely to gain new insights into the potential for energetic limitation to constrain attempts to protect, conserve, and restore aquatic ecosystems and species.

The productivity of rivers is a primary determinant of their capacity to retain and transform nutrients (Roberts and Mulholland 2007; Mulholland et al. 2008; Heffernan and Cohen 2010). The degree of synchrony or asynchrony between nutrient loading and productivity, and thus nutrient demand, may help explain a great deal of the variation in nutrient retention across rivers (Covino et al. 2018). Employing similar classification techniques to both productivity and nutrient data could help by identifying those times and places where rivers are most and least likely to be able to process excess nutrient loads. Such information could improve our ability to prevent or mitigate eutrophication issues in rivers. In nutrient-limited river systems, this same approach may allow identification of periods of time when nutrient supply or nutrient supply ratios limit ecosystem productivity.

As new metabolism records are produced that encompass a wider breadth of environmental conditions, we expect that the utility of this classification framework will continue to improve. Incorporating a greater diversity of productivity regimes and drivers will enhance our ability to differentiate classes and their drivers. The inclusion of more sites will also aid our power to derive information from the hierarchy of classification through the investigation of drivers most relevant for discerning productivity regimes at different hierarchical levels. For example, our results suggest that light limitation from terrestrial environments produces fundamentally different patterns of productivity. Because our classification framework is based on functional responses of GPP, which should reflect combinations of underlying drivers, we believe that it can also be used to test or validate conceptual organizational frameworks of stream ecosystems such as process domains (Montgomery 1999) or process zones (Thorp et al. 2006). As the number of sites grows, we should be able to resolve finer differences down the hierarchy of controls at sites to determine the relative importance of various drivers that may shape the seasonality of productivity.

As longer metabolism time series datasets become available, we can begin to ask questions about interannual variability in metabolism and we can detect metabolic change in response to external drivers with far more precision than has previously been possible. For example, this classification schema provides a more sophisticated way of thinking about aquatic and
terrestrial controls on productivity and allows us to begin to explore how shifts in vegetation type, phenology, or management may be related to the seasonality of productivity in rivers. We can also begin to ask questions about how commonly river ecosystems shift between regimes, what external drivers lead to these shifts, and the extent to which shifts are periodic, sustained, or reversible. Such insights will vastly improve our understanding of lotic ecosystems while also offering a suite of new tools for monitoring the impact of management decisions and climate change.

References

Acuña, V., I. Muñoz, A. Giorgi, M. Omella, F. Sabater, and S. Sabater. 2005. Drought and postdrought recovery cycles in an intermittent Mediterranean stream: Structural and functional aspects. J. N. Am. Benthol. Soc. 24: 919–933. doi:10.1899/04-078.1

Aghabozorgi, S., A. Seyed Shirkhorshidi, and T. Ying Wah. 2015. Time-series clustering—a decade review. Inf. Syst. 53: 16–38. doi:10.1016/j.is.2015.04.007

Appling, A. P., R. O. Hall Jr., C. B. Yackulic, and M. Arroita. 2018a. Overcoming equifinality: Leveraging long time series for stream metabolism estimation. J. Geophys. Res. Biogeosci. 123: 624–645. doi:10.1002/2017JG004140

Appling, A. P., R. O. Hall, M. Arroita, and C. B. Yackulic. 2018b. streamMetabolizer: Models for estimating aquatic photosynthesis and respiration. R package version 0.10.9. https://github.com/USGS-R/streamMetabolizer

Appling, A. P., and others. 2018c. The metabolic regimes of 356 rivers in the United States. Sci. Data 5: 180292. doi:10.1002/2017JG004140

Appling, A. P., and others. 2018d. Metabolism estimates for 356 U.S. rivers (2007-2017). U.S. Geological Survey. doi:10.5066/F70864KX

Arbelaitz, O., I. Gurruxtxaga, J. Muguerza, J. M. Pérez, and I. Perona. 2013. An extensive comparative study of cluster validity indices. Pattern Recogn. 46: 243–256. doi:10.1016/j.patcog.2012.07.021

Archfield, S. A., J. G. Kennen, D. M. Carlisle, and D. M. Wolock. 2014. An objective and parsimonious approach for classifying natural flow regimes at a continental scale. River Res. Appl. 30: 1166–1183. doi:10.1002/rra.2710

Berndt, D. J., and J. Clifford. 1994. Using dynamic time warping to find patterns in time series. AAAI technical report WS-94-03. Association for the Advancement of Artificial Intelligence.

Bernhardt, E. S., et al. 2017. The metabolic regimes of flowing waters. Limnol. Oceanogr 63: S99–S118. doi:10.1002/lio.10726

Blaszczyk, J. R., J. M. Delesantro, D. L. Urban, M. W. Doyle, and E. S. Bernhardt. 2018. Scoured or suffocated: Urban stream ecosystems oscillate between hydrologic and dissolved oxygen extremes. Limnol. Oceanogr. doi:10.1002/ino.11081

Boisvenue, C., and S. W. Running. 2006. Impacts of climate change on natural forest productivity—evidence since the middle of the 20th century. Glob. Chang. Biol. 12: 862–882. doi:10.1111/j.1365-2486.2006.01134.x

Britton, C. M., and J. D. Dodd. 1976. Relationships of photosynthetically active radiation and shortwave irradiance. Agric. Meteorol. 17: 1–7. doi:10.1016/0001-850X(76)90080-7

Chapra, S., and D. Di Toro. 1991. Delta method for estimating primary production, respiration, and reaeration in streams. J. Environ. Eng. 117: 640–655. doi:10.1061/(ASCE)0733-9372(1991)117:5(640)

Charrad, M., N. Ghazzali, V. Boiteau, and A. Niknafs. 2014. NbClust: An R package for determining the relevant number of clusters in a data set. J. Stat. Softw. 61: 36. doi:10.18637/jss.v061.i06

Demars, B. O. L., and others. 2011. Temperature and the metabolic balance of streams. Freshw. Biol. 56: 1106–1121. doi:10.1111/j.1365-2427.2010.02554.x

Downing, J. A., and others. 2012. Global abundance and size distribution of streams and rivers. Inland Waters 2: 229–236. doi:10.5268/IW-2.4.502

Field, C. B., M. J. Behrenfeld, J. T. Randerson, and S. Sabater. 2005. Drought and postdrought recovery cycles in an intermittent Mediterranean stream: Structural and functional aspects. J. N. Am. Benthol. Soc. 24: 919–933. doi:10.1899/04-078.1

Finlay, J. C. 2011. Stream size and human influences on ecosystem production in river networks. Ecosphere 2: 21. doi:10.1890/ES11-00071.1

Fisher, R. A. 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika 10: 507–521. doi:10.2307/2331838

Frisell, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. Environ. Manag. 10: 199–214. doi:10.1007/BF01867358

Fu, T.-C. 2011. A review on time series data mining. Eng. Appl. Artif. Intell. 24: 164–181. doi:10.1016/j.engappai.2010.09.007

Giorgino, T. 2009. Computing and visualizing dynamic time warping alignments in R: The dtw package. J. Stat. Softw. 37: 1–45. doi:10.18637/jss.v037.i01

Gomez-Velez, J. D., J. W. Harvey, M. B. Cardenas, and B. Kiel. 2015. Denitrification in the Mississippi River network
controlled by flow through river bedforms. Nat. Geosci. 8: 941–945. doi:10.1038/ngeo2567
Gu, L., W. M. Post, D. Baldocchi, T. A. Black, S. B. Verma, T. Vesala, and S. C. Wofsy. 2003. Phenology of vegetation photosynthesis, p. 467–485. In M. D. Schwartz [ed.], Phenology: An integrative environmental science. Tasks for vegetation science. Springer. doi:10.1007/978-94-007-0632-3_29
Hall, R. O., C. B. Yackulic, T. A. Kennedy, M. D. Yard, E. J. Rosi-Marshall, N. Voichich, and K. E. Behn. 2015. Turbidity, light, temperature, and hypodepeaking control primary productivity in the Colorado River, grand canyon. Limnol. Oceanogr. 60: 512–526. doi:10.1002/lo.10031
Heffernan, J. B., and M. J. Cohen. 2010. Direct and indirect coupling of primary production and diel nitrate dynamics in a subtropical spring-fed river. Limnol. Oceanogr. 55: 677–688. doi:10.4319/lo.2010.55.2.0677
Hill, W. R., and S. M. Dimick. 2002. Effects of riparian leaf dynamics on periphyton photosynthesis and light utilisation efficiency. Freshw. Biol. 47: 1245–1256. doi:10.1046/j.1365-2427.2002.00837.x
Homer, C. G., and others. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States—representing a decade of land cover change information. Photogramm. Eng. Remote. Sens. 81: 345–354. doi:10.3390/rs6087424
Izagirre, O., U. Agirre, M. Bermejo, J. Pozo, and A. Elosegi. 2008. Environmental controls of whole-stream metabolism identified from continuous monitoring of Basque streams. J. N. Am. Benthol. Soc. 27: 252–268. doi:10.1899/07-022.1
Jardine, T. D., and others. 2015. Does flood rhythm drive ecosystem responses in tropical riverscapes? Ecology 96: 684–692. doi:10.1890/14-0991.1
Keenan, T. F., and others. 2014. Net carbon uptake has increased through warming-induced changes in temperate forest phenology. Nat. Clim. Chang. 4: 598–604. doi:10.1038/nclimate2253
Keogh, E., and S. Kasetty. 2003. On the need for time series data mining benchmarks: A survey and empirical demonstration. Data Min. Knowl. Disc. 7: 349–371. doi:10.1023/A:1024988512476
Lamberti, G. A., and A. D. Steinman. 1997. A comparison of primary production in stream ecosystems. J. N. Am. Benthol. Soc. 16: 95–104. doi:10.2307/1468241
Lange, K., C. R. Townsend, and C. D. Matthaei. 2016. A trait-based framework for stream algal communities. Ecol. Evol. 6: 23–36. doi:10.1002/ece3.1822
Larsen, L. G., and J. Harvey. 2017. Disrupted carbon cycling in restored and unrestored urban streams: Critical timescales and controls. Limnol. Oceanogr. 62: S160–S182. doi:10.1002/lo.10613
Legendre, P., and L. F. Legendre. 2012. Numerical ecology. Elsevier.
Liao, T. W. 2005. Clustering of time series data—A survey. Pattern Recogn. 38: 1857–1874. doi:10.1016/j.patcog.2005.01.025
Lloyd, D. 1990. A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. Int. J. Remote Sens. 11: 2269–2279. doi:10.1080/01431169008955174
MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations, p. 281–297. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics. Fifth Berkeley Symposium on Mathematical Statistics and Probability. Univ. of California Press.
Maheu, A., N. Poff, and A. St-Hilaire. 2015. A Classification of Stream Water Temperature Regimes in the Conterminous USA. River Res. Applic. 32: 896–906. doi:10.1002/rra.2906
Marcarelli, A. M., C. V. Baxter, M. M. Mineau, and R. O. Hall. 2011. Quantity and quality: Unifying food web and ecosystem perspectives on the role of resource subsidies in freshwaters. Ecology 92: 1215–1225. doi:10.1890/10-2240.1
Menzel, A., and P. Fabian. 1999. Growing season extended in Europe. Nature 397: 659–659. doi:10.1038/17709
Menzel, A., and others. 2006. European phenological response to climate change matches the warming pattern. Glob. Chang. Biol. 12: 1969–1976. doi:10.1111/j.1365-2486.2006.01193.x
Meyer, J. L., and R. T. Edwards. 1990. Ecosystem metabolism and turnover of organic carbon along a Blackwater River continuum. Ecology 71: 668–677. doi:10.2307/1940321
Milly, P. C., K. A. Dunne, and A. V. Vecchia. 2005. Global pattern of trends in streamflow and water availability in a changing climate. Nature 438: 347–350. doi:10.1038/nature04312
Minshall, G. W. 1978. Autotrophy in stream ecosystems. BioScience 28: 767–771. doi:10.2307/1307250
Mitchell, K. E., et al. 2004. The multi-institution north American land data assimilation system (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. J. Geophys. Res. Atmos. 109: 1–32. doi:10.1029/2003JD003823
Montgomery, D. R. 1999. Process domains and the river continuum model of trends in stream metabolism. Freshw. Biol. 43: 397–410. doi:10.1111/j.1365-2486.1999.tb03598.x
Mueen, A., and E. Keogh. 2016. Extracting optimal performance from dynamic time warping, p. 2129–2134. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM.
Mulholland, P. J., and others. 2001. Inter-biome comparison of factors controlling stream metabolism. Freshw. Biol. 46: 1503–1517. doi:10.1046/j.1365-2427.2001.00773.x
Mulholland, P. J., and others. 2008. Stream denitrification across biomes and its response to anthropogenic nitrate loading. Nature 452: 202–205. doi:10.1038/nature06686
Noormets, A. 2009. Phenology of ecosystem processes: Applications in global change research. Springer. doi:10.1007/978-1-4419-0026-5
O’Connor, B. L., J. W. Harvey, and L. E. McPhillips. 2012. Thresholds of flow-induced bed disturbances and their
effects on stream metabolism in an agricultural river. Water Resour. Res. 48: 1–18. doi:10.1029/2011WR011488
Olden, J. D., and R. J. Naiman. 2010. Incorporating thermal regimes into environmental flows assessments: Modifying dam operations to restore freshwater ecosystem integrity. Freshw. Biol. 55: 86–107. doi:10.1111/j.1365-2427.2009.02179.x
Olden, J. D., M. J. Kennard, and B. J. Pusey. 2012. A framework for hydrologic classification with a review of methodologies and applications in ecohydrology. Ecohydrology 5: 503–518. doi:10.1002/eco.251
Parmesan, C. 2006. Ecological and evolutionary responses to recent climate change. Annu. Rev. Ecol. Evol. Syst 37: 637–669. doi:10.1146/annurev.ecolsys.37.091305.110100
Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421: 37–42. doi:10.1038/nature01286
Peeters, F., D. Straile, A. Lorke, and D. M. Livingstone. 2007. Earlier onset of the spring phytoplankton bloom in lakes of the temperate zone in a warmer climate. Glob. Chang. Biol. 13: 1898–1909. doi:10.1111/j.1365-2486.2007.01412.x
Poff, N. 1996. A hydrogeography of unregulated streams in the United States and an examination of scale-dependence in some hydrological descriptors. Freshw. Biol. 36: 71–79. doi:10.1046/j.1365-2427.1996.00073.x
Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegaard, B. D. Richter, R. E. Sparks, and J. C. Stromberg. 1997. The natural flow regime. Bioscience 47: 769–784. doi:10.2307/1313099
Poole, G. C. 2002. Fluvial landscape ecology: Addressing uniqueness within the river discontinuum. Freshw. Biol. 47: 641–660. doi:10.1046/j.1365-2427.2002.00922.x
Poole, G. C., and C. H. Berman. 2001. An ecological perspective on in-stream temperature: Natural heat dynamics and mechanisms of human-CausedThermal degradation. Environ. Manag. 27: 787–802. doi:10.1007/s002670010188
R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing.
Racault, M.-F., C. Le Quéré, E. Buitenhuis, S. Sathyendranath, and T. Platt. 2012. Phytoplankton phenology in the global ocean. Ecol. Indic. 14: 152–163. doi:10.1016/j.ecolind.2011.07.010
Ratanamahatana, C. A., and E. Keogh. 2005. Three myths about dynamic time warping data mining, p. 250. In Proceedings of the 2005 SIAM International Conference on Data Mining. SIAM.
Roberts, B. J., and P. J. Mulholland. 2007. In-stream biotic control on nutrient biogeochemistry in a forested stream, west fork of Walker branch. J. Geophys. Res. Biogeosci. 112: 1–11. doi:10.1029/2007JG000422
Roberts, B. J., P. J. Mulholland, and W. R. Hill. 2007. Multiple scales of temporal variability in ecosystem metabolism rates: Results from 2 years of continuous monitoring in a forested headwater stream. Ecosystems 10: 588–606. doi:10.1007/s10021-007-9059-2
Sakoe, H., and S. Chiba. 1978. Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. Acoust. Speech Signal Process. 26: 43–49. doi:10.1109/TASSP.1978.1163055
Sarda-Espinosa, A. 2017. dtwclust: Time series clustering along with optimizations for the dynamic time warping distance. R package version 4.1.1. https://CRAN.R-project.org/package=dtwclust
Sasaoka, K., S. Chiba, and T. Saino. 2011. Climatic forcing and phytoplankton phenology over the subarctic North Pacific from 1998 to 2006, as observed from ocean color data. Geophys. Res. Lett. 38: 1–6. doi:10.1029/2011GL048299
Schwartz, M. D., R. Ahas, and A. Aasa. 2006. Onset of spring starting earlier across the northern hemisphere. Glob. Chang. Biol. 12: 343–351. doi:10.1111/j.1365-2486.2005.01097.x
Simard, M., N. Pinto, J. B. Fisher, and A. Baccini. 2011. Mapping forest canopy height globally with spaceborne lidar. J. Geophys. Res. Bioges. 116: 1–12. doi:10.1029/2011JG001708
Sørensen, T. 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. Biol. Skr. 5: 1–34.
Stoy, P. C., G. G. Katul, M. B. Siqueira, J. Y. Juang, M. C. HR, H. S. Kim, A. C. Oishi, and R. Oren. 2005. Variability in net ecosystem exchange from hourly to inter-annual time scales at adjacent pine and hardwood forests: A wavelet analysis. Tree Physiol. 25: 887–902. doi:10.1093/treephys/25.7.887
Teng, W., H. Rui, R. Strub, and B. Vollmer. 2016. Optimal reorganization of NASA earth science data for enhanced accessibility and usability for the hydrology community. JAWRA J. Am. Water Resour. Assoc. 52: 825–835. doi:10.1111/j.1752-1688.12405
Thorp, J. H., M. C. Thoms, and M. D. Delong. 2006. The riverine ecosystem synthesis: Biocomplexity in river networks across space and time. River Res. Appl. 22: 123–147. doi:10.1002/rra.901
Uehlinger, U. 2000. Resistance and resilience of ecosystem metabolism in a flood-prone river system. Freshw. Biol. 45: 319–332. doi:10.1111/j.1365-2427.2000.00620.x
Uehlinger, U. 2006. Annual cycle and inter-annual variability of gross primary production and ecosystem respiration in a flood-prone river during a 15-year period. Freshw. Biol. 51: 938–950. doi:10.1111/j.1365-2427.2006.01551.x
USGS. 2017. National Water Information System Web interface. USGS.
Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The river continuum concept. Can. J. Fish. Aquat. Sci. 37: 130–137. doi:10.1139/f80-017
Ward, J. V., and J. Stanford. 1983. Serial discontinuity concept of lotic ecosystems, p. 29–42. In T. D. Fontaine, and S. M. Bartell [eds.], Dynamics of Lotic Systems. Ann Arbor Science, Ann Arbor.

White, M. A., F. Hoffman, W. W. Hargrove, and R. R. Nemani. 2005. A global framework for monitoring phenological responses to climate change. Geophys. Res. Lett. 32: 1–4. doi:10.1029/2004GL021961

Whittaker, R. H. 1962. Classification of natural communities. Bot. Rev. 28: 1–239. doi:10.1007/BF02860872

Williams, J. W., and S. T. Jackson. 2007. Novel climates, no-analog communities, and ecological surprises. Front. Ecol. Environ. 5: 475–482. doi:10.1890/070037

Williams, J. W., S. T. Jackson, and J. E. Kutzbach. 2007. Projected distributions of novel and disappearing climates by 2100 AD. Proc. Natl. Acad. Sci. USA 104: 5738–5742. doi:10.1073/pnas.0606292104

Wood, S. 2006. Generalized additive models: An introduction with R. CRC Press.

Xia, Y., and others. 2012. Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. J. Geophys. Res. Atmos. 117: D03109. doi:10.1029/2011JD016048

Xue, Z., P. Du, and L. Feng. 2014. Phenology-driven land cover classification and trend analysis based on long-term remote sensing image series. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. 7: 1142–1156. doi:10.1109/JSTARS.2013.2294956

Zhang, Y., and G. F. Hepner. 2017. The dynamic-time-warping-based k-means++ clustering and its application in phenoregion delineation. Int. J. Remote Sens. 38: 1720–1736. doi:10.1080/01431161.2017.1286055

Acknowledgments

The work described here was supported by working group funds through the USGS Powell Center Synthesis Center and by a NSF Macrosystems Program Grant (EF 1442439). We would also like to thank the anonymous reviewers for their comments which helped us refine and improve the content of the manuscript.

Conflict of Interest

None declared